Manufacturing & Service Operations Management

Optimal Algorithms for Assortment Selection Under Ranking-Based Consumer Choice Models

Dorothee Honhon, Sreelata Jonnalagedda, Xiajun Amy Pan,

Optimal Algorithms for Assortment Selection Under Ranking-Based Consumer Choice Models

Dorothee Honhon
Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands, d.b.l.p.honhon@tue.nl

Sreelata Jonnalagedda
Indian Institute of Management, Bangalore 560076, India, sreelata@iimb.ernet.in

Xiajun Amy Pan
University of Florida, Gainesville, Florida 32611, amypan@ufl.edu

A retailer's product selection decisions are largely driven by her assumptions on how consumers make choices. We use a ranking-based consumer choice model to represent consumer preferences: every customer has a ranking of the potential products in the category and purchases his highest ranked product (if any) offered in the assortment. We consider four practically motivated special cases of this model, namely, the one-way substitution, the locational choice, the outtree, and the intree preference models, and we study the retailer's product selection problem when products have different price and cost parameters. We assume that the retailer incurs a fixed carrying cost per product offered, a goodwill penalty for each customer who does not purchase his first choice and a lost sale penalty for each customer who does not find an acceptable product to buy. For the first three models, we obtain efficient solution methods that simplify to either a shortest path method or a dynamic program. For the fourth model, we construct an effective algorithm and show numerically that, in practice, it is much faster than enumeration. We also obtain valuable insights on the structure of the optimal assortment.

1. Introduction

Retailers face a routine and complex task of selecting the right products to offer from an increasing number of goods offered in the market place (Draganska and Jain 2005), in a way that addresses consumers' needs and also maximizes their profits. For example, in the yogurt category, Yoplait offers 57 different flavors in their main product lines whereas Dannon has 29.[1] Similarly, Ben and Jerry's has 50 different flavors whereas Häagen Dazs has 33 and Breyers has 29.[2] Given that these products differ in their selling prices and purchasing costs, the retailer may consider the following two extreme scenarios. First, it is possible to capture a greater segment of consumer population by making every consumer's first choice available in the product offering; but this would require a significant amount of space and prevent consumers from substituting to possibly more profitable items. Second, the retailer could offer only the most profitable item but

this may lead to reduced patronage and lost goodwill, when consumers are induced to substitute or leave the store empty-handed. Therefore, the retailer should leverage her knowledge of consumer substitution patterns and the product cost structure to strategically choose an assortment in order to divert the right amount of demand to more profitable products. However enumerating every possible subset of the 86 yogurt products and 112 flavors of ice cream would be too time-consuming.

In this work we develop efficient methods to improve a retailer's product selection process. Specifically, when all the products can be mapped onto a hierarchical ordering system such as a branched (outtree or intree), vertical (one-way substitution), and horizontal (locational) order, we obtain solution methods that boil down to a simple algorithm that uses dynamic program or the shortest path. The simple and effective algorithms we develop significantly ease a retailer's decision process and thus are pertinent. In addition, we provide some interesting insights about the structure of the optimal assortment; e.g., we show that it is not necessarily optimal to offer the most profitable product.

[1] *Sources*. http://www.yoplait.com/products/ and http://www.dannon.com/ourproducts.aspx.Information (retrieved June 3, 2011).

[2] *Sources*. http://www.benjerry.com/flavors/our-flavors/, http://www.haagen-dazs.com/products/default.aspx, and http://www.breyers.com/products/Default.aspx (retrieved June 3, 2011).

We use the ranking-based consumer choice model, wherein each customer is characterized by the list of products he is willing to buy in order of preference, as in Mahajan and van Ryzin (2001a, b), Smith et al. (2009), and Honhon et al. (2012a). This model is appealing for its ability to represent the underlying preference structure of consumer that is responsible for consumer choices. It is also very general in the sense that most of the existing consumer choice models, such as the locational choice model or the multinomial logit model, can be viewed as special cases thereof (under the assumption of fixed selling prices). Further, the ranking-based model does not suffer from the well-known issue of "independence from irrelevant alternatives" (for more details, see Anderson et al. 1992) that limits the applicability of the multinomial logit model.

Despite its appeal and generality, the ranking-based model is not without limitations. It is more difficult to estimate compared to the multinomial logit model. Although we do not address the question of estimation in this model we point the readers to recent papers that provide direction toward estimating such models; Farias et al. (2011), for example, propose a nonparametric approach to estimate the ranking-based model from limited data, and Yunes et al. (2007) successfully apply the ranking-based model to the design of product lines at Deere & Company. If conjoint survey results are available, as in Deere & Company, it is relatively straightforward to come up with consumer types and their ranking lists based on the part-worth utilities calculated from the survey (see Green and Krieger 1987 for illustration and Belloni et al. 2008 for a recent application).

Another limitation of the ranking-based model is that it does not easily lend itself to the study of endogenous pricing. Hence, we assume that selling prices are fixed. It is not uncommon for retailers to adopt an MSRP (manufacturer's suggested retail price) price strategy; evidence from the data collected by Carlton and Chevalier (2001) based on the fragrance product market shows that a number of stores (namely, upscale beauty and department stores) are charging the MSRP for the products they sell. Although contractually adhering to MSRP is rare, supplying exclusive products to retailers who adopt an across-the-board no-discounting policy is not unheard of (see Carlton and Chevalier 2001). Literature in the operations management area also lends support to the assumption of exogenous prices (see, for example, van Ryzin and Mahajan 1999, Smith and Agrawal 2000, Gaur and Honhon 2006).

Our problem is similar to the "*seller's* (*welfare*) *problem*" from the marketing literature described in Green and Krieger (1985) where a firm has to select products from a set of candidate items (called the "reference set") often from different manufacturers. This problem is related to the product line design problem wherein a manufacturer typically chooses from a continuum of product attributes over which consumer preferences are defined. Without any restrictions on the set of possible consumer types, this problem is known to be NP-hard (see Green and Krieger 1985), and a number of authors have suggested heuristics and solution methods (see, for example, Dobson and Kalish 1988, McBride and Zufryden 1988, Belloni et al. 2008, among others). Our contribution is most relevant in the context of the retail assortment selection problem and differs from the above in the explicit consideration of the impact of operational factors such as fixed, substitution, and lost sale penalty costs for four special cases of the famous problem. In the operations management literature, Li (2007) and Smith and Agrawal (2000) consider a similar setting as ours but Li (2007) uses the multinomial logit model to represent consumer preferences, and Smith and Agrawal (2000) use a model based on a substitution matrix. Other researchers use the same consumer choice models as ours but make different assumptions regarding substitution and prices. We detail these papers at the start of each subsection in §3.

The rest of this paper is organized as follows. We describe our model in §2 then present our solution methods for the four consumer choice model in §3. We conclude and discuss a possible extension of our work in §4. All proofs are in the electronic companion (available at http://msom.journal.informs.org/).

2. Model

We use bold characters to represent vectors and subscripts to denote their components; e.g., τ_j is the jth component of vector τ. Sets and matrices are denoted by capital letters, and $|S|$ denotes the cardinality of set S. Consider a product category consisting of n potential products, indexed 1 to n, where $\mathcal{N} = \{1, \ldots, n\}$ and let 0 denote the no-purchase option. To define the purchase behavior of consumers we define a *consumer type*, which is a vector of products that the customer is willing to purchase, arranged in decreasing order of preference. For example, a customer of type $(1, 2, 4)$ has product 1 as his first choice, product 2 as his second choice, product 4 as his third choice, and he never purchases products that do not belong in his type (3 and 5 to n). In general, a type τ is a vector $(\tau_1, \ldots, \tau_t)$ of product indices such that $\{\tau_1, \ldots, \tau_t\} \subseteq \mathcal{N}$ and $0 \leq t \leq n$. The "null" type is such that $t = 0$ and corresponds to customers who never purchase a product from the product category. Note that the type of a customer can result from a utility maximization procedure as in Mahajan and van Ryzin (2001a).

Let $\mathcal{T}$ be the set of all possible types; i.e., $\mathcal{T} = \{\tau = (\tau_1, \ldots, \tau_t): 0 \leq t \leq n, \tau_t \in \mathcal{N} \text{ and } \tau_1 \neq \tau_2 \neq \cdots \neq \tau_t\}$.

In addition, let α_τ be the proportion of customers of type $\tau \in \mathcal{T}$ in the customer population, such that $\sum_{\tau \in \mathcal{T}} \alpha_\tau = 1$, and $\mathcal{T}^+ = \{\tau \in \mathcal{T}: \alpha_\tau > 0\}$ be the set of consumer types that exist in the population. In practice, the number of types that exist in a population is clearly less than or equal to the theoretically possible number of types; i.e., $|\mathcal{T}^+| \le |\mathcal{T}| = \sum_{j=0}^{n} C_n^j j! = \sum_{j=0}^{n}(n!/(n-j)!)$. For example, for $n = 10$, there is a total of 9,864,100 possible types! In §3 we consider four possible values of $\mathcal{T}^+$, which correspond to four practically motivated consumer choice models.

In our *consumer-driven substitution* model, a customer of type τ for whom product $j \in S \subseteq \mathcal{N}$ is the kth choice, i.e., such that $\tau_k = j$, picks product j from set S if and only if products $\tau_1, \ldots, \tau_{k-1}$ do not belong to set S. Let $P_j^k(S)$ denote the proportion of customers who pick product j from assortment S as the kth choice. We have

$$
P_j^k(S) = \begin{cases} \sum_{\substack{\tau=(\tau_1,\ldots,\tau_t):\ t \ge k,\ \tau_k = j, \\ \tau_1,\ldots,\tau_{k-1} \notin S}} \alpha_{(\tau_1,\ldots,\tau_t)} & \text{if } j \in S \\ 0 & \text{otherwise.} \end{cases}
\tag{1}
$$

When $k > n - |S| + 1$, a product cannot be picked as a customer's kth choice because one of his first $k-1$-th choices must be offered in S. Let $P_0(S) = 1 - \sum_{j \in S}\sum_{k=1}^{n-|S|+1} P_j^k(S)$ denote the proportion of customers who do not pick anything from set S.

Let $\pi_j = (r_j - c_j)$ denote the profit margin on a product j, where r_j and c_j denote the selling price and variable cost of product j, respectively. Retailers generally incur a *fixed cost*, denoted by the parameter K, of carrying a product in their assortment. In addition, we assume that the retailer bears a *substitution penalty cost* for every customer who purchases a product that is not his first choice, and that the penalty is higher when he purchases a lower ranked product. More specifically, if a customer buys his kth favorite product, the penalty is $f(k)$ where f is a nondecreasing function. Note that, unless otherwise stated, it is possible that $\pi_j - f(k) < 0$ for some $j, k \in \{1, \ldots, n\}$, meaning that the retailer may lose money when a customer buys product j as their kth choice. Also let p denote the *lost sale penalty cost* that measures the disutility experienced by a customer who leaves the store empty-handed. We do not make any assumption regarding the relative values of $f(k)$ and p. However, one might argue that $p \ge f(n)$ makes sense for most product categories, i.e., the penalty for not satisfying a customer is larger than the penalty for making him substitute to his least favorite product.

Given the consumers' choice structure and the price and cost parameters of the products, the retailer's problem is to find the optimal assortment S^* such that $\Pi(S^*) = \max_{S \subseteq \mathcal{N}} \Pi(S)$, where $\Pi(S)$, the profit associated with the assortment S, is given by

$$
\Pi(S) = \sum_{j \in S}\sum_{k=1}^{n-|S|+1} P_j^k(S)[\pi_j - f(k)] - P_0(S)p - K|S|
$$
$$
= \sum_{j \in S}\sum_{k=1}^{n-|S|+1} P_j^k(S)[\pi_j + p - f(k)] - p - K|S|. \tag{2}
$$

We refer to $\sum_{k=1}^{n-|S|+1} P_j^k(S)[\pi_j + p - f(k)] - K$ as the *contribution* of product j to the profit of assortment S.

To make the exposition simple we ignore inventory effects. However, our results also apply in the presence of inventory considerations under the assumptions of assortment-based substitution and continuous store traffic (see part II of the electronic companion for a proof of the equivalence of the two settings).

In theory, it is always possible to find an optimal assortment by enumeration, that is, by computing the profit associated with every set $S \subseteq \mathcal{N}$ and looking for the highest profit value. For each assortment S, computing the profit requires the calculation of $P_j^k(S)$ for $j \in S$ and $k = 1, \ldots, n$. To calculate $P_j^k(S)$, we need to examine all the products in each customer type $\tau = (\tau_1, \ldots, \tau_t) \in \mathcal{T}^+$, which takes $O(|\mathcal{T}^+|n)$ because there are $|\mathcal{T}^+|$ types to consider and the maximum number of products in a type is n. This means that, if every possible consumer type exists in the population, i.e., $|\mathcal{T}^+| = |\mathcal{T}|$, the complexity of the enumeration method is $O(n(2n)^n)$ and that it decreases with the number of types in $\mathcal{T}^+$. However, the number of assortments to consider is always 2^n and thus the complexity of the enumeration method is always exponential in n. We find that it takes over 13 hours to obtain the optimal assortment via enumeration for a problem with $n = 20$ and intree preferences for which there are only 20 different types (see Table 2 in §3.4). As discussed in §1, retailers are generally looking at more than 20 potential products in a given product category. Moreover, Honhon et al. (2012b) show that simple heuristics (such as greedy) do not generally produce an optimal solution and can have very poor worst-case bounds. Hence, there is a need for efficient solution methods to find the optimal assortment.

In the next section we consider four practically motivated special cases of the ranking-based consumer choice models, namely, the one-way substitution model, the locational choice model, the outtree model, and the intree model, and we obtain efficient solution methods to find the optimal assortment.

3. Optimization

3.1. One-Way Substitution
We first consider a model that typically applies to vertically differentiated products, which differ in one

quality-defining attribute (Pan and Honhon 2012). The *one-way substitution* model is also referred to as the *downward* or *upward substitution model* in the literature. A downward substitution model (Bassok et al. 1999) is appropriate when higher quality products could be substituted for lower quality products, but not vice versa, such as circuits with different performance characteristics in the semiconductor industry and steel beams with different strength in the steel industry. Pentico (1974) uses dynamic programming to solve for the optimal inventory levels under downward substitution. They assume that the firm allocates products to customers and that customers are served in increasing order of their willingness to substitute. Bassok et al. (1999) generalize this problem to allow the retailer to choose the order in which to serve customers. As opposed to the above papers where firm-driven substitution is used, we use consumer-driven substitution.

In the one-way substitution model, the ordering of products is the same for all consumers but the heterogeneity in consumers arises from the fact that each consumer type may differ in their first choice and the number of products in their type. For example, if $n = 3$, customers can only be of the following types: $(1), (2), (3), (1, 2), (2, 3), (1, 2, 3)$. In general we have $\mathcal{T}^+ \subseteq \{\tau = (\tau_1, \ldots, \tau_t) \in \mathcal{T}: \tau_k = \tau_{k-1} + 1$ for $k = 2, \ldots, t\}$. Hence, $|\mathcal{T}^+| \leq n(n+1)/2$. In this setting, for notational convenience, we use (τ_1, τ_t) to denote type $(\tau_1, \ldots, \tau_t)$. With this model, the proportion of customers who buy product $j \in S$ as their kth choice for $k \leq j$ is $P_j^k(S) = \sum_{i=j}^n \alpha_{(j-k+1, i)}$ if $j-k+1, \ldots, j-1 \notin S$ and 0 otherwise. Note that $P_j^k(S)$ has a special structure in this model: to compute the proportion of customers who would buy product j, it is sufficient to know which product in S has the highest index lower than j.

LEMMA 1. *Let $S = \{s_1, \ldots, s_m\}$ with $s_1 < \cdots < s_m$. For $k = 1, \ldots, n$, we have*

$$P_{s_j}^k(S) = \begin{cases} P_{s_j}^k(\{s_j\}) & \text{if } j = 1, \\ P_{s_j}^k(\{s_{j-1}, s_j\}) & \text{if } 2 \leq j \leq m. \end{cases} \quad (3)$$

Lemma 1 indicates that the proportion of customers purchasing product s_j and thus the contribution of product s_j to the profit of assortment S depends only on the adjacent product s_{j-1}. This property enables us to develop a shortest path algorithm to identify the optimal assortment. To elucidate this method, we construct a graph $G = (V, A)$, where V is the set of nodes and A is the set of arcs. The set of nodes is $A = \{0, 1, \ldots, n, n+1\}$, where zero is the source node, $n + 1$ is the destination node, node i for $i = 1, \ldots, n$ corresponds to product i, and all the arcs from node $i \in V$ to $j \in V$, $i < j$, belong to set A. Following arc

(i, j) implies that product j is the next product to be added to the assortment after product i, meaning that products $i+1, \ldots, j-1$ are not added. The cost of arc (i, j) is proportional to contribution of product j to profit of the assortment as follows:

$c_{(i, j)}$

$$= \begin{cases} K - \sum_{k=1}^{j} P_j^k(\{j\})[\pi_j + p - f(k)] & \text{if } i = 0, 1 \leq j \leq n, \\ K - \sum_{k=1}^{j-i} P_j^k(\{i, j\})[\pi_j + p - f(k)] & \text{if } 0 < i < j \leq n, \\ p & \text{if } 0 \leq i < j = n+1. \end{cases}$$

We find the optimal assortment by solving the shortest path problem from the source node to the destination node.

PROPOSITION 1. *The problem of finding an optimal assortment reduces to a shortest path problem between nodes 0 and $n + 1$.*

LEMMA 2. *The complexity of the enumeration method for the one-way substitution model is $O(n^3 2^n)$. The complexity of the shortest path algorithm for one-way substitution is $O(n^4)$.*

Note that the shortest path method has been used by Alptekinoğlu and Corbett (2010) to identify trade-off between variety and lead time, and Pan and Honhon (2012) for vertically differentiated products in the absence of substitution and lost sale costs. Example 1 illustrates our solution method.

EXAMPLE 1. Let $n = 3$, $\mathcal{T}^+ = \{(1, 2, 3), (1, 2), (2, 3)\}$ and $\alpha_{(1, 2, 3)} = 1/4$, $\alpha_{(1, 2)} = 1/2$ and $\alpha_{(2, 3)} = 1/4$. Let $\pi_1 = 20$, $\pi_2 = 15$ and $\pi_3 = 10$. Let $f(k) = (k - 1)$ for $k = 1, \ldots, n$ and $p = 1.5$ and $K = 3$. We construct the graph for the shortest path problem as shown in Figure 1. Each path in the graph corresponds to an assortment and the length of each path corresponds to the profit for each assortment. For example, path $0 \to 1 \to 4$ corresponds to assortment $\{1\}$ with profit 11.625, and path $0 \to 1 \to 2 \to 3 \to 4$ corresponds to assortment $\{1, 2, 3\}$ with profit 9.75. Solving the shortest path

Figure 1   Graph for Example 1
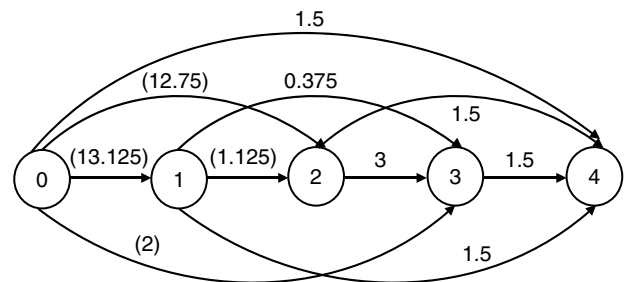
problem from 0 to 4 gives the optimal assortment $S^* = \{1, 2\}$ with the optimal profit of 12.75.

3.2. One-Dimensional Locational Choice Model

Having developed an efficient solution method for the one-way substitution model that applies to vertically differentiated products, we now focus on the *locational choice model* (or *Lancaster demand model* or *Hotelling's model*). This model applies to product categories in which products are horizontally differentiated on one attribute, e.g., shirts of different colors, yogurt with different flavor. Previous work on the locational choice model either studied a manufacturer's joint pricing and product line decisions (for example, Chen et al. 1998, Alptekinoğlu and Corbett 2008), or solely product line decision (such as Gaur and Honhon 2006) where the firm can choose from a continuum of locations.

We first present the locational choice model as it is traditionally presented, using the distribution of products and consumer locations on a horizontal line; we then show that it can be mapped to our ranking-based preference model. This one-to-one mapping allows us to present our results for the locational choice model with ease.

Let $\mathcal{L}$ be a one-dimensional attribute space. Customers are characterized by the location of their most preferred product on this attribute space. Let G be the distribution of customer locations on the attribute space. Let $x_j \in \mathcal{L}$ be the location of product j on the attribute space for $j = 1, \ldots, n$. A customer located at $x \in \mathcal{L}$ gets utility U(x,j) from buying product j,

$$U(x, j) = Z_j - r_j - d|x - x_j|, \qquad (4)$$

where Z_j is the reservation price for product j and $d > 0$ is the disutility associated with a distance of 1 between the customer and product locations. We assume that there are no two products with identical values of x_j, Z_j, and r_j. Without loss of generality, let the no-purchase utility be equal to 0. A customer is of type $\tau = (\tau_1, \ldots, \tau_t)$ if $U(x, \tau_1) \geq U(x, \tau_2) \geq \cdots \geq U(x, \tau_t) \geq 0$. Hence, we have $\alpha_\tau = \int_{x \in \mathcal{L}} I_{U(x, \tau_1) \geq U(x, \tau_2) \geq \cdots \geq U(x, \tau_t) \geq 0} \, dG(x)$, where $I(\cdot)$ is the indicator function. We show in the proof of Lemma 5 that $|\mathcal{T}^+| \leq (n^2 + 3n - 2)/2$. Let $L_j = (Z_j - r_j)/d$ be the maximum distance between x_j and a customer who gets a nonnegative utility from product j.

EXAMPLE 2. Let $n = 3$, $\mathcal{L} = [0, 1]$, $d = 10$, and G is uniform on $[0, 1]$. Let $(x_1, x_2, x_3) = (0.3, 0.2, 0.7)$, $(Z_1, Z_2, Z_3) = (75, 62, 206)$, $(r_1, r_2, r_3) = (72, 61, 203)$, and $(c_1, c_2, c_3) = (22, 21, 23)$. Let $f(k) = (k - 1)$ for $k = 1, \ldots, n$ and $p = 1.5$ and $K = 3$. This implies that $(L_1, L_2, L_3) = (0.3, 0.1, 0.3)$ and $(\pi_1, \pi_2, \pi_3) = (50, 40, 180)$. The attribute space is shown in Figure 2, where a customer location is displayed on the X-axis
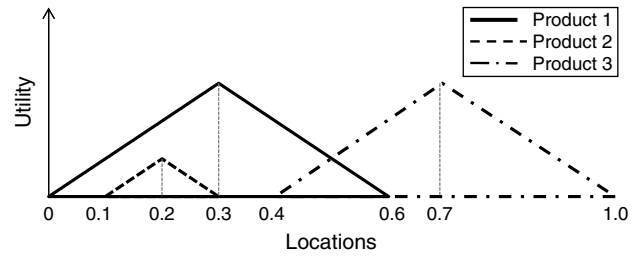
and the utility that a customer gets from products is displayed on the Y-axis. The proportion of customers of type τ, α_τ, is equal to 0.2, 0.2, 0.1, 0.1, 0.4, respectively, for types (1), (1, 2), (1, 3), (3, 1), and (3).

Suppose, without loss of generality, that the products are numbered such that $x_1 - L_1 \leq x_2 - L_2 \leq \cdots \leq x_n - L_n$ with ties broken arbitrarily. Consider $i, j \in \mathcal{N}$, such that $i < j$. If $x_i + L_i \geq x_j + L_j$, then every customer (weakly) prefers product i to product j. The following result shows that there exists an optimal assortment that does not contain two such products.

LEMMA 3. *Let the products be numbered such that $x_1 - L_1 \leq x_2 - L_2 \leq \cdots \leq x_n - L_n$. There exists an optimal assortment S such that for all $i, j \in S$ with $i < j$, $x_i + L_i < x_j + L_j$.*

It follows that we can consider only assortments $S = \{s_1, \ldots, s_m\}$ with $s_1 < \cdots < s_m$, such that $x_{s_j} + L_{s_j} < x_{s_{j+1}} + L_{s_{j+1}}$ for $j = 1, \ldots, m - 1$. As in the one-way substitution model, we show that the $P_j^k(S)$ function has an interesting property: it is sufficient to know the nearest neighbors (on the right and left on the attribute space) of a product to compute the proportion of customer who will purchase the product. We exploit this structure below to obtain an efficient method to find the optimal assortment.

LEMMA 4. *Let $S = \{s_1, \ldots, s_m\}$ with $s_1 < \cdots < s_m$ such that $x_{s_j} + L_{s_j} < x_{s_{j+1}} + L_{s_{j+1}}$ for $j = 1, \ldots, m - 1$. For $k = 1, \ldots, n$, we have*

$$P_{s_j}^k(S) = \begin{cases} P_{s_j}^k(\{s_j, s_{j+1}\}) & \text{if } j = 1, \\ P_{s_j}^k(\{s_{j-1}, s_j, s_{j+1}\}) & \text{if } 2 \leq j \leq m - 1, \\ P_{s_j}^k(\{s_{j-1}, s_j\}) & \text{if } j = m. \end{cases} \quad (5)$$

Given Lemma 4 we are able to model the problem of finding the best assortment as a shortest path problem. We construct a graph $G = (V, A)$, where V is the set of nodes and A is the set of arcs. Node V consists of pairs of nodes (i, j) such that $i < j$ and $x_i + L_i < x_j + L_j$. A node $(i, j) \in V$ indicates that product i can be offered along with j in the optimal assortment, according to Lemma 4. We also introduce two fictitious nodes: source node $(0, 0)$ and destination node $(n + 1, n + 1)$. An arc from node $(i, j) \in V$

to $(j', l) \in V$ belongs to set A if and only if $j = j'$, which means that product j is the next product to be offered in the assortment after product i and before product l, meaning that products $i+1, \ldots, j-1$ and $j+1, \ldots, l-1$ are not offered. The cost of the arc from (i, j) to (j, l) is proportional to the contribution of product j to the profit of the assortment, which can be computed as follows:

$$c_{(i,j),(j,l)}$$

$$= \begin{cases} 0 & \text{if } j = 0 \\ K - \sum_{k=1}^{n} P_j^k(\{j, l\})[\pi_j + p - f(k)] \\ & \text{if } 0 = i < j < l < n+1, \\ K - \sum_{k=1}^{n} P_j^k(\{i, j, l\})[\pi_j + p - f(k)] \\ & \text{if } 0 < i < j < l < n+1, \\ K - \sum_{k=1}^{n} P_j^k(\{i, j\})[\pi_j + p - f(k)] \\ & \text{if } 0 < i < j < l = n+1, \\ K - \sum_{k=1}^{n} P_j^k(\{j\})[\pi_j + p - f(k)] \\ & \text{if } 0 = i < j < l = n+1, \\ p & \text{if } j = n+1. \end{cases}$$

Proposition 2. *The problem of finding an optimal assortment reduces to a shortest path problem between* $(0, 0)$ *and* $(n+1, n+1)$.

Lemma 5. *The complexity of the enumeration method for the locational choice model is* $O(n^3 2^n)$ *and the complexity of the shortest path algorithm for one-dimensional locational choice model is* $O(n^6)$.

The following example illustrates the application of our shortest path method to the locational choice model.

Example 2 (Continued). To use the shortest path algorithm, we construct the graph shown in Figure 3. Note that node $(1, 2)$ does not belong to A because $x_1 + L_1 = 0.6 > x_2 + L_2 = 0.3$, which does not satisfy the condition of Lemma 3. Hence, there is no path corresponding to assortments $\{1, 2\}$ and $\{1, 2, 3\}$. Each path in the graph corresponds to an assortment and the length of each path corresponds to the profit for each assortment. For example, path $(0, 0) \rightarrow (0, 2) \rightarrow (2, 4) \rightarrow (4, 4)$ corresponds to assortment $\{2\}$ with profit 3.6, and path $(0, 0) \rightarrow (0, 2) \rightarrow (2, 3) \rightarrow (3, 4) \rightarrow (4, 4)$ corresponds to assortment $\{2, 3\}$ with profit 109.4. Solving the shortest path problem from $(0, 0)$ and $(4, 4)$ gives the optimal assortment $S^* = \{2, 3\}$ with the optimal profit of 116.

Figure 3   Graph for Example 2
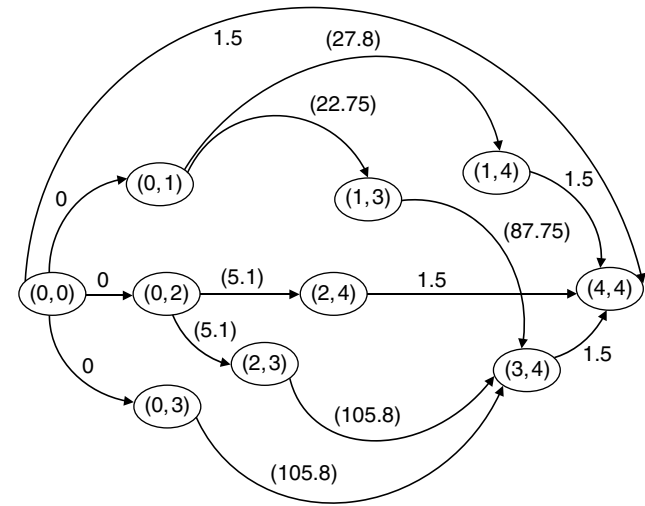
From Example 2 we see that it is impossible to eliminate *dominated* products, i.e., products that are less preferred than a more profitable product by all the customers, because they could be included in the optimal solution. To see this, note that every customer prefers product 1 to product 2 because $y_1 + L_1 = 0.6 > y_2 + L_2 = 0.3$ (graphically, the triangle corresponding to product 2 on Figure 2 is inside the one that corresponds to product 1) and that product 1 is more profitable than product 2 because $\pi_1 > \pi_2$. Yet, the optimal assortment contains product 2.

3.3.   Outtree Preferences

So far we have analyzed stylized models of consumer preferences that could be represented by a vertical or a horizontal preference ordering. We now analyze a model of consumer choice with *nested preferences*. *Nested preferences* models have been traditionally used with extensions of the multinomial logit model. The model we present here, called the *outtree preferences* model, is different in that it assumes that every customer has the same first choice. The outtree model applies to categories of products with nested sets of vertically differentiated attributes. For example, if the product category is toothpaste, product 1 on Figure 4 could be a toothpaste with whitening power and cavities and tartar protection, product 2 could be a toothpaste with whitening power and tartar protection, product 3 would be a toothpaste with cavities protection only, product 4 would be a toothpaste with whitening power only, and product 5 would be a toothpaste with tartar protection only. These products are often offered at the same price by retailers. The outtree preferences model also applies, for example, to all-purpose cleaners with or without the following properties: antibacterial, heavy-duty, grease cutting, deodorizing, etc. This model has not been studied by other researchers.

Figure 4   Example of Outtree
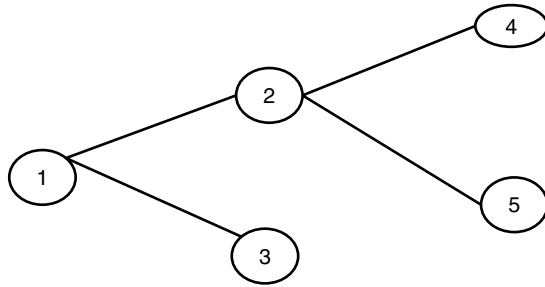
In the outtree preference model, the set of possible consumer types can be represented by an outtree, i.e., a directed graph in which there is a single initial node and a unique directed path from the initial node to any other node. Nodes represent products and each path from the initial node to a node represents a consumer type. Let $l(j)$ denote the *level* of node j, which is equal to the number of predecessors plus one. Let $\mathscr{P}_j$ be the set of all predecessors of node j. Let $\mathscr{S}_j$ be the set of *immediate* successors of node j and $\mathscr{Y}_j$ be the set of all successors of node j, plus j itself. Without loss of generality, let the products be numbered such that $j < k$ implies that $l(j) \leq l(k)$ for $j, k \in \{1, \ldots, n\}$, so that 1 is the root of the tree. We have $\mathscr{T}^+ \subseteq \{\tau = (\tau_1, \ldots, \tau_t) \in \mathscr{T} : \tau_1 = 1$ and $\tau_1 \to \tau_2 \to \cdots \to \tau_t$ is a directed path$\}$ so that $|\mathscr{T}^+| \leq n$.

Figure 4 shows an example of an outtree, where the possible types are (1), $(1, 2)$, $(1, 2, 4)$, $(1, 2, 5)$, $(1, 3)$ and we have $l(1) = 1$, $l(2) = l(3) = 2$ and $l(4) = l(5) = 3$.

Notice that product j is only ever purchased as a $l(j)$th choice by a customer, and thus the substitution penalty cost associated with product j is always $f(l(j))$. Also a product gets positive demand if and only if none of its predecessors are offered in the assortment; i.e., for $j \in S$,

$$P_j^{l(j)}(S)$$
$$= \begin{cases} P_j^{l(j)}(\{j\}) = P_j(\{j\}) = \displaystyle\sum_{\tau : \tau_{l(j)} = j} \alpha_\tau & \text{if } S \cap \mathscr{P}_j = \varnothing, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

We use this property to develop Algorithm 1 to obtain the optimal assortment. Let $V(j)$ be the maximum profit obtained when only products from $\mathscr{Y}_j$ can be included in the optimal assortment and $S^*(j)$ be the corresponding optimal assortment. The value function $V(j)$ is defined as

$$V(j)$$
$$= \max\left\{ P_j(\{j\})[\pi_j - f(l(j))] - K, \sum_{k \in \mathscr{S}_j} V(k) - \alpha_{(1, \ldots, j)} p \right\}.$$

Algorithm 1 (Dynamic Program for the Outtree Preferences Model)
for all $j = n, n-1, \ldots, 1$ **do**
   Set: $V(j) = \sum_{k \in \mathscr{S}_j} V(k) - \alpha_{(1, \ldots, j)} p$
   if $P_j(\{j\})[\pi_j - f(l(j))] - K > V(j)$ **then**
      $V(j) = P_j(\{j\})[\pi_j - f(l(j))] - K$ and $S^*(j) := \{j\}$
   else
      $S^*(j) := \bigcup_{k \in \mathscr{S}_j} S^*(k)$
   end if
end for

PROPOSITION 3. *Assortment $S^*(1)$ given by Algorithm 1 is optimal and $V(1)$ is the optimal profit value.*

LEMMA 6. *The complexity of the enumeration method for the outtree preferences model is $O(n^2 2^n)$. The complexity of Algorithm 1 is $O(n)$.*

We illustrate Algorithm 1 using Example 3.

EXAMPLE 3. Consider the outtree of Figure 4. Assume the five types are equally likely, i.e., $\alpha_\tau = 0.2$ for $\tau \in \mathscr{T}^+$. Let $\pi_1 = 4$, $\pi_2 = 8$, $\pi_3 = 17$, $\pi_4 = 30$, and $\pi_5 = 6$. Set $f(k) = k - 1$, $p = 1.5$, and $K = 3$. The optimal assortment is $\{3, 4\}$. Note that the optimal assortment contains more than one product but these products are such that they do not belong to the same type in $\mathscr{T}^+$.

3.4.   Intree Preferences
The natural counterpart to the *outtree preferences* model is *intree preferences* model. In this model, all customers have the same last choice but they differ in the number of products in their preference list and therefore in their first choice as well. The intree model applies to categories of products with nested sets of horizontally differentiated attributes. For example, if the product category is shampoo, product 5 on Figure 5 could be an all-purpose shampoo, products 3 and 4 could be shampoo for fine hair and thick hair, respectively, product 1 could correspond to shampoo for fine and dry hair, and product 2 to shampoo for fine and greasy hair. A customer with fine hair but neither dry nor greasy would have type $(3, 5)$. Other applications for the intree preferences model include lotions for different skin types and pet food for different age and type of pet (e.g., kitten versus senior cat, indoor versus outdoor cat, etc.). To the best of our knowledge this consumer choice model has not yet been studied in the literature.

Figure 5   Example of Intree
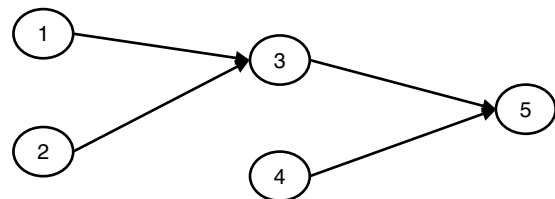
In the intree preference model, the set of possible consumer types can be represented by an intree, i.e., a directed graph in which a number of initial nodes are connected to a single end node and there is a unique directed path from any node to the end node. Nodes represent products and each path from a node to the end node in this tree represents a consumer type. Let $\mathcal{Y}_j$ and $\mathcal{P}_j$ be defined as in the outtree model. Also, let $l(j)$ be the *level* of node j, defined as the maximum number of nodes in a path to the end node minus the number of successors of j. Without loss of generality, let the products be numbered such that $j < k$ implies that $l(j) \leq l(k)$ for $j, k \in \{1, \dots, n\}$ so that the end node corresponds to product n. We have $\mathcal{T}^+ \subseteq \{\tau = (\tau_1, \dots, \tau_t) \in \mathcal{T}: \tau_t = n \text{ and } \tau_1 \to \tau_2 \to \cdots \to \tau_t \text{ is a directed path}\}$ and $|\mathcal{T}^+| \leq n$.

Figure 5 shows an example of an intree, where the possible types are (5), (4, 5), (3, 5), (1, 3, 5), (2, 3, 5) and we have $l(1) = l(2) = 1, l(3) = l(4) = 2$ and $l(5) = 3$.

In this model the proportion of customers buying product j is calculated as follows:

$$
P_j^k(S) = \begin{cases} \displaystyle\sum_{\substack{\tau = (\tau_1, \dots, \tau_t): t \geq l(n) - l(j) + 1 \\ \tau_{t-l(j)+1} = j}} \alpha_\tau & \text{if } S \cap \mathcal{P}_j = \varnothing, \\ 0 & \text{otherwise.} \end{cases} \tag{7}
$$

We make a few additional assumptions in order to obtain an effective solution method for this model. First we assume that $f(n) \leq \min_i \pi_i$, that is, even the least profitable product is sold at a profit to a customer as his nth choice. Also we assume that $f(k) = (k-1)b$ with $b \geq 0$, i.e., the substitution penalty cost is linear. These assumptions are not too restrictive; similar conditions on the cost parameters have been assumed by other researchers such as Pentico (1976) and Bassok et al. (1999).

If $K = 0$, Algorithm 2 reaches the optimal solution.

Algorithm 2 (Algorithm for the Intree Preferences Model When $K = 0$)
Set $S_0^* := \{n\}$.
for $j = n - 1, n - 2, \dots, 1$ **do**
  if $\pi_j - l(j)b = \max_{i \in \mathcal{Y}_j}\{\pi_i - l(i)b\}$ **then** $S_0^* := S_0^* \cup \{j\}$
  end if
end for

PROPOSITION 4. *Assortment S_0^* generated by Algorithm 2 is optimal when $K = 0$.*

When $K = 0$, it is optimal to offer product n because each customer prefers it to purchasing nothing, and no customer prefers it to another product and thus product n never cannibalizes a product with a higher profit margin. Other products are added if their profit margin (adjusted for the substitution penalty cost) is larger than that of all of their successors. Note that the

optimal assortment is independent of the α_τ values. Also the assortment generated by Algorithm 2 may include products that are not purchased by any customer, but this does not hurt profit because the fixed cost is equal to 0.

Finding the optimal assortment for $K > 0$ is more complex because a product should not only have a higher profit margin (adjusted for the substitution penalty cost) than its successors, it should also generate enough additional profit to recover the fixed cost K. We first compare the optimal assortment for $K = 0$ and $K > 0$.

LEMMA 7. *The optimal assortment for $K > 0$ is a subset of the optimal assortment when $K = 0$ as generated by Algorithm 2, with the same parameters other than K.*

This is because products that are not included in the optimal assortment when $K = 0$ are dominated in terms of profit by one of their successors and this does not change when the fixed cost is positive. However, products that are included in the optimal assortment when $K = 0$ may need to be removed when $K > 0$ because the extra profit they bring in comparison to their successors is not enough to recover the fixed cost K. We use Lemma 7 to construct Algorithm 3, which reaches the optimal solution when $K \geq 0$.

Algorithm 3 (Algorithm for the Intree Preferences Model When $K \geq 0$)
Run Algorithm 2 to obtain S_0^*.
Set: $N_{n+1} = 1, S_1 = \varnothing$.
for $j = n, n - 1, \dots, 1$ **do**
  $N_j = N_{j+1}$.
  if $j \in S_0^*$ **then**
    for $i = 1, \dots, N_{j+1}$ **do**
      if $\Delta^-(j, S_i) = \Pi(S_i \cup \mathcal{P}_j \cup \{j\}) - \Pi(S_i \cup \mathcal{P}_j) > 0$
        then $S_i := S_i \cup \{j\}$.
      else
        if $\Delta^+(j, S_i) = \Pi(S_i \cup \{j\}) - \Pi(S_i) > 0$
          then $N_j := N_j + 1, S_{N_j} := S_{N_j} \cup \{j\}$.
      end if
    end if
    end for
  end if
end for
Calculate $\Pi(S_i)$ for $l = 1, \dots, N_1$.
  Let $S^* = \arg\max_{l=1, \dots, N_1} \Pi(S_i)$.

PROPOSITION 5. *Assortment S^* generated by Algorithm 3 is optimal.*

The general idea behind the algorithm is as follows. The products are examined in reverse order of their index, i.e., from product n to product 1. After each iteration of the first "for" loop, a number of sets, called *candidate assortments*, are retained. For example, after looking at nodes n to $j + 1$, we have N_{j+1}

Table 1   Constructions of the Candidate Assortments in Example 4

Iteration	j	N_j	S_1	S_2	S_3
0		1	$\varnothing$		
1	5	2	$\varnothing$	{5}	
2	4	2	{4}	{4, 5}	
3	3	3	{3, 4}	{4, 5}	{3, 4, 5}
4	2	3	{3, 4}	{2, 4, 5}	{3, 4, 5}

candidate assortments denoted $S_1, \ldots, S_{N_{j+1}}$. The candidate assortments are such that $S_i \subseteq \{j+1, \ldots, n\}$ for $i = 1, \ldots, N_{j+1}$ and one of them is the optimal subset of $\{j+1, \ldots, n\}$ to offer. In the jth iteration of the algorithm, we study whether or not product j should be added to each of these candidate assortments. To examine whether or not product j should be added to set S_i for $i \in \{1, \ldots, N_{j+1}\}$, we need to evaluate the contribution of product j to the profit of an assortment S for $S \supseteq S_i$ and $S \backslash S_i \subseteq \{1, \ldots, j-1\}$, that is, after other products are added to set S_i. This contribution is equal to $\Pi(S \cup \{j\}) - \Pi(S)$, which depends on what products are in $S \backslash S_i$. We then calculate bounds as follows. A lower bound is obtained when all the predecessors of product j are in the assortment, i.e., $\mathscr{P}_j \subseteq S$, and is denoted by $\Delta^-(j, S_i)$. An upper bound is obtained when none of product j's predecessors are in the assortment, i.e., $S \cap \mathscr{P}_j = \varnothing$, and is denoted by $\Delta^+(j, S_i)$ (see the proof of Proposition 5 in the electronic companion for the calculation of $\Delta^-(j, S_i)$ and $\Delta^+(j, S_i)$). If the lower bound is positive then the contribution of product j is always positive, therefore product j should be added to S_i. In this case the ith candidate assortment becomes $S_i \cup \{j\}$. If the upper bound is negative then the contribution to profit of product j is always negative, therefore product j should not be added to S_i. In this case, the ith candidate assortment remains equal to S_i. In the remaining situation, i.e., when the lower bound is negative and the upper bound is positive, it is not

clear whether or not product j should be added to S_i. In that case, the candidate assortment S_i is kept as such and a new candidate assortment is created which is equal to $S_i \cup \{j\}$ so that both cases are considered. At the end of the algorithm, we have N_1 candidate assortments $S_1, \ldots, S_{N_1}$. The optimal assortment is the one with the highest profit among these.

We illustrate Algorithm 3 with Example 4.

EXAMPLE 4. Consider the intree of Figure 5. Assume the five types are equally likely, i.e., $\alpha_\tau = 0.2$ for $\tau \in \mathscr{T}^+$. Let $\pi_1 = 8$, $\pi_2 = 21$, $\pi_3 = 11.5$, $\pi_4 = 20$, and $\pi_5 = 2$. Set $f(k) = (k-1)0.2$, $p = 0.5$, and $K = 2$. Using Algorithm 2, we obtain $S_0^* = \{2, 3, 4, 5\}$. Table 1 shows how the candidate assortments are constructed by Algorithm 3.

We compare the optimal profit of the three candidate solutions and obtain 6.72, 3.28, and 5.22, respectively, for S_1, S_2, and S_3. Hence, the optimal assortment is $S^* = S_1 = \{3, 4\}$. Notice that the optimal assortment does not include the most profitable product, which is product 2.

LEMMA 8. *The complexity of the enumeration method for the intree preference model is $O(n^2 2^n)$. The complexity of Algorithm 2 is $O(n)$ and that of Algorithm 3 is $O(n^2 2^n)$.*

The theoretical complexity of Algorithm 3, which is used when $K > 0$, is the same as that of the enumeration method. This is because the number of candidate assortments may double at every iteration. However this is unlikely to be the case in practice; our numerical study demonstrates that Algorithm 3 is much faster than the enumeration method. Table 2 shows the average computation time of our algorithms and the enumeration method, respectively. For the intree preference model we also report the average number of candidate solutions, i.e., N_1 as given by Algorithm 3. The results demonstrate that

Table 2   Average Computation Time of Our Algorithms Compared to the Enumeration Method and Average Number of Candidate Assortments Generated by Algorithm 3

	One-way		Locational		Outtree		Intree		
n	Enum.	SP algorithm	Enum.	SP algorithm	Enum.	Algorithm 1	Enum.	Algorithm 3	Avg. N_1
5	0.012 sec	0.002 sec	0.011 sec	0.003 sec	0.010 sec	0.000 sec	0.008 sec	0.000 sec	1.93
10	0.802 sec	0.014 sec	0.676 sec	0.015 sec	0.582 sec	0.001 sec	0.605 sec	0.001 sec	3.94
15	77.37 sec	0.050 sec	70.47 sec	0.065 sec	63.39 sec	0.001 sec	63.85 sec	0.004 sec	6.72
20	13.06 hrs	0.156 sec	12.76 hrs	0.227 sec	13.10 hrs	0.002 sec	13.11 hrs	0.006 sec	9.60
25	43.89 days[a]	0.336 sec	42.43 days[a]	0.440 sec	42.21 days[a]	0.002 sec	46.98 days[a]	0.012 sec	12.26
30	18.09 yrs[a]	0.679 sec	17.89 yrs[a]	0.952 sec	18.25 yrs[a]	0.003 sec	21.75 yrs[a]	0.018 sec	14.92
35	>1,000 yrs[a]	1.265 sec	>1000 yrs[a]	2.027 sec	>1,000 yrs[a]	0.004 sec	>1,000 yrs[a]	0.025 sec	17.24
40	>1,000 yrs[a]	2.100 sec	>1,000 yrs[a]	3.282 sec	>1,000 yrs[a]	0.005 sec	>1,000 yrs[a]	0.035 sec	19.22
45	>1,000 yrs[a]	3.289 sec	>1,000 yrs[a]	5.710 sec	>1,000 yrs[a]	0.006 sec	>1,000 yrs[a]	0.046 sec	21.03
50	>1,000 yrs[a]	5.036 sec	>1,000 yrs[a]	8.967 sec	>1,000 yrs[a]	0.007 sec	>1,000 yrs[a]	0.060 sec	24.28

[a]Obtained by extrapolation.

Algorithm 3 is in practice much faster than the enumeration method (e.g., more than 6 million times faster for $n = 20$!) and can handle problems of a very large size.

In all cases we used Matlab 7.6 on a Dell Precision T5500 Workstation with a 64 bit Quad Core Intel® Xeon® Processor E5530 with 2.4 GHz and 6 GB of RAM. Average computation times were calculated based on 10,000 randomly generated problems, except for the enumeration method with $n = 20$, where we generated only 10 problems because of the long computation times, and for $n > 20$, where we used extrapolation with an exponential curve.

4. Conclusion

Table 3 provides a summary of the solution methods we have obtained for the four consumer choice models we considered. In every case, the enumeration method consists of listing all 2^n possible assortments and comparing their profit values. Therefore, it is always exponential in n but its exact complexity depends on the number of existing types. We provide an efficient algorithm for the one-way substitution, locational and outtree models. For the intree model, our solution method is efficient, i.e., it has a polynomial run time, when $K = 0$. When $K > 0$ our algorithm has the same complexity as the enumeration method but it is in practice significantly faster: more than six million times faster when $n = 20$!

In terms of managerial insights we show that the optimal assortment does not necessarily contain the most profitable product and that dominated products, i.e., products which are less preferred than a more profitable product by all customers, may be included in the optimal solution.

One limitation of our choice model is that it does not capture the strength of preferences. For example, two customers of type $(1, 2)$ may differ in how much they prefer product 1 to product 2. Customer A may strongly prefer product 1 to product 2 (though both products give him a positive utility), while customer B may have only a slight preference for product 1. If both customers end up buying product 2 (because product 1 is not available) it would be reasonable to assume a greater substitution penalty for customer A than customer B. Because the ranking-based model does not distinguish between these two customers, we are unable to capture this difference. Studying this novel aspect of consumer preferences is an avenue for future research which will require the design of a new consumer choice model.

Another limitation of our work is that we assume that selling prices are fixed. This assumption is valid when the retailer chooses not to deviate from the MSRP or any other set selling price. If that is not the case, we need to use a consumer choice model where the purchase probabilities are a function of the product selling prices. The ranking-based model is no longer adequate because the type of a customer, i.e., the ranking of the n possible products, is likely to change as the selling prices change. Hence, if for a given price, the set of possible types is consistent with an outree preference model, it may no longer be the case when the prices change. The case of endogenous pricing is better analyzed with a utility-based choice model involving the definition of reservation prices. However there are two special cases wherein the rankings of the n products for a customer do not change with their selling prices (though the number of acceptable products changes): (i) when customers have the same reservation prices for all the n products, and (ii) when customers have "lexicographic" type of preferences. Under lexicographic preferences, customers pick the highest ranked product in their type for which the price is lower or equal to their reservation price. Like in this paper, we plan to exploit the structure of the preference model to develop efficient methods to obtain the optimal prices and assortment in these two special cases. This is the subject of our ongoing research.

Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://msom.journal.informs.org/.

Table 3 Complexity of Solution Methods

Consumer choice model	Maximum no. of existing types	Complexity of enumeration	Complexity of our solution method
One way substitution	$n(n+1)/2$	$O(n^3 2^n)$	$O(n^4)$
Locational choice model	$(n^2 + 3n - 2)/2$	$O(n^3 2^n)$	$O(n^6)$
Outtree preferences	n	$O(n^2 2^n)$	$O(n)$
Intree preferences	n	$O(n^2 2^n)$	$O(n^2 2^n)$[a]

[a]Assuming $f(n) \leq \min_i \pi_i$ and $f(k) = b(k-1)$, $b \geq 0$. If $K = 0$, then $O(n)$.

References

Alptekinoğlu, A., C. J. Corbett. 2008. Mass customization vs. mass production: Variety and price competition. *Manufacturing Service Oper. Management* **10**(2) 204–217.

Alptekinoğlu, A., C. J. Corbett. 2010. Leadtime-variety tradeoff in product differentiation. *Manufacturing Service Oper. Management* **12**(4) 569–582.

Anderson, S., A. De Palma, J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* MIT Press, Cambridge, MA.

Bassok, Y., R. Anupindi, R. Akella. 1999. Single-period multi-product inventory models with substitution. *Oper. Res.* **47**(4) 632–642.

Belloni, A., R. Freund, M. Selove, D. Simester. 2008. Optimizing product line designs: Efficient methods and comparisons. *Management Sci.* **54**(9) 1544–1582.

Carlton, D. W., J. A. Chevalier. 2001. Free riding and sales strategies for the Internet. *J. Indust. Econom.* **49**(4) 441–461.

Chen, F., J. Eliashberg, P. Zipkin. 1998. Customer preferences, supply-chain costs, and product-line design. T.-H. Ho, C. S. Tang, eds. *Product Variety Management: Research Advances.* Kluwer Academic, Norwell, MA, 123–144.

Dobson, G., S. Kalish. 1988. Positioning and pricing a product line. *Marketing Sci.* **7**(2) 107–125.

Draganska, M., D. C. Jain. 2005. Consumer preferences and product-line pricing strategies: An empirical analysis. *Marketing Sci.* **25**(2) 164–174.

Farias, V. F., S. Jagabathula, D. Shah. 2011. A non-parametric approach to modeling choice with limited data. Working paper, Massachusetts Institute of Technology, Cambridge.

Gaur, V., D. Honhon. 2006. Product variety and inventory decision under a locational choice model. *Management Sci.* **52**(10) 1528–1543.

Green, P. E., A. M. Krieger. 1985. Models and heuristics for product line selection. *Marketing Sci.* **4**(1) 1–19.

Green, P. E., A. M. Krieger. 1987. A simple heuristic for selecting good products in conjoint analysis. *Appl. Management Sci.* **5** 131–153.

Honhon, D., V. Gaur, S. Seshadri. 2012a. A multi-supplier sourcing problem with a preference ordering of suppliers. *Production Oper. Management.* Forthcoming.

Honhon, D., X. A. Pan, S. Jonnalagedda. 2012b. Heuristics for the assortment planning problem under ranking-based consumer choice models. Working paper. University of Texas at Austin, Austin.

Hotelling, H. 1929. Stability in competition. *Econom. J.* **39**(153) 41–57.

Li, Z. 2007. A single-period assortment optimization model. *Production Oper. Management* **16**(3) 369–380.

Mahajan, S., G. van Ryzin. 2001a. Stocking retail assortments under dynamic consumer substitution. *Oper. Res.* **49**(3) 334–351.

Mahajan, S., G. van Ryzin. 2001b. Inventory competition under dynamic consumer choice. *Oper. Res.* **49**(5) 646–657.

McBride, R. D., F. S. Zufryden. 1988. An integer programming approach to the optimal product line selection problem. *Marketing Sci.* **7**(2) 126–140.

Pan, X. A., D. Honhon. 2012. Assortment planning for vertically differentiated products. *Production Oper. Management.* Forthcoming.

Pentico, D. W. 1974. The assortment problem with probabilistic demands. *Management Sci.* **21**(3) 286–290.

Pentico, D. W. 1976. The assortment problem with nonlinear cost functions. *Oper. Res.* **24**(6) 1129–1142.

Smith, J. C., C. Lim, A. Alptekinoğlu. 2009. New product introduction against a predator: A bilevel mixed-integer programming approach. *Naval Res. Logist.* **56**(8) 714–729.

Smith, S. A., N. Agrawal. 2000. Management of multi-item retail inventory systems with demand substitution. *Oper. Res.* **48**(1) 50–64.

van Ryzin, G., S. Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. *Management Sci.* **45**(11) 1496–1509.

Yunes, T. H., D. Napolitano, A. Scheller-Wolf, S. Tayur. 2007. Building efficient product portfolios at John Deere and Company. *Oper. Res.* **55**(4) 615–629.