An Algorithm and Demand Estimation Procedure for Retail Assortment Optimization

Marshall L. Fisher • Ramnath Vaidyanathan

OPIM Department, The Wharton School, 3730 Walnut St, Philadelphia, PA 19104, USA

fisher@wharton.upenn.edu • ramnathv@wharton.upenn.edu

June 2009

Abstract

We consider the problem of choosing, from a set of N potential SKUs in a retail category, K SKUs to be carried at each store so as to maximize sales or a defined profit function. Assortments can vary by store, subject to a maximum number of different assortments. We describe an approach in which we view a SKU as a set of attribute values, use sales history of the SKUs currently carried by the retailer to estimate the demand for attribute values and from this, the demand for any potential SKU, including those not currently carried by the retailer. We also introduce a model of substitution behavior, estimate the parameters of this model and consider the impact of substitution in choosing assortments. We use maximum likelihood estimation to fit the parameters of our model and describe several alternative heuristics for choosing SKUs. We describe application of this approach to optimize assortments for two real examples, snack foods and tires, which produced sales lifts of 20% *and* 32% *respectively.*

1 Introduction

A retailer's assortment is defined by the set of products carried in each store at each point in time.

The assortment planning process varies greatly across different retailers and product segments, but we can describe here a process that is common to a large number of retailers, particular in the segment commonly called hard goods and comprised of most non food and non apparel items.[1]

Assortment planning is conducted for a defined future planning horizon, which can be as short as a few hours in the case of a grocery retailer optimizing their assortment for morning breakfast shoppers, to as long as a year. Assortment planning considerations range from high level strategic issues, like whether a given assortment aligns with the retailer's brand position, to detailed operational decisions,

[1]This discussion is based on Levy and Weitz (2007) and conversations with several retail executives including Kevin Freeland, Executive VP, Advance Auto and former Senior VP for Inventory Management at Best Buy, Rob Price, Chief Marketing Officer of CVS and former Chief Marketing Officer of Wawa, and Bob DiRomualdo, CEO, Borders Group.

like which twenty digital still cameras, out of the thousands of possibilities, will a consumer electronics retailer carry in each of its stores.

Assortment planning begins with the strategic level questions of how much of the buying budget and store space is to be allocated to each product category and follows a process similar to that used by many companies for annual capital budgeting. Primary inputs to this process are a history of sales over time for each category that can be used to forecast category sales over the planning horizon supplemented with external inputs such as trade shows and advice from vendors. Based on a forecast of future category sales, a retailer would determine how much store space to allocate to each category, and hence how many SKUs to carry in each category.

Once a retailer has decided the number of SKUs to carry in a category, they face the operational problem of deciding which specific SKUs to carry. For example, if Best Buy decides to carry twenty digital still cameras, they then face the challenge of deciding which twenty to stock out of the thousands of possibilities available. The decision of which specific SKUs to carry is usually left to a buyer. Since a major purpose of assortment planning is to accommodate new products, there will be many products being considered for the assortment for which there is no sales experience. The lack of sales history for many of the potential stock keeping units (SKUs) a retailer might carry makes it hard to do analytic SKU level planning, so the buyer has little if any data or analytical tools to help them with the difficult decision of which specific SKUs to carry in a category.

Most retailers use the same assortment for all stores, except that in smaller stores they might eliminate some SKUs from the chain level assortment. Localizing assortments by store or store cluster is a high priority for many retailers. For example, Zimmerman (September 7, 2006), O'Connell (2008), McGregor (2008) and Zimmerman (October 7, 2008) describe recent efforts by Wal-Mart, Macy's, Best Buy and Home Depot to vary the assortment they carry at each store to account for local tastes. Retailer executives note many differences in sales patterns between stores, such as selling more large sizes in Iowa than Florida, more black in Manhattan than in Madison, Wisconsin and more expensive products in wealthier neighbourhoods. The goal of these localization efforts is to vary assortments by store or clusters of stores to take into account these differing purchase patterns. In the extreme, a retailer might carry a unique assortment in each store, but this is usually not practical because the administrative burden would be too high. For example, most retailers develop a diagram called a planogram that shows how all products would be displayed in a store, a process that is labor intensive. A planogram would need to be developed for each assortment, which means the administrative cost of each assortment is high. Most retailers now believe that more than one assortment is needed, to account for difference in consumer taste between stores, but that the number of different assortments must be limited to avoid overly high administrative costs. Despite the flurry of interest in localization amongst retailers reported in the business press and which we have encountered in our interaction with retailers, there have been no studies to document the level of benefits from localization or to provide tools to help a retailer determine the right degree of localization.

It's apparent that the ability to forecast store-SKU demand for all potential SKUs, including those with which a retailer has no prior sales experience, and to intelligently localize assortments by store,

would be valuable enhancements to current assortment planning practice, and the goal of this paper is to fill that void.

Our approach draws on the familiar idea from the marketing literature of viewing a SKU as a set of attribute values. We assume that each SKU that might be offered by a retailer in a category is defined by a set of attributes, each with a finite set of values and that a given customer has a preferred set of attribute values. We use prior sales to estimate the market share in each store of each attribute level and forecast the demand share for a SKU as the product of the demand shares for the attribute values of that SKU. For example, the percentage of customers who prefer a blue, Queen size, cotton sheet is estimated as the percent who prefer blue times the percent who prefer Queen times the percent who prefer cotton. We assume that if a customer does not find their ideal product in the assortment, then they buy the product in the assortment closest to their ideal with some probability and we also estimate these substitution probabilities. We then apply a greedy heuristic to these demand and substitution estimates to determine optimized assortments. Our process can control the degree of localization in the assortment to be any level between a single assortment for the chain to a unique assortment for each store.

We applied this approach to determine the snack cakes assortment at a regional convenience store chain and the tire assortment at a national retailer and found the increase in revenues to be 20% and 32% respectively, with 12% and 6% of this due to localizing the assortment by store. The factor of two differences in localization benefit between the two retailers is striking and we report some surprising and subtle reasons for this. The tire retailer implemented about a quarter of our recommended assortment changes and obtained a 6% revenue increase, which is significant given the traditional $3-4\%$ comparable store annual increases in this segment.

We do not consider inventory decisions, so our approach fits when inventory decisions are unrelated to assortment decisions as is the case for many slow movers, where the retailer carries a small amount of inventory, often just a single unit, e.g jewelry, auto parts, books and CDs. The convenience store retailer in our study carried a single facing of each product and the tire retailer four of each tire.

Despite the enormous economic importance of assortment planning, we are aware of only two other papers, Chong et al. (2001) and Kök and Fisher (2007), that formulate a decision support model for assortment planning, describe a methodology for estimating parameters and optimizing assortments and test the process on real data. While these papers are an excellent start, there is obviously a need for much more research on the vast topic of assortment planning. This paper seeks to extend these initial works in four important ways.

First, we forecast the demand for new SKUs that have not been carried before in any store, based on past sales of products currently carried. Chong et al. (2001) don't explicitly forecast new SKUs, although it would appear from their process that they could derive forecasts for new SKUs. However, they need consumer-level transaction data over multiple shopping trips and this detailed data is not available in most applications; the standard data available is store-SKU sales data. Kök and Fisher (2007) forecast how a SKU carried in some stores will sell in other stores in which it is not carried, but do not provide a way to forecast demand for a completely new SKU carried in no stores.

Secondly, we introduce a new demand model not previously considered in assortment decision support research. Both Chong et al. (2001) and Kök and Fisher (2007) use the multinomial logit demand model, which implicitly allows for customers to substitute if their first choice product is not available, in a way that assumes that substitution demand is divided over available products in proportion to those products' market share. This assumption fits a situation in which products are similar to each other, but vary in a tastes parameter, such as different flavors of yogurt or colors of apparel. By contrast, our demand model is a variant of the locational choice model which fits a situation in which some products are better substitutes for a given product than others. This was a very real feature of our applications; for example, the natural substitutes for a 14 inch tire are other 14 inch tires, not 15 inch tires. The locational choice approach allows us to estimate the probability a customer will substitute between specific product types. In one of our application we found the probability of substituting from Brand 1 to Brand 2 was 89%, but only 22% of substituting from Brand 2 to Brand 1. Hence Brand 2 customers were much more loyal. This is the only instance of which we are aware of measuring brand loyalty based on analysis of sales data.

The multinomial logit approach fits some real assortment problems and the locational choice approach fits others better, so both approaches are needed. Our providing an approach based on a locational choice demand model is analogous to the way in which, for insight models, Gaur and Honhon (2005) provided a locational choice complement to the multinomial logit based analysis of van Ryzin and Mahajan (1999).

Thirdly, we provide a thorough treatment of the important emerging topic of assortment localization. This includes incorporating a constraint on the amount of localization, which many retailers wish to impose to bound the administrative cost of localization, measuring how much assortment localization impacts revenue and understanding differences in the revenue impact of localization. Chong et al. (2001) don't deal with localization at all. Kök and Fisher (2007) allow a unique assortment for each store and hence they don't consider the constraint that a retailer may want to have a number of assortments considerably fewer than their number of stores, nor do they measure the impact on revenue of the degree of localization. We found that the revenue increase from localizing assortments by store varied considerably, from 12% in one case to only 6% in another. Moreover, we found surprisingly that the higher benefit for localization did not stem from greater variation in demand across stores but in the shape of the revenue-SKU count response curve; the retailer with lower benefits from localization had a steeper revenue response curve and hence was capturing a large fraction of total demand with their core assortment.

Last but not least, one of the retailers in our study has implemented a portion of our recommended assortment changes and we estimate that the revenue increase in revenue from these changes is larger than the annual same store revenue increases typically seen in this industry. This provides important validation of the value of an analytic approach to assortment planning.

Section 2 reviews relevant prior literature, Section 3 provides a precise formulation of the problem we solve, Section 4 describes our demand estimation process, Section 5 presents greedy heuristics for choosing assortments and Section 6 reviews the application of this process to two categories - snack

cakes and tires. Concluding remarks are offered in Section 7.

2 Literature Review

Assortment planning draws on a diverse literature in economics, operations management and marketing on consumer choice modeling, demand estimation and assortment optimization. We review here the prior research that is most related to the paper. For a more extensive review, the reader is directed to Kök et al. (2008).

2.1 Consumer Choice Modeling

Consumer choice models constitute the fundamental platform for assortment planning. They may be classified into two broad categories: (1) utility based choice models, and (2) exogenous demand models.

2.1.1 Utility Based Models

Utility based choice models assume that every customer associates a utility U_i with each product $i \in N$. In addition, there is a no-purchase option denoted $i = 0$, with associated utility U_0. When offered an assortment A, every customer chooses the option giving him the highest utility in $A \cup \{0\}$. The market share for each SKU $i \in A$ can then be evaluated once we know the distribution of utilities across the consumer population.

The Multinomial Logit (MNL) model is the most extensively studied utility-based model in the marketing and economics literature (Ben-Akiva and Lerman 1985, Anderson et al. 1992 , Guadagni and Little 1983). The MNL model assumes that the utilities U_i can be decomposed into a *deterministic* component u_i that represents the average utility derived by the population of customers, and a *random* component ξ_i that represents idiosyncrasies across customers. The ξ_i are assumed to be identical and independent Gumbel random variables with mean zero and scale parameter μ. Under these assumptions, the market share for each SKU $i \in A$ can be written as $\frac{\exp\left(\frac{u_i}{\mu}\right)}{\sum_{i \in A} \exp\left(\frac{u_i}{\mu}\right)}$.

The other commonly used utility based model is the Locational Choice model (LC) originally developed by Hotelling (1929) to study pricing and store location decisions of competing firms. Lancaster (1966, 1975) extended this work to a locational model of consumer product choice where every SKU $i \in N$ is represented as a bundles of attribute values $z_i = (z_i^1, z_i^2, ..., z_i^T) \in \mathbb{R}^T$. Each consumer has an ideal point $y \in \mathbb{R}^T$ that defines his most preferred attribute values. The utility this consumer associates to SKU i is $U_i^y = c - \tau \|y - z_i\|$, where c is the utility derived from his most preferred product y, and τ is the disutility associated with each unit of deviation from y. A consumer not finding his ideal product y in the assortment A would substitute the variant $j \in A$ that is located closest to his ideal point in the attribute space if $U_j^y > 0$, or else not purchase if $U_j^y \leq 0$.

2.1.2 Exogenous Demand Model

In the exogenous demand model, every consumer is assumed to have a favorite product i, and f_i is the share of consumers whose favorite product is i. A consumer whose favorite product is i buys it if $i \in A$; if $i \notin A$ they substitute to SKU $j \in A$ with probability α_{ij}. Under these assumptions, the market share of each SKU $i \in A$ is given by $q_i(A) = f_i + \sum_{j \notin A} f_j \alpha_{ji}$. The exogenous demand model has more degrees of freedom than the MNL and LC models and can accommodate extremely flexible substitution structures. However, as a consequence, it is also harder to empirically estimate, given the large number of parameters.

2.1.3 Comparison of Demand Models

All three demand models assume that customers have a favorite product and they buy that product if it is in the assortment. They also all assume that if a customer's favorite product is not in the assortment, they may substitute a different product. Where the models differ is in their assumptions about substitution behavior. The exogenous demand model is the most flexible model, allowing for any substitution structure, but it has many parameters and is hence difficult to estimate in practice. The MNL model assumes that the demand for a missing product which is not lost transfers to other products in the assortment in proportion to their popularity. By contrast, the Locational Choice model assumes that a given product may be more like some products than others and that substitution demand transfers to the product in the assortment that is most similar to a customer's preferred product. This assumption of the Locational Choice model most closely fits our applications. For example, in the tire assortment application, a 14" diameter tire is more like other 14" tires than it is like 15" tires. Accordingly, our demand model will be similar to the Locational Choice model, but we need to make several adaptations to that model to make it suitable for operational implementation.

2.2 Demand Estimation

Talluri and van Ryzin (2004) use sales transaction data (records of purchase time and product choice for each customer) to estimate demand in the context of airline revenue management. If customer arrivals, purchases and no-purchase outcomes are completely observed, then one can estimate the demand parameters using maximum likelihood methods. However, in practice only purchases are observed, and hence it is not possible to distinguish between a period with no arrivals, and a period with arrivals but no purchases. To overcome this problem Talluri and van Ryzin use the *Expectation-Maximization* (EM) algorithm of Dempster et al. (1977) to correct for the missing data. This method starts with arbitrary initial estimates of the demand parameters and uses Bayes rule to estimate the missing data. These estimates are now used to compute the conditional expected value of the likelihood (the expectation step), and the resulting expected log-likelihood function is maximized to generate new estimates for the demand parameters (the maximization step). This procedure is repeated until it converges. Anupindi et al. (1998) use a similar approach to estimate demand and substitution

probabilities for two products using sales transaction data from vending machines. Vulcano et al. (2009) use the EM algorithm to develop a procedure to estimate demand from sales transaction data, when the underlying substitution is governed by a MNL model. Kök and Fisher (2007) and Chong et al. (2001) also estimate demand as part of an assortment optimization process, and these papers are discussed in Section 2.3.

Fader and Hardie (1996) use an attribute based approach to estimate demand from sales transaction data. Consumer utility for a product is expressed in terms of its attributes. The attribute-level utilities are used to determine choice probabilities (based on an MNL model) and the likelihood of observing the given transaction data. They use maximum likelihood to estimate the parameters of the model. This results in a parsimonious estimation method as we only need to estimate attribute-level utilities, and has the added advantage of being able to estimate demand of new products. Batsell and Lodish (1981), Chiang (1991), Bucklin and Gupta (1992), Chintagunta (1993) use a similar model to investigate buying decisions of households. Bell et al. (2005) propose a new approach to obtaining (SKU)-level preferences from the underlying attribute preferences. The authors show that SKU-level parameters can be recovered by calculation from estimated attribute-level parameters, circumventing the need for direct estimation of the more complex SKU-level model.

2.3 Assortment Optimization

Assortment optimization research has been based on both stylized models intended to provide insight into structural properties of optimal assortments and decision support models intended to guide a manager planning assortments at a retailer.

2.3.1 Stylized Models

In a pioneering paper, van Ryzin and Mahajan (1999) study an assortment planning problem under a MNL consumer choice model and show that the optimal assortment consists of a certain number of the highest utility products. Mahajan and van Ryzin (2001) study the same problem allowing for stock-out substitution and develop heuristics based on a sample path approach. Maddah and Bish (2007) extend the van Ryzin Mahajan model by including pricing decisions. Cachon et al. (2005) study the van Ryzin and Mahajan (1999) model to understand the indirect benefits of an assortment in the presence of consumer search. Their key finding is that it might be optimal to add an unprofitable item to the assortment to prevent consumers from searching and hence reap indirect benefits.

Gaur and Honhon (2005) show that for a locational choice model, the products in the optimal assortment are located far apart from each other in the attribute space, indicating maximum differentiation, with no substitution between products in the assortment. This implies that the most popular product may not be carried in the optimal assortment, contrasting the results of van Ryzin and Mahajan (1999).

2.3.2 Decision Support Models

Green and Krieger (1985) formulate a product line design (PLD) problem in which there are m consumer segments indexed by i, and n products, indexed by j. Every consumer segment i has a utility associated with product j denoted by u_{ij}. A consumer chooses from all available products the one that maximizes his utility. Green and Krieger then formulate the problem of which k products out of the n should a firm select so as to maximize (1) consumer welfare or (2) firm profits, and propose solution heuristics. Green and Krieger 1987a,b, McBride and Zufryden 1988, Dobson and Kalish 1988, Kohli and Sukumar 1990 extend this line of research. Green and Krieger (1992) apply their approach to the product line positioning of a new dietary food supplement using data from conjoint-analysis, a statistical technique to elicit consumer valuations for product attributes by asking them to compare pairs of products. Belloni et al. (2008) compare the performance of different heuristics for product line design and find that the greedy and the greedy-interchange heuristics perform extremely well.

Smith and Agrawal (2000) use an exogenous demand model to formulate an assortment planning problem as an integer program. Their paper is noteworthy in that they model substitution demand and consider the optimal amount of inventory to stock, in a decision support setting. They solve a number of small problems by complete enumeration to demonstrate how assortment and stocking decisions depend on the nature of assumed substitution behavior, and also propose a heuristic to solve larger problems. We build on their exogenous demand model by developing a methodology for estimating model parameters and applying it to real data.

Chong et al. (2001) develop an assortment modeling framework based on the Guadagni and Little (1983) brand-share model. They account for product similarity within a category by incorporating new measures. They use consumer-level transaction data over multiple shopping trips to empirically estimate the parameters of the model and use a local improvement heuristic to suggest an alternative assortment with higher revenue. It is important to note that although their model can implicitly predict SKU level demand, their explicit focus is on brand-shares, and hence they don't create or measure the accuracy of SKU level demand forecasts. They report an average 50% mean squared error of predicting brand choice at the customer level across different product categories.

Kök and Fisher (2007) use an exogenous demand model to study a joint assortment selection and inventory planning problem in the presence of shelf-space constraints. They note that the constrained shelf space implies that average inventory per SKU is inversely related to the breath of the assortment and focus in particular on the tradeoff between assortment breadth vs in-stock level for the SKUs carried. They consider substitution and assume that substitution demand accrues to available products in proportion to their original market share. They provide a process for estimating demand and substitution rates and apply their method at a large Dutch grocery retailer.

Caro and Gallien (2007) is one of the first papers to incorporate demand learning in a dynamic assortment planning problem setting. The objective of their paper is to help a retailer optimally modify their product assortment over time to maximize overall profits across the selling horizon. They formulate the assortment optimization problem as a multiarmed bandit, and use Bayesian learning methods

to update demand. They propose a closed-form dynamic index policy using dynamic programming techniques and show that their policy is near optimal for a range of numerical experiments.

3 Problem Formulation

We seek optimal assortments for a retail category over a defined planning horizon. For the sake of concreteness, we assume that the goal is to maximize revenue, since this was the primary concern in our two applications, although it should be clear that adapting our approach to maximizing any linear profit function, such as unit sales or dollar gross margin, is straightforward. We define the following parameters.

$N = \{1, 2, ..., n\}$	Index set of all possible SKUs a retailer could carry in this category
$M = \{1, 2, ..., m\}$	Index set of all stores
K	Maximum number of SKUs per assortment
L	Maximum number of assortments
D_i^s	Number of customers at store s for whom SKU i is their preferred product
p_i	Price of SKU $i \in N$

The parameters K and L would be specified by the retailer. K would follow directly from current macro level assortment planning practices which determine, as described in the Introduction, the number of SKUs to be carried in a category. For expositional simplicity, K does not vary by store, but it would be easy to modify our process to enable K varying by store, and in fact we do this in our computational work. L would be chosen to lie between 1 and m and to trade off the greater revenue that come with larger L against the administrative simplicity that comes with smaller L. As will be seen, our solution approach makes it easy to solve this problem for all possible values of K and L, thus providing the retailer with rich sensitivity analysis to guide their choice of these parameters.

We define an assortment to be a set $A \subseteq N$ with $|A| \leq K$ and let $a(s)$ denote the assortment assigned to store s. A solution to the assortment optimization problem is completely defined by the portfolio of assortments A_l, $l = 1, 2, ..., L$ and $a(s)$, $s \in M$.

Our demand model needs more extensive discussion. We assume that a consumer shopping this category in store s has a most preferred SKU $i \in N$, but might be willing to substitute other SKUs if $i \notin A_{a(s)}$. We define $R(i) = \{j_1, j_2, ..., j_{i^*}\} \subseteq N$ to be the set of acceptable substitutes for a consumer whose first preferred choice is SKU i, rank ordered so that j_1 is their most preferred substitute and j_{i*} their least preferred. If a consumer with most preferred SKU i finds $i \in A_{a(s)}$ when they shop the store, then we assume they buy it. Otherwise, they buy the substitute ranked highest in $R(i)$ that is in the assortment, $j(i, A_{a(s)}) = \arg\min_k \{j_k \in R(i) \cap A\}$ with probability $\pi^s_{ij(i,A_{a(s)})}$. We allow substitution probabilities to vary by store because we found in our applications that they did in fact vary by store. For example, we will see in Section 6 that the willingness of a consumer to substitute to a higher priced product varies by store and is correlated with median income in the zip code in which the store is located. The revenue $R_s(A_{a(s)})$ earned by store s using assortment $A_{a(s)}$ can then be written as

$$\left(\sum_{i \in A_{a(s)}} p_i D_i^s + \sum_{i \notin A_{a(s)}, j = j(i, A_{a(s)})} D_i^s \pi_{ji}^s p_j \right)$$

The first term in the expression for $R_s(A_{a(s)})$ is the revenue from customers whose most preferred SKU is contained in the assortment and the second term is the expected substitution revenue from customers whose most preferred SKU was not in the assortment.

We can now state the assortment optimization problem as choosing A_l, $|A_l| \leq K$, $l = 1, 2, ..., L$ and $a(s)$ for all $s \in M$ to maximize $\sum_{s \in M} R_s(A_{a(s)})$.

4 Estimating Model Parameters

In this section, we describe a methodology for estimated the parameters in the formulation presented in section 3: $D_i^s, \pi_{ij}^s, R(i)$ and p_j for all i, j and s. D_i^s and π_{ij}^s are estimated at the store level, but for expositional simplicity we will drop the store superscript in the discussion that follows.

Our primary input for estimation will be store-SKU sales of products currently carried by the retailer during a prior history period. Let A denote the assortment carried in a particular store and x_i the sales of SKU $i \in N$, during a history period. We describe below how to use the x_is to estimate the model parameters.

4.1 Estimating demand and substitution probabilities

The central idea behind our approach for estimating demand and substitution probabilities is to view a SKU as a collection of attribute values, use historical sales data to estimate the demand share of each attribute value, and finally estimate the demand share of any SKU as the product of the demand shares of its attribute values. For example, the attributes for a bed sheet might be size, color, fabric type and thread count. More generally, we have encountered three types of attributes in the various examples we have considered.

1. Functional fit, attributes that define whether a given product works or not for a particular consumer need. Examples include most automobile parts and the size of tires, batteries, shoes, apparel and sheets There is no substitution across these attributes. Someone looking for a twin sheet is not going to buy a queen sheet.

2. Price/value, such as the good-better-best segmentation frequently used by retailers. A higher thread count sheet is regarded as a better product and accordingly carries a higher price. This attribute has a logical ordering in that the closer the price of a product to a customer's preferred price/value point, the more preferred it is as a substitute.

3. Taste, such as flavor, color and fabric type.

We introduce some notation to formalize the notion of a SKU as a bundle of attribute values. Let T denote the number of attributes, Ω the set of T attributes, ϕ_t the number of values for attribute t and Φ_t the set of values for attribute t. Each SKU $i \in N$ now corresponds to a vector of attribute values $(1_i, 2_i, ..., T_i)$, where $t_i \in \Phi_t$. The total number of potential SKUs in this product category is given by $n = \prod_{t \in \Omega} \phi_t$.

Let f_i denote the fraction of consumers whose first choice is product i and f_{ti_t} the fraction of consumers whose first choice on attribute t is the value i_t. We will estimate the demand share f_i of a SKU $i = (1_i, 2_i, ..., T_i)$ as $f_i = \Pi_{t \in \Omega} f_{tt_i}$. The approximation inherent in this estimation method is that a consumer's preferences for the various attributes are independent, which clearly may not be true. For example, a college student shopping for *twin size* bed-sheets might have a different color preference than an urban homemaker shopping for *queen size* sheets. None the less, this attribute decomposition assumption is extremely useful in that it enables estimation of the demand for the large number of SKUs not carried by the retailer. Also, this approach reduces the number of parameters to be estimated since there are usually fewer attribute values than SKUs, and this is important since the amount of data available for estimation tends to be sparse.

In an application we can check the accuracy of this approximation by comparing demand estimates with actual sales for the SKUs currently carried. We can also examine sales data to see if the assumption is violated (e.g. how different is the color sales mix on twin vs. queen sized sheets). If necessary, we can mitigate the assumption by combining attributes that interact strongly into a single attribute, although this increases the number of parameters that need to be estimated.

In a similar vein, we assume that the probability that a consumer with most favorite SKU i will purchase SKU $j \in A$ can be decomposed in terms of substitution probabilities across the SKU's constituent attribute values, i.e. $\pi_{ij} = \prod_{t \in \Omega} \pi_{t,t_i t_j}$.

To estimate $R(i)$, we first assume that consumers have an importance rank ordering of the attributes and follow a hierarchical process while making their purchasing decisions. We have found that attribute rankings can usually be obtained from a combination of market-research studies and managerial intuition. In judging how close a substitute product is to their first choice, we assume that consumers follow a lexicographic approach in which differences between more important attributes dominate differences between less important attributes. For example, if the sheet attributes are ranked size, fabric, color and thread count and a consumer's ideal product is a twin, cotton, blue, 200 thread count sheet, then a twin, cotton, red, 100 thread count sheet is farther from their ideal point than a twin, cotton, blue, 300 thread count sheet.

The approach we follow to rank ordering attribute values depend on the type of attribute. For *functional fit* attributes, where we assume no substitution across attribute values, any product with a different attribute value from a consumer's ideal point is in essence infinitely far from the ideal point and can appear at the end of $R(i)$, although in this case, the ranking is irrelevant, since the probability of substitution is assumed to be 0. For attributes that can be quantified, such as *price/value,* then alternative products can be ranked based on how much their value differs from the ideal value. Some *taste* attributes can also be treated this way. For example color can be quantified by the RGB scale by

the intensity of red, green and blue. For taste attributes that can't be quantified, we rely on market research studies and buyer intuition to estimate relative attractiveness of substitute products for a given preferred product.

Given our assumptions, we can write the probability $F_j(A)$ that a consumer would purchase $j \in A$, when offered assortment A, as $F_j(A) = f_j + \sum_{i \notin A} f_{ij(i,A)} \pi_{ij(i,A)}$. Let $F(A) = \sum_{j \in A} F_j(A)$ denote the probability that a customer shopping in this category would make a purchase from assortment A. Then, the likelihood of observing sales data $x = \{x_i\}_{i \in A}$ is given by

$$L = K \prod_{j \in A} \left[\frac{F_j(A)}{F(A)} \right]^{x_j}$$

where the proportionality constant is $K = \frac{(\sum x_{j \in A})!}{\prod_{j \in A} x_j!}$. The maximum likelihood estimates (MLE) for the parameters (f, π) can be obtained by maximizing the log-likelihood function

$$LL(f, \pi) = \sum_{j \in A} x_j \log F_j(A) - \left(\sum_{j \in A} x_j \right) \log F(A) \tag{1}$$

subject to the constraints

$$\sum_{i_t \in \Phi_t} f_{t i_t} = 1, \quad \forall t \in \Omega \tag{2}$$

$$\pi_{t,t_i t_j} \in [0,1] \quad \forall t \in \Omega, \forall t_i,\ t_j \in \Phi_t \tag{3}$$

Given the complex nature of the log-likelihood function, it is not possible to derive analytical results. Hence we resort to numerical optimization methods based on gradients after transforming the problem into an unconstrained optimization problem, by reparametrizing f_{t,t_i} and $\pi_{t,t_i t_j}$ as

$$f_{t,t_i} = \frac{\exp(\widehat{f}_{t,t_i})}{\sum_{t_j \in \Phi_t} \exp(\widehat{f}_{t,t_j})}, \ \ t_i \neq \phi_t \text{ and } \widehat{f}_{\phi_t} = 1 \tag{4}$$

$$\pi_{t,t_i t_j} = \frac{\exp(\widehat{\pi}_{t,t_i,t_j})}{1 + \exp(\widehat{\pi}_{t,t_i t_j})} \tag{5}$$

Inspecting the log-likelihood function in detail reveals that it may not always be concave. We illustrate this using the following example. Consider a product with two attributes $t = \{1, 2\}$ and attribute levels $\Phi_1 = \{A, B, C\}, \Phi_2 = \{a, b, c\}$. The assortment offered consists of the SKUs $\{Aa, Ab, Ac, Ba, Bb, Cb, Cc\}$ with sales given in Table 1. We assume that the only possible substitutions are from Bc to Bb and from Ca to Cb. We wish to estimate the six attribute shares and two substitution probabilities given by $\gamma = \{f_A, f_B, f_C, f_a, f_b, f_c, \pi_{Bc-Bb}, \pi_{Ca-Cb}\}$. A necessary condition for the log-likelihood function to be concave is that for any two points γ_1, γ_2,

$$LL\left(\frac{\gamma_1 + \gamma_2}{2} \right) \geq \frac{LL(\gamma_1) + LL(\gamma_2)}{2}$$

Let us denote $\gamma(\alpha) = \{0.20, 0.30, 0.50, 0.40, 0.50, 0.10, \alpha, 0.30\}$. If we now let $\gamma_1 = \gamma(0)$ and $\gamma_2 = \gamma(1)$, we find that the log-likelihoods computed using Equation (1) are $LL(\gamma_1) = -1959.15$, $LL(\gamma_2) = -2181$ and $LL\left(\frac{\gamma_1+\gamma_2}{2}\right) = -2072.94$. This implies that $L\left(\frac{\gamma_1+\gamma_2}{2}\right) - \frac{LL(\gamma_1)+LL(\gamma_2)}{2} = -2.87$ and hence the log-likelihood function is not concave. In fact, such non-concave demand functions are commonly encountered in models with substitution (e.g. Netessine and Rudi 2003). The absence of concavity implies that numerical optimization methods may not always converge to the global maximum. We handle this issue by running the optimization algorithm from several randomly generated starting points. This does not guarantee global convergence but lowers the chances of the algorithm getting stuck at a local maximum. Mahajan and van Ryzin (2001) use a similar approach to compute the optimal inventory levels using the sample path gradient algorithm.

Once we obtain the MLE for the demand shares and substitution probabilities, we can compute the total demand for the product category by dividing total sales by the total market share captured. In other words

$$D = \frac{\sum_{i \in A} x_i}{F(A)}$$

Then D_i is computed as $f_i D$.

Our demand model most closely resembles the LC model, but with three important differences: (1) the probability of purchasing the closest substitute from the assortment is either 0 or 1 (depending on no-purchase utility) in the LC model, while we have an explicit parameter, (2) consumers are assumed to be distributed in a continuous space in the LC model, while we discretize this space and restrict consumers to occupy only those positions at which products are located and (3) all attribute values in the LC model have a numeric value which allows the calculation of distance between products and identification of the nearest product to a given ideal point, whereas the nearest product to an ideal point in our model is identified via the sets $R(i)$, which can be thought of as engendering a distance metric for attributes like taste that can't be located in a space. In fact one could view our approach as adding the enhancements needed to make the LC model operational in practice.

4.2 Setting Prices for New SKUs

We endeavored to set prices on SKUs not currently carried by the retailer in a way that would be consistent with their current pricing policy. We assume that prices on existing SKUs were set in relationship to the value of the SKU to a consumer and that consumer value is correlated with attribute values. Hence, we regressed $\log(p_i)$ on the attribute values $\Phi_t, \forall t \in \Omega$ to obtain the pricing equation:

$$\log(p_i) = \alpha_0 + \sum_{t \in \Omega} \sum_{f_j \in \Phi_t} \beta_{tt_j} \mathbf{I}\{t_j = t_i\} \tag{6}$$

This is a hedonic pricing equation, and has been extensively used in the economics (Rosen 1974, Goodman 1998, Pakes 2003).

5 Heuristics for Choosing Assortments

If there were no substitution, then the assortment problem could be easily solved by a greedy algorithm that chose SKU's in decreasing order of their revenue contribution, but substitution makes the objective function nonlinear, because the contribution of a SKU depends in part on it's substitution demand, which depends on which other SKUs are in the assortment. As a result, the assortment problem would be complex to solve optimally. We still apply the greedy algorithm in this case, although optimality cannot be guaranteed when demand substitution is present.

In this section we define various greedy heuristics for choosing the assortments $A_l, l = 1, 2, ..., L$ and the specification $a(s), s \in M$ of the assortment assigned to store s. We first define a greedy heuristic for finding a single assortment $A^G(S)$ for a specified subset of stores S. This heuristic will then be used as a subroutine in forward and reverse greedy heuristics for finding a portfolio of L assortments and assignment of stores to these assortments. In the statement of Greedy(S) below we define $R_s(\Phi) = 0$.

Greedy(S)

1. INITIALIZE $A^0 = \Phi$, $k = 1$
2. WHILE $(k \leq K)$ DO
   (a) $j_k = \arg\max_{j \notin A^{k-1}} \sum_{s \in S} R_s \left(A^{k-1} \cup \{j\} \right)$
   (b) $A^k = A^{k-1} \cup \{j_k\}$
   (c) $k = k + 1$
   (d) END WHILE
3. RETURN $A^G(S) = A^K$

Forward Greedy

In the forward greedy heuristic, we first apply Greedy(S) $m+1$ times with $S = M$ and $S = \{s\}, s \in M$, initialize $A_1 = A^G(M)$, and assign all stores to this assortment. If $L > 1$, we identify the assortment $A^* \in \mathbf{D} = \{A^G(\{s\} \mid s \in M\}$ to add that would maximize the incremental revenue gain. To identify A^*, we calculate the incremental revenue gain from adding each assortment $A \in \mathbf{D}$ by reassigning stores to their revenue maximizing assortment in $A_1 \cup A$ and calculating the increase in revenue due to the reassignment. We then choose as A^* the assortment that gave the greatest revenue in this process. At any point in the algorithm, we have a portfolio of l assortments and l store clusters defined by the assignment of stores to assortments. As long as $l < L$, we add to this portfolio the assortment that leads to the highest increase in revenue and reassign stores to the enhanced set of assortments.

1. INITIALIZE

(a) $l = 1$, $A_1 = A^G(M)$

(b) $\mathbf{C}^l = \{A_i \mid i = 1, 2, ..., l\}$

(c) $\mathbf{D} = \left\{A^G(\{s\}) \mid s \in M\right\}$

(d) $a(s) = 1, \forall s \in M$

2. WHILE $(l < L)$ DO

(a) $A^* = \arg\max_{A \in \mathbf{D}} \sum_{s \in M} R_s\left(A_{a^*(s)}\right) - R_s\left(A_{a(s)}\right)$
where $a^*(s) = \arg\max_{i: A_i \in \mathbf{C}^l \cup A} R_s(A_i)$

(b) $l = l + 1$

(c) $A_l = A^G(\{s^*\})$

(d) $a(s) = \arg\max_{i \leq l} R_s(A_i)$

(e) $S_i = \{s \in M \mid a(s) = i\}, i = 1, 2, ..., l$

(f) $A_i = A^G(S_i), i = 1, 2, ..., l$

END WHILE

3. RETURN $A_l, l = 1, 2, ..., L$, $a(s), s \in M$

Reverse Greedy

In the reverse greedy heuristic, we first apply Greedy(S) m times with $S = \{s\}$ for all $s \in M$, initialize $A_i = A^G(M)$ for $i = 1, 2, ..., m$ and set $a(s) = s$ for all $s \in M$. If $L < m$, we identify the single assortment $A^* \in \mathbf{D} = \{A_i \mid i = 1, 2, ..., m\}$ to delete that would minimize the incremental revenue loss. We calculate the incremental revenue loss from deleting any assortment $A \in \mathbf{D}$ by reassigning stores to their revenue maximizing assortment in $\mathbf{D} - A$ and calculating the loss in revenue due to the reassignment. At any point in the algorithm, we have a portfolio of l assortments and l store clusters defined by the assignment of stores to assortments. As long as $l > L$, we delete one assortment from this portfolio that leads to the least loss in revenue and reassign stores to the reduced set of assortments.

1. INITIALIZE

(a) $l = m$

(b) $\mathbf{D} = \left\{A^G(\{s\}) \mid s \in M\right\}$

(c) $a(s) = s, \forall s \in M$

2. WHILE $(l > L)$ DO

(a) $A^* = \arg\min_{A \in \mathbf{D}} \sum_{s \in M} R_s\left(A_{a(s)}\right) - R_s\left(A_{a^*(s)}\right)$
where $a^*(s) = \arg\max_{i: A_i \in \mathbf{D} - A} R_s(A_i)$

(b) $\mathbf{D} = \mathbf{D} - A^*$

(c) $l = l - 1$

(d) $a(s) = \arg\max_{i:A_i \in \mathbf{D}} R_s(A_i)$

(e) $S_i = \{s \in M \mid a(s) = i\}, i = 1, 2, ..., l$

(f) $A_i = A^G(S_i), i = 1, 2, ..., l$

(g) $\mathbf{D} = \{A_i \mid i = 1, 2, ..., l\}$

END WHILE

3. RETURN $A_i, i = 1, 2, ..., L$, $a(s), s \in M$

The solution $A_i, i = 1, 2, ..., L$ and $a(s), s \in M$ provided by the Forward Greedy and Reverse Greedy heuristics can be further improved by an iterative procedure of reassigning stores to assortments and re-assorting each cluster of stores by using greedy to select the assortment. This iteration between store assignment and assortment selection continues as long as the incremental revenue ΔR exceeds a predefined threshold ϵ. This iterative procedure can also be carried out within the WHILE-DO loop in Step 2, but does not provide any significant improvement in our applications.

4. WHILE $(\Delta R > \epsilon)$ DO

(a) $R = \sum_{s \in M} R_s(a(s))$

(b) $S_i = \{s \in M \mid a(s) = i\}, i = 1, 2, ..., L$

(c) $A_i = A^G(S_i), i = 1, 2, ..., L$

(d) $a(s) = \arg\max_{i \leq L} R_s(A_i), \forall s \in M$

(e) $\Delta R = \sum_{s \in M} R_s(a(s)) - R$

END WHILE

Interchange heuristics (Manne 1964, Teitz and Bart 1968, Achabal et al. 1982) have been defined for a variety of problems and can also be applied here to the results of Greedy(S), Forward Greedy and Reverse Greedy. An interchange heuristic applied after Greedy(S) would test all exchanges of an $i \in A^G(S)$ with a $j \notin A^G(S)$ to see if they increased revenue, and continue to make revenue increasing interchanges as long as they could be discovered. The output of either the Forward or Reverse Greedy's defines L sets of stores $S_i, i = 1, 2, ..., L$ determined by the variables $a(s)$. Interchange applied to this solution would interchange two stores in different store clusters and reapply Greedy(S) to each of these clusters to determine if revenue increased. Such revenue increasing interchanges would be made as long as they could be discovered.

6 Results

This section describes the application of our approach to two examples: planning the snack cakes assortment for a regional convenience chain and the tire assortment for a national tire retailer.

6.1 Regional Convenience Chain

The retailer offered snack cakes in 60 flavors, two brands $\{B_1, B_2\}$, and several different package sizes in 140 stores. We restricted our analysis to the top 23 flavors that accounted for 95% of revenue. Although there were several different package sizes, what mattered from a consumer's perspective was whether the size was single-serve or family size, and hence we grouped sizes into *Single Serve* (S) and *Family Size* (F). Further, because the retailer advised us that brand shares and willingness to substitute varied by size, we combined brand and size to obtain a single attribute called *Brand-Size*. This resulted in 92 possible SKUs (23 flavors, 4 brand-sizes), of which 52 were being offered by the retailer in at least one store. The number of SKUs offered across stores varied between 24 and 52, and averaged 40.3.

The set of attributes, Ω, is given by $\Omega = \{v, b\}$, where v refers to *Flavor* and b to *Brand-Size*. An internal market research study on the industry comissioned by the retailer showed that *Flavor* was the most important attribute for a consumer purchasing from this category. Hence, we assumed that the probability of substituting across flavors is negligible and could be set to 0. The retailer also believed that there is negligible substitution between sizes S and F, so this substitution was assumed to be 0. Thus $R(j)$ was comprised of a single SKU with the same size and flavor as j, but a different brand. We need to estimate the 23 flavor shares, 4 brand-size shares, and 4 substitution probability parameters $\{\pi_{B_1S-B_2S}, \pi_{B_2S-B_1S}, \pi_{B_1F-B_2F}, \pi_{B_2F-B_1F}\}$. Let v_j and b_j denote the flavor and brand-size, respectively, of SKU j. Also, recall based on the notation defined in Section 4.1, that f_{b,b_j} denotes the share of demand of brand-size b_j, and f_{v,v_j} denotes the share of demand of flavor v_j. The sales share of SKU $j \in A$ at any store can then be expressed as

$$F_j(A) = f_{b,b_j} f_{v,v_j} + \sum_{i \in A^c \cap R(j)} f_{b,b_i} f_{v,v_i} \pi_{b_i,v_i-b_j,v_j}$$

where $A^c = N - A$. We can then write Equations (1)-(3) from Section 4.1 as maximizing

$$LL(f, \pi) = \sum_{j \in A} x_j \log F_j(A) - \left(\sum_{j \in A} x_j \right) \log F(A) \tag{7}$$

subject to the constraints

$$\begin{aligned} \sum_{v_j \in \Phi_v} f_{v,v_j} &= 1 \\ \sum_{b_j \in \Phi_b} f_{b,b_j} &= 1 \qquad (8) \\ \pi_{b,b_i b_j} &\in [0,1] \quad , b_j \in \{B_1S, B_2S, B_1F, B_2F\} \qquad (9) \end{aligned}$$

The Lagrangian can be written as

$$\begin{aligned} H &= \sum_{j\in A} x_j \log F_j(A) - \sum_{j\in A} x_j \log F(A) - \lambda_1 \left(\sum_{v_j \in \Phi_v} f_{v,v_j} - 1 \right) - \lambda_2 \left(\sum_{b_j \in \Phi_b} f_{b,b_j} - 1 \right) \\ &\quad - \sum_{b_j \in \Phi_b} \mu_{b_j} \pi_{b_j,b_j^*} - \sum_{b_j \in \Phi_b} \gamma_{b_j} (1 - \pi_{b_j,b_j^*}) \end{aligned} \tag{10}$$

Applying first-order conditions with respect to the flavor shares f_{v,v_i}, we get

$$\frac{\partial H}{\partial f_{v,v_i}} = \frac{1}{f_{v,v_i}} \left[\sum_{j\in A, v_j = v_i} x_j - \frac{\sum_{j\in A} x_j}{F(A)} F_j(A) \right] - \lambda_1 = 0 \tag{11}$$

Multiplying Equation 11 by f_{v,v_i}, and adding the equations across all values of i_v, gives us

$$\lambda_1 = \sum x_j - \frac{\sum_{j\in A} x_j}{F(A)} \sum_{j\in A} F_j(A) = 0 \tag{12}$$

Hence, we get

$$\frac{\sum_{j\in A, v_j = v_i} F_j(A)}{F(A)} = \frac{\sum_{j\in A, v_j = v_i} x_j}{\sum_{j\in A} x_j} \tag{13}$$

which on simplification yields

$$f_{v,v_i} = \frac{\sum_{j\in A, v_j = v_i} x_j}{\sum_{j\in A} x_j} \frac{F(A)}{\sum_{j\in A, v_j = v_i} \frac{F_j(A)}{f_{v,v_j}}} \tag{14}$$

Note that $\frac{F_j(A)}{f_{v,v_j}} = f_{b,b_j} + \sum_{i\in A^c \cap R(j)} f_{b,b_j} \pi_{b_i,v_i - b_j,v_j}$ and can be expressed in terms of the brand-size shares and substitution probabilities, and $F(A)$ can be calculated by using the fact that $\sum_{v_i} f_{v,v_i} = 1$. This reduces the number of parameters to be estimated from 31 to 8, as the log-likelihood function can now be written as a function of the brand-size shares (f_b) and substitution probabilities (π) alone. As was described in Section 4, we use numerical methods based on gradients from randomly generated starting points to maximize this log-likelihood function.

We used a hedonic regression as described in Section 4.2 to assign price to SKUs not currently offered. The regression R^2 is 85.5% and confirms our assumption that prices on existing SKUs were set to be correlated with attribute values that determine consumer value. We normalized the estimated prices by a scaling factor so as to equalize the estimated revenue at the chain level to the actual revenue, which will facilitate comparison of new optimized assortments with current revenue.

We used Store-SKU sales data for the six month period from July 2005 to December 2005 to estimate model parameters at each of the 140 stores. Tables 2 and 3 show the average parameter estimates across all stores.

We measured the overall estimation error across all stores, by computing the Sales-Weighted Mean-Absolute-Deviation (MAD) of estimated Store-SKU sales shares from actual Store-SKU sales shares, as given by $\frac{\sum_{s\in M}\sum_{i\in A(\{s\})}\left|\frac{x_i^s}{D_s}-F_i^s(A)\right|x_i^s}{\sum_{s\in M}\sum_{i\in A(\{s\})}\frac{x_i^s}{D_s}}$. We measure MAD in terms of sales shares because unit sales are significantly influenced by overall growth or shrinkage in the category, whereas sales shares are not. Moreover, our assortment choices are determined solely by sales shares, so these are the parameters important to our analysis. The MAD for this retailer was calculated to be 16.4% at the Store-SKU level and 6.2% at the Chain-SKU level.

To validate our results, we used Store-SKU sales data for the six month period from July 2007 to December 2007. For this period, we had data for only 54 of the 140 stores in the chain. We used the previously estimated parameters to compute the share of sales for each SKU for the validation period (July 2007 to December 2007), at the store and chain levels and compared it with the actual sales shares. The sales-weighted MAD of predicted sales shares from the actual sales shares was calculated to be 40.1%[2] at the store-SKU level and 25.8% at the chain-SKU level. Note that our chain SKU validation error of 25.8% and new SKU error of 16.2% compares favorably to the 30.7% MAPE for chain sales of two new SKUs reported by Fader and Hardie (1996), the only prior reporting of which we are aware, of the errors of forecasts for new SKUs based on sales data.

One of the recommended SKUs that was added to the assortment was Butterscotch in Brand 2 Single Serve. The MAD of the predicted sales shares from the actual sales shares at the chain level for the newly introduced SKU was 16.2%.

Sources of error affecting both the calibration and validation samples include random fluctuation in sales and the approximation of representing SKU shares as the product of attribute shares. Additional sources of error for the validation sample included changes in relative prices across brand-sizes, which affect share and substitution probability estimates, and a steady increase in the demand shares of some newer flavors.

Figure 1 show the results of applying the forward greedy heuristic[3] described in Section 5 to compute optimized assortments at the chain and store level, varying the number of SKUs in an assortment. Figure 1 shows revenue as a function of SKU count, for $L = 1$ and $L = m$. These revenue curves represent the two extremes of assortment localization with $L = 1$ corresponding to no localization and $L = m$ corresponding to a unique assortment for each store. These curves are a free by-product of Greedy(S) described in Section 5, as each iterative step generates an assortment with $k \leq K$ SKUs, thereby letting us compute revenues as a function of SKU count.

To quantify the potential improvement in revenue, we compare the assortments we generated for $L = 1$ and $L = m$ to the current assortment, which had a revenue of \$6.19 million. Because SKU count varied somewhat by store, in computing the revenue of the $L = 1$ assortment, we first generated the greedy

[2]The 54 stores in this analysis had a higher MAD of 24.6% at the Store-SKU level (and 10.3% at the Chain-SKU level) in the calibration sample, as compared to 16.4% at the Store-SKU level (and 6.2% at the Chain-SKU level) for the whole chain. This suggests that if we use data for the whole chain, then the Store-SKU level MAD in the validation sample, may proportionally come down from 40.1% to 26.7%.

[3]The results obtained by applying the reverse greedy heuristic were not significantly different.

solution for a SKU count equal to the maximum SKU count across all stores, and then for store s with SKU count K_s, we used the first K_s SKUs chosen by greedy. The revenue of this solution was \$8.01 million, a 29.2% increase over the current revenue. The revenue for store-specific assortments with SKU counts of K_s for each store was \$8.75 million, a 41.4% increase over the current revenue. We thus see that maximum localization adds 12.2% beyond the 29.2% achieved by chain level optimization.

These revenue estimates are based on the calibration sample parameter estimates and hence impacted by estimation errors described previously. To determine how much these estimated revenue improvements would be eroded during implementation due to forecast errors, we used parameter estimates based on the validation sample Store-SKU sales data (July 2007 to December 2007) to estimate the revenue that would have been achieved had our recommended assortment changes been implemented in the validation period. Because we were only able to obtain data on 54 stores during our validation period (vs. 140 stores in the calibration period), we compared the revenue lift of these 54 stores based on the calibration sample with the lift based on the validation sample. The calibration period revenue of the current 54 store assortment was \$2.43 million and the store-optimal assortment estimated revenue based on the calibration period was \$3.44 million, which is a 41.5% increase. Recomputing revenue estimates for these two assortments during the validation period gives \$2.5 million and \$3.0 million, a 20% increase. We see that the revenue improvement has eroded by half because the calibration period parameter estimates are an imperfect representation of the validation period reality, although a 20% revenue increase is still economically very significant. The calibration period revenue of the single chain-wide assortment was \$2.74 million indicating that localization would lead to an incremental 12.8% gain in revenues. The corresponding number calculated based on the validation period was \$2.69 million indicating a localization lift of 7.6%.

Figure 1 shows that higher revenue can be achieved from store specific assortments. However, this retailer believed that it would be unmanageably complex to have more than 6 assortments for the chain, because for each assortment they needed to develop a diagram (called a Planogram) showing how the product would be displayed in the store.

To quantify the benefit of a realistic localization, we applied our assortment heuristics for $L = \{1, 2, ..., 6\} \cup \{m\}$, keeping $K = 40$ for all stores. Table 4 show revenue as a function of L. We note that complete localization increases revenue to \$8.11 million compared to the revenue of \$7.38 million for $L = 1$.However, 76.7% of this increase can be achieved with just 6 different assortments, suggesting that a small amount of localization can have a big impact.

Table 5 compares the chain-wide optimal assortment with the retailer's current assortment at a representative store. The result that was initially most surprising to the retailer is that the optimal assortment completely drops *Brand 1* in the *Family Size.* This is easily explained by looking at Table 3 which shows that 89% of consumers are willing to switch from B_1 to B_2 in the family-size segment. Hence, by not offering B_1F, which accounts for 6% of primary demand, the retailer only loses $6\% * 11\% \sim 0.7\%$ of demand, which is more than made up by carrying B_2S which has a higher base demand (27%) and higher loyalty ($\sim 78\%$), in more flavors. The 6 least popular flavors Pineapple/Cheese, Blueberry, Marshmallow, Apple, Cream, and Glazed, are also dropped since the revenue

lost is small, and this frees up room in the assortment to add additional brand-sizes in the more popular flavors. More generally, the number of Brand-Sizes offered is directly linked to the sales rank of a flavor. These results made sense to the retailer and we are in discussions with them to pilot these recommendations at a group of stores.

6.2 National Tire Retailer

Tire attributes include brand, size, mileage warranty, price, speed rating and load limit. However, these attributes are not independent of each other. For example, size is positively correlated with load limit, while mileage warranty is correlated with speed rating. Based on discussions with management and analysis of attribute data, we concluded that brand, size and mileage warranty were the fundamental defining attributes of a tire relevant to assortment planning.

The retailer offered several nationally advertised brands that they believed were equivalent to the consumer and which we denote National (N) and treat as one brand. They also offered three house brands of decreasing quality, which we denote as *House 1* ($H1$), *House 2* ($H2$) and *House 3* ($H3$), where $H1$ is the highest quality and most expensive house brand. There were a large number of distinct mileage warranties offered, but some of these varied only slightly and hence were believed by the retailer to be equivalent to consumers. Therefore, we aggregated the mileage warranties into three levels of *Low* ($15,000 - 40,000$ miles), *Medium* ($40,001 - 60,000$ miles) and *High* ($> 60,000$ miles), denoted L, M and H, respectively. We combined brand and warranty into a single attribute to account for interaction between these attributes (for example, national brands were always offered only with high or medium warranty, while $H3$ tires were offered only with low warranty) to identify the following six brand-warranty combinations: NH, NM, $H1H$, $H2H$, $H2M$, $H3L$. Sixty four distinct tire sizes were offered, resulting in $64 * 6 = 384$ distinct possible tire SKUs that could be offered. This retailer carried 122 of these 384 possible SKUs in at least one of their stores. The number of SKUs offered across stores varied between 93 and 117, and averaged 105.2. The assortment offered also varied slightly across the stores, indicating some localization.

The set of attributes, Ω, is given by $\Omega = \{s, b\}$, where s refers to *Size* and b refers to *Brand-Warranty*. We were advised by the retailer that customers do not substitute across sizes, since tire size is uniquely determined by the car (for example, a 14" diameter tire cannot be used on a 15" wheel). Thus $R(j)$ is comprised of SKUs with the same size as j but different brand warranties. Table 6 shows for each brand-warranty, the rank ordered substitute brand-warranties and Table 7 depicts the qualitative likelihood of substitution across brand-warranty levels, where $\{S, L, M\}$ denote substitution levels *Somewhat Likely/Likely/Most Likely* with probabilities $\{\pi_S, \pi_L, \pi_M\}$. Both of these tables were constructed based on estimates provided by the Vice President of the tire category for the retailer.

We used sales data at the Store-SKU level for the six month period from July 2004 to December 2004 to fit the model and estimate its parameters. We need to estimate 64 size shares, 6 brand-warranty shares, and 3 substitution probabilities. The estimation procedure followed was equivalent to the procedure described for snack cakes if we view *Size* as being equivalent to *Flavor* (in that there is no

substitution across *Size* or *Flavor*) and *Brand-Warranty* being equivalent to *Brand-Size*. In particular, we can express log-likelihood as a function of the brand-warranty shares and substitution probabilities and hence need to estimate only 6 brand-warranty shares and 3 substitution probabilities for each store.

This process worked for 319 of the retailer's 574 stores, but at 255 stores there was insufficient data to determine all 6 brand-warranty shares. In particular, in these stores there was no size in which brand-warranties $H2M$ and $H3L$ were both offered, so it was not possible to identify the split of demand between $H2M$ and $H3L$. Hence, we had to modify our estimation approach for these stores. From the parameter estimates in the 319 stores with sufficient data to estimate all parameters, we observed that the share of $H3L$ at a store is correlated with median household income ($R^2 = 0.15, p < 0.10$). We regressed the $H3L$ share against median income for the stores at which we could estimate all the parameters and used the regression estimate for the share of $H3L$ at other stores. We then used MLE to estimate the remaining parameters. Of the 255 stores with insufficient data, there were 52 stores where we could neither identify the share of $H3L$ nor $H2M$. For these stores, in addition to estimating $H3L$ shares, we also estimated $H2M$ shares by regressing it against income. As before, we used MLE to estimate the remaining parameters. Figure 2 shows that the estimated share of $H3L$ at each store is negatively correlated with median income level in the zip code in which the store is located, which is to be expected. It also shows that the share of $H2H$ and $H2M$ are positively correlated with median income, which is also to be expected. In addition to supporting the parameter estimation process as described above, these results provide confirming demographic evidence to support the reasonableness of our parameter estimates.

Tables 8 and 9 show the average across all stores of the estimated brand warranty demand shares and substitution probabilities.

We used the hedonic regression described in Section 4.2 to assign price to SKUs not currently offered. The regression R^2 is 96.32% which supports our assumption that prices on existing SKUs are based on attribute values. As before, we normalized the estimated prices by a scaling factor so as to equate the estimated revenue of the current assortment to the actual revenue of this assortment.

Table 8 also gives sales share estimates which can be compared with the demand share estimates. The most interesting comparisons are the demand share estimate for $H3L$, which is much higher than the sales share, and for $H2M$, which is much lower than the sales share. The reason for this appears to be that the retailer offered $H3L$ in many fewer sizes than $H2M$; $H3L$ is offered in only 15 of the 64 sizes, versus 52 sizes in which $H2M$ is offered. But for those sizes where $H3L$ and $H2M$ are both offered, $H3L$ outsells $H2M$ by $40:1$ on an average, indicating that it is strongly preferred over $H2M$. The retailer offered $H3L$ in fewer sizes because they preferred to sell the higher priced $H2M$ and believed that their sales staff could convince customers to trade up to this tire. The substitution estimate of 45% shows that many customers did in fact trade up, and this explains the high sales share for $H2M$ relative to its demand share. However, the 55% of the 61% of customers preferring $H3L$ represented more than 34% of demand that was being lost due to the meager offering of $H3L$ in the current assortment, suggesting that there was substantial opportunity to increase sales by re-assorting.

The same analysis holds even if we compare the impact on revenues. The average price of $H3L$ and $H2M$ in the sizes where both were offered was \$28 and \$36 respectively. Suppose that there were 100 consumers shopping the store and consider the two alternatives of offering $H3L$ alone or $H2M$ alone. Offering $H2M$ would capture $(5\% * 100 + 45\% * 61\% * 100) * \$36 = \$1168$ in revenues while offering $H3L$ would capture $(61\% * 100) * \$28 = \$1708$ implying 46% additional revenue.

We computed the sales-weighted MAD of sales shares predicted based on the parameter estimates from the actual sales shares, and found it to be 13.6% at the store-SKU level and 4.5% at the chain-SKU level.

To validate our results, we used Store-SKU sales data for the next six month period from January 2005 to June 2005. We used the previously estimated parameters to compute Store-SKU sales shares and compared them with actual sales shares for the validation period (January 2005 to June 2005). The sales-weighted MAD of predicted sales shares from the actual sales shares was 38.2% at the Store-SKU level and 21.1% at the chain-SKU level. We note that our 21.1% chain level error is reasonable when compared with the 30.7% MAPE for chain sales of two new SKUs reported by Fader and Hardie (1996).

Figure 3 show the results of applying the greedy heuristic described in Section 5 to compute optimized assortments at the chain and store level, varying the maximum number of SKUs in the assortment. Figure 3 shows revenue as a function of SKU count for $L = 1$ and $L = m$ and Table 10 summarizes the store-optimal assortment for a representative store in terms of changes (additions and deletions) over the current assortment.

To quantify the potential improvement in revenue, we compare the assortments we generated for $L = 1$ and $L = m$ to the current assortment, which had a revenue of \$80.2 million. Because SKU count varied somewhat by store, in computing the revenue of the $L = 1$ assortment, as in the snack cakes example, we first generated the greedy solution for a SKU count equal to the maximum SKU count across all stores, and then for store s with SKU count K_s, we used the first K_s SKUs chosen by greedy. The revenue of this solution was \$104.1 million, a 30.1% increase over the current revenue. The revenue for store specific assortments with SKU counts of K_s for each store was \$108.7 million, a 35.9%[4] increase over the current revenue. We thus see that maximum localization adds 5.8% beyond the 30.1% achieved by chain level optimization.

In contrast to the snack cakes example, where a unique assortment per store was not feasible due to planogramming complexities, it is completely feasible here for the retailer to offer store-specific assortments, since the tires are not actually displayed at the store. Hence, we did not compute revenue lifts for values of L between 1 and m as we did with snack cakes.

While localizing assortments by store added only 5.8% to revenue in aggregate, there were meaningful differences between store assortments. Table 11 shows demand estimates for two different stores A and B, and also gives the median income for the zip code in which these stores are located. Store

[4]We also optimized the assortment to maximize gross margins using approximate gross margin data by brand. The increase in gross margins was 32%, still a significant number.

B customers with 50% higher income than Store A customers, buy more of the prestigious National brand and less of the low quality House 3 Brand. Table 12 shows the SKUs which are unique to each store in the store localized assortments. Here again Store B's specific SKUs are mostly the National brand and high mileage warranties, whereas Store A's specific SKUs are the lowest two brands and mostly Low or Medium mileage warranty.

We want to emphasize that these SKU differences are not a result of explicitly considering income. Rather, the causal link leading to higher priced tires being carried in the wealthier area store is high income leading to the purchase of higher quality tires leading to higher demand estimates for the better tires, leading to the revenue maximization algorithm to choosing more of these tires.

Since one of the key research questions addressed in our paper is the impact of localized assortments on the potential revenues, we sought to understand why localization had a significantly lower impact on tires than snack cakes (5.8% vs. 12% revenue impact). We first thought that the store to store variation in demand patters for snack cakes was greater than for tires, but found that store variation, as measured by the coefficient of variation of demand share of SKUs across stores, was nearly equal for the two cases, and in fact slightly higher for tires. We then realized that the percent of maximum demand captured by the single chain optimal assortment would also affect the benefits of localization. If a retailer carried such a broad assortment that the single chain assortment captured nearly all of the potential demand, then there would be very little improvement potential left for localization. The snack cakes assortment offers 40% of the potential SKUs, while the tire assortment offers 27% of the potential SKUs, so it might seem that the snack cakes assortment is broader. However, as shown by the Pareto curves in 4, (which show cumulative percent of maximum revenue captured by a given percent of SKUs in the chain optimal assortment, where the SKUs have been sorted in decreasing order of contribution to revenue), the 27% of tire SKUs offered capture 82% of maximum potential revenue, whereas the 40% of snack cake SKUs offered capture just 68% of maximum potential revenue. Thus for tires, only 18% of maximum revenue is left to be captured by localization, whereas 32% is left to be captured by localization in the snack cakes, nearly twice as much.

Thus the benefits of localization depend not only on the amount of demand variation across stores, but on the percentage of demand captured by the single chain optimal assortment. This is an important principle for retailers to consider in choosing which categories to localize. The natural inclination is for a retailer to localize their most important categories, but we have seen cases where the retailer carried such a broad assortment in their most important categories that there was little to be gained from localization.

We also performed a sensitivity analysis to determine how robust our results were to the assumptions made about substitution probabilities. We varied the estimated substitution probabilities by increasing/decreasing them by a factor of 2 and computed the revenues obtained from the chain-wide optimal assortment. Our analysis revealed that the optimal assortment and sales lift were sensitive to only one substitution parameter, the substitution probability from $H3L$ to $H2M$. This makes sense because the demand shares of other brand-warranties are low, and the fraction of customers who substitute if these low share options are not offered, has little impact on revenue. Table 13 shows how the lift

in revenues varied as a function of this average substitution probability. Note that for the base case, where the average substitution probability from $H3L$ to $H2M$ is 0.45, we obtain a 30.1% increase in revenues, whereas when this probability is 1, then the increase reduces to 9%. We can interpret this 9% as the revenue gains excluding the effect of adding $H3L$ SKUs.

The retailer decided to test a portion of our recommendations by adding eleven of the 47 SKUs we had recommended be added to the assortment, ten $H3L$ SKUs and one $H1H$. Given the lead time involved in procuring these new tires, the changes to the assortment were implemented only in July 2005.

The validation analysis that we conducted earlier used data for the period January 2005 to July 2005 during which none of the eleven new SKUs were included in the assortment. Hence, our first objective was to test the performance of our demand estimation procedure for forecasting sales of the eleven new tires introduced into the assortment in the period July 2005 to December 2005. To achieve this, we first recalibrated the model by using sales data from January 2005 to June 2005 to estimate demand and substitution parameters available immediately prior to the July to December 2005 implementation period. We then used the revised demand estimates to forecast sales shares, at each store, for the newly introduced SKUs, for the period July 2005 to December 2005. Table 14 shows a comparison of the predicted vs. actual chain sales shares for the new SKUs. The sales-weighted MAD across all SKUs is 17%, which compares favorably to the 30.7% MAPE for chain sales of two new SKUs reported by Fader and Hardie Fader and Hardie (1996).

Our second objective was to estimate the revenue lift that the retailer achieved by implementing a portion of our recommendations. Estimating the change in revenue from the current to the implemented assortment was complicated because, in addition to adding eleven new SKUs to the assortment, the retailer deleted more than eleven SKUs in each store. The number of SKUs deleted varied somewhat by store but averaged 24 SKUs deleted. To achieve a fair comparison, in a store where N_s SKUs had been deleted and 11 added, we used the greedy heuristic to choose $N_s - 11$ SKUs that were in the current assortment but not in the implemented assortment and added them to create a modified implementation assortment that had the same SKU count as the pre implementation assortment at each store and that was used in evaluation. We then used parameters estimated for the calibration, validation and implementation periods to estimate revenue for the current assortment, the modified implementation assortment and the two optimized assortments. Table 15 gives revenue estimates and percentage improvement over the current baseline assortment for these periods.

Note that in the validation period, we would have estimated a 13.1% revenue increase (from \$74.9 million to \$84.7 million) from the implemented assortment, whereas the actual increase was 5.8% (from \$72.3 million to \$76.5 million) due to change in parameters over time. We note that a 5.8% improvement is large relative to what retailers typically achieve through enhancements to existing stores. For example Canadian Tire reports achieving a $3-4\%$ annual revenue increase in existing stores during $2005-2009$ and is targeting the same increase through 2012 (Canadian Tire Corporation Limited 2007).

Similar to the analysis conducted for Snack Cakes, we find that the higher MAD for the validation

and implementation periods can be explained partly by sales trends that cause the parameter values to change. For example, the aging over time of the car models that use a particular tire impacts the demand for that tire. The demand for a tire type initially increases as cars that use that tire age and need replacement tires, but eventually declines as those cars becomes old enough that they begin to exit the population. Moreover, the retailer changed relative prices of the six brand-warranties from the calibration to the demand periods, which impacted the six demand shares. Table 16 shows how changes in relative prices across brand-warranties relates to systematic changes in their demand shares over time, which is clearly not captured in our current model. In particular, it is quite striking in this Table that as the price difference between $H2M$ and $H3L$ narrowed from 43.7% to 41.9% to 22.9% over the three periods, the demand split between these two brand-warranties shifted from 7%/70% to 29%/27%.

7 Conclusions

We have formulated a model of the problem of choosing a profit or revenue maximizing assortment, devised a demand model for estimating parameters and a heuristic for choosing assortments and applied all of this to two real problems – choosing a snack cake assortment for a regional convenience chain and a tire assortment for a national retailer. Several important conclusions follow from this research.

1. Optimizing a retailer's assortment can have an enormous impact on revenue. Revenue increases of 20% and 32% were indicated for the snack cakes and tires examples respectively, based on the validation sample. The tire retailer who implemented a small portion of our recommendations achieved a 6% revenue increase, which is significant given the 3-4% comparable store sales increases typical for this segment.

2. Product attributes in these applications were of three types: 1) technical specifications, such as tire size, on which there could be no substitution 2) price-value, such as varying brand qualities, where we would expect substitution between neighboring qualities and 3) taste attributes such as flavor.

3. Forecast accuracy was adequate to achieve real benefits, but improving forecast accuracy would be a useful focus of future research. Our validation sample forecast errors for the snack cakes example were 25.8% at the chain SKU level and 40.1% at the store SKU level. For the tire example they were 21.1% at the chain SKU level and 38.2% at the store-SKU level. The chain level errors seem reasonable in light of the 30.7% chain level error for two SKUs reported by Fader and Hardie (1996). None the less, the errors were great enough to reduce the revenue increase by about half from the fit to validation samples.

   Sources of error include randomness in sales, our assumption that attribute preferences are independent across attributes, drift in parameters over time and price changes. One implication

is that the assortment optimization process should be repeated frequently, perhaps twice per year, to adjust assortments as parameters change. Also note that chain level assortment optimization is based on the more accurate chain level forecasts, whereas localization by store requires store level forecasts and the higher store level errors is a barrier to assortment localization. In some applications it could happen that the benefits of localization are illusory because they are based on noise in the store sales data rather than real differences.

4. The assortment offered can distort sales away from true demand. In the tire application, the lowest price brand-warranty level had a demand share of 60% but a sales share of only 5%, because the retailer offered very few of these tires. Adding more of this brand warranty to the assortment was a big source of the revenue increase attained.

5. Substitution can be measured, can vary significantly and have a major impact on the optimal assortment. In our snack cakes example, in the family size, the probability of substituting from Brand 1 to Brand 2 was 89%, versus only a 22% probability of substituting from Brand 2 to Brand 1. This resulted in a complete replacement of Brand 1 by Brand 2 in family size of the optimized assortment.

6. We were able to use demographic data to confirm our parameter estimates in the tire example. In particular, the share of the lowest price tire and unwillingness to substitute up to a higher price tire were correlated with median income in the store area. Moreover, differences between store assortments were also correlated with income, with the higher income stores carrying more of the high priced tires. In some instances, we also used demographic data to assist in estimating parameters.

7. The benefit from localizing assortments by store varied considerably, from 6% to 12%. The difference did not stem from differences in demand variation across stores but in the shape of the revenue-SKU count response curve; the retailer with lower benefits from localization had a steeper revenue response curve and hence was capturing a large fraction of total demand with their core assortment. Given that the benefits of localization are varied and uncertain, the process presented here provides a useful test a retailer can apply to discern if localization of assortments is worth the substantial effort.

We hope this paper will provide a framework that will stimulate much needed additional research. We have four suggestions for future research.

1. An obvious enhancement of our approach would be to use multiple time periods of sales data so as to model trend in attribute values. For example, if the category was cameras and an attribute was analog or digital, then one might estimate trend in this attribute from sales history. However, even this enhancement may be insufficient to capture the attribute demand dynamics of highly volatile categories like fashion apparel.

2. There may be interaction between attribute values. In the case of tires, the demand for H3L will be higher for a size tire that goes on an older, inexpensive car than for a tire that goes on a new, luxury car. This could be incorporated into our approach by, for example, making the demand for H3L in a particular size depend in part on the average book value of the cars that size tire fits. However, very complicated and significant interaction between attributes would limit the applicability of our approach.

3. We did not have access to detailed sales and inventory data over time, but if we had, it could be valuable to fashion a more aggressive estimation process that took advantage of the varying assortments presented to a customer at a store each day because of stock outs.

4. Differences in price and quality is an important attribute, sometimes called 'good, better, best' by retailers. However, as the price changes over time in the tire example showed, 'good, better, best' is not a precisely defined attribute as the price differentials between these three levels may change over time or may be different for different size tires. Hence, it would be fruitful to incorporate relative price differences into the attribute definition.

Acknowledgments

We thank Professors Abba Krieger and David Bell from the Wharton School for numerous valuable comments and helpful suggestions with our research. We also thank Carol Jensen, Tamara H. Jermyn, and Amanda O'Brien from Wawa, Inc. for their support.

8 Tables

Table 1: Example to show that the Likelihood Function may not be Concave

	a	b	c	Total
A	100	1000	200	1300
B	300	150		450
C		100	400	500

Table 2: Snack Cakes: Demand Share Estimates (Averaged across Stores)

Brand Size	**Sales Share**	**Demand Share**
B_1S	67%	61%
B_2S	24%	27%
B_1F	4%	6%
B_2F	5%	6%

Table 3: Snack Cakes: Substitution Probability Estimates (Averaged across Stores)

	B_1S	B_2S	B_1F	B_2F
B_1S	1	18%	0	0
B_2S	26%	1	0	0
B_1F	0	0	1	89%
B_2F	0	0	22%	1

Table 4: Snack Cakes: Impact of Amount of Localization on Revenue

L	**Revenues ($ million)**
1	7.38
2	7.62
3	7.75
4	7.86
5	7.92
6	7.94
...	....
$m = 140$	8.11

Table 5: Snack Cakes: Assortment Changes for a Representative Store

	Single Serve		**Family Size**		**Demand**	**Brand**
Flavor	**Brand 1**	**Brand 2**	**Brand 1**	**Brand 2**	**Rank**	**Sizes**
Cinnamon	-	-	x	-	1	3
Chocolate	-	-	x	-	2	3
Peanut Butter	-	+	x	+	3	3
Butterscotch	-	+	x	+	4	3
Butter	-	-		-	5	3
Vanilla	-	+		-	6	3
Raspberry	+	-		-	7	3
Fudge	-	-		+	8	3
Honey	-	-		+	9	3
Buttercream	+	+	x	+	10	3
Choc. Chip	-		x	+	11	2
Cherry/Cheese	+	x		+	12	2
Cheese	-	-		x	13	2
Coconut	-				14	1
Oatmeal/Raisin	-				15	1
Jelly	-				16	1
Vanilla/Chocolate	+	x			17	1
Pineapple/Cheese		x			18	0
Blueberry		x			19	0
Marshmallow	x				20	0
Apple		x		x	21	0
Cream				x	22	0
Glazed					23	0
Retained	-	25				
Deleted	x	15				
Added	+	15				

Table 6: Tires: Rank-Order of Preferences $R(i)$

j_b	j_b^*
NH	$\{NM, H1H, H2H\}$
NM	$\{NH, H1H, H2H\}$
$H1H$	$\{H2H, H2M\}$
$H2H$	$\{H2M, H1H\}$
$H2M$	$\{H2H, H1H\}$
$H3L$	$\{H2M\}$

Table 7: Management's estimate of the most likely substitution probabilities

	NH	NM	$H1H$	$H2H$	$H2M$	$H3L$
NH	1	S	S	S	0	0
NM	L	1	S	S	0	0
$H1H$	0	0	1	L	S	0
$H2H$	0	0	S	1	S	0
$H2M$	0	0	S	L	1	0
$H3L$	0	0	0	0	M	0
S =*Somewhat Likely,* L =*Likely,* M = *Most Likely*						

Table 8: Tires: Demand Share Estimates (Averaged across Stores)

Brand Warranty	**Sales Share**	**Demand Share**
National High	1%	4%
National Medium	1%	3%
House 1 High	3%	4%
House 2 High	26%	24%
House 2 Medium	45%	5%
House 3 Low	24%	61%

Table 9: Tires: Substitution Probability Estimates (Averaged across Stores)

Substitution Likelihood	**Probability**
Somewhat Likely	2%
Likely	6%
Most Likely	45%

Table 10: Tires: Assortment Changes for a Representative Store

Brand Warranties	**SKUs Added**	**SKUs Deleted**
NH	11	1
NM		5
H1H	1	8
H2H	19	
H2M	2	32
H3L	14	1
Total	47	47

Table 11: Demand Shares and Median Household Income for Two Stores

Brand	**Warranty**	**Store** A	**Store** B
National	High	2%	6%
National	Medium	1%	4%
House 1	High	2%	4%
House 2	High	19%	29%
House 2	Medium	6%	6%
House 3	Low	71%	51%
HH Income		$\$47,368$	$\$77,194$

Table 12: Tires: Store Specific SKUs for Two Stores

Store A			Store B		
Size	**Brand**	**Warranty**	**Size**	**Brand**	**Warranty**
205/70R14	House 3	LOW	P155/80R13	House 2	HIGH
P185/65R15	House 2	MEDIUM	P155/80R13	House 3	LOW
P195/65R15	House 2	MEDIUM	P175/65R14	House 2	MEDIUM
P195/70R14	House 2	MEDIUM	P175/65R14	National	HIGH
P205/55R16	House 2	MEDIUM	P175/65R14	National	MEDIUM
P215/70R14	House 2	MEDIUM	P175/70R13	National	HIGH
P215/70R14	National	HIGH	P185/60R14	House 2	HIGH
P215/75R15	House 2	HIGH	P185/65R14	House 1	HIGH
P225/60R16	House 2	MEDIUM	P185/65R14	National	HIGH
P225/60R16	House 1	HIGH	P185/65R14	National	MEDIUM
P225/70R14	House 2	MEDIUM	P185/65R15	National	HIGH
P225/70R15	House 2	MEDIUM	P185/70R13	House 2	HIGH
P235/55R16	House 2	MEDIUM	P185/70R14	National	HIGH
P235/60R14	House 2	HIGH	P195/60R14	House 2	HIGH
P235/60R15	House 2	HIGH	P195/60R14	National	HIGH
P235/60R15	House 2	MEDIUM	P195/60R15	House 1	HIGH
P235/60R16	House 2	MEDIUM	P195/60R15	National	HIGH
P235/60R16	House 1	HIGH	P195/60R15	National	MEDIUM
P235/70R15	House 2	HIGH	P195/65R15	National	HIGH
P235/70R15	House 2	MEDIUM	P195/70R14	National	MEDIUM
P235/75R15	House 2	HIGH	P205/60R15	Futura	MEDIUM
P235/75R15	House 2	LOW	P205/60R15	National	HIGH
P245/60R15	House 2	HIGH	P205/60R15	National	MEDIUM
P255/60R15	House 2	HIGH	P205/65R15	House 1	HIGH
P255/70R15	House 2	HIGH	P205/65R15	National	MEDIUM
P255/70R15	House 2	MEDIUM	P205/70R15	National	MEDIUM
P255/70R15	National	HIGH	P215/60R16	National	HIGH
P275/60R15	House 2	MEDIUM	P215/65R15	National	HIGH
P275/60R15	National	HIGH	P215/70R15	National	MEDIUM
P295/50R15	House 2	HIGH	P235/60R16	National	MEDIUM

Table 13: Tires: Revenue Lift vs. α_M

α_M	**Revenues ($ million)**	**% Increase**
0.00	337.7	322.1
0.23	133.1	66.4
0.45	104.1	30.1
0.73	95.2	19.0
1.00	87.2	9.0

Table 14: Tires: Actual vs. Predicted Chain Level Sales of New SKUs

			% Share of Total Sales		
Size	**Brand**	**Warranty**	**Actual**	**Predicted**	**% Error**
P225/60R16	House 3	LOW	3.0	3.4	10.9
P215/70R15	House 3	LOW	3.1	3.5	12.6
P205/65R15	House 3	LOW	3.2	4.4	37.1
P205/70R15	House 3	LOW	2.8	2.7	5.1
P195/65R15	House 3	LOW	2.2	2.4	7.6
P215/65R15	House 3	LOW	1.0	1.2	25.3
P205/55R16	House 3	LOW	1.2	1.1	6.2
P215/60R16	House 3	LOW	1.1	1.3	19.4
P215/70R14	House 3	LOW	0.9	1.2	37.2
195/70R14	House 2	High	0.8	1.0	29.5

The sales-weighted MAD across all new SKUs is 17%

Table 15: Tires: Revenue Estimates for Current, Implemented and Optimized Assortments (percentage improvement over current revenues given in parenthesis)

Revenues ($$ million)	**Parameter Estimates Used**		
Assortment	**Jul 04 - Dec 04**	**Jan 05 - Jun 05**	**Jul 05 - Dec 05**
Current Assortment (Jul 2004 - Dec 2004)	80.2	74.9	72.3
Modified Implementation (Jul 2005 - Dec 2005)	90.7 (13.1%)	84.7 (13.1%)	76.5 (5.8%)
Recommended Optimal Assortment, Chain	104.1 (29.8%)	94.3 (25.9%)	81.4 (12.6%)
Recommended Optimal Assortment, Store	108.2 (34.9%)	99.2 (32.4%)	83.6 (15.6%)

Table 16: Tires: Price Changes and Impact on Demand Shares for a Representative Store

	Share of Demand (Price in $$)		
Brand Warranty	**Jul - Dec 20004**	**Jan - Jun 2005**	**Jul - Dec 2005**
NH	2 (73.9)	5 (69.7)	8 (77.6)
NM	2 (58.9)	3 (51.4)	6 (50.9)
H1H	3 (59.8)	10 (61.1)	8 (58.6)
H2H	16 (49.6)	21 (53.5)	22 (56.5)
H2M	7 (43.3)	10 (45.7)	29 (46.5)
H3L	70 (30.1)	51 (32.2)	27 (37.9)
H2M-H3L % Price Difference	43.7	41.9	22.9

9 Figures

Figure 1: Snack Cakes: Impact of Chain-wide Assortment and Fully Localized Assortments

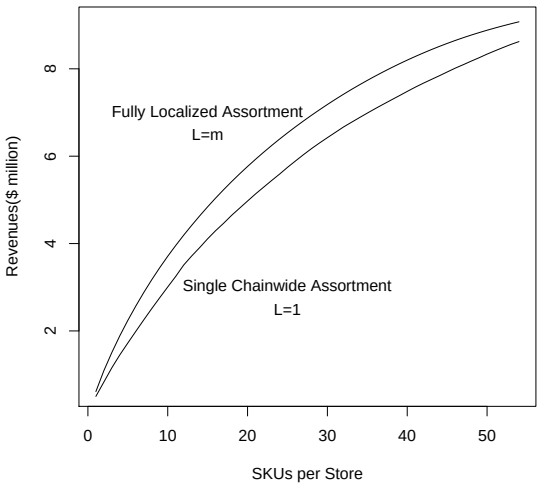


Figure 2: Share of $H3L$ ($H2H, H2M$) is Negatively (Positively) Correlated with Income

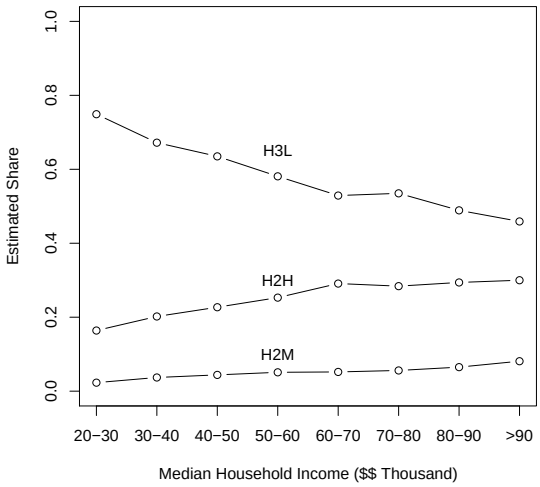


Figure 3: Tires: Impact of Chain-wide and Fully Localized Assortments

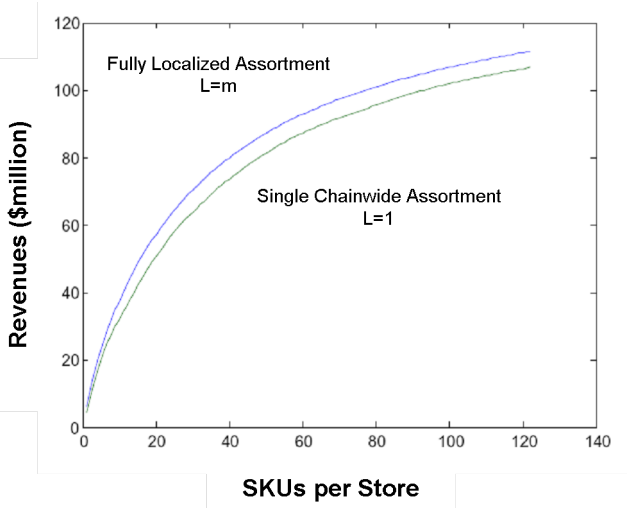
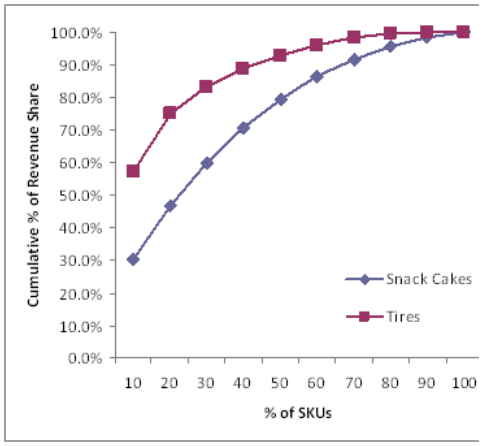


Figure 4: Cumulative Revenue Share vs. No. of SKUs



References

Achabal, D.D., W.L. Gorr, V. Mahajan. 1982. Multiloc: A multiple store location decision model. *Journal of Retailing* **58**(2) 5–25.

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

Anupindi, R., M. Dada, S. Gupta. 1998. Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Science* **17**(4) 406–423.

Batsell, Richard R, Leonard M Lodish. 1981. A model and measurement methodology for predicting individual consumer choice. *Journal of Marketing Research* **18**(1) 1–12.

Bell, D.R., A. Bonfrer, P.K. Chintagunta. 2005. Recovering SKU-level preferences and response sensitivities from market share models estimated on item aggregates. *Journal of Marketing Research* **42**(2) 169–182.

Belloni, A., R. Freund, M. Selove, D. Simester. 2008. Optimizing Product Line Designs: Efficient Methods and Comparisons. *Management Science* **54**(9) 1544.

Ben-Akiva, M., S.R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.

Bucklin, R.E., S. Gupta. 1992. Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *Journal of Marketing Research* **29**(2) 201–215.

Cachon, G., C. Terwiesch, Y. Xu. 2005. Retail assortment planning in the presence of consumer search. *Manufacturing & Service Operations Management* **7**(4) 330.

Canadian Tire Corporation Limited. 2007. Annual report.

Caro, F., J. Gallien. 2007. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science* **53**(2) 276–276.

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

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

Chong, J. K., T. H. Ho, C. S. Tang. 2001. A modeling framework for category assortment planning. *Manufacturing & Service Operations Management* **3**(3) 191–210.

Dempster, AP, NM Laird, DB Rubin. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* **39**(1) 1–38.

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

Fader, Peter S., Bruce G. S. Hardie. 1996. Modeling consumer choice among skus. *Journal of Marketing Research* **33**(4) 442–452.

Gaur, V., D. Honhon. 2005. Assortment planning and inventory decisions under a locational choice model. *Management Science* **52**(10) 1528–1543.

Goodman, A. C. 1998. Andrew court and the invention of hedonic price analysis. *Journal of Urban Economics* **44**(2) 291–298.

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

Green, P.E., A.M. Krieger. 1987a. A consumer-based approach to designing product line extensions. *Journal of Product Innovation Management* **4**(1) 21–32.

Green, P.E., A.M. Krieger. 1987b. A simple heuristic for selecting 'good' products in conjoint analysis. *Applications of Management Science* **5** 131–153.

Green, P.E., A.M. Krieger. 1992. An application of a product positioning model to pharmaceutical products. *Marketing Science* **11**(2) 117–132.

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

Hotelling, H. 1929. Stability in competitition. *Economic Journal* **39** 41–57.

Kohli, R., R. Sukumar. 1990. Heuristics for Product-Line Design Using Conjoint Analysis. *Management Science* **36**(12) 1464–1478.

Kök, G., M. L. Fisher. 2007. Demand Estimation and Assortment Optimization under Substitution: Methodology and Application. *Operations Research* **55**(6) 1001–1021.

Kök, G., M. L. Fisher, R. Vaidyanathan. 2008. Assortment planning: Review of literature and industrial practice. Invited chapter in Retail Supply Chain Management, Eds. N. Agrawal, S.A. Smith. Kluwer Publishers.

Lancaster, K. 1975. Socially Optimal Product Differentiation. *The American Economic Review* **65**(4) 567–585.

Lancaster, K. J. 1966. A new approach to consumer behavior theory. *Journal of Political Economy* **74** 132–157.

Maddah, B., EK Bish. 2007. Joint pricing, assortment, and inventory decisions for a retailer's product line. *Naval Research Logistics* **54**(3) 315.

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

Manne, A.S. 1964. Plant location under economies-of-scale-decentralization and computation. *Management Science* **11**(2) 213–235.

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

McGregor, Jena. 2008. At Best Buy, marketing goes micro. *BusinessWeek* .

Netessine, S., N. Rudi. 2003. Centralized and competitive inventory models with demand substitution. *Operations Research* 329–335.

O'Connell, Vanessa. 2008. Reversing field, Macy's goes local. *Wall Street Journal* .

Pakes, Ariel. 2003. A reconsideration of hedonic price indexes with an application to PC's. *The American Economic Review* **93**(5) 1578–1596.

Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy* **82**(1) 34–55.

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

Talluri, K., G. van Ryzin. 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science* **50**(1) 15–33.

Teitz, M.B., P. Bart. 1968. Heuristic methods for estimating the generalized vertex median of a weighted graph. *Operations Research* **16**(5) 955–961.

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

Vulcano, G., G. van Ryzin, R. Ratliff. 2009. Estimating primary demand for substitutable products from sales transaction data. Working Paper, NYU Stern.

Zimmerman, A. September 7, 2006. To boost sales, Wal-Mart drops one-size-fits-all approach. *Wall Sreet Journal Online* A1.

Zimmerman, Ann. October 7, 2008. Home Depot learns to go local. *Wall Street Journal* .