University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

Fall 12-21-2011

Retail Demand Management: Forecasting, Assortment Planning and Pricing

Ramnath Vaidyanathan
University of Pennsylvania, ramnath.vaidya@gmail.com

Follow this and additional works at: http://repository.upenn.edu/edissertations

Part of the Management Sciences and Quantitative Methods Commons

Retail Demand Management: Forecasting, Assortment Planning and Pricing

Abstract

In the first part of the dissertation, we focus on the retailer's problem of forecasting demand for products in a category (including those that they have never carried before), optimizing the selected assortment, and customizing the assortment by store to maximize chain-wide revenues or profits. We develop algorithms for demand forecasting and assortment optimization, and demonstrate their use in practical applications. In the second part, we study the sensitivity of the optimal assortment to the underlying assumptions made about demand, substitution and inventory. In particular, we explore the impact of choice model mis-specification and ignoring stock-outs on the optimal profits. We develop bounds on the optimality gap in terms of demand variability, in-stock rate and consumer heterogeneity. Understanding this sensitivity is key to developing more robust approaches to assortment optimization. In the third and final part of the dissertation, we study how the seat value perceived by consumers attending an event in a stadium, depends on the location of their seat relative to the field. We develop a measure of seat value, called the Seat Value Index (SVI), and relate it to seat location and consumer characteristics. We apply our methodology to a proprietary dataset collected by a professional baseball franchise in Japan. Based on the observed heterogeneity in SVI, we provide segment-specific pricing recommendations to achieve a service level objective.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Managerial Science and Applied Economics

First Advisor
Marshall L. Fisher

Keywords
assortment planning, retailing, empirical research, pricing, seat value, demand estimation

Subject Categories
Management Sciences and Quantitative Methods

RETAIL DEMAND MANAGEMENT: FORECASTING, ASSORTMENT PLANNING AND PRICING

Ramnath Vaidyanathan

A DISSERTATION

in

Operations and Information Management

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2011

Supervisor of Dissertation

Marshall L. Fisher
UPS Professor of Operations and Information Management

Graduate Group Chairperson

Eric T. Bradlow
Vice Dean and Director, Wharton Doctoral Programs

Dissertation Committee
Abba Krieger, Professor of Statistics and Operations Research
Serguei Netessine, Professor of Operations Management
Senthil Veeraraghavan, Associate Professor of Operations Management

ACKNOWLEDGEMENTS

ABSTRACT

RETAIL DEMAND MANAGEMENT: FORECASTING, ASSORTMENT PLANNING AND PRICING

Ramnath Vaidyanathan

Marshall Fisher

In the first part of the dissertation, we focus on the retailer's problem of forecasting demand for products in a category (including those that they have never carried before), optimizing the selected assortment, and customizing the assortment by store to maximize chain-wide revenues or profits. We develop algorithms for demand forecasting and assortment optimization, and demonstrate their use in practical applications.

In the second part, we study the sensitivity of the optimal assortment to the underlying assumptions made about demand, substitution and inventory. In particular, we explore the impact of choice model mis-specification and ignoring stock-outs on the optimal profits. We develop bounds on the optimality gap in terms of demand variability, in-stock rate and consumer heterogeneity. Understanding this sensitivity is key to developing more robust approaches to assortment optimization.

In the third and final part of the dissertation, we study how the seat value perceived by consumers attending an event in a stadium, depends on the location of their seat relative to the field. We develop a measure of seat value, called the Seat Value Index (SVI), and relate it to seat location and consumer characteristics. We apply our

methodology to a proprietary dataset collected by a professional baseball franchise in Japan. Based on the observed heterogeneity in SVI, we provide segment-specific pricing recommendations to achieve a service level objective.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation is focused on retail demand management. The dissertation has three parts. In the first part, comprising of Chapter 3, we consider the retailers 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 revenue or profit. Assortments can vary by store, subject to a maximum number of different assortments. We introduce a model of substitution behavior, in case a customer's first choice is unavailable and consider the impact of substitution in choosing assortments for the retail chain. We view a SKU as a set of attribute levels, apply maximum likelihood estimation to sales history of the SKUs currently carried by the retailer to estimate the demand for attribute levels and substitution probabilities, and from this, the demand for any potential SKU, including those not currently carried by the retailer. We specify several alternative heuristics for choosing SKUs. We apply this approach to optimize assortments for three real examples: snack cakes, tires and automotive appearance chemicals. A portion of our recommendations for tires and appearance chemicals were implemented and produced sales increases of 5.8% and 3.6% re-

spectively, which are significant improvements relative to typical retailer annual comparable store revenue increases. We also forecast sales shares of 1, 11 and 25 new SKUs, for the snack cakes, tires and automotive appearance chemicals applications, respectively with MAPEs of 16.2%, 19.1% and 28.7%.

In the second part, comprised of Chapter 4, explores the sensitivity of the assortment to the key assumptions made in practice. Most assortment planning papers assume that the retailer knows (a) the customer arrival process, (b) the customer choice process, and (c) the demand parameters. In addition, the impact of stock-out substitution is ignored to keep the models analytically tractable. Clearly, the optimal assortment and the optimal profits are sensitive to these underlying assumptions about demand, substitution and stock-outs. We particularly explore the impact of choice model mis-specification and ignoring stock-out substitution, on the optimal assortment and profits. We develop bounds on the optimality gap of profits in terms of demand variability, in-stock rate and assortment size.

In the third and final part, comprised of Chapter 5, we study how the seat value perceived by consumers attending an event in a theater/stadium, depends on the location of their seat relative to the stage/field. We develop a measure of seat value, called the Seat Value Index, and relate it to seat location and consumer characteristics. We implement our analysis on a proprietary dataset that a professional baseball franchise in Japan collected from its customers, and provide recommendations. We find that customers seated in symmetric seats on left and right fields might derive very different valuations from the seats. We also find that the more frequent visitors to the stadium report extreme seat value less often when compared to first-time visitors. Our findings and insights remain robust to the effects of price and game related factors. Our research quantifies the significant influence of seat location on

the ex-post seat value perceived by customers. Utilizing the heterogeneity in seat values at different seat locations, we provide segment-specific pricing recommendations based on a service-level objective that would limit the fraction of customers experiencing low seat value to a desired threshold.

Chapter 2

Literature Review

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

2.1   Consumer Choice Modeling

Consumer choice models constitute the fundamental platform for assortment planning, and may be classified as (1) utility based models, and (2) exogenous demand models.

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 S, every customer chooses the option giving him the highest utility in $S \cup \{0\}$. The market share for each SKU $i \in S$ 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 $i \in S$ can be written as

$$q_i(S) = \frac{\exp\left(\frac{u_i}{\mu}\right)}{\sum_{i \in S} \exp\left(\frac{u_i}{\mu}\right)} \, .$$

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 levels $z_i = (z_i^1, z_i^2, ..., z_i^T) \in \mathbb{R}^T$. Each consumer has an ideal point $y \in \mathbb{R}^T$ that defines his most preferred attribute levels. 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 substitutes the variant $j \in A$ that is located closest to his ideal point in the attribute space if $U_j^y > 0$, or declines to purchase if $U_j^y \leq 0$.

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 $i \in S$; if $i \notin S$ they substitute to SKU $j \in S$ with probability α_{ij}. Under these assumptions, the market share of each SKU $i \in S$ is given by $q_i(S) = f_i + \sum_{j \notin S} f_j \alpha_{ji}$. The exogenous demand model has more degrees of freedom than the MNL and LC models and can accommodate extremely flexible substitution structures.

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.

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.

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.

Estimation is parsimonious as they only need to estimate attribute-level utilities, and their model 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 to investigate buying decisions of households. Bell et al. (2005) describe a method for obtaining (SKU)-level preferences from estimated attribute level parameters, circumventing the need for direct estimation of the more complex SKU-level 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 2.3.2.

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 retail assortments.

2.3.1 Stylized Models

The stylized model research began with a pioneering paper by van Ryzin and Mahajan (1999). They 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. Cachon et al. (2005), Caro and Gallien (2007), and Maddah and Bish (2007) extend the van Ryzin Mahajan model in various ways.

Gaur and Honhon (2006) show that for a locational choice model, the products in the optimal assortment are located far 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).

Although different papers focusing on assortment planning make different sets of assumptions, there is a common underlying structure across all of them. All stylized models assume that customers arrive according to a stochastic process, looking to purchase an item from the set $N \cup \{0\}$, where $N = \{1, 2, \ldots, n\}$ is the set of all potential products that can be offered in the category and $\{0\}$ represents the no-purchase option. On arriving at the store they observe the assortment $S \subseteq N$ carried by the retailer and choose to purchase item $i \in S$ with probability q_i. [1] The retailer's objective is then to select the optimal assortment $S \subseteq N$ and set inventory levels that maximize his expected profits.

This set up suggests that we can decompose an assortment planning model into four separate modules: (1) customer arrival process, (2) customer choice process, (3) item level demand distribution, and (4) joint optimization of assortment and inventory levels.

Customer Arrival Process

Customer arrival processes in a stochastic setting can be modeled using different probability distributions. A common assumption made to ensure model tractability is that the arrival process is independent of the assortment carried by the retailer.

---

[1]The choice probabilities are governed by the customer choice model and assumptions made about how customers react to stock-outs.

Let D denote the total number of customer arrivals in a cycle, and $\psi(d) = \mathbf{P}\{D = d\}$ be its probability mass function (or probability density function if we assume a continuous distribution). For example, Smith and Agrawal (2000) model arrivals as a negative binomial distribution, while van Ryzin and Mahajan (1999) and Gaur and Honhon (2006) assume that arrivals follow a normal distribution with mean λ and standard deviation $\sigma \sqrt{\lambda}$.

Customer Choice Process

As pointed out earlier, consumer choice models constitute the fundamental platform on which assortment planning models are built. A detailed discussion of the different models used in practice is presented in Section 2.1. The central idea running across all choice models is that customers purchase their favorite product if carried in the assortment and substitute to a different product, when it is not available. The difference arises from the assumption made about substitution behavior.

The end-goal of modeling customer choice is to derive an expression for the choice probabilities q_i. Note that the assortment on offer changes dynamically as the products carried in the assortment may stock out with time. This leads to a complex set of time-varying choice probabilities which makes the models highly intractable. A simplifying assumption commonly used is that consumers substitute only when their favorite product is not carried in the assortment, and not when their favorite product is in the assortment carried, but stocked-out at the time of her visit. The reader is directed to Kök et al. (2008) for more details. This assumption makes the choice probabilities q_i independent of the inventory decision.

Item Level Demand Distribution

The item level demand distribution can now be derived by observing that an arriving customer chooses to purchase item i with probability $q_i(S)$ and not purchase it with probability $1 - q_i(S)$. Hence, the conditional distribution of demand for item $i \in S$ can be written as $\mathbf{P}\left(D_i = d_i \mid D = d\right) = \binom{d}{d_i}$

The probability distribution for demand for item $i \in S$ can then be obtained by computing $\psi_i(d_i \mid q_i) = \sum_{d=d_i}^{\infty} \binom{d}{d_i} q_i^{d_i}(1 - q_i)^{d-d_i}$. The mean and standard deviation of the item level demand can be expressed in terms of the mean (λ) and standard deviation (σ) of the total demand, and written as $\mu_i = \lambda q_i$ and $\sigma_i = \sqrt{\lambda q_i} \times \sqrt{1 + h_i \left(\frac{\sigma^2}{\lambda} - 1\right)}$

For certain choices of the customer arrival distribution $\psi(d)$, $\psi_i(d_i \mid q_i)$ takes a simple form. For example, if $\psi(d)$ is Poisson with rate λ, then $\psi_i(d_i \mid q_i)$ is Poisson with rate λq_i. Similarly, when $\psi(d)$ is a Negative Binomial Distribution (NBD), the item level demands follow a thinned NBD.

In order to obtain closed form expressions for the expected profits, it is common to approximate the item level demand using a normal distribution. The approximation is exact when customer arrivals are Poisson.

Joint Optimization of Assortment and Stock Levels

Assortment optimization includes both the selection of items to stock (assortment selection) and deciding the inventory levels for each item in the assortment (inventory optimization). As shown by van Ryzin and Mahajan (1999), if the item level demands can be approximated by a normal distribution, the assortment optimization problem is separable across assortment selection and inventory optimization.

11

If S denotes the assortment offered, then the optimal inventories for all items $i \in S$ are given by the newsvendor model, and can be written as $x_i^* = \lambda q_i(S) + z_i \sqrt{\lambda q_i}$, where $z_i = \Phi^{-1}\left(1 - \dfrac{c}{p}\right)$, and Φ is the cumulative distribution function of the standard normal distribution. The expected profits from assortment S can be expressed as $\mathbf{E}\left[\Pi^A\left(\mathbf{x}^*(S)\right)\right] = (p - c)\lambda \sum_{i \in S} q_i(S) - p\phi(z) \sum_{i \in S} \sqrt{\lambda q_i(S)}$. The assortment optimization problem now reduces to selecting the assortment S that maximizes the total expected profits.

The structure of the optimal assortment S is known only for two cases. For the MNL model, van Ryzin and Mahajan (1999) show that the optimal assortment is of the form $S = \{1, 2, \ldots, k^*\}$ for some value of $k^* \leq n$, where the items are arranged in decreasing order of their popularities. For the LC model, Gaur and Honhon (2006) show that the optimal assortment S consists of a number of non-overlapping variants (across which consumers don't substitute), which can be determined by a simple line search for a single parameter.

2.3.2 Decision Support Models

Decision support research began with Green and Krieger (1985), who formulate a product line design 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, 1992), McBride and Zufryden 1988, Dobson and Kalish (1988) and 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 and an integer programming formulation of assortment planning. 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.

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 grocery shopping trips to estimate the parameters of the model and use a local improvement heuristic to suggest an alternative assortment with higher revenue. 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.

13

They note that the constrained shelf space implies that average inventory per SKU is inversely related to the breadth of the assortment and assess 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, as in the MNL model. They provide a process for estimating demand and substitution rates and apply their method to data from 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 multi-armed 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.

Chapter 3

Demand Estimation and Assortment Optimization[1]

3.1 Introduction

A retailer's assortment is the set of products they carry in each store at each point in time. Retailers periodically review and revise the assortment for each category of products they carry to take account of changes in customer demand over time as well as new products introduced by suppliers. This periodic assortment reset seeks to choose a set of SKUs to carry in the new assortment to maximize revenue or profit over a future planning horizon, subject to a shelf space constraint, which can often be expressed as an upper bound on the number of SKUs carried.

If a customer's most preferred product is not in a retailers assortment, they may elect to buy nothing or to purchase another product sufficiently similar to their most

---

[1]This chapter is based on Fisher and Vaidyanathan 2011. An Algorithm and Demand Estimation Procedure for Retail Assortment Optimization, Working Paper

preferred product that they are willing to buy it. This possibility of substitution must be taken into account in both assortment optimization and in estimation. In estimation, substitution probabilities need to be estimated and the sales of a SKU to customers who most preferred that SKU must be distinguished from sales to customers who preferred a different SKU but substituted when they didn't find their preferred SKU in the assortment.

Most retailers use the same assortment for all stores, except that in smaller stores they might eliminate some SKUs. Recently however, localizing assortments by store or store cluster has become a high priority for many retailers. For example, Zimmerman (September 7, 2006), O'Connell (April 21, 2008), McGregor (May 15, 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. In the extreme, a retailer might carry a unique assortment in each store, but most retailers claim that this is administratively too complicated. For example, retailers develop a diagram called a planogram showing how all products should 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. Despite the flurry of interest in localization 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.

The assortment planning process varies greatly across different retailers and product categories. Retail product categories are commonly segmented into into apparel, grocery, and everything else, usually called hard goods.[2] An analytic approach

---

[2]This discussion is based on Fisher and Raman (2010) and conversations with several retail

16

to apparel assortments is challenging because rapidly changing tastes make sales history of limited value. Assortment planning is most developed in the grocery segment (where it is usually called category management), due in part to Nielsen/IRI who enlists households to record over time their grocery purchases in all stores. Much academic research on grocery consumer behavior has relied heavily on national panel data. Among other things, the data allows one to model substitution behavior by observing what a customer buys, if anything, when a product they purchase every week is unavailable due to a stockout. Many grocery retailers are now engaged in SKU rationalization efforts aimed at reducing SKU count with minimal impact on revenue.

The approach described in this chapter best fits hard goods, where many retailers conduct an annual review of their various categories aimed at identifying SKUs to delete and add to the assortment in each category. Deletion decisions are easier, since sales data is available to indicate the popularity of existing SKUs, but current industry practice (for example, the household purchase data available in grocery is not available in hard goods) provides little if any hard data with which to forecast the sales of SKUs that might be added to the assortment, and hence a category manager is forced to rely on intuition and the representations of suppliers as to the merits of new products they are introducing. This also makes it impossible to forecast the revenue impact of strategic changes such as assortment localization.

It is 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 intelli-

gently localize assortments by store, would be valuable enhancements to current assortment planning practice, and the goal of this chapter is to provide those enhancements.

Our approach follows the marketing literature in viewing a SKU as defined by a set of attribute levels and assuming that a given customer has a preferred set of attribute levels. 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 levels of that SKU. We assume that if a customer does not find their ideal product in the assortment, they buy the product in the assortment closest to their ideal with some probability and we also estimate these substitution probabilities. We apply various heuristics to these demand and substitution estimates to determine optimized assortments. Our process can control the degree of localization by limiting the number of different assortments to be any level between a single assortment for the chain to a unique assortment for each store.

We applied this approach to the snack cakes category at a regional convenience store chain, the tire assortment at a national tire retailer and the appearance chemicals category of a major auto aftermarket parts retailer. The tire and auto parts retailers implemented portions of our recommended assortment changes and obtained revenue increases of 5.8% and 3.6% respectively, significant improvements given traditional comparable store annual increases in these segments.

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 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 chapter extends these papers, and hence the exisiting literature, in four ways.

1. We provide a thorough treatment of the important emerging topic of assortment localization. We allow a constraint on the number of different assortments so as to bound the administrative cost of localization and measure how the amount of localization impacts revenue. We compare and explain differing levels of localization benefits across our three applications. Chong et al. (2001) don't deal with localization. Kök and Fisher (2007) allow a unique assortment for each store but don't provide a constraint on the number of assortments or assess how localization impacts revenue.

2. 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 non-grocery 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.

3. 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 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 taste 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. As another example, in our snack cakes 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.

The multinomial logit approach better 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 (2006) provided a locational choice complement to the multinomial logit based analysis of van Ryzin and Mahajan (1999).

4. Two of the retailers in our study 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. We believe this is the first implementation validation of an analytic approach to assortment planning.

Section 3.2 provides our demand model and a formulation of the assortment optimization problem. Section 3.3 describes estimation of the demand model and heuristics for assortment selection. Section 3.4 presents results for the three applications. Section 3.5 analyzes the application results to understand differences in localization benefit and to assess the performance of the heuristics. Section 3.6 offers some concluding remarks.

3.2 Problem Formulation and Demand Model

We seek optimal assortments for a retail category over a specified future planning horizon. For concreteness, we assume that the goal is to maximize revenue, since this was the primary concern in our three applications, although it is straightforward to adapt our approach to maximizing other functions, such as unit sales or dollar gross margin. We define the problem parameters in Table 3.1.

The parameters K and L would be specified by the retailer. 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 to tradeoff the greater revenue that comes 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.

Table 3.1   Definition of Problem Parameters

Notation	Description
$N = \{1, 2, ..., n\}$	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 different assortments
D_i^s	Number of customers at store s for whom SKU i is their most preferred product
p_i	Price of SKU $i \in N$

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

Table 3.2   Description of Mathematical Notation

Notation	Description
A	Number of attributes
a	Attribute type index
N_a	Number of levels of attribute a, $a = 1, 2, \ldots, A$
f_{au}^s	Fraction of customers at store s who prefer level u of the attribute a, $u = 1, 2, \ldots, N_a$, $a = 1, 2, \ldots, A$
π_{auv}^s	Probability that a customer at store s who's first choice on attribute a is u is willing to substitute to v, defined for all $a, u,$ and v.

Our demand model assumes that a consumer shopping this category in store s has a most preferred SKU $i \in N$, but might be willing to substitute to other SKUs if

$i \notin S_{a(s)}$. We view a SKU as a collection of attribute levels, use historical sales data to estimate the demand share of each attribute level, and finally estimate the demand share of any SKU as the product of the demand shares of its attribute levels.

We introduce some notation to formalize this approach in Table 3.2. By definition, $\pi_{avv}^s = 1$. Moreover, π_{auv}^s can be 0 if attribute level v is not a feasible substitute for u. For example, size is an attribute of a tire and a 14 inch tire is not a feasible substitute for a customer with a 15 inch wheel.

The fraction of customers who most prefer SKU i with attribute levels $i_1, i_2, \ldots, i_A$ is defined to be $f_i^s = \Pi_{a=1}^{a=A} f_{ai_a}^s$. If a customer's most preferred SKU i with attributes $i_1, i_2, \ldots i_A$ is not in the assortment, they are willing to substitute to SKU j with attributes $j_1, j_2, \ldots j_A$ with probability $\pi_{ij}^s = \Pi_{a=1}^{a=A} \pi_{ai_a j_a}^s$. If a customer with most preferred SKU i finds $i \in S_{a(s)}$ when they shop the store, then we assume they buy it. Otherwise, they buy the best substitute for i in S, defined to be $j(i, S) = \arg \max_{i \in S} \Pi_{a=1}^{A} \pi_{ai_a j_a}^s$.

In using store sales data to estimate the parameters of our model, we first estimate D^s, the total unit demand in store s, as total unit sales divided by the share of demand captured, as defined by demand shares and substitution probabilities, and then set $D_i^s = f_i^s D^s$.

The revenue earned by store s using assortment $S_{a(s)}$ can then be written as

$$R_s(S_{a(s)}) = \left(\sum_{i \in S_{a(s)}} p_i D_i^s + \sum_{i \notin S_{a(s)}} D_i^s \pi_{ij(i, S_{a(s)})}^s p_{j(i, S_{a(s)})} \right)$$

The first term in this expression 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.

The assortment optimization problem is to choose S_l, $| S_l | \leq K$, $l = 1, 2, \ldots, L$ and $a(s)$ for all $s \in M$ to maximize $\sum_{s \in M} R_s(S_{a(s)})$.

Our demand model most closely resembles the locational choice model, but with three important differences: (1) in the locational choice model, the probability of purchasing the closest substitute from the assortment is either 0 or 1 (depending on the no-purchase utility), while we allow general substitution probabilities, (2) consumers are assumed to be distributed in a continuous space in the locational choice model, while we allow customer locations to be restricted to discrete locations in the attribute space, and (3) all attribute levels in the locational choice 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 substitution probabilities, which can be thought of as inducing a distance metric for attributes that can't be located in a space. These three enhancements were needed to make the locational choice model operational in the applications we consider.

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 3.4 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.

Our demand model implies that a consumer's preferences for the various attributes are independent, which may not be true. For example a college student shopping for twin size bed-sheets might have a different color preference than a suburban homemaker shopping for queen size sheets, so the color and size attributes for sheets would interact.

Our defense of this assumption is three-fold: (1) all prior publications we are aware of that use attributes in demand estimation make a similar assumption. Fader and Hardie (1996) is typical of the approach followed in the literature. They assume that the utility for a SKU is a linear function of its attributes and then use this utility in a Multinomial Logit model to determine SKU demand shares. They state that "*In both the marketing and economics literature, it is common to assume an additive utility function*" and note that this implies no interaction between attributes. (2) in our applications, we check the accuracy of this approximation by comparing demand estimates with actual sales for the SKUs currently carried and find that forecasts based on this model are accurate compared to previously published research. (3) if there is significant interaction between attributes, we demonstrate ways to modify our demand model to take this into account. In the snack cakes application (Section 3.4.1), the attributes are flavor, package size (single serve or family size) and brand. Package size and brand interact since one brand is stronger in single serve and another in family size. We deal with this by combining brand and size into a new attribute brand-size. In the tire application (Section 3.4.2), the attributes are size, brand (4 brands) and mileage warranty (low, medium, high). Brand and mileage warranty interact because a given brand does not offer all warranty levels, and so we combine brand and warranty level to create the attribute brand-warranty.

In the tires example, there is also an interaction between size and brand-warranty. A tire with a given size attribute level fits a defined set of car models of a certain age and value. The six brand-warranty levels correspond to different price points and quality. There is a clear interaction between brand-warranty shares across price points and the age and value of the cars a size tire fits. We show how to deal with this by partitioning the sizes into a finite number of homogenous segments (which

are latent) and allowing the brand-warranty shares to be conditional on segment membership.

3.3 Analysis

We describe our methods for estimating model parameters and choosing assortments.

3.3.1 Estimating Demand and Substitution Probabilities

We use Maximum Likelihood Estimation to estimate demand and substitution probabilities. Our primary input for estimation is store-SKU sales of products currently carried by the retailer during a prior history period. Parameters are estimated at the store level, but for expositional simplicity we will drop the store superscript in the discussion that follows. We describe here a generic, broadly applicable approach; in Section 3.4, we will exploit special structure of the applications to refine this approach. Let S denote the assortment carried in a particular store and x_i the sales of SKU $i \in S$, during a history period.

We can write the probability $F_j(S)$ that a customer purchases $j \in S$, as $F_j(S) = f_j + \sum_{i \notin S, j=j(i,S)} f_i \pi_{ij}$. Let $F(S) = \sum_{j \in S} F_j(S)$ denote the probability that a customer shopping in this category makes any purchase from assortment S. Then, assuming that each consumer purchase is an independent random draw, the likelihood of observing sales data $x = \{x_i\}_{i \in S}$ is given by

$$LH(f, \pi) = C \prod_{j \in S} \left[\frac{F_j(S)}{F(S)} \right]^{x_j},$$

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

$$LLH(f, \pi) = \sum_{j\in S} x_j \log F_j(S) - \left(\sum_{j\in S} x_j \right) \log F(S) \tag{3.1}$$

subject to the constraints

$$\sum_{u=1}^{N_a} f_{au} = 1, \quad a = 1, 2, \ldots, A \tag{3.2}$$

$$f_i = \Pi_{a=1}^{a=A} f_{ai_a} \quad i = 1, 2, \ldots, n \tag{3.3}$$

$$\pi_{ij} = \Pi_{a=1}^{a=A} \pi_{ai_a j_a} \quad i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, n \tag{3.4}$$

$$f_{au}, \pi_{auv} \in [0, 1] \quad \forall a, u, \text{ and } v \tag{3.5}$$

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_{au} and π_{auv} as

$$f_{au} = \frac{\exp(\widehat{f}_{au})}{\sum_{u=1}^{N_a} \exp(\widehat{f}_{au})}, \quad u = 1, 2, \ldots, N_a - 1 \text{ and } \widehat{f}_{aN_a} = 1 \tag{3.6}$$

$$\pi_{auv} = \frac{\exp(\widehat{\pi}_{auv})}{1 + \exp(\widehat{\pi}_{auv})}, \quad \forall a, u, \text{ and } v \tag{3.7}$$

We found examples showing that the log-likelihood function may not be concave, which implies that numerical optimization methods may not converge to a 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 MLE estimates for demand shares and substitution probabilities, as noted in the previous section, we estimate total demand for the product category as $D = \frac{\sum_{i \in S} x_i}{F(S)}$, and D_i as $f_i D$.

3.3.2 Estimating 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 related to attribute levels. Hence, we regressed the log of price on attribute levels to obtain the pricing equation:

$$\log(p_i) = \alpha_0 + \sum_{a=1}^{A} \sum_{u=1}^{N_a - 1} \beta_{au} z_{iau}, \quad i = 1, 2, \ldots, n \tag{3.8}$$

where z_{iau} is a dummy variable taking the value one if SKU i has level u of attribute a, and zero otherwise.

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

3.3.3 Heuristics for Choosing Assortments

If there were no substitution, then the assortment problem could be optimally 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 is complex to solve optimally. Hence, we define greedy and interchange heuristics for choosing the assortments $S_l, l = 1, 2, \ldots, L$ and the specification $a(s), s \in M$ of the assortment assigned to store s.

For assortment planning, Kök and Fisher (2007) use a greedy heuristic and Chong et al. (2001) an interchange heuristic. For the product line design problem, Green and Krieger (1985) use greedy and interchange. Belloni et al. (2008) find that for the product line design problem, greedy and interchange together find 98.5% of optimal profits on average for randomly generated problems, and 99.9% for real problems.

We first define a greedy heuristic for finding a single assortment $S^G(T)$ for a specified subset of stores T. In the statement of Greedy(T) below, we define $R_s(\varnothing) = 0$.

We also use an interchange heuristic which starts with a given assortment and tests whether interchanging a SKU which is not in the assortment with a SKU in the assortment would increase revenue. Any revenue increasing interchanges are made as they are discovered. The process continues until a full pass over all possible interchanges discovers no revenue increasing interchanges. We apply the interchange heuristic both starting with the greedy assortment and starting with

Algorithm 3.1 Greedy(T)

---

1. INITIALIZE $S^0 = \emptyset$, $k = 1$
2. WHILE $(k \leq K)$ DO

   (a) $j_k = \arg\max_{j \notin S^{k-1}} \sum_{s \in T} R_s \left(S^{k-1} \cup \{j\} \right)$

   (b) $S^k = S^{k-1} \cup \{j_k\}$

   (c) $k = k + 1$

   END WHILE

3. RETURN $S^G(T) = S^K$

---

random assortments.

Algorithm 3.2 Forward Greedy for Finding L Assortments

---

1. INITIALIZE

   (a) $l = 1$, $S_1 = S^G(M)$

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

   (c) $\mathbf{E} = \{S^G(\{s\}) \mid s \in M\}$

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

2. WHILE $(l < L)$ DO

   (a) $S^* = \arg\max_{S \in \mathbf{E}} \sum_{s \in M} R_s \left(S_{a^*(s)} \right) - R_s \left(S_{a(s)} \right)$

       where $a^*(s) = \arg\max_{i: S_i \in \mathbf{C}^l \cup S} R_s(S_i)$

   (b) $l = l + 1$

   (c) $S_l = S^G(\{s^*\})$

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

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

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

   END WHILE

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

To find a portfolio of L assortments and assignments of stores to assortments, we have two alternative heuristics, a forward and reverse greedy. In the forward greedy heuristic, we first apply Greedy(T) $m + 1$ times with $T = M$ and $T = \{s\}, s \in M$, initialize $S_1 = S^G(M)$, and assign all stores to this assortment. If $L > 1$, we

identify the assortment $S^* \in E = \{S^G(s) \mid s \in M\}$ to add that would maximize the incremental revenue gain. To identify S^*, we calculate the incremental revenue gain from adding each assortment $S \in E$ by reassigning stores to their revenue maximizing assortment in $S_1 \cup S$ and calculating the increase in revenue due to the reassignment. We then choose as S^* the assortment that gives the greatest revenue increase 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.

In the reverse greedy heuristic, we first apply Greedy(T) m times with $T = \{s\}$ for all $s \in M$, initialize $S_i = S^G(M)$ for $i = 1, 2, ..., m$ and set $a(s) = s$ for all $s \in M$. If $L < m$, we identify the single assortment $S^* \in \mathbf{E} = \{S_i \mid i = 1, 2, ..., m\}$ to delete that would minimize the revenue loss. We calculate the incremental revenue loss from deleting any assortment $S \in \mathbf{E}$ by reassigning stores to their revenue maximizing assortment in $\mathbf{E} - S$ 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.

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 us- ing 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

Algorithm 3.3 Reverse Greedy for Finding L Assortments

1. INITIALIZE

    (a) $l = m$

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

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

2. WHILE $(l > L)$ DO

    (a) $S^* = \arg\min_{S \in \mathbf{E}} \sum_{s \in M} R_s \left(S_{a^*(s)} \right) - R_s \left(S_{a(s)} \right)$

    where $a^*(s) = \arg\max_{i:S_i \in \mathbf{E}-S} R_s(S_i)$

    (b) $\mathbf{E} = \mathbf{E} - S^*$

    (c) $l = l - 1$

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

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

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

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

    END WHILE

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

Algorithm 3.4 Iterative Procedure to Improve Heuristic Solution

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

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

$T_i = \{ s \in M \mid a(s) = i \}, i = 1, 2, \ldots, L$

$S_i = S^G(T_i), i = 1, 2, \ldots, L$

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

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

END WHILE

WHILE-DO loop in Step 2, but does not provide any significant improvement in

our applications.

3.4 Results

3.4.1 Regional Convenience Chain

This 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*, indexed 1 to 4 for SB_1, SB_2, FB_1 and FB_2 in order.

Thus there were 23 *Flavor* attribute levels, 4 *Brand-Size* attribute levels and 92 possible SKUs, 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. An internal market research study on the industry commissioned 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. The substitute for a particular brand-size is the other brand in the same size. We define $i(j)$ to be the brand-size that would substitute to j if $i(j)$ is not in the assortment and set $i(1) = 2, i(2) = 1, i(3) = 4$ and $i(4) = 3$. We need to estimate the 23 flavor shares f_{1v}, 4 brand-size shares f_{2b}, and 4 substitution probability parameters $\pi_{12}, \pi_{21}, \pi_{34}$ and π_{43}.

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. We describe here a refinement of the estimation procedure described in Section 3.3.1 that exploits some special structure of this application, namely that there are two attributes with no substitution across one of them. Estimation was done at the store level, but for simplicity in the discussion below, we drop the store subscript. As before, let x_j denote the total sales of SKU j at a particular store during July 2005 to December 2005 and let v_j and b_j denote the flavor and brand-size, respectively, of SKU j. Given an assortment S, the sales share of SKU $j \in S$ at any store can then be expressed as

$F_j(S) = f_{1v_j} f_{2b_j}$, if $i(b_j) \in S$ and $F_j(S) = f_{1v_j} \left(f_{2b_j} + f_{2i(b_j)} \pi_{i(b_j)b_j} \right)$, if $i(b_j) \notin S$.

We can then write Equations (3.1) - (3.7) from Section 3.3.1 as maximizing

$$LL(f, \pi) = \sum_{j \in S} x_j \log F_j(S) - \left(\sum_{j \in S} x_j \right) \log F(S) \qquad (3.9)$$

subject to the constraints

$$\sum_{v=1}^{23} f_{1v} = 1,$$

$$\sum_{b=1}^{4} f_{2b} = 1 \qquad (3.10)$$

where all variables are $\in [0, 1]$.

The Lagrangian can be written as

$$H = \sum_{j \in S} x_j \log F_j(S) - \sum_{j \in S} x_j \log F(S) - \lambda_1 \left(\sum_{v=1}^{23} f_{1v} - 1 \right) - \lambda_2 \left(\sum_{b=1}^{4} f_{2b} - 1 \right)$$

$$- \sum_{b=1}^{4} \mu_b \pi_{i(b)b} - \sum_{b=1}^{4} \gamma_b (1 - \pi_{i(b)b}) \qquad (3.11)$$

34

Applying first-order conditions with respect to the flavor shares f_{1v}, we get

$$\frac{\partial H}{\partial f_{1v}} = \frac{1}{f_{1v}} \left[\sum_{j \in S, v_j = v} x_j - \frac{\sum_{j \in S} x_j}{F(S)} F_j(S) \right] - \lambda_1 = 0 \qquad (3.12)$$

Multiplying Equation 3.12 by f_{1v}, and adding the equations across all values of v, gives us

$$\lambda_1 = \sum x_j - \frac{\sum_{j \in S} x_j}{F(S)} \sum_{j \in S} F_j(S) = 0 \qquad (3.13)$$

Hence, we get

$$\frac{\sum_{j \in S, v_j = v} F_j(S)}{F(S)} = \frac{\sum_{j \in S, v_j = v} x_j}{\sum_{j \in S} x_j} \qquad (3.14)$$

which on simplification yields

$$f_{1v} = \frac{\sum_{j \in S, v_j = v} x_j}{\sum_{j \in S} x_j} \frac{F(S)}{\sum_{j \in S, v_j = v} \frac{F_j(S)}{f_{1v_j}}} \qquad (3.15)$$

Note that $\frac{F_j(S)}{f_{1v_j}} = f_{2b_j} + f_{2i(b_j)} \pi_{i(b_j)b_j}$, and can be expressed in terms of the brand-size shares and substitution probabilities, and $F(S)$ can be calculated by using the fact that $\sum_{v=1}^{23} f_{1v} = 1$. 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_{2b}) and substitution probabilities ($\pi_{i(b)b}$) alone.

As was described in Section 3.3, we transform the variables using Equations (3.6) and (3.7) to impose the constraints that they lie between 0 and 1, and use numerical methods based on gradients from randomly generated starting points to maximize this log-likelihood function. Tables 3.3 and 3.4 show the average parameter

estimates across all stores.

<table>
<tr><td colspan="3">Table 3.3 Snack Cakes: Demand Share Estimates (store average)</td></tr>
<tr><th>Brand Size</th><th>Sales (%)</th><th>Demand (%)</th></tr>
<tr><td>$B_1 S$</td><td>67</td><td>61</td></tr>
<tr><td>$B_2 S$</td><td>24</td><td>27</td></tr>
<tr><td>$B_1 F$</td><td>4</td><td>6</td></tr>
<tr><td>$B_2 F$</td><td>5</td><td>6</td></tr>
</table>

<table>
<tr><td colspan="5">Table 3.4 Snack Cakes: Substitution Probability Estimates (store average)</td></tr>
<tr><th></th><th>$B_1 S$</th><th>$B_2 S$</th><th>$B_1 F$</th><th>$B_2 F$</th></tr>
<tr><td>$B_1 S$</td><td>1</td><td>18%</td><td>0</td><td>0</td></tr>
<tr><td>$B_2 S$</td><td>26%</td><td>1</td><td>0</td><td>0</td></tr>
<tr><td>$B_1 F$</td><td>0</td><td>0</td><td>1</td><td>89%</td></tr>
<tr><td>$B_2 F$</td><td>0</td><td>0</td><td>22%</td><td>1</td></tr>
</table>

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 $\dfrac{\sum_{s \in M} \sum_{j \in S^s} \left| \frac{x_j^s}{\sum_j x_j^s} - F_j^s(S^s) \right| x_j^s}{\sum_{s \in M} \sum_{j \in S^s} \frac{x_j^s}{\sum_j x_j^s}}$. 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.

We used a hedonic regression as described in Section (3.3.2) 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 levels that determine consumer value. We multiplied the estimated prices by a scale factor of 0.97 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.

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%[3] at the store-SKU level and 25.8% at the chain-SKU level.

One new SKU was added to the assortment in the July-December 2007 period,

Butterscotch in Brand 2 Single Serve. The MAD and MAPE3[4] of the predicted sales

shares from the actual sales shares at the chain level for the newly introduced SKU

was 16.2%. The 16.2% MAPE 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.

Sources of error affecting both the calibration and validation samples include ran-

dom fluctuation in sales and the approximation of representing SKU shares as the

product of attribute shares. Additional sources of error for the validation sample

include changes in relative prices across brand-sizes, which affect share and substi-

tution probability estimates, and a steady increase in the demand shares of some

newer flavors.

Figure 3.1 shows the results of applying the forward greedy heuristic described

in Section 3.3.3 to compute optimized assortments at the chain level, varying the

number of SKUs in the assortment from 1 to 92. Figure 3.1 also shows results for the

---

[3]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%.

[4]We define MAPE as $\dfrac{1}{\sum_{s\in M}\sum_{j\in S^s} x_j^s} \sum_{s\in M}\sum_{j\in S^s} \left| \dfrac{\frac{x_j^s}{\sum_j x_j^s} - F_j^s(S^s)}{\frac{x_j^s}{\sum_j x_j^s}} \right| x_j^s$

tires and appearance chemicals examples which are discussed in Sections 3.4.2 and 3.4.3. Figure 3.1 shows the percentage captured of the maximum possible revenue if all SKUs were offered, as a function of K expressed as a percentage of n. Note that maximizing revenue for a given value of K is equivalent to maximizing this percentage of maximum revenue captured.

Figure 3.1   Revenue vs. Percent of Maximum Possible SKUs in the Assortment
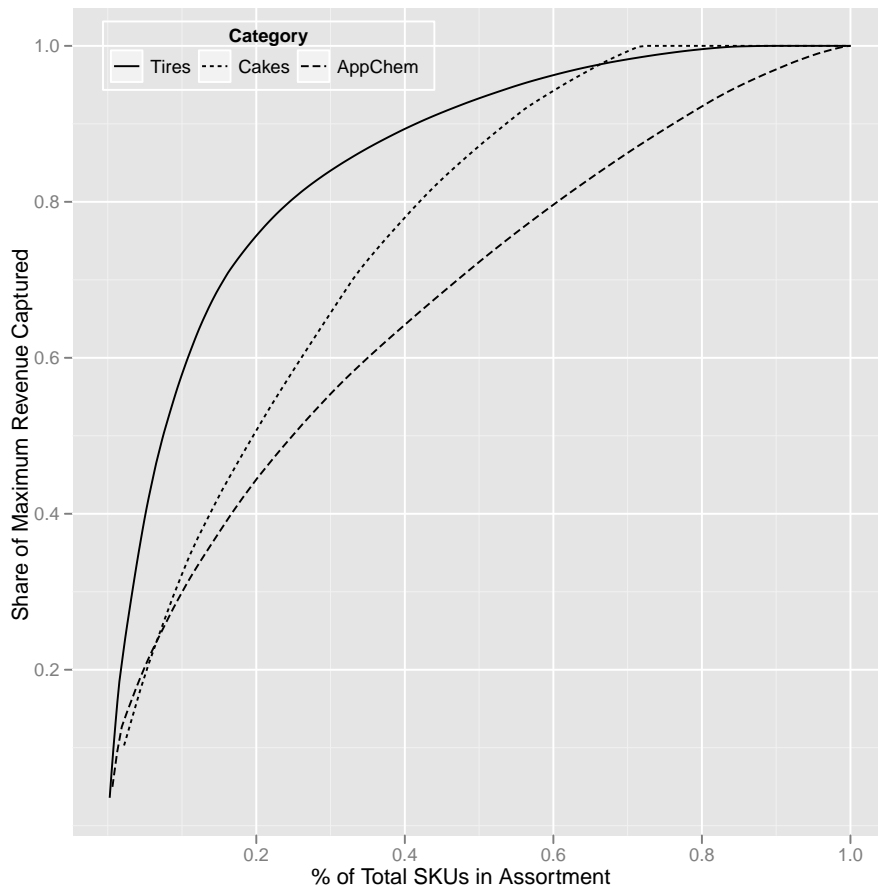
To quantify the potential improvement in revenue, we compare the assortments we generated for $L = 1$ and $L = m$ to the current assortment, which had a revenue of \$6.19 million. Because SKU count varied somewhat by store, in computing the revenue of the $L = 1$ assortment, we first generated the greedy 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 $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 are estimates of revenue improvement based on the calibration sample. If these assortment changes were implemented, we would expect the actual improvement to be less because of differences in the calibration sample sales and sales during the period of implementation. 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

39

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%.

Table 3.5   Snack Cakes: Number of Brand-Sizes by Flavor for the Optimal Assortment at a Representative Store

Flavor	Demand Rank	# Brand Sizes in Optimal Assortment
Cinnamon	1	3
Chocolate	2	3
Peanut Butter	3	3
Butterscotch	4	3
Butter	5	3
Vanilla	6	3
Raspberry	7	3
Fudge	8	3
Honey	9	3
Buttercream	10	3
Choc. Chip	11	2
Cherry/Cheese	12	2
Cheese	13	2
Coconut	14	1
Oatmeal/Raisin	15	1
Jelly	16	1
Vanilla/Chocolate	17	1
Pineapple/Cheese	18	0
Blueberry	19	0
Marshmallow	20	0
Apple	21	0
Cream	22	0
Glazed	23	0

Table 3.5 provides data for an optimized assortment at a representative store and shows that the assortment has an intuitive property; the higher the sales rank of a flavor, the more brand-sizes are carried. Moreover, the same brand-sizes were carried for all flavors that had the same number of brand-sizes. If one brand-size were offered, it is B1S, if two are offered they are B1S and B2F and if three offered, they are B1S, B2F and B2S.

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 3.4 which shows that 89% of consumers are willing to switch from B_1 to B_2 in the family-size segment[5]. 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 more brand-sizes in other flavors. This result made sense to the retailer, who told us that Brand 1 was strongest in single serve, but Brand 2 was by far the strongest in family size, and that's why so many customers were willing to substitute from Brand 1 to Brand 2. They found this the most interesting finding of the study, as they believed substitution rates varied, but had previously had no way to measure the exact rates.

While using a unique assortment from each store adds 12.2% to revenue, 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 level of localization, we applied our assortment

---

[5]Beswick and Isotta (2010), page 2, reports a very similar finding for an orange juice study. For the leading brand, only 21% are willing to substitute to another brands, but for the second brand, 85% are willing to substitute

heuristics for $L = \{1, 2, ..., 6\} \cup \{m\}$, keeping $K = 40$ for all stores.[6] Table 3.6 shows 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.

3.4.2   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

---

[6]We only report results obtained using the forward greedy heuristic as the results based on the reverse greedy heuristic were not significantly different.

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.

We were advised by the retailer that customers do not substitute across sizes; for example, a 14″ diameter tire cannot be used on a 15″ wheel. Table 3.7 is based on estimates provided by the Vice President of the tire category for the retailer and depicts the qualitative likelihood of substitution across brand-warranty levels. We let $\{\pi_S, \pi_L, \pi_M\}$ denote the substitution probabilities *somewhat likely, likely* and *most likely*.

Table 3.6   Snack Cakes: Impact 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 3.7   Tires: Management's Estimate of the Most Likely Substitution Probabilities

From	To					
	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	1

S =Somewhat Likely, L =Likely, M = Most Likely

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 $H3L$. 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 3.2 shows that the estimated share of $H3L$ at each store is negatively correlated and the share of H2H and H2M are positively correlated with median income

level in the zip code in which the store is located, which is 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.

Figure 3.2 Share of $H3L$ **(**$H2H, H2M$**) is Negatively (Positively) Correlated with Income**
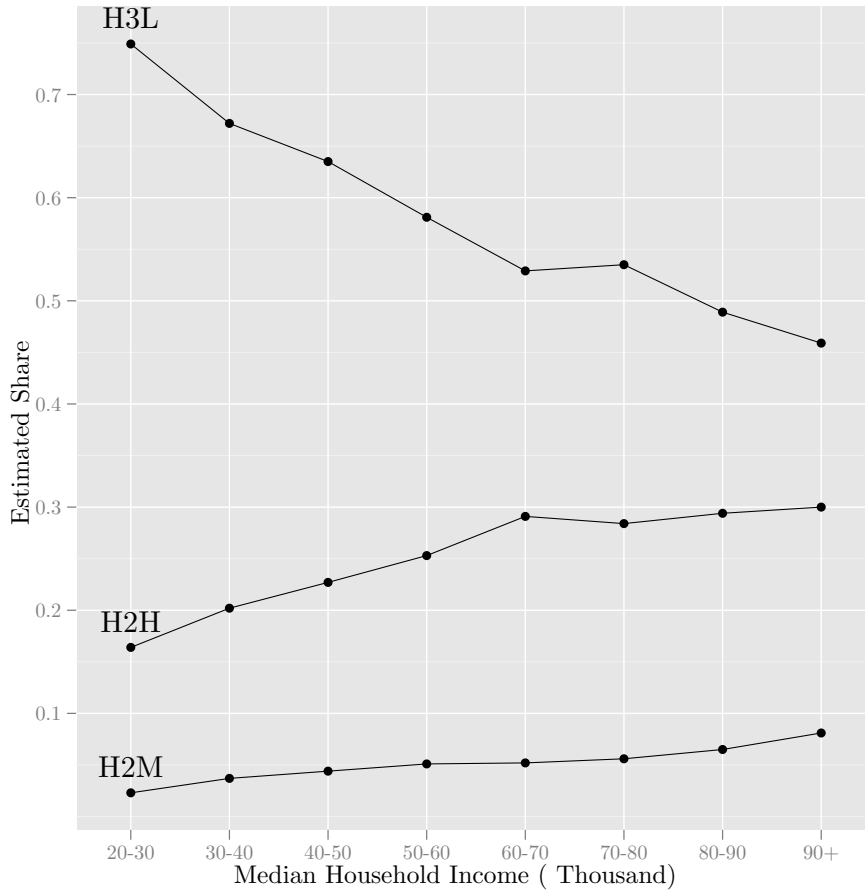
Tables 3.8 and 3.9 show the average across all stores of the estimated brand warranty demand shares and substitution probabilities. The sales-weighted MAD of sales shares predicted based on the parameter estimates from the actual sales shares is

13.6% at the store-SKU level and 4.5% at the chain-SKU level.[7]

<table>
<tr><td colspan="3">Table 3.8 Tires: Demand Share Estimates (store average)</td></tr>
<tr><td>Brand Warranty</td><td>Sales (%)</td><td>Demand (%)</td></tr>
<tr><td>NH</td><td>1</td><td>4</td></tr>
<tr><td>NM</td><td>1</td><td>3</td></tr>
<tr><td>H1H</td><td>3</td><td>4</td></tr>
<tr><td>H2H</td><td>26</td><td>24</td></tr>
<tr><td>H2M</td><td>45</td><td>5</td></tr>
<tr><td>H3L</td><td>24</td><td>61</td></tr>
</table>

Table 3.8 Tires: Demand Share Estimates (store average)

Brand Warranty	Sales (%)	Demand (%)
NH	1	4
NM	1	3
H1H	3	4
H2H	26	24
H2M	45	5
H3L	24	61

Table 3.9 Tires: Substitution Probability Estimates (store average)

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

Table 3.8 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 average, indicating that it is strongly preferred over *H2M*. The retailer offered *H3L* in fewer sizes because they preferred

---

[7]As discussed in Section 3.2, there is some interaction between size and brand-warranty of a tire. One way to account for this interaction is to use a latent class model (Fader and Hardie, 1996, Kamakura and Russell 1989). In a latent class model, we assume that there are several homogenous segments of sizes, and the brand-warranty shares vary across size segments, but are the same within each segment. For example, a size segment could include tires that fit old cars, and the brand-warranty shares reflect this in that the shares of less expensive tires are relatively higher. The estimation problem then reduces to maximizing the likelihood function by jointly estimating the probabilities of segment membership for each size and the associated brand-warranty shares, which can be achieved by using the Expectation Maximization algorithm. Each size is assigned to the size segment for which it has the highest segment membership probability. The choice of the optimal number of segments is made by the Bayesian Information Criterion (BIC), which penalizes the likelihood function for the addition of segments. We estimated a latent class model for a subset of 20 stores. We decided on using two size segments based on BIC. The average share of H3L was 34.2% and 71.3% for the two segments. We also find that the sales-weighted MAD of sales share estimates improved from 13.6% to 11.3%.

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 $H3L$ who did not substitute represents 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.

We can see that offering H3L in only a few sizes hurts revenue. 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 used the hedonic regression described in Section (3.3.2) 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 levels. As before, we multiplied the estimated prices by a scale factor of 1.05 so as to equate the estimated revenue of the current assortment to the actual revenue of this assortment.

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.

Figure 3.1 shows the results of applying the greedy heuristic described in Section

(3.3.3) to compute optimized assortments at the chain level, varying the maximum number of SKUs in the assortment, and Table 3.10 summarizes the store-optimal assortment for a representative store in terms of additions and deletions to 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%[8] 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.

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 com-

---

[8]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.

Table 3.10   Tires: Assortment Changes for a Representative Store

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

Table 3.11   Tires: Revenue Lift vs. Substitution Probability π_M

π_M	Revenues ($ million)	Increase (percent)
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

puted 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 3.11 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 retailer used the same assortment in all stores.

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 re-calibrated 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 3.12 shows a comparison of the predicted vs. actual chain sales shares for the new SKUs. The sales-weighted MAD across all SKUs is 17%, and the MAPE is 19.1%, which compares favorably to the 30.7% MAPE for chain sales of two new SKUs reported by 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,

Table 3.12   Tires: Actual vs. Predicted Chain Level Sales of New SKUs

		Sales Share (per cent)		
Size	**Brand Warranty**	**Actual**	**Predicted**	**Error**
P225/60R16	H3L	3.0	3.4	10.9
P215/70R15	H3L	3.1	3.5	12.6
P205/65R15	H3L	3.2	4.4	37.1
P205/70R15	H3L	2.8	2.7	5.1
P195/65R15	H3L	2.2	2.4	7.6
P215/65R15	H3L	1.0	1.2	25.3
P205/55R16	H3L	1.2	1.1	6.2
P215/60R16	H3L	1.1	1.3	19.4
P215/70R14	H3L	0.9	1.2	37.2
195/70R14	H2H	0.8	1.0	29.5

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

validation and implementation periods to estimate revenue for the current assortment, the modified implementation assortment and the two optimized assortments. Table 3.13 gives revenue estimates and percentage improvement over the current baseline assortment for these periods.

Table 3.13   Tires: Revenue Estimates for Current, Implemented and Optimized Assortments (percentage improvement over current revenues given in parenthesis)

		Revenues ($$ million)		
Assortment		**Jul - Dec 04**	**Jan - Jun 05**	**Jul - Dec 05**
Current	Jul 04 - Dec 04	80.2	74.9	72.3
Implemented	Jul 05 - Dec 05	90.7 (13.1)	84.7 (13.1)	76.5 (5.8)
Recommended	Chain Optimal	104.1 (29.8)	94.3 (25.9)	81.4 (12.6)
Recommended	Store Optimal	108.2 (34.9)	99.2 (32.4)	83.6 (15.6)

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 3.14 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%.

Table 3.14   Tires: Price Changes and Impact on Demand Shares (representative store)

Brand Warranty	Share of Demand (Price in $$)		
	Jul - Dec 04	Jan - Jun 05	Jul - Dec 05
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

3.4.3   Major Auto Aftermarket Parts Retailer

This retailer examines performance of each of their product categories once a year on a rotating schedule and considers changes in the assortment. We were asked in early May, 2009 to apply our methodology for the annual assortment reset for the appearance chemicals category, a category comprised of liquids and pastes for washing, waxing, polishing, protecting, etc. all surfaces of an auto, including the body, tires, wheels, windshield and other glass, and various interior surfaces.

We worked with the appearance chemicals category manager and other members of her team, as well as a staff team that supported category management. On June 1 we received store-SKU sales history on the 160 SKUs currently carried in this category at 3236 stores for the period May 1, 2008 to April 30, 2009. We applied the methodology described in this chapter and reported final results to the senior management of the retailer on July 25. These recommendations were accepted for implementation with the few modifications described below. The new assortment

was implemented during 1/17/2010 – 1/23/2010 and subsequent sales results have been tracked. We describe below the details of this application.

The retailer used a market research firm, NPD, that had assigned to each SKU in the appearance category the attributes (1) segment (defined by the surface of the car treated and what is done to that surface), (2) 9 brands and (3) 3 quality levels, denoted as one of the three levels good, better or best, where 'good' is the lowest quality and 'best' is the highest. We appended package size, denoted small (S) or large (L), to the segment attribute to create 45 segment/size attribute levels. We combined brand and quality to create a second attribute, brand/quality. Because some brands didn't offer all quality levels, there were 17, not 27 brand/quality combinations. In some cases there were two package sizes that differed slightly and were classified as S or L, so that two SKUs occupied the same cell of the attribute matrix. Consequently, the 160 SKUs currently offered corresponded to 130 cells of the 45 x 17 matrix of possible attribute levels. Of the 45 x 17 – 130 = 635 attribute combinations not carried by the retailer, only 24 were available in the market.

We applied the methodology described in this chapter to estimate demand for available SKUs not carried. No substitution parameters were used in the model for the following reason. If a brand-quality level were not offered for a product, it seemed likely that some of the demand for that brand-quality would transfer to several other brands. To estimate this effect we would have needed instances of different stores with varying numbers of brand-quality levels offered for the same product, and this data was not available. The Mean Absolute Deviation of forecasts of existing SKUs across all 3,236 stores was 20.2%.

After parameter estimation, we applied the greedy heuristic to the problem of choosing 130 out of a potential 154 attribute combinations so as to maximize pre-

dicted revenue. We worked first on a 2,183 store 'warm up' case and then the full 3,236 store case, generating up to five different assortments, five being the greatest number of assortments the retailer specified they would consider, given the administrative load of multiple assortments. The estimation process took approximately 5 seconds per store on an Intel Core 2 Duo 2 GHz processor. The computation time for the greedy algorithm to generate five different assortments for all stores on the same computer was approximately 75 minutes.

Table 3.15   Estimated Revenue Increases (%) vs. Number of Store Clusters

No. of Clusters	2183 Stores (warm up)	3236 Stores (implementation)
1	14.1	11.8
2	14.4	11.9
3	14.6	12.0
5	14.9	

Table 3.15 shows the estimated revenue increase for various cases. Based on results of the 2,183 store case, the retailer concluded that at most three store clusters would be used, so these were the cases run for the 3,236 store case. A single assortment resulted in a 11.8% increase in revenues, while store specific assortments lead to a 14.2% increase in revenues, implying a localization lift of 2.4%. The two cluster solution was selected for implementation. In the revised assortments, 20 SKUs in cluster 1 with the lowest estimated revenue were replaced by 20 new SKUs. In cluster 2, 19 existing SKUs were replaced.

Table 3.16 shows the distribution of revenue for the top 24 segment/sizes by segment/size, brand and quality level for the two clusters and some demographic data,

including the percentage of people in the categories 'suburban' or 'urban-bilingual' in the zip codes in which the stores of each cluster are located. Noteworthy differences for cluster 2 are a higher demand for tire related products, higher demand for brand 2, lower demand for brand 5 and a higher percentage of urban/bilingual. Table 3.17 shows these same data for the two stores with highest and lowest percentage suburban. The differences noted above persist, and to a much great degree.

The retailer largely adopted our recommendations. They used exactly the assignment of stores to clusters in our recommendations. They also adopted our recommendations on which attribute combinations to add to the assortment, although in some instances more than one SKU in the market corresponded to the same attribute combination, with the result that the number of SKUs added exceeded the number of attribute combinations added. Twenty two SKUs were added to cluster one and twenty five to cluster two. The sales-weighted MAD across the new SKUs added is 26.8%, and the MAPE is 28.7%, which again compares favorably to the 30.7% MAPE for chain sales of two new SKUs reported by Fader and Hardie (1996). The choice of which SKUs to delete differed from our recommendations in the number of SKUs deleted and in which SKUs were deleted. Seventeen SKUs were deleted from cluster one and twenty three from cluster two. The choice of which SKUs to delete was guided by factors other than year to date revenue; for example, one car wash SKU was deleted due to a history of quality issues.

With respect to localization, the retailer regarded cluster 1 as a base case that was representative of the chain as a whole and cluster two as a subset of stores differentiated by the higher level of tire related purchases, a preference for brand two, and a higher percentage of 'urban/bilingual'. This store segmentation was compelling for the retailer and agreed with more qualitative market research inputs they had

56

Table 3.16 Cluster Statistics for Top 25 SKUs for Appearance Chemicals

Revenue Distribution by Product type - package size, Quality and Brand

Segment - package size	Cluster 1	Cluster 2	Total	Quality	Cluster 1	Cluster 2	Total
Tire Dressings / Shines TRIGGER L	7.1%	12.6%	8.8%	Good	12%	14%	13%
Multi-Purpose Protectants S	7.2%	8.4%	7.6%	Better	60%	67%	62%
Washes L	7.4%	6.6%	7.1%	Best	28%	20%	25%
Tire Dressings / Shines AEROSOL S	5.0%	7.4%	5.8%	**Total**	**100%**	**100%**	**100%**
Tire Cleaners TRIGGER L	4.6%	7.4%	5.5%				
Liquid Wax S	5.5%	3.6%	4.9%	**Brand**	**Cluster 1**	**Cluster 2**	**Total**
Wash and Wax S	4.6%	4.4%	4.6%	Brand 1	18%	20%	19%
Washes S	4.9%	3.8%	4.6%	Brand 2	7%	11%	8%
Multi-Purpose Protectants L	4.0%	3.8%	4.0%	Brand 3	6%	6%	6%
Tire Foams (multi-purpose) AEROSOL S	3.1%	4.1%	3.4%	Brand 4	19%	16%	18%
Carpet / Upholstery Cleaners AEROSOL S	3.2%	3.0%	3.1%	Brand 5	12%	7%	10%
Spray Wax S	3.0%	2.6%	2.9%	Brand 6	6%	7%	7%
Paste Wax S	3.0%	2.2%	2.8%	Brand 7	2%	2%	2%
Wheel Care / Cleaner. ALL TRIGGER S	2.7%	2.8%	2.7%	Brand 8	11%	10%	11%
Spray Detailers L	3.0%	2.1%	2.7%	Brand 9	19%	21%	19%
Leather Cleaners / Conditioners TRIGGER S	2.7%	2.0%	2.5%	**Total**	**100%**	**100%**	**100%**
Wheel Care / Cleaner. ALL TRIGGER L	2.4%	2.5%	2.4%				
Rubbing / Polishing Compounds S	2.1%	1.5%	1.9%				
Tire Dressings / Shines BOTTLE GEL S	1.9%	1.2%	1.7%		**Demographic statistics**		
Scratch Removers S	1.8%	1.1%	1.6%		**Cluster 1**	**Cluster 2**	**Total**
Rubbing / Polishing Compounds L	1.9%	0.9%	1.6%	Income Index	0.95	0.80	0.90
Glass Cleaners TRIGGER L	1.5%	1.6%	1.5%	% Suburban	85%	62%	78%
Plastic / Lens Cleaners, Polishes & Repair S	1.6%	1.3%	1.5%	% Urban/bilingual	16%	42%	24%
Glass Cleaners AEROSOL S	1.6%	1.4%	1.5%	Store count	2263	973	15%

Table 3.17 Statistics for Stores with Maximum and Minimum Percent Suburban

Revenue Distribution by Product type - package size, Quality and Brand

Segment - package size	Store A	Store B	Quality	Store A	Store B
Tire Dressings / Shines TRIGGER L	5.5%	16.6%	Good	13%	18%
Multi-Purpose Protectants S	5.3%	10.6%	Better	57%	66%
Washes L	7.8%	8.5%	Best	31%	16%
Tire Dressings / Shines AEROSOL S	4.6%	8.8%	Total	**100%**	**100%**
Tire Cleaners TRIGGER L	6.5%	5.3%			
Liquid Wax S	8.1%	2.3%	**Brand**	**Store A**	**Store B**
Wash and Wax S	2.8%	5.0%	Brand 1	16%	18%
Washes S	4.5%	3.4%	Brand 2	6%	15%
Multi-Purpose Protectants L	4.8%	3.2%	Brand 3	5%	6%
Tire Foams (multi-purpose) AEROSOL S	3.4%	3.0%	Brand 4	16%	12%
Carpet / Upholstery Cleaners AEROSOL S	2.3%	3.2%	Brand 5	17%	4%
Spray Wax S	4.1%	2.0%	Brand 6	7%	10%
Paste Wax S	2.2%	3.0%	Brand 7	2%	2%
Wheel Care / Cleaner. ALL TRIGGER S	2.3%	2.8%	Brand 8	10%	10%
Spray Detailers L	5.1%	1.2%	Brand 9	22%	23%
Leather Cleaners / Conditioners TRIGGER S	3.2%	1.2%	**Total**	**100%**	**100%**
Wheel Care / Cleaner. ALL TRIGGER L	3.4%	1.0%			
Rubbing / Polishing Compounds S	0.5%	2.5%		**Demographic statistics**	
Tire Dressings / Shines BOTTLE GEL S	1.5%	1.8%		**Store A**	**Store B**
Scratch Removers S	1.4%	1.6%	Income Index	0.98	0.88
Rubbing / Polishing Compounds L	2.0%	1.1%	% Suburban	98%	8%
Glass Cleaners TRIGGER L	2.0%	1.1%	% Urban/bilingual	1%	91%
Plastic / Lens Cleaners, Polishes & Repair S	1.5%	1.3%	Location	Marietta, OH	Lithonia, GA

received. In addition to assortment changes, their localization efforts included giving more prominent display and signage for tire products and brand 2 in the cluster 2 stores.

To evaluate the impact of these changes, we had available sales by cluster of the new assortment for the 27 week period January 1, 2010 – July 8, 2010, and of the previous assortment for a comparable period in 2009. In the discussion below, we refer to these as 2010 and 2009 sales, while recognizing they were for only a portion of these years.

SKUs can be segmented into three groups: kept SKUs that were in both the 2009 and 2010 assortments, deleted SKUs that were in the 2009 assortment but not the 2010 and added SKUs that were in the 2010 assortment but not the 2009 assortment. To evaluate the impact of the assortment changes we compared 2010 kept plus added revenue to 2010 kept revenue plus an estimate of what 2010 revenue would have been for SKUs deleted. We are thus comparing the new assortment revenue to an estimate of what the old assortment would have sold in 2010.

We needed to deal with the fact that more SKUs were added than deleted. Twenty two SKUs were added to cluster 1 vs. seventeen deleted and twenty five SKUs were added to cluster 2 vs. twenty three deleted. The retailer had a fixed amount of shelf space allocated to this category and accommodated the increase in SKU count by reducing the shelf space assigned to some of the existing SKUs. They therefore did not view the increase in SKU count as a cause for concern. Still, reducing the space for some existing SKUs might have caused greater stock outs, reducing revenue in a way we could not capture. Thus, to make a more rigorous evaluation of benefits, we used the twenty two and twenty five SKUs for clusters one and two, respectively, with lowest revenue in the 2009 evaluation period in estimating deletion revenue,

thereby equalizing the add and delete counts. Revenue is affected by a variety of factors other then assortment, including weather, the economy and competitive activity. We measured the impact of these other factors by the ratio of 2010 to 2009 revenue for kept SKUs and estimated the 2010 revenue of the deleted SKUs as their 2009 revenue times this factor.

The newly added SKUs were introduced some time after January 1, 2010 and hence were not on sale for the entire January 1 – July 8, 2010 period and moreover took some time to build to a steady state level of sales. Examining the weekly sales data of the added SKUs, we observed that it took them 7 weeks to achieve a steady steady sales rate. Hence, we used added SKU revenue for weeks 8 – 27 scaled by 27/20 as our estimate of added SKU revenue for the period January 1 – July 8, 2010.

The result of these calculations showed a 3.6% revenue increase due to the revised assortment. In addition, there may have been some improvement due to the localized product display and signage in cluster two stores that we were not able to measure. The retailer's appearance chemicals team agreed with our assessment of benefits and believed that the re-assortment exercise had been a success.

3.5   Analysis of Results

3.5.1   Understanding Localization Revenue Lift

The Localization Lift, defined as the revenue increase from using store specific assortments vs. a single assortment for the chain was 12.2%, 5.8% and 2.4% for the snack cakes, tires and appearance chemicals examples described in the previous sections. As we sought to understand what features of the problem data cause these

differences in Localization Lift, our first thought was that Localization Lift must be driven by demand variation across stores. We thus calculated a coefficient of demand variation (COV) defined as $\sqrt{\sum_{i \in N} \sigma_i^2} / \sum_{i \in N} \mu_i$, where μ_i and σ_i denote the mean and standard deviation of revenue shares of SKU i across all stores. COV for snack cakes tires and appearance chemicals was 17%, 11% and 10% respectively. These values shed some light on variation in Localization Lift in that the highest COV matches the highest lift, for snack cakes, but leave open the drivers of variation in Localization Lift between tires and appearance chemicals, where the Localization Lift varies by a factor of three while the COV's are nearly equal.

To better understand this issue, we examined the data more closely and made two observations. First, SKUs can be segmented into three groups: (1) those with such high demand that they were carried in every store optimal assortment, (2) those with such low demand that they were in no store optimal assortment and (3) the remainder. While there may be substantial variation in demand across stores for the first groups, none of this variation impacts Localization Lift.

For example, in the case of appearance chemicals, we see in Table 3.16 that the best selling segment-size for the chain is Tire Dressings/Shines TRIGGER L. Table 3.17 shows substantial difference in the sales rate between two stores for Tire Dressings/Shines TRIGGER L. The best selling single SKU in this segment-size, as defined by brand and quality level, accounted for 5.5% of revenue in Store A and 16.6% in Store B, a 3.3 to 1 difference. Yet even though the SKU sold much worse in Store A, with a revenue share of 5.5%, it clearly made sense to have this SKU in a revenue maximizing assortment for Store A, and hence this difference in sales rate had no impact on the Localization Lift.

Secondly, we noticed substantial variation in the breadth of assortment carried,

from 40 out of 92 possible SKUs for snack cakes to 130 out of 154 possible SKUs for appearance chemicals. A broader assortment means that a single chain optimal assortment captures a greater fraction of potential demand, leaving less room for improvement from assortment localization.

The example in Table 3.18 is designed to illustrate how these patterns can occur. The example compares COV and Localization Lift for two demand cases, and within each case, for K equal to 3 or 4. The price of all SKUs is \$1, so unit demand and revenue are the same. Moreover, total demand is equal for the two stores, so there is no distinction between demand units and demand share. We also assume there is no substitution. The variance column gives the variance in demand across the two stores for each SKU. COV, as we have defined above is the square root of total variance divided by total demand.

In all cases the chain optimal assortment is SKUs 1 through K. For $K = 3$, the optimal assortment for store 1 is SKUs 1, 2 and 3 and for store 2, SKUs 1, 2 and 4. For $K = 4$, SKUs 1 to 4 are an optimal assortment for both stores.

Considering the case $K = 3$, note that Case 1 has the highest COV but the lowest Localization Lift. This happens because almost all of the inter store demand variation occurs for SKUs that are in either both store optimal assortments or neither and hence none of this variation impacts Localization Lift. By contrast, in Case 2, all of the inter store demand variation occurs for SKUs 3 and 4, the very SKUs that differ in the store optimal assortments.

Note also that when K=4, the lift in both cases drops to 0, demonstrating the impact that breadth of assortment can have on lift.

Motivated by what we saw in the data for the three applications and by the features

demonstrated by the example in Table 3.18, we defined two additional metrics which we hypothesize would be correlated with Localization Lift. COV Select (K) is the coefficient of variation for SKUs which are in some but not all store optimal assortments. This metric is a function of K because store optimal assortments depend on K. We re-label our original coefficient of variation metric as COV – All to emphasize its difference with COV – Select (K).

Table 3.18   Example to Illustrate Drivers of Localization

SKU	Demand Store 1	Store 2	Variance	Demand Store 1	Store 2	Variance
1	50	100	1250	75	75	0
2	100	50	1250	75	75	0
3	40	50	50	50	0	1250
4	50	40	50	0	50	1250
5	40	0	800	25	25	0
6	0	40	800	25	25	0
7	40	0	800	25	25	0
8	0	40	800	25	25	0
9	40	0	800	25	25	0
10	0	40	800	25	25	0
11	40	0	800	25	25	0
12	0	40	800	25	25	0
Totals	400	400	9000	400	400	2500
COV			11.9%			6.4%

K	Optimal Chain	Localized	Localization Lift	Optimal Chain	Localized	Localization Lift
3	390	400	2.6%	350	400	14.3%
4	480	480	0	400	400	

We also define Chain Optimal Share (K) to be the share of the total potential revenue $\sum_{s\in M}\sum_{i\in N}p_i D_i^s$ achieved by a single, chain optimal assortment with K SKUs. So far we haven't discussed the impact of substitution. We simply note that an increase in willingness to substitute increases Chain Optimal Share (K) and thus decreases Localization Lift.

Table 3.19 gives Localization Lift and the three metrics for the three applications. We observe that the highest Localization Lift for cakes can be explained by the low value of Chain Optimal (K) and high values of COV – All and COV Select (K). We also note that the difference in Localization Lift between tires and appearance chemicals can be explained by the high value of Chain Optimal Share (K) for appearance chemicals.

We also used the solutions to the three applications for K varying from 1 to n to create Figure 3.3 showing Localization Lift versus Percent of Maximum Total SKUs in the Assortment and Chain Optimal Share (K). We note that Localization Lift varies with Chain Optimal Share (K) as we have hypothesized.

Table 3.19   Explaining Localization Lift

Category	Lift	COV - All	COV - Select	Chain Optimal Share
Cakes	0.122	0.17	0.23	0.71
Tires	0.058	0.11	0.12	0.80
Appearance Chemicals	0.024	0.10	0.10	0.93

To further investigate the effect of demand variation and share captured by the chain optimal assortment on localization lift, we used the existing data to create 100 additional problem instances for each of the three applications by randomizing

(a) the number of stores in the chain, (b) the actual stores sampled and (c) K, the maximum number of SKUs allowed in the assortment.

Figure 3.3   Localization Lift versus Maximum SKUs in the Assortment and Chain Optimal Share
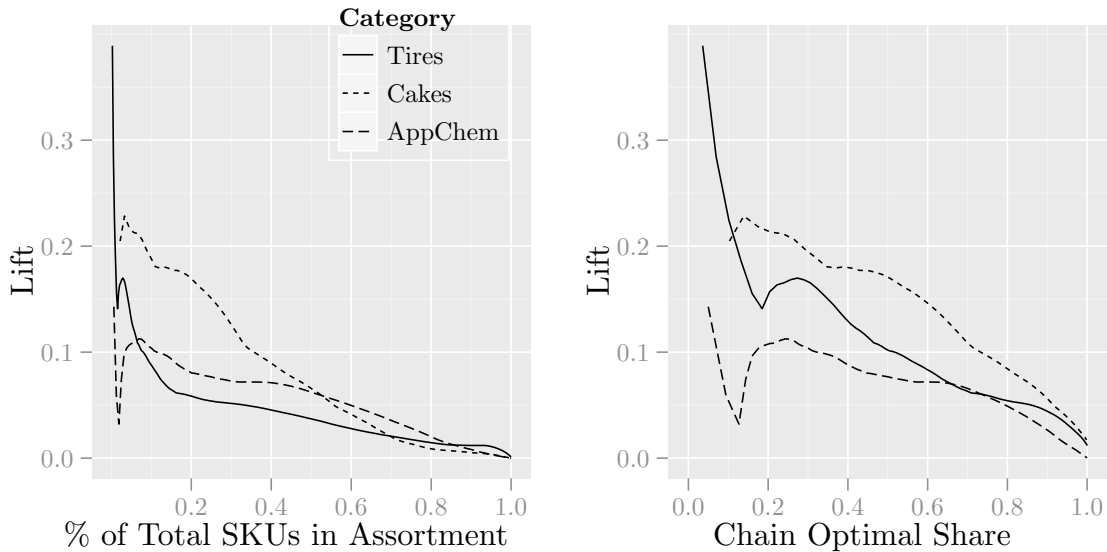
Table 3.20 gives the range used for each application in randomly generating store count and K. We then calculated Localization Lift, COV - All, COV - Select and Chain Optimal Share (K) and regressed Localization Lift against the three dependent variables; the results are summarized in Table 3.20. The regression results confirm our hypothesis that Localization Lift depends mainly on Chain Optimal Share and COV - Select, which are highly significant, and not on COV - All, which is insignificant.

Table 3.20   Regression of Localization Lift

Coefficient	Snack Cakes	Tires	Appearance Chemicals
(Intercept)	0.147***	0.113***	0.186***
Chain Optimal Share	-0.095***	-0.166***	-0.157***
COV Select	0.045*	0.475***	0.339***
COV All	-0.024	0.232	-0.000
Simulation Details			
Minimum K	10	10	5
Maximum K	40	100	40
Minimum Stores	10	20	80
Maximum Stores	30	60	240
# of Instances Simulated	100	100	100

3.5.2   Performance of Heuristics

We consider the quality of the solutions produced by the greedy and interchange heuristics. Both of these heuristics have been used previously for assortment optimization, greedy in Kök and Fisher (2007) and interchange in Chong et al. (2001). Belloni et al. (2008) computationally evaluate greedy and interchange for product line design, a problem similar in structure to assortment optimization, and find that on real problems greedy achieves 98.4% of maximum profit on average and greedy followed by interchange, 99.9%.

Note that if there is no substitution, greedy finds an optimal solution since the assortment problem is then maximization of a linear function subject to an upper bound on the sum of the variables. Thus greedy found optimal solutions for appearance chemicals where there was no substitution, so we will restrict our

attention here to the snack cakes and tires applications.

Note that Table 3.5 shows an interesting property for the assortment for a typical snack cakes store. The flavors are sorted by demand rank and as we move down the list of flavors, the number of brand-sizes offered steadily decreases. The following Theorem will establish conditions under which this property always holds for an optimal assortment. We'll use the theorem to generate a sample of large problems with known optimal solution against which we can test the effectiveness of the greedy and interchange heuristics.

Theorem 1. Consider an assortment planning problem for a store s with the following characteristics.

1. $A = 2$

2. $\pi_{1uv} = 0$ for all u and v, i.e., no substitution across levels of attribute 1

3. If $p(u,v)$ denotes the price of a SKU with levels u and v for attributes 1 and 2 respectively, then there exists constants p_{1u} and p_{2v}, $u = 1, 2, \ldots N_1$ and $v = 1, 2, \ldots, N_2$ such that $p(u,v) = p_{1u}p_{2v}$.

4. The levels of attribute 1 are indexed so that $p_{1u}f_{1u} \geq p_{1,u+1}f_{1,u+1}$, $u = 1, 2, \ldots, N_1 - 1$.

Given an assortment S, let $S(u)$ denote the set of levels of attribute 2 that are present in the SKUs in S that have level u of attribute 1, where $S(u)$ is the empty set if there is no SKU in S with level u of attribute 1.

Then there exists an optimal assortment satisfying $|S(u)| \geq |S(u+1)|$, $u = 1, 2, \ldots, N_1 - 1$.

Proof. Let D denote total unit demand at the store, $k(u) = |S(u)|$, and for $v \notin S(u)$, $w(v, S(u)) = \arg\max_{w \in S(u)} \pi_{2vw}$. $w(v, S(u))$ is the level of attribute two that is the best substitute in the assortment S for a customer desiring level v of attribute 2. By the assumption of the theorem, we can write the assortment optimization problem for store s as

$$\max D \sum_{u=1}^{N_1} p_{1u} f_{1u} \max_{S(u), |S(u)|=k(u)} \left[\sum_{v \in S(u)} p_{2v} f_{2v} + \sum_{v \notin S(u)} p_{2w(v,S(u))} f_{2v} \pi_{2vw(v,S(u))} \right]$$

Note that the expression in square brackets maximized in the choice of $S(u)$ does not depend on u. So letting

$$Z(k) = \max_{S(u), |S(u)|=k(u)} \left[\sum_{v \in S(u)} p_{2v} f_{2v} + \sum_{v \notin S(u)} p_{2w(v,S(u))} f_{2v} \pi_{2vw(v,S(u))} \right],$$

we can express assortment revenue as $D \sum_{u=1}^{N_1} p_{1u} f_{1u} Z(k(u))$.

If the theorem is violated, then there is a u such that $k(u) < k(u+1)$. But then $Z(k(u)) \leq Z(k(u+1))$, because $Z(k(u+1))$ is the optimal value of a maximization problem that is less constrained than that which determines $Z(k(u))$. This, together with assumption (4) of the theorem implies

$$p_{1u} f_{1u} Z(k(u+1)) + p_{1,u+1} f_{1,u+1} Z(k(u)) \geq p_{1u} f_{1u} Z(k(u)) + p_{1,u+1} f_{1,u+1} Z(k(u+1)).$$

So we can revise the solution by assigning the set $S(k(u+1))$ for attribute level u and the set $S(k(u))$ for attribute level $u+1$ without reducing revenue. The revision removes this violation of the theorem. Repeated application of this step will produce an optimal solution satisfying the condition of the theorem.

The snack cakes and tires applications satisfy conditions 1 and 2 of the theorem. With prices as determined by the hedonic regression, they also satisfy 3, and 4 is easily satisfied by indexing attribute 1 as required. In the applications, we used real, not estimated, prices for current SKUs, so the problems as we solved them did not exactly satisfy the conditions of the theorem. However, we subsequently used the theorem to find optimal solutions to store level assortment problems for all stores for the cakes and tires applications with estimated prices and then tested the heuristics against these known optima.

To find the optimal solutions using the theorem, we first note that because the expressions in square brackets in the proof doesn't depend on the level u of the first attribute, the same set is optimal for all u with a common value of $k(u)$. Moreover, the number of levels for the second attribute (4 for cakes, 6 for tires) is small enough that we can enumerate all subsets of these attribute levels, and for each possible cardinality k, identify the optimal subset and associated values $Z(k)$. Then the assortment optimization problem reduces to finding optimal values of $k(u)$, $u = 1, 2, \ldots, N_1$, satisfying $k(u) \geq k(u+1)$ and $\sum_{u=1}^{N_1} k(u) = K$. The number of feasible values of $k(u)$ is small enough to allow for complete enumeration.

Using this approach we found optimal assortments with estimated prices for each of the 140 stores in the snack cakes application and the 574 stores in the tires applications. This gave us a sample of 714 problems of realistic size and known optimal assortments on which to test the performance of our heuristics. The results of applying greedy to all of these problems are reported in Table 3.21 and show that greedy finds near optimal solutions and performs similarly on our applications to what Belloni et al. (2008) found for the product line design problem.

We also tested the interchange heuristic, starting both with the greedy assortment

Table 3.21   Performance of Greedy Heuristic

Product	N_1	N_2	Average Performance (%)	Finds Optimal Solution (%)	Average Time per Trial (seconds)
Cakes	23	4	97.2	74.6	0.14
Tires	64	6	98.5	80.3	0.34

and random assortments, but in no case found solutions that improved on the greedy assortment.

3.6   Conclusions

We have formulated a process for finding optimal assortments, comprised of a demand model, and estimation approach and heuristics for choosing assortments. We have applied this process to real data from three applications and shown that the approach produces accurate forecasts for new SKUs. Our recommendations were implemented in two of the cases. We measured the impact based on actual sales and found the assortment revisions had produced revenue increases of 5.8% and 3.6%, which are significant relative to typical comparable store increases in these product segments.

We note the following observations from this research.

1. Forecast accuracy for new SKUs was adequate to achieve significant benefits in implementation. The only prior reported results for forecasts of new retail SKUs is Fader and Hardie (1996), who reported an average MAPE of 30.7%

for two new grocery SKUs. We found a MAPE of 16.2% for one new snack cakes SKU, 19.1% for 11 new tires SKUs, and 28.7% for 25 new appearance chemical SKUs, somewhat improving on the results of Fader and Hardie (1996). Nonetheless, the errors were great enough to reduce the revenue increase by about half from the fit to validation samples, so improving forecast accuracy would be a useful focus of future research.

2. Sales is not true demand, but demand distorted by the assortment offered. We don't see demand for SKUs not offered, and the sales of SKUs offered may be increased above true demand due to substitution. The impact of these effects can be significant. 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 the lowest price brand-warranty in few sizes. Adding more of this brand warranty to the assortment was a big source of the revenue increase attained.

3. Substitution can be measured, can vary significantly and have a major impact on the optimal assortment. In the 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.

4. We were able to use demographic data to confirm our parameter estimates in the tire and appearance chemicals examples. 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. In some instances, we also

used demographic data to assist in estimating parameters.

5. The benefit from localizing assortments by store varied considerably, from 2% to 12%. We showed that this difference is not driven by demand variation across stores, but by the percent of maximum revenue captured by a chain optimal assortment and by variation in demand for those SKUs that vary across store optimal assortments.

6. A limited amount of localization can capture most of the benefits of maximum localization. In the snack cakes example, going from 1 assortment to 6 provided 77% of the benefit as going from 1 assortment to 140 assortments.

7. There may be interaction between attribute levels not captured by our simplest demand model. In the case of tires, the demand for the least expensive brand-warranty level 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. We showed that this could be incorporated into our approach through latent class analysis.

Chapter 4

A Sensitivity Analysis of Assortment Planning Models

4.1 Introduction

From Section 2.3.1, we know that the solution to the assortment optimization problem depends on assumptions made about (a) consumer response to stock-outs, and (b) the underlying choice process. Mahajan and van Ryzin (2001) use a few examples to illustrate the effects of ignoring stock-out substitution on the optimal assortment and profits. Similarly, Gaur and Honhon (2006) show that the structure of the optimal assortment can be vastly different based on the choice model assumed. The objective of this chapter is to investigate the sensitivity of the optimal assortment and expected profits to these two key modeling assumptions, in a more systematic fashion.

Customer response to stock-outs is a key assumption central to the assortment optimization problem. Most models assume that customers do not substitute in the

event of stock-outs, in order to obtain closed form expressions for the expected profits and keep the optimization problem tractable. A stream of recent research papers have focused on developing heuristic approaches to the assortment optimization problem in the presence of stock-outs. However, a key question that remains to be asked is how well does the solution that ignores stock-out substitution, perform in its presence. While, current assortment literature sheds some light on this issue, there is need to investigate the effect of stock-out substitution on the optimal assortment and expected profits, in a more systematic fashion, especially focusing on the effect of key problem parameters. We use the same model setup described in **?** to investigate the performance of a heuristic ignoring stock-out substitution, on the optimal assortment and expected profits. In addition, we also study the performance of another heuristic proposed by Mahajan and van Ryzin (2001), which makes the extreme assumption that the probability of substitution under a stock-out equals one. Our key contribution is a systematic investigation of these simple heuristics under a wide range of problem parameters, to provide deeper insights into their performance in the presence of stock-out substitution.

The process by which customers make their choices is an important ingredient while modeling assortments. Two commonly used choice models are the multinomial logit model (MNL) and the locational choice model (LC). The focus of assortment research has generally been on optimizing the assortment given complete knowledge of the choice process. However, in many practical situations, the choice process is not known clearly. In such cases, an incorrectly specified choice model can have an adverse impact on the optimal solution. For example, Farias et al. (2011) show how such model mis-specifications can lead to sub-optimal solutions. Our objective is to investigate the effect of choice model mis-specification on the optimal assortment

and expected profits. More specifically, we study the locational choice model proposed by Gaur and Honhon 2006 and analyze the deviation from optimality when an MNL model is used.

The rest of this Chapter is organized as follows. In Section 4.2, we investigate the impact of ignoring stock-out substitution on the optimal expected profits. We study the effects of using an MNL model to optimize the assortment, when the underlying choice behavior follows the LC model in Section 4.3. Finally, in Section 4.4 we summarize the key insights obtained and provide conclusions.

4.2 Impact of Stock-Out Substitution

4.2.1 Substitution Behavior

A key consideration in assortment planning models is substitution behavior of customers. There are two types of substitution considered in the assortment planning literature.

1. **Assortment Based Substitution (A)** assumes that customers make their purchase decisions based on the assortment offered at the start of the period. Hence, they DO NOT substitute across products in the assortment in the event of a stockout. This is also referred to as *static substitution*.

2. **Stockout Based Substitution (S)** assumes that customers make their purchase decisions based on the assortment available to them at the time of their visit. Hence they DO substitute across products in the assortment in the event of a stockout. This is also referred to as *dynamic substitution.*

While dynamic substitution is a more realistic model of reality, it makes the analysis intractable since item level demands are a function of the inventory vector and computing the expected profits analytically, even for simple choice models is not possible. Mahajan and van Ryzin (2001) show that the expected profit function under dynamic substitution is not even quasi-concave, and hence global optimization may be difficult. Assuming assortment based substitution simplifies the analysis considerably, as it allows the decomposition of the joint assortment selection and inventory management problems across products. However, using the solution to a static model when in fact dynamic substitution is present, comes at a cost as it misses the impact of substitution due to stock-outs, and thereby provides a sub-optimal solution.

4.2.2 Heuristic Methods

Given the complex nature of the dynamic substitution problem, several heuristics have been proposed to solve it. Mahajan and van Ryzin (2001) propose a *sample path gradient algorithm* to compute the optimal assortment under dynamic substitution. They observe that allocating inventories based on the assumption of assortment-based substitution tends to ignore two effects (1) the excess demand from stock-out substitution which provides an incentive to increase inventory, and (2) the reduction in underage cost which creates an incentive to reduce inventory. Hopp and Xu (2008) approximate the expected profits under dynamic substitution using a fluid network model and a service-inventory mapping and solve for the optimal assortment using this static approximation. Honhon et al. (2010) study the dynamic substitution problem under a general choice model specified in terms of customer types, where each type is defined by a rank ordered set of products that they are willing to

purchase.[1] They approximate the expected profits under dynamic substitution using a *fixed proportion heuristic,* where the proportion of customer arrivals of each type is assumed to be constant at all points in time. They solve the modified problem to optimality using a dynamic programming approach which has a complexity of $O(8^n)$. Honhon and Seshadri (2011) further show that the *fixed proportion heuristic* proposed by Honhon et al. (2010) provides an upper bound on the optimal expected profit under dynamic substitution.

In addition to the sample path gradient heuristic, Mahajan and van Ryzin (2001) propose two simple newsvendor heuristics to the joint assortment planning and inventory management problem. They use it as a baseline to benchmark the performance of the sample path gradient algorithm.

Independent Newsvendor Heuristic (I)

The independent newsvendor heuristic makes the simplifying assumption that demand for each item in the assortment is independent of the current on-hand inventory levels. This assumption is true when the substitution behavior is assortment-based or static, and customers do not substitute on account of stock-outs.

Under this assumption, if $\mathbf{x} = (x_1, x_2, \ldots x_n)$ is the starting inventory vector, then the expected profit made by the retailer on item j is given by

$$\mathbb{E}\left[p \times \min\left\{x_j, D_j\right\} - cx_j\right],$$

---

[1] Their model assumes that every customer is defined by a type, which is a sequence of products that he is willing to purchase, arranged in decreasing order of preference. If there are n products, then there are a total of $|T| = \sum_{j=0}^{n} \dfrac{n!}{(n-j)!}$ customer types. The model specification is completed by specifying the proportion of customers of each type in the population. This choice model nests the popularly used MNL and Locational Choice models, since they can be obtained by imposing certain constraints on the possible customer types and their proportions in the population.

where x_j represents the inventory of item j and D_j is its demand. Hence, the assortment optimization problem can be written as

$$\max_{\mathbf{x}, S=\{i | x_i > 0\}} \sum_{j \in S} \mathbb{E}\left[p \times \min\{x_j, D_j\} - cx_j \right]$$

Conditional on the assortment S, the optimization problem decomposes into $|S|$ independent newsvendor problems. If we assume that the item level demands can be approximated by a normal distribution with mean and variance $\lambda q_j(S)$, where $q_j(S)$ is the share captured by item j, then the optimal stocking levels for items in the assortment, can be determined as

$$x_j^{*I}(S) = \lambda q_j(S) + z \sqrt{\lambda q_j(S)},$$

for all $j \in S$, where $z = \Phi^{-1}\left(1 - \dfrac{c}{p}\right)$ and Φ denotes the cumulative distribution function of a standard normal distribution. The optimal profits associated with assortment S can be written as

$$\mathbf{E}\left[\Pi^A(\mathbf{x}^{*I}(S))\right] = (p - c) \sum_{j \in S} \lambda q_j(S) - p\phi(z) \sum_{j \in S} \sqrt{\lambda q_j(S)}.$$

Note that the superscript A in the equation for expected profits implies that they are calculated under the assumption of assortment-based substitution.

In the general case, choosing the optimal assortment S is a hard problem and enumerating all 2^n possible candidates might be the only possible solution. However, simpler solutions exist for a handful of cases satisfying some conditions. For example, when consumer choice is governed by the MNL model, all items have the same price and cost parameters, and there are no capacity constraints on number of items

or inventory, **?** show that the optimal assortment S^* is of the form $S^* = \{1, 2, \ldots, k\}$, where the items are indexed in decreasing order of their popularities. This reduces the search for the optimal S^* from 2^n to n candidates. Under the same set of conditions, Gaur and Honhon (2006) show that solving the static substitution problem for the locational choice model can be reduced to a simple line search for a single parameter in a bounded interval.

Pooled Newsvendor Heuristic (P)

The pooled newsvendor heuristic assumes that customers treat the entire assortment as a single product and freely substitute across them in the event of a stockout. It is an extreme case of dynamic substitution, where consumers substitute on account of stock-outs with probability equal to one. Mahajan and van Ryzin (2001) refer to this substitution behavior as *complete substitution*, which we denote as C.

Under the assumption of *complete substitution*, the customer is faced with a single product whose demand is approximately normal with mean and variance $\lambda q(S)$, where S is the assortment offered and $q(S) = \sum_{j \in S} q_j(S)$. The optimal aggregate inventory and total expected profit for assortment S are given by

$$
\begin{aligned}
x_{tot}^{*P}(S) &= \lambda q(S) + z\sqrt{\lambda q(S)} \\
\mathbb{E}\left[\Pi^C\left(x_{tot}^{*P}(S)\right)\right] &= (p - c)\lambda q(S) - p\phi(z)\sqrt{\lambda q(S)}
\end{aligned}
$$

Since $\mathbb{E}\left[\Pi^C\left(x_{tot}^{*P}(S)\right)\right]$ is increasing in $q(S)$, it is maximized for $S = N$. Hence, it would be optimal to offer $S = N$, and set aggregate inventory to $x_{tot}^{*P}(N) = \lambda q(N) + z\sqrt{\lambda q(N)}$. A reasonable heuristic to determine item level inventory levels would be to allocate the total inventory in proportion to their market shares. In

other words

$$x_j^{*P} = x_{tot}^{*P}(N)\frac{q_j(N)}{q(N)}$$

Despite the fact that these two simplistic newsvendor heuristics clearly miss out on providing an optimal solution under dynamic substitution, several papers have observed that they perform well in practice over a wide range of problem parameters. Gaur and Honhon (2006) find based on a series of numerical simulations, that the independent newsvendor heuristic leads to expected profits which are at most 1.44% away from the optimal solution on an average. Honhon et al. (2010) find that the independent newsvendor heuristic performs surprisingly well compared to the more sophisticated dynamic programming and sample path gradient heuristics, with an average optimality gap of 0.5% from the best solution. They also observe that the *naive* independent newsvendor heuristic occasionally outperforms the more complicated sample path gradient heuristic.

Similarly, Mahajan and van Ryzin (2001) observe that "*.... the simple Pooled Newsboy heuristic performs remarkably well in the equal-margin case. Perhaps treating an entire category as if it were a single variant and then performing a simple allocation of the aggregate inventory is a reasonable way to manage such assortments in practice*".

Given the simplicity of the independent and pooled newsvendor heuristics, and the fact that they perform surprisingly well in practice, it is important to delve deeper into understanding their performance. In particular, our research objective is to (a) systematically explore the performance of these two newsvendor heuristics under dynamic substitution over a wide range of parameters, (b) develop analytical bounds on their optimality gap in terms of key parameters and (c) provide recommendations as to when can one use either of the two heuristics, without losing

significant profits.

4.2.3 Bounds on Expected Profits

We denote the three types of substitution behavior we defined earlier by A (Assortment Based Substitution), D (Dynamic Substitution) and C (Complete Substitution). We denote the three types of substitution behavior we defined earlier by A (Assortment Based Substitution), D (Dynamic Substitution) and C (Complete Substitution). Let us denote by $\mathbb{E}\left[\Pi^{SU}(\mathbf{x})\right]$, the expected profits obtained under substitution behavior SU, and starting inventory vector $\mathbf{x}$, where $SU = A, D$ or C. Let $\mathbf{x}^{*H}$ denote the inventory vector obtained using heuristic H, where $H = I$ represents the independent newsvendor heuristic and $H = P$ represents the pooled newsvendor heuristic. The assortment under heuristic H is defined by $S^{*H} = \left\{i \in N \mid x_i^{*H} > 0\right\}$, and consists of all items with a positive inventory. We now develop bounds on the expected optimal profits $\mathbb{E}\left[\Pi^D\left(x^{*D}\right)\right]$ under dynamic substitution, where $\mathbf{x}^{*D}$ is the optimal inventory vector, and $S^{*D} = \left\{i \in N \mid x_i^{*D} > 0\right\}$ is the optimal assortment.

$$\mathbb{E}\left[\Pi^A\left(x^{*I}\right)\right] \leq \mathbb{E}\left[\Pi^D\left(x^{*D}\right)\right] \leq \mathbb{E}\left[\Pi^C\left(x^{*P}\right)\right]$$

Assortment Based Substitution assumes that consumers do not substitute in the event of a stock-out. Hence, the expected sales and profits for a given inventory vector $\mathbf{x}$, considering the impact of *Dynamic Substitution*, are higher than that computed based on *Assortment Based Substitution*. Hence, we have

$$
\begin{aligned}
\mathbb{E}\left[\Pi^D\left(x^{*D}\right)\right] &\geq \mathbb{E}\left[\Pi^D\left(x^{*I}\right)\right] \\
&\geq \mathbb{E}\left[\Pi^A\left(x^{*I}\right)\right].
\end{aligned}
$$

Under the assumption of *Complete Substitution*, consumers substitute in the event of a stock-out, with probability equal to one. Hence, for any inventory vector, the expected sales and expected profits under complete substitution are higher than that obtained under dynamic substitution. Hence, we get

$$
\begin{aligned}
\mathbb{E}\left[\Pi^D\left(x^{*D}\right)\right] &\leq \mathbb{E}\left[\Pi^C\left(x^{*D}\right)\right] \\
&\leq \left[\mathbb{E}\Pi^C\left(x^{*P}\right)\right].
\end{aligned}
$$

Note that Proposition 4.2.3 provides us with a bound on the optimality gap of the independent newsvendor heuristic. If we let ϵ_I^{UB} denote the optimality gap of the independent newsvendor heuristic, where the superscript UB refers to the fact that it is a bound, we can write

$$
\begin{aligned}
\epsilon_I^{UB} &= \frac{\mathbb{E}\left[\Pi^D\left(x^{*D}\right)\right] - \mathbb{E}\left[\Pi^D\left(x^{*I}\right)\right]}{\mathbb{E}\left[\Pi^D\left(x^{*D}\right)\right]} \\
&\leq \frac{\mathbb{E}\left[\Pi^C\left(x^{*P}\right)\right] - \mathbb{E}\left[\Pi^A\left(x^{*I}\right)\right]}{\mathbb{E}\left[\Pi^C\left(x^{*P}\right)\right]}
\end{aligned}
$$

This bound can be computed easily based on the solution provided by the independent and pooled newsvendor heuristics. In fact, we can show that if S^{*I} represents the assortment under the independent newsvendor heuristic, then

$$
\epsilon_I^{UB} \leq \frac{\left(1 - \sqrt{\dfrac{q(S^{*I})}{q(N)}}\right) - \dfrac{\phi(z)}{\Phi(z)}\dfrac{1}{\sqrt{\lambda q(N)}}\left(1 - \sum_{j\in S^{*I}}\sqrt{\dfrac{q_j(S^{*I})}{q(N)}}\right)}{1 - \dfrac{\phi(z)}{\Phi(z)}\dfrac{1}{\sqrt{\lambda q(N)}}},
$$

where $q(S) = \dfrac{\sum_{j\in S} v_j}{v_0 + \sum_{j\in S} v_j}$. Note that the bound depends on (1) $z = \Phi^{-1}(r)$, where r

81

is the critical in-stock rate, (2) the arrival rate λ, (3) the overall market penetration of the category $q(N)$ and (4) the relative market penetration of the optimal assortment under static substitution, $\dfrac{q(S)}{q(N)}$.

For certain choice models like the MNL model, we can compute this bound analytically, since it is easy to solve for S^{*I}, and we have closed form expressions for the choice probabilities $q_j(S)$.

4.2.4 Factors Impacting Heuristics

We start with an example from Honhon et al. (2010) with $n = 5$. Consumer choice follows an MNL model with utilities $\mathbf{u} = \{v_i \mid i \leq n\}$, and the exponential utilities, $\mathbf{v} = \exp(\mathbf{u})$ are given by $\mathbf{v} = (11, 1, 1, 1, 1)$ and $v_0 = 1$. The mean demand is $\lambda = 100$, price of items is $p = 10$, while cost is $c = 5$.

The optimal assortment under the independent newsvendor heuristic consists of 1 product with market share given by $q_1\left(S^{*I}\right) = \dfrac{11}{12}$. The critical newsvendor fractile is given by $r = 1 - \dfrac{c}{p} = 0.5$, which gives us $z = \Phi^{-1}(r) = 0$. Since, $q(N) = \dfrac{15}{16}$, we can compute the bounds on expected profits as $\mathbb{E}\left[\Pi^A\left(\mathbf{x}^{*I}\right)\right] = 420$, and $\mathbb{E}\left[\Pi^C\left(\mathbf{x}^{*C}\right)\right] = 431$. Using Proposition 1, we can compute a bound on the optimality gap of the independent newsvendor heuristic as 2.6%.

The inventory vectors (rounded) based on the independent and pooled newsvendor heuristics are calculated to be $\mathbf{x}^{*I} = (91, 0, 0, 0, 0)$, and $\mathbf{x}^{*C} = (69, 6, 6, 6, 6)$. In order to obtain a sharper bound on the optimality gap, we used simulations to compute the actual profits obtained under dynamic substitution. We generated a sample path by simulating the aggregate demand $D \sim Pois(\lambda)$, and the random utilities under an MNL model, given $\mathbf{v}$. We determined for each sample path and starting

inventory vector, the products purchased by each of the D customers, and computed the total profit. We estimated expected profits by computing the mean profit over 1000 sample paths, using a common set of sample paths across all heuristics to ensure a fair comparison. Table 4.1 shows the computed bounds and the expected profits simulated under dynamic substitution. In addition to the independent and pooled newsvendor heuristics, we also include for comparison, the expected profits obtained using the sample path gradient heuristic (Mahajan and van Ryzin 2001) and fixed proportion heuristics (Honhon et al. 2010).

We can make several interesting observations based on Table 4.1. First, the lower and upper bounds we computed are close to each other. In fact, the upper bound we obtain is the same as that obtained by Honhon and Seshadri (2011), albeit using a more sophisticated dynamic programming approach. Second, the independent newsvendor heuristic does reasonably well with an optimality gap of 2.5% from the upper bound. Third, the pooled newsvendor heuristic does better than all the other heuristics bringing the expected profits within 1% of the upper bound.

Why does the Pooled Newsvendor Heuristic perform so well for this example? The answer lies in the implicit connection between category penetration and degree of substitution for the MNL model. Since the MNL model assumes that substitution probabilities are proportional to the market shares of items, the probability that a customer looking for product $i \in S$ would substitute to product $j \in S - \{i\}$, when product i is not in stock is given by $\alpha_{ij} = \dfrac{v_j}{\sum_{k \in S - \{i\}} v_k + v_0}$. Hence, the average probability that a customer would be willing to substitute to another available product in the assortment in the event of a stockout can be written as $\bar{\alpha}_S = \dfrac{1}{|S|} \sum_{j \in S} \sum_{i \in S - \{j\}} \dfrac{v_j}{\sum_{k \in S - \{i\}} v_k + v_0}$. This can be approximated as

83

Table 4.1   Expected Profits under Dynamic Substitution across Heuristics

Heuristic	Expected Profits	Value ($)
Lower Bound	$\mathbb{E}\left[\Pi^A\left(x^{*I}\right)\right]$	420
Independent Newsvendor	$\mathbb{E}\left[\Pi^D\left(x^{*I}\right)\right]$	421
Pooled Newsvendor	$\mathbb{E}\left[\Pi^D\left(x^{*P}\right)\right]$	427
Fixed Proportion	$\mathbb{E}\left[\Pi^D\left(x^{*F}\right)\right]$	426
Sample Path Gradient	$\mathbb{E}\left[\Pi^D\left(x^{*S}\right)\right]$	383
Upper Bound (C)	$\mathbb{E}\left[\Pi^C\left(x^{*P}\right)\right]$	431

$\bar{\alpha}_S \approx \dfrac{1}{|S|}\dfrac{\sum_{j\in S}\sum_{i\in S-\{j\}} v_j}{\sum_{k\in S} v_k + v_0} = \dfrac{|S|-1}{|S|}\theta_S$, where θ_S is the total market penetration of assortment S. For this example $\theta_N = \dfrac{15}{16}$, implying that the probability of substitution in the event of a stockout is $\bar{\alpha}_S \approx \dfrac{4}{5}\times\dfrac{15}{16} = 0.75$. This explains the superior performance of the Pooled Newsvendor Heuristic which assumes that $\alpha_S = 1$. Before delving into a more systematic evaluation of the heuristics, we simulate two other scenarios studied by Honhon et al. (2010).

In the first scenario, Honhon et al. (2010) investigate the impact of relative popularity of the products by varying the utilities $(v_1, v_2, \ldots v_5)$, such that $\sum_{i=1}^{5} v_i = 15$ and $v_0 = 1$. This assumption implies a total category penetration of $\theta = \dfrac{\sum_{i=1}^{5} v_i}{\sum_{i=1}^{5} v_i + v_0} = \dfrac{15}{16}$. In addition, they set $p = 10$, $c = 5$ and $\lambda = 100$. We express the heterogeneity of each assortment in terms of its Gini Coefficient (Atkinson (1970)), G. $G = 0$, when consumers are extremely heterogeneous implying that all products are equally popular, and $G = 1$, when consumers are completely homogenous with all of them
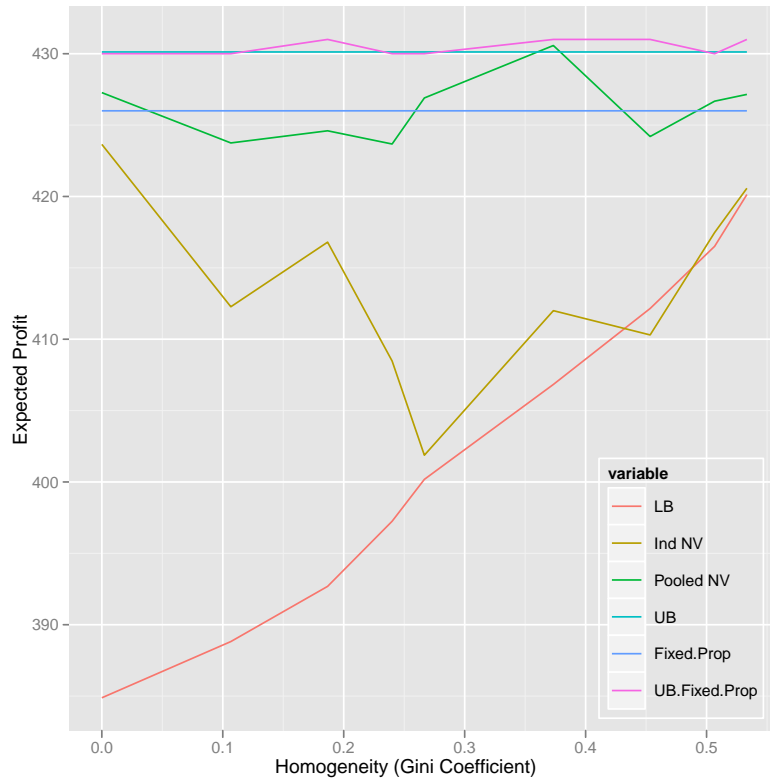
Figure 4.1   Impact of Consumer Homogeneity on Heuristic Performance

preferring the same product.

From Figure 4.1, we observe that the upper bounds we obtain are close to those obtained by Honhon and Seshadri (2011). The simple pooled newsvendor heuristic continues to perform remarkably well and in fact betters the more sophisticated heuristics at times. We also note that the independent newsvendor heuristic has an average optimality gap of 3.8% from the upper bound, and as pointed out by Honhon et al. (2010), its performance deteriorates at intermediate levels of consumer heterogeneity.

In the second scenario, Honhon et al. (2010) investigate the impact of the optimal in-stock rate by varying the overage costs. They set $\mathbf{v} = (7, 2, 2, 2, 2)$, $v_0 = 1$, $c_u = 5$

and $\lambda = 100$, while varying c_o from 1 to 9 in steps of 1. This assortment has a gini coefficient of $G = 0.27$. Figure 4.2 shows that (a) our upper bound is very close to the one computed by Honhon and Seshadri (2011), (b) the pooled newsvendor performs extremely well, and (c) the average optimality gap of the independent newsvendor heuristic is 5.2% which steadily reduces as the in-stock rate increases.
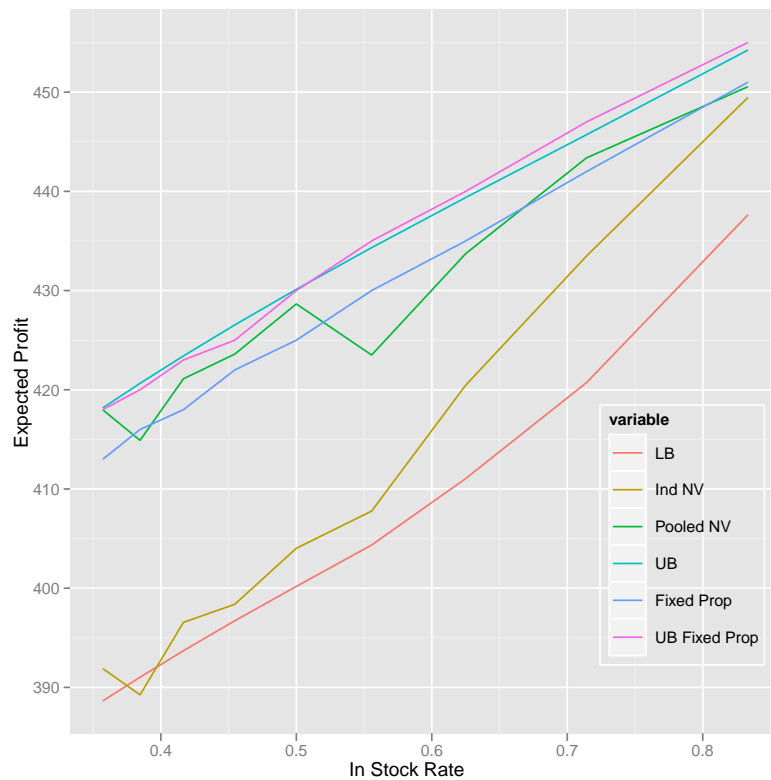
Figure 4.2   Impact of In Stock Rates on Heuristic Performance

4.2.5   Performance of Heuristics

Based on the preliminary results analyzing heuristic performance, we note that there are three key parameters that seem to be driving the optimality gap: (1) Degree of Substitutatbility θ, (2) Degree of Homogeneity G, and (3) In-Stock Rate r.

In addition, we also consider a fourth parameter, the arrival Rate λ.

To systematically investigate the effect of these parameters on heuristic performance, we consider the same setup as used by **?**. Accordingly, we assume that (a) consumer choice follows the MNL model, where $\{u_i \mid i \in N\}$ represent the deterministic utilities associated with the products, with $u_0 = 0$ being the utility of no-purchase, (b) all items have the same price and cost parameters and (c) there are no capacity constraints on inventory or the number of products that can be offered. Under these assumptions, we know that the market share captured by product i when all products in N are offered, is given by $f_i = \dfrac{v_i}{1 + \sum_{i \in N} v_i}$, where $v_i = \exp(u_i), \forall i \in N$. Let the products be labelled in decreasing order of their utilities. Hence $f_1 \geq f_2 \geq \ldots \geq f_N$. Let $F_i = \sum_{j=1}^{i} f_j$ represent the cumulative market share captured by products $\{1, 2, \ldots, i\}$.

The shape of F_i is concave increasing and the curvature depends on the homogeneity of consumer preferences. For example F_i would be a straight line if consumers were extremely heterogeneous making all products equally popular. In order to systematically simulate a wide range of values for the degree of homogeneity G, we let

$$F_i = \theta \left\{ 1 - \left(1 - \frac{k}{n} \right)^{\frac{1+G}{1-G}} \right\},$$

where θ represents total market penetration and G is the Gini Coefficient measuring the heterogeneity of preferences.

Figure 4.3 shows how $\dfrac{F_i}{\theta}$ varies as a function of the fraction of products $\dfrac{k}{n}$ for different values of G. The Gini Coefficient, G, by definition is twice the area above the 45 degree line. It measures homogeneity of preferences across products in the category. Note that $G = 0$, implies complete heterogeneity in consumer preferences

leading to equal shares for all products, while higher values of G imply higher homogeneity in preferences. For instance, $G = 0.75$ leads to the familiar 80-20 rule where 20% of the products capture 80% of the total.
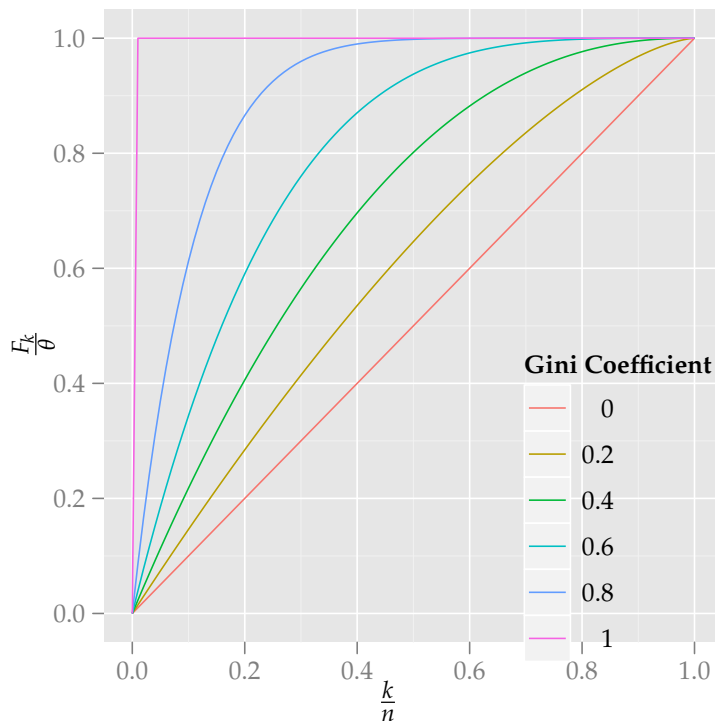
Figure 4.3   Specification of Market Shares

We define the optimality gap of heuristic H with respect to the upper bound as

$$\epsilon_H^{UB} = \frac{\mathbb{E}\left[\Pi^C\left(\mathbf{x}^{*P}\right)\right] - \mathbb{E}\left[\Pi^D\left(\mathbf{x}^{*H}\right)\right]}{\mathbb{E}\left[\Pi^C\left(\mathbf{x}^{*P}\right)\right]}$$

In addition to the computed bounds and solutions based on the two newsvendor heuristics, we also generated solutions to the assortment optimization problem based on the sample path gradient algorithm of Mahajan and van Ryzin (2001) and the static approximation proposed by Hopp and Xu (2008). We denote the

highest expected profits across all heuristics by $\mathbb{E}\left[\Pi_{max}^D\right]$. We use this as a proxy for optimal profits under dynamic substitution, since there is no known algorithm that can guarantee global optimality for this problem. We define the optimality gap of heuristic H with respect to the highest expected profits as

$$\epsilon_H^{OPT^*} = \frac{\mathbb{E}\left[\Pi_{max}^D\right] - \mathbb{E}\left[\Pi^D\left(\mathbf{x}^{*H}\right)\right]}{\mathbb{E}\left[\Pi_{max}^D\right]}$$

We simulate the performance of the various heuristics using the same approach described in Section 4.2.4, and compute the optimality gap over a range of these four parameters as shown in Table 4.2.

Table 4.2 Parameter Values

Parameter	Values
λ	{100, 200, 300, 400}
θ	{0.5, 0.6, 0.7, 0.8, 0.9}
r	{0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9}
G	{0, 0.2, 0.4, 0.6, 0.8}

Table 4.3 summarizes the average optimality gap of the two newsvendor heuristics. We observe that both the independent and pooled newsvendor heuristics perform reasonably well with average optimality gaps of 5.7% and 5.3% respectively, with respect to the upper bound. The bound on the optimality gap computed based on Proposition 1 is 11.0%, roughly twice that the actual gap. When measured with respect to the highest expected profits, the average optimality gaps are 1.8% and 1.3% respectively. This suggests that the simple newsvendor based heuristic perform extremely well on an average.

Table 4.3 **Summary of Heuristic Performance**

Heuristic	ϵ_H^{UB}	ϵ_H^{OPT*}
	%	%
Independent Newsvendor	5.7	1.8
Pooled Newsvendor	5.3	1.3

In order to analyze the optimality gap of the heuristics, we compute the average optimality gap by parameter for each heuristic. Figure 4.4 shows the average optimality gap over different values of θ, r and G for $\lambda = 100$ for the two newsvendor heuristics. We can make several interesting observations on the optimality gap, which are summarized in Table 4.4.

1. **Degree of Substitutability** (θ): The optimality gap with respect to the highest expected profits increases(decreases) with θ for the independent newsvendor (pooled newsvendor) heuristic. This makes intuitive sense, since the independent newsvendor heuristic assumes that $\theta = 0$, while the pooled newsvendor heuristic assumes that $\theta = 1$. The optimality gap with respect to the upper bound decreases with θ for both heuristics. However, we know that for low values of θ, the independent newsvendor heuristic is close to optimal. Hence, we conclude that this behavior is on account of the upper bound being a gross overestimate of expected profits for low values of θ which manifests in the computed optimality gap. This observation is consistent with that reported by Honhon and Seshadri (2011).

2. **In Stock Rate** (r): The optimality gap for the independent newsvendor heuristic decreases with increasing in-stock rate. This occurs since for low values of
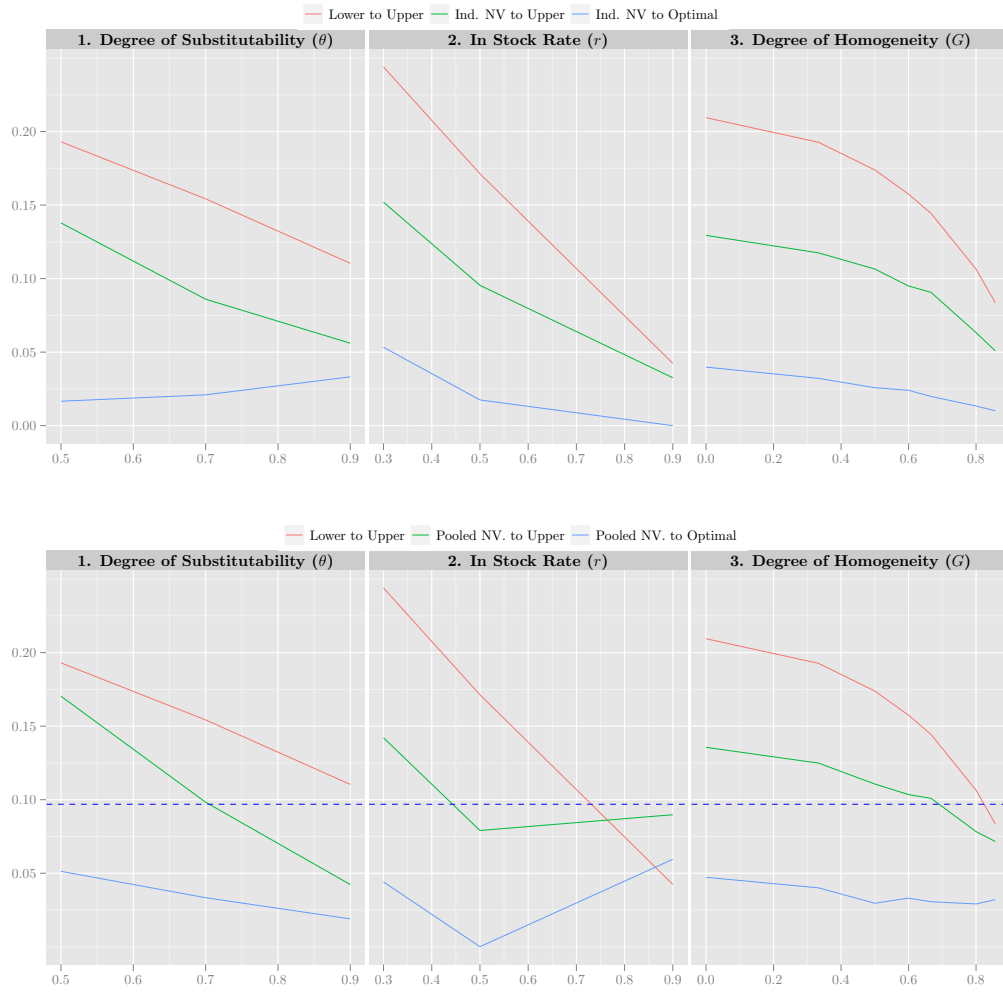
Figure 4.4   Optimality Gap of Newsvendor Heuristics

r, the extent of stock-outs is reduced and as a result the effect of stock-out substitution on the expected profits is smaller. The pooled newsvendor heuristic seems to perform best for intermediate values of r. This behavior is a little puzzling and needs deeper investigation to understand what is driving it.

3. **Degree of Homogeneity** (G): The optimality gap for both heuristics decrease with increase in the degree of homogeneity. For $G = 0$, consumer preferences are extremely heterogeneous. In this case, the independent newsvendor

heuristic selects a much smaller assortment than is optimal as it ignores the effects of dynamic substitution. The pooled newsvendor heuristic misses the mark as well as it assumes consumers regard the whole category as a single product, which is not the case when $G = 0$. On the other hand, when $G = 1$, the optimal assortment consists of just one product, and both heuristics choose this assortment, thereby reducing the optimality gap to zero.

The optimality gap is also decreasing in the arrival rate, λ. This is a direct consequence of the assumption of Poisson arrivals, which implies that the coefficient of demand variation, defined as $\frac{\mu}{\sigma}$, equals $\lambda^{-0.5}$. With increasing λ, the coefficient of demand variation reduces, thereby lowering the effects of any demand-supply mismatch on expected profits. In the extreme case when $\lambda^{-0.5} \to 0$, aggregate demand is deterministic, as a result of which the extent of stock-outs and its effects on expected profits is minimal. As a consequence, any assumption made regarding stock-out substitution is rendered inconsequential as a result of which both heuristics perform extremely well.

Table 4.4   **Effect of Key Parameters on Optimality Gap**

Parameter	Description	Newsvendor Heuristic	
		Independent	Pooled
θ	Degree of Substitutability	↑	↓
r	In Stock Rate	↓	↓↑
G	Degree of Homogeneity	↓	↓
λ	Arrival Rate	↓	↓

One can further understand the relative performance of the newsvendor heuristics by plotting the expected profits across combinations of parameters. Figure 4.5 shows

the expected profits normalized by the upper bound for a selected combination of values of θ, r and G holding $\lambda = 100$. There are several interesting observations we can make. First, the pooled newsvendor heuristic outperforms the independent newsvendor heuristic for low r and high θ, whereas it is the other way round for high r and low θ. In fact, the performance of the independent newsvendor heuristic is close to optimal for high r and low θ. Second, the upper and lower bounds gets tighter as the value of r, θ and G increase.
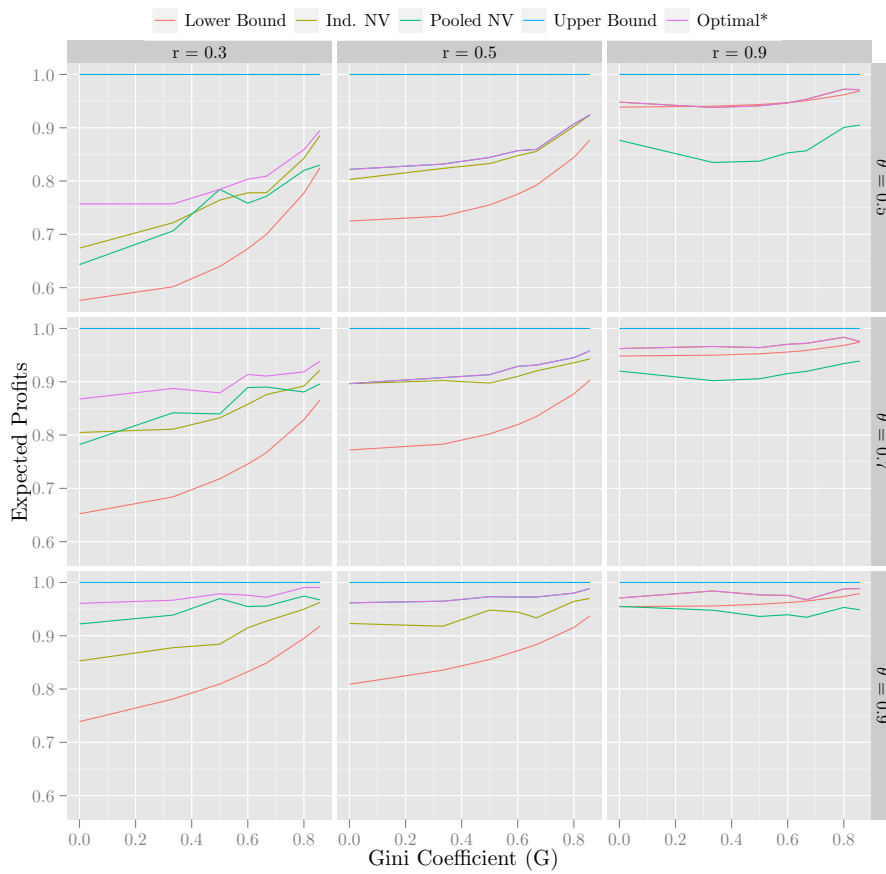
Figure 4.5   Expected Profits Relative to the Upper Bound

4.3 Mis-specification of Choice Model

The optimal assortment and expected profits depend on the choice probabilities $q_i(S)$, which in turn are driven by the choice model considered. In this section, we investigate the impact of misspecified choice model on the optimal assortment and expected profits. Since the MNL model is the most popularly used model in literature, we restrict ourselves to cases where the underlying choice model is different from MNL, and we incorrectly use an MNL model to choose the assortment. Moreover, since we are interested in isolating the effect of the choice model on optimal profits, we ignore the presence of stock-out substitution.

We first consider the Locational Choice (LC) Model discussed in Gaur and Honhon (2006). In some ways, the LC model is a polar opposite of the MNL, since substitution is restricted to products that are similar to each other.

Following Gaur and Honhon (2006), we assume that (a) demand is poisson with rate λ, (b) consumers are distributed on the $[0,1]$ line with continuous probability distribution F, and (c) the coverage interval of each product is given by L (this is the interval containing most preferred goods of all customers who obtain a non-negative utility from the product). Gaur and Honhon (2006) show that if $\mathbf{b} = (b_1, b_2, \dots, b_k)$ is the assortment of products offered, then the expected profits are given by

$$\mathbb{E}\left[\Pi^{LC}(\mathbf{b})\right] = (p-c)\sum \lambda q_j(\mathbf{b}) - p\phi(z)\sqrt{\lambda q_j(\mathbf{b})},$$

where, for $j = \{1, 2, \ldots, k\}$,

$$
\begin{aligned}
q_j(\mathbf{b}) &= F(b_j^+) - F(b_j^-), \\
b_j^+ &= \min \left\{ b_j + L, \frac{b_j + b_{j+1}}{2} \right\}, \\
b_j^- &= \max \left\{ b_j - L, \frac{b_j + b_{j-1}}{2} \right\}.
\end{aligned}
$$

Let $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ denote the set of all possible products that can be offered. Hence, the demand share of product j is given by $q_j(\mathbf{b})$. Let us assume that the values of $q_j(\mathbf{b})$ are known. Consequently, if we use the MNL model to recover utilities from the observed values of $q_j(\mathbf{b})$, we get $u_j = \ln \left(\dfrac{q_j(\mathbf{b})}{1 - q_j(\mathbf{b})} \right)$.

Let us now investigate the effects of using an MNL model to optimize the assortment. The optimal MNL assortment can be determined by arranging the products in decreasing order of their utilities u_j, and selecting the top k products for a certain value of k. We can solve for the optimal value of k through enumeration. Let us denote the optimal MNL assortment by $\mathbf{b}_{MNL}$.

The optimal assortment under the LC model consists of all products if their coverage intervals $[b_j \pm L], j = \{1, 2, \ldots, n\}$ are non-overlapping. If there is an overlap, then for small values of n, we can solve for the optimal assortment through enumeration. Let us denote the optimal LC assortment by $\mathbf{b}_{LC}$.

The impact of the mis-specification can now be easily computed as

$$
\epsilon = \frac{\mathbb{E} \left[\Pi^A(\mathbf{b}_{LC}) \right] - \mathbb{E} \left[\Pi^A(\mathbf{b}_{MNL}) \right]}{\mathbb{E} \left[\Pi^A(\mathbf{b}_{LC}) \right]}
$$

Following Gaur and Honhon (2006), we assume that F follows a beta distribution

with symmetric parameters $\gamma_1 = \gamma_2 = \gamma$. Table 4.5 shows the parameter values considered for this exercise. We find that the average optimality gap is 27%. This

Table 4.5   Parameter Values for the LC Model

Parameter	Values
λ	{100, 200, 300, 400, 500}
r	{0.5, 0.55, 0.6,..., 0.99}
γ	{2, 3, 4, ...,10}
$\mathbf{b}_N$	{0.05, 0.25, 0.45, 0.65, 0.85}
L	{0.10, 0.11, ..., 0.20}

suggests that the use of an incorrect choice model affects the optimal assortment profits significantly. Digging deeper into the performance of the optimal solution, we find that the optimality gap is decreasing in γ, λ, L as shown in Figure 4.6. This can be explained as follows:

1. **Degree of Substitutability** (L): Once again, we observe that the average error ϵ is decreasing in L. When substitution probabilities are high, there are two effects in play. For a high value of L, the MNL model stocks fewer variants as compared to the LC model, thereby increasing ϵ. However, for higher values of L, the MNL assortment is able to capture a market share comparable to that captured by the LC model, which tends to decrease ϵ. In our study, the second effect dominates, thereby leading to a downward sloping curve for ϵ.

2. **Degree of Homogeneity** (γ): We note that the average error ϵ is decreasing in γ. When consumers have heterogenous preferences (low values of γ), the MNL model tends to stock fewer variants as compared to the LC model. This

occurs since the substitution behavior embedded within the MNL assumes that a small subset of the most popular variants will capture demand from the other variants not carried through substitution. However, since true substitution under the LC model is more localized, the MNL assortment does not achieve the market share it expects. As a consequence, for high levels of heterogeneity, the MNL assortment profits are highly sub-optimal as compared to the LC assortment. As consumer preferences become more and more homogenous, this difference reduces, as the most popular variants do end up capturing the lion's share of demand, thereby countering the substitution effect.

3. **Arrival Rate** (λ): Higher value of λ lead to better performance of the MNL. This occurs since higher λ leads the MNL model to increase the number of variants stocked bringing it in line with the number stocked by the LC model, thereby reducing the loss in profits.

4.4 Conclusions

In this chapter, we explored two important issues in assortment modeling. We investigated the impact of (1) ignoring stock-out substitution and (2) using an incorrectly specified choice model, on the optimal assortment and profits. We quantified their effects in terms of the maximum percentage gap from the optimal solution and studied its variation across a wide range of values for key parameters specifying the problem. Our research revealed several interesting insights.

First, we find that an incorrectly specified choice model has a much higher impact on the optimal assortment profits as compared to ignoring stock-out substitution.
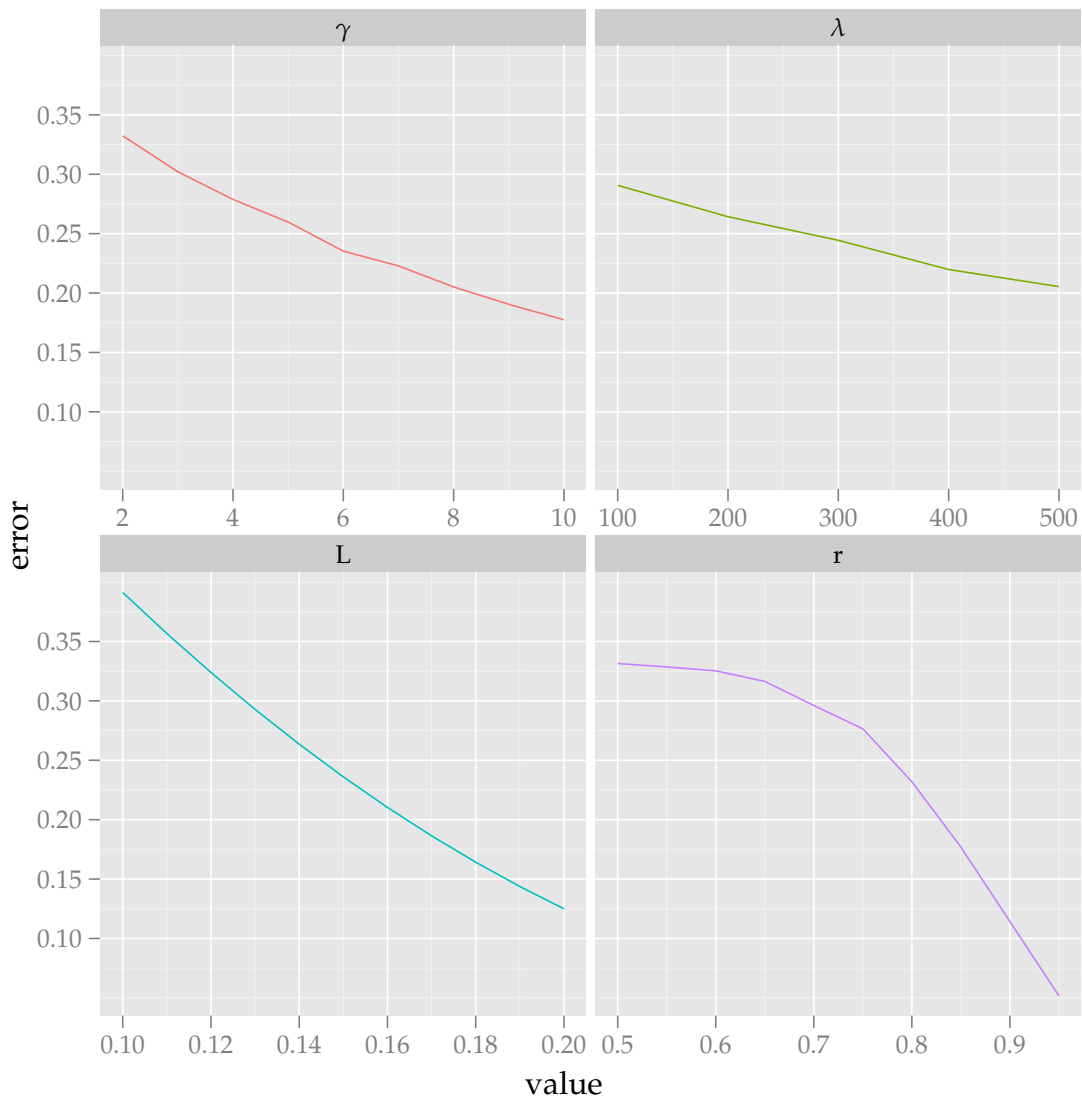
Figure 4.6   Impact of Choice Model Mis-specification

This is significant, since traditionally, much of the focus in OM literature has been on the latter. A more recent stream of literature by Farias et al. (2011) and Besbes and Saure (2010), focus on developing robust approaches to assortment optimization when the choice model is unknown. Our results support this direction of research.

Second, we find that ignoring stock-out substitution does not reduce the optimal

assortment profits significantly. In fact, the independent newsvendor heuristic performs extremely well in spite of the fact that it does not incorporate the impact of stock-out substitution. The independent newsvendor heuristic performs best when the contribution of stock-out substitution to the expected optimal profits is minimal. This happens when the in-stock rate is high, causing stock-outs to be minimal in the first place, or when the degree of substitutability is low, in which case little substitution takes place even when stock-outs actually occur.

Third, we find that the pooled newsvendor heuristic performs beyond our expectations in spite of the fact that it makes unrealistic assumptions on the extent of substitution in the event of a stock-out. Specifically, the solution proposed by the pooled newsvendor heuristic, of offering all available products in the assortment, seems highly impractical for several real situations. One explanation to this seemingly anomalous solution is that although the pooled newsvendor heuristic allocates the total inventory across all products, it does so in proportion to their market shares. Hence, for several products with very low market shares, the inventory on offer could be less than one, which is almost equivalent to it not being offered. A more careful investigation of the pooled newsvendor heuristic needs to be carried out to understand the dynamics of its performance before using it in practice.

Finally, we provide analytical bounds on the optimality gap of the independent newsvendor heuristic. This is handy in practice, since a retailer can simply plug in the parameters of their assortment problem and get a rough estimate of the maximum value of incorporating dynamic substitution into the optimization algorithm.

There are several directions in which this work can be extended. First, the analytical bounds we compute are not tight. The upper bound makes the assumption that substitution probabilities equal one when there are stock-outs. Future work can be

done to sharpen this bound by making assumptions on the maximum probability of a customer substituting in the event of stock-outs. Second, we only explored the impact of one type of choice model mis-specification. It would be interesting to investigate this impact for other commonly used choice models like the nested logit model, exogenous demand model etc. This would help in the development of robust approaches to assortment optimization. Finally, it is important to study the impact of using a Poisson process to model customer arrivals. While the Poisson distribution is a reasonably good way to model customer arrivals for a number of retail product categories, it suffers from the fact that the coefficient of demand variation decreases by the negative square root of the mean arrival rate. Hence, it is important to explore the effect of this assumption on the optimal assortment and expected profits, when demand is actually over-dispersed.

Chapter 5

Measuring Seat Value in Stadiums and Theaters[1]

5.1 Introduction

Theaters and sports stadiums have several characteristics that are well suited to Revenue Management (RM) methods. There are many different customer segments (e.g. season ticket holders, families, students) each with varying usage patterns and willingness to pay. The value experienced by a consumer attending an event depends on several factors, such as the location of his seat, the popularity of the event, and other consumer-related attributes (see Talluri and van Ryzin (2004) for more details). However, there has been limited research on how the value experienced by consumers in such settings is influenced by the aforementioned factors.

---

[1]This chapter is based on Veeraraghavan and Vaidyanathan 2011. Measuring Seat Value in Stadiums and Theaters. *Forthcoming in Production and Operations management*

According to Talluri and van Ryzin (2004) *"fear of negative customer reactions and consequent loss of customer goodwill are the main reasons firms seem to be avoiding bolder demand management strategies."* This fear is not unfounded; Anderson et al. (2004) find a positive association between customer satisfaction and long-run financial performance of firms in retail settings. Hence, it is imperative to develop a systematic understanding of seat value experienced by consumers in order to be able to improve ticket selling strategies. This is our main research objective.

The value of a seat in a stadium/theater is a function of the experience they offer consumers, and could be driven significantly by the location of the seat relative to the stage or playing field. For instance, front row seats in a theater are valued higher as they offer a better view of the performance. This is in stark contrast to airline seats, where seat value in the same travel class is less sensitive to seat location,[2] as airline seats primarily serve as a conduit for transporting a person from an origin to a destination. Consequently, for the most part, the price of a ticket in economy class indicates how much a person values the trip, more than how much he values the seat itself. However, theater/stadium seats might be thought of as experience goods. It is unclear how consumer valuations are distributed across different attributes. Moreover, the dependence of seat value on the location of the seat can be fairly complex. For example, in theaters, seats in the middle of a row might be preferred over seats toward the end of a row further forward, and seats at the front of second-level sections are sometimes preferred to seats at the back of first-level sections (Leslie 2004). This ordering of seat value by location is only understood subjectively by theaters and stadiums. However, there has been little

---

[2] Although there are differences between aisle seats and middle seats, most seats in the same travel class (business or economy) are perceived to provide comparable valuations for consumers. Of late, these seat value differences based on seat location are gaining attention. See *www.seatguru.com*.

research on developing a measure of seat value in these settings. Measuring seat value and developing a better understanding of how it is driven by seat location would assist theaters and stadiums in formulating their ticket selling strategies.

The relationship between seat value and seat location is not well understood. This has been a focus of subjective discussions recently. We briefly discuss one such case. In 2006, the Oakland Athletics decided to reduce the capacity of McAfee Coliseum (where their home games are played) by covering several of their upper deck seats with tarpaulin sheets, thus reducing the stadium capacity from 44,000 seats to about 34,077 seats (Urban 2005). The Oakland A's announced that the decision was made in order to provide an "intimate" experience to those in attendance, in a smaller field. In fact, when the team moves to a newer field for the 2012 season, they plan to play in a stadium that has lesser capacity (32,000) than the currently used tarpaulin-covered stadium. Bnet.com quoted "...*the fans who are feeling slighted most are the lower-income brackets who feel the third deck was their last affordable large-scale refuge for a seat behind home plate, even one so high.*" The team management contended that people liked the upper deck mostly because of availability, and perhaps not so much because of the view (Steward 2006). One article in Slate Magazine criticized the move, stating "*Some of us want to sit far away*" (Craggs 2006). Thus, the seat value perceived by consumers seated at the upper deck was not only unclear, but also varied among different fans. So is it true that the consumers seated in the upper deck valued those seats highly? Were the upper deck seats being underpriced? How did the seat value perceived by consumers attending the game differ across seat locations? These are some of the questions that will be addressed by our research.

In addition to seat location, there are a number of other factors that might affect the seat value perceived by a customer. For instance, in the case of a sports stadium,

the nature of the opposing team, the age of the customer, or whether the customer is a regular or an infrequent visitor, might affect her valuation of the seat. For most theaters and stadiums, understanding heterogeneity in customer valuations is the key to increasing revenues. A clear understanding of the seat valuations would lead to the creation of better "fences" that would provide theaters and stadiums with an opportunity to manage their revenues and customer base better. Our research sheds more light on the key factors influencing seat value in these settings.

Our research on non-traditional industries (theater and sports) complements current RM literature by (1) developing a measure of seat value (Seat Value Index), (2) establishing the critical relationship between the Seat Value Index and seat locations, and (3) providing *segment-specific* recommendations that would help the firm achieve a service-level objective such as a *"desired level of seat value"*.[3] We apply this research methodology to a proprietary dataset collected by a professional baseball franchise in Japan, from a survey of its customers. Based on the findings from the dataset, we provide various measures by which stadiums/theaters can improve customer satisfaction through better handling of ticket pricing, seat rationing, and seating layout decisions. Since RM practices are not employed on a large scale in these areas of interest, our research fills a gap, both in theory and practice.

To our knowledge, ours is the first attempt to study the distribution of consumer seat value and its dependence on seat location in theater/stadium environments. Revenue management practice hinges on the ability to price-discriminate, which is possible only if there is heterogeneity in seat value. Based on service-level objectives, we provide pricing recommendations that a firm may use to improve

---

[3]This notion is analogous to "fill-rate" measures employed in retail settings. While focusing on a desired fill-rate might be sub-optimal for short-run profit maximization, it improves availability, leading to long-run benefits. Quantity adjustments are more difficult in stadiums/theaters, but price adjustments to "satisfice" value can be made.

positive experience from the repeated consumption of the good. We apply our model to a dataset collected by a Japanese baseball franchise and find evidence for heterogeneity in seat value at the stadium. Using our model, we *quantify* this heterogeneity in terms of customer attributes and their seat locations. Pursuant to the results from applying our method, we provide some segment-specific pricing recommendations.

In the following Section §5.2, we position our research with respect to the existing literature. In Section §5.3, we discuss our research design, methodology and its application to a proprietary dataset. In Section §4, we test the robustness of our results to game effects, prices and seat location. In Section §5, we provide segment-specific pricing recommendations and discuss insights from our analysis. We conclude this chapter by summarizing the key ideas of our methodology and charting future research directions.

5.2   Literature Review

We analyze seat value perceived by consumers, and the key implications it has for pricing in sports stadiums and theaters. Most of the literature in the sports and entertainment industry has been about secondary markets and ticket pricing in scalping markets (See Courty (2000) for a comprehensive survey). The only paper similar to our research is Leslie (2004) which studies the profit implications of price-discrimination based on exogenously defined seat quality and consumers' income levels for a Broadway theater. In contrast to Leslie (2004), we measure seat value based on *consumer perceptions*.

Our research also contributes to an evolving literature on consumer behavior and

empirical modeling in Revenue Management. Shugan and Xie (2000) show that advanced selling mechanisms can be used effectively to improve firm profits as long as (a) consumers have to purchase a product ahead of their consumption, and (b) their post-consumption valuation is uncertain. Xie and Shugan (2001) provide guidelines for when and how sellers should advance sell in markets with capacity constraints. Dana (1998) shows that advance-purchase discounts can be employed effectively in competitive markets, if consumers' uncertain demand for a good is not resolved before the purchase of the good. Su (2007) finds that heterogeneity in consumer valuations, along with waiting time behavior, influences pricing policies of a monopolist. Gaur and Park (2007) consider consumer learning in competitive environments. While most of this literature is analytical, we take an empirical approach to analyze seat values as perceived by customers, and study its implications for revenue management decisions in the sports/theater business.

There has been recent interest in modeling Revenue Management decisions in non-traditional settings. Roels and Fridgeirsdottir (2009) consider a web publisher who can manage online display advertising revenues by selecting and delivering requests dynamically. Popescu and Rudi (2008) study revenue management in stadiums where experience is often dictated by the collective experience of others around a patron.

Methodologically, our approach is related to the literature employing ordinal models to study the antecedents and drivers of customer satisfaction. Kekre et al. (1995) study the drivers of customer satisfaction for software products by employing an ordinal probit model to analyze a survey of customer responses. Bradlow and Zaslavsky (1999) use a Bayesian ordinal model to analyze a customer satisfaction survey with 'no answer' responses. Rossi et al. (2001) propose a hierarchical

approach to model customer satisfaction survey data that overcomes reporting heterogeneity across consumers. We use an ordinal logit model similar to the aforementioned papers, taking into account heterogeneity in reporting (across customers) and heterogeneity in the distribution of seat values (across seat locations).

Anderson and Sullivan (1993) note that relatively few studies investigate the antecedents of satisfaction, though the issue of post-satisfaction behavior is treated extensively. They note that disconfirmation of expected valuation causes lower satisfaction and affects future consumption. While previous considerations about a product might affect how consumers value the experience, we mainly focus on how product attributes such as seat location, and personal attributes such as gender, age and frequency of visits affect customer valuations.

Homburg et al. (2005) show that customer satisfaction has a strong impact on willingness to pay. Ittner and Larcker (1998) provide empirical evidence that financial performance of a firm is positively associated with customer satisfaction and customer value perception. We use seat value measures reported by consumers in a survey to recommend changes that would help the firm (a baseball franchise in our context) achieve a chosen service objective on seat value. Hence we believe that this objective would improve customer goodwill, which in turn would lead to better long-run performance.

5.3 Objectives and Methodology

5.3.1 Research Issues

The focus of our research is to understand how the seat value perceived by a customer in a stadium/theater varies based on the location of her seat relative to the stage/field. Since we are interested in post-consumption seat value perceived by customers in attendance, we do not consider the underlying trade-offs made while arriving at the purchase and seat choice decisions. Therefore, we only model the ex-post *net valuations* realized by consumers, in order to understand how they differ based on seat location.

To derive sharper insights, we assume that consumers are forward-looking and have rational expectations, i.e. that they do not make systematic forecasting errors about what valuations they might receive from attending a game or seeing a show. The rational expectations assumption is widely employed in empirical research in economics (Muth 1961, Lucas and Sargent 1981, Hansen and Sargent 1991) and marketing literature (for example, Sun et al. 2003). Accordingly, we assume that every consumer has some belief on the distribution of possible valuations that she could realize, conditional on her covariates. Furthermore, the ex-ante distribution of valuations for a rational consumer is identical to the ex-post distribution of valuations realized by the consumer population with identical covariates. Note that rational expectations does not imply that consumers are perfectly informed about their true valuations.

5.3.2 Methodology

Seat Value

We define the value perceived by a consumer as the valuation realized from her event experience *net* of the price paid (consistent with Zeithaml 1988). We note that the exact valuation realized from the experience cannot be easily quantified, and therefore the value perceived is *latent*. However, the consumer would be able to translate her latent value perceived on some graded scale. In other words, although she cannot describe the *exact* worth of the show she attended, she can usually confirm if the value she perceived was low, medium or high. We define *Seat Value Index* (SVI) as an ordinal measure that captures the post-consumption latent value perceived by a consumer. Let V_i denote the SVI reported by a respondent i. It takes values in $\{1, 2, \ldots, J\}$, $J \in \mathbb{N}$, where $V_i = 1$ corresponds to the lowest SVI (low net value), and $V_i = J$ represents the highest SVI (high net value).

Service Objective

In many operational contexts, firms that seek to improve customer service adopt a service level measure such as *fill rate* or *in-stock* probability (Cachon and Terwiesch 2008). Such decisions are based on the belief that improving availability of products reduces the incidence of costs that might be associated with stock-outs, and the resultant loss of goodwill. For instance, firms would hope to set prices such that it keeps the fraction of customers experiencing low seat value to an acceptable level at each seat location. Such service level measures that focus on limiting the fraction of customers facing inferior service experience, is commonly applied in several industries. Call centers choose staffing level according to an 80/20 rule (or, some

variation thereof) that focuses on limiting the fraction of customers that face waiting times exceeding a certain threshold.

While a newsvendor can adjust quantities of goods produced based on the chosen service level objective, in many RM scenarios, the quantities are unchangeable (for example, the number of seats in a theater cannot be adjusted easily). In such cases, prices are the main lever by which RM firms can attain their service objective. However, in many revenue management scenarios, especially in stadiums/theaters, the value of the product is intrinsically linked to the experience. For example, it is possible that customers who experience low value might switch to other services, or balk from visiting again. Firms would hope to set prices such that the fraction of customers experiencing low seat value could be limited to acceptable levels. Such an objective would be consistent with the models of customer behavior linked to service/stockout experiences considered in previous Operations Management settings (For example, see Hall and Porteus 2000, Gans 2003, Gaur and Park 2007).

Several RM firms desire to limit the fraction of customers experiencing low seat value in order to mitigate the loss of goodwill or to reduce switching. Hence, we consider a service-level objective that aims to set prices to maximize revenues while keeping the probability of a customer reporting low SVI to a maximum threshold level, α_l, at some seat location l. For expositional ease, we shall assume that $\alpha_l = \alpha$ across all seat locations. This clearly need not be the typical case. A theater might be willing to impose more stringent constraints on certain sections of the arena compared to other sections. Therefore, under our service level objective for a particular seat category l, the firm would like to set some price p_l^* under the constraint

$$\Pr[SVI \leq j | p_l^*] \leq \alpha_l \tag{5.1}$$

The choice of α_l and j are flexible, and could be based on the long term objective of the firm.

We only consider static price adjustments in our setting, since such schemes are consistent with industry practice where we apply our model. It is very common that theaters and sports stadiums announce prices for the entire season; the number of price changes are extremely limited within the selling horizon. Utilizing the service level objective we elaborated, the firm can increase or decrease prices suitably to achieve a desired level of seat value.

Modeling SVI

We could treat SVI as continuous and estimate a multiple regression model using ordinary least squares (OLS). However, this approach is flawed as (1) the OLS estimates are inefficient and the predictions cannot be restricted to the interval $[1, J]$ (Kmenta 1986), and (2) the regression estimates will roughly correspond to the correct ordered model only if differences in value between two consecutive indices are identical. For additional discussion on the limitations of OLS regressions, see Judge et al. (1980).

Alternately, we could treat SVI as categorical and employ a multinomial logit or probit model. This overcomes the limitations of OLS regression, but is still inefficient as it throws away valuable information by ignoring the ordinal nature of SVI. Hence, the appropriate model for our purpose is an *ordinal regression* model that takes into account the categorical nature of the data as well as the ordered information contained.

A respondent i derives her SVI, $V_i \in \{1, 2, \ldots, J\}$, by categorizing her post-consumption

latent net value realized, V_i^*, into buckets defined by the thresholds $\{\tau_i^0, \tau_i^1, ..., \tau_i^J\}$, where it is understood that $\tau_i^0 = -\infty$ and $\tau_i^J = +\infty$. Hence, respondent i reports her SVI as $V_i = j$, if and only if $\tau_i^{j-1} < V_i^* \leq \tau_i^j$, for $j = 1, 2, \ldots, J$. This mapping between a respondent's experienced net value and reported SVI is illustrated in Figure 5.1 for the case $J = 3$.

The net value experienced by the customer is $V_i^* = \mathbf{x}_i^T \beta + \epsilon_i$, where $\mathbf{x}_i$ is a vector of consumer and seat characteristics (excluding a constant). β is the associated vector of parameters, and ϵ_i is a stochastic term that captures the idiosyncratic value derived from the experience. The model is completed by specifying the distribution of ϵ_i and constraints on the thresholds, τ_i^j (required for identifiability). The reader is directed to Liu and Agresti (2005) for a detailed overview and survey of ordinal data analysis. We provide a more detailed description of the ordinal regression models used, in Section 5.3.5

5.3.3 Description of Baseball Dataset

We now illustrate our research issue based on the data from a professional league baseball franchise (equivalent of Major League Baseball) in Japan. The franchise is located in a mid-small city, and hence could not rely on conventional streams of revenue such as broadcasting, merchandizing and advertising. The franchise management decided to focus on ticket sales as it saw an upside potential in considering improvements in pricing and seating layouts.

As the team was a recently established franchise, the management conducted a survey to better understand the traction for the team among its fans. The survey was designed by the team based on inputs from various departments and team
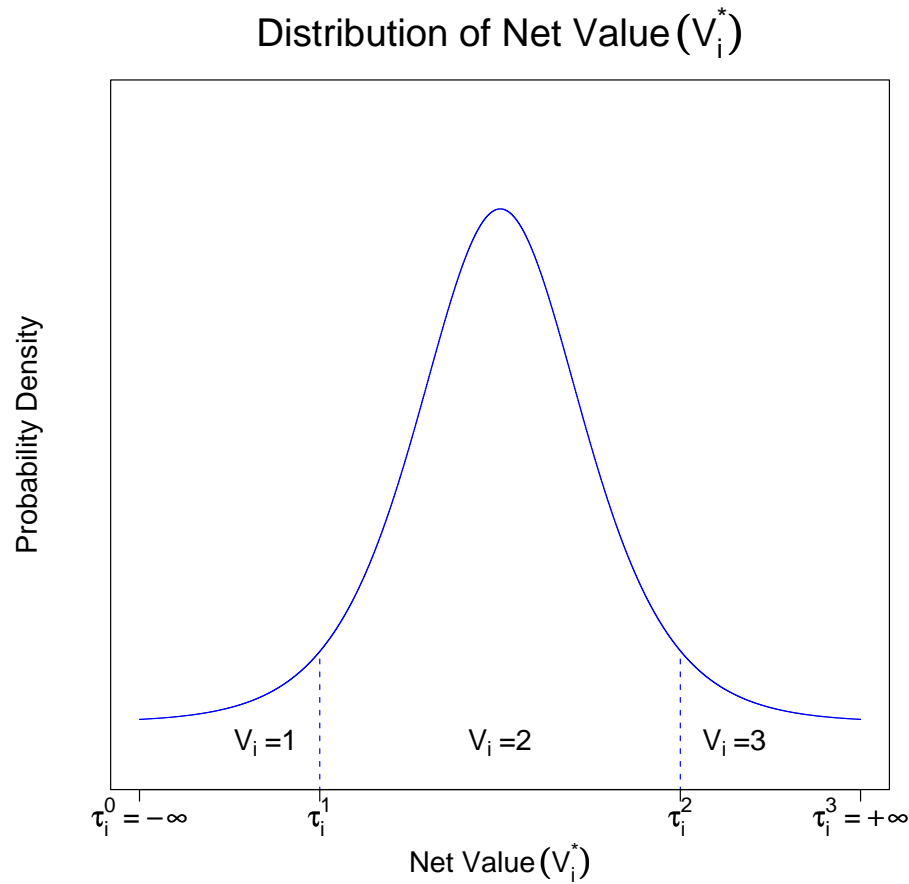
Figure 5.1 **Mapping between Net Value, V_i^* (latent), and Seat Value Index, V_i**

executives in the franchise. The survey was administered to a random sample of consumers at the franchise's stadium on a weeknight game. Only one response was obtained from each consumer.

In the survey, respondents were asked to report the net worth of the seats they sat in as Low, Medium or High. This corresponds to the Seat Value Index (SVI) measure which was defined before as a quantification of a respondent's realized net value. In addition, customers were asked to report their age[4], gender, hometown,

---

[4]We treat age as a continuous variable in our model in order to preserve the order information contained. We tested an alternate specification treating age as categorical, but rejected it in favor of the continuous specification based on the AIC values of the two models. In addition, we also

seat, frequency of visits to the stadium and preference for visiting teams. Table 5.1 provides more details on these variables and how we treat them in our models.

Table 5.1   Description of Variables in the Dataset

Variable	Values	Treatment
SVI	Low, Medium, High	Ordinal (1-3)
Age	$0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60+$	Continuous (1-7)
Gender	Male, Female	Categorical
Hometown	City, Prefecture, Outside	Categorical
Seat	$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$ (see Figure 5.2)	Categorical
Frequency	First Time, Once, Thrice, Five Times, All Games	Continuous (1-5)
Visiting Team	Team 1, Team 2, Team 3, Team 4, Team 5	Categorical

The experience and the resulting value perceived are highly dependent on the location of the seat from which a respondent watched the game. However, this information is not clearly captured by the explanatory variable *Seat*. For example, customers seated in locations 2 and 7 have almost identical views, but this linkage is not apparent in the current coding of the *Seat* variable. Hence, we represented each seat in terms of three location attributes given by **Side** = {1st Base, 3rd Base, Backnet, Field, Grass}, **InOut** = {Infield, Outfield} and **Deck** = {Upper, Lower}.

5.3.4   Preliminary Analysis

From a total of 1397 respondents, 259 responses were dropped due to missing information, resulting in $N = 1138$ responses.[5] A preliminary analysis revealed

considered an alternate continuous specification for age, where each age-group is represented by its mid-point. We find that our results remain largely unchanged.

[5] According to Bradlow and Zaslavsky 1999, there are two possible causes for respondents not reporting a satisfaction score like SVI: (1) the respondent does not consider the satisfaction score as salient, or (2) the respondent considers it salient, but has a mild opinion and hence does not voice it.
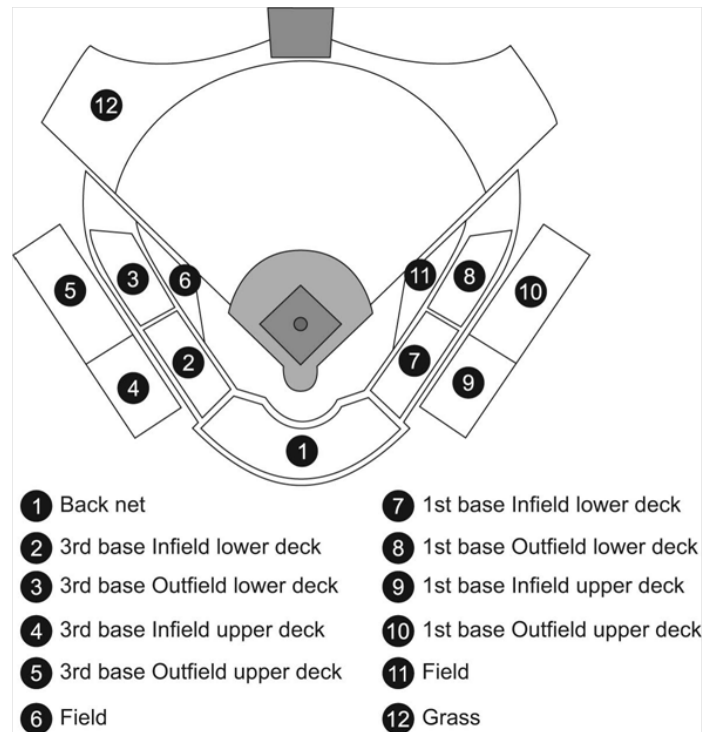
Figure 5.2   Stadium Seating Layout.

that the frequency distribution of SVIs was skewed towards the right, as shown in Figure 5.3. This implies that a higher proportion of consumers reported a low SVI, which underlines the further need for studying seat value.

Figure 5.3 also reveals some cursory insights. The seat value index reported by older respondents seems to be more homogeneous. Customers seated in Grass seats report higher SVI, while respondents seated at Backnet seem to have a lower SVI. Infield and Lower Deck seats seem to have a higher proportion of respondents reporting low SVI as compared to Outfield and Upper Deck seats. Finally, the season regulars attending all games seem to have more homogeneous SVIs as compared to the first-timers. We now discuss the regression methodology adopted and the

---

We analyzed the responses missing information using a series of auxillary regressions, but did not find any systematic patterns of non-response.

estimation of model parameters.
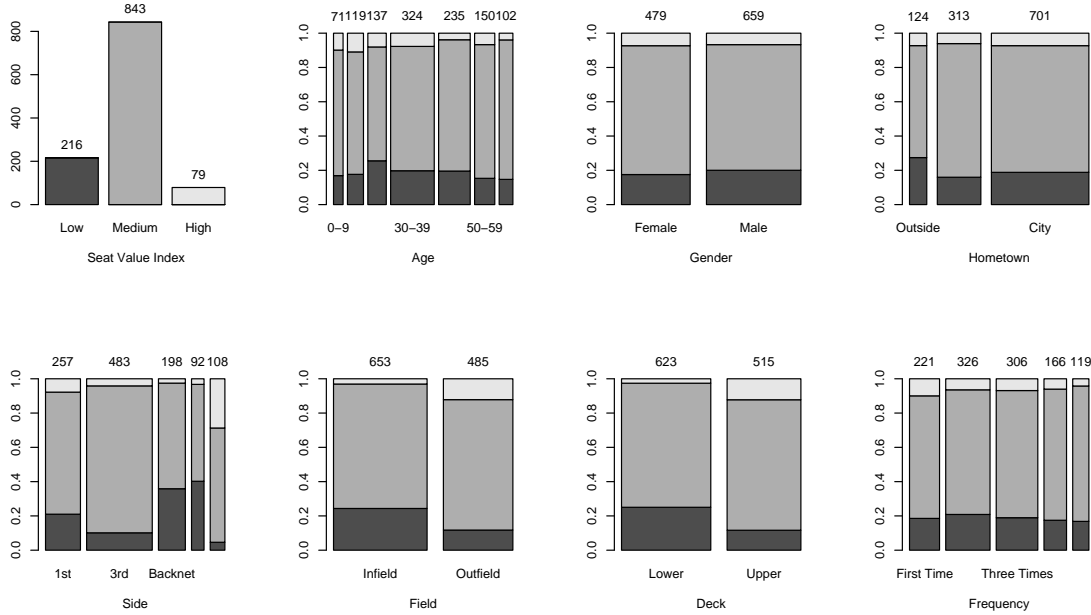
Figure 5.3   Distribution of Seat Value Indices.

5.3.5   Estimation of Parameters

Let V_i denote the SVI reported by respondent i, $i = 1, 2, \ldots, N$. Note that V_i can take the rank-ordered values $j = 1, 2, 3$ corresponding to *Low, Medium* and *High*, respectively.

Given that our response variable is ordinal, we follow McCullagh (1980) and use ordinal regression to model our data. The standard ordinal regression model assumes that all consumers use the same response thresholds while reporting their SVI (i.e. $\tau_i^j = \tau^j, \forall i$) and experience independent identically distributed idiosyncratic value (i.e. ϵ_i are iid). The usual choices for the distribution of ϵ_i are the normal or logistic

distributions, which would lead us to the standard ordinal probit model (McKelvey and Zavoina 1975), or the standard ordinal logit model (McCullagh 1980), respectively. McCullagh (1980) shows that the ordinal probit and logit models are qualitatively similar and that the fits are indistinguishable for any given data set; hence the selection of an appropriate distribution should be primarily based on ease of interpretation. We use the logistic error distribution as it allows us to interpret the regression coefficients in terms of log-odds.

Given these assumptions, the cumulative probability distribution of V_i can be written as

$$
\begin{aligned}
\Pr(V_i \leq j \mid \mathbf{x}_i) &= \Pr(V_i^* \leq \tau^j \mid \mathbf{x}_i) = \Pr(\mathbf{x}_i^T \beta + \epsilon_i \leq \tau^j \mid \mathbf{x}_i) \\
&= \Pr(\epsilon_i \leq \tau^j - \mathbf{x}_i^T \beta) = \frac{e^{(\tau^j - \mathbf{x}_i^T \beta)}}{1 + e^{(\tau^j - \mathbf{x}_i^T \beta)}} \\
&= \Lambda(\tau^j - \mathbf{x}_i^T \beta) \quad \forall j = 1, 2, \ldots, J-1,
\end{aligned}
\tag{5.2}
$$

where $\mathbf{x}_i$ is a vector of covariates consisting of *Age, Gender, Hometown, Side, InOut, Deck, Frequency* and *Team 1*, β is the associated vector of parameters, and Λ is the cumulative distribution function of the logistic distribution. Note that $\mathbf{x}_i^T \beta$ expands to

$\mathbf{x}_i^T \beta = \beta_1 \text{Age}_i + \beta_2 \text{Male}_i + \beta_3 \text{City}_i + \beta_4 \text{Prefecture}_i + \beta_5 \text{3rdBase}_i + \beta_6 \text{Backnet}_i + \beta_7 \text{Field}_i + \beta_8 \text{Grass}_i + \beta_9 \text{Outfield}_i + \beta_{10} \text{UpperDeck}_i + \beta_{11} \text{Frequency}_i + \beta_{12} \text{Team1}_i.$[6]

For the probabilities to be well-defined, the threshold parameters need to satisfy

---

[6]Note that the actual price paid may have an effect on consumer valuations and the ex-post survey scores reported. While our approach can incorporate price into the regression, our dataset lacks granular price data at the consumer level. Therefore, we do not explicitly consider price in our model. Instead, we study the effects of seat price on SVI and test the robustness of our model to price effects in Section 5.4.4. We find that our conclusions remain unchanged even when price dependencies are considered. We thank an anonymous reviewer for pointing out this aspect.

the condition

$$\tau^1 < \tau^2 < \ldots < \tau^{J-1}. \tag{5.3}$$

We can now express the log-likelihood function for the standard ordinal logit model as

$$LL(\beta, \tau \mid \mathbf{V}, \mathbf{X}) = \sum_{i=1}^{i=N} \sum_{j=1}^{J} \mathbf{I}(V_i = j) \log \left\{ \Lambda(\tau^j - \mathbf{x}_i^T \beta) - \Lambda(\tau^{j-1} - \mathbf{x}_i^T \beta) \right\}, \tag{5.4}$$

where $\mathbf{I}$ is the indicator function, $\mathbf{V} = \{V_i\}_1^N$ and $\mathbf{X} = \{\mathbf{x}_i\}_1^N$. The parameters of the model are estimated by maximizing the log-likelihood in Equation (5.4) subject to the constraints in condition (5.3).

Prior to running the regression model, we first tested for the usual symptoms of multi-collinearity (Greene 2003): (1) high standard errors, (2) incorrect sign or implausible magnitude of parameter estimates, and (3) sensitivity of estimates to marginal changes in data. We found no evidence of these symptoms in our dataset. We computed the Variance Inflation Factors (VIF) for every covariate and found all of them to be less than two (i.e. $\max(VIF) < 2$), which again suggests that multi-collinearity is not an issue. In addition, we added random perturbations to the independent variables and re-estimated the model (Belsley 1991). We determined the changes to the coefficients of those variables to be insignificant on repeated trials, thus further supporting that multicollinearity might not be a significant concern.

We use the *OLOGIT* routine in STATA 10.0 to estimate the parameters of the model using the maximum likelihood approach. The results are summarized in Table 5.2.

The standard ordinal logit model is equivalent to estimating $J - 1$ logistic regressions of the form $\Pr(V_i \leq j \mid \mathbf{x}_i) = \Lambda(\tau^j - \mathbf{x}_i^T \beta^j)$, with the assumption that the slope

coefficients are identical across all equations, i.e. $\beta^j = \beta, j = 1, 2, \ldots J - 1$. If we rewrite Equation (5.2) in terms of the odds of $\{V_i \leq j\}$, we get $\text{Odds}(V_i \leq j \mid \mathbf{x}_i) = \exp(\tau^j - \mathbf{x}_i^T \beta)$. Hence, for two different response levels j_1 and j_2, we find that the ratio of odds, given by

$$\frac{\text{Odds}(V_i \leq j_1 \mid \mathbf{x}_i)}{\text{Odds}(V_i \leq j_2 \mid \mathbf{x}_i)} = \exp(\tau^{j_1} - \tau^{j_2}), \tag{5.5}$$

is independent of the covariate $\mathbf{x}_i$.

Equation (5.5) is often referred to as the Proportional Odds property and implies that all respondents have the same ratio of odds of reporting a low SVI to odds of not reporting a high SVI. While it might be reasonable to assume that customers sitting in different seats might inherently have the same propensity to find higher (or lower) value, one would expect that customers 'learn' their valuation through repeated visits to the stadium/theater, and hence would have different odds ratios based on the number of prior visits. Hence, we need to investigate the validity of the implicit proportional odds assumption made by the standard ordinal logit model, before using it to make any inferences.

The standard approach to test if the proportional odds property holds, is to use a Likelihood Ratio Test (LRT) to compare the standard ordinal logit model (SOLM) with an expanded ordinal logit model (EOLM), that allows the slope coefficients (β) to depend on the threshold levels. The null hypothesis being tested is H_0 : $\beta^j = \beta, j = 1, 2, \ldots, J - 1$. The test statistic $-2\{\ln(SOLM) - \ln(EOLM)\}$ has a $\chi^2_{(k)}$ distribution, where k is the number of additional parameters in the expanded model.

The Likelihood Ratio Test is an omnibus test that the slope coefficients (β) are equal

119

across threshold levels for all the explanatory variables in the regression. This test only indicates whether or not the slope coefficients are different across threshold levels for some explanatory variables, but does not allow us to identify them specifically. In order to pinpoint the explanatory factors that drive this deviation from proportional-odds, we need a more detailed procedure, and we achieve this by employing a Wald test developed by Brant (Brant 1990).

The Brant test validates the proportional odds property for each covariate individually. The main idea of this test is to fit separate logistic regressions, $\Pr(V_i \leq j \mid \mathbf{x}_i) = \Lambda(\tau^j - \mathbf{x}_i^T \beta^j)$, for each of the $J-1$ threshold levels. These unconstrained estimates generally will not correspond to the overall maximum likelihood estimates for the integrated model as they could violate the monotonicity property for thresholds in i (as expressed in Equation 5.3). If the threshold monotonicity property is violated, they lead to negative values for fitted probabilities. Hence the standard likelihood based procedures, such as the likelihood ratio test, cannot be conducted based on the 'separate' unconstrained regressions. Nevertheless, we could use the 'separate' logistic regression estimates and the asymptotic covariance matrix to construct a test statistic that checks for the equality of the β^js and thus identify those variables for which the proportional-odds property do not hold.

We applied the likelihood ratio test and found that the standard ordinal logit model is strongly rejected in favor of an expanded model that allows for the slope coefficients to differ across threshold levels ($\chi^2_{(12)} = 46.74, p < 0.0001$). Consequently, we conducted the Brant test, to find that the proportional-odds property is violated for the coefficients β_1 (Age), β_5 (Side) and β_{10} (Deck).[7] To rule out the possibility

---

[7]A likelihood ratio test confirms that a partially constrained model that allows only for β_1, β_5 and β_{10} to depend on j cannot be rejected in favor of an unconstrained model that allows all the β's to depend on j ($\chi^2_{(9)} = 6.33, p = 0.71$).

of a misspecified link, we applied the Brant test to ordinal models with different link functions (probit, log-log and complementary log-log), but still found the same violations of the proportional-odds property.

The proportional-odds property in Equation (5.5) is a direct consequence of specifying that all consumers use the same response thresholds τ and have independent and identically distributed idiosyncratic value (ϵ). Violation of the proportional-odds property for a subset of covariates, detected by the Brant test, suggests that there is some inherent heterogeneity across consumers and seat locations. While ignoring such heterogeneity in exchange for a simple parsimonious model might be acceptable for some applications, in several revenue management settings, the inherent heterogeneity in the consumer population could be the key driver for pricing strategies.

Hence, we consider two different modifications to the standard ordinal logit model that account for this heterogeneity.

1. First, we consider a generalized threshold model that addresses the possibility of customers using different thresholds in reporting their responses, by relaxing the assumption that the thresholds, τ_i^j, are identical for all respondents.

2. Second, we consider a heteroskedastic model that addresses the inherent differences in the distribution of net value across seat locations, by allowing the variance of the idiosyncratic value term, ϵ_i, to systematically vary across respondent groups.

We now discuss these two sources of heterogeneity and the modeling strategies that can account for them.

		Standard	Generalized		Heteroskedastic
Variables			**Ordinal Logit Models**		
		Standard	**Generalized**		**Heteroskedastic**
		$j = 1, 2$	$j = 1$	$j = 2$	$j = 1, 2$
Threshold: Low-Medium	τ^1	-1.215***	-0.761**		-0.748***
Threshold: Medium-High	τ^2	3.387***	2.071***		2.067***
Age	β_1^j	0.048	0.127**	-0.172**	0.034
Male	β_2^j	-0.019	-0.026		-0.034
City (vs. Outside)	β_3^j	0.083	0.029		0.011
Prefecture (vs. Outside)	β_4^j	0.192	0.166		0.102
3rd Base (vs. 1st Base)	β_5^j	0.428**	0.873***	-0.727**	0.145
Backnet (vs. 1st Base)	β_6^j	-0.730***	-0.678***	-0.678***	-0.440***
Field (vs. 1st Base)	β_7^j	-0.893***	-0.824***		-0.509***
Grass (vs. 1st Base)	β_8^j	1.816***	1.206***		0.919***
Outfield	β_9^j	0.215	0.211		0.171
Upper Deck	β_{10}^j	0.246	0.066	0.947***	0.263**
Frequency	β_{11}^j	-0.126**	-0.093	-0.234**	-0.081**
Team 1	β_{12}^j	0.249*	0.250*		0.185**
Age	γ_1				-0.075***
3rd Base (vs. 1st Base)	γ_5				-0.324***
Upper Deck	γ_{10}				0.208***
Frequency	γ_{11}				-0.057*
Log Likelihood	LL	-748.12	-727.18		-726.27
Likelihood Ratio χ^2	LR	149.02	190. 90		192.72
No. of Parameters		14	18		18
McFadden Pseudo R^2		9.06%	11.60%		11.71%

Table 5.2   Parameter Estimates for All Models

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Heterogeneity in Response Thresholds: Generalized Threshold Model

It is not uncommon for people to use different thresholds in reporting their ordinal

responses. The generalized threshold ordinal logit model retains the idea that

consumers realize their net value from a common distribution, $V_i^* \sim \Lambda(\mathbf{x}_i^T \beta, \frac{\pi^2}{3})$, but assumes that they use systematically different thresholds, τ_i^j, while reporting their net value. A common approach to model generalized thresholds is to make the threshold parameters linear (Maddala 1983, Peterson and Harrell 1990) or polynomial functions of the covariates.[8] We choose the linear specification and accordingly let $\tau_i^j = \tilde{\tau}^j + \mathbf{x}_i^T \delta^j$, where $\mathbf{x}_i$ is the set of covariates and $\delta^j, j = 1, 2$, are vectors of the associated parameters that capture the effect of the covariates in shifting the thresholds. Substituting the expression for τ_i^j in place of τ^j in Equation (5.2), we can write the equations for the generalized ordinal logit model

$$\Pr(V_i \leq j \mid \mathbf{x}_i) = \Lambda(\tilde{\tau}^j - \mathbf{x}_i^T \beta^j), \; \beta^j = \beta - \delta^j \; \forall j = 1, 2. \tag{5.6}$$

According to the generalized threshold ordinal logit model, the net effect of any covariate k, β_k^j on SVI, is a combination of two effects (a) the real effect (β_k) and (b) the threshold-shifting effect (δ_k^j). It is the threshold-shifting effect (δ_k^j) that leads to the manifestation of unequal slopes detected by the Brant test. Thus, two groups of customers might have identical distributions of net value, but the distributions of their reported SVIs might differ because of different reporting thresholds. Figure 5.4 illustrates this case for two customers, *A* and *B*, seated at identical locations.

From the results of the Brant test, we infer that the covariates *Age*, *3rd Base* and

---

[8]For example, despite having the same level of 'true' health, older people may report their health differently from younger people. This phenomenon of subgroups of population using systematically different thresholds when assessing some latent quantity is referred to as *Response Category Threshold Shift* or *Reporting Heterogeneity*. It is also possible that some respondents are biased and answer questions on latent factors (such as the value of a seat) by comparing themselves with a reference group or a situation, that may be unobservable to the researcher (*Scale of Reference Bias* Groot 2000). In addition, respondents could display systematic biases in using different portions of the scale, e.g. the lower and upper ends. For instance, some discerning consumers attending a play might be quite strict on reporting 'high' responses (hard to please critics). This is referred to as *Scale-Usage Heterogeneity* (Rossi et al. 2001).
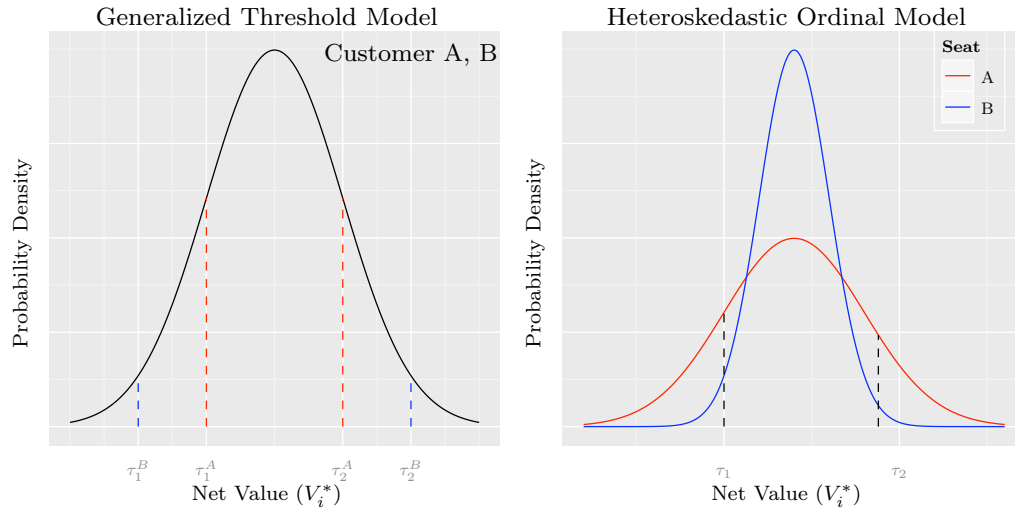
Figure 5.4 Generalized Threshold Model vs. Heteroskedastic Ordinal Model

Upper Deck could be driving the shift in thresholds. In addition, we believe that repeated visits help respondents learn the true value of the game experience and would induce them to use different thresholds. Accordingly, we let the thresholds depend on the subset of covariates $\mathbf{z}_i = \{$Age, 3rd Base, Upper Deck, Frequency$\}$, and set $\delta_k^j = 0, j = 1, 2$ for $k \notin \mathbf{z}_i$.

We estimate the parameters of this generalized threshold model using the *GOLOGIT2* routine (Williams 2006a) in STATA 10.0. The results are summarized in Table 5.2. We observe that in addition to *Side* and *Frequency*, *Age* also becomes a significant predictor now. A standard measure of fit for ordinal regression models is the McFadden pseudo-R^2 which is defined as $1 - \frac{LL_{Model}}{LL_{Null}}$, where LL_{Model} refers to the model log-likelihood. It indicates the improvement in likelihood due to the explanatory variables over the intercepts-only (null) model. We find the pseudo-R^2 for the generalized threshold model to be 11.60%.[9]

---

[9]This value needs to be interpreted with caution as it is not directly comparable to the R^2 obtained in OLS, which is a measure of the proportion of variance in the responses explained by the predictors. In fact, it is possible to obtain low values for the pseudo-R^2, even when the explanatory power of

Heterogeneity in Net Value Distribution: Hetetoskedastic Ordinal Logit (McCullagh and Nelder 1989)

In the previous subsection, we considered customers using different thresholds to report different levels for the same realized experience. However, it is also possible that the distribution of values, ϵ_i, realized by different consumer groups might, themselves, be different. Consumers seated in different locations could have different variabilities in their experience depending on their seat location. Such occurrences are very likely in several Revenue Management settings. It is likely that consumers seated in some sections such as dress circles may have smaller differences in the value experienced than those consumers seated at farther sections of the same theater. Therefore, we believe that it is important for firms to account for such systematic differences in the variance of the distribution of idiosyncratic value, to obtain meaningful parameter estimates.[10]

We capture the dependence of the error variance on the covariates using a skedastic function $h(.)$ that scales the iid ϵ_is in the standard ordinal logit model. Mathematically, we write $V_i^* = \mathbf{x}_i^T \beta + h(\mathbf{z}_i)\epsilon_i$, where $\mathbf{z}_i$ is the vector of covariates upon which the residual variance depends. Following Harvey (1976), we parametrize $h(.)$ as an exponential skedastic function given by $h(\mathbf{z}_i) = \exp(\mathbf{z}_i^T \gamma)$. We can now rewrite Equation (5.2) to obtain the defining set of equations for the heteroskedastic ordinal

---

the model is good (Hauser 1978). Hence we analyzed more detailed fit statistics in Section 5.4.1 to support the predictive power of the model. When we compared the actual number of respondents at a given seat location reporting a particular SVI, with those predicted by the model, we observed a high degree of correlation. This suggested that the model provides a pretty good fit.

[10] Ignoring systematic differences in variances across seat locations might lead to incorrect conclusions in some cases. For instance, consider two identical groups of consumers in a theater, who are seated at locations A and B, who have the same mean idiosyncratic value, but group A has twice the variance realized by group B, i.e., $\beta_A = \beta_B$, but $\sigma_A = 2\sigma_B$. This case is illustrated in Figure 5.4. Now, if we assumed that variances are equal at both locations, it would lead us to the erroneous conclusion that $\hat{\beta}_A = 0.5\hat{\beta}_B$, where $\hat{\beta}_i$ is an estimate of the true parameter β_i. Hence, accounting for heteroskedasticity is critical.

logit model as

$$\Pr(V_i \leq j \mid \mathbf{x}_i) = \Lambda \left(\frac{\tau^j - \mathbf{x}_i^T \beta}{\exp(\mathbf{z}_i^T \gamma)} \right) \ \forall j = 1, 2. \tag{5.7}$$

The heteroskedastic ordinal logit model belongs to a larger class of models known as location-scale models, and the reader is directed to McCullagh and Nelder (1989) for more details.[11]

Since the explanatory variables *Age*, *3rd Base* and *Upper Deck* violated the Brant test, we include these covariates in the expression for variance of idiosyncratic value. In addition, we also include the covariate *Frequency* in the variance expression, as we believe that repeated visits should help respondents learn the "true value" of the game experience, and consequently reduce the residual variation in their net value perceived. We estimate the parameters of the heteroskedastic ordinal logit model using the *OGLM* routine (Williams 2006b) in STATA 10.0.

From the results summarized in Table 5.2, we observe that the covariates *Frequency*, *Side* (except *3rd Base*) and *Upper Deck* have significant β coefficients. All the γ coefficients included in the variance equation are significant. We can draw several interesting inferences from these results.

Controlling for heteroskedasticity, we find that respondents at the third base have the same average net value as respondents at the first base, as $\hat{\beta}_5$ is not significant. However, the respondents seated on the third base side have *significantly less* variance in the net value realized (standard deviation is 1-$\exp(\hat{\gamma}_5)$ = 28% lower) as compared to those seated on the first base side. This could be due to the location of the home team dugout and/or the relative incidence of foul balls/home runs on the

---

[11]Note that the heteroskedastic ordinal logit model does not display proportional odds for the covariates in $\mathbf{z}_i$. This can be seen by writing out the expression for log-odds of $V_i \leq j$ conditional on $\mathbf{x}_i$, and observing that the effect of the covariates $\mathbf{z}_i$ on the log-odds is now dependent on the threshold level j: $\log(\mathrm{Odds}(V_i \leq j \mid \mathbf{x}_i)) = \frac{\tau^j - \mathbf{x}_i^T \beta}{\exp(\mathbf{z}_i^T \gamma)}$.

left field. Figure 5.4 details a comparison of reported SVIs for a customer located on the first base side and the third base side.

We find that the net value experienced by respondents seated at the upper deck has a higher mean ($\hat{\beta}_{10} = 0.263, p = 0.04$), as well as a higher variance ($\hat{\gamma}_{10} = 0.208, p = 0.0408$), when compared to the net value experienced by respondents seated at the lower deck. The net value experienced by customers visiting more frequently has a lower mean ($\hat{\beta}_{11} = -0.081, p = 0.028$) and a lower variance ($\hat{\gamma}_{11} = -0.058, p = 0.074$). Age of a respondent does not affect the mean of net value experienced, but older respondents tend to have lower variance in the net value experienced.

The current dataset has only one response for each consumer. Hence, it is not possible to econometrically distinguish between the Generalized Threshold Model and the Heteroskedastic Model. The observed deviation from proportional-odds could be a manifestation of consumers using different thresholds, or of the value distribution being heteroskedastic across seat locations. Hence, the applicability of either model must depend on the appropriate interpretation. For example, it is more likely that heterogeneity across consumers is explained by thresholds, while heterogeneity across seat locations is better explained by differences in the idiosyncratic value distribution. We interpret our results accordingly.

5.3.6 Achieving the Service Objective

Let us now consider the aforementioned service-level objective that we discussed before, where the firm aims to set prices such that the probability of a customer reporting low SVI is limited to a maximum threshold level, α, at all seat locations l.

127

In Lemma 1, we derive an expression for the price change at each seat location that would help the firm achieve this objective, using the heteroskedastic ordinal logit model specification.

Lemma 1 *Let* $\mathbf{x}_l$ *denote the vector of covariates for a customer seated at location l. Let* α, β, γ *and* $\mathbf{z}_l$ *be defined as in the heteroskedastic ordinal logit model, and* θ *denote the price elasticity of* V_l^*. *To limit the probability of this customer reporting SVI=1 at seat location l to a threshold* α, *the required price change* Δp_l *is given by*

$$\Delta p_l = \frac{1}{\theta} \left\{ -\tau^1 + \mathbf{x}_l^T \beta + \Lambda^{-1}(\alpha) \exp(\mathbf{z}_l^T \gamma) \right\} \tag{5.8}$$

Proof : At current prices, the probability of a typical customer reporting SVI as low is given by

$$\Pr(V_l^* \leq \tau^1) = \Lambda \left(\frac{\tau^1 - \mathbf{x}_l^T \beta}{\exp(\mathbf{z}_l^T \gamma)} \right) \tag{5.9}$$

Increasing the ticket price for seat location l by Δp_l would change this probability to

$$\Pr(V_l^* - \theta \Delta p_l \leq \tau^1) = \Lambda \left(\frac{\tau^1 + \theta \Delta p_l - \mathbf{x}_l^T \beta}{\exp(\mathbf{z}_l^T \gamma)} \right).$$

Equating this to α, we can calculate the desired price change Δp_l shown in Equation (5.8).

We apply the results of this lemma in Section §5.3.8 to derive price changes for a baseball franchise. Note that we could allow the service-level thresholds to differ across seat locations by specifying different αs.

5.3.7 Calculating Marginal Probabilities

The main purpose of our model is to predict the probability that a consumer seated at a particular seat location reports a certain SVI. In order to manage SVI, it is crucial to understand how these probabilities of a consumer reporting a certain SVI change with seat location and other covariates. Regression coefficients only explain the mean effects. In contrast, marginal probabilities measure how a change in a covariate impacts the distribution of the response variable.[12] Hence we calculated the marginal probabilities of the impact of different covariates on SVI. While measuring the marginal probability effects of any covariate, we define a *typical customer* for every covariate by fixing the rest of the covariates at their mean (or their mode for categorical covariates).

We use the MFX2 routine in STATA 10.0 to estimate the marginal probability effects and the results are summarized in Table 5.3, and interpreted in Section §5.5. Note that both the generalized threshold and heteroskedastic models provide comparable marginal probability estimates. Therefore, irrespective of the non-proportional-odds model considered, we obtain the same qualitative insights. As indicated before, we employ the threshold interpretation for consumer attributes (such as age, gender, frequency of visit, etc.), and the heterogeneity interpretation for all seat attributes.

In addition to the calculation of marginal probabilities for a typical customer, we also calculate the marginal probabilities for different customer segments (Age, Geography, Frequency of Visits). We discuss these results and their implications for *segment-specific* pricing in Section §5.5.1.

---

[12]If we let x_{il} denote the value of the lth covariate for respondent i, then the marginal probability effect is given by $\frac{\partial \Pr(V_i=j|\mathbf{x}_i)}{\partial x_{il}}$ for a continuous covariate and $\Delta \Pr(V_i = j \mid \mathbf{x}_i)$ for a categorical covariate.

Table 5.3   **Marginal Probability Effects** [a]

SVI	Variable	Standard	Generalized	Heteroskedastic
	Age	-0.006	-0.017**	-0.024***
	3rd Base	-0.056**	-0.114***	-0.102***
	Backnet	0.114***	0.107**	0.116***
Low	Field	0.151***	0.139**	0.143***
	Grass	-0.149***	-0.118***	-0.136***
	Upper Deck	-0.033	-0.009	-0.013
	Frequency	0.017**	0.013	0.006
	Team 1	-0.034*	-0.035*	-0.043**
	Age	0.004	0.024***	0.030
	3rd Base	0.035**	0.143***	0.131***
	Backnet	-0.086**	-0.084**	-0.090**
Medium	Field	-0.120**	-0.114**	-0.120**
	Grass	-0.023	0.039**	0.011*
	Upper Deck	0.021	-0.032	-0.036
	Frequency	-0.011**	-0.003	0.008
	Team 1	0.022*	0.025	0.030**
	Age	0.002	-0.007**	-0.009**
	3rd Base	0.021**	-0.029**	-0.031**
	Backnet	-0.029***	-0.023***	-0.026***
High	Field	-0.031***	-0.025***	-0.027***
	Grass	0.172***	0.079***	0.120***
	Upper Deck	0.012	0.041***	0.048***
	Frequency	-0.006**	-0.010**	-0.013**
	Team 1	0.012*	0.010*	0.013**

[a]*Gender, Hometown and InOut did not have significant effects.*

* p<0.1, ** p<0.05, *** p<0.01

5.3.8 Calculating Price Adjustments to Achieve the Service Objective.

Suppose that the franchise wants to keep the probability of a specific customer reporting SVI = Low to a threshold α at all seats. The current probability of a specific customer seated at location l reporting SVI = Low, can be calculated using Equation (5.9). We can then use Equation (5.8) to calculate the price change required at each seat location, that would equate the probability of this customer reporting SVI = Low, to the threshold value α. The parameters (β, γ) are known from the regression estimates, while the price elasticity of SVI (θ) can be estimated using the price variation observed across seat locations.

We now illustrate this calculation for a typical customer of the franchise (Age=4.22, Gender=Male, Hometown=City, Frequency=2.68) and a threshold of $\alpha = 15\%$. Table 5.4 summarizes the current service levels and the price changes (Δp_l) that achieve the threshold service level of $\alpha = 15\%$ for a typical customer. Note that the franchise might be interested in achieving this service objective for different consumer segments. We discuss this in Section 5.5.1.

5.4 Model Validation and Analysis of Effects

In this section, we validate our empirical results using various robustness checks. Specifically, we study game related effects with an additional dataset and the effects of price on seat value. In addition, we compare the effect of seat specific attributes (such as seat location) vs. customer specific attributes (such as age) on SVI.

Table 5.4 Price Increase to Keep $\Pr(SVI = Low)$ **to** $\alpha = 15\%$ **for a Typical Consumer**

Seat	Location	Section	Row	$\Pr(V_l^* \leq \tau^1)$	$\theta \Delta p_l$	Δp_l **(Yen)**
1	Backnet	Infield	Lower	34.7%	-0.689	-718
2	3rd	Infield	Lower	10.3%	0.197	205
3	3rd	Outfield	Lower			
4	3rd	Infield	Upper	9.7%	0.278	290
5	3rd	Outfield	Upper			
6	1st	Infield	Lower	20.8%	-0.249	-259
7	1st	Outfield	Lower			
8	1st	Infield	Upper	19.3%	-0.236	-246
9	1st	Outfield	Upper			
10	Field	Infield	Lower	37.2%	-0.758	-790
11	Grass	Outfield	Upper	6.8%	0.683	711

Frequency of Visits						
Additional Visit				20.8%	-0.269	-280

5.4.1 Model Validation

The standard approach to validate regression models is to estimate the model parameters on a calibration sample and validate those results on a hold-out sample. Accordingly, we constructed a calibration sample and a validation sample by randomly splitting our data-set into two equal parts. We measured the predictive accuracy of our model using an R^2 measure (see Equation 5.10), and find that $R_H^2 = 57.1\%$, which implies that the model significantly improves prediction accuracy over a naive model. Figure 5.5 shows a comparison of the actual number of respondents at each seat location reporting a particular SVI, with the expected numbers predicted by the model for the hold-out sample. These predictions generally match the distribution of the SVI for various seat locations.

While the R^2 is an indirect measure of predictive accuracy computed at a highly

Figure 5.5   Number of Respondents by Seat Location and SVI: Actual vs. Predicted

disaggregated level, a more direct measure is the accuracy of the predicted service level, $\Pr(SVI = 1)$. Computing the predicted service level for the hold-out sample, we find that while the actual service level is 19.6%, our model predicts a service level of 19.2%, thereby providing further confidence on the predictive power of our model.

We now briefly describe how we calculate our R^2 measure of predictive accuracy.

Calculating R^2: To calculate a measure of predictive accuracy, we ran the heteroskedastic ordinal logit model (M) on the calibration sample (C) to obtain estimates of the parameters β, γ and $\tau^j, j = 1, 2$. We then computed the expected number of respondents reporting SVI= $j \in 1, 2, 3$ at each seat location l, for the hold-out sample (H), using the following expressions.

$$E_{lH}^M[SVI = j] = \sum_{i \in H, Location=l} \Pr(SVI_i = j | \mathbf{x_i}).$$

A naive model (N) would estimate the expected number of respondents reporting SVI= j at seat location l as

$$E_{lH}^N[SVI = j] = \frac{n_{lH}}{|C|} \sum_{i \in C} \mathbf{I}(SVI_i = j),$$

where n_{lH} is the number of respondents in the hold-out sample, seated at location l.

If we let n_{lH}^j be the number of respondents in the hold-out sample seated at location l reporting SVI= j, then we can calculate the squared error of predicting n_{lH}^j using

the ordinal logit model (M) as

$$\epsilon_H^M = \sum_{l \in AllLocations} \sum_{j \in 1,2,3} \left(n_{lH}^j - E_{lH}^M[SVI = j] \right)^2.$$

We can compute an R^2 measure of predictive accuracy by comparing the ratio of ϵ_H^M to the squared errors of the naive model

$$\epsilon_H^N = \sum_{l \in AllLocations} \sum_{j \in 1,2,3} \left(n_{lH}^j - E_{lH}^N[SVI = j] \right)^2,$$

which gives us

$$R_H^2 = 1 - \epsilon_H^M / \epsilon_H^N. \tag{5.10}$$

5.4.2 Game Effects

Clearly, SVI is influenced by the actual game/event and hence it is important to consider the robustness of our results to variations across games. For instance, the outcome of the game, the composition of the playing teams, or the weather could have affected the seat value distribution customers reported. However, this limitation could be easily overcome by surveying consumers from multiple games and employing the same methodology to analyze the collected data and explore specific recommendations.

While the ideal way to test this would be to conduct the same survey across multiple games, record key game related attributes (result, attendance, visiting team, etc.) and use them as control variables in the regression equation, for reasons beyond our control, the franchise chose to vary some aspects of the survey across multiple games. For instance, a survey conducted during a different game included many of

the same questions as before (Age, Gender and Seat Location), but did not capture a few variables like Hometown and Frequency. We decided to combine the data from these two surveys to check the robustness of our results, especially the relationship between SVI and Seat Location, to inter-game variations.

We modify our regression equation for the HOLM by including only the common covariates across the two surveys and adding a fixed effects parameter to control for difference in valuations across games. The modified regression equation can be written as

$$
\begin{aligned}
V_i^* = {} & \beta_1 \text{Age}_i + \beta_2 \text{Male}_i + \beta_5 \text{3rdBase}_i + \beta_6 \text{Backnet}_i + \beta_7 \text{Field}_i + \beta_8 \text{Grass}_i + \\
& \beta_9 \text{Outfield}_i + \beta_{10} \text{UpperDeck}_i + \beta_{13} \text{Game}_i + \sigma_i \epsilon_i,
\end{aligned}
$$

where ϵ_i is a standard logistic random variable, and σ_i is a heteroskedastic variance scaling factor given by

$$
\sigma_i = \exp(\gamma_1 \text{Age}_i + \gamma_5 \text{3rdBase}_i + \gamma_{10} \text{UpperDeck}_i + \gamma_{13} \text{Game}_i).
$$

Note that the parameter γ_{13} captures differences in the variance of the distribution of seat values across the games. Table 5.5 shows a comparison of the parameter estimates obtained using the combined dataset with those obtained from the single game.

Note that all our verifiable conclusions hold even after we control for variations across games. Customers seated on the 3rd Base continue to experience lower variance in the seat value perceived ($\gamma_5 = -0.319, p < 0.001$), while the means show no statistically significant differences. Similarly, customers seated on the Upper Deck continue to have higher mean valuations ($\beta_{10} = 0.303, p < 0.001$) as

135

well as higher variance ($\gamma_{10} = 0.140, p < 0.05$). This suggests that our findings might be robust across games.

Table 5.5   Comparison of Parameter Estimates across Datasets

Variable		Single Game	Combined
Age	β_1	0.034	-0.007
Male	β_2	-0.034	-0.146
City (vs. Outside)	β_3	0.011	
Prefecture (vs. Outside)	β_4	0.102	
3rd Base (vs. 1st Base)	β_5	**0.145**	**0.080**
Backnet (vs. 1st Base)	β_6	-0.440***	-0.126
Field (vs. 1st Base)	β_7	-0.509***	-0.449**
Grass (vs. 1st Base)	β_8	0.919***	1.055***
Outfield	β_9	0.171	-0.153
Upper Deck	β_{10}	0.263**	0.303***
Frequency	β_{11}	-0.081**	
Team 1	β_{12}	0.185**	
Game	β_{13}		0.055
Age	γ_1	**-0.075*** **	**-0.062** **
3rd Base (vs. 1st Base)	γ_5	**-0.324*** **	**-0.319*** **
Upper Deck	γ_{10}	**0.208*** **	**0.140** **
Frequency	γ_{11}	-0.057*	
Game	γ_{13}		0.854***
Log Likelihood	LL	-726.27	-2043.87
Likelihood Ratio χ^2	LR	192.72	364.82
No. of Parameters		16	18
McFadden Pseudo R^2		11.71%	8.19%

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

It is interesting to note that while the mean valuations across games are not significantly different ($\beta_{13} = 0.055, p = 0.64$), the variances are significantly different

$(\gamma_{13} = 0.854, p < 0.001)$. In other words, the shape of the distribution of seat values is significantly influenced by the game. For instance, the first survey was conducted during a game that the home team lost, while the second survey was conducted during a game that the home team won. The result of the game could explain a portion of the difference in variances. Nevertheless, even after controlling for differences across the games, our seat value results remain largely unchanged.

Repeating the validation analysis discussed in Section 5.4.1, we find that even for the combined dataset, the model significantly improves the predictive power over the naive model.

5.4.3   Seat Location Effects

The experience in such entertainment settings is clearly a function of the product (the game in our context), the consumer and her seat location. Hence it is important to investigate how much of SVI is accounted for by each of these factors (game attributes, consumer attributes and seat location attributes). We study the relative impact of each of these three factors in influencing SVI, by following a three-step approach:

1. First, we ran several heteroskedastic ordinal regressions using a combination of these three factors as explanatory variables, both on the original dataset as well as the combined dataset.

2. Second, we measured the ability of each of these models to predict the number of consumers reporting a particular SVI at each seat location, using the R^2 defined in Equation (5.10).

3. Third, we compared the computed R^2 across the different models to understand the contribution of each of the three factors in predicting SVI.

Applying this analysis to the original dataset, we find that while the model consisting of both consumer and seat-location factors had an R_H^2 of 57.1%, a *major portion of SVI is accounted for by seat location attributes* (with an R^2 of 56.1%), while consumer attributes have almost insignificant predictive power ($R^2 = 6\%$).

To investigate how seat location factors influence SVI once we control for game related attributes, we applied the same analysis to the combined dataset. We find that while all three factors combined together have an R_H^2 of 56.1%, *seat location attributes still account for a major portion of the SVI*, with an R_H^2 of 38.5%, even after controlling for game related factors (See Table 5.6).

Table 5.6 Predictive Accuracy of Different Models.

Explanatory Variable Attributes			Dataset R_H^2	
Consumer	Seat	Game	First	Combined
✓			6.0%	9.1%
	✓		56.1%	38.5%
		✓		17.3%
✓	✓		57.1%	51.1%
✓		✓		17.7%
	✓	✓		55.6%
✓	✓	✓		56.1%

The analysis summarized in Table 5.6 emphasizes that seat location factors explain a significant portion of SVI. Game and Consumer attributes do matter, but explain a smaller portion. This finding underscores the importance of seat location factors in influencing seat value. It also strengthens the case for the need for studies like

ours that shed more light on the drivers of seat value. Finally, firms have reasonable control over seat location factors, and hence can take advantage of these findings to manage SVI.

5.4.4   Seat Price Effects

In order to properly estimate the relationship between seat value and seat location, we need to further isolate the effect of the location-dependent price variable. We address this issue by studying the relationship between SVI and Seat Location, controlling for the price variable. To achieve this, we consider three versions of the Heteroskedastic Ordinal Logit Model.

1. The original model described in Section 5.3.5 that does not include ticket price.

2. A model that included the ticket price for each seat section in addition to all the other covariates.

3. A model that includes ticket price for each seat section, but excludes all the seat location attributes.

The motivating question behind this analysis is to determine the extent to which the introduction of ticket prices impact our results. From Table 5.7, we observe that seat location attributes continue to explain a significant portion of SVI even after controlling for ticket price, as can be seen by comparing the McFadden Pseudo-R^2 of Models (b) and (c). In fact, adding seat location attributes to Model (c), which uses only ticket price, increases the pseudo-R^2 from 5.5% to 12.4%. Finally, we find that most of our results and inferences made in Section 5.3.5 continue to hold.

1. The effect of Age on SVI remains almost unchanged, as seen by the β and γ coefficients in Models (a) and (b).

2. Frequency of Visits have almost the same effect on SVI as before. The estimates for both the mean effect and the variance effect remain almost unchanged.

Table 5.7   Parameter Estimates for Heteroskedastic Ordinal Model with and without Price

Variable		Heteroskedastic Ordinal Logit Models		
		Only Seats[a]	Seats + Price[b]	Only Price[c]
Threshold: Low-Medium	τ^1	-0.748***	-2.710***	-1.340***
Threshold: Medium-High	τ^2	2.067***	0.090	1.240***
Age	β_1^j	0.034	0.030	-0.020
Male	β_2^j	-0.034	-0.030	-0.020
City (vs. Outside)	β_3^j	0.011	0.009	0.010
Prefecture (vs. Outside)	β_4^j	0.102	-0.101	-0.100
3rd Base (vs. 1st Base)	β_5^j	0.145	0.110	
Backnet (vs. 1st Base)	β_6^j	-0.440***	4.642***	
Field (vs. 1st Base)	β_7^j	-0.509***	-0.311	
Grass (vs. 1st Base)	β_8^j	0.919***	0.722***	
Outfield	β_9^j	0.171	-0.221	
Upper Deck	β_{10}^j	0.263**	-0.042	
Frequency	β_{11}^j	-0.081**	-0.071***	-0.053
Team 1	β_{12}^j	0.185**	0.184***	0.224**
Price	β_{13}^j		-0.01***	-0.001***
Age	γ_1	-0.075***	-0.080***	-0.071***
3rd Base (vs. 1st Base)	γ_5	-0.324***	-0.367***	
Upper Deck	γ_{10}	0.208***	0.223***	
Frequency	γ_{11}	-0.057*	-0.062***	-0.072***
McFadden Pseudo R^2		11.71%	12.4%	5.5%

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3. The asymmetry that we identified previously still holds, as can be seen from the γ coefficient for the 3rd Base. The mean effect still stays insignificant.

4. Consumers still find Grass seats very valuable, as seen from the β coefficient.

5. The mean effect of Backnet has changed significantly (β is now positive). This might be because price affects SVI non-linearly, or that Backnet customers are significantly different. The variance effect that we identified, on the other hand, remains almost unchanged.

6. The β parameter corresponding to the Upper Deck is no longer significant. However, the heterogeneity effects still persist. In fact, the parameter estimates show no significant change. ($\gamma_{10}^a = 0.208, \gamma_{10}^b = 0.223$)

The surveys did not ask consumers for the actual price that they paid, as the franchise felt that consumers might be more biased in their responses if price related information was asked. Hence, we only had seat prices at each section. The absence of variation in price across consumers seated at the same location renders any regression involving location and seat prices susceptible to the effects of multicollinearity. This also makes it difficult to isolate the effects of price from seat location. Hence, we study the impact of different prices paid by consumers, by adding a random noise term to perturb the ticket price specified for each seat section. Accordingly, the price paid by consumer i for a seat in section l was modeled as $p_{il} = p_l * (1 - \psi_i)$, where p_l is the ticket price specified for section l and ψ_i is the noise term distributed uniformly over $[0, m]$. Based on conversations with the franchise management on the range of discounts provided to consumers, we varied m from 5% to 20%. We repeated the analysis discussed above with these prices, and find that our results remain unchanged.

5.5 Pricing Recommendations and Insights

Based on robustness checks in Section §5.4, we are able to underline the importance of seat location in influencing consumer experience. Hence it is appropriate to consider seat-location specific prices for each consumer segment.

5.5.1 Segment Specific Pricing

In Sections 5.3.7 and 5.3.8, we discussed the calculation of marginal probability effects for a typical consumer and the price changes across seat locations required to achieve a service-level objective of $\alpha = 15\%$. However, the firm could engage in more targeted pricing schemes based on how the marginal probabilities varied across consumer segments. We now calculate the marginal probability effects for different seat locations for each consumer segment based on age groups (Table 5.8) and visiting frequencies (Table 5.9).

From the marginal probability tables, we infer that customers in the age group 40 - 49 years and 50 - 59 years tend to have the highest probabilities of reporting low SVI for the Backnet and Field seats, as compared to a similar seat on the 1st Base side. Hence, the franchise could offer reduced prices for these customers for the Backnet and Field seats.

We also infer that the regulars to the games have a much higher propensity to report a low SVI for the pricier Backnet and Field seats. Given that it is important for the franchise to manage the satisfaction levels of its most loyal customers, the franchise could offer discounts for multi-game tickets for selected stadium seats on the Backnet and Field, and set prices such that the dissatisfaction levels are below an appropriate threshold.

It is also interesting to note that for the Grass and 3rd Base seats, the first-timers are more likely to report a low SVI. Hence the franchise can encourage people to start watching games in the stadium by offering special discounts to newcomers, on the Grass and 3rd Base seats, or reserving a portion of these seats at lower prices for the first-timers.

Table 5.8   Marginal Probability of SVI = Low by Age

Age	3rd Base	Backnet	Field	Grass
0-9	-10.6%	12.4%	14.6%	-17.4%
10-19	-10.8%	13.0%	15.2%	-17.0%
20-29	-10.8%	13.5%	15.8%	-16.3%
30-39	-10.6%	13.8%	16.3%	-15.3%
40-49	-10.2%	14.0%	16.6%	-14.2%
50-59	-9.6%	14.0%	16.7%	-12.8%
60+	-8.7%	13.8%	16.6%	-11.2%

Table 5.9   Marginal Probability of SVI = Low by Frequency of Visits

Frequency	3rd Base	Backnet	Field	Grass
First Time	-10.0%	12.0%	14.2%	-13.6%
Once	-10.3%	13.1%	15.5%	-14.4%
Thrice	-10.7%	14.3%	16.9%	-15.4%
Five Times	-11.1%	15.6%	18.5%	-16.6%
All Games	-11.5%	17.2%	20.2%	-18.0%

The recommended segment-specific price changes for each seat section are summarized in Table 5.10 (for consumer segments based on age) and Table 5.11 (for consumer segments based on frequency of visits).

5.5.2   Actionable Recommendations

We now develop more concrete and actionable pricing recommendations that would help a franchise achieve a specified threshold service level on any given set of seat products that they might make available. For this analysis, we ignore substitution effects associated with the price changes. In Lemma 1, we derived an expression for the price change at each seat location that would help the firm achieve its service

Table 5.10   Price Change Percentage to set Pr(SVI=Low) to $\alpha = 15\%$

Age	3rd Base	Backnet	Field	Grass
0-9	-6.0%	-14.8%	-52.7%	25.1%
10-19	-1.0%	-13.0%	-46.6%	37.7%
20-29	3.8%	-11.3%	-40.9%	49.7%
30-39	8.3%	-9.7%	-35.5%	61.1%
40-49	12.7%	-8.1%	-30.3%	71.8%
50-59	16.8%	-6.7%	-25.4%	82.1%
60+	20.8%	-5.3%	-20.8%	91.8%

Table 5.11   Price Change Percentage to set Pr(SVI=Low) to $\alpha = 15\%$

Frequency	3rd Base	Backnet	Field	Grass
First Time	12.0%	-8.9%	-33.1%	66.1%
Once	10.5%	-9.1%	-33.8%	64.6%
Thrice	8.8%	-9.4%	-34.6%	62.9%
Five Times	6.9%	-9.7%	-35.6%	60.8%
All Games	5.0%	-10.0%	-36.7%	58.4%

level objective of keeping $\Pr(SVI_l \leq 1 \mid p_l^*)$ to a threshold α. If we let p_l denote the current seat price, then we can use Lemma 1 to calculate the new price p_l^* to be charged at each seat location as:

$$p_l^* = p_l + \frac{1}{\theta} \left\{ \mathbf{x}_l^T \beta - \tau^1 + \ln\left(\frac{\alpha}{1-\alpha}\right) \exp(\mathbf{z}_l^T \gamma) \right\}$$

This equation prices each seat location for a specific consumer whose characteristics are known. However, we can use this equation to price any set of seat products that a baseball firm could make available. For example, a firm is interested in setting a single price for each seat location such that the service level constraint is met. We can derive the new price p_l^* to charge consumers by taking a weighted average of the new prices derived using Lemma 1 over the distribution of consumer characteristics. Alternately, if the firm wants to provide targeted prices for specific consumer segments (e.g. Age, Frequency, Age-Frequency combination), then the new price to charge each segment can be derived by taking a weighted average of the new price over the distribution of the remaining consumer characteristics.

We now illustrate the application of this method in calculating seat prices. First, we calculate the location specific seat prices that the firm should set in order to achieve the service level objective for each seat location. The results are summarized in Table 5.12. From the seat location specific prices calculated in Table 5.12, we observe that the seat prices across 3rd base and 1st base are asymmetric. In fact, the seats located on the 3rd base command a 33% premium on average as compared to those on the 1st base. Moreover, as one would expect, seats on the lower deck continue to be priced higher than those on the upper deck.

Now, the firm can do better by setting targeted prices for specific consumer segments. For instance, the firm can target specific age groups such as students, regulars and retirees. The price to charge each group for a particular seat location are calculated as in Table 5.12. Note that, as expected, student tickets are heavily discounted across seat locations, while retirees are made to pay a premium[13].

In addition, the firm might also consider targeting consumers based on their frequency of visits by setting different prices for five game packs and season passes. From the seat prices based on frequency, as summarized in Table 5.12, we observe that season passes are discounted, compared to single game tickets. An interesting thing to note is that the maximum discount for season passes occurs for 3rd base tickets, which suggests that the firm stands to gain by exploiting the asymmetry in more than one way.

While we have illustrated price calculations for some specific instances of variable pricing, our method is general enough to accommodate more complex forms. For

---

[13]Based on the raw data, we observe that retirees on an average reported higher SVIs. Hence, our result of charging higher prices to retirees is consistent with our service level objective. In practice, there might be several other considerations that drive ticket pricing, which might make firms offer discounted tickets to retirees.

example, the firm might want to offer price bundles based on combination of age and frequency of visits. In this case, we can integrate the consumer specific seat prices across the distribution of remaining consumer characteristics (Gender, City, Team1 etc.) to derive the best price for each bundle that achieves a given service level.

Table 5.12   Seat Prices by Consumer Segment to Achieve Service Level Objective of $\alpha = 15\%$

				Price (Yen)					
Seat	Location	Section	Row	Base	Student	Regular	Retirees	5 Game	Seasons
1	Backnet	Infield	Lower	6980	6670	6930	7270	6940	6920
2	3rd	Infield	Lower	2420	2160	2360	2630	2380	2340
3	3rd	Outfield	Lower	1650	1470	1670	1940	1620	1580/
4	3rd	Infield	Upper	1740	1480	1710	2030	1700	1670
5	3rd	Outfield	Upper	1850	1660	1890	2200	1820	1790
6	1st	Infield	Lower	1980	1660	1920	2260	1950	1920
7	1st	Outfield	Lower	1260	970	1230	1570	1230	1200
8	1st	Infield	Upper	1230	870	1170	1570	1220	1200
9	1st	Outfield	Upper	1270	1030	1330	1720	1270	1260
10	Field	Infield	Lower	1490	1190	1450	1790	1450	1430
11	Grass	Outfield	Upper	2000	1670	1970	2350	2000	1980

The ideal way to test the impact of our recommendations would have been to offer the new prices to consumers and observe the resulting distribution of SVIs. However, that approach was not feasible, in our case, as it required the franchise to implement price changes across the board, and conduct the survey post implementation. Hence, we used the demographic profile of consumers in our validation sample to calculate the achieved service levels, assuming that consumers had paid these set prices. From Table 5.13, we clearly observe that the new prices achieve a

service level very close to the threshold of $\alpha = 15\%$ that we set out to achieve.

Table 5.13 Service Levels Achieved in the Validation Sample based on New Prices

				Pr(SVI = 1 | Validation Sample)					
Seat	Location	Section	Row	Base	Student	Regular	Retirees	5 Game	Seasons
1	Backnet	Infield	Lower	0.143	0.146	0.150	0.163	0.152	0.127
2	3rd	Infield	Lower	0.162	0.162	0.155	0.155	0.161	0.141
3	3rd	Outfield	Lower	0.150	0.150	0.144	0.146	0.131	0.127
4	3rd	Infield	Upper	0.140	0.137	0.155	0.166	0.152	0.115
5	3rd	Outfield	Upper	0.151	0.144	0.143	0.121	0.129	0.124
7	1st	Infield	Lower	0.151	0.151	0.155	0.144	0.161	0.132
8	1st	Outfield	Lower	0.180	0.169	0.167		0.167	0.112
9	1st	Infield	Upper	0.170	0.163	0.157	0.156	0.165	0.162
10	1st	Outfield	Upper	0.160	0.147	0.154		0.145	0.119
11	Field	Infield	Lower	0.141	0.149	0.151	0.155	0.151	0.126
12	Grass	Outfield	Upper	0.148		0.144	0.172	0.148	

5.5.3 Insights

Based on the results obtained, we gather several interesting insights on the net value perceived by consumers who attended the game. Our results help quantify seat value in terms of seat location characteristics and consumer attributes. Furthermore, we also characterize the distribution of SVIs that helps us determine the probability of customers reporting low SVI. We make several recommendations based on our empirical results and the service objective considered, and these are being implemented by the franchise.

1. **Seats are Asymmetric:** We find that consumers seated on opposite sides of the ball park report *asymmetric* SVIs. Thus, the distribution of SVIs reported by customers

seated on the third base side is significantly different from that of customers seated on the first base side. In fact, *any* customer located on the third base side has a *lower* probability of reporting low Seat Value Index as compared to an identical customer seated in a symmetric location on the first base side. This asymmetry is intriguing. Although every professional baseball team prices its tickets identically for left field and right fields, there are several underlying asymmetries in the game/ballpark that could possibly explain this difference in perceived value. First, the incidence of foul balls is generally higher in right field, which could influence how customers respond to their experience of the game. Second, for the stadium of the franchise we study, the location of the home-team dugout was on the third base side, which possibly provided higher value for some of the fans. Third, weather related factors like sunlight, wind, etc. can affect the viewing experience across seat locations. Finally, in many professional ball parks, although the prices are always symmetric, the views from the seats are not. In fact, to many players and baseball fans, the fundamental asymmetries in the design of a ballpark add to the idiosyncratic charm of the game (Maske 1992).

Asymmetric seat values provide the franchise with an opportunity to price tickets differently while maintaining identical probabilities of experiencing low seat value on both sides of the stadium. Our recommendations would initiate differential pricing across symmetric locations and achieve two goals. First, they eliminate the *inherent asymmetry* in net value perceived (and SVIs). Secondly, they also help the franchise achieve a certain desired level of customer service. The franchise is currently implementing our recommendation of pricing the single-game tickets asymmetrically for the upcoming season.

2. **Value of Seat Locations for Consumer Segments (based on Age):** Conventional

wisdom provides some guidelines on valuable seat locations in a baseball stadium. For example, Backnet seats are considered quite valuable to customers. In the introduction, we raised the question: "Do the customers seated at the upper deck value those seats highly?". Equipped with our analysis, we can now summarize the value perceived by customers at those seats, and compare our findings with common notions of seat value. Moreover, we can do this analysis across each consumer segment.

Upper Deck Seats: First, we consider upper deck seats that are generally inexpensive, and located further away from the playing field. Our results suggest higher mean SVI for customers seated at the upper deck. While higher mean seat values are interesting in their own right, our analysis of marginal probabilities reveals a subtler insight. For example, considering a customer in the age group 30-39, we find that he has the same probability of reporting his SVI as Low (or Medium) whether he is seated at the lower deck or the upper deck. However, the probability of reporting SVI as High increases as he moves from a lower deck seat to a similar upper deck seat. In other words, the higher value perceived at the upper deck is almost entirely driven by a significantly higher proportion of customers reporting their seat value as high. Thus, our results argue for the continued availability of upper deck seats for customers.

Backnet Seats: Backnet seats are often considered to be the best seats in the stadium. However, it is unclear how the franchise should price them across consumer segments.

Our analysis implies that the franchise can offer age based discounts as summarized in Table 5.10. For consumers in the age group 10-19, the recommended segment-specific prices are 13% lower than the current single ticket Backnet prices, whereas

for the age group 30-39, the recommended prices are 9.7% lower than current prices (see Table 5.10). In effect, according to our segment-specific pricing scheme, high-school/college students (in age group 10-19) should receive roughly a 5% discount for single-ticket Backnet prices, compared to the age group 30-39. The franchise can achieve the service-level objective of $\alpha = 15\%$ by suitably *discounting* the backnet seat prices, as indicated in Table 5.10.

Grass Seats: Grass seats located further in the outfield are similar to upper deck seats. Customers perceive significantly higher value at grass seats. In contrast to upper deck seats, this higher value is driven by a mean shift in its distribution. Conducting a segment-specific pricing analysis similar to that carried out for the Backnet seats, we find that we can increase the grass seat ticket prices and still keep the probability of low SVI within 15%.

3. **Segment Specific Prices based on Frequency of Visits:** We find that repeated visits to the ballpark reduce the probability that a customer would report extreme SVI. For example, we find that a customer visiting the ball-park for the eighth time has an 8% lower probability of reporting SVI = High as compared to a first-time visitor (Table 5.3 shows that the probability of reporting SVI = High reduces by 1% for every additional visit). Looking at the results of the Generalized Threshold Model in greater detail, we infer that a likely explanation for the reduced tendency of the more frequent customers to report extreme SVIs is that they use stricter thresholds ($\hat{\delta}_{11}^2 - \hat{\delta}_{11}^1 = \hat{\beta}_{11}^1 - \hat{\beta}_{11}^2 = 0.141$).[14] In other words, for the same experience and realization of net value, the more frequent customers are less likely to be 'surprised' and are therefore less likely to respond with extreme reactions.

---

[14]By "stricter thresholds", we mean that consumers use a higher threshold to report Low SVI, and a lower threshold to report High SVI (i.e., consumers are more likely to respond "Low", and less forthcoming to respond "High").

While the difference in reporting thresholds seems to be the most likely explanation for the observed distribution of SVIs, one cannot rule out the possibility that the distribution of net values might be heteroskedastic with respect to frequency of visits. In fact, the results of the heteroskedastic model would suggest that the distribution of realized values for the more frequent customers does have a lower variance ($\hat{\gamma}_{11} = -0.057$) as well as a lower mean ($\hat{\beta}_{11} = -0.081$). This would support the notion that baseball games are experience goods with residual uncertainty that decreases with repeated visits to the ballpark.

Based on the segment-specific pricing analysis, we can recommend price discounts for each seat location based on frequency of visits. These prices are summarized in Table 5.11. First, we find that the recommended price discounts increase with increasing frequency of visits. Second, we observe that the recommended ticket prices can be higher or lower compared to the current prices. For example, at the 3rd Base, the recommended prices are $5 - 12\%$ higher, whereas at the Backnet, they are $9 - 10\%$ lower than the current prices. This leads to a subtler third insight, that the price discounts offered to a season regular (relative to a first-timer) can be as high as 6% (for the 3rd Base) and as low as 2% (for the Backnet).

5.6   Conclusions and Future Direction

In this chapter, we first developed Seat Value Index, a measure of net value perceived by a consumer after attending an event. Then, we established the relationship between the SVIs reported by consumers and their seat locations. Finally, we provide directions that would help the firm achieve a *"desired level of seat value"* by suitably increasing or decreasing ticket prices in each segment. The key steps of our

approach and methodology can be summarized as:

1. Capture on some ordinal scale, the net value perceived by consumers, using a survey instrument.

2. Design a Seat Value Index (SVI) measure.

3. Investigate how the Seat Value Index is influenced by consumer characteristics, seat location attributes and event-related factors, using a series of Ordinal Logit Models. Deviation from proportional-odds (verified using Brant test) suggests the presence of heterogeneity in the model, which can be incorporated in two ways: (i) The Generalized Threshold Model, which assumes that consumers use different thresholds, and (ii) The Heteroskedastic Model, which assumes that the variance of the idiosyncratic value term differs across covariates.

4. Estimate current service-levels as the probability of a given customer seated at a particular location reporting SVI=Low. Then, optimize the prices to achieve the aforementioned probability threshold acceptable to the firm (as derived in Lemma 1).

We illustrated the application of our methodology by applying it to two survey datasets collected by a professional league baseball franchise in Japan. Our findings provide a characterization of seat value perceived by consumers in a stadium based on their age, location of the seat, and the number of visits. We showed that a careful study of the interactions between SVI and the explanatory variables, specifically accounting for systematic heterogeneity in response thresholds and distributions of seat value across customer segments, reveals some relatively unexpected dependencies (asymmetries, etc.). Detailed analysis reveals that the seat location plays a

152

crucial role in how seat values are distributed, which enables us to consider pricing based on individual segments.

The insights on seat value that we derive here provide the crucial initial steps in planning how seats should be sold, and how to price tickets based on segment-specific and consumer-specific information for different sections of the stadium/theater.

Limitations: Finally, our research is not without limitations, which is typical while exploring empirical RM aspects. The first limitation is that our pricing recommendations ignore the effects of substitution. Pricing changes might modify the valuations and choices that consumers make. Hence, it is important for the franchise to keep this limitation in mind and further estimate the changes in demand or customers' future valuations. This could be achieved by perturbing prices and observing the resulting demand and re-evaluating customers' responses.

The second limitation is that consumer responses to price changes might change the optimal assortment of different ticket categories both in prices and capacity offered at that price. The assortment decision can be studied with additional data on how customers arrived at their revealed preferences. Analyzing Capacitated Multinomial Logit assortment problems is a challenging stream of research. For example, see Rusimevichientong et al. (2010), and references therein. Due to paucity of data on how consumers chose their seats, we did not model the optimal assortment decision.

A third limitation is around the design of the survey. Most customers in the survey reported SVI = 2. Although this may be a natural response of consumers in our context, we cannot rule out the possibility that respondents avoided using extreme response categories (referred to as central tendency bias). Future work can focus on improved survey design and better measurement of consumer responses in order

153

to counter these biases.

A fourth limitation is that SVI is clearly influenced by the actual price paid by consumers. However, we were unable to incorporate seat prices directly into our model and study its effects in detail, as our dataset lacked granular prices at the consumer level. This presents an opportunity for future work, where more granular price data could be gathered to simultaneously study the impact of price and seat location on consumer valuations.

Furthermore, Neelamegham and Jain (1999) argue that modeling customers' expectations (through emotional stimulation and latent product interest) before the choice is made, and modeling post choice evaluations (determined by consumers' post consumption experience) are both important in modeling the consumption of experience goods. Thus our findings on post-consumption perceived value, combined with the decision-models of customers' revealed preferences, would allow firms to explore the impact of subsequent decisions in greater detail.

Finally, we were unable to incorporate individual level heterogeneity in our model on account of data limitations. However, in the presence of a richer data-set, we could estimate a Hierarchical Bayes model (Bradlow and Zaslavsky 1999) that would allow us to incorporate individual level heterogeneity and obtain more robust estimates of the effects of covariates.

Nevertheless, we hope that our analysis of differing seat values provides sports franchises and theater establishments with the first steps in analyzing customer perceptions of different seats, and factoring those perceptions while making their pricing decisions. In a variety of sporting events/performances, the attending customers value their experience differently based on their seat locations. Although some seats might appear similar, they might provide different valuations for long-

time patrons who have a well-developed sense about which seats have better value. Exploring such non-obvious differences in the value perceived by customers located in different seats provides sports and theater establishments with an opportunity to improve their customer base through more efficient pricing, or better selling mechanisms.

Chapter 6

Concluding Remarks

In this dissertation, we focused on three issues affecting retail demand management. First, we addressed the issue of demand estimation, assortment optimization and assortment localization in product retailing. Second, we investigated the sensitivity of the optimal assortment and expected profits on the key assumptions made about the choice model and substitution under stock-outs. Finally, we explored the issue of pricing of seats in entertainment settings to achieve a service level objective on customer satisfaction.

In Chapter 3, we formulated a process for finding optimal assortments, comprised of a demand model, estimation approach and heuristics for choosing assortments. We applied this process to real data from three applications and showed that the approach produces accurate forecasts for new SKUs. Our recommendations were implemented in two of the cases. We measured the impact based on actual sales and found the assortment revisions had produced revenue increases of 5.8% and 3.6%, which are significant relative to typical comparable store increases in these product segments. Our research provides a framework to stimulate much needed

additional research, and there are four concrete directions in which we can take this forward.

First, an obvious enhancement of our approach would be to extend the demand estimation methodology to multiple time periods. This would incorporate sales data over multiple time periods allowing us to model sales trends across 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. Nevertheless, it is an important enhancement to bridge the gap between theory and practice, since it is critical to capture trends.

Second, as we observed in several of the applications in this chapter, 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.

Third, we did not have access to detailed sales and inventory data over time. Given that our approach to demand estimation utilized 'holes' in the assortment to measure substitution, stock-outs constitute a natural experiment to quantify the effects of how demand shifts when a product is not available in the assortment. Detailed inventory data over time would thereby 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.

Finally, 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. Estimating a demand model with price effects would also allow joint optimization of the assortment offered and price.

In Chapter 4, we investigated the impact of commonly made assumptions of choice model and effect of stock-outs, on the optimal assortment and profits. We derived analytical bounds on the optimal profits in the presence of stock-out substitution and simulated the actual expected profits from using simple newsvendor based heuristics, and concluded that for a wide range of problem parameters in practice, the effect of ignoring stock-out substitution is not significant. On the other hand, we found that incorrectly specified choice models impacts profits significantly. Using an MNL choice model when the true underlying substitution structure is governed by a LC model leads to sub-optimal profits, often leading to a loss of more than 25%.

There are several interesting ways in which this work can be extended. First, we only studied the impact of choice model misspecification for the case where the underlying choice model was governed by locational choice and incorrectly assumed to be MNL. It would be useful to explore the effects of misspecification over a wider class of underlying choice models and incorrect specifications. This would enable us to develop a deeper understanding of the sensitivity of the optimal assortment and expected profits to choice models, and help develop more robust

approaches to assortment optimization.

Second, it would be interesting to study sensitivity of the assortment profits to assumptions made about the arrival distribution. More specifically, it would be worthwhile investigating the impact on optimal profits when the underlying demand process is over-dispersed, which is the case for several retail categories.

Third, we studied the effects of choice model misspecification and ignoring stockouts separately. In practice, these would occur simultaneously. Hence, interesting future work could target enhancing the approach developed here to incorporate these effects simultaneously and quantifying their impact on the optimal assortment profits. Understanding the drivers of assortment profits would allow retailers to focus their efforts on these high-impact areas.

In Chapter 5, we developed a measure of seat value, called Seat Value Index (SVI) and related it to consumer and seat location attributes. We derived interesting insights on how SVI is affected by these factors for a baseball stadium located in Japan. Using our models of SVI, we provided insightful and actionable recommendations aimed at helping such firms achieve a service level objective of keeping seat value perceived by consumers above a threshold value.

There are several ways in which we can extend this work in the future. First, it would be interesting to investigate game effects on seat value by gathering multi-game data, since game effects do significantly affect seat value. This would be extremely useful in uncovering other sources of seat value. For instance, analyzing seat value across multiple games, one might find that seats on the side of the home team's dugout carry higher value for customers, and a franchise could exploit this knowledge to extract value. This would require gathering survey data over multiple games and multiple stadiums and can be supplemented using secondary market

information on ticket prices from sources like StubHub.

Second, it would be useful to incorporate the effects of substitution across seats. Customers making their purchase do choose across seat locations and any recommendations on price change should take into account how it impacts choice. We can model this as a capacitated assortment optimization problem, and solve for optimal seat prices that would maximize revenue while keeping the probability of a low seat value below a threshold.

Bibliography

Anderson, E. W., C. Fornell, S. K. Mazvancheryl. 2004. Customer satisfaction and shareholder value. *Journal of Marketing* **68**(4) 172–185.

Anderson, Eugene W, Mary W Sullivan. 1993. The antecedents and consequences of customer satisfaction for firms. *Marketing Science* **12**(2) 125–143.

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.

Atkinson, A.B. 1970. On the measurement of inequality. *Journal of economic theory* **2**(3) 244–263.

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.

Belsley, D. A. 1991. *Conditioning diagnostics: collinearity and weak data in regression*. Wiley New York, NY.

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

Besbes, O., D. Saure. 2010. Product assortment and price competition under multinomial logit demand. Working Paper.

Beswick, P., M. Isotta. 2010. Making the right choices: SKU rationalization in retail. Oliver Wyman Working Paper.

Bradlow, E. T., A. M. Zaslavsky. 1999. A hierarchical latent variable model for ordinal data from a customer satisfaction survey with" no answer" responses. *Journal of the American Statistical Association* **94**(445) 43–44.

Brant, Rollin. 1990. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics* **46**(4) 1171–1178.

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. 2008. *Matching Supply with Demand: An Introduction to Operations Management*. McGraw-Hill, Boston, MA.

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.

Courty, P. 2000. An economic guide to ticket pricing in the entertainment industry. *Louvain Economic Review* **66**(1) 167–192.

Craggs, Tommy. 2006. Upper cuts: Baseball's war on the cheap seats. Slate.com.

Dana, J.D., Jr. 1998. Advance-purchase discounts and price discrimination in competitive markets. *Journal of Political Economy* **106**(2) 395–422.

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.

Farias, V.F., S. Jagabathula, D. Shah. 2011. A nonparametric approach to modeling choice with limited data. MIT Working Paper.

Fisher, M., A. Raman. 2010. *The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance*. Harvard Business School Press.

Gans, N. 2003. Customer loyalty and supplier quality competition. *Management Science* **48**(2) 207–221.

Gaur, V., D. Honhon. 2006. Assortment planning and inventory decisions under a locational choice model. *Management Science* **52**(10) 1528–1543.

Gaur, V., Y.H. Park. 2007. Asymmetric consumer learning and inventory competition. *Management Science* **53**(2) 227–240.

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.

Greene, W. H. 2003. *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ.

Groot, Wim. 2000. Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics* **19**(3) 403–420.

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

Hall, J., E. Porteus. 2000. Customer service competition in capacitated systems. *Manufacturing & Service Operations Management* **2**(2) 144–165.

Hansen, L. P., T. J. Sargent. 1991. *Rational expectations econometrics*. Westview Press Boulder.

Harvey, A. 1976. Estimating regression models with multiplicative heteroskedasticity. *Econometrica* **44** 461–465.

Hauser, J. R. 1978. Testing the accuracy, usefulness, and significance of probabilistic choice models: An information theoretic approach. *Operations Research* **26**(3) 406–21.

Homburg, C., N. Koschate, W. D. Hoyer. 2005. Do satisfied customers really pay more? A study of the relationship between customer satisfaction and willingness to pay. *Journal of Marketing* **69**(2) 84–96.

Honhon, D., V. Gaur, S. Seshadri. 2010. Assortment planning and inventory decisions under stockout-based substitution. *Operations Research* **58**(5) 1364–1379.

Honhon, D., S. Seshadri. 2011. Fixed versus random proportions demand models for assortment planning. Working Paper, Univ. of Texas at Austin.

Hopp, W.J., X. Xu. 2008. A static approximation for dynamic demand substitution with applications in a competitive market. *Operations Research* **56**(3) 630–645.

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

Ittner, C.D., D.F. Larcker. 1998. Are nonfinancial measures leading indicators of financial performance? an analysis of customer satisfaction. *Journal of Accounting Research* **36** 1–35.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lutkepohl, T.C. Lee. 1980. *The Theory and Practice of Econometrics*. Wiley, New York, U.S.A.

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

Kekre, S., M.S. Krishnan, K. Srinivasan. 1995. Drivers of customer satisfaction for software products: Implications for design and service support. *Management Science* **41**(9) 1456–1470.

Kmenta, J. 1986. *Elements of Econometrics, 2nd Edition*. Macmillan Publishing Company, New York.

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.

Leslie, Phillip. 2004. Price discrimination in Broadway theater. *The RAND Journal of Economics* **35**(3) 520–541.

Liu, Agresti. 2005. The analysis of ordered categorical data: An overview and a survey of recent developments. *TEST* **14**(1) 1–73.

Lucas, R. E., T. J. Sargent. 1981. *Rational expectations and econometric practice*. University of Minnesota Press, Minneapolis MN.

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

Maddala, G.S. 1983. *Qualitative and Limited Dependent Variable Models in Econometrics*. Cambridge: Cambridge University Press.

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

Maske, Mark. 1992. "For players, park will need figuring". Washington Post, 31st March 1992.

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

McCullagh, P., J.A. Nelder. 1989. *Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton, FL.

McCullagh, Peter. 1980. Regression models for ordinal data. *Journal of the Royal Statistical Society. Series B (Methodological)* **42**(2) 109–142.

McGregor, Jena. May 15, 2008. At Best Buy, marketing goes micro. *BusinessWeek* .

McKelvey, R.D., W. Zavoina. 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* **4**(1) 103–120.

Muth, John F. 1961. Rational expectations and the theory of price movements. *Econometrica* **29**(3) 315–335.

Neelamegham, Ramya, Dipak Jain. 1999. Consumer choice process for experience goods: An econometric model and analysis. *Journal of Marketing Research* **36**(3) 373–386.

O'Connell, Vanessa. April 21, 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.

Peterson, Bercedis, Frank E. Harrell. 1990. Partial proportional odds models for ordinal response variables. *Applied Statistics* **39**(2) 205–217.

Popescu, Ioana, Nils Rudi. 2008. Revenue management for fans (or collective experience goods). INSEAD Working Paper.

Roels, G., K. Fridgeirsdottir. 2009. Dynamic revenue management for online display advertising. *Journal of Revenue and Pricing Management* **8**(5) 452–466.

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

Rossi, Peter E, Zvi Gilula, Greg M Allenby. 2001. Overcoming scale usage heterogeneity: A Bayesian hierarchical approach. *Journal of the American Statistical Association* **96**(453) 20–31.

Rusimevichientong, P., Z.J.M. Shen, D.B. Shmoys. 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research* (6) 1666–1680.

Shugan, S. M., J. Xie. 2000. Advance pricing of services and other implications of separating purchase and consumption. *Journal of Service Research* **2**(3) 227–239.

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

Steward, Carl. 2006. A's committed to intimate 34,000 park in '06. Oakland Tribune, January 29, 2006.

Su, Xuanming. 2007. Intertemporal pricing with strategic customer behavior. *Management Science* **53**(5) 726–741.

Sun, B., S. A. Neslin, K. Srinivasan. 2003. Measuring the impact of promotions on brand switching when consumers are forward looking. *Journal of Marketing Research* **40**(4) 389–405.

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

Urban, Mychael. 2005. A's to offer more fan-friendly stadium. MLB Press Release, December 21, 2005.

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.

Williams, R. 2006a. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal* **6**(1) 58–82.

Williams, R. 2006b. OGLM: Stata module to estimate ordinal generalized linear models. Working Paper.

Xie, J., S. M. Shugan. 2001. Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Science* **20**(3) 219–243.

Zeithaml, V. A. 1988. Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing* **52**(3) 2–22.

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* .