A Modeling Framework for Category Assortment Planning

Juin-Kuan Chong • Teck-Hua Ho • Christopher S. Tang
The NUS Business School, National University of Singapore, Singapore 119260
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104
Anderson Graduate School of Management, University of California at Los Angeles, Los Angeles, California 90095
fbacjk@nus.edu.sg • hoteck@wharton.upenn.edu • ctang@anderson.ucla.edu

The complexity of managing a category assortment has grown tremendously in recent years due to the increased product turnover and proliferation rates in most categories. It is an increasingly difficult task for managers to find an effective assortment due to uncertain consumer preferences and the exponential number of possible assortments. This paper presents an empirically based modeling framework for managers to assess the revenue and lost sales implication of alternative category assortments. Coupled with a local improvement heuristic, the modeling framework generates an alternative category assortment with higher revenue.

This framework, which consists of a category-purchase-incidence model and a brand-share model, is calibrated and validated using 60,000 shopping trips and purchase records. Specifically, the purchase-incidence model predicts the probability of an individual customer who purchases (and who does not purchase) from a given product category during a shopping trip. The no-purchase probability enables us to estimate lost sales due to assortment changes in the category. The brand-share model predicts which brand the customer chooses if a purchase incidence occurs in the category. Our brand-share model extends the classical Guadagni and Little model (1983) by utilizing three new brand-width measures that quantify the similarities among products of different brands within the same category.

We illustrate how our modeling framework is used to reconfigure the category assortment in eight food categories for five stores in our data set. This reconfiguration exercise shows that a reconfigured category assortment can have a profit improvement of up to 25.1% with 32 products replaced. We also demonstrate how our modeling framework can be used to gauge lost sales due to assortment changes. We find the level of lost sales could range from 0.9% to 10.2% for a period of 26 weeks.
(*Retailing; Product Assortment; Brand Reconfiguration; Purchase Incidence; Brand Share; Logit Model*)

1. Introduction

Due to intense competition and rapidly changing consumer tastes, many retailers experience an ever increasing turnover rate in most product categories. As Table 1 reveals, as many as one third of the products on the shelf were replaced in a two-year period. For instance, in the ice-cream category depicted in Table 1, 129 ice-cream products were introduced and 118 were removed during the two-year period. In some cases, we actually witness a substantial net increase in product variety (e.g., spaghetti sauce and yogurt). With the high turnover and huge proliferation, the complexity of managing category assortment increases vastly.

1523-4614/01/0303/0191$05.00
1526-5498 electronic ISSN

Table 1 Basic Data Description for the Product Categories

Category	Coffee	Frozen Pizza	Hotdogs	Ice Cream	Potato Chips	Regular Cereal	Spaghetti Sauce	Yogurt
Category Activities Information								
Total number of brands over five stores	47	40	38	37	29	35	41	15
Average number of brands observed per visit	22.2	15.3	14.5	15.1	12.1	14.3	17.6	5.6
Total number of SKUs over five stores	391	337	128	421	285	242	194	288
Number of products added in the two-year period	113	109	47	129	93	114	70	107
Number of products removed in the two-year period	135	96	32	118	77	75	36	51
Product Attributes Description								
Package size								
Total number	49	145	11	9	32	73	30	7
Example	26 oz	22 oz	16 oz	64 oz	6.5 oz	12 oz	30 oz	6 oz
	8 oz	20 oz	12 oz	16 oz	7 oz	18 oz	26 oz	8 oz
	24 oz	17 oz	40 oz	32 oz	6 oz	15 oz	14 oz	32 oz
Flavor/Ingredient								
Total number	90	71	15	145	31	45	31	74
Example	Regular	Sausage	Beef	Vanilla	Regular	Corn	Plain	Plain
	Columbian	Cheese	Chicken & Pork	Neapolitan	BBQ	Wheat Bran	Italian Garden	Strawberry
	Kenya	Deluxe	Pork & Turkey	Chocolate	Sour Cream & Onion	Rice	Tomato & Herb	Raspberry

The increased complexity in category management has direct profit implications. The most direct impact on cost is the additional inventory cost due to (1) inventory obsolescence (because of high turnover rate) and (2) either more frequent stockout or more buffer stock being carried (because of limited shelf space or higher demand uncertainty). The Efficient Consumer Response (ECR) initiative is an effort to reduce the additional inventory cost through information sharing. Under the ECR initiative, retailers share point-of-sale information with manufacturers in return for lower wholesale prices. The point-of-sale information allows the manufacturers to better gauge the end-user demand and to streamline their production and distribution processes. This helps to reduce the chance of inventory obsolescence and to lower buffer stock requirement. The reader is referred to Cachon and Fisher (1997, 2000) for empirical studies on how ECR can lead to inventory cost reduction, and Lee et al. (2000) for an analytical study that quantifies the value of such information sharing.

The risk of having an ''ineffective'' category assortment also increases with the complexity of the category management task. An ineffective assortment is one that does not cater to the needs of its customers, and thus may affect the revenue of a retailer. A category assortment which is perceived to offer low variety may affect the store traffic negatively, resulting in reduced store revenue (c.f., Hoch et al. 1999).

In this paper, we are concerned with the direct effects a category assortment has on its *category revenue.* There are two ways an ineffective assortment may adversely affect the category revenue: (1) the ineffective assortment can shift demand from high-margin brands to low-margin brands, and (2) the ineffective assortment may reduce the total category sales. Any modeling framework that assesses the impact of assortment on revenue would have to capture its impact on brand-share distribution as well as on category sales changes. The modeling framework that we propose in this paper does so through two components: a brand-share model with three brand-level assortment (brand-width) measures embedded and a purchase-incidence model with a category-level assortment measure incorporated. Our brand-share model extends the classic brand-share model of Guadagni and Little (1983) by incorporating three brand-width measures that capture the similarities and differences among products of different brands within the category. Our purchase-incidence model is based on the standard purchase-incidence model (e.g., Chiang 1991 and Chintagunta 1993) with a category-level assortment measure built in.

One may argue for carrying a complete category assortment, and in fact some retailers do so to eliminate the risk of not fulfilling the needs of the customers (c.f., Ho and Tang 1999). However, this approach faces two major challenges: (1) more products may actually confuse consumers (c.f., Kahn 1999), and (2) more products need more shelf space.[1] Therefore, a more realistic solution seems to be replacing some products with others. In this paper, we use the modeling framework together with a local improvement heuristic to generate alternative category assortments that are profit improving while keeping the assortment size constant.

This paper contributes to the existing literature in category management in three ways:

(1) We introduce three brand-width measures to characterize the brand-level assortment and explain how they capture consumer preferences for different assortments. We show these measures have predictive power for purchase incidence and brand share.

(2) We develop and estimate a hierarchical modeling framework of purchase incidence and brand choice using an extensive panel-level data set spanning eight food categories. We show that our model fits and predicts better than the benchmark modeling framework that combines the standard purchase-incidence model with Guadagni and Little's brand-share model.

(3) We couple the modeling framework with a local improvement heuristic to form a tool for reconfiguring category assortment. We use this tool to reconfigure the category assortments in the five stores for all eight categories in our data set. We show that category profit can increase as much as 25.1% as a result of assortment reconfiguration. This demonstration shows the promise of the modeling framework for category configuration in other frequently bought product categories. In addition, we use our modeling framework to quantify the amount of lost sales for not offering a complete list of products available in the market. We find that the amount of lost sales ranges from 0.9% to 10.2%.

This paper is organized as follows: In §2, the basic building block of our modeling framework, the three brand-width measures, is discussed first before we present the framework. Specifically, we represent each brand as a tree and derive three brand-width measures that capture the underlying characteristics of category assortment. Section 3 presents the modeling framework and shows how these brand-width measures are incorporated into the purchase-incidence model and the brand-share model. In §4, we use 60,000 shopping trips and purchase records to calibrate and validate our purchase-incidence model and the brand-share model. Section 5 illustrates how we couple the modeling framework with a local improvement heuristic to reconfigure category assortment for category profit improvement. In addition, we present an approach for estimating lost sales due to changes in category assortments. Section 6 summarizes the contributions and suggests future research directions.

2. Brand-Width Measures

Before presenting a hierarchical modeling framework of purchase-incidence and brand-choice decisions in §3, we shall develop three brand-width measures to be embedded in the modeling framework. This section is organized as follows. In §2.1, we represent a

[1]As reported in Quelch and Kenny (1994), the number of products increased by 16% per year between 1985 and 1992 while shelf space expanded by only 1.5% per year during the same period.

product category as a "product tree." Each brand is a subtree of this product tree and each product occupies an end node of this tree.[2] In §2.2, we use the historical purchase records of a consumer to estimate the importance weight of a particular node for the consumer. We use the importance weights to generate the three brand-width measures in §2.3.

2.1. Product Tree

Consider a product category that has several *salient attributes.* For example, in the ice-cream category, flavor and packaging size are two salient attributes that customers use to identify products. For estimation purposes we assume that there are only two salient attributes: package size and flavor. The same approach, however, can be used when there are more than two salient attributes.[3]

Consider a product category that has J brands, where brand j is comprised of N_j products or stockkeeping units (SKUs), for $j = 1, \ldots, J$ (see the Appendix for a list of notations used). Because most consumer products have a discrete number of levels for each attribute, we can represent the product structure of each brand j as a tree. Because there are two salient attributes for the product category, the tree has two layers and each layer represents an attribute. At each layer, different branches correspond to different levels of an attribute that the brand possesses. Since each SKU can be specified by a combination of different levels of two attributes, we represent each SKU as an end node of the tree, where the path between the root node and the end node specifies the combination of two attribute levels that the SKU possesses. Different SKUs of the same brand may share the same path if they have the same attribute combinations.

Let us consider a hypothetical example in which a store carries only two brands of ice cream: Haagen-Dazs and Breyer's. The product structure of the Haagen-Dazs brand in the store is depicted in Figure 1. As shown, the store carries four SKUs of Haagen-Dazs, where SKU 3 and SKU 4 possess the same combination of attribute levels. For example, SKU 3 could represent [16 oz., vanilla] with ground vanilla beans while SKU 4 could represent [16 oz., vanilla] without ground vanilla beans. This subtree structure forms a part of a larger product tree that represents the entire product category. The larger product tree is the union set of all products offered by all brands, regardless of whether a product is carried in the store. Figure 2 shows the category product tree for the ice-cream category, which has six nodes consisting of two package sizes and three flavors. Note that the store carries nine SKUs and that some nodes (e.g., [8 oz., vanilla]) are offered by both brands while others (e.g., [8 oz., strawberry]) are offered by only one brand.

Figure 1 The Product Tree for Haagen-Dazs Brand

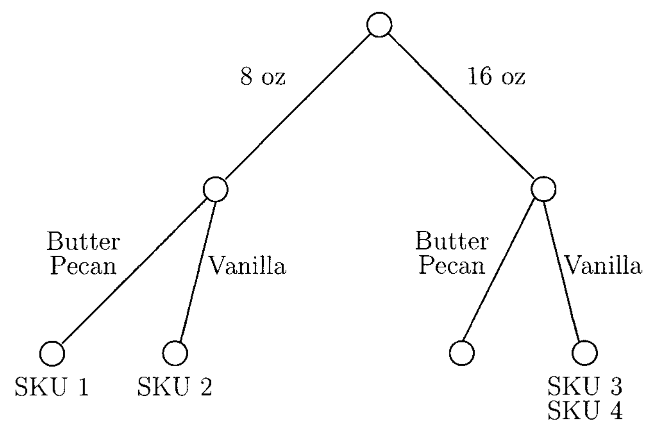
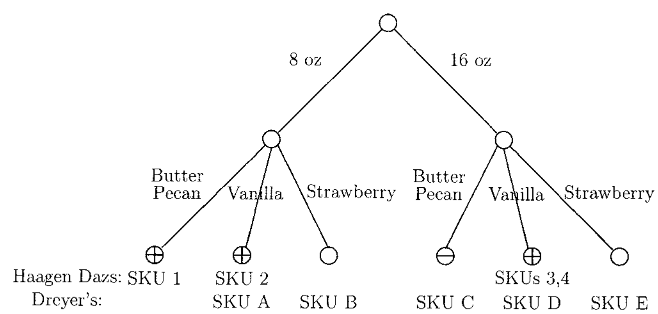


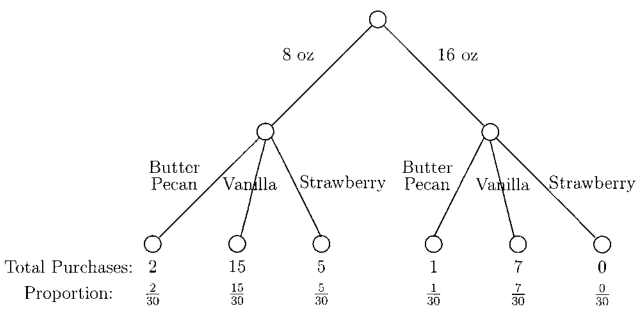
Figure 2 The Category Product Tree for the Ice-Cream Category



[2]The use of a tree structure to represent products is prevalent in marketing literature (e.g., Tversky and Sattath 1979, Moore et al. 1986, Kannan and Wright 1991).

[3]The use of multiple attributes to represent a product is common in the literature. For example, see Kannan and Wright (1991), Fader and Hardie (1996), and Ho and Chong (2000).

Figure 3 Purchase Record of Consumer *i* for the Ice-Cream Category



2.2. Importance Weights

A consumer may find each node to have a different degree of "importance." We measure the importance of a node to a consumer using his historical purchase frequency of the node. The importance weight of the node is then defined relative to the importance of other nodes. Specifically, the importance weight of a node k to consumer i during trip t, denoted by $r_i(k, t)$, is given by:

$$r_i(k, t) = \frac{b_i(k, t-1)}{\sum_{k'} b_i(k', t-1)}, \tag{2.1}$$

where $b_i(k, t-1)$ is the number of times consumer i purchases product node k in all stores prior to trip t.

Essentially, $r_i(k, t)$ corresponds to the relative frequency that consumer i purchased product node k prior to trip t. To illustrate, let us consider Figure 3. Consumer i has purchased ice cream 30 times prior to trip t, where the node [8 oz. butter pecan] was purchased twice, [8 oz., vanilla] 15 times, etc. Thus, the importance weight of node [8 oz., vanilla] to consumer i prior to trip t is equal to r_i([8 oz., vanilla], t) $= 15/30 = 0.5$.

2.3. Measure Specification

Before we introduce our brand-width measures, let us utilize the product tree for the ice-cream category (Figure 2) to classify the product nodes of Haagen-Dazs into three types: *distinct nodes*, *extensible nodes*, and *nonextensible nodes.* In Figure 2, distinct nodes of Haagen-Dazs are marked with a "+" sign, extensible nodes are marked with a "−" sign, and nonextensible nodes are left blank. Our three brand-width measures are defined according to these three types of nodes as follows.

The set of distinct nodes consists of distinct combinations of attribute levels occupied by current Haagen-Dazs products. For example, as shown in Figure 2, SKUs 3 and 4 share the same combination of attribute levels; hence they would only count as a single distinct combination or node. Therefore, Haagen-Dazs has three distinct nodes. Let $\mathcal{S}_j^G(t)$ be the set of distinct nodes that brand j carries in the store that consumer i visited during trip t.[4] It is easy to check from Figure 2 that $\mathcal{S}_{\mathit{Haagen\text{-}Dazs}}^G(t) = \{$[8 oz., butter pecan], [8 oz., vanilla], [16 oz., vanilla]$\}$.

The set of extensible nodes corresponds to products that Haagen-Dazs does not carry, although all of their component attribute levels are currently offered by Haagen-Dazs in the store. For example, as shown in Figure 2, Haagen-Dazs does not offer [16 oz., butter pecan]. However, Haagen-Dazs offers both 16 oz. and butter pecan in products that it currently carries in the store. Let $\mathcal{S}_j^R(t)$ be the set of extensible nodes for brand j during trip t. That is, we have $\mathcal{S}_{\mathit{Haagen\text{-}Dazs}}^R(t) = \{$[16 oz., butter pecan]$\}$.

The set of nonextensible nodes corresponds to products with at least one component attribute level that could not be found in Haagen-Dazs products currently offered by the store. In other words, nonextensible nodes correspond to products that possess unique attribute levels not offered by the brand. For example, as shown in Figure 2, Haagen-Dazs does not offer the product [8 oz., strawberry], and Haagen-Dazs does not offer any product with strawberry flavor in this store. In other words, strawberry is a unique attribute level that Haagen-Dazs does not offer. Therefore, [8 oz., strawberry] is a nonextensible node for Haagen-Dazs. Let $\mathcal{S}_j^E(t)$ be the set of nonextensible nodes for brand j during trip t. It is easy to check from Figure 2 that $\mathcal{S}_{\mathit{Haagen\text{-}Dazs}}^E(t) = \{$[8 oz., strawberry], [16 oz., strawberry]$\}$. Based on the definition of these three types of nodes for any brand j, it is

[4]Note that only one store is associated with a trip t. Therefore, the store identity can be derived from the trip index t. Any reference to a trip t necessarily implies a reference to the store visited. Also, we should have used the subscript i to indicate that the trip is made by customer i. We suppress i for simplicity.

easy to see that $\mathcal{S}_j^G(t) \cup \mathcal{S}_j^R(t) \cup \mathcal{S}_j^E(t)$ covers all product nodes for the entire product category during trip t.

By using the classification of the product nodes defined above, we now define a brand-width measure for each type of product node. Consider any node k in $\mathcal{S}_j^G(t)$ (i.e., distinct by brand j during trip t). Because multiple brands may occupy the same node, let $m_i(k, t)$ denote the total number of brands that offer node k, as observed by consumer i during trip t. We use $m_i(k, t)$ to adjust the relative importance weight for a brand. Presumably, the potency of an importance weight is reduced when more brands offer the same attribute combination.

We sum the total adjusted importance of all distinct nodes of brand j for consumer i at trip t to generate the first brand-width measure, denoted by $G_{ij}(t)$.[5] Specifically, we have:

$$G_{ij}(t) = \sum_{k \in \mathcal{S}_j^G(t)} \frac{r_i(k, t)}{m_i(k, t)}. \tag{2.2}$$

The brand-width measure $G_{ij}(t)$ quantifies the attraction of brand j to consumer i during trip t. Thus, a brand that has a higher brand-width measure $G_{ij}(t)$ should have a higher brand share.

We now turn our attention to developing the brand-width measures generated from the extensible nodes $\mathcal{S}_j^R(t)$ and the nonextensible nodes $\mathcal{S}_j^E(t)$. Since $\mathcal{S}_j^R(t)$ and $\mathcal{S}_j^E(t)$ are the nodes that are not carried by the brand, the two associated brand-width measures quantify the disappointment level of brand j to consumer i at trip t. They capture the opportunity loss in brand share for not carrying those product nodes.

By using the same approach for defining $G_{ij}(t)$, we define the brand-width measures associated with the extensible node (the nonextensible nodes), denoted by $R_{ij}(t)(E_{ij}(t))$, as equal to the total importance weights of all extensible nodes (nonextensible nodes) of brand j for consumer i at trip t. In this case, we have the brand-width measures for the extensible nodes:

$$R_{ij}(t) = \sum_{k \in \mathcal{S}_j^R(t)} r_i(k, t), \tag{2.3}$$

and the nonextensible nodes:

$$E_{ij}(t) = \sum_{k \in \mathcal{S}_j^E(t)} r_i(k, t). \tag{2.4}$$

Let us illustrate these measures with a numerical example. During trip t, consumer i visits a store that offers two brands of ice cream and SKUs as indicated in Figure 2. Recall from the definition of distinct nodes that the set $\mathcal{S}_{\textit{Haagen-Dazs}}^G(t)$ = {[8 oz., butter pecan], [8 oz., vanilla], [16 oz., vanilla]}. In this case, $G_{i,\textit{Haagen-Dazs}}(t) = 2/30 + 15/2 \cdot 30 + 7/2 \cdot 30$. Similarly for the set of extensible nodes $\mathcal{S}_{\textit{Haagen-Dazs}}^R(t)$ = {[16 oz., butter pecan]}, $R_{i,\textit{Haagen-Dazs}}(t) = 1/30$. For the nonextensible nodes, we have $\mathcal{S}_{\textit{Haagen-Dazs}}^E(t)$ = {[8 oz., strawberry], [16 oz., strawberry]}, which results in $E_{i,\textit{Haagen-Dazs}}(t) = 5/30$.

3. Modeling Framework

In this section, we present a hierarchical modeling framework of purchase and brand-choice decisions. With the three brand-width measures embedded, this modeling framework captures the impact of category assortment on individual consumers' purchase and brand-choice decisions. These individual-level responses are aggregated to form the category-level sales volume and brand-share distribution useful for a retailer to evaluate the profit implication of a category assortment.

We assume that consumer i adopts a two-step hierarchical decision process during a shopping trip. Specifically, she must first decide whether or not to buy a product from a particular category. If the decision is positive, then she must choose a specific brand from the category. The purchase decision can depend on (1) her inventory at home, and (2) the attractiveness of the category in terms of preferred choice, price, etc. Thus, the probability that consumer i purchases brand j during trip t can be defined as follows:

$$\mathrm{Prob}_{ij}(t) = Pc_i(t) \cdot Pr_{ij}(t), \tag{3.1}$$

where $Pc_i(t)$ is the probability that consumer i makes a purchase in the category during trip t and $Pr_{ij}(t)$ is

[5]Since our panelists shop at multiple stores, they will see different product assortments depending on which store they visit on trip t. Thus, the summation signs in Equations (2.2)–(2.4) are over nodes that are offered by the store visited by consumer i during trip t.

the conditional probability that brand j is chosen, given that a purchase is made. $Pc_i(t)$ is the category-purchase-incidence probability and $Pr_{ij}(t)$ is the brand-choice probability. We model the assortment impact on these two probabilities by building the three brand-width measures into the brand-share and purchase-incidence models.

3.1. Brand-Share Models

Consider a consumer i who visits a store to buy ice cream during trip t. There are J brands of ice cream available for her to choose. Each brand j is perceived to offer a utility $U_{ij}(t)$ during trip t, where:

$$U_{ij}(t) = V_{ij}(t) + \epsilon_{ij}(t).$$

The term $V_{ij}(t)$ corresponds to consumer i's deterministic utility obtained from buying brand j during trip t, and $\epsilon_{ij}(t)$ represents the stochastic term of her utility. We assume that the error terms $\epsilon_{ij}(t)$, $\forall i, j, t$ are independent and identically distributed with a double exponential (Gumbel) distribution (i.e., $F(\epsilon_{ij}(t)) = \exp(e^{-\epsilon_{ij}(t)})$, $\forall i, j, t$). If we assume that consumer i would select the brand that maximizes her utility, then she will choose brand j during trip t with probability $Pr_{ij}(t)$ (McFadden 1974, Ben-Akiva and Lerman 1985) where:[6]

$$\begin{aligned} Pr_{ij}(t) &= \mathrm{Prob}[U_{ij}(t) > U_{ij'}(t), \forall\, j' \neq j, j' \in \mathcal{J}_i(t)] \\ &= \frac{e^{V_{ij}(t)}}{\sum_{j' \in \mathcal{J}_i(t)} e^{V_{ij'}(t)}}. \end{aligned} \tag{3.2}$$

Note that $\mathcal{J}_i(t)$ corresponds to the set of brands available in the store during trip t.[7]

[6]Note that this is a brand-choice probability conditioned on the event that a category purchase is made.

[7]Note that $Pr_{ij}(t)$ is the choice probability of an aggregate entity (brand) which consists of individual units (SKUs). Hence the utility of the aggregate entity can be seen as derived from the individual units. In particular,

$$U_{ij}(t) = \max_{k \in S_j^G(t)} u_{ih}(t),$$

where $u_{ih}(t) = v_{ih}(t) + \epsilon_{ih}(t)$ is consumer i's utility for SKU h at trip t, $v_{ih}(t)$ is the deterministic component of the utility, and $\epsilon_{ih}(t)$ is the stochastic component. Ben-Akiva and Lerman (1985) showed that an adjustment term needs to be incorporated into the utility $U_{ij}(t)$ to account for the aggregation effect. For the detailed derivation of the exact adjustment term, the reader is referred to Chapter 9 of Ben-Akiva and Lerman (1985). Our brand-width measure $G_{ij}(t)$ aggregates the impact of individual SKUs for brand j; hence, our model specification accounts for the aggregation effect.

3.1.1. Model 1: The Guadagni and Little Model (GL). In the Guadagni and Little model, the deterministic utility $V_{ij}(t)$ is specified as:

$$V_{ij}(t) = \alpha_j + \beta_L L_{ij}(t) + \beta_P P_j(t) + \beta_D D_j(t) + \beta_{AD} AD_j(t),$$

where α_j is an intercept term specific to brand j. α_j is assumed to be stationary over time and constant across all consumers. In addition, $L_{ij}(t)$ represents consumer i's purchase experience of brand j up to but not including trip t, and β_L is the corresponding parameter. According to Guadagni and Little (1983), this purchase experience corresponds to *brand loyalty*, which can be expressed as the exponentially weighted average of past purchases made to brand j by consumer i as follows:[8]

$$L_{ij}(t) = \phi L_{ij}(t-1) + \begin{cases} (1-\phi) & \text{if brand } j \text{ is bought on trip } t-1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

The term $L_{ij}(t-1)$ is the loyalty of consumer i towards brand j on trip $t-1$, and ϕ is a smoothing constant bounded between zero and one. The above specification of purchase experience implies that, if a brand was frequently bought in the past it would have a higher value of $L_{ij}(t)$. Next, $P_j(t)$, $D_j(t)$, $AD_j(t)$ represent the price, display, and advertising features of brand j during trip t, respectively; and β_P, β_D, β_{AD} are the corresponding parameters. The term $\beta_P P_j(t) + \beta_D D_j(t) + \beta_{AD} AD_j(t)$ controls for the marketing environment that varies over time.

3.1.2. Model 2: The Brand-Configuration Model (BC). Our brand-configuration model extends the (GL) model by adding the three brand-width measures to the deterministic utility $V_{ij}(t)$ as follows:

$$\begin{aligned} V_{ij}(t) = \alpha_j &+ \beta_L L_{ij}(t) + \beta_P P_j(t) + \beta_D D_j(t) + \beta_{AD} AD_j(t) \\ &+ \beta_G G_{ij}(t) + \beta_R R_{ij}(t) + \beta_E E_{ij}(t), \end{aligned} \tag{3.4}$$

[8]Hence, the brand-loyalty (purchase experience) variable is bounded between 0 and 1.

where β_G, β_R, β_E are the corresponding parameters. Because the term $G_{ij}(t)$ measures the attractiveness of a brand generated by the distinct nodes, we expect that $\beta_G > 0$. Similarly, the terms $R_{ij}(t)$, $E_{ij}(t)$ measure the disappointment level of a brand generated by the nodes that the brand does not carry (i.e., the extensible nodes and the nonextensible nodes). Also, since the brand carries partial substitutes for each extensible node, we would anticipate the opportunity loss in brand share due to these nodes is somewhat less than that of the nonextensible nodes. For this reason, we expect $\beta_E < \beta_R < 0$.

3.2. The Purchase-Incidence Model

We follow the traditional marketing approach in specifying the purchase-incidence probability (e.g., Chiang 1991 and Chintagunta 1993). In particular, we set:

$$Pc_i(t) = 1 - \exp[-\rho \cdot A_i(t)], \tag{3.5}$$

where $A_i(t)$ captures the category attraction to consumer i on trip t and ρ is a scale parameter estimate that converts the attraction into incidence probability. To model the behavior that the purchase-incidence probability increases when the category attraction increases, we expect the parameter ρ to be nonnegative. The above functional form ensures that the probability $Pc_i(t)$ lies between 0 and 1.

The category attraction $A_i(t)$ is a function of two components that are likely to influence a consumer's purchase incidence: (1) the consumer's preference for the category vis-à-vis the category tree and (2) the consumer's inventory level of the product category. In our model, we assume that the consumer's preference for the category can be modeled as the sum of her preferences for the individual brands. Specifically, we assume that consumer i's preference for the category at time t is equal to $\Sigma_{j\in \mathcal{J}(t)} \exp[V_{ij}(t)]$, which is essentially the denominator of brand-choice model (3.1). In addition, we assume that the consumer's category preference exhibits decreasing return to scale. Therefore, we use $\ln\{\Sigma_{j\in \mathcal{J}(t)} \exp[V_{ij}(t)]\}$ as the first argument for the category attraction $A_i(t)$. Observe that $V_{ij}(t)$ captures the assortment changes through three brand-width measures $G_{ij}(t)$, $R_{ij}(t)$, and $E_{ij}(t)$. Hence, any assortment change that reduces $V_{ij}(t)$ via these three brand-width measures will reduce the purchase incidence probability. (To our knowledge, previous brand-share models such as (GL) have not captured this impact of assortment change.) Hence, $\ln\{\Sigma_{j\in \mathcal{J}(t)} \exp[V_{ij}(t)]\}$ is a summary measure for the category configuration.

The consumer's inventory level can be estimated as follows. Denote consumer i's household inventory level after the last shopping trip at time r as $Q_i(r)$. If consumer i bought any product during that trip, $Q_i(r)$ will reflect those purchases. We assume that consumer i has a linear consumption rate of θ_i. Therefore, at any time t before the next shopping trip, consumer i's inventory level can be estimated as $Q_i(r) - \theta_i(t - r)$. The category attraction $A_i(t)$ can be specified as follows:

$$A_i(t) = \exp\left(\eta \ln\left[\sum_{j\in \mathcal{J}(t)} \exp(V_{ij}(t))\right] + \gamma[Q_i(r) - \theta_i(t-r)]\right), \tag{3.6}$$

where γ measures the sensitivity of category attraction to the household inventory level and η measures the sensitivity to category configuration measure. We take the exponentiation on the sum of the two components to ensure that the category attraction $A_i(t)$ is nonnegative.

Combining (3.5) and (3.6), the purchase-incidence probability can be expressed as follows:

$$Pc_i(t) = 1 - \exp\left(-\rho \cdot \left[\sum_{j\in \mathcal{J}(t)} \exp(V_{ij}(t))\right]^{\eta} \times \exp(\gamma[Q_i(r) - \theta_i(t-r)])\right). \tag{3.7}$$

When consumer i's inventory level is high (i.e., when $[Q_i(r) - \theta_i(t - r)]$ is high), the purchase-incidence probability should be low. Hence, we expect γ to be nonpositive. When the category attraction is low (i.e., when $\Sigma_{j\in \mathcal{J}} \exp(V_{ij}(t))$ is low), the purchase-incidence probability should also be low. As such, we expect η to be positive.

By specifying the brand-share probability $Pr_{ij}(t)$ in

(3.2) and the purchase-incidence probability $Pc_i(t)$ in (3.7), we have completed the specification of our model framework as stated in (3.1). In the next section, we provide empirical evidence to show that our model framework (PI + BC), which couples the standard purchase-incidence (PI) model with our (BC) brand-share model, has better explanatory and predictive power than a benchmark (PI + GL) model framework that uses the (GL) brand-share model.

4. Estimation and Results

In this section, we first describe the data set and briefly discuss the estimation methodology. Then we present the empirical results.

4.1. Data Description

The scanner panel data set is drawn from a single IRI market in a metropolitan area in the United States.[9] It contains shopping information from 548 households over a two-year period (June 1991–June 1993). In addition, the data set contains purchasing information in eight food categories at five stores located within a two-mile radius.[10] These eight food categories are: coffee, frozen pizza, hot dogs, ice cream, potato chips, regular cereal, spaghetti sauce, and yogurt.[11] The data set also contains information regarding product availability at each store on a weekly basis, as well as marketing information such as price of SKUs at each store, advertising features, and in-store display on a weekly basis.

The input variables for our brand-share model are defined as follows. First, the price of each SKU is computed according to the price per basic unit (e.g., price per oz.). To compute $P_j(t)$, the price of brand j in week t, we compute the average price of all SKUs belonging to the brand weighted by their respective market shares.[12] Similarly, the variable $AD_j(t)$ (the advertising feature) and the variable $D_j(t)$ (the in-store display) are weighted averages of zero-one variables that indicate whether these SKUs are advertised and on store display. For the brand-width measures, we utilize the data description files to identify the corresponding brand name, package size, and flavor of each SKU.[13] Due to stockout, product addition, or product deletion, the product tree structure may vary from week to week because the set of SKUs associated with each brand varies from week to week.[14]

For each consumer, we keep track of her every shopping trip, whether she bought in a category and what brand she bought. The historical product preference of consumer i and the product offering of each brand at trip t allow us to compute her brand-width measures for brand j during trip t. Also, we estimate θ_i for consumer i using her average consumption rate during the calibration period.

4.2. Estimation Methodology

To estimate the parameters of our models, we use the method of maximum likelihood, which is asymptotically efficient.[15]

The likelihood of observing consumer i's behavior during trip t, denoted by $\mathcal{L}_{i,t}$, can be expressed as:

$$\mathcal{L}_{i,t} = (1 - Pc_i(t))^{1-Bc_i(t)} \cdot Pc_i(t)^{Bc_i(t)} \cdot \prod_j Pr_{ij}(t)^{B_{ij}(t)}, \quad (4.1)$$

where $B_{ij}(t)$ equals 1 if consumer i chooses brand j during trip t, and 0 otherwise. $Bc_i(t)$ equals 1 if consumer i makes a purchase within the category during trip t, and 0 otherwise. Thus, the total log-likelihood can be simplified as follows:

[9]We are grateful to Professor David Bell for providing us with the data set. The data set used here represents a portion of the "Basket" data set from Information Resources, Inc.

[10]Since a majority of the panelists shop at more than one of the five stores, we cannot estimate the model at the store level.

[11]We choose to estimate our model on food products because these categories have higher variety. In addition, the phenomenon of variety seeking is more prevalent in food products (e.g., McAlister 1982) and complicates the task of product planning for these categories.

[12]For example, Chiang (1991), and Wagner and Taudes (1986) used the same approach to compute the weekly price of a brand.

[13]Examples of the different package sizes and flavors for each category are given in Table 1.

[14]Note that the weekly tree structure is derived from the weekly data from all five stores, not from the consumer purchase record.

[15]In most product categories, our data set has in excess of 3,000 purchases and 20,000 shopping trips. Hence, we should have a sufficient sample size to benefit from the asymptotic property.

$$\mathcal{TLL} = \sum_i \sum_t \{[1 - Bc_i(t)]\ln[1 - Pc_i(t)] + Bc_i(t)\ln[Pc_i(t)]\} + \sum_i \sum_j \sum_t B_{ij}(t)\ln Pr_{ij}(t). \quad (4.2)$$

To avoid singularity in our estimation for the GL model or the BC model as specified in §3.1, we must fix one of the intercepts (α's) to zero. (Specifically, we arbitrarily choose to set α_J to zero where brand J is the brand that has the lowest brand share.) The loyalty-smoothing parameter ϕ is estimated all other parameters.[16] In addition, we use a nonlinear optimization routine with analytical gradient to perform the maximization.

4.3. Calibration and Validation Results

We divide the data set for initialization, calibration, and validation purposes as follows. The first 13 weeks of data are used for initialization, the next 65 weeks are reserved for in-sample calibration, and the last 26 weeks are used for out-of-sample validation purposes. Table 1 details the breakdown of the total number of shopping trips in-sample and out-of-sample made by the panelists in each of the eight categories.

The top half panel of Table 2 shows the in-sample calibration results. We report the total log-likelihoods for the coupled model frameworks, as well as the hit rate and the mean squared deviation for the best-fitted brand-share models.[17] Since the (PI + GL) model is nested within the (PI + BC) model, we can test whether the former can be rejected in favor of the latter by conducting the log-likelihood ratio test. These log-likelihood ratio test statistics are defined as $LR(PI + BC)$, where $LR(PI + BC) = -2(\mathcal{TLL}_{(PI+BC)} - \mathcal{TLL}_{(PI+GL)})$ and are also reported in Table 2.[18] The log-likelihood ratio test suggests that the (PI + GL) model can be rejected in all categories in favor of the (PI + BC) model. In terms of hit rate, the (BC) brand-share model outperforms the (GL) model in six out of eight categories. The (BC) model performs better in seven out of eight categories in mean square deviation.

The bottom half panel of Table 2 shows the out-of-sample validation results. We observe a similar pattern in results. The (PI + BC) model performs better in log-likelihood in every category except regular cereal. In terms of hit rate, (BC) is at least as good as (GL) in every category. The (BC) model has a lower mean square deviation in six out of eight categories.

Table 3 reports the parameter estimates.[19] For most categories, our brand-width measures have significant impact on brand share as predicted. In §3.1, we

[16]This parameter introduces an element of nonconcavity into the likelihood function. To mitigate the possibility of getting a local optimum due to the nonconcavity, we ran test estimation using three different initial values for ϕ at 0.25, 0.50, and 0.75 and picked the best of the three. In general, all three initial values seem to converge to a common parameter value. Also, in a previous version we estimated the brand-share submodel allowing for two segments of shoppers. The two-segment model fits slightly better than the single-segment model. We abandon the two-segment model because it is cumbersome and difficult to generate reliable estimates.

[17]A brand share model's hit rate is the proportion of times the model's most likely prediction matches the actual brand choice by the consumer. The mean squared deviation is computed as follows:

$$\frac{\sum_i \sum_t \sum_j [B_{ij}(t) - Pr_{ij}(t)]^2}{\sum_i \sum_t \sum_j B_{ij}(t)}.$$

[18]We also check the (PI + BC) model for multicollinearity. We use a measure suggested by Belsley et al. (1980) to detect any multicollinearity that might exist among the three brand-width measures. The measure ranges from 1 to ∞. A value of one indicates that the brand-width measures are completely independent. The larger the correlation among brand-width variables, the higher the value. Belsley et al. (1980) suggest that potential problems might arise if the value exceeds 20. The value of the multicollinearity measure varies from 1.40 to 2.54 for all eight product categories. Thus, we conclude that multicollinearity is not a problem for our model.

[19]The logit model assumes the independence of the irrelevant alternatives property, which may not be tenable in some choice settings. This property suggests that the ratio of two brand-choice probabilities remains constant regardless of the composition of the choice menu as long as it contains the two brands. To test if this property holds, we formulate a more general model than (BC). In this general model, we allow the ratio to vary according to menu size. Specifically, we model choices in large-menu settings with one set of parameters and choices in small-menu settings with another set of parameters (e.g., consider the price coefficients; we estimate β_P^S for small menu and β_P^L for large menu). The division of large-menu and small-menu settings is consumer specific. We take each consumer's two-year purchase occasions and compute the average number of brands the consumer sees in the two years. Those purchase occasions with menu size above this average are considered large menu and those below small menu. We test the significance in difference between the general model and (BC), and conclude that the difference is not significant with the log-likelihood ratio test. Therefore, we are satisfied that the IIA property holds with our model. We thank one anonymous reviewer for raising this important issue.

Table 2 In-Sample and Out-of-Sample Performances

Category	Coffee	Frozen Pizza	Hotdogs	Ice Cream	Potato Chips	Regular Cereal	Spaghetti Sauce	Yogurt
In-Sample Calibration								
Total shopping trips	32,223	23,670	26,138	34,510	30,877	41,041	22,925	28,288
Total shopping trips ended with category purchase	4,359	3,396	2,577	4,351	4,395	8,262	2,701	7,949
The Brand Configuration (PI + BC) Model Framework								
Total log-likelihood	−16,811	−13,846	−11,437	−17,774	−17,962	−28,579	−10,496	−21,025
Hit rate (BC)	0.6563	0.5495	0.5615	0.6045	0.5424	0.6184	0.6586	0.7653
Mean squared deviation (BC)	0.4590	0.5898	0.5734	0.5298	0.6039	0.5174	0.4416	0.3345
The Benchmark (PI + GL) Model Framework								
Total log-likelihood	−16,853	−13,885	−11,513	−17,861	−17,984	−28,772	−10,558	−21,131
Hit rate (GL)	0.6561	0.5518	0.5572	0.5980	0.5392	0.6122	0.6524	0.7668
Mean squared deviation (GL)	0.4601	0.5930	0.5837	0.5375	0.6048	0.5214	0.4506	0.3319
Log-likelihood ratio (PI + BC versus PI + GL)	85	77	152	174	44	387	124	213
Out-of-Sample Validation								
Total shopping trips	14,018	11,650	11,458	14,927	13,430	17,093	11,094	12,885
Total shopping trips ended with category purchase	1,509	1,412	927	1,623	1,698	3,040	1,085	3,189
The Brand Configuration (PI + BC) Model Framework								
Total log-likelihood	−6,838	−6,243	−4,518	−6,891	−7,264	−11,709	−4,772	−9,073
Hit rate (BC)	0.5984	0.5949	0.5879	0.6168	0.5524	0.5609	0.6590	0.7736
Mean squared deviation (BC)	0.5547	0.5679	0.5636	0.5233	0.5959	0.5727	0.4777	0.3386
The Benchmark (PI + GL) Model Framework								
Total log-likelihood	−6,898	−6,260	−4,557	−6,948	−7,270	−11,677	−4,803	−9,114
Hit rate (GL)	0.5944	0.5914	0.5879	0.6131	0.5518	0.5605	0.6535	0.7736
Mean squared deviation (GL)	0.5626	0.5734	0.5710	0.5346	0.5965	0.5671	0.4839	0.3327

Table 3 Parameter Estimates

Category	Coffee	Frozen Pizza	Hotdogs	Ice Cream	Potato Chips	Regular Cereal	Spaghetti Sauce	Yogurt
The Brand Configuration (PI + BC) Model Framework								
Purchase incidence								
ρ	0.0331	0.0356	0.0345	0.0240	0.0748	0.0450	0.0242	0.0843
	(0.0011)	(0.0004)	(0.0023)	(0.0005)	(0.0012)	(0.0018)	(0.0008)	(0.0030)
γ	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
η	1.4033	1.7286	1.7552	1.7379	1.8150	1.4695	1.6540	2.2794
	(0.0055)	(0.0150)	(0.0235)	(0.0177)	(0.0276)	(0.0141)	(0.0255)	(0.0417)
Brand choice								
ϕ	0.7794	0.6402	0.7621	0.7228	0.7973	0.8621	0.7117	0.3419
	(0.0873)	(0.0570)	(0.0122)	(0.0443)	(0.0381)	(0.0223)	(0.0384)	(0.0441)
β_L	6.6979	4.0628	4.6303	4.7064	4.5073	7.3409	3.9042	2.9768
	(0.0178)	(0.0356)	(0.0258)	(0.0147)	(0.0162)	(0.0771)	(0.0324)	(0.0221)
β_P	−2.8791	−0.8736	−0.9387	−1.7623	−1.1783	−0.6749	−0.7192	−1.9883
	(0.2055)	(0.0118)	(0.0404)	(0.0240)	(0.0126)	(0.0237)	(0.0240)	(0.0251)
β_D	1.2948	0.9361	0.9135	1.4964	0.8580	0.8309	0.9026	00.5629
	(0.0428)	(0.0167)	(0.0162)	(0.0192)	(0.0177)	(0.0281)	(0.0122)	(0.0821)
β_{AD}	0.7650	0.4848	0.6353	1.0467	0.2196	0.0604	0.8625	0.2875
	(0.0271)	(0.0189)	(0.0305)	(0.0087)	(0.0194)	(0.0367)	(0.0390)	(0.0391)
β_G	0.3847	1.3552	0.4068	0.6314	0.7512	0.2229	1.0346	0.4660
	(0.0260)	(0.0329)	(0.0580)	(0.0835)	(0.0181)	(0.0334)	(0.0379)	(0.0180)
β_R	−0.0946	0.9580	−0.3082	−0.1985	0.1974	−0.0632	−0.0904	0.3406
	(0.0163)	(0.0248)	(0.0233)	(0.0401)	(0.0173)	(0.0235)	(0.0128)	(0.0102)
β_E	−0.6427	0.6416	−0.9306	−0.8695	0.0491	−0.2820	−0.3593	−0.5381
	(0.0258)	(0.0267)	(0.0437)	(0.0788)	(0.0310)	(0.0231)	(0.0184)	(0.0171)
The Benchmark (PI + GL) Model Framework								
Purchase incidence								
ρ	0.0290	0.0505	0.0255	0.0197	0.0842	0.0444	0.0269	0.1183
	(0.0007)	(0.0014)	(0.0013)	(0.0005)	(0.0016)	(0.0006)	(0.0006)	(0.0039)
γ	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
η	1.3897	1.7732	1.7468	1.7124	1.8031	1.4783	1.6128	2.0720
	(0.0075)	(0.0302)	(0.0166)	(0.0122)	(0.0274)	(0.0061)	(0.0218)	(0.0404)
Brand choice								
ϕ	0.7768	0.6771	0.7520	0.7320	0.8001	0.8610	0.7011	0.5240
	(0.0710)	(0.0129)	(0.0489)	(0.0086)	(0.0428)	(0.0282)	(0.0097)	(0.0179)
β_L	7.3318	4.5013	5.4647	5.3667	4.6695	6.1177	4.7913	2.9965
	(0.0164)	(0.0130)	(0.0308)	(0.0368)	(0.0185)	(0.0148)	(0.0541)	(0.0214)
β_P	−2.5686	−0.8676	−0.8551	−1.8090	−1.1876	−0.4482	−0.8753	−2.3287
	(0.1489)	(0.0150)	(0.0324)	(0.0436)	(0.0115)	(0.0184)	(0.0243)	(0.0258)
β_D	1.4770	0.9129	0.9905	1.4467	0.8548	0.9770	0.9055	0.6322
	(0.0376)	(0.0210)	(0.0112)	(0.0199)	(0.0207)	(0.0405)	(0.0164)	(0.0588)
β_{AD}	0.9246	0.4848	0.6669	1.0514	0.2211	0.1879	0.8312	0.2472
	(0.0138)	(0.0296)	(0.0200)	(0.0314)	(0.0202)	(0.0515)	(0.0394)	(0.0464)

Note. Standard error in parentheses.

conjecture that the parameter associated with the distinct nodes will be positive (i.e., $\beta_G > 0$), and that the parameters associated with the extensible nodes and nonextensible nodes will be negative (i.e., $\beta_R < 0$ and $\beta_E < 0$). This conjecture is mostly confirmed by the estimated parameters reported in Table 3. Note that β_G is positive in every category. β_R and β_E are negative in a majority of the product categories.

Note that the estimated parameters for the purchase incidence model (ρ, γ, and η) are all consistent with the prediction. For instance, $\hat{\rho}$ and η are positive in every category. These results suggest that consumers are likely to buy from a category if the category attraction $A_i(t)$ is higher. Similarly, γ is negative in all categories. This implies that consumers are less likely to buy when their inventory levels are high.

In sum, our (PI + BC) model fits and predicts the observed purchase-incidence and brand choice better than the standard (PI + GL) model. In the next section, we demonstrate how our modeling framework can be used to reconfigure a category assortment to enhance profits and to quantify the degree of lost sales due to assortment changes.

5. Assortment Reconfiguration and Lost Sales

We formulate the category assortment reconfiguration problem as a constrained profit-maximization problem. In what follows, we show how the shelf-space constraint and the profit function are derived. Let us consider the assortment reconfiguration problem for store s in week v. For notational convenience, we shall suppress the store subscript for all variables.

Let $\mathcal{J}(v)$ be the set of brands associated with the category in week v. Suppose the amount of shelf space allocated to the category L is given. The category manager must select the SKUs to carry so that the total alloted shelf space does not exceed L. Let $\mathcal{C}(j)$ denote the set of all *existing* SKUs associated with brand j that the category manager can choose from, and let w_{jk} be the width of one shelf facing of SKU $k \in \mathcal{C}(j)$. For each SKU $k \in \mathcal{C}(j)$, let f_{jk} be the number of facings SKU k has on the shelf and $X_{jk}(v)$ be the binary decision variable that indicates whether the store carries SKU k during week v.[20] So, $X_{jk}(v) = 0$ corresponds to the case in which SKU k is not carried by the store in week v. Since the shelf space occupied by SKU k is equal to its number of facings (f_{jk}) multiplied by its width (w_{jk}), a *feasible assortment configuration,* $\{X_{jk}(v),\ k \in \mathcal{C}(j),\ j \in \mathcal{J}(v)\}$, must satisfy the following constraint in week v:

$$\sum_{j\in\mathcal{J}(t)} \sum_{k\in\mathcal{C}(j)} w_{jk}\cdot f_{jk}\cdot X_{jk}(v) \le L. \tag{5.1}$$

Next we consider the profit function. First, note that the expected sales of brand j generated from a consumer i in week v is given by $q_i\cdot PC_i(v)\cdot Pr_{ij}(v)$, where q_i is the average quantity bought by consumer i on a given trip when he buys from the product category.[21] The total expected sales of brand j in week v can be derived by summing over all consumers, $\Sigma_{\text{all } i}\, q_i\cdot Pc_i(v)\cdot Pr_{ij}(v)$.

Notice that the purchase-incidence probability $Pc_i(v)$ and the conditional brand-share probability $Pr_{ij}(v)$ vary with the assortment configuration. Given a feasible assortment configuration, the category manager can construct a product tree for the whole category (and the product trees for different brands within the category) as described in §2.2. These product trees allow the category manager to determine consumer i's brand-width measures of brand j in week v; namely, $G_{ij}(X_{jk}(v))$, $R_{ij}(X_{jk}(v))$, and $E_{ij}(X_{jk}(v))$ as specified in §2.3.[22] We use these three brand-width measures and other marketing variables to compute consumer i's brand-share probability, $Pr_{ij}(v)$ as given in (3.2), and her purchase-incidence probability $Pc_i(v)$ as given in (3.7).

Since the average price and cost of brand j depend on the assortment carried by the brand, we define $[P_j(X_{jk}(v),\ k \in \mathcal{C}(j)) - C_j(X_{jk}(v),\ k \in \mathcal{C}(j))]$ to be the "average" profit margin of brand j in week v. For notational convenience, we shall abbreviate the average

[20]We treat the number of facings as a parameter because we do not model how it affects sales. A more general model can capture this and allow number of facings to be a decision variable.
[21]Note that the total quantity bought by consumer i over a time period will increase if $Pc_i(v)$ increases as a result of reconfiguration.
[22]The three measures are a function of brand configuration. We now express them as a function of $X_{jk}(v)$ to remind the reader that the decision variables affect these three measures.

profit margin of brand j to $[P_j(X_{jk}(v)) - C_j(X_{jk}(v))]$. In this case, the total profit associated with the entire product category can be expressed as:

$$\sum_{j \in \mathcal{J}(v)} \left(\{P_j[X_{jk}(v)] - C_j[X_{jk}(v)]\} \sum_{i \in \mathcal{B}(v)} q_i \cdot Pc_i(v) \cdot Pr_{ij}(v) \right),$$

where $\mathcal{B}(v)$ is the customer base for the store in week v. Thus, the assortment reconfiguration problem can be formulated as follows:

Problem (AR).

$$\max_{X_{jk}(v)=\{0,1\}, k \in \mathcal{C}(j), j \in \mathcal{J}(v)}$$

$$\times \sum_{j \in \mathcal{J}(v)} \left(\{P_j[X_{jk}(v)] - C_j[X_{jk}(v)]\} \sum_{i \in \mathcal{B}(v)} q_i \cdot Pc_i(v) \cdot Pr_{ij}(v) \right)$$

$$\text{subject to} \sum_{j \in \mathcal{J}(v)} \sum_{k \in \mathcal{C}(j)} w_{jk} \cdot f_{jk} \cdot X_{jk}(v) \le L.$$

Notice that $Pr_{ij}(v)$ and $Pc_i(v)$ are nonlinear functions of the brand-width measures $G_{ij}(X_{jk}(v))$, $R_{ij}(X_{jk}(v))$, and $E_{ij}(X_{jk}(v))$, which vary with the category assortment $\{X_{jk}(v),\ k \in \mathcal{C}(j),\ j \in \mathcal{J}(v)\}$. Consequently, it is very difficult to express the profit function in an analytical form of $X_{jk}(v)$.

To maximize the total profit for the entire category, the category manager must find an optimal solution to problem (AR) (i.e., $X^*_{jk}(v) = 0$ or 1 for $k \in \mathcal{C}(j)$ and for $j \in \mathcal{J}(v)$). Since profit function cannot be expressed in an analytical form of $X_{jk}(v)$, it is very difficult to find an optimal assortment configuration unless one conducts an exhaustive search. However, the required computational effort would be prohibitively expensive. To elaborate, consider a hypothetical case in which a product category (ice cream) consists of five brands, where each brand has 20 possible SKUs to be selected. There are a total of 100 decision variables $X_{jk}(v)$ and the exhaustive search would require 2^{100} evaluations of the complex objective function of problem (AR). In view of this challenge, we propose a local improvement heuristic for assortment reconfiguration. Our intent is to illustrate how our modeling framework can be used to reconfigure a category assortment for profit enhancement.

5.1. A Local Improvement Heuristic

Our local improvement heuristic can be described as follows. First, we rank the brands in descending order according to their current profit margins. Each brand-level assortment is reconfigured one brand at a time according to the order of this ranking.

For each brand, we consider a SKU that is *not* currently on the shelf as a potential replacement for another SKU currently on the shelf.[23] We will make this pairwise interchange only when it does not violate the allocated shelf space and when the total category profit increases as a result of this replacement. We reconfigure a brand by using this pairwise interchange repeatedly until no further category profit improvement can be made. A more detailed procedure of our local-improvement heuristic can be explained as follows:

(1) Sort the brands in descending order according to the average brand profit margin of the existing configuration so that $[P_1(X_{jk}(v)) - C_1(X_{jk}(v))] \ge [P_2(X_{jk}(v)) - C_2(X_{jk}(v))] \ge \ldots$. Call this list the *brand list*. Compute the current category profit associated with this assortment configuration.

(2) Pick the first brand on the *brand list* to reconfigure. If the list is empty, stop.

(3) Define a feasible pair of interchange as a pair of SKUs l and k such that $X_{j'k} = 1$ and $X_{jl} = 0$ (i.e., SKU k is on the shelf while SKU l is not), and $w_{jl} \cdot f_{jl} \le w_{j'k} \cdot f_{j'k}$ where f_{jl} is the largest feasible value allowable by the constraint. Enumerate every feasible pair and for each feasible pair of interchange, compute the resulting improvement in category profit if the interchange is executed. If no feasible pair exists, go to Step 6.

(4) Find a feasible pair that produces the maximum improvement in category profit.

(5) If the maximum improvement in category profit is at least 0.01%, then we accept this pairwise interchange of SKUs (i.e., replace SKU k^* with SKU l^* with $X_{j'k^*} = 0$ and $X_{jl^*} = 1$), update the brand configuration and category profit, and go to step 3. Else, continue to next step.

(6) Delete the current brand from the *brand list* and go to step 2.

[23]In the case where each brand has a fixed shelf-space allocation, the SKU being replaced also belongs to the same brand.

5.2. Reconfiguration Exercise

We perform the reconfiguration exercise on the eight product categories in the five stores from our data set. The reconfiguration exercise is done on the last 26 weeks of data using the modeling framework we calibrated on the first 78 weeks. We faced two challenges in data availability. Specifically, our data set does not contain (1) profit margin for the brands, and (2) shelf-space allocation for stores. Our approach to resolve these challenges is as follows:

• We obtain the profit margin for the brands from some stores of Dominick's Finer Food in the Chicago area during the same time period (September 1989–May 1997).[24] We are able to obtain the profit margin data only for regular cereal. For other categories, we generate the profit margins as follows. First, we construct a frequency distribution of profit margin using the data for regular cereal.[25] For each SKU in other categories, we randomly draw a profit margin from this distribution and assign it to this SKU. We assume that profit margin of the SKU is the same across the five stores.

• We collect the total shelf space (L), the width of each SKU (w_{jk}), and the number of facings for each SKU (f_{jk}) for regular cereal from three different stores in California (two Ralph's stores and one Albertson's store).[26] Based on our store visits, most SKUs seem to have the same number of facings: Large stores have a minimum of two facings and smaller stores have one facing. For other categories besides regular cereal, we randomly generate facing data as follows. First, we assign one facing to each SKU. The top 20% SKUs (based on past sales volume) are randomly selected and assigned one or two additional facings.[27]

[24]The data are available from the marketing group of the University of Chicago Graduate School of Business.

[25]Note that profit margin here is expressed as x cents of profit per dollar. We multiply this data by the price of the SKU to obtain an absolute profit margin figure for use in the optimization.

[26]Because of a lag of seven years between the IRI data set and the shelf-space data, a few SKUs that existed in the IRI data set were no longer available at the stores we visited. For completeness, we shall assume that each of these SKUs has one facing at the store.

[27]Rather than arbitrarily decide a cutoff point, we apply the 80/20 rule, where 80% sales volume is usually accounted for by 20% of the SKUs.

We use this local improvement heuristic to reconfigure the regular cereal category (by using the actual data from three sources: IRI, Dominick's Finer Food, and our store visit). In addition, we use our local improvement heuristic to reconfigure the remaining seven categories (with randomly generated facing and profit margin). For each of the 26 weeks, we reconfigure the category assortment to achieve higher profit for each of the five stores. By comparing the weekly category profits before and after reconfiguration, we compute the percentage profit improvement due to reconfiguration. For ease of comparison, we report the percentage profit improvement *per replacement* in Table 4.

To trace the source of profit improvement, we report the total sales volumes and profits, before and after the reconfiguration, for the top three brands and the rest of the category in Table 5. As noted in the introduction, profit improvement comes from two sources: (1) higher share for higher-margin brands, and (2) higher category sales. From Table 5, we can see (from the before and after sales volume rows for the total category sales) that four categories (frozen pizza, regular cereal, spaghetti sauce, and yogurt) achieve higher profit through higher category sales. We see a reduction in category sales in the other four categories (coffee, hotdogs, ice cream, and potato chips) along with a shift of sales volume from the top two brands to higher-margin brands in the rest of the category, resulting in higher category profit.

Based on the results reported in Tables 4 and 5, and our own observations when running our reconfiguration exercises, we conclude that:

• Significant improvement in category profit can be achieved through small changes in category assortment. The reasonable profit improvement reported in Table 4 does not require a major revamp in the category assortment. The number of replacements in reconfiguration exercises ranges from 2 to 32 with an average replacement per brand of less than 1. Since our heuristic enables the category manager to realize sizable profit improvement, a more sophisticated optimization technique may provide an even higher profit improvement.[28] If the 0.4–1.3% per replacement im-

[28]One caveat is in order here. Since the margin of regular cereal is

Table 4 Half-Yearly Profit Improvement from Category Assortment Reconfiguration[1]

Category	Coffee	Frozen Pizza	Hotdogs	Ice Cream	Potato Chips	Regular Cereal	Spaghetti Sauce	Yogurt
Store 1								
% improvement	4.4%	10.2%	8.2%	21.1%	12.6%	10.2%	3.0%	5.2%
Improvement per replacement	0.3%	0.9%	1.7%	1.1%	1.0%	1.0%	1.1%	1.0%
Avg number of replacements	13	11	5	19	13	10	3	5
Store 2								
% improvement	4.4%	13.5%	0.5%	17.9%	11.5%	5.2%	7.5%	5.6%
Improvement per replacement	0.3%	0.4%	0.2%	1.5%	0.6%	0.4%	0.9%	0.4%
Avg number of replacements	17	31	2	12	18	13	9	14
Store 3								
% improvement	7.1%	25.1%	12.6%	15.0%	11.4%	4.9%	5.9%	7.0%
Improvement per replacement	0.3%	0.8%	1.8%	0.8%	0.8%	0.4%	1.2%	0.7%
Avg number of replacements	22	32	7	18	14	12	5	10
Store 4								
% improvement	12.3%	10.6%	1.6%	6.9%	19.1%	10.8%	10.1%	6.2%
Improvement per replacement	0.5%	1.2%	0.5%	0.6%	1.1%	0.7%	2.0%	1.0%
Avg number of replacements	24	9	3	11	17	16	5	6
Store 5								
% improvement	13.9%	11.6%	7.5%	5.1%	24.3%	11.6%	10.0%	4.0%
Improvement per replacement	0.5%	0.8%	2.7%	0.7%	1.3%	1.3%	2.4%	0.8%
Avg number of replacements	27	14	3	8	19	9	4	5

Note 1. Total category profit for the half-year period from week 79 to 104.

provement of regular cereal is of any guide, it does show that the reconfiguration exercise is worthwhile for the five stores.

• Product variety is key for improving category profit. Specifically, we observe that the heuristic provides solutions where different brands offer different products with different attribute-level combinations to meet heterogeneous consumer needs. In particular, we see a general trend of reallocating shelf space from the top brand to other brands with products of unique attribute-level combination to achieve higher category profit.

In summary, we have illustrated how a category manager can use our modeling framework and heuristic to achieve higher category profit through assortment reconfiguration. Our modeling framework can also be used as an integral component in a complete store assortment planning system.

high and we use it to randomly generate margins for other categories, there may be a tendency to inflate the increases in profit.

5.3. Lost Sales Assessment

Our previous discussion has focused on the profit implication of category assortment. Profit may not be the only objective that a retailer considers. Lost sales is also an important measure that retailers give attention to. In this section, we use our modeling framework to quantify lost sales due to assortment changes. Assortment changes may be due to product addition, product deletion, or stockout.[29]

We define lost sales as the difference between the actual category sales of a store and the expected sales associated with a reference assortment. We choose the union of all products offered by the store over the two-year period to be the reference assortment and call it the ''complete assortment'' (versus the ''actual

[29]Note that our data set only captures weekly changes in assortment. Hence, any stock-out event that lasted less than a week will not be recorded. This coarser level of data could result in underestimation of lost sales. We thank Professor Ananth Raman for this observation.

Table 5 Average Weekly Sales Volume and Profit for the Top Three Brands Before and After Category Assortment Reconfiguration

Category		Coffee	Frozen Pizza	Hotdogs	Ice Cream	Potato Chips	Regular Cereal	Spaghetti Sauce	Yogurt
Top Brand									
Before	Sales volume	18.32	6.85	10.01	10.10	23.14	76.98	7.02	173.94
	Profit ($)	1.60	6.96	6.98	1.43	26.62	0.33	2.64	91.17
After	Sales volume	18.27	6.85	9.61	9.64	23.07	77.19	6.70	172.34
	Profit ($)	1.63	7.10	6.84	1.39	27.09	0.34	2.57	92.18
Second Brand									
Before	Sales volume	9.36	5.56	7.66	7.24	9.40	55.52	9.34	131.53
	Profit ($)	1.17	6.63	4.39	5.44	12.08	0.34	3.30	74.60
After	Sales volume	9.36	5.60	7.64	7.19	9.40	55.52	9.84	131.76
	Profit ($)	1.20	6.81	4.47	5.52	12.33	0.35	3.55	76.26
Third Brand									
Before	Sales volume	12.09	3.64	6.44	1.17	6.60	16.46	3.52	27.43
	Profit ($)	1.80	3.76	2.66	0.36	7.85	0.07	0.89	10.40
After	Sales volume	12.10	3.65	6.54	1.19	6.61	16.44	3.53	27.15
	Profit ($)	1.83	3.84	2.75	0.37	8.02	0.07	0.91	10.50
Others									
Before	Sales volume	31.83	23.93	18.56	37.39	26.47	37.66	17.96	66.03
	Profit ($)	5.05	29.91	20.06	14.37	32.56	0.17	10.23	31.17
After	Sales volume	31.83	23.92	18.66	37.72	26.46	37.63	17.95	71.92
	Profit ($)	5.15	30.50	20.74	14.78	33.18	0.17	10.44	34.77
Total									
Before	Sales volume	71.59	39.98	42.66	55.90	65.60	186.62	37.84	398.93
	Profit ($)	9.62	47.26	34.08	21.60	79.10	0.91	17.07	207.35
After	Sales volume	71.56	40.01	42.46	55.74	65.54	186.78	38.02	403.17
	Profit ($)	9.81	48.25	34.80	22.06	80.62	0.93	17.47	213.71

Note. Average weekly sales volume and profit computed for the last six months.

assortment''). This allows us to determine the impact of a partial assortment on sales volume.

In a product category, let q_i be consumer i's expected purchase quantity per trip. Let the purchase incidence probability of consumer i during trip t under complete product assortment be $\overline{Pc_i(t)}$. Finally, let $Bc_i(t)$ equal 1 if consumer i makes a purchase within the category during trip t and 0, otherwise. In this case, the expected lost sales at a store s can be expressed as follows:

$$\sum_i \sum_{t \text{ at store } s} \{q_i \cdot \overline{Pc_i(t)}\} - \sum_i \sum_{t \text{ at store } s} q_i \cdot Bc_i(t). \quad (5.2)$$

Notice that the first term represents the expected category sales when store s carries the complete assortment during trip t and observe that the second term corresponds to the actual category sales when store s carries the actual assortment during the same trip. Therefore, the difference between these two terms is the expected lost sales at store s for carrying the actual assortment instead of the complete assortment.

Using the modeling framework we estimated in the previous sections, we compute the expected lost sales for the eight categories in our data set.[30] We focus on the expected lost sales for the last 26 weeks for all five stores.

Table 6 reports the expected lost sales as a per-

[30]Although only the purchase incidence probability is used here, the brand share probability has to be estimated simultaneously since it provides the category-level assortment measure as an input for the purchase incidence probability.

Table 6 Lost Sales Assessment for the Two-Year Period

Category		Coffee	Frozen Pizza	Hotdogs	Ice Cream	Potato Chips	Regular Cereal	Spaghetti Sauce	Yogurt
Store 1	Current sales	1,170	1,266	711	1,379	1,639	2,244	1,181	2,267
	Potential sales[1]	1,187	1,293	763	1,429	1,702	2,254	1,204	2,360
	Lost sales	17	27	52	50	63	10	23	93
	% lost sales	1.45%	2.06%	6.82%	3.48%	3.72%	0.44%	1.88%	3.93%
Store 2	Current sales	1,348	1,427	1,190	1,821	1,962	2,715	1,233	3,935
	Potential sales	1,364	1,449	1,312	1,865	2,021	2,731	1,273	4,148
	Lost sales	16	22	122	45	59	16	41	213
	% lost sales	1.16%	1.50%	9.29%	2.40%	2.93%	0.60%	3.21%	5.14%
Store 3	Current sales	2,019	1,403	1,286	1,922	1,998	3,787	881	2,929
	Potential sales	2,042	1,418	1,444	1,949	2,083	3,813	902	3,010
	Lost sales	23	14	158	27	85	26	20	82
	% lost sales	1.13%	1.02%	10.93%	1.37%	4.08%	0.67%	2.26%	2.71%
Store 4	Current sales	900	422	335	533	310	1,712	325	1,201
	Potential sales	903	437	398	535	326	1,738	340	1,260
	Lost sales	3	15	63	2	15	26	15	59
	% lost sales	0.33%	3.43%	15.81%	0.41%	4.62%	1.48%	4.31%	4.65%
Store 5	Current sales	697	491	245	490	380	1,216	324	1,112
	Potential sales	707	495	278	497	406	1,238	338	1,155
	Lost sales	10	4	33	7	26	22	13	42
	% lost sales	1.45%	0.83%	11.74%	1.39%	6.43%	1.77%	3.94%	3.68%
Total	Current sales	6,134	5,010	3,768	6,145	6,289	11,674	3,945	11,444
	Potential sales	6,203	5,092	4,196	6,275	6,538	11,774	4,057	11,933
	Lost sales	69	82	427	130	249	100	112	489
	% lost sales	1.12%	1.61%	10.18%	2.08%	3.80%	0.85%	2.76%	4.10%

Note 1. Potential sales is calculated assuming the store has the full portfolio for each brand that it carries.

centage of the expected category sales when store *s* carries complete assortment. By examining Table 6, we observe the following:

- The expected lost sales ranges from 0.85% (regular cereal) to 10.18% (hotdogs).
- Lost sales is more category dependent than store dependent. No single store appears to be good at managing lost sales for all categories. For instance, store 4 is relatively good at managing lost sales for coffee but is poor in managing lost sales for frozen pizza.

6. Summary

In this paper, we have introduced three brand-width measures that capture consumers' historical preferences for the assortment of a brand versus other brands. Specifically, these measures capture the similarities and differences among products within a brand and across different brands.

These brand-width measures are embedded into a hierarchical modeling framework consisting of two empirical models of consumer shopping behaviors (1) a model for predicting purchase incidences, and (2) a model for predicting brand share. The purchase-incidence model follows the standard purchase-incidence model and the brand-share model extends the traditional Guadagni and Little model. Using an extensive panel-level data set that involves more than 60,000 shopping trips spanning eight food categories, we have shown that our modeling framework (PI + BC) fits and predicts better than the standard modeling framework (PI + GL). In addition, our modeling framework can be used to quantify the impact of product assortment on category profit and lost sales. Specifically, we have illustrated how our modeling framework can be used to reconfigure category assortment for higher profit and to estimate potential lost sales.

There are several ways to extend the research presented in this paper. First, we have ignored the impact of number of facings on demand. A more general model can include number of facings as one of the independent variables in the brand-share model. This will allow us to analyze the allocation of shelf space to brands and products. Second, we have not captured the inventory costs into our model. A more general modeling framework should take them into consideration.[31] Including these inventory costs is likely to lower the assortment of the entire product category and it is important to determine which brand will be affected the most. Third, one can model the consumer choice at the stock-keeping-unit level instead of the brand level. Such an approach avoids the potential aggregation bias of a brand-choice model elegantly and allows one to examine the issue of product substitution at the stock-keeping-unit level rather than at the brand level (c.f., Ho and Chong 2000).

Acknowledgments

The authors thank Leroy Schwarz, Ananth Raman, and two anonymous reviewers for their helpful comments. They are especially grateful to Marshall Fisher for his many insightful suggestions which have greatly improved the quality of this manuscript. This research is partially supported by UCLA James Peters Research Fellowship and Wharton—SMU Research Center.

Appendix. Notations

We document the notational convention used in this paper. In particular, we present (1) the indexing convention, and (2) the naming convention, as well as the list of variables arranged in categories.

Indexing Convention

We have the following indexing convention:

(1) i indexes the consumers where $i = 1, \ldots, I$.
(2) j indexes the brands where $j = 1, \ldots, J$.
(3) k indexes the nodes in a product tree.
(4) t indexes the trip made by a consumer.
(5) v indexes the calendar weeks. It is used in §5.
(6) s indexes the stores.

Naming Convention

We use different types of notation for different categories of variables. The following is the list of variables arranged in each category of variables.

(1) We use capital letters for our input and independent variables. They are:
  (a) $Pr_{ij}(t)$ is the brand-choice probability of consumer i for brand j on trip t.
  (b) $Pc_i(t)$ is the purchase-incidence probability of consumer i on trip t.
  (c) $U_{ij}(t)$ is the total utility of consumer i for brand j on trip t.
  (d) $V_{ij}(t)$ is the deterministic component of utility of consumer i for brand j on trip t.
  (e) $P_j(t)$ is the average price of brand j on trip t. It is a function of brand configuration. In §5, we express it as a function of brand configuration.
  (f) $C_j(t)$ is the average cost of brand j on trip t. It is a function of brand configuration. In §5, we express it as a function of brand configuration.
  (g) $D_j(t)$ is the display indicator variable of brand j on trip t.
  (h) $AD_j(t)$ is the advertising indicator variable of brand j on trip t.
  (i) $L_{ij}(t)$ is the brand-loyalty variable of consumer i for brand j on trip t.
  (j) $G_{ij}(t)$ is the positive brand-configuration effect from the set of distinct nodes of brand j for consumer i on trip t. It is a function of brand configuration. In §5, we express it as a function of brand configuration.
  (k) $R_{ij}(t)$ is the negative brand-configuration effect from the set of extensible nodes of brand j for consumer i on trip t. It is a function of brand configuration. In §5, we express it as a function of brand configuration.
  (l) $E_{ij}(t)$ is the negative brand-configuration effect from the set of nonextensible nodes of brand j for consumer i on trip t. It is a function of brand configuration. In §5, we express it as a function of brand configuration.
  (m) $B_{ij}(t)$ is the indicator variable that consumer i bought brand j on trip t.
  (n) $Bc_i(t)$ is the indicator variable that consumer i made a purchase in the category on trip t.
  (o) $Q_i(t)$ is the total units of products that consumer i bought on trip t.
(2) We use small Greek alphabets for our parameter estimates and stochastic variables.
  (a) α_j is the intercept for brand j.
  (b) ϕ is the decaying/smoothing constant for the brand-loyalty variable.
  (c) β_P is the parameter estimate for price variable.
  (d) β_D is the parameter estimate for display variable.
  (e) β_{AD} is the parameter estimate for advertising variable.
  (f) β_L is the parameter estimate for the brand-loyalty variable.

[31]Quelch and Kenny (1994) discussed the cost associated with different assortments of products within a brand. Extensive field research on the auto industry by Fisher and Ittner (1999) and Fisher et al. (1995) shows that proliferated product lines can increase both overhead and variable production costs.

(g) β_G is the parameter estimate for the positive brand-configuration effect of distinct nodes.
(h) β_R is the parameter estimate for the negative brand-configuration effect of extensible nodes.
(i) β_E is the parameter estimate for the negative brand-configuration effect of nonextensible nodes.
(j) $\epsilon_{ij}(t)$ is the stochastic component of utility of consumer i for brand j on trip t.
(k) θ_i is the consumption rate of consumer i.
(l) ρ, γ, and η are the scale parameters for the purchase incidence probability $Pc_i(t)$.

(3) We use calligraphic style capital letters to denote a set.
(a) $\mathcal{C}(j)$ is the set of SKUs in brand j.
(b) $\mathcal{S}_j^G(t)$ is the set of distinct nodes of brand j on trip t.
(c) $\mathcal{S}_j^R(t)$ is the set of extensible nodes of brand j on trip t.
(d) $\mathcal{S}_j^E(t)$ is the set of nonextensible nodes of brand j on trip t.
(e) $\mathcal{J}_i(t)$ is the set of brands available to consumer i on trip t.
(f) $\mathcal{B}(v)$ is the customer base for a store in week v.

(4) We use small letters for any other types of variables.
(a) $r_i(k, t)$ is the relative weight consumer i places on node k on trip t.
(b) $b_i(k, t)$ is the total purchases consumer i made in node k up to trip t.
(c) $m(k, t)$ is the number of brands occupying node k on trip t.
(d) q_i is the average quantity of consumer i purchased each trip the consumer makes a category purchase.
(e) f_{jk} is the number of shelf facings of a SKU k where $k \in \mathcal{C}(j)$.
(f) w_{jk} is the width of one shelf facing of a SKU k.

References

Belsley, D., E. Kuh, R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley, New York.

Ben-Akiva, M., S. R. Lerman. 1985. *Discrete Choice Analysis.* MIT Press, Cambridge, MA.

Cachon, G., M. Fisher. 1997. Campbell Soup's continuous replenishment program: Evaluation and enhanced decision rules. *Production Oper. Management* **6**(3) 266–276.

——, ——. 2000. Supply chain inventory management and the value of shared information. *Management Sci.* **46**(8) 1032–1048.

Chiang, J. 1991. A simultaneous approach to the whether, what and how much to buy questions. *Marketing Sci.* **10**(4) 297–315.

Chintagunta, P. K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Sci.* **12**(2) 184–208.

Fader, P., B. Hardie. 1996. Modeling consumer choice among SKU's. *J. Marketing Res.* **33** 442–452.

Fisher, M., C. Ittner. 1999. The impact of product variety on automobile assembly operations: Empirical evidence and simulation analysis. *Management Sci.* **45**(6) 771–786.

——, A. Jain, J. MacDuffie. 1995. Strategies for product variety: Lessons from the auto industry. B. Kogut, E. Bowman, eds. *Redesigning the Firm.* Oxford University Press, New York, 116–154.

Guadagni, P., J. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2**(3) 203–238.

Ho, T., J. Chong. 2000. A parsimonious model of SKU choice. Working paper 99-020, Marketing Department, University of Pennsylvania, Philadelphia, PA.

——, C. Tang. 1999. *Product Variety Management: Research Advances.* Kluwer Academic Publishers, New York.

Hoch, S., E. Bradlow, B. Wansink. 1999. The variety of an assortment. *Marketing Sci.* **18**(4) 527–546.

Kahn, B. E. 1999. Variety: From the consumer's perspective. T. Ho, C. Tang, eds. *Research Advances in Product Variety Management.* Kluwer Academic Publishers, New York.

Kannan, P., G. Wright. 1991. Modeling and testing structured markets: A nested logit approach. *Marketing Sci.* **10**(1) 58–82.

Lee, H., K. So, C. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Sci.* **46**(5) 626–643.

McAlister, L. 1982. A dynamic attribute satiation model of variety-seeking behavior. *J. Consumer Res.* **9** 141–150.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. P. Zarembka, ed. *Frontiers in Econometrics.* Academic Press, New York, 105–142.

Moore, W. L., E. A. Pessemier, D. R. Lehmann. 1986. Hierarchical representations of market structures and choice processes through preference trees. *J. Bus. Res.* **14**(5) 371–386.

Quelch, J., D. Kenny. 1994. Extend profits, not the product lines. *Harvard Bus. Rev.* **72**(5) 153–160.

Tversky, A., S. Sattath. 1979. Preference trees. *Psych. Rev.* **86**(6) 542–573.

Wagner, U., A. Taudes. 1986. A multivariate polya model of brand choice and purchase incidence. *Marketing Sci.* **5**(5) 219–244.

The consulting Senior Editors for this manuscript were Marshall Fisher and Ananth Raman. This manuscript was received January 25, 2000, and was with the authors 339 days for 4 revisions. The average review cycle time was 61.25 days.