

MANAGEMENT OF PRODUCT VARIETY IN RETAIL OPERATIONS

Abdullah Gürhan Kök

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

2003

---

---

Dissertation Supervisor

---

Graduate Group Chairperson

COPYRIGHT

Abdullah Gürhan Kök

2003

To Mom and Dad,

# Acknowledgements

I would first and foremost thank my advisors Marshall Fisher and Gérard Cachon. I would like to thank Marshall Fisher for his mentorship, guidance and financial support throughout my doctoral studies. He has been a great example of passion and execution in research. He provided me with the chance to work on applied research projects, which had a tremendous impact on my taste in research. I take great pride in being his student and having him as a professional and personal role model. I would like to also thank Gérard Cachon for being a mentor to me in research and other professional matters. I would like to thank him for spending so much time with me in front of the white board in his office. He guided me through many roadblocks and many of my frustrations. Without him, this dissertation would not have been possible. My learnings from him have an enormous impact on my future research agenda and my view of the academic world. I feel immensely privileged to have been his student.

I would like to thank other members of dissertation committee: David Bell, Morris Cohen, and Ananth Raman for their comments on this dissertation which has improved it significantly. I would like to thank Ananth Raman also for sharing his experiences and valuable opinions since my first year at Wharton.

The first part of this dissertation is based on a research project with AHOLD, N.V. of the Netherlands. Thanks go to Frank Jansen and Robert van Lunteren of Ahold R&D Department and Marshall Fisher for providing me with the chance to work on real problems and the satisfaction to know that some of my research will contribute directly to the improvement of a business process.

I have been part of a wonderful group of people at the Operations and Information Management Department at The Wharton School of the University of Pennsylvania. I would like to thank the faculty for my formal and informal education. Special thanks to Ziv Katalan for arranging the flexible teaching assistantship for one semester in my final year. Many thanks to the staff of the OPIM Department: Kim Watford, Marjorie Wyler, Barbara Ayrapetyan, and Cynthia Anderson, for their friendship and administrative support.

I would like to thank Yi Xu and Justin Ren for walking down this path with me from the beginning till the end. We had uncountably many lunches and dinners together. I want to also thank my colleagues Noel Watson, Vishal Gaur, Sergei Savin, Manu Goyal, and Fuqiang Zhang and my roommates Tolga Emre and Volkan Isler for their friendship.

I would like to thank my family: My parents, my sister, and my brother. Being away from you was the hardest part of this journey I undertook. Dad, I wouldn't be the man I am today if it were not for your example and I can only hope to live up to the standards set by your example. Mom, your infinite love and support had me going all these years. Thanks for always being there for me and believing in me more than I ever did. Tugba and Emre, I love you. Grandma, thanks for including me in your prayers. Finally, I would like to thank Cagla for her love and care: SHSCS.

## ABSTRACT

### MANAGEMENT OF PRODUCT VARIETY IN RETAIL OPERATIONS

Abdullah Gürhan Kök

Marshall L. Fisher and Gérard P. Cachon

In the first part of the dissertation, we study the assortment planning problem where the retailer selects the set of products to be carried at each store and the inventory level of each selected product. We introduce a novel methodology of demand estimation for multiple products under demand substitution using sales data from multiple stores. We obtain a chain-wide substitution rate and for each store, we obtain demands of products carried in the store as well as products previously not carried in the store. We also describe a heuristic optimization procedure for localizing the assortment to each store and establish new structural properties (based on the heuristic solution).

In the second part, we study the interaction of variety decisions across different merchandise categories and their impact on the long run attractiveness of the stores to the customers. We characterize the equilibrium and optimal variety levels under decentralized and centralized management regimes, respectively. We show that category management, which is an example of a decentralized regime for controlling assortment, chooses variety levels that are lower than the optimal solution in at least one category and the profit loss due to decentralization is significant. We describe a simple heuristic that recovers most of this loss.

In the third part of the dissertation, we explore the impact of competition between suppliers selling their products through a retailer under wholesale price contracts and revenue sharing contracts. Our model with competing contracts brings new perspectives to the area of supply chain coordination with contracts. Contracts serve as a tool for competition rather than achieving system optimality. Under wholesale price contracts, we show that introducing competition at the suppliers can be desirable as it offsets some of the double marginalization effects. We find that the revenue sharing contracts coordinate the system when the retailers are centrally managed, but lead to more than system optimal inventory when retailers are decentralized. Also revenue sharing contracts lead to more inventory than wholesale price contracts.



# Contents

|          |                                                                     |           |
|----------|---------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                                 | <b>1</b>  |
| 1.1      | Product variety issues and the scope of this dissertation . . . . . | 2         |
| 1.2      | Retailing industry . . . . .                                        | 12        |
| <b>2</b> | <b>Literature Review</b>                                            | <b>15</b> |
| <b>3</b> | <b>Joint Assortment Selection and Inventory Planning</b>            | <b>21</b> |
| 3.1      | Related literature . . . . .                                        | 23        |
| 3.2      | Problem structure . . . . .                                         | 29        |
| 3.2.1    | Modeling consumer demand . . . . .                                  | 29        |
| 3.2.2    | Estimation of Demand . . . . .                                      | 34        |
| 3.2.3    | Description of the inventory system . . . . .                       | 37        |
| 3.2.4    | Assortment optimization . . . . .                                   | 39        |
| 3.3      | Application . . . . .                                               | 41        |

|          |                                                           |           |
|----------|-----------------------------------------------------------|-----------|
| 3.3.1    | Data . . . . .                                            | 42        |
| 3.3.2    | Assessing the quality of the estimation . . . . .         | 43        |
| <b>4</b> | <b>Assortment Planning under Demand Substitution</b>      | <b>49</b> |
| 4.1      | Model of Substitution . . . . .                           | 51        |
| 4.1.1    | Suggested model . . . . .                                 | 54        |
| 4.2      | Estimating Substitution . . . . .                         | 56        |
| 4.2.1    | Estimation of assortment-based substitution . . . . .     | 56        |
| 4.2.2    | Estimation of stockout-based substitution . . . . .       | 62        |
| 4.3      | Assortment Optimization under Substitution . . . . .      | 63        |
| 4.3.1    | Comparison with other optimization methods . . . . .      | 67        |
| 4.3.2    | Structural Properties . . . . .                           | 69        |
| 4.3.3    | Numerical study . . . . .                                 | 74        |
| 4.4      | Implementation and Results . . . . .                      | 76        |
| 4.4.1    | Results and financial impact . . . . .                    | 77        |
| 4.4.2    | Comparison with alternative method . . . . .              | 78        |
| 4.4.3    | Impact of delivery frequency on the system cost . . . . . | 79        |
| 4.5      | Discussion . . . . .                                      | 79        |

|          |                                                                     |            |
|----------|---------------------------------------------------------------------|------------|
| <b>5</b> | <b>Category Coordination in Retail Assortment Planning</b>          | <b>81</b>  |
| 5.1      | Related literature . . . . .                                        | 84         |
| 5.2      | Model Basics . . . . .                                              | 87         |
| 5.3      | Analysis . . . . .                                                  | 93         |
| 5.3.1    | Category management . . . . .                                       | 93         |
| 5.3.2    | Optimal solution . . . . .                                          | 100        |
| 5.3.3    | Comparison of the centralized and decentralized solutions . . . . . | 102        |
| 5.3.4    | Numerical examples . . . . .                                        | 104        |
| 5.4      | Properties of the Optimal and Equilibrium Assortments . . . . .     | 105        |
| 5.4.1    | Depth vs. breadth . . . . .                                         | 105        |
| 5.4.2    | Expanding the assortment: Adding new categories . . . . .           | 108        |
| 5.4.3    | Which categories should have more variety? . . . . .                | 109        |
| 5.5      | Aggregation Heuristic . . . . .                                     | 111        |
| 5.6      | Numerical Results . . . . .                                         | 113        |
| 5.7      | Discussion . . . . .                                                | 116        |
| <b>6</b> | <b>Supplier Competition in Retail Supply Chains</b>                 | <b>119</b> |
| 6.1      | Related Literature . . . . .                                        | 123        |
| 6.2      | Model Basics . . . . .                                              | 127        |
| 6.3      | Model C: All suppliers and retailers centralized . . . . .          | 130        |

|          |                                                                       |            |
|----------|-----------------------------------------------------------------------|------------|
| 6.4      | Model D: Competing supplier-retailer pairs . . . . .                  | 130        |
| 6.5      | Model CD: Centralized suppliers, competing retailers . . . . .        | 131        |
| 6.5.1    | Wholesale price contracts . . . . .                                   | 131        |
| 6.5.2    | Revenue sharing contracts . . . . .                                   | 133        |
| 6.6      | Model DD: Competing suppliers and competing retailers . . . . .       | 135        |
| 6.6.1    | Wholesale price contracts . . . . .                                   | 135        |
| 6.6.2    | Revenue sharing contracts . . . . .                                   | 139        |
| 6.7      | Model CC: Centralized suppliers, centralized retailers . . . . .      | 140        |
| 6.7.1    | Wholesale price contracts . . . . .                                   | 140        |
| 6.7.2    | Revenue sharing contracts . . . . .                                   | 141        |
| 6.8      | Model DC: Supplier competition through centralized retailer . . . . . | 141        |
| 6.8.1    | Wholesale price contracts . . . . .                                   | 141        |
| 6.8.2    | Revenue sharing contracts . . . . .                                   | 143        |
| 6.9      | Numerical Examples . . . . .                                          | 146        |
| 6.10     | Summary and Discussion of Results . . . . .                           | 148        |
| 6.11     | Future research . . . . .                                             | 150        |
| <b>7</b> | <b>Summary of Contributions</b>                                       | <b>153</b> |

|          |                                                                            |            |
|----------|----------------------------------------------------------------------------|------------|
| <b>8</b> | <b>Appendices</b>                                                          | <b>157</b> |
| 8.1      | Appendix A: Sum of two Gumbel distributed random variables . . . . .       | 157        |
| 8.2      | Appendix B: Example of a replenishment system with convex operating cost   | 159        |
| 8.3      | Appendix C: Proof of second order properties of the contract parameters in |            |
|          | Model DC : . . . . .                                                       | 160        |
| 8.3.1    | Wholesale price contracts . . . . .                                        | 160        |
| 8.3.2    | Revenue sharing contracts . . . . .                                        | 161        |

# List of Tables

|                                                                                                                               |     |
|-------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1. Summary of merchandise categories in data .....                                                                          | 179 |
| 3.2. Estimation quality for purchase incidence models for all stores .....                                                    | 179 |
| 3.3. Estimation quality for the choice model for subcategory 10.....                                                          | 179 |
| 3.4. Comparison of the <i>choice</i> model with other alternative models .....                                                | 180 |
| 4.1. A sample of substitution estimates from different weeks in the planning period .....                                     | 180 |
| 4.2. Iterative optimization example .....                                                                                     | 180 |
| 4.3. Performance of the optimization method in comparison to the optimal solution.....                                        | 181 |
| 4.4. Performance of the optimization method in comparison to the optimal solution for different<br>levels of shelf space..... | 181 |
| 4.5. Comparison of existing and suggested assortments.....                                                                    | 181 |
| 5.1. Allocation of basket shoppers in numerical experiments.....                                                              | 182 |
| 5.2. Summary of numerical results .....                                                                                       | 182 |

|                                                                                                                                                                                                                                                                              |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.3. Performance of the category management and aggregation heuristic in comparison to the optimal solution for a set of examples .....                                                                                                                                      | 183 |
| 6.1. Comparison of equilibrium under wholesale price and revenue sharing contracts for two symmetric suppliers selling through a centralized retailer. $d_i \sim \text{Normal}(100,30)$ , $b=0.05$ , $p_i=100$ , $c_i=10$ for $i=1,2$ . ....                                 | 183 |
| 6.2. Comparison of equilibrium under wholesale price and revenue sharing contracts for asymmetric suppliers selling through a centralized retailer. $d_1 \sim \text{Normal}(200,30)$ , $d_2 \sim \text{Normal}(100,30)$ , $b=0.05$ , $p_i=100$ , $c_i=10$ for $i=1,2$ . .... | 183 |
| 8.1. Resulting choice probabilities using the sum of Gumbel random variables and its multinomial logit approximation.....                                                                                                                                                    | 184 |

# List of Figures

|                                                                                                                                         |     |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1. A sample realization of inventory trajectory over time .....                                                                       | 185 |
| 3.2. Average of mean absolute deviation (MAD) values of day-store-SKU level $PQ$ estimates.<br>.....                                    | 185 |
| 4.1. Demand of variants under single and multiple substitution attempts when variants 8,9 and 10<br>are left out of the assortment..... | 186 |
| 4.2. Sales vs. Number of variants.....                                                                                                  | 187 |
| 4.3. Sales vs. Total original demand .....                                                                                              | 187 |
| 4.4. Substitution estimation scheme .....                                                                                               | 188 |
| 4.5. Estimates of observed demand and original demand for a subcategory.....                                                            | 188 |
| 4.6. Percentage improvement in category gross profits by the suggested assortment over the<br>current assortments.....                  | 189 |

|                                                                                                            |     |
|------------------------------------------------------------------------------------------------------------|-----|
| 5.1. A two category example with unique category management equilibria .....                               | 190 |
| 5.2. A two category example with three category management equilibria.....                                 | 190 |
| 5.3. A two category example where the unique category management equilibrium is at zero..                  | 190 |
| 5.4. Average profit loss due to decentralization.....                                                      | 191 |
| 5.5. Average decrease in variety level due to decentralization .....                                       | 191 |
| 8.1. Cumulative density functions of the sum of Gumbel random variables and its Gumbel approximation ..... | 192 |

# Chapter 1

## Introduction

This dissertation focuses on the management of product variety in retailing. The dissertation has three parts. Chapter 2 reviews the literature related to all three parts. In the first part, comprised of Chapters 3 and 4, we study the assortment planning problem where the retailer selects the set of products to be carried at each store and the inventory level of each selected product. We introduce a novel methodology for demand estimation for multiple products under demand substitution using sales data from multiple stores. We also describe an optimization procedure for localizing the assortment to each store. We apply our methodology at a multi-billion dollar retailer and measure its financial impact. In the second part, comprised of Chapter 5, we look at the assortment planning problem from a different perspective. In an attempt to understand the long run effects of product variety in retailing, we study the interaction of variety decisions across different merchandise categories and their impact on the long run attractiveness of the stores to the customers. The third part, comprised of Chapter 6, explores the impact of competition between different

brands (manufacturers) on retailers' inventory decisions and the manufacturers' profits in a two-stage supply chain. We characterize and compare the solutions under wholesale price contracts and revenue sharing contracts.

## **1.1 Product variety issues and the scope of this dissertation**

The product assortment decision is probably one of the most important decisions faced by the retailers. This decision has several aspects. Retailers have to decide what categories to carry, how much space to allocate for each category. Within each category retailer needs to decide what products to carry along with the inventory levels for each product. Related to these, the pricing policies across and within categories need to be decided.

Product variety has been studied extensively from manufacturer's point of view. The retailer's problems relating to product variety are quite different from the manufacturer's. Manufacturer's is a product line design problem, where the number, quality and pricing of products should be determined in order to determine its position relative to its competitors. The retailer's problem is choosing subsets of products from the product lines of multiple suppliers in order to attract customers from the target segments. The cost structures related to expanding an assortment can also be substantially different in retailing and manufacturing. Product line decisions for a manufacturer involves longer periods of planning and possibly major investments in design and production processes.

The conventional wisdom in retailing when managing product variety has been 'more is better'. Due to heterogeneous and uncertain customer preferences, variety seeking behavior by consumers, and competition between brands, retailers have increased product selections

in all merchandise categories. For example, the average number of stock-keeping units (SKUs) carried by a grocery store increased from 6,000 to 30,000 in one generation (Food Marketing Institute Report, 1993). Quelch and Kenny (1994) report that the number of products in the market place increased by 16% per year between 1985 and 1992 while shelf space expanded only by 1.5% per year during the same period. Managing product variety requires more attention and finesse than before because of these trends. With fixed shelf space, increasing variety means less store space for each product, which translates to higher operational costs due to either lower availability or higher replenishment costs. As a result of the excessive levels of variety, many retailers are switching to a strategy called “efficient assortments”, which primarily seeks to find profit maximizing level of variety by eliminating low-selling products (Kurt Salmon Associates, 1993). Many studies report that reduction in variety resulted in higher gross profits and sometimes even higher sales.

In the first part of the dissertation, we study the joint assortment selection and inventory planning problem, defined as the selection of the products in a category and setting the inventory levels for each product. This problem can be thought as a financial portfolio selection problem: selecting stocks with uncertain returns under a budget (or in our case a shelf space) constraint. Limited shelf space creates a tradeoff between breadth of assortment and in-stock rates of products in the assortment. We study this tradeoff in Chapter 3 under the assumption that there is no substitution between products by consumers, that is the demand of each product is independent of demand and availability of other products. We develop a model of demand estimation in this case using daily sales data at store-product level. We also describe an assortment optimization algorithm that optimally

selects the products and their inventory levels. The algorithm accommodates many realistic constraints such as batch sizes, delivery lead times, facing sizes, perishability, etc. The no-substitution model serves as a benchmark and a basic building block of the model with substitution.

The degree of importance of the tradeoff between breadth of assortment and in-stock rates of products depends on the customers' willingness to accept another product as a substitute if their favorite product is not available. From a category sales perspective, it doesn't matter whether a customer bought their favorite product or a substitute. If customers are always willing to accept any substitute, the optimal assortment is to carry one product and benefit from the inventory pooling effects. Clearly, the reality lies between no substitution and full substitution cases. Assessing how likely consumers are to accept one of the available variants as a substitute and incorporating this into assortment planning is therefore essential for retailers. Chapter 4 presents a methodology for assortment planning under demand substitution.

Although there had been significant research in inventory and assortment decisions under substitution, there is little research on the estimation of input parameters. We introduce a relatively simple model of demand substitution and discuss its strengths and limitations. We present a novel methodology for demand estimation under the assumed substitution structure. The methodology is based on the following observation. Define a subcategory as a substitutable group of products. Consider stores  $A$  and  $B$ , with the same base demand rates for all products in a subcategory. Store  $B$  carries every possible variant, assortment  $N$ . Store  $A$  carries assortment  $S \subset N$ . If the category sales at store  $A$  is more than the

sales of products in  $S$  at store  $B$ , it means that at store  $A$  some of the customers who came for products in  $N \setminus S$  purchased products from  $S$  as substitutes. We use data from multiple stores with different assortments to find out how much more an assortment really sold than its contribution at stores with full assortment. For a given rate of substitution, we can theoretically compute how much more  $S$  would sell at store  $A$  than its contribution to sales at store  $B$ . Combining these, we find the substitution rate that minimizes the error between the theoretical and observed differences. Using the estimated substitution rate and the observed demands, we reverse engineer the true demand of each product at each store. We obtain a chain-wide substitution rate for each subcategory of products and for each store, we obtain demands of products carried in the store as well as products previously not carried in that store.

Under substitution, the effective demand to a product depends on the inventory levels of other products because of the spill over from the unavailable products. Hence the assortment optimization problem is a nonlinear, nonseparable, discrete resource allocation problem with integer variables. We describe an iterative heuristic that basically solves a series of nonlinear knapsack problems. On average, our method seems to find solutions within 0.5% of the optimal solution. We demonstrate the value of considering substitution explicitly in assortment planning with numerical examples.

The only known structural property of optimal assortments is related to demand rates (van Ryzin and Mahajan, 1999). We extend this result to the shelf space constrained problem. We establish new properties that relate which products are included in the assortment as well as their inventory levels to demand rates and other characteristics of the

products such as per-unit margin, per-unit volume, batch size, and lead time.

We applied our method at a large retail chain in the grocery industry and found the optimal assortment for each store. Comparing results with the existing assortments suggest more than 50% increase in profits.

The contributions of this paper are as follows.

1. We discuss the applicability of existing assortment selection models and propose a methodology that is applicable even for problems with large number of products.
2. We develop an estimation scheme for the substitution rate, which can be used in understanding the roles of categories better and as an input to assortment planning algorithms.
3. We develop an iterative optimization heuristic for a nonlinear, nonseparable, discrete resource allocation problem and assess the performance of the heuristic and establish new structural properties of the optimal assortment based on its solution.
4. We demonstrate the value of taking consumer behavior into account and the value of an operational optimization model in a grand scheme.

All departments in a store influence whether customers shop at that store, hence assortment in a department influences how many customers walk by other departments. In other words, product variety decisions in seemingly independent merchandise categories may interact with each other due to the role of multiple categories in attracting customers together. In the first part of the dissertation, we assume that the original demand to

categories are exogenously given and will not be affected by our decisions. This may not be true in the long run, the decisions across different categories affect the attractiveness of the stores to the customers, and as a result of the decreased number of customers visiting the store, the demand to the retailer's products decreases.

The second part of the dissertation studies the assortment planning problem with multiple merchandise categories. The categories interact with each other due to limited shelf space and the effects of basket shoppers, consumers who purchase from more than one category in a shopping trip. We present a model where the retailer chooses the number of variants to offer in each category and the consumers of multiple categories make their store choice between the retail store and an exogenous outside alternative to maximize their total shopping utility. We investigate the interaction of the category variety decisions at the retail store under centralized and decentralized management regimes. Interestingly, the variety levels of two categories may act as strategic complements or substitutes depending on their levels in comparison to the aggregate outside alternative. They act as complements at low variety levels, whereas as substitutes at high variety levels.

The common practice of category management is an example of a decentralized regime for controlling assortment because each category manager is charged with maximizing profit of his or her assigned category. Hence, a category manager typically optimizes given the assumption that the variety offered by other categories is constant. This assumption can lead the category managers to choose variety levels that correspond to a heuristic equilibrium. We show that there exists at least one heuristic equilibrium (and maybe more) under category management. We show that the optimal variety is never an equilibrium

and the equilibrium variety levels are mostly lower than optimal in at least one category. In a numerical study we demonstrate that the profit loss due to category management can be significant, 28% on average and the category management provides 45% less variety than optimal. More importantly, the performance gets worse as the number of categories increases.

We discuss the tradeoff between the breadth (number of categories) and the depth (variety in each category) of assortments in this setting. We classify categories into two types: *basket categories* have a high co-occurrence rate in consumer baskets with other categories and *independent categories* attract customers on their own. We present conditions where basket categories should be preferred to independent categories in assigning more variety or expanding an assortment.

Realizing the difficulties in implementing the centralized solution, we propose a heuristic that makes the variety decision for a category by not only considering the category profits but also considering its impact on other categories in an aggregate way. The Aggregation Heuristic reduces both data requirements and computation time to polynomial order and recovers most of the loss due to the decentralized regime. We describe the heuristic in detail and discuss its performance.

Major contributions of the second part can be summarized as follows.

1. Variety levels of categories can be both strategic complements and strategic substitutes depending on their levels in comparison to competition.
2. Category management never achieves the optimal solution and mostly provides less

variety than the optimal.

3. We propose a decentralized heuristic that is simple to implement and recovers most of the loss due to category management.

The first two parts of the dissertation concentrate on the retailer's assortment planning problem. A retailer is bound to select its assortment from the products available from the manufacturers. Manufacturers compete with each other through retailers' assortment. Suppliers offer quantity discounts on total sales volume as an incentive for the retailer to carry their line of products rather than the competitors'. The same tactic can be used to manipulate inventory decisions of a retailer who sells one product from each supplier. In the case of two symmetric manufacturers, the retailer ideally would like to choose the same number of products (stocking quantity) from each. With quantity discounts however, the retailer may choose to carry more of one brand (more inventory of one product) and less of the other.

We extend the retailer's assortment planning problem to include the price competition between suppliers, i.e. each supplier use price or discounts as a competitive tool to effect retailer's assortment planning decisions in her favor. Consider the following two-echelon supply network. There are multiple suppliers each selling a single product to retailers selling one product each. Retailers can be under centralized or decentralized control. Multiple retailers under centralized control is equivalent to a single retailer selling multiple products. Suppliers offer contracts to the retailer(s), whose terms are the wholesale prices per unit under a price-only contract and supplier's share of revenues under revenue sharing contracts. Given a set of contracts, retailer(s) choose the stocking quantities of the products

to meet the single period stochastic demand. The products are strategic substitutes, that is demand of a product is negatively effected by other products' stocking quantities. This negative externality between products can be due to consumers switching to other products when their first choice is not available (substitution) or when consumer's choice behavior is effected by the shelf space (i.e., a product with larger presence in the store may get a larger fraction of demand, hence reducing demand to other products). Under different centralization/decentralization configurations we study two contract types: wholesale price contracts and revenue sharing contract types.

We characterize the equilibria under wholesale price contracts and we are able to compare them when products are symmetric. We find that horizontal integration of either echelon (i.e., centralization at the retailer level or the supplier level) increases the wholesale prices and reduces the total stock level in the system regardless of the other echelon being centralized or decentralized. Vertical integration of supplier-retailer pairs on the other hand increases the total stock in the system. Horizontal integration eliminates the competition effect whereas vertical integration eliminates the double marginalization effect. Hence their effect on total inventory is in different directions.

Introducing competition always leads to higher inventories under wholesale price contracts. Competition leads to higher inventory than system optimal when supplier-retailer pairs are vertically integrated, resulting in less profits. However, it may be desirable to introduce competition and increase the inventory in the system. For example when there is a centralized retailer who buys from suppliers with a wholesale price contract, having decentralized suppliers leads to an inventory level that is closer to (but still smaller than)

the system optimal in comparison to having centralized suppliers. The retailer wins in two ways from competition between suppliers: lower wholesale prices and increased sales.

We present many interesting findings regarding revenue sharing contracts. Revenue sharing contracts eliminate the double marginalization effect always in the decentralized retailers case and when revenue share of all products are identical in the centralized retailer case. Surprisingly, it also eliminates the competition effect at the supplier level. Competition leads to more inventories when it is at the retail level, but has no effect when it is at the supplier level. Hence, the system optimal is reached with revenue sharing contracts when there is a centralized retailer and more than system optimal inventory is held when the retailers are decentralized. We also show that revenue sharing contracts leads to higher inventory than wholesale price contracts in any of the configurations. We demonstrate with numerical examples that independent suppliers selling through a centralized retailer are significantly better off under revenue sharing contracts than the wholesale price contracts and the retailer is much worse off.

This paper breaks new grounds in the area of supply chain coordination with contracts and retail inventory management. Following is a brief summary of the differentiating features of our paper.

1. The fundamental difference of our model is that the products are supplied by different suppliers, not by the same supplier. Suppliers compete with each other through a centralized retailer's optimization decisions or equilibria of a noncooperative game between retailers.

2. Contracts serve as a tool for competition rather than achieving system optimality.

There isn't much research on why some contracts are more prominent in practice.

Ours is a first attempt in understanding the impact of different contract types on the profits of the supply chain players.

3. To the best of our knowledge, ours is the first paper where supply chain coordinating contracts are employed in a multiple competing products setting. In supply chain coordination with contracts, usually contract parameters can be arbitrarily chosen where the realistic outcomes are the result of bargaining power or negotiations between the supplier and the retailer. In our model, the contract parameters emerge as the result of the competition between suppliers.

## 1.2 Retailing industry

Retailing is one of the largest industries in the United States, generating 9% of the gross domestic product<sup>1</sup>. According to the National Retail Federation (a worldwide voice for retailers), retailing accounted for \$3.2 trillion in annual sales in the United States alone in 2000, and its 23 million workers is second only to the services sector. It is a very dynamic and complex environment, with a huge number of products, changing customer preferences and different supply chain structures. Many successful retailers follow different paradigms, in terms of product or customer focus: Category killers such as Best Buy and Toy's R Us have been great success stories. On the other hand, companies like Cost-Co and Sam's

---

<sup>1</sup>The largest industries are services and manufacturing, accounting for 40.7% and 16% of the economic output, respectively (Yuskavage and Strassner, 2003).

Club that carry every category under the sun have also been very successful. While Wal-Mart follows a growth strategy to build on its size advantage, Texas based HEB Groceries owes its success to its focusing locally and understanding the needs of the customers better than national chains.

The retail industry has been going under tremendous transformation over the last decades, with changing firm structures, shorter life cycles of products, and less flexible supply chains (i.e., longer leadtimes, outsourced production). Percentage of markup sales increase every year, indicating a mismatch between supply and demand. The recent advance of information technology enabled retailers to collect data at all points of the supply chain, especially point-of-sale data. Making use of this readily available data in rational decision making in choosing the product mix, timings and quantities of purchase orders and transforming the judgemental and experiential decision making practices to analytical, optimization-based systems has been a recent and successful trend in retailing. See Fisher et. al. (2000) for a discussion of the transformation in the retailing industry.

Retailing is an ideal research laboratory. There are clear levers to make changes in the system, such as order quantities, timing of orders, and selection of products. The response time to such changes is shorter than it is in manufacturing organizations. This is due to shorter operating cycles of retailers. For example, financial results of changes in assortment and inventory decisions for a quarter can be observed by the end of that quarter and maybe earlier. On the other hand, product line decisions for a manufacturer involves longer periods of planning and execution. It is also easier to measure the operational effectiveness of retailers, since retailers are mostly investors in a portfolio of inventory. Inventory turns

and gross margin are standard and easy to observe metrics and good indicators of operational performance. In manufacturing operational performance depends on productivity in production processes among other things.

With its momentum of transformation, abundance of data, a need for analytical management, and clear ways to measure impact, retailing industry, presents many challenging opportunities to develop new theories of operations management and apply existing theories.

## Chapter 2

# Literature Review

In this chapter, we describe the background relevant to all three parts of the dissertation.

We also devote a section to review the specific literature in each part.

### **Consumer choice models**

Consider an individual who has to choose a single alternative from a finite choice of alternatives, e.g., {coffee, tea, milk, none of the above} in a given category. Economic theory supposes that the individual has perfect discriminatory power and unlimited information, allowing herself to rank all of the alternatives in a well defined and consistent manner. It is always possible to construct a deterministic utility function that represents the individual's preferences. The individual then chooses the alternative that provides her the maximum utility. From an economic point of view, we are interested in determining the fraction of the population choosing a given alternative. The population can be divided into subpopulations such that each subpopulation is homogenous with respect to the distribution of their

utility from each alternative. That is the utility of a choice alternative to the individuals of a homogenous alternative are independent draws from the identical distribution. The utility to an individual,  $U$ , is decomposed into two parts: one part representing the observable characteristics and the other part representing the unobservable heterogeneity among individuals in the subpopulation. The utility from alternative  $i$  can be written as

$$U_i = u_i + \varepsilon_i$$

It is assumed that members of the subpopulation differ only with respect to the unobservable characteristics ( $\varepsilon_i$ ). Even though the choice behavior of a specific individual is deterministic, the modeler is unable to predict the characteristic of a given individual and she can at best predict the individual's choice up to a probability function. Therefore  $\varepsilon_i$  is modeled as a random variable with a zero mean. This random utility model has another commonly accepted interpretation in which the randomness is due to the psychological theory that an individual can be inconsistent in her choices over time. For an extensive discussion of the two interpretations and comparison of stochastic utility models with other demand generating models, see Chapter 3 in Anderson et. al. (1992).

Depending on the distribution of  $\varepsilon_i$ , closed form formulas can be obtained for the choice probabilities. The most well known is the Multinomial Logit (MNL) formula. Assuming that  $\varepsilon_i$  are i.i.d according to the double exponential distribution, the resulting choice probability for alternative  $i$  is given by

$$p_i = \frac{e^{u_i/\mu}}{\sum_i e^{u_i/\mu}}$$

where  $\mu$  is a positive constant and determines the variance of  $\varepsilon_i$ . We discuss this model in more detail in Chapters 3 and 5.

Using the MNL framework, Guadagni and Little (1983) model the choice between alternatives given that a household purchases something from that category. The observable part of the utility ( $u_i$ ) is specific to different subpopulations (segments) of households and linked to marketing variables such as price, promotion and product characteristics. The model is calibrated with household panel data and proves to have very good explanatory and predictive power. There is an extensive list of papers in marketing that follow up Guadagni and Little's work. Extensions such as Chiang (1991), Bucklin and Gupta (1992), and Chintagunta (1993) are particularly interesting. In addition to the choice between alternatives in a product category, they also investigate whether to buy, and how much to buy decisions of households. In these papers, whether-to-buy decision is modeled as a binary choice between the no-purchase alternative and the resulting utility from the choice between alternatives in that category. We discuss these models formally in Section 3.2.1.

Linear random utility models have proved to be instrumental in estimating consumer demand in academic papers and in practice. An alternative approach that is commonly used in modeling product differentiation is the address approach. Although this approach is commonly used in equilibrium analysis of competing differentiated products (firms), it hasn't been used at all in demand estimation. Address approach can be traced back to Hotelling (1929) and is generalized by Lancaster (1966). A product is viewed as a composition of its attributes and characterized by a vector whose components indicate how much of each attribute is embodied in one unit of it. Consumers have their ideal point in the attribute space and choose the product that is closest to their ideal point. The density of consumers at each point in the attribute space is deterministic and known. This model

have many similarities with the choice models with stochastic utility. For an extensive discussion of the address approach and its relation to stochastic utility models, the reader is referred to Chapter 4 in Anderson et. al. (1992).

The choice models in this dissertation are based on the linear random utility models.

### **Behavioral aspects of variety**

The consumer choice models assume that the customers are perfectly knowledgeable about their preferences and the product offerings. Therefore, a consumers are always better off when they choose from a broader set of products. However, empirical studies show that consumer choice is affected by their perception of variety level rather than the real variety level. The perception can be influenced by the space devoted to a category, the presence or absence of a favorite item (Broniarczyk et. al., 1998), the arrangement of the assortment and the presence of repeated items (Hoch et. al., 1999; Simonson, 1999). Moreover, confusion or complexity due to higher variety may cause dissatisfaction of consumers and decrease sales (Huffman and Kahn, 1998). Kahn (1995) have shown that consumers of certain categories are variety seeking and would like to try different products in repeated purchases, rather than having one favorite product and sticking with it. We don't consider any of these aspects in this dissertation.

### **Models of product differentiation**

Although not directly related to this dissertation, it is worth mentioning that there is a huge economics literature on oligopoly competition between single product firms. Most

are based on the address approach simply referring to Hotelling's work, but also some are based on the stochastic utility models and some on other demand generating models, such as the constant elasticity of substitution (CES) model. The goal is to gain insights from equilibrium analysis regarding location of products and prices. Anderson et. al. (1992) provides an excellent summary and discussion of this literature. Extensions include analysis of number of firms in the market, competition between vertically differentiated products (products with different quality levels), and competition between multi-product firms. The extension most relevant to us is the theory of market segmentation by Moorthy (1984) that provides insights about a monopolist's product line design problem: choosing a subset of products from a vertically differentiated continuum of products and pricing each in a market with different customer classes. Consider cars as a product with a single attribute, say engine size. Monopolist's problem is to choose what size engines to put in the cars and how to price the cars. Still this is assortment planning from the manufacturer's perspective. The results are insightful to pricing policies and cannibalization effects. This theory is also extended to competition between two firms' product lines by Moorthy (1987).



## Chapter 3

# Joint Assortment Selection and Inventory Planning

Merchandising and inventory planning are two fundamental functions in retailing. A merchandising manager's primary responsibility is to customize the assortment at each store by choosing a subset of the products from a chain-wide menu of available products. Customization is necessary because stores differ in size, which effects the shelf space assigned to a category, and the preferences of the customer base. Merchandising managers need to consider the following tradeoff in their decision. With a broader assortment (i.e., more variants in a category), the store carries the favorite product of more customers, but due to limited shelf space<sup>1</sup>, average inventory for the products in the assortment is lower, resulting in lower in-stock rate and more lost sales. It is simply not possible to carry everything and

---

<sup>1</sup>Quelch and Kenny (1994) report that the number of products in the market place increased by 16% per year between 1985 and 1992 while shelf space expanded only by 1.5% per year during the same period.

have them available all the time.

Consider a consumer who can not find her favorite variant, because it is either temporarily stocked out or not carried at the store at all. She can *i*) buy one of the available items from that category (substitute), *ii*) not purchase anything in that category, but does the rest of her shopping (lost sales to the category), *iii*) decide to shop at another store (lost customer). We focus on the first two alternatives because of the following reasons. The second and third alternatives are the same from a category sales perspective. A consumer's shopping basket typically includes many items and it is reasonable to assume that absence of one variant in one category will not drive the customer away. This is especially true for the grocery industry.

A consumer's willingness to accept a substitute has strong implications for assortment planning. In the extreme case i.e., full substitution, the optimal assortment is to carry one item and benefit from inventory pooling effects. In the other extreme, i.e., no substitution, demand of products are independent from each other. If there is some level of substitution, demand for each product depends on the availability of all the variants, which depends on whether a variant is in or out of the assortment, and its inventory level.

We start with analyzing the case of no substitution, i.e., the demand of each product is independent of the demand and availability of other products. The no-substitution model serves as a benchmark and a basic building block of the model with substitution. In this chapter, we develop a model of demand estimation in the no-substitution case using daily sales data at the store-product level. Then, we describe an assortment optimization algorithm that selects optimally products and their inventory levels at each store. This

optimization tool accommodates many realistic constraints such as batch sizes, delivery lead times, facing sizes and perishability of products. In the next chapter, we discuss assortment planning under substitution and an application of our methodology to a large retail chain.

### **3.1 Related literature**

We review the literature related to Chapters 3 and 4 in this section. Two types of substitution needs to be defined before starting. Assortment-based substitution is the switch to an available variant by a consumer when her original favorite product is not carried in the store. That is the substitution is due to the assortment decision. Stockout-based substitution is the switch to an available variant by a consumer when her original favorite is carried in the store, but it is stocked-out at the time of her shopping

#### **Multi-item inventory models**

Multi-item inventory problem under a single resource constraint is studied extensively in the operations literature and a solution using Lagrangian multipliers is presented in various textbooks, e.g., Hadley and Whitin (1963), Nahmias and Schmidt (1984). Downs et. al. (2001) describe a heuristic approximation to the multi-period version of this problem with lost sales. In these models, the demand of products are not dependent on others' inventory levels (i.e., no substitution), and the inventory policy is a rather simple base stock policy.

There are several inventory models with multiple products and stock-out based substitution, first introduced by McGillivray and Silver (1978). Netessine and Rudi (2003)

investigate the case with  $n$  products under centralized and decentralized regimes. Substitution is modeled by letting the unsatisfied demand for a product flow to other products in deterministic proportions. Decentralized regime carries more inventory than the centralized regime. Parlar (1985) and Avsar and Baykal-Gursoy (2002) study the infinite horizon version of this problem under centralized and competitive scenarios respectively. These inventory models focus on the stocking decisions given a selection, but not the selection of the products. In addition, these models assume unconstrained resources. Even in this case, it is not possible to obtain an explicit solution to the problem. A final difference is that our inventory model is a rather complex multi-period problem with positive leadtime, batch deliveries and perishability.

In multi-item inventory literature, it is common to model the demand of products as a multi-variate random process with given correlation structure. With substitution, demand of products depend on each other based on the chosen order quantities rather than an exogenous correlation structure.

### **Joint assortment selection and inventory planning**

Assortment planning is defined as the selection of the products in a category and setting the inventory levels for each product. van Ryzin and Mahajan (1999) establish structural properties of the optimal assortments in a stochastic single period model with a utility-based choice (MNL). In their model, consumers can substitute if their favorite variant is not carried in the store (assortment-based substitution), but the sales is lost if their favorite variant is carried but temporarily unavailable (no stockout-based substitution).

They prove that the optimal assortment always consists of a certain number of the most popular products. In other words, optimal assortment is one of  $\{1\}, \{1, 2\}, \dots, \{1, 2, \dots, n\}$ , where products are indexed from most popular to least popular. This result reduces the search space considerably, from  $n!$  possible assortments to  $n$ . Since they consider only assortment-based substitution, for each of these  $n$  assortments, it is easy to compute the resulting demand rates for the products that are 'in' and solve for the newsvendor quantities. Mahajan and van Ryzin (2001) develop a stochastic sample path optimization method for the same model under consumer substitution that is both assortment-based and stockout-based. Both these papers do not consider resource constraints such as limited shelf space or budget. Agrawal and Smith (2000) study the assortment planning problem with a hybrid substitution structure and multi-period base-stock inventory model. The solution methodologies in the last two papers have limited applicability for large problems and the parameter estimation has not been discussed. We discuss these papers in more detail in Section 4.1.

Rajaram (2001) studies a single period assortment planning problem in fashion retailing. He develops a heuristic based on Lagrangian relaxation, which has an average performance gap of 1%. He does not consider substitution between products. Jayaraman, et. al. (1998) optimizes assortment and inventory decisions under assortment-based substitution (when favorite variant is not carried in the store). They use the deterministic EOQ model for each product as a building block of a nonlinear program.

## **Consumer response to stockouts**

Although assortment or inventory planning with substitutable products attracted some attention in the literature, there are few papers that study the estimation of the substitution behavior. Consumer response to stockouts are shown to be negative from a brand perspective (Schary and Christopher, 1979) as well as a category sales perspective (Bell and Fitzimons, 1999). Anupindi et. al. (1998) describe a method for estimating consumer demand with stockout-based substitution (when favorite variant is temporarily unavailable). They assume a Poisson arrival process for all the products. Based on the sales data and the stocked out items, they write down the likelihood function in a parameterized way. Finally they solve for the arrival rates and substitution probabilities by maximizing the likelihood function. Their model requires partial or complete inventory data. Our estimation procedure is for assortment-based substitution, which enables us to estimate the demand rates of products in a store that are never carried in that particular store. To our knowledge, there is no other reported research on estimation of substitution behavior.

## **Resource allocation models**

The simplest resource allocation problem is the Knapsack problem with a linear objective function and continuous decision variables. The nonlinear integer problem is of interest to us in this paper. It can be transformed to a linear 0-1 integer problem if the cost function is increasing and concave (Hochbaum 1995). The resulting linear knapsack can be solved using one of many algorithms available, e.g., Martello and Toth (1990).

The objective function that arise in the inventory model with case packs and perish-

bility is not necessarily concave. The integer knapsack problem with nonconvex objective function is usually solved with branch and bound methods and using convex functions underestimating the nonconvex function for lower bounding purposes. See Hurst and Tuy (1990) and Benson et. al. (1990) for algorithms and convergence results. Also see Bretthauer and Shetty (2002) for a review of nonlinear resource allocation problems.

Due to substitution, our objective function is a nonseparable function of the inventory levels. A few papers studied the nonseparable problem with 0-1 variables (Gallo et. al. 1980, Caprara et. al. 1999) by developing methods that directly solve the nonseparable problem. Dussault et. al. (1986) and Klastorin (1990) approximate the nonseparable problem with a series of separable problems. At each node of the branch and bound method, the relaxed nonseparable problem is solved via an iterative gradient based method. Our approach is different. At each iteration, we approximate the nonseparable function with a separable function (based on the current solution) and find the next integer solution.

### **Shelf space allocation models**

In an influential paper Corstjens and Doyle (1981) suggest a method for allocating shelf space to categories. They perform store experiments to estimate the sales function of the form  $\alpha_i s_i^\beta \prod s_i^{\delta_{ij}}$ , where  $s_i$  is the space allocated to product  $i$ ,  $\beta_i$  is own space elasticity, and  $\delta_{ij}$ 's are the cross-space elasticities. Cost functions of the form  $\gamma_i s_i^{\tau_i}$ , are also estimated from the experiments. The details of the experiments are not described. The problem of profit maximization with a shelf space constraint is solved within a geometric programming framework. Their results are significantly better than commercial algorithms that allocate

space proportional to sales or to gross profit by ignoring interdependencies. The estimation and optimization procedures can not be applied to large problems, hence they elect to work with product groups rather than SKUs. Bultez and Naert (1988) apply the Corstjens and Doyle (1981) model at brand level assuming symmetric cross elasticities within product groups (similar to symmetric substitution within subcategories in our model). Their model is tested at four different Belgian supermarket chains, leading to encouraging results. These two papers do not explicitly model the assortment selection or the inventory side of the problem. They work with the sales function both in estimation and optimization.

An interesting paper by Borin and Farris (1995) report the sensitivity of the shelf space allocation models (a variant of Corstjens and Doyle, 1981) to forecast accuracy. They compare the solution with correct parameters to that with incorrect parameter estimates. Even when the error in parameter estimates are 24%, the net loss in category return on inventory is just over 5% compared to the optimal allocation based on true estimates. This proves the robustness of these models to estimation errors. Similar to the shelf space allocation papers, but with an inventory theoretic perspective, Urban (1998) models the own and cross product effects of displayed inventory on the demand rate in a mathematical program and solves for shelf space allocation and optimal order-up-to quantities. He reports that on average a greedy heuristic yields solutions that are less than 1% worse than a solution obtained by genetic programming.

## 3.2 Problem structure

The number of products carried in a retail store can be very large. In the grocery industry, supermarkets often carry more than 30,000 stock keeping units (SKUs). At the top level of the hierarchy SKUs are divided into three groups, chilled products, dry goods, and groceries. Each group then is divided into merchandising categories, such as wines, bread spreads, butter & margarine, etc. There may be multiple subcategories of products in a given category. A subcategory is defined as a group of variants such that the difference between products within a subcategory is minimal but the difference between subcategories is significant. For example, the subcategories in the butter & margarine category include deep-fry fat, regular butter, healthy butter, margarines, etc.

### 3.2.1 Modeling consumer demand

Our model of consumer purchase behavior is based on three decisions, (1) whether or not to buy from a subcategory (purchase incidence), (2) which variant to buy (choice) given purchase incidence, and (3) how many units to buy (quantity). This hierarchical model of choice is similar to Bucklin and Gupta (1992) that models the first two decisions with an additional focus on the segmentation of customers and Chintagunta (1993) that models all three decisions. Both papers work with household panel data, whereas we work with daily sales data<sup>2</sup>. The hierarchy in the model does not assume that consumer decisions

---

<sup>2</sup>There is a recent stream of marketing papers that work with sales data. Nijs et. al. (2001) and Pauwels et. al. (2002) study the medium and long term effect of price promotions with vector-auto-regressive models. Pauwels et. al. (2002) models purchase incidence, choice and quantity decision to be interdependent and

necessarily proceed in that order. Recall that in this chapter we assume that there is no substitution by consumers. We want to find out the demand vector given that all customers buy their favorite variant.

We start with describing the *choice* model. Suppose the set of variants in a subcategory  $s$  is denoted  $N = \{1, 2, 3, \dots, J\}$  and store  $i$  carries an assortment  $S_i \subseteq N$ . For now, we don't include the subscript  $s$  denoting the subcategory, because the demand estimation takes place within a subcategory and is totally independent from other subcategories. Later in the optimization part, we start including the subcategory subscript. Given the purchase incidence, i.e., given that a customer has decided to purchase from a subcategory, a customer chooses to purchase one of the variants in  $S_i$ . We use the multinomial logit<sup>3</sup> (MNL) to model this choice process. MNL models consumers as rational utility maximizers. Each customer visiting the store associates a utility  $U_j$  with variants  $j \in S_i$ . A customer chooses the variant with the highest utility among the set of available choices,  $\{U_j : j \in S_i\}$ .

There are two main characteristics of the utility-based choice behavior. First, the utilities are not observable by the retailer; we can only observe the resulting choice behavior. solves the three regression models simultaneously. Our model is a special case of their model : no auto-regression and no interdependence between three models. We choose to work with the simpler model for the following reasons. Our data set is from a much shorter time period and we are not interested in the extent of covariation in the linked time series. These papers work with aggregate data (week-brand level) to be able to solve the simultaneous problems, whereas we estimate sales at day-SKU level. Finally, our contribution is more in the use of the estimates rather than the estimation scheme per se.

<sup>3</sup>MNL is widely used in economics and marketing literatures. See for example Guadagni and Little (1983) and a whole stream of following papers for applications in marketing, Ben-Akiva and Lerman (1985) for applications in travel industry, and Anderson et.al. (1992) for MNL based models of product differentiation.

Second, the utility of the same variant may be different for different consumers due to heterogeneity of preferences. MNL models these characteristics as random components in the utility values. The utility is decomposed into two parts, the deterministic component of the utility  $u_j$  (that can be decomposed further to understand the effects of different explanatory variables) and a random component  $\varepsilon_j$ . The random component is modeled as a Gumbel random variable. Also known as Double exponential or Extreme value Type-I, it is characterized by the distribution  $Pr\{X \leq \varepsilon\} = \exp(-\exp - ((\varepsilon/\mu) + \gamma))$ , where  $\gamma$  is Euler's constant (0.57722). Its mean is zero, and variance is  $\mu^2\pi^2/6$ . A higher  $\mu$  implies a higher degree of heterogeneity among the customers. The Gumbel random term assumption proves to be very useful in leading to a simple form of choice probabilities. The probability that  $j^{th}$  variant will be chosen by a customer in subcategory  $s$  is

$$p_j = \frac{e^{u_j/\mu}}{\sum_{k \in S_i} e^{u_k/\mu}} \quad (3.1)$$

Notice that increasing the deterministic utility of a variant  $k$  will cause a cannibalization of sales from all other variants,  $j \in S_i \setminus \{k\}$ . A known weakness of the MNL model is its Independence of Irrelevant Alternatives (IIA) property. This property says that the ratio of choice probabilities of two alternatives is independent of the other alternatives in the choice process. IIA property should not hold in cases where there are subgroups of similar products in the choice set, for example when choice is made within a brand or size rather than the whole group. Consider two products from different brands. If brand loyalty is important, adding a new product from the first brand can cannibalize the sales of its sister product more than the rival product. IIA does not capture this important aspect of consumer choice. In marketing, MNL model is shown in numerous papers to work very

well despite its IIA property. Using nested logit models (first brand then SKU or first size then SKU) is one way to eliminate this problem. However, it requires the knowledge of key attributes and their hierarchy for the consumers and makes the estimation problem much harder.

Note that our model assumes that a customer chooses one product only from a subcategory. There are cases where customers choose a bunch of products (e.g., different flavors of yogurt). Since we are working with sales data, we only see the total sales for each product. From the estimation model's point of view, one customer who bought two flavors of yogurt is equivalent to two customers who bought one flavor each.

The *purchase incidence* model explains the probability that a customer in the store buys anything from the category of interest. Let  $V$  be the utility of a customer for making a purchase in subcategory  $s$ .  $V$  is also composed of two parts, a deterministic part  $v$  and a random component that is distributed according to logistic distribution. The alternative to making a purchase is the no-purchase option that will yield a utility of zero. The utility maximizing customer will choose to purchase from the subcategory if  $V > 0$ . This leads to the simple binary logit function. Namely, the probability of purchase incidence in subcategory  $s$  is

$$\pi = \frac{e^v}{1 + e^v} \quad (3.2)$$

We model the regression equations for estimating  $V$  by using the sum of the variables that make up the  $u_j$ 's. The details can be seen in the next section. Bucklin and Gupta (1992) use a nested structure to estimate  $V$ . Theoretically,  $V$  is the utility from the chosen variant in the choice process, that is  $V = \max(U_j : j \in S_i)$ . Since the Gumbel distribution

is closed under maximization, the maximum utility - from the chosen variant- is also distributed Gumbel with mean  $\ln \sum_{j \in S_i} e^{u_{jit}}$ , (Ben-Akiva and Lerman, 1985). Bucklin and Gupta (1992) use this expression as one of the explanatory variables in  $V$  and this nested structure necessitates a sequential procedure for estimation of *choice* and *purchase incidence* models. Our approach doesn't use the nested structure but tries to imitate it by using the variables that go in the  $u_j$ 's in estimating the  $V$  as well, so that when  $u$ 's increase so does  $V$  (as it would in the nested model).

Also note that both probability expressions in (3.1) and (3.2) are the same for all customers at a given date and store. We do not explicitly consider heterogeneity in customers' deterministic utilities because we are working with sales data rather than household panel data. As it is very common in papers analyzing panel data, it is easy to adopt the above formulas to the case of multiple classes of consumers.

Finally we estimate the average number of units a customer buys from the product she chose ( $q_j$ ) using an ordinary linear regression. The details of which is coming in the next section.

$\pi p_j$  is the probability that a customer will purchase the  $j^{th}$  variant in subcategory  $s$ . Combining this with the average quantity of units that a customer will buy from product  $j$  ( $q_j$ ), we can estimate the average demand a customer will generate. Let  $(PQ)_j$  denote this quantity. Clearly,

$$(PQ)_j = \pi p_j q_j$$

and the demand of product  $j$  at store  $i$  on day  $t$  is

$$D_{jit} = N_{it}(PQ)_{jit}$$

where  $N_{it}$  is the estimate of number of customers that will visit store  $i$  at a given day  $t$ .

Variety level has two implications in this model. Once the  $p_j$ 's are estimated, dropping products from the assortment  $S_i$  will result in not meeting the demand of those products (due to no-substitution assumption) because probability that a sales is made in subcategory  $s$  is  $\sum_{j \in S_i} p_j$ . Second, in a nested model,  $V = \max(U_j : j \in S_i)$  also increases with number of products, hence penetration to the category increases with variety.

### 3.2.2 Estimation of Demand

In this section we describe the mechanics of the estimation of demand based on the model in the preceding section. Specifically, we describe the regression equations for estimation of  $\pi_s$ ,  $p_{sj}$ , and  $q_{sj}$  for all  $j$  in a subcategory. We use the following notation throughout the paper. Subscripts  $s$  is dropped because the models are specific to each subcategory.

$$i \in Stores = \{1, \dots, K\}$$

$$t \in Dates$$

$$s \in Subcategories$$

$$j_s \in SKUs \text{ in } s = \{1, \dots, J_s\}$$

$S_i$  : the set of variants in subcategory  $s$  at store  $i$

$$D_l^t = \begin{cases} 1 & \text{if observation } t \text{ is from day } l \text{ of the week, } l=1..6 \\ 0 & \text{otherwise} \end{cases}$$

$$A_{sjit} = \begin{cases} 1 & \text{if there is a promotion of product } j \text{ of subcategory in store } i \text{ at date } t \\ 0 & \text{otherwise} \end{cases}$$

$$E_t^l = \begin{cases} 1 & \text{if event } l \text{ is on date } t \\ 0 & \text{otherwise} \end{cases}$$

There are 13 events such Christmas, Easter, etc. and one generic holiday dummy (i.e.,  $E_t^0 = \sum_{l=1}^{13} E_t^l$ ).

$T_t$  = temperature for period  $t$

$HDI_t$  = Human discomfort index for period  $t$ . It is a linear combination of temperature, humidity, wind chill, and hours of sunshine.

$\bar{A}_{it} = \frac{\sum_{j \in S_i} A_{jit}}{|S_i|}$ , average promotion activity in the subcategory  $s$  in store  $i$  in period  $t$ .

### Purchase Incidence (Subcategory Penetration) Model

The data we have is sales history for all products and the number of customers visited each store on a given date. Using this data, we compute the total sales in a subcategory. Dividing that by the number of customers gives us the observed penetration rate to the subcategory. The regression equation that is used for estimating the purchase incidence is obtained by inverting the binary logit function given in (3.2).

$$\ln \left( \frac{\pi_{it}}{1 - \pi_{it}} \right) = v_{it} = \alpha_{0i}^\pi + \alpha_{1i}^\pi T_t + \alpha_{2i}^\pi HDI_t + \sum_{l=1}^6 \gamma_{il}^\pi D_t^l + \alpha_{4it}^\pi \bar{A}_{it} + \sum_{l=1}^{14} \beta_{il}^\pi E_t^l \quad (3.3)$$

## Choice Model

Cooper and Nakanishi (1984) prove that using a log-centered transformation of data is equivalent to the original MNL formula given by (3.1). They also suggest the use of this method when the number of choice experiments in the data set is large, because the original method requires a dummy variable for each choice experiment (days in our case).

$$\ln \left( \frac{p_{jit}}{\bar{p}_{it}} \right) = u_{jit} = \delta_{jik}^C + \sum_{l=1}^6 \theta_{ik}^C I_{jk} HDI_t + \alpha_{1i}^C (R_{jit} - \bar{R}_{it}) + \alpha_{2i}^C (A_{jit} - \bar{A}_{it}), \text{ for all } j \in S_i \quad (3.4)$$

where

$$\begin{aligned} \bar{p}_{it} &= \left( \prod_{j \in S_i} p_{jt} \right)^{\frac{1}{|S_i|}}, \text{ geometric mean of } p_{jit} \text{'s} \\ I_{jk} &= \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}, \text{ variant dummy} \\ R_{jit} &= \text{selling price of article } j \text{ in store } i \text{ in period } t \\ \bar{R}_{it} &= \frac{\sum_{j \in S_i} R_{jit}}{|S_i|}, \text{ average price in the subcategory} \end{aligned}$$

Intercepts are SKU specific. Notice that the price and promotion coefficients are common to all the SKUs in a subcategory. In a choice model, if the explanatory variable has the same value for all the choice alternatives (e.g., weather, holiday event, etc.), the coefficients have to be specific to each SKU, otherwise it will not effect the shares at all, because an increase of same magnitude in utilities of all alternatives cancel each other out in the share formula given by (3.1). Therefore, we need SKU specific coefficients for event and HDI variables, which increases the number of explanatory variables.

## Units Models

From the data, we observe the total number of units sold in a day and total number of customers who purchased them, but we don't know how many units each customer purchased. In other words, we know the average number of units that customers bought, but not the individual observations. Commonly in marketing, variables such as number of units purchased or number of web pages visited by a customer are modeled using Poisson regression. Since we don't have the individual observations, we can not follow the same approach. We fit the following ordinary linear regression to the data of average units purchased for each SKU-store combination.

$$q_{jit} = \alpha_{0ji}^Q + \alpha_{1ji}^Q A_{jit} + \alpha_{2ji}^Q HDI_t + \sum_{l=1}^{14} \beta_{jil}^Q E_t^l, \text{ for all } j \in S_i \quad (3.5)$$

### 3.2.3 Description of the inventory system

In this section, we describe the specifics of an inventory system that is typical in the grocery industry. Different inventory systems can easily be incorporated into our methodology following a similar approach.

All the products in a category are subject to the same delivery schedule and fixed leadtimes. There is no backroom, therefore orders are directly delivered to the shelves. Orders are delivered in batches since the cases that come from the wholesalers are never broken down until they arrive at the shelves. Shelves are divided to facings. SKUs in a category share the same shelf area but not the same facing, i.e., only one kind of SKU can be put in a facing. Capacity of a facing depends on the depth of the shelf and the unit

volume of a SKU. The decision variable is the number of facings for each SKU, which is effectively a discrete version of the maximum inventory level.

Since shelf space is very valuable, the replenishment system tries to fill the facings assigned to a SKU, therefore there is no minimum order quantity. However this doesn't mean that an order will be placed at every ordering time for each SKU. The system is constrained by the case sizes, an order can be placed only if there is enough room for one case of the product. Often the case size is larger than the capacity of a facing and an order can not be placed even if one facing is totally empty.

Consider a periodic review stochastic inventory model with lost sales. There is no closed form solution to such models due to the lost sales assumption. Moreover, we want to incorporate the additional operational constraints such as constant delivery lead-time and batch deliveries, because capacity of facings and case sizes vary significantly across products and significantly affect the returns from inventory. Another operational reality is the perishability of some products, such as bread, milk, etc. These products simply go bad after a certain shelf life and have to be salvaged. Below is a formal description of our inventory model. The subcategory subscript  $s$ , store subscript  $i$  and date subscript  $t$  are dropped for convenience.

Each product has a certain daily demand rate,  $d_j$ , number of products per facing (facing capacity),  $f_j$ , and case size,  $b_j$ . We need a few definitions to describe how the inventory model works. At each review point an ordering decision is made. We define a period as the time between two review points. Inventory level ( $IL$ ) is on-hand inventory. Inventory position ( $IP$ ) is the inventory level plus the orders placed but not delivered yet.

Each product in the category is assigned a number of facings,  $n_j$ , which corresponds to the maximum level of inventory,  $n_j f_j$ . At a review moment, if the inventory position is lower than the reorder point  $(n_j f_j - b_j)$ , an integral number of cases is ordered to take the inventory level as close as possible to the maximum without exceeding it. In other words, the number of cases to be ordered ( $k$ ) is chosen so that inventory position ( $IP$ ) rises back to the  $(n_j f_j - b_j, n_j f_j]$  interval. Mathematically,  $k^* = \min\{k : IL + kb_j > n_j f_j - b_j\}$ .

Figure 3.1 illustrates a possible realization of inventory level of an imperishable product for four review periods. One case of the product is ordered in periods 1 and 2. An order of two cases is placed on period 3 and no orders are placed in period 4. It is also seen that the item is out of stock for a while during period 3.

Since we are not able to obtain closed form or approximate expressions for this complicated inventory model, we use simulation to estimate the expected gross profit for given number of facings. Gross profit is defined as sales times unit margin minus the cost of disposed products (if the product is perishable). The simulations results are averages from multiple replications. In each replication, first 10 periods are the warm-up periods and statistics are recorded for the last 250 periods. We continue to replicate until a 95% confidence interval is reached.

### 3.2.4 Assortment optimization

We obtain the expected profit versus facings curves for each SKU via simulation. The objective is to assign number of facings to each product to maximize total expected gross profits from one or more categories subject to a shelf space constraint. We obtain the

optimal number of facings by solving the following mathematical program

$$\begin{aligned}
& \max \quad \sum_s \sum_j GP_{sj}(n_{sj}) \\
& s.t. \quad \sum_s \sum_j n_{sj} w_{sj} \leq ShelfSpace \\
& \quad \quad n_{sj} \in \{0, 1, 2, ..\}, \text{ for all } s, j
\end{aligned} \tag{AP0}$$

where  $w_{sj}$  is the width of a facing for SKU  $j$ .

This is basically a knapsack problem with a nonlinear objective function. The optimization is done by an enhanced greedy heuristic and the solution is improved locally by searching profitable facing exchanges between SKUs. The enhancement to the greedy heuristic is as follows. Some of the SKUs have case sizes larger than the capacity of their facings. If only one facing is assigned to such items, the reorder point is negative, so no inventory is carried and sales of the item is zero. Nevertheless, second or third facings can create positive returns. Regular greedy heuristic never assigns the first facing to these products because the marginal benefit of the first facing is zero. Also for perishable products or products with high demand rate, gross profit may be nonconcave due to the discrete case size - facing size relations: If the inventory level is frequently zero at the end of the period, (either because unsold units are salvaged or all inventory is sold), and the number of cases that fit in the given space doesn't increase with an assigned facing, that last facing may not be effective. The gross profit will be like a step function that jumps every time a new case is able to fit to the space defined by number of facings. The enhanced greedy heuristic overcomes these problems by also considering two-step jumps. The performance of the suggested heuristic versus the optimal solution is also included in the discussion of

the results.

The objective function is not necessarily concave, because the individual gross profit curves can be quasi-concave. The resulting mathematical program has a nonlinear and nonconcave objective function with integer decision variables. Our proposed heuristic is not guaranteed to find the optimal solution. Bretthauer and Shetty (2002) reviews branch and bound based methods that are guaranteed to find the optimal solution of nonconcave, separable problems.

### **3.3 Application**

Albert Heijn is the leading supermarket chain in the Netherlands. They are known for their high quality, high variety, and reasonable prices. The stores are smaller than regular US supermarkets. The customer base is also different than typical US customer. Customers visit the stores very often and usually walk or bike to the store, therefore the size of the shopping baskets are smaller.

At Albert Heijn, each category of products belongs to a delivery group. Each delivery group has a fixed schedule of delivery times for every store. The replenishment orders has to be placed 18 hours earlier for daytime deliveries and 9 hours earlier for evening deliveries. The backroom is used for only a few bulky and fast moving products, and orders are delivered directly to shelves.

As mentioned before, our methodology is custom designed for the grocery industry. The inventory model and the optimization algorithm can be adopted to any other system easily.

The estimation model we developed relies on the following key characteristics of the grocery industry and particularly Albert Heijn.

- At Albert Heijn, assortment changes are done twice a year. Therefore, at a particular store, all consumers see the same assortment (choice set) throughout the data collection period. This can be fixed by directly maximizing the likelihood function rather than using the log-centered transformation.
- The service levels are quite high, so that lost sales is negligible, i.e., demand is equal to sales.
- We know the number of customers that made purchases from all products in a subcategory, the total units sold and the total number of customers who made transactions at the store on a given day. In the grocery industry it is safe to assume that the customers who made transactions in the store is a good proxy for the number of customers that visited the store. This may not be applicable to other industries. Due to these characteristics,  $D_{sjit} = N_{it}\pi_s p_{sj} q_{sj}$  decomposition is possible. In this study, we assume that we can estimate  $N$  accurately. Simple regression equations yield very good predictions. Albert Heijn is able to estimate  $N$  with remarkable accuracy (with less than 1% error) using auto-regressive techniques.

### 3.3.1 Data

Sales data is collected through a period of 20 weeks from 37 stores for seven merchandise categories. The categories are cereals, bread spreads, butter & margarine, canned fruits,

canned vegetables, cookies, and banquet sweets. There were 114 subcategories in these seven categories. The number of subcategories and SKUs in each category can be seen in Table 3.1. The size of subcategories varies from 1 to 29 SKUs, with an average of 7.7 and a standard deviation of 5.7.

### 3.3.2 Assessing the quality of the estimation

Note that we applied each of the three regression models (*purchase incidence*, *choice* and *quantity*) to 4218 store-subcategory combinations. Choice and quantity models involve  $|S_i|$  equations for each subcategory. The final forms of the regression models are decided based on a preliminary analysis. In the preliminary analysis we included 3 stores and 12 subcategories. We kept the variables that are significant most of the time and dropped the others. We also dropped week dummies, although they were significant all the time. Since we had only 20 weeks of data, the week variable was creating high levels of collinearity with holiday and weather variables. The seasonality effects are captured by weather and a generic holiday dummy.

Below we report the results for 34 SKUs from 3 subcategories at all 37 stores. For each SKU, we compute mean absolute deviation (*MAD*) and average bias of each store model and report the average across stores. The models are calibrated on data from 15 weeks and five weeks is hold out as a test sample.

### **Purchase Incidence ( $\pi_{sit}$ ) Rates**

Table 3.2 below summarizes the average performance of the purchase incidence model on 37 stores. Average bias is low and  $MAD$  is around 22%, which is fairly low for estimating the penetration to subcategories at store-day level.

### **Conditional Choice Probabilities ( $p_{sjit}$ ) using weekly shares**

Table 3.3 presents the average accuracy of the weekly share estimates for the SKUs in subcategory 10 across all stores. The range of  $MAD$  is 15-53% for 8 of the 9 SKUs in subcategory 10. We are better at estimating SKUs with higher shares. Notice also that bias is negative for lower share SKUs and positive for higher share SKUs. The results are similar for the other subcategories.

### **Quantity model ( $q_{sjit}$ ) using weekly sales data**

In the quantity model the variance of the observations is very small and  $R^2$  values for the regression equations range mostly from 5%-50%. However, the model's predictive power is much better. Range of  $MAD$  is 9% to 20% and average bias ranges from 0 to 1% for fit and test samples.

### **Analysis of $(PQ)_{sjit}$ on daily data**

Notice that we used the weekly shares within subcategories and weekly average quantities for calibration. Now we combine the three models to estimate the  $(PQ)$  values at day-store-SKU level. Each point in Figure 3.2 represents a SKU. Data is from the fit sample.

Notice that the accuracy of the estimation is higher for SKUs with higher shares. *MADs* for most SKUs are between 40%-80%. This is a remarkably successful accuracy for demand rates with such high variance (due to events and randomness). For example, coefficient of variation of  $(PQ)$  for SKUs in subcategory 10 ranges from 0.53 to 1.35, and it is larger than 1 for 6 out of 9 SKUs.

On average across store-SKUs, we observe the following.

- *MADs* are 5% higher on the test samples (compared to fit samples)
- On the fit sample, bias ranges from -50% to 12%. 19 out of 34 SKUs within -10% to 10% range.
- On the test sample, bias ranges from -35% to 38%. 22 out of 34 SKUs within -10% to 10% range.
- Large negative bias is observed mostly for small share items.

### **Analysis of daily demand estimation $D_{sjit}$**

Since Albert Heijn is able to estimate the number of customers visiting the store with remarkable accuracy ( $R^2 > 99\%$ ) using autoregressive models, the quality of estimation of  $D_{sjit}$  closely follows the one of  $(PQ)_{sjit}$ .

### **Comparison with other models**

The hierarchical model we described basically estimates  $(PQ)$  in three steps. One might argue that instead of running three models and introducing estimation errors multiple times,

it could be better to directly estimate  $(PQ)$ . We can simply estimate the daily  $(PQ)$  values with the average  $(PQ)$  values across stores for each SKU and benchmark our results. The averages yield 10%-20% higher  $MAD$  than our model for the SKUs in subcategory 10 across all stores. A more sophisticated approach is what has been done in the past at Albert Heijn. We compare the results of our model with an improved version of the Albert Heijn model. The improvement to Albert Heijn model is including a cross promotion variable to capture the impact of promotion activities of other SKUs in the subcategory. We again drop subscript  $s$  for convenience.

$$(PQ)_{jit} = \alpha_{0ji} + \alpha_{1ji}T_t + \alpha_{2ji}HDI_t + \sum_{l=1}^6 \gamma_{jil}D_t^l + \alpha_{4ji}\bar{A}_{jit} + \sum_{l=1}^{14} \beta_{jil}E_t^l \quad (3.6)$$

$$\ln\left(\frac{PQ_{jit}}{1 - PQ_{jit}}\right) = \alpha_{0ji} + \alpha_{1ji}T_t + \alpha_{2ji}HDI_t + \sum_{l=1}^6 \gamma_{jil}D_t^l + \alpha_{4ji}\bar{A}_{jit} + \sum_{l=1}^{14} \beta_{jil}E_t^l \quad (3.7)$$

where

$$\bar{A}_{jit} = \sum_{k \in S_i, k \neq j} A_{kit}, \text{ total promotion activity in the subcategory except SKU } j$$

Equation (3.7) is a logistic regression version of equation (3.6). Let *directPQ* denote model (3.6) and *logitPQ* denote model (3.7). In these models, the cross promotion variable is the only link between SKUs in a subcategory. Recall that in the *choice based* model, all independent variables of each SKU impact the sales of all SKUs in the subcategory through the MNL formula.

Table 3.4 presents average  $MAD$  and bias figures of the three estimation models across all stores and SKUs. *logitPQ* model is clearly outperformed by the other models in terms of both  $MAD$  and bias. In the fit sample, *logitPQ* yields a 4% less  $MAD$  than the choice model. However, the lower  $MAD$ s all come from the low share SKUs. So the choice model

has predicted the larger share SKUs better in the fit sample case too. We observe that average *MADs* of the choice model and the *directPQ* model are similar in the fit sample and 4% lower by the *directPQ* model in the test sample. However, average bias results for the test sample favor the *directPQ* model.

There are two major problems with the *directPQ* model. First, it is a linear model, therefore can not model the diminishing rates of return on prices, weather, etc. Second, the predictions can be negative. For example, approximately 5% of the predictions for subcategory 73 are negative.

We conclude that the *choice based* model is better than the (*PQ*) models in terms of quality of the estimation. The theoretical advantages of the *choice based* model can be listed as follows.

- *Choice based* model imitates customer's choice behavior in a subcategory: It is consistent with the idea of a subcategory (that the items are substitutable, SKUs can cannibalize each other's shares).
- The interpretation of own and cross elasticities are clear.
- In contrast to the *directPQ* and *logitPQ* models, demand for SKUs in a subcategory are not independent.



## Chapter 4

# Assortment Planning under Demand Substitution

Recall that a consumer who can not find her favorite variant can *i*) buy one of the available items from that category (substitution), *ii*) not purchase anything in that category, but does the rest of her shopping (lost sales to the category). This chapter introduces a methodology for discovering the rate of substitution inherent to a category and incorporating that into demand estimation and assortment optimization.

Although there had been significant research in inventory and assortment decisions under substitution, there is little research on the estimation of input parameters. We introduce a relatively simple model of demand substitution and discuss its strengths and limitations. We present a novel methodology for demand estimation under the assumed substitution structure. We use data from multiple stores with different assortments to find out how much more an assortment really sold than its contribution at stores with full assortment.

For a given rate of substitution, we can also theoretically compute this difference. Combining these, we find the substitution rate that minimizes the error between the theoretical and observed differences. Using the substitution rate and the observed demands, we can compute the demand of each product at each store with reverse engineering. We obtain a chain-wide substitution rate and for each store, we obtain demands of products carried in the store as well as products previously not carried in the store.

Under substitution, the effective demand to a product depends on the inventory levels of other products because of the spill over from the unavailable products. Hence the assortment optimization problem is a nonlinear, nonseparable, discrete resource allocation problem with integer variables. We describe an iterative heuristic that basically solves a series of nonlinear knapsack problems. On average, our method seems to find solutions within 0.5% of the optimal solution.

The only known structural property of optimal assortments is related to demand rates (van Ryzin and Mahajan, 1999). We extend this result to the shelf space constrained problem. We establish new properties that relate which products are included in the assortment as well as their inventory levels to demand rates and other characteristics of the products such as per-unit margin, per-unit volume, batch size, and lead time.

We applied our method at a large retail chain in the grocery industry. Comparing results with the existing assortments suggest more than 50% increase in profits. We present our results and discuss the significance of joint optimization and explicit consideration of substitution in the context of our application. Using this system we can also obtain realistic estimates of the impact of changes in delivery frequency and total shelf space.

## 4.1 Model of Substitution

A customer's favorite variant may not be available at the time of her shopping because that variant is either stocked out temporarily or not carried at the store. In the case of unavailability, customers may be willing to accept a similar product as a substitute. This is called consumer-driven substitution and is observed in many retail settings.

There are two ways of modeling consumer-driven substitution that are effectively quite similar. We discuss the studies in the literature and our model in the contexts of these two paradigms. Suppose the set of potential products are  $N = \{1, 2, \dots, J\}$  and a store  $i$  carries assortment  $S_i \subseteq N$ .

The utility-based model of demand assumes that consumers choose the utility maximizing alternative from  $S_i \cup \{0\}$ . This choice set includes the available products and the no-purchase option. If a customer's first choice is to buy a product that is not available, the second highest utility alternative is chosen. Same procedure is repeated until either a chosen product is available or no-purchase option is chosen.

van Ryzin and Mahajan (1999) use the utility-based approach, namely the Multinomial Logit (MNL), to model the consumer demand. Consumers choose from  $S_i \cup \{0\}$  to maximize their utilities. If a consumer's choice is not available due to a stockout, the sale is lost. They call this a *static* substitution, since customers choose from  $S_i$  rather than  $N$ , people who would normally choose a product from  $N \setminus S_i$  may choose something from  $S_i$  (*static* substitution) but not allowed to substitute if their choice is stocked out (*dynamic* substitution). Notice that the probability of accepting something from  $S_i$  among the customers of

$N \setminus S_i$  is  $\sum_{j \in S_i} e^{u_j} / \sum_{j \in S_i \cup \{0\}} e^{u_j}$ , which increases with larger  $S_i$ .

Mahajan and van Ryzin (2001) use an extension of the above approach to characterize the optimal inventory levels for a given assortment. In this paper, their demand model allows the customers to choose their second, third favorite until either a chosen product is available or no-purchase option is chosen. They call this *dynamic* substitution since the customer's final choice depends on the inventory levels at the moment of her shopping.

In both models of substitution, utility of no-purchase option with respect to the utility of the products determines the rate of substitution. Consider a dynamic substitution example, where  $S = \{1, 2\}$  and  $u_0 = u_1 = u_2$ . Two thirds of the customers are willing to buy something from the category. Only half of the consumers whose favorite is stocked out will switch to the other variant as a substitute and the other half will prefer no-purchase alternative to the other variant. In this example the penetration to the category (purchase incidence) is  $2/3$  and the average substitution rate is 50%. These two quantities are linked via the deterministic utilities. It is not possible with this model to have two categories with the same penetration rate but different substitution rates. It wouldn't be reasonable to expect all categories to have the same substitution rates. One way to control for substitution rate is to vary the value of  $u_0$ , but that is not desirable since it also determines the initial penetration rate to the category. Also there is no way to distinguish between the types of categories (i.e. high/low substitution) in an estimation procedure. This model fails to capture notably the most important aspect of consumer-driven substitution (i.e., rate of substitution depends on the specifics of the category) and don't work concurrently with estimation procedures.

Recall the definition of substitution types at this point. Assortment-based substitution is the switch to an available variant by a consumer when her original favorite product is not carried in the store. Stockout-based substitution is the switch to an available variant by a consumer when her original favorite is carried in the store, but it is stocked-out at the time of her shopping.

The second paradigm, a probabilistic model of demand, is mostly used in operational models and assumes that a customer knows exactly what her favorite product is among all products in that subcategory (set of substitutable products). If her first favorite is not available (either temporarily stocked out or permanently removed from the assortment), she chooses to substitute with another product with a certain probability. She also knows what her second favorite product is. If the second choice is not available either, she substitutes with a certain probability and choose her third favorite and so on. Probability matrices govern the substitution rates between variants. The matrices at each stage can be the same or different. The substitution probability matrix  $(\alpha_{ij})$  can be used for *assortment-based* or *stockout-based* substitutions.

The probabilistic model is common in inventory models with demand substitution (see Netessine and Rudi ,2003 and references therein). Smith and Agrawal (2000) use the probabilistic model to consider both assortment-based and stockout-based substitution. They allow for one substitution attempt only. Customers choose from the set  $N$ , and if the item they choose is not available for any reason, they have the chance to substitute according to the substitution matrix. If the substitute is not available either, they are not allowed to repeat the procedure. Both stockout and assortment-based substitutions can

happen, but only for one round. In a sense this model stands between the van Ryzin and Mahajan (1999) and the Mahajan and van Ryzin (2001) models.

#### 4.1.1 Suggested model

Neither the papers with utility-based models nor the ones with the probabilistic models discuss how the parameters of their models can be estimated from sales or panel data. We want to define a model of substitution that can be used consistently both in estimation of demand (and substitution behavior) and optimization of the assortment. Our model of substitution behavior is a probabilistic model with one substitution attempt only (similar to Smith and Agrawal, 2001) and follows the assumptions listed below.

(A1) Every customer chooses her favorite variant from the set  $N$

(A2) If this favorite is not available for any reason, the second favorite is chosen following the substitution probability matrix.

(A3) At this point either the product is available and the sale is made, or the sale is lost.

(A4) No more attempts of substitution are allowed.

We assume a simple structure that determines the substitution probability matrix. Probability of substitution is  $\delta$  and then the second favorite is chosen with deterministic proportions of the demand rate. More specifically, let  $d_j$  denote the base demand for the products  $N = \{1..J\}$ . The probability of substituting product  $i$  for  $j$  is

$$\alpha_{ji} = \delta \frac{d_i}{\sum_{k \in N \setminus \{j\}} d_k}$$

This simple, one-parameter model is quite powerful; it captures many important realistic aspects of substitution.

- $\delta$  single handedly determines the intensity of substitution, hence we are able to differentiate between categories with low and high substitution rates.
- With this demand model penetration rate to the category depends on the assortment  $S$ :  $\sum_{k \in S} d_k$  is willing to buy one of the available items in the assortment (base demand).
- $\delta \sum_{k \in N \setminus S} d_k \left( \sum_{k \in S} d_k / \sum_{k \in N \setminus \{j\}} d_k \right)$  will be accepted as substitution demand in the first round of substitution. Hence the total demand to the category is

$$\sum_{k \in S} d_k \left( 1 + \delta \sum_{k \in N \setminus S} d_k / \sum_{k \in N \setminus \{j\}} d_k \right)$$

- $\alpha_{ji} > \alpha_{jk}$  if  $d_i > d_k$ . This is certainly a desirable property and also consistent with what would happen in a utility-based framework (i.e., items with higher utilities - that had higher shares of base demand - will get higher shares of the substitution demand in the second round).
- Suppose that a store doesn't carry the whole assortment, i.e.,  $N \setminus S \neq \emptyset$ . Then the substitution from any product  $j$  demand to products in  $N \setminus S \left( \delta \sum_{i \in N \setminus S} d_i / \sum_{k \in N \setminus \{j\}} d_k \right)$  would be lost since another round of substitution is not allowed. The realized substitution rate from variant  $j$  would be  $\sum_{i \in S} \alpha_{ji} = \delta \sum_{i \in S} d_i / \sum_{k \in N \setminus \{j\}} d_k$ . Notice that the realized substitution rate is increasing in the set  $S$ . This is definitely consistent with the fact that a consumer who can not find her favorite variant in the store is more likely to accept a substitute, as the set of potential substitutes gets larger.

- Note that assumptions A1-A4 are similar to the probabilistic models, however our one-parameter model features many desirable properties of the utility-based models.
- One can argue that limiting the number of substitution attempts may be too restrictive. Figure 4.1 presents an example where the effective demands of a three-attempts-substitution model with rate 0.5 are approximated almost perfectly with a single-attempt-substitution model with rate 0.58. Since the single-attempt model should have a higher rate, the approximation may not work well if the multiple-attempts model has a very high substitution rate.

Based on the above observations, we believe that the described model of substitution is a reasonable representation of reality. In the next section, we discuss the sales implications of this model and use the implications to develop an estimation methodology.

## 4.2 Estimating Substitution

### 4.2.1 Estimation of assortment-based substitution

#### Implications of the substitution model on sales data

Suppose that a store carries assortment  $S \neq N$  with 100% service rate and the category has a positive substitution rate, i.e.,  $\delta > 0$ . Clearly the observed demands for variants in  $S$  are greater than the true demands, since they include the substitution demand from the variants not in  $S$ . The challenge is to identify these two parts of the observed demand so that we would know the true demand of the products in  $S$  and as well as in  $N \setminus S$ .

Consider the following hypothetical example where variants are numbered in decreasing demand rates. Based on the original demand rates, we compute and plot the total sales for no substitution and substitution cases. In Figure 4.2, total sales increases in the no-substitution case at a diminishing rate due to decreasing sales rate of individual variants. The curvature of the sales curve with positive substitution rate is due to two effects, diminishing sales and substitution. We can decompose the total effect into these two effects by transforming the x-axis from number of variants to associated cumulative market share. In Figure 4.3, the x-axis is the total original demand for the first  $n$  products. If the first  $n$  products were carried, given enough inventory, observed sales under no-substitution would be equal to the original demand rates for those  $n$  products. No-substitution sales is the  $45^\circ$  line and substitution sales is a concave function above it. The deviation from the  $45^\circ$  line is determined by the substitution rate ( $\delta$ ) of the subcategory.

The generalization of this idea necessitates plotting the observed sales versus the cumulative sales at no-substitution. We can plot these curves for each subcategory using data from multiple stores. Each store is a point on the euclidean plane where its coordinates are  $x$  = what would its sales be with its' current assortment if there was no substitution and  $y$  = observed sales of the subcategory. Consider store  $A$  with less than full assortment. The  $y$  coordinate can be estimated from the regression models for specific to store  $A$ . Regressions help clean out the data from the effect of promotions, special holidays, weather, etc. The  $x$  coordinate however is not known. Nevertheless, there are stores with full assortment and no assortment-based substitution takes place in those stores. Hence the observed demands should be equal to the true demands in these stores. We can obtain

estimates of the original demand by applying the *incidence-choice-quantity* models on all such stores together (i.e., regression coefficients would be the same across stores). We call these models the *national* models. We can then use the original demand estimates to compute what stores with less than full assortment would sell with no substitution. That is simply the sum of the demands of the variants in the assortment. Once we place the stores on the graph, finding the  $\delta$  value that fits the data best is trivial.

This approach certainly is crude and needs some tuning before use. First of all, we need to use the penetration rate to the subcategory rather than the sales to eliminate the scale differences between stores. Secondly, this approach implicitly makes the assumption that the choice behavior is the same across all stores because the regression coefficients in the national model are not store-specific. We could minimize the effects of this assumption by clustering the stores or using the store properties in the regression equations and incorporating these in the computation of the coordinates of the stores. Stores have five properties: store size, percentage of customers who bike to the store, percentage of customers who drive to the store, average sales per customer, and sales share in the chain. Since these five properties were highly collinear, we included only the first two in the *national* models.

## The Methodology

We start with using our regression models to estimate subcategory sales per customer with two methods. First method is the *store* models, applying the *purchase incidence*, *choice* and *quantity* models to each store separately. Second method is the *national* models, applying the same models to the data from all stores with full assortment. Let superscripts

$i$  and  $n$  denote the estimations from the *store* model  $i$  and the *national* model, respectively.

Let superscript  $n$  denote the estimations from the *national* model (e.g.,  $p_{ji}$  is the choice probability of product  $j$  in store  $i$  as estimated from that store model  $i$ , whereas  $p_{ji}^n$  is the same thing but estimated from national model).

(y-axis)  $(PQ)_i^y = \pi_i \sum_{j \in S_i} p_{ji} q_{ji}$ ,  $PQ$  at store  $i$  given the *store* model  $i$  and assortment  $S_i$

(x-axis)  $(PQ)_i^x = \pi_i^n \sum_{j \in S_i} p_{ji}^n q_{ji}^n$ ,  $PQ$  at store  $i$  given the *national* model and assortment  $S_i$

Let's also define

$(PQ)_i^{full} = \pi_i^n \sum_{j \in N} p_{ji}^n q_{ji}^n$ ,  $PQ$  at store  $i$  given the *national* model and assortment  $N$

The observed sales per visiting customer at store  $i$  is  $(PQ)_i^y$ . However, according to the national model and a given substitution rate  $\delta$ , the sales at store  $i$  should be

$$(\widehat{PQ})_i^y = (PQ)_i^x + \delta \frac{(PQ)_i^x}{(PQ)_i^{full}} \left( (PQ)_i^{full} - (PQ)_i^x \right)$$

The best substitution rate estimate is the one that minimizes the total squared error of estimation across stores and time (subscript  $t$ ).

$$\delta^* = \arg \min_{0 \leq \delta \leq 1} \sum_i \sum_t \left( (\widehat{PQ})_{it}^y - (PQ)_{it}^y \right)^2$$

Figure 4.4 nicely illustrates the main idea of this method and summarizes the methodology.

## Computation of the true demand rates

Now that we have an estimate of the substitution rate, we need to adjust the demand rates of all variants in  $N$  for each store  $i$ . This includes two tasks: deflating the demand rate of the variants already in the assortment  $S_i$ , and estimating a positive demand rate for the variants that are in  $S_i$ . For variants in  $S_i$ , we use the estimate from the *store* model as much as possible, so that the less predictive *national* model does not hurt the quality of estimation much. We take a variant's demand rate based on the *store* model only and deflate it according to the substitution rate. Basically, we reconstruct the demand vector for the case that the store carries full assortment. Needless to say, if a store is already carrying full assortment, no computation is necessary, the *store* model is fully adequate. The formulas for the adjusted demand rates are given below.

$$d_j = \begin{cases} (totPQ)_i \frac{(PQ)_i^x}{(PQ)_i^{full}} \frac{\pi_i p_{ii} q_{ji}}{(PQ)_i^y} N_i & j \in S_i \\ (totPQ)_i \frac{\pi_i^n p_{ii}^n q_{ji}^n}{(PQ)_i^{full}} N_i & j \notin S_i \end{cases}$$

where

$$(totPQ)_i = \frac{(PQ)_i^x}{\left( \frac{(PQ)_i^x}{(PQ)_i^{full}} + \delta \sum_{j \notin S_i} \left( \frac{\pi_i^n p_{ii}^n q_{ji}^n}{(PQ)_i^{full}} \frac{(PQ)_i^x}{(PQ)_i^{full} - \pi_i^n p_{ji}^n q_{ji}^n} \left( (PQ)_i^{full} - (PQ)_i^x \right) \right) \right)}$$

Observe that, as one would expect, if the estimated substitution rate were zero, there is no deflation of demand but just inclusion of the demand of the new variants. Figure 4.5 provides an example of adjustment of demand rates for a particular store-subcategory-date.

## Verification

This procedure uses a combination of results from many regression models estimates a substitution rate. Namely, we use 114 regression models for each subcategory (i.e., *incidence*, *choice*, *quantity* models for each of the 37 stores and the national model). There is no straightforward metric to measure accuracy of this procedure, such as  $R^2$  or  $MAD$ . Nevertheless, we know how much error reduction is achieved by choosing the error minimizing substitution rate. This gives us an idea about how significant the substitution is in that particular subcategory. Table 4.1 provides a sample of the estimation results and the percentage error reduction for a select set of subcategories. Out of the 114 subcategories in our data, 47 had all variants in all stores, and there was no assortment substitution possible. 35 among the remaining 67 had positive substitution rate estimates and 16 of them had more than 5% error reduction.

For verification purposes, we compare the estimates from different periods of time (see Table 4.1) and observe that the estimates are consistent for both significant ( $>5\%$ ) and insignificant error reductions.

We devise another verification test. We made up 35 fake subcategories, by choosing some SKUs from subcategories that are fully represented in all stores and some SKUs that are not carried in all stores. Assuming that the subcategory definitions are accurate, i.e., there are no interactions between SKUs across subcategories and there are no correlations between stores' SKU selections in different subcategories, ideally, our procedure should estimate substitution rate to be zero for all 35 fake subcategories. Out of the 35 segments only 5 had positive substitution rates with more than 5% error reduction. We believe

that this supports the substitution estimation scheme described in this section as a reliable methodology.

#### **4.2.2 Estimation of stockout-based substitution**

Stockout-based substitution happens when a consumer’s favorite product is temporarily out of stock. Tracking this behavior necessitates perfect inventory data with time stamp of sales and replenishments. Anupindi et. al. (1998) describe a method to estimate substitution between two soda types using data from vending machines. They model the customer arrivals as a Poisson process and maximize the likelihood function using the Expectation-Maximization (EM) algorithm. They also describe an implementation of the method when there are more than two products.

We applied a similar method to the inventory data for five subcategories from one store. Given the substitution structure suggested in §4.1.1, we elected to apply a sequential procedure (rather than simultaneous) to estimate original demand rates and a substitution rate. First, the shares of the variants are estimated using the choice model. Second, as transactions are read from inventory data, the likelihood function is computed with substitution rate as a variable. Maximum likelihood estimate then is obtained using a nonlinear maximization technique. The resulting estimates of the substitution rate are 1.0 for all subcategories.

The most important drawback of this procedure is that it does not produce any demand estimates for SKUs that are not carried in a store. In any case, it can not be of any use at Albert Heijn due to the following reasons. First, inventory data is not accurate. It is

hard to decide what is a stockout and what is not, when there are times inventory level is negative, sometimes down to -12. It is possible to apply a similar methodology with the substitution structure given in Section 4.1.1. However, this requires availability of accurate data of inventory transactions. Raman et. al. (2001) reports that inventory data at retailers are often wildly inaccurate. Secondly, service levels at Albert Heijn are very high and stockouts are rare. Thirdly, the inventory data is available for only their pilot stores but not any other stores.

We use the rate estimated for assortment-based substitution also for the stockout-based substitution. It is worth mentioning that consumers probably respond to assortment changes and stockouts in a different way. Therefore, using the same rate for both types of substitution may not be most accurate. On the other hand because of high service levels, demand from stockout-based substitution can be small compared to that from the assortment-based substitution.

### **4.3 Assortment Optimization under Substitution**

We modify the assortment optimization problem under no substitution (AP0) that was described in §3.2.4 to incorporate substitution. Allocating a number of facings to a product implies a maximum inventory level, whose expected profit is computed via simulation. The objective is to assign number of facings to each product to maximize total expected gross profits from one or more categories subject to a shelf space constraint. We obtain the

optimal number of facings by solving the following mathematical program.

$$\begin{aligned}
& \max \quad \sum_s \sum_j GP_{sj}(\delta_s, \mathbf{n}_s, \mathbf{d}_s) \\
& s.t. \quad \sum_s \sum_j n_{sj} w_{sj} \leq ShelfSpace \\
& \quad \quad n_{sj} \in \{0, 1, 2, \dots\}, \text{ for all } s, j
\end{aligned} \tag{AP(\mathbf{d})}$$

where  $\mathbf{n}_s = (n_{s1}, n_{s2}, \dots)$ ,  $\mathbf{d}_s = (d_{s1}, d_{s2}, \dots)$ , and  $w_{sj}$  is the width of a facing for SKU  $j$  in  $s$ .

Notice that the gross profit function  $GP_{sj}$  is not only a function of the facings of that SKU ( $n_{sj}$ ), but also the facings of all other SKUs in subcategory  $s$ . The reason of course is substitution. This is basically a knapsack problem with a nonlinear and nonseparable objective function, whose coefficients need to be calculated for every combination of the decision variables. Even if we rule out the stockout-based substitution, we need to consider 'in' and 'out' of the assortment values for all SKUs leading to  $\prod_s 2^{(\#SKUs \text{ in } s)}$  combinations.

We suggest the following iterative heuristic that solves a series of separable problems. Let  $\mathbf{d}$  be the matrix  $(\mathbf{d}_1, \dots, \mathbf{d}_s, \dots)$ . We set  $\delta_s = 0, \forall s$  and solve  $AP(\mathbf{d})$  with true demand rates resulting in a particular facings allocation,  $\mathbf{n}^0$ . At iteration  $t$ , given  $\mathbf{n}_s^{t-1}$  and  $\delta_s$ , we can recompute the effective demand rates  $\mathbf{d}_s^t$  for each subcategory using the following formula

$$d_{sj}^t = \begin{cases} D_s \delta_s d_{sj}^0 / \sum_{k \neq j} d_{sk}^0 & \text{if } n_{sj}^{t-1} > 0 \\ d_{sj}^0 & \text{if } n_{sj}^{t-1} = 0 \end{cases} \tag{4.1}$$

where  $D_s = \sum_{j|n_{sj}^{t-1}=0} d_{sj}^0 + \sum_{j|n_{sj}^{t-1}>0} LostSales_{sj}(d_{sj}^0, n_{sj}^{t-1})$ . The first sum is the redistributed demand from *assortment-based* substitution and the second one is from *stockout-based* substitution. We then solve  $AP(\mathbf{d}^t)$  for  $\delta_s = 0, \forall s$ . Below is a formal description of

our iterative heuristic. Figure 2 presents an example that illustrates the working of the heuristic without the local search.

### Iterative Heuristic

1. Set  $t = 0$ . Solve  $AP(\mathbf{d}^0)$  and record  $n_{sj}$  for all  $s$  and  $j$ .
2. Compute  $\mathbf{d}^{t+1}$  using equation (4.1)
3.  $t \leftarrow t + 1$
4. Solve  $AP(\mathbf{d}^t)$  for  $\delta_s = 0, \forall s$ .
5. If  $\mathbf{n}^t \neq \mathbf{n}^{t-1}$ , the facing allocation has changed, GO TO 2. If not, then the procedure converged, GO TO 6.
6. Record the solution  $\mathbf{n}^t$  as the 'best'.

### LOCAL SEARCH:

7. Randomly pick a subcategory  $s$  that was not picked before and has positive variety, i.e.,  $\sum_j n_{sj} > 0$ . If none left, STOP.
8. Drop the last SKU of that subcategory that is added to the assortment and mark that SKU to not enter again.
9. Apply 2 through 5. When the procedure converges, if the solution is better than 'best', record the current solution as the best. If not, revert to the 'best' solution.
9. GO TO 7.

We solve Step 3 as follows. When  $\delta_s = 0$ ,  $GP_{sj}$  depends on  $n_{sj}$  and  $d_{sj}$  only, i.e., the objective function is separable in the decision variables. The optimization is done by an enhanced greedy heuristic. The enhancement to the greedy heuristic is as follows. Some of the SKUs have case-sizes larger than the capacity of their facings. If only one facing is assigned to such items, the reorder point is negative, so no inventory is carried and sales of the item is zero. Nevertheless, second or third facings can create positive returns. Regular greedy heuristic never assigns the first facing to these products because the marginal benefit of the first facing is zero. Also for perishable products or products with high demand rate, gross profit may be nonconcave due to the discrete casesize-facingsize relations: If the inventory level is frequently zero at the end of the period, (either because unsold units are salvaged or all inventory is sold), and the number of cases that fit in the given space doesn't increase with an assigned facing, that last facing may not be effective. The gross profit will be like a step function that jumps every time a new case is able to fit to the space defined by number of facings. The enhanced greedy heuristic overcomes these problems by also considering two-step jumps. The performance of the enhanced greedy heuristic is also included in the discussion of the results.

**Example:**

This example illustrates the iterative heuristic without the local search. We have two subcategories. First has two variants and the second has three variants. Substitution rate is 1. All five products have the same price, cost and facings size. The facing size is equal to five units, which means that the inventory level is five times the allocated facings. For simplicity, assume that the leadtime is zero, the items are not perishable

and there is no stockout-based substitution. Demands are random variables with Poisson distribution, whose means are given below. The gross profit return per period for each facing,  $E[\min\{5n_{sj}, d_{sj}\}]$ , is calculated numerically. Our objective is to allocate five facings to these products to maximize gross profits.

Table 4.2 presents the data and expected sales for original and updated demand rates. In the first iteration when the substitution effects are not considered, the allocation vector is (2,1,1,1,0). After the second iteration, the products in the assortment has not changed, so the algorithm is converged. The final allocation is (1,1,2,1,0). The sales vector for (2,1,1,1,0) is (4.95+3.8, 4.12, 4.97, 4.2, 0) with a total of 22.04. The sales vector for (1,1,2,1,0) is (4.95, 4.12, 4.97+4.15, 4.2, 0) with a total of 22.39. The improvement in total sales is 1.6%, which can be significant in low margin businesses. For example, with a gross margin of 25% and the firm's net profit of 3% of the revenues, the impact of 1.6% increase in sales is equivalent to a 13% increase in the firm's net profit.

### 4.3.1 Comparison with other optimization methods

Iterations in the Iterative Heuristic is a way to seek consistency between the input demand curves and the optimal solution. Since the application of greedy assumes independent products (i.e., no substitution), the resulting solution (availability of products) may imply a demand vector different than the inputted demand vector. Iteration (update-solve) helps us to reach consistency. It is worth noting here that it is different than the kind of iteration in a nonlinear continuous optimization method such as the Newton-Raphson method, where iteration is based on the gradient of the objective function. In other words, we don't know

the true value of the objective function, until the iterative heuristic converges. Therefore, the solution will not necessarily improve at each iteration. The local search enables us to jump away from the current solution by perturbing the solution for one segment and observing the final impact of that perturbation.

For nonseparable resource allocation problems, another iterative method that is guaranteed to reach to the optimal solution for convex problems is suggested by Dussault (1986) and Klastorin (1990). They simply solve the integer problem with the branch and bound method, but at each node of the branch and bound method, they solve the relaxed (i.e., continuous variables) nonseparable quadratic optimization problem with a series of separable problems. When the nonseparable objective function is replaced with its Taylor approximation, the quadratic structure yields a separable approximation, which can be solved easily. At the next step, Taylor approximation is updated based on the current solution. The iteration of separable problems is in the continuous domain. Our method basically applies this idea in the discrete space. Application of the greedy heuristic with no-substitution assumption corresponds to the solution of the approximate separable objective function and the updating of the demand vector based on the current solution corresponds to the recomputing of the Taylor approximation. The difference is that our method incorporates the integrality at each iteration, whereas Dussault (1986) and Klastorin (1990) impose the integrality constraints at the highest level of hierarchy in the Branch and Bound method. Another difference is that the quadratic problem's Taylor approximation is clearly separable, whereas we impose separation at the input level (i.e., demand vector) rather than the objective function, because it is hard to characterize the shape of the dependence at the

objective function level. None of these methods are designed for nonconvex problems and may not reach the global optimal solution.

### 4.3.2 Structural Properties

The iterative heuristic is based on a greedy heuristic. Therefore we can find properties of the resulting solution by exploiting the way the greedy heuristic works. First we note that the gross profit curve for a product depends on demand, margin and operational constraints. Demand level and per-unit margin effects the maximum gross profit a product can generate if sufficient inventory is held. Operational constraints, such as case sizes, delivery leadtime effect the curvature of the gross profit function. A product with a smaller case size has a higher slope of the gross profit curve for low inventory levels, resulting in reaching the maximum possible gross profit with less inventory. Below we state these observations formally.

**Lemma 4.3.1** *Consider two non-perishable products A and B in subcategory s. Let  $\mathbf{n}_{others}$  denote the vector of facing allocations for all products in s other than A and B. The following holds for any replenishment system and substitution rate  $\delta_s$ .*

i) *For products A and B: All else is equal, the original demand rates are different. If  $d_A > d_B$ , then*

$$GP_A(n) - GP_A(n-1) > GP_B(n) - GP_B(n-1), \text{ for } n \geq 1 \text{ and any } \mathbf{n}_{others}$$

ii) *For products A and B: All else is equal, the per unit gross margins (gm) are different.*

If  $gm_A > gm_B$ , then

$$GP_A(n) - GP_A(n-1) > GP_B(n) - GP_B(n-1), \text{ for } n \geq 1 \text{ and any } \mathbf{n}_{others}$$

**Proof.** First note that for given  $\mathbf{n}_{others}$  and  $\delta_s$ ,  $GP_j()$  is a function of  $n_j$  and  $D_j$  (the effective demand to a product) only.

i) Since the substitution demand is allocated to the products proportional to the original demand rates, the mean of effective demands for products  $A$  and  $B$  follow the order of the mean of original demands, i.e.,  $D_A > D_B$  for any  $\mathbf{n}_{others}$ . Therefore we can compare the gross profit functions based on the original demand rates.

The number of facings  $n$  implies a maximum inventory level (a basestock level)  $nf$  where  $f$  is the capacity per facing. Clearly increasing the demand rate of a product cannot decrease sales under any inventory policy, i.e.,  $GP_A(n) > GP_B(n)$ . Moreover, it cannot decrease the marginal benefit of increasing the maximum inventory level.

ii) Since the substitution demand is allocated to the products that are in the assortment proportional to the original demand rates, the mean of effective demands for products  $A$  and  $B$  are the same for any  $\mathbf{n}_{others}$ . GP is sales times gross margin. Since expected sales is the same for  $A$  and  $B$  and  $gm_A > gm_B$ , we have  $GP_A(n) - GP_A(n-1) = (gm_A/gm_B)(GP_B(n) - GP_B(n-1)) > GP_B(n) - GP_B(n-1)$ . ■

Note that the above lemma implies that also  $GP_A(n) > GP_B(n)$  for any  $n$  under both cases (*i* or *ii*).

**Lemma 4.3.2** *Consider two products  $A$  and  $B$  in subcategory  $s$ . Let  $\mathbf{n}_{others}$  denote the vector of facing allocations for all products in  $s$  other than  $A$  and  $B$ .*

1.  $\lim_{n \rightarrow \infty} GP_A(n) = \lim_{n \rightarrow \infty} GP_B(n)$  for any  $\mathbf{n}_{others}$ .

2. If exactly one of the four conditions (i or iv) is met,

i) For products  $A$  and  $B$ : All else is equal, the batch sizes are such that  $b_A = 1$  and

$$b_B > 1$$

ii) Replenishment system is a multi-period basestock system with batch sizes of 1.

For products  $A$  and  $B$ : All else is equal, the coefficients of variation of demand are such that  $cv_A < cv_B$ .

iii) Replenishment system is a continuous review system (e.g., a  $(Q,r)$  policy). For

products  $A$  and  $B$ : All else is equal, the coefficients of variation of demand are such that  $cv_A < cv_B$ .

iv) Replenishment system is a continuous review system (e.g., a  $(Q,r)$  policy). For

products  $A$  and  $B$ : All else is equal, the leadtimes are such that  $l_A < l_B$ .

then, the following hold for any substitution rate  $\delta_s$ .

$$GP_A(n) > GP_B(n) \text{ for } n \geq 1 \text{ and any } \mathbf{n}_{others}.$$

**Proof.** Since the substitution demand is allocated to the products proportional to the original demand rates, the mean of effective demands for products  $A$  and  $B$  are the same for any  $\mathbf{n}_{others}$ . The first result is obvious: given infinite inventory levels, sales for both products will be equal to the expected effective demand rates.

2.i) Product  $A$ 's replenishment system can always replicate product  $B$ 's ordering decisions. However since it is more flexible, at any ordering point,  $A$  orders more than or equal

to the order size of  $B$ , resulting in more inventory for  $A$ . This means lower lost sales and higher gross profits.

2.ii) At every period, the initial inventory level is the same for both products. Notice that the mean effective demand for products  $A$  and  $B$  are the same. Expected sales is mean demand minus lost sales during each period. Lost sales in a period increases with demand variance. Hence sales of product  $A$  is larger than product  $B$ .

2.iii and iv) In a continuous review system,  $r + Q$  is the maximum inventory level corresponding to  $n$  facings. It is known in the inventory literature (see Zheng 1992 for details) that the inventory position ( $y$ ) is uniformly distributed between  $(r, r + Q]$ . The expected inventory costs is an average of newsvendor costs with inventory level  $y$ , i.e.,  $\int_r^{r+Q} G(y) dy$ , where the demand for the newsvendor is the total demand during the leadtime. As leadtime or the demand variance increases,  $G(y)$  increases, resulting in a higher gross profit for product  $A$  than product  $B$ . ■

**Proposition 4.3.3** *Consider two products in the same subcategory subject to the same operational characteristics. If exactly one of the following conditions is met,*

i) *All else is equal.  $d_A \geq d_B$ .*

ii) *All else is equal.  $gm_A \geq gm_B$ .*

iii) *All else is equal.  $w_A \leq w_B$ .*

*then  $n_A \geq n_B$  in the final solution of the Iterative Heuristic.*

**Proof.** At every iteration  $t$ , the demand rates are computed based on  $\mathbf{n}^{t-1}$  and greedy heuristic is applied to the data. At each iteration, the value of the assigning the  $n^{th}$  facing

to product  $j$  per unit shelf space is  $(GP_j(n) - GP_j(n-1))/w_j$ . By Lemma 4.3.1, the marginal value is higher for product  $A$  than  $B$ . Hence the greedy heuristic will never assign facing  $n$  to product  $B$  before product  $A$ . ■

The implications of the first part of this proposition is clear, an allocation algorithm based on demand rates should work fairly well when products are differentiated by the demand rates only. Part i) proposition for the shelf space constrained problem is similar to the property of optimal assortments in the unconstrained problem in van Ryzin and Mahajan (1999).

**Proposition 4.3.4** *Consider two products in the same subcategory subject to the same operational characteristics. If exactly one of the four conditions in Lemma 4.3.2 is met, then the following is true. In the final solution of the Iterative Heuristic if product  $B$  is included in the assortment then so is  $A$  (i.e.,  $n_B > 0 \implies n_A > 0$ ).*

**Proof.** At every iteration  $t$ , the demand rates are computed based on  $\mathbf{n}^{t-1}$  and greedy heuristic is applied to the data. At each iteration, the value of the assigning the 1<sup>st</sup> facing to product  $j$  per unit shelf space is  $GP_j(1)/w_j$ . By Lemma 4.3.2, the marginal value is higher for product  $A$  than  $B$ . Hence the greedy heuristic will never assign the first facing to product  $B$  before product  $A$ . ■

We suspect that generalizing condition i) in Lemma 4.3.2 to  $b_A < b_B$  will yield the same result for most inventory systems. In any case, when one of the conditions of the Lemma hold, due to limited shelf space, if  $A$  is not included in the assortment, neither is  $B$ . Because the maximum level of the gross profit curve of  $A$  is higher and the slope is higher

for low inventory levels, the profit impact of first facing is higher for  $A$ , resulting in a higher rank in the ordered input list to the Greedy heuristic. However, if both products are in the assortment, the inventory level of  $B$  can be larger than  $A$ . The reason is that gross profit of  $A$  reaches its maximum level quickly with the early inventory allocations, whereas it takes more inventory for  $B$  to reach its maximum. In such cases, allocation heuristics based on demand rates perform poorly. A reasonable rule of thumb can be derived from this discussion. First high demand rate products shall be included in the assortment, then more inventory shall be allocated to the products that have more restrictive operational constraints.

The above propositions also hold for products from different subcategories under no-substitution. Under substitution however, the results cannot be extended to products from different subcategories, because the effective demand rates depend on the availability of other products from the respective subcategories.

### 4.3.3 Numerical study

In the simple example in the previous section, explicitly considering substitution increased sales an extra 1.6%. It is of importance to investigate this impact with larger examples. We focus on a category with 14 segments and study three sets of examples. First set includes single segment problems with up to 7 SKUs in each segment. Second set includes 13 two-segment problems with up to a total of 8 SKUs between the two segments. Third set includes 4 three-segment problems with up to a total of 8 SKUs between them. We consider the combination of the following settings resulting in 6 instances for each of the 31 problems.

$$Substitution\ rate = \{0, 1\} \quad Shelf\ space = \{1/6, 1/3, 2/3\} \times (full.service.shelf.space)$$

where *full\_service\_shelf\_space* is set such that all products have a service level of 99.9% or more. We solve for the optimal solution via enumeration and compare the heuristic solution to the optimal gross profits. In order to reduce running time, we approximate the gross profit functions with the newsvendor formula rather than using simulations.

In the no-substitution case, the initial greedy solution is the final solution of the heuristic. The average percentage gap with the optimal solution is 0.1% with a maximum of 3.2%. The greedy heuristic found the optimal solution 85 times out of the 93 problems. The results for the substitution case are summarized in Table 4.3. The greedy heuristic (or step 1 of the Iterative Heuristic) has an average gap of 3.0% and a maximum gap of 24.1%, finding the optimal solution in 14 out of the 93 problems. Iterative Heuristic reduces average gap to 0.5%, the maximum gap to 12.1% and finds the optimal solution in 40 instances (26 more instances).

The 2.5% difference on average between the Iterative Heuristic solution and the Greedy solution is simply the value of explicitly considering substitution in the optimization.

Table 4.4 presents the average gap for different levels of shelf space. Average gap decreases as resource (shelf space) is less restrictive. Among 62 problems that has 2/3 of the *full\_service\_shelf\_space*, the maximum gap of the iterative heuristic is 0.1%. We suspect that the practical examples in food retailing will be more similar to such high shelf space cases.

## 4.4 Implementation and Results

Our methodology is particularly designed for the grocery industry and is being applied at Albert Heijn, a large retail chain in the Netherlands. The grocery industry has distinctive characteristics. Supermarkets often carry more than 30,000 stock keeping units (SKUs), so there are very similar products in subcategories such as butters, coffee, washing detergents, hence substitution by consumers is likely to occur frequently. Products are mostly fast moving products in the middle of their life cycles. There is limited backroom for storage, so most of products are delivered from warehouses to the shelves. Since, one truck delivers many products to multiple stores, a single category does not have much influence on the delivery schedule. Therefore we can assume a fixed delivery schedule for a category at a particular store. The assortment optimization part of our methodology can be easily adopted to settings with other characteristics as long as the profit impact of including a product in the assortment can be estimated. We comment on the applicability of the substitution estimation scheme to other industries later in Chapter 4.

There are other operational constraints that needed to be incorporated to the optimization module. Minimum number of SKUs in a subcategory, minimum number of facings in a subcategory, minimum and maximum number of facings for particular SKUs are specified by merchandising managers. It is not difficult to incorporate these constraints to ( $AP0$ ) or ( $AP_s$ ). The delivery schedule for each category of products is known. A period is defined as the time between two deliveries. We compute the demand for each SKU for each period. For each subcategory, we choose the period that has the highest total demand across SKUs. We call this the peak-load period. The optimization is done based on the

peak-load period demands. For non-perishable items, the assigned facings are filled as much as possible (following the inventory policy) at all times, even in the non-peak-load periods. For perishable items however, after assigning the facings, we solve for maximum inventory levels that can be a fraction of the assigned facings. Moreover, Albert Heijn has a real time demand estimation - inventory control policy in place for products that have a shelf life of a few days or less that works within the facing limits defined by our method. As a final note on the implementation details, the merchandising managers can overwrite the recommended substitution rates.

The estimation and optimization system is tested for seven categories with data from 37 stores. Now the results are being implemented at 3 pilot stores in the Netherlands.

#### **4.4.1 Results and financial impact**

Table 4.5 presents the gross profit figures for the existing and suggested assortments for two categories and 10 stores. The improvement in the gross profits optimized over the peakload periods is 12.3%. However, this improvement may not be realized at non-peakload periods such as Monday through Wednesday. Weekly gross profits are estimated based on daily demand rates and service levels from the peakload periods for high demand days and 100% service level for the low demand dates. The total improvement in weekly gross profit estimates is 5.9%. An average retailer has a gross margin of 25% and pretax income of 3% of its revenues. An 5.9% increase in gross margin is a 1.5% increase in revenues, which increases the pretax income to %4.5. In other words, the impact of our methodology on a retail chain is a 50% increase in profits. Based on Albert Heijn's \$5B yearly sales, the

impact of the bottom line is an additional \$75 million..

The improvements have come from three types of changes to the existing assortment: radical changes in total space for subcategories, addition/deletion of SKUs within subcategories, facing exchanges between SKUs in subcategories. One of the reasons for facing exchanges between SKUs is due to Proposition 4.3.4. Facing allocations have been usually made based on sales (demand rates). Therefore a high demand item always gets more facings. However, the maximum inventory levels are usually much higher than the demand rate, due to positive leadtime and case packs. Sometimes a high demand item has a smaller case pack and does not need several facings to reach its sales plateau. These facings can be assigned to lower demand items with larger case packs. This benefit could not have been recognized if operational aspects had not been incorporated in the gross profit estimation.

Figure 4.6 illustrates the decline in percentage improvement with the increase in total shelf space after removing the two outlier points that has nearly 40% improvement. The improvement by our method is more in smaller stores where shelf space is scarcer.

#### **4.4.2 Comparison with alternative method**

A simple alternative method is devised to create a benchmark other than Ahold's current solution. The shelf space is allocated to products proportional to their average sales at that store. The continuous numbers are rounded down to discrete facings and resulting excess space is distributed to the products that had highest reminders in the round-down procedure. The Proportional Allocation heuristic outperformed the current Ahold solution by 4.4 % over the optimized peakload period and by 1.5% in weekly gross profits. Recall

that our suggested solution resulted in 12.3% and 5.9% improvements respectively.

It is surprising that a simple heuristic such as the Proportional Allocation heuristic outperforms the existing Ahold assortments. We should keep in mind that we make these comparisons based on the demand estimates obtained via our method. The results of these comparisons can be different if based on other demand estimates. There may also be other constraints merchandising managers use in their planning that were not formally part of their system (hence not communicated to us).

#### **4.4.3 Impact of delivery frequency on the system cost**

Using our system, we can generate realistic estimates of the impact of changes in the system and make more reliable cost-benefit analysis. For example, by removing every other scheduled delivery from the delivery list, we decreased the average delivery frequency by half. The impact on gross profits is a decrease of 1.5%. Gaur (2001) reports that increasing delivery frequency to twice a day from once a day increases distribution costs at Albert Heijn by 33%. Assuming symmetry, we estimate that savings in distribution costs from reducing delivery frequency by half is 33% as well. If the share of distribution costs in total revenues is less than 4.5%, it pays to decrease the delivery frequency by half.

### **4.5 Discussion**

We have presented a methodology for assortment planning under substitution, including a demand estimation procedure and an optimization algorithm. The demand estimation

procedure under substitution is the first attempt in the literature for understanding the degree of assortment-based substitution by consumers. We present a novel and intuitive method, which basically computes the imputed substitution rate by comparing the sales levels from different stores with different assortments. With this method, we obtain a substitution rate that governs consumers' choice process and the demand estimates for not only the products carried in the store, but also the products previously not carried in the store. We also present an optimization method to solve the nonlinear, nonseparable resource allocation problem, which seems to find solutions within 0.5% of the optimal solution. We have applied our methodology at a large retail chain and the improvement over their current assortments was enough to increase profits by 50%. The improvement demonstrates the benefits of joint assortment selection and inventory planning while considering substitution and operational constraints.

## Chapter 5

# Category Coordination in Retail Assortment Planning

There are three main reasons for a retailer to carry a product. First, the retailer makes money from that product. Decisions regarding these products are usually made via a marginal analysis. Second, the product pulls traffic into the store. Loss leaders are good examples of this group. Third, the product makes other products look good. Having a broad enough assortment makes customers comfortable buying what they like from this retailer, because they know that they have seen a large enough sample from the products in the market place. Existing assortment planning models assume that all categories of products fall into the first group. However, one needs to be careful in making decisions about a category based on marginal analysis, because marginal analysis does not take long run benefits into account. If a retailer reduces variety in all categories, then the overall attractiveness of the retailer would decrease and customers may defect to other

retailers. This may invalidate the assumptions the marginal analysis was based on, e.g., the assumption that the number of customers that visit the store is exogenous. This concern is particularly relevant with respect to basket shoppers, consumers who purchase from more than one category in a shopping trip. In this chapter, we incorporate the first two of the above three reasons simultaneously into an assortment planning model with multiple merchandising categories. The level of variety in a category impacts its own demand (first reason), and also the demand to the other categories (second reason). The categories interact with each other due to the effects of basket shopping consumers and limited shelf space. The retailer chooses the number of variants to offer in each category. Consumers' expected utility in each category is increasing with broader assortments. Consumers of multiple categories make their store choice between the retail store and an exogenous outside alternative based on their total basket utilities. Our model would be applicable in types of retail stores where there is a significant level of basket shopping, such as food retailing (large number of items in each consumer basket) or consumer electronics (small number of items in a basket). In addition, our model assumes that consumers choose between stores that offer a similar set of categories e.g., Best-Buy and Circuit City. However, if the customers' basket consist of electronics items and clothing items, the choice between Best-Buy and Wal-Mart does not fit in our framework.

We investigate the interaction of the category variety decisions at the retail store under centralized and decentralized management regimes. The advantage of the decentralized regime is its simplicity in solution and execution and its low level of data requirements. The research questions that we are interested in are: What is the cost of having these

advantages, i.e., what is the loss due to decentralization compared to the optimal solution? Do decentralized solutions have a consistent bias? Is there a way to have the best of both worlds, i.e., are there simple and efficient management regimes?

The common practice of category management is an example of a decentralized regime for controlling assortment because each category manager is charged with maximizing profit with his or her assigned category. Hence, a category manager typically optimizes given the assumption that the variety offered by other categories is constant. As one would expect, the category managers increase their variety levels together at low variety levels (i.e., act as strategic complements). But at higher levels of variety, they try to get a free ride by decreasing the variety level (i.e., act as strategic substitutes) counting on the large number of customers attracted to the store by the other categories. Having strategic complementarity and substitutability in the same problem is an interesting observation in itself.

We show that there exists at least one equilibrium (and maybe more) under category management. We show that the optimal variety is never an equilibrium and the variety levels are mostly lower than optimal in at least one category in any equilibrium. We also provide numerical examples to illustrate the number of category management equilibria. In a numerical study we demonstrate that the profit loss due to category management can be significant, 28% on average and the category management provides 45% less variety than optimal. More importantly, the performance gets worse as the number of categories increases. We discuss the trade off between the breadth (number of categories) and the depth (variety in each category) of assortments in this setting. We classify categories into two types: *basket categories* have a high co-occurrence rate in consumer baskets with other

categories and *independent categories* attract customers on their own. We show that under situations where the retailer is less attractive than the outside alternative, it is optimal to add a basket category to an existing assortment rather than an independent category. We present conditions where basket categories should have more variety than independent categories, and show that the equilibrium solution may do exactly the opposite.

The conventional wisdom in retailing when managing product variety has been 'more is better'. Many recent studies report that variety levels have become so excessive that reduction in number of products does not decrease sales (Dreze et. al. 1994, Broniarczyk et. al. 1998, Boatwright and Nunes 1998). Our model shows however category based analysis may be misleading, and the intuition of retailers 'more is better' can hold to be true, at least under cases with significant amount of basket shopping.

Finally, we propose a simple heuristic that mimics the centralized regime. Each category manager finds the centralized solution of a two-category problem where one category is their own and the other is an aggregate category that represents all the other categories. The computational complexity of the data estimation and solution procedures for the centralized regime are exponential in the number of categories. Our Aggregation Heuristic reduces both requirements to polynomial time and produces solutions very close to the optimal solution.

## 5.1 Related literature

### Multi-category consumer choice models

Most research on marketing has been on single category choice decisions. Recently, researchers have started to examine multiple category purchases in a single shopping occasion by modeling the dependency across multi category items explicitly . See Russell et. al. (1997) for a review of perspectives on multiple category consumer choice models. A group of papers study cross category correlations in preferences and interested in how consumers compose their shopping basket at a given trip (i.e., select the categories that they will make purchases in). We assume that basket shoppers know what categories are in their shopping list apriori and make their store choice based on the expected utility of the basket. The most relevant work to our discussion addresses the interrelated problems of store choice and market baskets. Bell and Lattin (1998) show that consumers make their store choice based on the total basket utility. Utility from a category is affected by prices and the set of products (variety) in Bell and Lattin (1998), but their analysis focus on the impact of prices. Bodapati and Srinivasan (1999) relates feature advertising to store traffic effects using a nested logit framework. Bell, Ho and Tang (1998) use market basket data to analyze consumer store choices and explicitly consider the roles of fixed and variable costs. In these papers and ours, consumers first assign a utility to an anticipated market basket and subsequently use this utility to determine store choice.

An intuitive reason for basket shopping is that consumers want to consolidate their shopping locations as much as possible due to the cost of using multiple retailers (locations) such as search and travel costs. Klemperer (1992) explores this issue with a model of competition between two multi-product firms. He shows that at equilibrium each consumer makes all her shopping from one firm only.

The bundle pricing literature is relevant to basket shopping since customers purchase bundles (multiple products grouped and priced together). Noble and Gruca (1999) devote a section to bundle pricing in their review of industrial pricing literature. Again, the focus in this area is pricing and the analysis is on product basis rather than categories of products.

### **Impact of variety on operational costs**

In a two-echelon supply chain with multiple retailers (products), Eppen (1979) and Zipkin (1995) show that inventory cost increases in proportion to the square root of the number of retailers (products). As an extension to Zipkin (1995), Thonemann and Bradley (2002) study the impact of variety on supply chain costs due to increased setup times. None of these papers include the benefits of variety (having the inventory at the retailer in the case of Eppen 1979) in their models. In our model total demand increases with variety, since the propensity of a consumer's visit to the store increases with variety. In contrast to these papers, we argue that operational costs increase convexly with variety, due to fixed shelf space and the increase in total demand.

### **Endogenous demand**

Although not relevant to product variety, there are papers in the operations management literature that model the dependence of demand to operational performance. Examples include an inventory model with service level dependent demand in Dana and Petruzzzi (2001), a clearance pricing model with inventory level dependent demand in Achabal and Smith (1998), papers modeling competition between firms such as the time based competition papers (Lee and Li, 1994; Lederer and Li, 1997), and firms competing on quality of their service (Hall and Porteus, 2000; Gans, 2002).

## 5.2 Model Basics

Consider a retailer that can potentially carry multiple categories of goods. The retailer chooses the breadth and depth of the assortment that she would like to offer, i.e., what categories to carry and what level of variety to provide in each category. The breadth of the assortment is one of the main factors that determine the attractiveness of the store to the consumers. Consumers in the market are willing to buy something from one or more categories. The consumers choose between the retailer and an outside alternative to maximize their expected total shopping utility.

Suppose that the set of categories is denoted  $J = \{1, 2, \dots, N\}$  and the set of variants in each category is denoted  $S_j = \{1, 2, \dots, I_j\}$ . The retailer chooses her assortment in all categories,  $X = (X_1, X_2, \dots, X_J)$ , where  $X_j \subset S_j$ , for all  $j$ . The assortment at the outside alternative is  $Z = (Z_1, Z_2, \dots, Z_J)$  where  $Z_j \subset S_j$ , for all  $j$ . Let  $t$  be the index to the elements of the power set of  $N$ ,  $\{B : B \subseteq N\}$ .  $B_t \subset N$  denotes a shopping basket that contains categories  $j \in B_t$ . Consumers of each basket type  $t$  want to patronize all of the categories in  $B_t$ . Consumers do all their shopping at the store of their choice, either the retailer or the outside alternative. Consumers are rational expectation maximizers and have perfect information about the service levels at the retail store and the outside alternative. Consumers of type  $t$  evaluate the attractiveness of the retail store  $A_t(X)$  and the outside alternative  $A_t(Z)$ . As suggested by the Market-Share Theorem (Bell, Keeney and Little, 1975), the share of type  $t$  consumers that choose the retail store is given by

$$s_t(X) = \frac{A_t(X)}{A_t(X) + A_t(Z)}$$

This intuitive and well accepted share formula can be obtained with many different formulations. Examples include Kotler’s fundamental theorem on marketing effort, multiplicative competitive interactions model, and the multinomial logit model. We use the multinomial logit<sup>1</sup> (MNL) framework to describe a certain form of the attractiveness function but later continue the analysis with general attractiveness functions with certain properties.

In each merchandise category  $j$ , the retailer chooses the assortment from the variants  $\{1, \dots, I_j\}$ . A consumer’s utility from variant  $i$  in category  $j$  is  $u_{ji} = v_{ji} + \varepsilon$  where  $\varepsilon$  are *i.i.d.* distributed Gumbel with mean zero and variance  $\pi^2\mu^2/6$ , i.e.,  $F(x) = \exp[-e^{-(x/\mu+\gamma)}]$  and  $\gamma$  is Euler’s constant ( $\gamma \approx 0.5722$ ). Without loss of generality, we can order the variants in a category in decreasing expected utilities, i.e.,  $v_{j1} \geq v_{j2} \geq \dots \geq v_{jI_j}$ . Van Ryzin and Mahajan (1999) show in a one category setting that it is optimal to either add the variant with highest utility to an existing set of variants or add nothing at all<sup>2</sup>. This leads to the intuitive result that the optimal assortment in a category is of the form  $\{1, \dots, x_j\}$  where  $0 \leq x_j \leq I_j$ . Consistent with this result, suppose that the retailer chooses  $x_j$  variants with the highest expected utilities to be offered as category  $j$ . Similarly,  $z_j$  highest utility variants are carried by the outside alternative. This simplifies the decision space for each category from the set of variants  $X_j$  to an integral number  $x_j$ . We can hence revise our

---

<sup>1</sup>MNL is widely used in economics and marketing literatures. See for example Guadagni and Little (1983) and a stream of follow-up papers for applications in marketing, Ben-Akiva and Lerman (1985) for applications in travel industry, and Anderson et.al. (1992) for MNL based theoretical models of product differentiation.

<sup>2</sup>The cost structure in van Ryzin and Mahajan (1999) is different than this model. In fact, showing this structure for optimal assortments is almost trivial in our setting.

market share equation as follows.

$$s_t(\mathbf{x}) = \frac{A_t(\mathbf{x})}{A_t(\mathbf{x}) + A_t(\mathbf{z})}$$

where variety levels at the retailer and the outside alternative are given by  $\mathbf{x} = (x_1, x_2, \dots, x_J)$ , and  $\mathbf{z} = (z_1, z_2, \dots, z_J)$  where  $0 \leq x_j, z_j \leq I_j$ , for all  $j$ .

A customer of a category chooses the variant that provides the maximum utility, i.e., utility from a category with  $x_j$  variants is  $U(x_j) = \max\{u_i : i = 1, 2, \dots, x_j\}$ . Since Gumbel distribution is closed under maximization, the maximum utility -from the chosen variant  $i$ - is also Gumbel distributed with mean  $\mu \ln \sum_{i=1}^{x_j} e^{v_i/\mu}$  and scale parameter  $\mu$  (Ben-Akiva and Lerman 1985). Let's consider the customer type with only category  $j$  in their list, i.e.,  $B_t = \{j\}$ . The store choice of these consumers are affected only by the variety level in category  $j$ . MNL leads to the following simple share equation.

$$s_t(x_j) = \frac{\exp\left(\frac{E[U(x_j)]}{\mu}\right)}{\exp\left(\frac{E[U(x_j)]}{\mu}\right) + \exp\left(\frac{E[U(z_j)]}{\mu}\right)} = \frac{A(x_j)}{A(x_j) + A(z_j)}$$

where  $A(x_j) = \sum_{i=1}^{x_j} e^{v_i/\mu}$ . Note that  $A(x_j)$  is nonnegative, increasing and concave in  $x_j$ . In words, attractiveness of a category is increasing in the variety level at a decreasing rate.

Now consider the customers with more than one category in their basket, the basket shoppers. A basket shopper type  $t$  buys one unit of one variant from each category in her basket ( $j \in B_t$ ). Her total basket utility is the sum of utilities of the categories in her basket. She chooses the store that maximizes her basket utility, and buys the whole basket from the store of her choice. The total basket utility is  $L_t(\mathbf{x}) = \sum_{j \in B_t} U(x_j)$ .  $L_t(\mathbf{x})$  does not follow a Gumbel distribution since Gumbel is not closed under addition. The choice probability of the consumers can not be specified in closed form with the true distribution

of  $L_t$ . To be able to use the MNL formula, we need to show that the distribution of  $L_t$  can be approximated with a Gumbel distribution. In Appendix A, we compare the true distribution of  $L_t$  and a Gumbel distribution whose first two moments are equal to that of  $L_t$ . We show that the Kolmogorov-Smirnov statistic, the maximum distance between the two distributions is fairly small across different parameter values. We also report the difference in the choice probabilities resulting from these distributions to illustrate the quality of such an approximation. As for empirical support for the approximation, Bell and Lattin (1998) use the MNL framework to estimate the basket utilities from consumer basket data and report that their model fits the store choice data well.

Let  $G_t(\mathbf{x})$  be a Gumbel random variable whose mean and variance are the same with  $L_t(\mathbf{x})$ , i.e.,

$$E[G_t(\mathbf{x})] = \sum_{j \in B_t} \mu \ln \sum_{i=1}^{x_j} e^{v_i/\mu} = \mu \ln \prod_{j \in B_t} \sum_{i=1}^{x_j} e^{v_i/\mu}$$

and  $V(G_t) = |B_t| \pi \mu^2 / 6$ . Hence the scale parameter of  $G_t$  is  $\mu \sqrt{|B_t|}$ . Consumer type  $t$  chooses the maximum of  $G_t(\mathbf{x})$  and  $G_t(\mathbf{z})$  to decide where to shop. This leads to the following expected share for the retailer:

$$s_t(\mathbf{x}) = \frac{\exp\left(\frac{E[G_t(\mathbf{x})]}{\mu \sqrt{|B_t|}}\right)}{\exp\left(\frac{E[G_t(\mathbf{x})]}{\mu \sqrt{|B_t|}}\right) + \exp\left(\frac{E[G_t(\mathbf{z})]}{\mu \sqrt{|B_t|}}\right)} = \frac{A_t(\mathbf{x})}{A_t(\mathbf{x}) + A_t(\mathbf{z})}$$

where  $A_t(\mathbf{x}) = \prod_{j \in B_t} A(x_j)^{\frac{1}{\sqrt{|B_t|}}}$ . Similar to  $A(x_j)$ ,  $A_t(\mathbf{x})$  is also increasing and concave in  $x_j$  for all  $j$ .

The number of type  $t$  consumers in the market is randomly distributed with mean  $\lambda_t$ . Suppose that the cost of providing a variety level  $x_j$  is  $c_j(x_j)$  and  $c_j(x_j)$  is increasing and

convex<sup>3</sup> in  $x_j$  (see Appendix B for a discussion of the convexity assumption). Let  $p_j > 0$  be the per-unit gross profit for all variants in category  $j$ <sup>4</sup>. Expected profit in category  $j$  is

$$\pi^j(\mathbf{x}) = p_j \sum_{t|j \in B_t} \lambda_t s_t(\mathbf{x}) - c_j(\mathbf{x})$$

We compare two modes of management. In the centralized system, total store profits are optimized over all category variety levels.

$$\max_{\mathbf{x}} \left\{ \Pi(\mathbf{x}) = \sum_{j=1}^N \pi^j(\mathbf{x}), \text{ s.t. } 0 \leq x_j \leq I_j, \forall j \right\} \quad (\text{OPT})$$

The dominant paradigm in retailing, category management, optimizes profits of each category. Since basket shoppers' store choice decision depends on the variety levels of other categories, category  $j$ 's optimal variety level depends on the vector of variety levels of the other categories denoted  $x_{-j}$ . Hence a game theoretic situation arises where category  $j$  will employ a best response, defined as follows: the strategy  $x_j(x_{-j})$  is category  $j$ 's best response to  $x_{-j}$  if

$$x_j(x_{-j}) = \arg \max_{x_j} \{ \pi^j(x_j, x_{-j}), \text{ s.t. } 0 \leq x_j \leq I_j \}, \forall j \quad (\text{CM})$$

The category management regime can be interpreted as a non-cooperative game between the managers of categories. Alternatively, it can be interpreted as a heuristic optimization algorithm where each category's variety level is chosen based on an optimization of category

---

<sup>3</sup>Convexity assumption is sufficient for the best response functions to be well-defined. In the case of concave cost functions, the analysis of the game would be complicated and even existence of equilibrium may not be guaranteed. For an example of a game with concave cost functions, see Cachon and Harker (2002).

<sup>4</sup>The assumption of identical prices of all variants in a category is restrictive. However, it may be realistic in certain cases. For example, a different color/size shirts of the same style have the same price tag.

profits given all the other categories' variety levels and categories are optimized iteratively until convergence. This scheme can lead to a vector of variety levels that correspond to a heuristic equilibrium, as described by Cachon and Kok (2002). A heuristic is at an equilibrium when the actions are optimal given the assumptions and the assumptions of the heuristic are consistent with the outcome of the action. In our model, at equilibrium each category chooses its optimal level of variety assuming other categories' variety levels fixed and other categories' optimal variety levels calculated in the same way are consistent with the assumed fixed levels.

Note that in the marketing models on multi-category choice based on panel data, consumers dynamically choose their shopping baskets at each trip. Our model is not dynamic, it is rather a snap shot in time. The consumers with apriori fixed basket types can therefore be interpreted as a realization of the dynamic choice processes. A more realistic alternative model to ours would have consumers who do not commit to purchase the whole basket. Consumers can make their store choice based on the apriori baskets but make category purchase decisions at the store. Qualitatively, our results will hold with this alternative model too. Because, the main driver of our results, consumers' store choice based on the variety levels of multiple categories, exists in that model.

In the next section, we characterize and analyze the solutions to decentralized and centralized regimes.

## 5.3 Analysis

The assortment decision  $\mathbf{x}$  is a vector of integers. Hereafter, we work with the continuous version of the problem because the continuous version has the following advantages. First, differentiability makes the derivation and presentation of the results easier and allows us to do comparative statics. Secondly, we show later in this section that the category management game is neither supermodular nor submodular, therefore we can not even guarantee the existence of an equilibrium under integrality constraints. Continuous version of this game is more amenable to analysis.

### 5.3.1 Category management

The game  $(CM)$  is neither supermodular nor quasi-supermodular. It is not supermodular because the cross partials can be negative. We can show that the quasi-supermodularity condition  $f(\mathbf{x}' \wedge \mathbf{x}'') < f(\mathbf{x}') \implies f(\mathbf{x}'') < f(\mathbf{x}' \vee \mathbf{x}'')$  does not hold for all  $\mathbf{x}$ : Suppose that increasing  $x'_1$  to  $x''_1$  is profitable when  $x_2 = x'_2$ . But this does not necessarily imply that increasing  $x'_1$  to  $x''_1$  is profitable when  $x_2 = x''_2 > x'_2$ .

Define  $\pi_{x_j}^j(\mathbf{x}) \equiv \frac{\partial \pi^j(x_j, x_{-j})}{\partial x_j}$ . Differentiating category  $j$  profits with respect to  $x_j$ , we get

$$\pi_{x_j}^j(\mathbf{x}) = \sum_{t|j \in B_t} p_j \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_j} - c'_j(x_j) \quad (5.1)$$

where

$$\frac{\partial s_t(\mathbf{x})}{\partial x_j} = \frac{A_t(\mathbf{z})}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^2} \frac{\partial A_t(\mathbf{x})}{\partial x_j}$$

and

$$\frac{\partial A_t(\mathbf{x})}{\partial x_j} = I_{\{j \in B_t\}} A_t(\mathbf{x}) A(x_j)^{-1} A'(x_j) \frac{1}{\sqrt{|B_t|}}$$

The second order derivative with respect to  $x_j$  is

$$\pi_{x_j x_j}^j(\mathbf{x}) = \sum_{t|j \in B_t} p_j \lambda_t \frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2} - c_j''(x_j)$$

where

$$\frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2} = \frac{A_t(\mathbf{z})}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^3} \left( -2 \left( \frac{\partial A_t(\mathbf{x})}{\partial x_j} \right)^2 + (A_t(\mathbf{x}) + A_t(\mathbf{z})) \frac{\partial^2 A_t(\mathbf{x})}{\partial x_j^2} \right) \quad (5.2)$$

and

$$\frac{\partial^2 A_t(\mathbf{x})}{\partial x_j^2} = I_{\{j \in B_t\}} \frac{1}{\sqrt{|B_t|}} A_t(\mathbf{x}) A(x_j)^{-1} \left( A(x_j)^{-1} A'(x_j)^2 \left( \frac{1}{\sqrt{|B_t|}} - 1 \right) + A''(x_j) \right)$$

Clearly,  $A_t(\mathbf{x})$  is concave in  $x_j$ , therefore  $s_t(\mathbf{x})$  is concave in  $x_j$ , for all  $t$ . Also  $c''(x_j) > 0$  by assumption. Consequently, the category profit  $\pi^j(\mathbf{x})$  is concave in  $x_j$ . When the payoff of each player is concave in player's action and the action space is compact, the existence of an equilibrium is guaranteed (Theorem 1.2. in Fudenberg and Tirole, 2000). The following theorem states this result and the equilibrium conditions.

**Theorem 5.3.1** (*Existence*) *A Nash equilibrium in (CM) exists and any Nash equilibrium of (CM) is characterized by the following conditions.*

$$\sum_{t|j \in B_t} p_j \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_j} - c_j'(x_j) = 0, \quad j = 1..N. \quad (5.3)$$

Best response functions  $x_j(x_{-j})$  are defined by (5.3). From the Implicit Function Theorem,

$$\frac{\partial x_j(x_{-j})}{\partial x_k} = - \frac{\pi_{x_j x_k}^j(\mathbf{x})}{\pi_{x_j x_j}^j(\mathbf{x})}$$

The sign of  $\partial x_j(x_{-j})/\partial x_k$  depends on the cross partial, which is

$$\pi_{x_j x_k}^j(\mathbf{x}) = \sum_{t|\{j,k\} \subseteq B_t} p_j \lambda_t \frac{\partial^2 s_t(\mathbf{x})}{\partial x_k \partial x_j}$$

where

$$\begin{aligned}\frac{\partial^2 s_t(\mathbf{x})}{\partial x_k \partial x_j} &= \frac{A_t(\mathbf{z})}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^3} \left( -2 \frac{\partial A_t(\mathbf{x})}{\partial x_k} \frac{\partial A_t(\mathbf{x})}{\partial x_j} + (A_t(\mathbf{x}) + A_t(\mathbf{z})) \frac{\partial^2 A_t(\mathbf{x})}{\partial x_k \partial x_j} \right) \\ &= \frac{A_t(\mathbf{z})}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^3} \frac{1}{|B_t|} A_t(\mathbf{x}) A(x_j)^{-1} A(x_k)^{-1} A'(x_j) A'(x_k) (A_t(\mathbf{z}) - A_t(\mathbf{x}))\end{aligned}\quad (5.4)$$

$A_t(\mathbf{z}) - A_t(\mathbf{x}) > 0$ , in  $X_t^+ = \left\{ \mathbf{x} \in \mathfrak{R}^N, \prod_{j \in B_t} A(x_j) < \prod_{j \in B_t} A(z_j) \right\}$ . Let  $X_{jk}^+ = \{ \mathbf{x} \in X_t^+, \forall t | \{j, k\} \subseteq B_t \}$ . In  $X_{jk}^+$ ,  $\pi_{x_j x_k}^j(\mathbf{x})$  is positive, hence  $\pi^j(\mathbf{x})$  is supermodular in  $(x_j, x_k)$ . Similarly let  $X_t^- = \left\{ \mathbf{x} \in \mathfrak{R}^N, \prod_{j \in B_t} A(x_j) > \prod_{j \in B_t} A(z_j) \right\}$  and  $X_{jk}^- = \{ \mathbf{x} \in X_t^-, \forall t | \{j, k\} \subseteq B_t \}$ . In  $X_{jk}^-$ ,  $\pi^j(\mathbf{x})$  is submodular in  $(x_j, x_k)$ . Notice that  $\mathfrak{R}^N - (X_{jk}^+ \cup X_{jk}^-)$  is not empty, unless there is only one shopping basket with  $j$  and  $k$  in it. We can not say anything about the sign of  $\pi_{x_j x_k}^j(\mathbf{x})$  in this region, there may be a surface cutting the region into two parts, where the sign is positive in one and negative in the other, or the sign may be alternating more than once.

Two categories  $j$  and  $k$  act as complements in  $X_{jk}^+$  and as substitutes in  $X_{jk}^-$ . It is interesting to observe that the relation of the categories is not uniform, i.e., complements everywhere or substitutes everywhere, but rather depends on the attractiveness of all the shopping baskets that contain these two categories. This relation depends on the variety levels of categories  $j$  and  $k$  and other categories that co-exist in the shopping baskets with  $j$  and  $k$ . It is not surprising that the total effect is attractiveness of the baskets weighted by the demand rate.

In the  $N = 2$  case, since there is only one basket type with both  $j$  and  $k$  in it, we can tell exactly the sign of  $\partial x_j(x_{-j})/\partial x_k$  in any region in  $\mathfrak{R}^2$ . The categories act as complements, for  $A(x_j)A(x_k) < A(z_j)A(z_k)$  and substitutes otherwise.

However, we can not partition the action space as clearly in the  $N > 2$  case, since we do not know the slope of the best response functions in  $\mathfrak{R}^N - (X_{jk}^+ \cup X_{jk}^-)$ . Nevertheless, in this complicated setting, we are able to characterize two distinct regions where the categories are complements and substitutes. In the region where categories act as complements, we can say more about the game and its' characteristics by using the theory of supermodular games.

**Theorem 5.3.2** (*Restricted game*) *Redefine the action space of the category management game (CM) such that each category's profit maximization problem is*

$$\max_{x_j} \{ \pi^j(x_j, x_{-j}), \text{ s.t. } 0 \leq x_j \leq z_j \}, \forall j \quad (\text{CM1})$$

*This game is supermodular. The equilibrium set is non-empty and has a largest and a smallest element. The largest element is the Pareto best and the smallest element is the Pareto worst equilibrium. The largest and the smallest equilibria are increasing in  $\mathbf{p}$  and  $\boldsymbol{\lambda}$  and decreasing in  $\mathbf{z}$  and  $\mathbf{c}$ .*

**Proof.** The payoff function  $\pi^j(\mathbf{x})$  is supermodular in  $(x_j, x_k)$  for  $\mathbf{x} \in X_{jk}^+, \forall j, k$ . Let  $X^+ = \bigcap_{j,k \in J} X_{jk}^+$ . Clearly payoff functions of all categories are supermodular in all pairs of variables in  $X^+$ . Also note that  $\{\mathbf{x} \leq \mathbf{z}\} \subseteq X_{jk}^+, \forall j, k$ . Hence  $\{\mathbf{x} \leq \mathbf{z}\} \subseteq X^+$ . The feasible action space  $\{\mathbf{0} \leq \mathbf{x} \leq \mathbf{z}\}$  is a non-empty, convex and compact set. Supermodularity of  $\pi^j$  in  $(x_j, x_k)$  for all  $j, k$  implies that  $\pi^j(x_j, x_{-j})$  has increasing differences in  $\{\mathbf{0} \leq \mathbf{x} \leq \mathbf{z}\}$ . Therefore (CM1) is a supermodular game. The payoff function  $\pi^j$  is continuous in  $x_j$  for each  $x_{-j}$  in  $\{\mathbf{0} \leq \mathbf{x} \leq \mathbf{z}\}$ . Then by Theorem 4.2.1. of Topkis (1998), the set of equilibrium points is a non-empty complete lattice and a greatest and a least equilibrium point exist.

Since the payoff to a player is increasing in other players' strategies, i.e.,  $\frac{\partial \pi^j(\mathbf{x})}{\partial x_k} \geq 0$  for all  $k \neq j$ , the largest element is the Pareto best and the smallest element is the Pareto worst equilibrium. This result follows from a simple stepwise improvement argument, see Section 2.2.3. of Vives (1999). We parameterize the game (*CM1*) by  $(\mathbf{p}, \boldsymbol{\lambda}, -\mathbf{z}, -\mathbf{c})$  and trivially show that  $\pi^j$  is continuous in  $x_{-j}$  for every  $(\mathbf{p}, \boldsymbol{\lambda}, -\mathbf{z}, -\mathbf{c})$  and has increasing differences in  $x_j$  and  $(\mathbf{p}, \boldsymbol{\lambda}, -\mathbf{z}, -\mathbf{c})$ . By Theorem 4.2.2. of Topkis (1998) we obtain the monotonicity results. ■

The action space defined in Theorem 5.3.2 is quite realistic, because we can not expect the retail store to be more attractive than the aggregate outside alternative. As long as the retail store attracts no more than half of the customers in the market, the results of Theorem 5.3.2 hold. Even in the original game with the unrestricted strategy space, if the cost of providing variety is too high and all the equilibria are in  $X^+$ , the results of Theorem 5.3.2 apply.

We are not able to characterize the number of equilibria neither in the original game (*CM*) or the restricted game (*CM1*). Numerical examples show that there can be multiple equilibria even in the  $N = 2$  case. Figure 5.1 provides an example that has a unique equilibrium and Figure 5.2 an example with three equilibria. We are also able to find examples with multiple equilibria for  $N > 2$ . Notice in Figure 5.1 that the best response functions are increasing-decreasing in the other category's variety level, which corresponds to the complementary categories in  $X^+$  and substitute categories in  $X^-$ .

The category management game (*CM*) can have multiple equilibria, but we are able

to characterize that there is at most one symmetric equilibrium in  $X^-$ . Since all the best response functions  $x_j(x_{-j})$  are decreasing in all components of  $x_{-j}$  in  $X^-$ , there can not be another symmetric equilibrium larger than any other symmetric equilibria in  $X^-$ . Therefore, there can be at most one symmetric equilibrium in the submodular region of the game (CM). The following proposition shows that a symmetric equilibrium is locally unique if it falls in a certain region of the action space.

**Proposition 5.3.3** *The symmetric equilibria  $\mathbf{x}$  in the game CM with  $A_t(\mathbf{x}) > \frac{(2\tau-1)}{(2\tau+1)}A_t(\mathbf{z}), \forall t$ , where  $\tau = \frac{2}{\sqrt{|B_t|}}$ , if there exists any, can not be part of a continuum of equilibria.*

**Proof.** As we mentioned before, the game (CM) has multiple equilibria, but we use the uniqueness conditions as an argument in our proof. Showing that a best reply map  $(x_j(x_{-j}))_{j=1..N}$  is a contraction is sufficient for a game to have a unique equilibrium. A sufficient condition for a mapping to be a contraction is that the Jacobian of  $\pi_{x_j}^j = 0, j = 1..N$  to be diagonally dominant, i.e.,  $\pi_{x_j x_j}^j(\mathbf{x}) + \sum_{k \neq j} |\pi_{x_j x_k}^j(\mathbf{x})| < 0, \forall \mathbf{x} \in X$ , (Section 2.5., Vives 1999). We have,

$$\sum_{k \neq j} |\pi_{x_j x_k}^j(\mathbf{x})| = \sum_{k \neq j} \sum_{t|\{j,k\} \subseteq B_t} p_j \lambda_t \left| \frac{\partial^2 s_t(\mathbf{x})}{\partial x_k \partial x_j} \right| \text{ and } \pi_{x_j x_j}^j(\mathbf{x}) = \sum_{t|j \in B_t} p_j \lambda_t \frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2} - c_j''(x_j)$$

Hence, a sufficient condition for diagonal dominance is to show that  $\left| \frac{\partial^2 s_t(\mathbf{x})}{\partial x_k \partial x_j} \right| < -\frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2}, \forall t$ .

Let  $\tau = \frac{1}{\sqrt{|B_t|}}$ ,  $\Phi = A_t(\mathbf{x}) + A_t(\mathbf{z})$  and  $\Delta = A_t(\mathbf{z}) - A_t(\mathbf{x})$ . From (5.2) and (5.4), we have

$$\begin{aligned} -\frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2} &= \frac{A_t(\mathbf{z})A_t(\mathbf{x})A(x_j)^{-1}\tau}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^3} (A(x_j)^{-1}A'(x_j)^2(2\tau A_t(\mathbf{x}) + (1-\tau)\Phi) - \Phi A''(x_j)) \\ &= \frac{A_t(\mathbf{z})A_t(\mathbf{x})A(x_j)^{-1}\tau}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^3} (A(x_j)^{-1}A'(x_j)^2(\Phi - \tau\Delta) - \Phi A''(x_j)) \end{aligned} \quad (5.5)$$

$$\left| \frac{\partial^2 s_t(\mathbf{x})}{\partial x_k \partial x_j} \right| = \frac{A_t(\mathbf{z})A_t(\mathbf{x})A(x_j)^{-1}\tau^2}{(A_t(\mathbf{x}) + A_t(\mathbf{z}))^3} A(x_k)^{-1}A'(x_j)A'(x_k) |\Delta| \quad (5.6)$$

Sufficient condition  $\left| \frac{\partial^2 s_t(\mathbf{x})}{\partial x_k \partial x_j} \right| < -\frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2}, \forall t$  does not hold in general but it holds when  $A(x_j) = A(x_k)$  and  $A'(x_j) = A'(x_k)$ . In  $X^+$ , we need  $(\Phi - \tau\Delta) > \tau\Delta$ . which is equivalent to  $A_t(\mathbf{x}) > \frac{(2\tau-1)}{(2\tau+1)}A_t(\mathbf{z}) \approx A_t(\mathbf{z})/6$ . Notice that this sufficiency condition is quite loose since all we need is the sum over the basket types to satisfy it not the individual entries. Moreover the nonnegative term  $-\Phi A''(x_j)$  is not used for the proof. In  $X^-$ ,  $|\Delta| = -\Delta$ , therefore (5.6) is equal to the second term in (5.5), and the first and third terms of (5.5) are positive. Hence the sufficiency condition is satisfied in  $X^-$  too. Due to the slacks not used in proving the sufficiency condition, we can argue (or formally prove) that diagonal dominance of the Jacobian is satisfied not only at the diagonal but also in the neighborhood of the diagonal.

Using this result, we define new games on smaller compact cubes around the diagonal of  $\mathfrak{R}^N$ . Since the best response map is a contraction there is a unique equilibrium in that cube. The implication of this result on the original game ( $CM$ ) is that there can not be a continuum of equilibria or equilibria too close to each other around the diagonal. ■

To give an idea, the coefficient  $\frac{(2\tau-1)}{(2\tau+1)}$  in the proposition is approximately 0.3 at its largest for  $|B_t| = 2$ . Therefore the proposition applies to a large part of the strategy space.

In a symmetric game, i.e., the categories have the same price, cost, competition, and demand vectors, we can rule out asymmetric equilibria in the supermodular region of the game. The following proposition establishes this result. This result can not be extended to the submodular regions of the game and the reason can be clearly seen in the proof.

**Proposition 5.3.4** (*Restricted game*) *There can not be asymmetric equilibria in the game*

(CM1) with symmetric data across all categories.

**Proof.** The proof is by contradiction. Let  $\mathbf{x}$  be asymmetric equilibrium with  $x_j > x_k$  for some  $j$  and  $k$ . By definition of equilibrium  $\mathbf{x}$  satisfies 5.3 for all categories including  $j$  and  $k$ .  $\sum_{t|j \in B_t} p \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_j} - c'(x_j) = 0$  and  $\sum_{t|j \in B_t} p \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_k} - c'(x_k) = 0$ . let's compare  $\frac{\partial s_t(\mathbf{x})}{\partial x_j}$  and  $\frac{\partial s_t(\mathbf{x})}{\partial x_k}$ .

$$\frac{\partial s_t(\mathbf{x})}{\partial x_j} = \frac{\partial s_t(x_j, x_{-j})}{\partial x_j} < \frac{\partial s_t(x_k, x_{-j})}{\partial x_k} < \frac{\partial s_t(x_k, x_{-k})}{\partial x_k} = \frac{\partial s_t(\mathbf{x})}{\partial x_j}$$

The first inequality follows from the concavity of  $s_t()$  and  $x_j < x_k$ . The second inequality follows from the supermodularity of  $s_t()$  and  $x_{-j} > x_{-k}$ . Also we have  $c'(x_j) > c'(x_k)$ . Hence the LHS of the equation for  $j$  is always strictly smaller than the one for  $k$ . As a result, both conditions can not be satisfied at the same time and  $\mathbf{x}$  can not be an equilibrium.

■

### 5.3.2 Optimal solution

Differentiating  $\Pi(\mathbf{x})$  with respect to  $x_j$  gives us

$$\frac{\partial \Pi(\mathbf{x})}{\partial x_j} = \sum_{t|j \in B_t} p_j \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_j} - c'_j(x_j) + \sum_{\substack{k=1 \\ k \neq j}}^N \sum_{t|j, k \in B_t} p_k \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_j} \quad (5.7)$$

Equation (5.7) captures the marginal effect of  $x_j$  on own and cross category sales. Also we have,

$$\frac{\partial^2 \Pi(\mathbf{x})}{\partial x_j^2} = \sum_{t|j \in B_t} p_j \lambda_t \frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2} - c''_j(x_j) + \sum_{k=1, k \neq j}^N \sum_{t|j, k \in B_t} p_k \lambda_t \frac{\partial^2 s_t(\mathbf{x})}{\partial x_j^2} \quad (5.8)$$

Since  $s_t(\mathbf{x})$  is concave in  $x_j$ , for all  $j$  and  $t$ , the first order condition (5.7) characterize the unique optimal variety level  $x_j$  for fixed  $x_{-j}$ . However, we are not able to show that the joint

optimization problem is unimodular, therefore satisfying the system of equations  $\frac{\partial \Pi(\mathbf{x})}{\partial x_j} = 0$  for  $j = 1, \dots, N$  is not sufficient for optimality. It is easy to find numerical examples even in  $N = 2$  case where  $\Pi(\mathbf{x})$  is not jointly concave. Hence we can not guarantee that the solutions to the first order conditions would yield only the global or local optimal solutions.

Let's define  $\kappa$  as the maximum  $p_j/(p_j + p_k)$  ratio among all pairs  $(j, k)$ , i.e.,  $\kappa = \max\{p_j/(p_j + p_k), (j, k) \in N \times N\}$ . Also recall that  $\tau = \frac{2}{\sqrt{|B_t|}}$ .

**Proposition 5.3.5** *The optimization problem (OPT) has at most one symmetric stationary point in  $X^- = \bigcap_{j,k \in J} X_{jk}^-$ . Moreover, any symmetric stationary point of  $\Pi(\mathbf{x})$ , with  $\frac{\tau(\kappa+1)+1}{\tau(\kappa+1)-1} > \frac{A_t(\mathbf{x})}{A_t(\mathbf{z})} > \frac{\tau(\kappa+1)-1}{\tau(\kappa+1)+1}$ , if there exists any, is a local optimal solution and can not be part of a continuum of local optima.*

**Proof.** Following the same line of analysis we did for the game (CM), we can show that  $\frac{\partial \Pi(\mathbf{x})}{\partial x_k \partial_j} < 0$  in  $X^-$ . Therefore, the solution to the first order condition for category  $j$  for fixed  $x_{-j}$  is decreasing in  $x_{-j}$  in  $X^-$ . Decreasing functions can intersect at the diagonal at most once. Hence there can be at most one symmetric solution in  $X^-$  to the first order conditions of (OPT). Similar to Proposition 5.3.3, we can show that the Jacobian of  $\frac{\partial \Pi(\mathbf{x})}{\partial x_j} = 0$  is diagonally dominant in  $X^+$  if  $A_t(\mathbf{x}) > \frac{\tau(\kappa+1)-1}{\tau(\kappa+1)+1} A_t(\mathbf{z})$  is satisfied and in  $X^-$  if  $A_t(\mathbf{x}) < \frac{\tau(\kappa+1)+1}{\tau(\kappa+1)-1} A_t(\mathbf{z})$ . A symmetric diagonally dominant matrix with negative diagonal entries is negative definite (Corollary 7.2.3 of Horn and Johnson, 1985). Hence, in this interval  $\Pi(\mathbf{x})$  is jointly concave and the symmetric stationary points of  $\Pi(\mathbf{x})$  are local maxima. ■

Clearly,  $\kappa \geq 2$ , for any price vector. Even when all categories have identical prices,

$\kappa = 2$  and the interval for  $\frac{A_t(\mathbf{x})}{A_t(\mathbf{z})}$  is  $[0.36, 2.78]$  as its largest case. As  $\kappa$  increases the interval shrinks.

**Theorem 5.3.6** (*Restricted problem*) Suppose  $x_j$  is restricted to be in  $[0, z_j], \forall j$  in the optimization problem (*OPT*). Let (*OPT1*) denote this restricted version. i) The set of stationary points of (*OPT1*) defined by  $\frac{\partial \Pi(\mathbf{x})}{\partial x_j} = 0$  has a smallest and a largest element that are monotone increasing in  $\mathbf{p}$  and  $\boldsymbol{\lambda}$  and decreasing in  $\mathbf{z}$  and  $\mathbf{c}$ . ii) The set of optimal points  $\arg \max_{x \in X^+} \Pi(\mathbf{x})$  is non-empty, has a smallest and a largest element that are monotone increasing in  $\mathbf{p}$  and  $\boldsymbol{\lambda}$  and decreasing in  $\mathbf{z}$  and  $\mathbf{c}$ .

**Proof.** The existence of a largest and a smallest equilibrium and the monotonicity of the set in the parameters follow from applying the theory of supermodular games to the embedded game defined by the first order conditions, similar to the proof of Theorem 5.3.2. However, we can not guarantee whether the largest stationary point is a local maximum, let alone global maximum. The result on the set of optimal points follow from the optimization of a supermodular function (Theorem 2.8.3 in Topkis, 1999). ■

### 5.3.3 Comparison of the centralized and decentralized solutions

Let  $X^d$  (where  $d$  denotes decentralized) be the set of Nash equilibria in the category management game (*CM*) and  $X^c$  (where  $c$  denotes centralized) be the set of stationary points of the profit function in (*OPT*). The following theorems show that the category management regime never finds the (local) optimal solution as long as there are basket shoppers.

**Theorem 5.3.7** *Suppose that there exists at least one consumer type with more than one category in their basket and positive demand rate, i.e.,  $\lambda_t > 0$ . The set of Nash equilibria  $X^d$  in the category management game (CM) does not contain the global optimal solution to (OPT).  $X^d$  does not contain any of the local optimal solutions either. If data is symmetric across all categories, then there is a symmetric stationary point larger than the largest symmetric equilibrium.*

**Proof.** Compare the first order conditions to the optimization problem (OPT) and the game (CM):

$$\Pi_{x_j}(\mathbf{x}) - \pi_{x_j}^j(\mathbf{x}) = \sum_{k=1, k \neq j}^N \sum_{t|j, k \in B_t} p_k \lambda_t \frac{\partial s_t(\mathbf{x})}{\partial x_j} > 0$$

if there is one basket type with more than one category and positive demand rate. Therefore, no point  $\mathbf{x}$  can satisfy both set of conditions, i.e.,  $X^d \cap X^c = \emptyset$ . For fixed vector of variety levels of other categories,  $x_{-j}$ , the solution to the first order conditions for category  $j$  ( $x_j^c(x_{-j})$ ) is always larger than the best response of category  $j$  in the game ( $x_j^d(x_{-j})$ ). In the symmetric game, let the largest symmetric equilibrium have variety level  $x^s$  for all categories. It is sufficient to consider a an arbitrary category. We have  $x_j^c(x_{-j}^s) > x_j^d(x_{-j}^s) = x^s$ . We also know that  $x_j^c(x_{-j}^s)$  does not go to infinity since it is decreasing in  $x > z$ . Therefore,  $x_j^c(x_{-j}^s)$  will cross the 45degree line at least once more and the intersection is a stationary point of (OPT), i.e.,  $\exists x^1 : x_j^c(x_{-j}^1) = x^1 > x^s$ . ■

**Theorem 5.3.8 (Restricted Problem)** *The global optimal solution of the optimization problem (OPT1),  $\mathbf{x}^c$ , has strictly higher variety in at least one category than any of the Nash*

equilibria of the game (CM1), including the Pareto best equilibrium, i.e.,  $\forall x^d \in X^d, \exists j, x_j^c > x_j^d$ .

**Proof.** Let  $\mathbf{x}^c$  be the optimal solution to (OPT1) and  $\mathbf{x}^d$  be any equilibrium of the game (CM1). By contradiction, suppose that  $\mathbf{x}^c < \mathbf{x}^d$ .

$$\Pi(\mathbf{x}^c) = \sum_j \pi^j(x_j^c, x_{-j}^c) < \sum_j \pi^j(x_j^c, x_{-j}^d) < \sum_j \pi^j(x_j^d, x_{-j}^d) = \Pi(\mathbf{x}^d)$$

This is a contradiction with the optimality of  $\mathbf{x}^c$ . The first inequality follows from  $\frac{\partial \pi^j(\mathbf{x})}{\partial x_k} \geq 0$  for all  $k \neq j$  and  $\mathbf{x}$ , and the second inequality follows from the definition of the equilibrium, i.e.,  $x_j^d = \arg \max_{x_j} \pi^j(x_j, x_{-j}^d)$ . ■

Due to supermodularity of profit functions in the restricted strategy space, discrete versions of Theorems 5.3.2, 5.3.6 and 5.3.8 can be easily established with the following slight modification in Theorem 5.3.8. The optimal solution has higher (not strictly higher) variety than the equilibrium solution in at least one category.

#### 5.3.4 Numerical examples

It is worthwhile to present here some pathological examples where category management performs rather poorly. Section 5.6 is devoted to numerical analysis and further explores the magnitude of the loss due to category management. Here we are interested in demonstrating potential outcomes rather than the average and worst case analysis of the magnitude of the loss. Figure 5.2 presents the optimal solution and the three equilibria of the category management game. Among the three equilibria, the zero equilibrium (EQ1) and the largest

one (EQ3) are stable, (EQ2) is not. Even the best category management equilibrium (EQ3) provides significantly less variety than the optimal solution. What is even worse is that category management may choose to eliminate both categories (EQ1: zero equilibrium), although there is a profitable positive variety equilibrium. Figure 5.3 presents an example where the unique equilibrium of category management is a zero equilibrium, i.e., eliminating both categories, while the optimal solution has positive variety for both categories, i.e., category management has no choice but to shut down the store in a case where the store can make positive profits at the optimal solution.

## 5.4 Properties of the Optimal and Equilibrium Assortments

In this section, we investigate the effect of perturbations on the optimal and equilibrium solutions and make inferences about interesting questions on how to allocate resources to categories. We assume that the optimal solution,  $\mathbf{x}^c$ , is an interior solution, so that it will be responsive to the small changes in the parameters. Similarly, we consider the equilibria that are interior solutions.

### 5.4.1 Depth vs. breadth

Consider the scenario with  $N$  categories but focus on two categories  $j$  and  $k$ , where the total demand for a category remains constant, but the proportion of basket shoppers is increasing. We can achieve this simply by increasing the demand rate for basket type with  $j$  and  $k$  in it, and decreasing the single category shoppers' rate at an equal rate. Let subscripts  $j$  and  $k$  denote the individual category shoppers, and  $jk$  denote the basket type

with  $j$  and  $k$  in it and  $\lambda_{jk} = \phi, \lambda_j - \phi, \lambda_k - \phi$ . Note that  $\phi = 0$  represents the case where there are no basket shoppers.

By Implicit Function Theorem (IFT), the sign of  $\frac{\partial x_j^c}{\partial \phi}$  is the same as the sign of  $\frac{\partial^2 \Pi(\mathbf{x})}{\partial \phi \partial x_j} \Big|_{\mathbf{x}^c} = -p_j s'_j(x_j) + (p_j + p_k) \frac{\partial s_{jk}(x_j, x_k)}{\partial x_j} \Big|_{\mathbf{x}^c}$ , which can be positive or negative. Therefore, a shift in mix of customers toward more basket shoppers may result in an increase or a decrease in the variety levels in the categories of interest. What is interesting is that the direction of change in the variety levels remains the same as the shift continues, i.e., if the first shift from  $\phi = 0$  to arbitrarily small  $\varepsilon$  causes the variety levels to increase (decrease), increasing  $\phi$  all the way to  $\min(\lambda_j, \lambda_k)$  would keep increasing (decreasing) the optimal variety levels. The question that needs to be answered then is which direction it takes first and there is no clear answer to that. It seems from numerical examples that the optimal variety levels usually increase with more basket shoppers, but decrease only in examples with lower category costs ( $\mathbf{c}$ ) and lower competition ( $\mathbf{z}$ ). Similarly in the equilibrium solution  $\mathbf{x}^d$ , the direction of change depends on  $-s'_j(x_j) + \frac{\partial s_{jk}(x_j, x_k)}{\partial x_j} \Big|_{\mathbf{x}^d}$ , which also can be positive or negative. Although it seems intuitive that the equilibrium variety levels should go up (down) as proportion of basket shoppers increase in the supermodular (submodular) region of the game, i.e.,  $X_{jk}^+(X_{jk}^-)$ , counter examples can be easily generated. In most examples, with high enough  $\mathbf{z}$  and  $\mathbf{c}$ , the optimal variety levels increase and equilibrium variety levels decrease with the proportion of basket shoppers.

It is possible to interpret the composition of a category at different levels, TVs, VCRs, DVD players can be considered as categories for an electronics retailer and bikes, board games, baby dolls for a toy retailer. The above result suggests that with the increase of

basket shoppers, assuming the competition level is high enough, the retailers must choose higher levels of variety in each of these categories. At a broader level, electronics and toys can be considered as categories themselves. Looking at the retail industry with the broader category interpretation in mind, we see that there three types of retailers: category killers, discount retailers and the likes of Cost-Co. Category killers such as Best Buy, PetSmart, and Toys-R-us have a very high level of variety in one category. The discount retailers such as Wal-Mart and K-Mart offer more than one category and relatively less variety in each category. At the extreme, Cost-Co offers almost every possible category, but has little variety in each. Notice that our model does not necessarily explain the low variety levels in the multi-category retailers. The relative positions of these retailers can not be explained just by variety. Different business models of these retailers determine many other important factors such as the target consumer segments, pricing strategies, levels of information and support services, and the cost structures based on their market power and operational implications of their choices.

What happens to the optimal profit? By the envelope theorem,  $\frac{d\Pi(\mathbf{x}^c)}{d\phi} = \frac{\partial\Pi(\mathbf{x})}{\partial\phi}\Big|_{\mathbf{x}^c} = -p_j s_j(x_j^c) - p_k s_k(x_k^c) + (p_j + p_k) s_{jk}(\mathbf{x}^c)$ . Assuming identical prices and symmetric categories  $j$  and  $k$ , we can show that  $\frac{d\Pi(\mathbf{x}^c)}{d\phi} > 0$  iff  $\mathbf{x}^c > \mathbf{z}$ . This means that regardless of the direction of change in the variety levels in the optimal solutions, the optimal profit is increasing (decreasing) with the ratio of the basket shoppers if the competition is weak (strong), i.e., optimal variety levels of the retailer is less (more) than the variety levels at the outside alternative. The same result can be stated for the equilibrium solution, that is  $\frac{d\Pi(\mathbf{x}^d)}{d\phi} > 0$  iff  $\mathbf{x}^d > \mathbf{z}$ . Notice that these properties are local and hold only in the neighborhood of the

optimal and equilibrium solution.

#### 5.4.2 Expanding the assortment: Adding new categories

Suppose that we are considering to add a new category to the existing assortment,  $S$ . There are two candidates  $k$  and  $l$ , where  $k$  is a basket category relative to  $l$ . Ideally we would like to compare the optimal variety levels and profits of the  $S \cup \{k\}$  and  $S \cup \{l\}$ . That comparison is analytically not possible, because we can not obtain closed form expressions for the optimal variety level or the profits. Nevertheless, we can get around this problem by making local comparisons. We start with  $l$  that is a slight perturbation of  $k$ , by taking the solution to the assortment with  $S \cup \{k\}$  and perturbing the demand rates of baskets that include  $k$  in the direction of  $l$ . If the profit response to the perturbation is positive, we can infer that exchanging  $l$ , an independent category, with  $k$ , a basket category is a profitable decision.

Assume that the demand rates of all basket types including  $k$  or  $l$  are the same except for the basket which has  $j$  and  $k$  in it, whose demand rate is denoted  $\lambda_{jk}$ . Let  $\lambda_{jl} = \lambda_{jk} + \delta$  and  $\lambda_l = \lambda_k - 2\delta$ . Notice that the total demand remains the same, but more customers of  $l$  are basket shoppers compared to  $k$ . Clearly,  $x_j^c$  or  $x_j^d$  is increasing with  $\delta$ , because  $\lambda_{jk}$  is increasing. The direction of change in  $x_k^c$  and  $x_k^d$  is similar to the § 5.4.1.

In this setting, we are more interested in the profit impact of the perturbation. By the Envelope Theorem,  $\frac{d\Pi(\mathbf{x}^c)}{d\delta} = \frac{\partial\Pi(\mathbf{x})}{\partial\delta} \Big|_{\mathbf{x}^c} = (-2p_k s_k(x_k) + (p_j + p_k)s_{jk}(x_j, x_k)) \Big|_{\mathbf{x}^c}$ . Assuming identical prices,  $\frac{d\Pi(\mathbf{x}^c)}{d\delta} > 0$  iff  $\mathbf{x}^c > \mathbf{z}$ . Therefore, under less competitive situations, i.e.,  $\mathbf{x}^c > \mathbf{z}$ , it is more profitable to include the basket category in the assortment. If

$\mathbf{x}^c < \mathbf{z}$  (strong competition), then it is better to diversify the assortment by including the independent category instead of the basket category. The intuition behind this result is the following. When the competition is strong, a retailer cannot attract basket shoppers by increasing variety in one category only. It is better to go after the customers of the independent category, and be a niche player, because single category customers are more responsive to changes in the variety level of a category. Again analogous results can be stated for the equilibrium profits.

### 5.4.3 Which categories should have more variety?

In § 5.4.1, we increased the ratio of basket shoppers. In § 5.4.2, we compared adding a basket category or an independent category, by converting basket shoppers to individual category shoppers. Now we perturb two categories  $k$  and  $l$  in the assortment at the same time. Let's decrease  $\lambda_k$ , increase  $\lambda_{jk}$ , increase  $\lambda_l$  and decrease  $\lambda_{jl}$  by  $\delta$ . Note that the net change in total demand for each category is zero. With  $\delta > 0$ ,  $k$  becomes more of a basket category (similar to § 5.4.2) and  $l$  an independent category.

$\frac{\partial x_k^c}{\partial \delta}$  has the same sign with  $\left. \frac{\partial \Pi(\mathbf{x})}{\partial \delta} \right|_{\mathbf{x}^c} = -p_k s'_k(x_k^c) + (p_j + p_k) \frac{\partial s_{jk}(x_j^c, x_k^c)}{\partial x_k}$  and  $\frac{\partial x_l^c}{\partial \delta}$  has the same sign with  $p_l s'_l(x_l^c) - (p_j + p_l) \frac{\partial s_{jl}(x_j^c, x_l^c)}{\partial x_l}$ . Hence  $x_k^c$  and  $x_l^c$  are more likely to move in reverse directions, but there may also be cases where they move in the same direction. If the categories are symmetric,  $x_k^c = x_l^c$ , and they move in the reverse direction with increasing  $\delta$ .

$x_k^c$  increases with  $\delta$  when  $p_k s'_k(x_k^c) < (p_j + p_k) \frac{\partial s_{jk}(x_j^c, x_k^c)}{\partial x_k}$ . In § 5.3.1, we showed that  $\frac{\partial s_{jk}(x_j, x_k)}{\partial x_k}$  is an increasing-decreasing function of  $x_j$  with its' peak at  $x_j = z_j z_k / x_k^c$ . Hence

there may be an interval of  $x_j^c$  say,  $(\underline{x}_j^c, \bar{x}_j^c)$ , where the inequality is satisfied and  $x_k^c$  increases with  $\delta$ , i.e. the optimal variety level of the basket category increases and the optimal variety of the independent category decreases. For  $x_j^c < \underline{x}_j^c$  and  $x_j^c > \bar{x}_j^c$ , the reverse is true. The intuition behind this result is the following. If  $x_j$  is too low, it is not possible to attract enough many basket shoppers by increasing the variety of basket category, so it is better to invest in the independent category. If  $x_j$  is in the interval, the investment in the basket category pays off. As  $x_j$  gets too large, there is no need to invest in category  $k$  since  $j$ 's variety level is attracting many basket customers to category  $k$  as well, and the returns in the independent category is higher.

Similar analysis yields that the equilibrium variety levels display the same properties, but the  $x_k^d$  increases with  $\delta$  when  $p_k s'_k(x_k^d) < p_k \frac{\partial s_{jk}(x_j^d, x_k^d)}{\partial x_k}$ . This implies that the interval of  $x_j^d$  that  $x_k^d$  increases with  $\delta$  is a subset of  $(\underline{x}_j^c, \bar{x}_j^c)$ . Therefore, there may be cases where it is optimal to increase the variety of the basket category and decrease the variety of the independent category, but the equilibrium solution does the reverse.

In this section, we investigated the sensitivity of the variety levels to the proportion of basket shoppers in a category. In doing so, we kept everything constant and changed as little as possible from the initial setup in order to isolate the main effect. It is clear that when the demand rates of many categories change simultaneously, the net effect will be a combination of the effects discussed above.

## 5.5 Aggregation Heuristic

The category management equilibrium solution will emerge as the outcome of a natural iterative process. Whether a heuristic optimizes categories in an iterative manner, or category managers react to changes to maximize their own profits, the process will converge to one of the category management equilibria. We will see in the numerical results section that the loss due to decentralization can be significant. On the other hand it is not easy to implement the centralized solution, because unlike category management we need to know beforehand what the optimal solution is and impose that solution on all categories.

Finding the optimal solution involves estimating the number of customers of all basket types and  $N$ -dimensional optimization of an ill-behaved function. In order to make our model work, the demand for each type of basket shoppers needs to be estimated. Given the large number of categories, e.g., in food retailing, number of types of basket shoppers increase exponentially and it may not be possible to estimate the demands accurately for each basket type. Even when the data is available, the amount of time required for estimation is exponential in  $N$ . In the optimization part, since the objective function is not well-behaved, the computational requirement of a search procedure is also exponential.

In this section we introduce a decentralized heuristic that promises the best of both worlds (i.e., the simplicity of category management and the profit level of the optimal solution). The main weakness of category management is that it fails to recognize the intercategory effects of variety decisions and underestimates the marginal benefit of variety. A category manager does not need to (and probably cannot) know the specifics of the intercategory dependences. Incorporating an aggregate representation of all other categories

into her decision mechanism will push her choice of variety level in the direction of the optimal solution.

Our aggregation heuristic does just that. From each category's perspective, it simplifies the world to a two-category problem, where there is the category of interest,  $j$ , and an aggregate category that represents all other categories,  $A_j$ . We estimate the parameters of the aggregate category with a weighted average of categories. Each category manager solves their two-category problem and implement the variety level of their own category. Below is a formal description of the heuristic.

For each category  $j=1,\dots,N$

1. Transform the  $N$ -category problem to a 2-category problem (denote  $(j,A)$ )
  - 1.a.  $j \leftarrow j, \lambda_j \leftarrow \lambda_j, p_j \leftarrow p_j, c_j \leftarrow c_j, z_j \leftarrow z_j,$
  - 1.b.  $A_j \leftarrow \{1, \dots, N\} \setminus \{j\}$
  - 1.c.  $\lambda_{A_j} \leftarrow \sum_{t, j \notin B_t} \lambda_t$
  - 1.d.  $\lambda_{j,A_j} \leftarrow \sum_{t, j \in B_t, |B_t| > 1} \lambda_t$
  - 1.d.  $p_{A_j} \leftarrow \frac{\sum_{t, j \in B_t, |B_t| > 1} \lambda_t p_t}{\lambda_{A_j}}$
  - 1.e.  $c_{A_j} \leftarrow \frac{\sum_{t, j \in B_t, |B_t| > 1} \sum_{k \in B_t} \lambda_t c_k}{\sum_{t, j \in B_t, |B_t| > 1} \sum_{k \in B_t} \lambda_t}$
  - 1.f.  $z_{A_j} \leftarrow \frac{\sum_{t, j \in B_t, |B_t| > 1} \sum_{k \in B_t} \lambda_t z_k}{\sum_{t, j \in B_t, |B_t| > 1} \sum_{k \in B_t} \lambda_t}$
2. Find the optimal solution  $(x_j^h, x_{A_j}^h)$  to (OPT) with above parameters..
3. Set variety level of category  $j$  to  $x_j^h$ .

It is important to note that this heuristic is not iterative, it simply solves  $N$  separate

2-category problems. The computational requirements for this heuristic is in the order of  $N^3$  ( $N$  two dimensional optimization problems). The data requirements for this formal heuristic is the same as finding the optimal solution, i.e., requires  $\lambda_t$  for all  $t$ . However, in a practical application of this method,  $\lambda_{j,A_j}$  can be estimated simply by the number of customers whose baskets consist of category  $j$  and something else and  $\lambda_{A_j}$  is the number of customers that are not interested in category  $j$ . Therefore the input length to the heuristic is in the order of  $N$ . We report the performance of the Aggregation Heuristic in Section 5.6. We suspect that other aggregation techniques other than the proposed weighted average will perform similar satisfactory results.

## 5.6 Numerical Results

This section reports a numerical study to further explore the category coordination problem. In particular, we are interested in the impact of problem size, and other design parameters on the magnitude of profit loss due to category management in comparison to the optimal solution. We also wish to compare the variety levels of the *basket* and *independent* categories.

We generate 84 scenarios from the combination of the following parameters. In these 84 scenarios, the data for all categories is symmetric and we consider only the symmetric equilibria.

$$\begin{aligned} p_j &= 1 & z &= \{5, 10\} \\ A(x) &= x & N &= \{2, 3, 5\} \\ c_j(x) &= cx & c &= \{2, 4\} \end{aligned}$$

We assume a total demand of  $\lambda_T = 100$  for each category and divide into the baskets containing that category according to the basket ratio vectors given in Table 5.1. For example, the vector  $\{.2, .6, .2\}$  for  $N = 3$  implies that 20% of the demand goes to baskets of size one, 60% goes to baskets of size two, and 20% to baskets of size three. Now for category 1 for example there are two baskets of size two,  $\{1,2\}$  and  $\{1,3\}$  and 60 units of demand is equally divided among them. Resulting demand vector is  $\lambda_1 = 20, \lambda_{12} = 30, \lambda_{13} = 30$  and  $\lambda_{123} = 20$ . The demand vectors for other categories and for larger  $N$  are computed similarly.

Out of the 84 instances, category management had a zero equilibrium in 14 instances as unique equilibrium and resulted in 100% performance loss. 25 instances out of the remaining 70 instances, category management had multiple equilibria one of which was the zero equilibrium. In comparing the category management equilibria to the optimal profits, we took into account only the best equilibrium in the 25 multi-equilibria cases. Table 5.2 presents summary data on the performance of category management equilibria. The profit loss due to decentralization on average is 28% and the variety level provided by category management is on average 45% less than the optimal variety level. Figures 5.4 and 5.5 illustrate the average profit loss and the average decrease in variety level with respect to  $N$ . The performance of decentralized regime, deteriorates with  $N$ ,  $z$ ,  $c$ , and proportion of basket shoppers. That is with more categories, tougher competition, higher costs of variety and more basket shoppers, the coordination of categories becomes more and more important.

In 5.4.1 we noted that the optimal and equilibrium variety levels may go in reverse

direction as the ratio of basket shoppers increases. Indeed, equilibrium variety levels decreased and optimal variety increased in all of the 84 instances. In terms of the impact of more basket shoppers on profits, equilibrium profits increased in 38 out 84 instances whereas optimal profits increased for 68 of the instances. This means that a retailer prefers more basket shoppers mostly but may not take advantage of it, if it employs category management techniques.

It is not obvious which one to choose between *basket* and *independent* categories in expanding assortments or revising variety levels of existing categories. Our results show that the *basket* categories are preferred 81% of the time in the optimal solution and 45% of the time in the equilibrium. We also observe that *basket* categories are preferred more with higher  $N$ , lower  $c$  and  $z$ .

We briefly discuss some of the other issues that can be numerically explored. Recall that Proposition 5.3.4 ruled out asymmetric equilibria in the symmetric restricted game. We have not found any asymmetric equilibria in the unrestricted regions of the problems with two symmetric categories. In the instances with two asymmetric categories, the category with higher price, lower cost, or lower competition has higher variety level. The profit loss due to category management decreases in the cases with strictly concave attraction function  $A(x)$  and strictly convex cost function  $c(x)$  compared to the cases with linear functions.

Table 5.3 presents the performance of the Aggregate Heuristic in comparison to category management and optimal solution for a set of examples. In all examples, a major portion of the profit loss due to decentralization is recovered. The maximum loss with Aggregation Heuristic is 2.3% in comparison to 43% of the category management. The

average loss decreases to 0.5% from 22%. The variety levels are closer to the optimal. Note also that the heuristic variety levels can be higher than the optimal solution in some instances. This remarkable improvement is achieved by understanding the weakness of category management and giving it a nudge in the right direction.

## 5.7 Discussion

Historically, the conventional wisdom in retailing when managing product variety has been ‘more is better’. Our model supports the intuition of the retailers by showing that category based analysis may be misleading and provides less variety than optimal.

We study the assortment planning problem with multiple categories where basket shopping consumers choose to patronize a retailer or an outside alternative based on their total utility across categories of interest. In this setting, the benefits of variety in a category depends on other categories’ variety levels. We investigate these interactions under centralized and decentralized management regimes. The first lesson from this research is that variety levels of categories are not always complements or always substitutes. Rather, they may act as strategic complements or substitutes depending on their levels in comparison to the outside alternative. They act as complements at low variety levels and substitutes at high variety levels. This is the first example in the literature where the type of interaction between two decisions may change based on their levels.

The main conclusion from this research is that category management regime, which optimizes each category independently, leads to a suboptimal heuristic equilibrium with less variety than optimal. We show in a numerical study that the profit loss due to

decentralization is 28% on average and variety is 45% less than the optimal. We observe that the performance get worse as the number of categories increases, making the recognition of the inefficiency more and more important. We show that putting more categories together and having a higher proportion of basket shoppers yields higher profits if the competition is weak. Contrary to intuition, this may also lead to higher variety in both categories in the optimal solution. We classify categories into *basket* and *independent* categories and characterize the conditions for which type of categories should be preferred in adding new categories to an existing assortment or in assigning more variety. Finally, we note that equilibrium decisions may go in the reverse direction of the optimal decisions.

Understanding what category management has to offer and unfolding its weakness enabled us to develop the Aggregation Heuristic. It simplifies the  $N$ -category problem to  $N$  2-category problems and works similar to the decentralized regime. In short, it reduces the computational requirements of the centralized solution substantially and produces solutions that recovers most of the profit loss due to decentralization.

The results presented in this paper can be interpreted from other perspectives than variety in retailing. Variety in our model can be thought as a proxy for any determinant of the attractiveness of a category, such as service levels, pricing, etc. Also merchandise categories can be thought of different business units, for example consumers choose which bank to patronize based on the attractiveness of a bank's deposits, lending, investment services. Analogous results in those cases can be derived using similar models and analyses.

The variety level at the outside alternative is assumed to be exogenously given in our model. In reality outside alternative is multiple strategically acting firms. One interesting

extension to our model is to incorporate competition, although it may be difficult to obtain analytical results. Another interesting extension is include pricing across and within categories.

A final lesson to highlight from this research is that modeling the demand side more carefully by taking consumer choice process into account helped us get new and interesting results on a well studied problem. We suspect that this would be the case with many other operations management problems.

## Chapter 6

# Supplier Competition in Retail

## Supply Chains

The first two parts of the dissertation concentrate on the retailer's assortment planning problem. A retailer is bound to select its assortment from the products available from the manufacturers. Manufacturers compete with each other through retailers' shelf space and employ several tactics to influence the retailer's assortment decisions in their favor. Suppliers employ quantity discounts on total sales volume as an incentive for the retailer to carry more of their product rather than the competitors'. In the case of two symmetric manufacturers, the retailer ideally would like to choose the same stocking quantity from each. With quantity discounts however, the retailer may choose to carry more inventory of one product and less of the other.

The third part of the dissertation focuses on supplier competition in a retail supply chain. We extend the retailer's assortment planning problem to include the price competition

between suppliers, i.e. each supplier use price or discounts as a competitive tool to effect retailer's assortment planning decisions in her favor. Consider the following two-echelon supply network. There are multiple suppliers each selling a single product to retailers selling one product each. Retailers can be under centralized or decentralized control. Multiple retailers under centralized control is equivalent to a single retailer selling multiple products. Suppliers offer contracts to the retailer(s), whose terms are the wholesale prices per unit under a price-only contract and supplier's share of revenues under revenue sharing contracts. Given a set of contracts, retailer(s) choose the stocking quantities of the products to meet the single period stochastic demand. The products are strategic substitutes, that is demand of a product is negatively effected by other products' stocking quantities. This negative externality between products can be due to consumers switching to other products when their first choice is not available (substitution) or when consumer's choice behavior is effected by the shelf space (i.e., a product with larger presence in the store may get a larger fraction of demand, hence reducing demand to other products).

This supply network provides a framework to study different industry models in practice. Consider the following centralization/decentralization schemes: *i)* All suppliers and retailers are controlled by a central planner. *ii)* Each product is sold through a jointly owned supplier-retailer pair. *iii)* A single supplier sells to independent retailers (similar to a distribution system). *iv)* All suppliers and retailers are independent. *v)* A single supplier sells all products to a single retailer (one-to-one supply chain with multiple products). *vi)* Multiple suppliers sell to a common retailer. We study these models under two contract types, wholesale price contracts and revenue sharing contract types. We study revenue

sharing contract types rather than quantity discounts because effectively they are similar, and revenue sharing contracts are nicely parameterized whereas there can be infinitely many forms quantity discount contracts.

We characterize the equilibria under wholesale price contracts and we are able to compare them when products are symmetric. We find that horizontal integration of either echelon (i.e., centralization at the retailer level or the supplier level) increases the wholesale prices and reduces the total stock level in the system regardless of the other echelon being centralized or decentralized. Vertical integration of supplier-retailer pairs on the other hand increases the total stock in the system. Horizontal integration eliminates the competition effect whereas vertical integration eliminates the double marginalization effect. Hence their effect on total inventory is in different directions.

Introducing competition always leads to higher inventories under wholesale price contracts. Competition leads to higher inventory than system optimal when supplier-retailer pairs are vertically integrated, resulting in less profits. However, it may be desirable to introduce competition and increase the inventory in the system. For example when there is a centralized retailer who buys from suppliers with a wholesale price contract, having decentralized suppliers leads to an inventory level that is closer to (but still smaller than) the system optimal in comparison to having centralized suppliers. The retailer wins in two ways from competition between suppliers: lower wholesale prices and increased sales.

We present many interesting findings regarding revenue sharing contracts. Revenue sharing contracts eliminate the double marginalization effect always in the decentralized retailers case and when revenue share of all products are identical in the centralized retailer

case. Surprisingly, it also eliminates the competition effect at the supplier level. Competition leads to more inventories when it is at the retail level, but has no effect when it is at the supplier level. Hence, the system optimal is reached with revenue sharing contracts when there is a centralized retailer and more than system optimal inventory is held when the retailers are decentralized. We also show that revenue sharing contracts leads to higher inventory than wholesale price contracts in any of the configurations. We demonstrate with numerical examples that independent suppliers selling through a centralized retailer are significantly better off under revenue sharing contracts than the wholesale price contracts and the retailer is much worse off.

Our paper is novel in many ways. The fundamental difference of our model is that the products are supplied by different suppliers, not by the same supplier. Suppliers compete with each other through a centralized retailer's optimization decisions or equilibria of a noncooperative game between retailers. Secondly, in our setting supply chain coordination is not the main issue. The suppliers are interested in maximizing their own profits. Contracts serve as a tool for competition rather than achieving system optimality. A supplier may even prefer the retailer to order more than system optimal quantity of her product, because she is basically stealing share from the other suppliers. Thirdly, to the best of our knowledge, ours is the first paper where supply chain coordinating contracts are employed in a multiple competing products setting. In supply chain coordination with contracts, usually contract parameters can be arbitrarily chosen where the realistic outcomes are the result of bargaining power or negotiations between the supplier and the retailer. In our model, the contract parameters emerge as the result of the competition between suppliers.

For example, the equilibrium revenue share of the retailer provides a lower bound to the outcome of practical and more complicated scenarios. Because even when the retailer has no negotiation power, the supplier will not offer a revenue share lower than the equilibrium outcome of our setting.

## 6.1 Related Literature

### Single product models

Consider a supplier selling to a single retailer. Each of the parties base their decisions on their own profit margins rather than the total margin. This is called “double marginalization” and leads to less than optimal sales for the total chain. Jeuland and Shugan (1983) study a ‘simple manufacturer-retailer channel’ facing a deterministic demand function. They show that uncoordinated and independent channel members’ decisions result in a higher price paid by consumers. They introduce quantity discount schedule as a coordination mechanism and derive the form of the quantity discount schedule that results in optimal channel profits.

The operations equivalent of this model is a single supplier selling to a single retailer facing uncertain demand (a newsvendor). Lariviere and Porteus(2001) study this model when there is no coordinating mechanism. Given a wholesale price  $w$ , selling price  $p$  and random demand  $D$ , the retailer orders the newsvendor quantity,  $q(w)$  that satisfies  $p \Pr\{q < D\} = w$ . The supplier’s profit  $\pi^s = (w - c)q(w)$  where  $c$  is the production/procurement cost per unit. In previous studies,  $\pi^s$  is shown to be unimodal in  $w$  when  $D$  follows a uniform distribution, but not for other distributions. Lariviere and Porteus (2001) present a different

approach, they maximize  $(w(q) - c)q$  over  $q$  rather than maximizing  $(w - c)q(w)$ . These two optimization problems are equivalent since there is a one-to-one mapping between  $w$  and  $q$  through the retailer's optimality equation. In this case, the suppliers profit function can be shown to be unimodal in  $q$  for demand distributions with Increasing Generalized Failure Rate (IGFR). The generalized failure rate is  $xf(x)/(1 - F(x))$  where  $f$  and  $F$  are the density and cumulative density functions, respectively. IGFR condition is less restrictive than the increasing hazard rate condition, and it is satisfied by many distributions such as uniform, normal, and gamma. Unimodality guarantees the optimality of the solution to the first order conditions.

Cachon (2002) reviews the literature on coordinating the newsvendor with contracts. He describes five coordinating mechanisms: buy-back contracts, revenue sharing contracts, quantity flexibility contracts, sales rebate contracts, and quantity discount schedules. In our paper, suppliers implement revenue-sharing contracts. Revenue sharing contracts are shown to be equivalent to buy back contracts. Cachon and Lariviere(2000) provide an analysis of these contracts in more detail. Revenue sharing contracts are also effectively similar to quantity discount contracts: the retailer pays a fixed price for all units purchased and gives a share of each unit sold back to the supplier. The supplier's expected revenue share from the last unit sold decreases with the order quantity. Hence effective price paid to the supplier (wholesale price plus the revenue share) is decreasing in the order quantity.

In our paper, the suppliers compete with each other based on the contract terms. A supplier can get the retailer order more of his product with a lower wholesale price in the price-only contract or a higher retailer share in the revenue sharing contract. The double

marginalization effect is eliminated when all suppliers offer the same share to the retailer. To the best of our knowledge, ours is the first paper where supply chain coordinating contracts are employed in a multiple competing contracts setting.

There is a large number of papers on design of optimal quantity discount schedules where order sizes impact the inventory carrying and setup costs at the retailer and the supplier (Monahan, 1984; Lee and Rosenblatt, 1986). Weng (1995) combines the quantity discount models with the Jeuland and Shugan (1983) model and shows that quantity discount schedules alone are not sufficient for coordination with price dependent demand. In our model, the operational costs (due to the inventory risks) are accrued only at the retail level.

### **Revenue sharing contracts**

Revenue sharing contracts are able to coordinate one-to-one supply chains when the retailer makes a quantity decision facing demand uncertainty or a pricing decision facing a downward sloping demand curve or both the quantity and pricing decisions. They also coordinate multiple competing retailers buying from a single supplier. In all these settings an arbitrary allocation of profits across the supplier and the retailer(s) is possible. A contract so powerful must be ubiquitously used in practice. That is not the case however. For revenue sharing to work two conditions must be met. One, marginal cost of production must be relatively low. Two, the administrative burden for the supplier to monitor the revenues of the retailer(s) must be small enough. Video rental industry clearly meets these two conditions. Copying a tape is cheap (the production costs are already sunk) and all video rental companies use computerized point-of-sales systems that can be linked to the suppliers. Cachon and Lariviere (2001) describe the transformation of the video

rental industry by adopting revenue sharing mechanisms. The authors also suggest digital music industry and advanced medical equipment (e.g., laser surgery to correct myopia) as other industries for which revenue sharing would work well. McDowell and Lee (2003) describe Indian government's revenue sharing agreements with suppliers in their bid to increase the coverage and quality of the wireless services across the country. There is no reported implementation of revenue sharing contracts in other retail industries. But quantity discount schemes are common in all forms of retailing and recall that quantity discount schemes are theoretically equivalent to revenue sharing contracts. Recall that in our model, multiple suppliers sell through a common retailer. Video rental industry provides an example of our model as well, since almost all studios working with Blockbuster Video adopted revenue sharing mechanisms. It is not reported however whether studios use the revenue shares as a competitive tool to increase their share of the business at the retailer.

### **Multi-product models**

Staelin and McGuire (1983) study two channels selling two competing products. They compare different vertical integration levels and conclude that suppliers are better off not owning the retailer when products are highly substitutable, because they avoid direct competition. Choi (1999) studies price competition between two suppliers selling their products through a common retailer. To the best of our knowledge, this is the only paper that studies the competition between two firms through a retailer. The retailer decides on the selling prices of the products in a differentiated Bertrand competition. The retailer(s) in our model decide on the stocking quantities of the products facing stochastic demand. Choi (1999)

compares the equilibria under different leadership scenarios. We compare the equilibria of different horizontal and vertical integration levels under different contract types.

Inventory models with competing products are studied extensively. Lippman and McCardle(1997), Netessine and Rudi (2003), Mahajan and van Ryzin (2001b) study single period inventory models competing based on stocking quantities and show that decentralized retailers stock more than a centralized retailer. These papers study the problem at the retail level. Netessine and Zhang (2003) extends this research by studying products with positive externalities (i.e., complementary products) and show that the results are reversed.

The study of multi-product inventory models at the retail level is extended to the analysis of distribution systems where a single supplier sells her products to retailers under inventory or price competition. Examples include Bernstein and Federgruen (2001, 2003), Netessine and Zhang (2003), Mahajan and van Ryzin (1999), Anupindi and Bassok (1999). The main difference of our model is that the products are supplied by different suppliers, not by the same supplier. Chen (2001) studies the procurement of a single retailer from multiple suppliers with an auction theoretic perspective. Bernstein and Vericourt (2002) present a model where each retailer chooses her supplier based on backorder penalties offered by the suppliers.

## 6.2 Model Basics

The retailer's decision  $q_i$  are the stocking quantities of products  $i = 1..n$ . The retailer faces the stochastic demand vector  $D = (D_1, .., D_n)$ . Either due to consumer driven substitution

or because of externalities the effective demand to each product is linked to the stocking quantities of the other products. We specify three examples to illustrate this dependency.

(A1) Consumer driven substitution: Each product faces stochastic demand  $d_j$ , the unmet demand  $(d_j - q_j)^+$  flows to other products according to a substitution probability matrix  $(\alpha_{ji})$ . Hence the effective demand to product  $i$  is  $D_i = d_i + \sum_j \alpha_{ji}(d_j - q_j)^+$ , where  $(d_1, \dots, d_n)$  is drawn from a known multivariate distribution. This probabilistic substitution model is very common in multi-item inventory and assortment planning literatures.

(A2) Additive model:  $D_i = d_i - g_i(q_{-i})$ , where  $(d_1, \dots, d_n)$  is drawn from a known multivariate distribution and  $g_i$  is a nondecreasing function that is jointly concave in  $q_j$  for all  $j \neq i$ .

(A3) Multiplicative model:  $D_i = d_i g_i(q_{-i})$ , where  $(d_1, \dots, d_n)$  is drawn from a known multivariate distribution and  $g_i$  is a nondecreasing function that is jointly concave in  $q_j$  for all  $j \neq i$ .

Let  $\mathbf{q} = (q_1, \dots, q_n)$ . The profit function of the retailer from product  $i$  is

$$\pi_i(\mathbf{q}) = p_i E[\min\{q_i, D_i\}] - w_i q_i$$

All three demand models generate revenues with the following properties, which we think are essential to our understanding of the problem. Revenues from product  $i$  is increasing in  $q_i$ , decreasing in  $q_j$ . The marginal benefit of increasing  $q_i$  is decreasing in  $q_j$  for all  $j \neq i$ , i.e., products are strategic substitutes.

Demand model (A1) does not generate a revenue function that is jointly concave in  $\mathbf{q}$  (see Netessine and Rudi (2003) for a counter example). Models (A2) and (A3) satisfy joint

concavity by their definition of  $g_i$ . Hence under (A2) and (A3), a centralized retailer's first order conditions have a unique solution (vector of quantities) that is the global optimal. In the decentralized retailers case however, the game between retailers have a unique equilibrium under all demand models, (A1), (A2) and (A3) (See Netessine and Rudi, 2003 and Netessine and Zhang, 2003). The analysis in this chapter is built on demand models (A2), but can accommodate demand models (A1) and (A3) with slight modifications. Later in this chapter we focus on the two-product case with the specific demand model  $D_i = d_i - bq_j$ ,  $b \in (0, 1)$ .

Suppliers buy or manufacture these products at a cost of  $c_i$  per unit, and sell it to the retailers according to a contractual agreement. In this paper, we consider price-only and revenue sharing contracts. Following is an outline of the alternative ownership (centralization/decentralization) configurations. We analyze each of these models under both contract types. NR and NZ stand for Netessine and Rudi (2003) and Netessine and Zhang (2003), respectively.

|    | Setup for $n$ products     | Analysis                           |
|----|----------------------------|------------------------------------|
| C  | Centralized $n$ products   | See NR and references therein      |
| D  | Decentralized $n$ products | See NR and references therein      |
| CD | Centralized suppliers,     | Wholesale price contracts (see NZ) |
|    | Decentralized retailers    | Revenue sharing contracts (NEW)    |
| DD | Decentralized suppliers,   | Wholesale price contracts (NEW)    |
|    | Decentralized retailers    | Revenue sharing contracts (NEW)    |
| CC | Centralized suppliers,     | Wholesale price contracts (see NZ) |
|    | Centralized retailers      | Revenue sharing contracts (NEW)    |
| DC | Decentralized suppliers,   | Wholesale price contracts (NEW)    |
|    | Centralized retailers      | Revenue sharing contracts (NEW)    |

### 6.3 Model C: All suppliers and retailers centralized

The joint owner solves the optimization problem

$$\max_{\mathbf{q}} \sum_i p_i E[\min\{q_i, D_i\}] - c_i q_i$$

$\mathbf{q}^C$  is the unique solution to

$$p_i \Pr\{q_i \leq D_i\} - \sum_{j \neq i} p_j \Pr\{D_j \leq q_j\} b - c_i = 0 \text{ for all } i \quad (6.1)$$

### 6.4 Model D: Competing supplier-retailer pairs

Supplier  $i$  and retailer  $i$  are jointly owned. Similarly supplier  $j$  and retailer  $j$  are jointly owned. This is equivalent to competing newsvendors. Each supplier-retailer pair maximizes

their own profits given others' stocking decisions.

$$\max_{q_i} p_i E[\min\{q_i, D_i\}] - c_i q_i \text{ for all } i$$

The unique equilibrium solution  $\mathbf{q}^D$  is the solution to

$$p_i \Pr\{q_i \leq D_i\} - c_i = 0 \text{ for all } i \quad (6.2)$$

Unlike the centralized solution, the equilibrium solution  $\mathbf{q}^D$  is unique for all demand models (See NR and NZ). It is easy to show that  $\mathbf{q}^D$  is larger than  $\mathbf{q}^C$  in at least component.

## 6.5 Model CD: Centralized suppliers, competing retailers

### 6.5.1 Wholesale price contracts

This configuration is similar to distribution systems, i.e., single supplier and competing retailers. However, the difference is that supplier sells different products to each retailer. Given wholesale prices  $\mathbf{w} = (w_1, \dots, w_n)$ , the retailers maximize their own profits. That is

$$\max_{q_i} p_i E[\min\{q_i, D_i\}] - w_i q_i \text{ for all } i$$

The unique equilibrium solution  $\mathbf{q}^{CD}(\mathbf{w})$  is the solution to

$$p_i \Pr\{q_i \leq D_i\} - w_i = 0 \text{ for all } i \quad (6.3)$$

The suppliers maximize their profits given the retailers' equilibrium order quantities characterized by (6.3).

$$\max_{\mathbf{w}} \sum_i (w_i - c_i) q_i^{CD}(\mathbf{w}) \text{ for all } i \quad (6.4)$$

Netessine and Zhang (2003) (hereafter NZ) show that for two symmetric products and symmetric wholesale prices, supplier's problem has a unique solution for IGFR distributions and linear demand function  $D_i = d_i - bq_j$ . It is possible to generalize their result to  $n$  products, and also prove it for a multiplicative demand model in addition to the linear demand model. With symmetric solution for symmetric products, the optimality conditions are equivalent to the one in single supplier-single retailer contracts studied by Lariviere and Porteus (2003). In a symmetric solution, the retailers' equilibrium solution becomes  $p \Pr\{q \leq D\} - w = 0$ . There is a one-to-one mapping between wholesale price  $w$  and stocking quantity  $q$ . The supplier's optimization problem is

$$\max_w \sum_i (w - c) q^{CD}(\mathbf{w}) = \max_w n(w - c) q^{CD}(\mathbf{w}) = \max_q n(w(q) - c) q.$$

Substituting the retailers' equilibrium condition into this optimization problem, we get

$$\max_w n(p \Pr\{q \leq D\} - c) q$$

The first order optimality condition for additive demand model (i.e.,  $D_i = d_i - b \sum_{j \neq i} q_j$ ) is

$$pP + pqP' - c = 0$$

$$pP(1 + qP'/P) - c = 0$$

$$p \Pr\{q \leq D\} \left( 1 - \frac{qf(D)}{\Pr\{q \leq D\}} (1 + nb) \right) = 0$$

The expression in the parenthesis is decreasing in  $q$  for any  $b \geq 0$  if  $d_i$  follows an IGFR distribution with density function  $f$ . Hence the whole expression is decreasing at points where the first order condition holds. Hence, the supplier's profit function is unimodal in the  $q$ , or equivalently  $w$ . For a multiplicative model (i.e.,  $D_i = d \prod_{j \neq i} q_j^{-b}$ , the last term becomes  $(1 + nb)q^{nb}$ .

Recall that in the symmetric problem there is a one-to-one mapping between  $w$  and  $q$ . Clearly, there exists a wholesale price that induces the retailers to order the system optimal quantities. NZ derives that wholesale price  $w$  such that  $\mathbf{q}^{CD}(\mathbf{w}) = \mathbf{q}^C$ . This is an interesting result as the double marginalization effect and the retail level competition cancel each other and the retailers choose the supply chain optimal stocking quantities. However, this doesn't mean that the supplier will choose that wholesale price. How  $\mathbf{q}^{CD}(\mathbf{w}^{CD})$  compares with  $\mathbf{q}^C$  depends on which one of the double marginalization and competition effects dominates.

### 6.5.2 Revenue sharing contracts

The supplier offers a revenue sharing contract, where the retailer pays  $(1 - \phi_i)p_i$  for each unit sold. Similar to the revenue sharing contracts in one-to-one settings, the supplier can get the retailer to order higher than they would with a wholesale price contract, by setting  $w_i = \phi_i c_i$ . To see this consider the game between the retailers

$$\max_{q_i} \phi_i p_i E[\min\{q_i, D_i\}] - \phi_i c_i q_i \text{ for all } i$$

The equilibrium conditions are

$$\phi_i(p_i \Pr\{q_i \leq D_i\} - c_i) = 0 \text{ for all } i \quad (6.5)$$

which is equivalent to (6.2).

**Proposition 6.5.1** *In model CD with revenue sharing contracts  $(\phi_i, \phi_i c_i)$ , retailers' stocking quantities are  $\mathbf{q}^{CD} = \mathbf{q}^D \geq \mathbf{q}^C$  and the manufacturer chooses  $\phi_i$  to allocate profits arbitrarily.*

**Proof.** The equilibrium conditions (6.5) for  $\mathbf{q}^{CD}$  are equivalent to that of  $\mathbf{q}^D$  given in (6.2) for any value of  $\phi$ .  $\mathbf{q}^D < \mathbf{q}^C$  is due to previous studies and can be derived by examining (6.1) and (6.2). ■

An alternative would be to devise a contract with transfer payments where the supplier effectively shares the total revenue from all retailers with retailers. This will lead the retailers to carry the system optimal inventory. However it is not realistic to expect the retailers to accept such an arrangement where other retailers get a share of their sales.

Clearly, there exists a revenue sharing contract that produces the system optimal quantities by setting  $w_i \geq \phi_i c_i$  and reducing the incentives of retailers to carry inventory. For example, for  $\phi_i = \phi_j = 1$ , the system is equivalent to the case of wholesale prices. NZ derives the coordinating wholesale price, denoted  $w^C$ .

**Proposition 6.5.2** *In model CD with revenue sharing contracts, there exists a set of contracts  $(\phi_i, \phi_i w_i^*)_{i=1..n}$  that will lead to retailers stocking  $\mathbf{q}^{CD} = \mathbf{q}^C$ .*

**Proof.** Consider arbitrary vector  $\phi$ , the supplier offers each retailer  $(\phi_i, \phi_i w_i^C)$ . The retailers compete through profit functions

$$\max_{q_i} \phi_i p_i E[\min\{q_i, D_i\}] - \phi_i w_i^C q_i \text{ for all } i$$

Since  $\phi_i$  does not effect the outcome of the game, this game is equivalent to

$$\max_{q_i} p_i E[\min\{q_i, D_i\}] - w_i^C q_i \text{ for all } i$$

This is exactly the retailers' game under wholesale price contracts (see equation 6.3) with  $w^C$ . NZ shows that the solution to the above set of equations is  $\mathbf{q}^C$ . ■

In this case, Arbitrary allocation of profits is not possible. The supplier makes positive profits for any  $\phi_i$ . If she sets  $\phi_i$  arbitrarily close to zero, the retailers will order  $\mathbf{q}^C$  and the supplier's profit

$$p_i E[\min\{q_i, D_i\}] - c_i q_i - \phi_i (p_i E[\min\{q_i, D_i\}] - w_i^C q_i)$$

which is equal to the system optimal profit for  $\phi_i = 0$ . At the other extreme, For  $\phi_i = 1$  however, the retailers get only a fraction of the system optimal profits, the supplier's profit is

$$w_i^C q_i^C - c_i q_i^C > 0$$

## 6.6 Model DD: Competing suppliers and competing retailers

### 6.6.1 Wholesale price contracts

Given wholesale prices  $\mathbf{w} = (w_1, \dots, w_n)$ , the retailers maximize their own profits. Retailer's quantity selection game is the same as Model *CD*. The unique equilibrium solution  $\mathbf{q}^{DD}(\mathbf{w})$  is the solution to equilibrium conditions (6.3). The suppliers maximize their profits given the retailers' ordering function  $\mathbf{q}^{DD}(\mathbf{w})$ .

$$\max_{w_i} (w_i - c_i) q_i^{DD}(\mathbf{w}) \text{ for all } i$$

**Proposition 6.6.1** *Consider the case of two products with IGFR demand distributions.*

*There exists an equilibrium  $\mathbf{w}^{DD}$  in the suppliers' pricing game characterized by*

$$(w_i - c_i) \frac{\partial q_i^{DD}(\mathbf{w})}{\partial w_i} + q_i^{DD}(\mathbf{w}) = 0 \text{ for } i = 1, 2 \quad (6.6)$$

**Proof.** To show existence of an equilibrium (i.e., solution to (6.6)), we need to show that each supplier's profit is unimodal in its own wholesale price. We start with evaluating  $\partial q_i(\mathbf{w})/\partial w_i$ . By the Implicit Function Theorem,

$$\begin{vmatrix} \frac{\partial q_i}{\partial w_i} \\ \frac{\partial q_j}{\partial w_i} \end{vmatrix} = - \begin{vmatrix} -p_i f_{D_i}(q_i) & -p_i f_{D_i}(q_i)b \\ -p_j f_{D_j}(q_j)b & -p_j f_{D_j}(q_j) \end{vmatrix}^{-1} \begin{vmatrix} \frac{\partial R_i}{\partial w_i} \\ \frac{\partial R_j}{\partial w_i} \end{vmatrix}$$

Let  $\gamma = p_i f_{D_i}(q_i)$  and  $\delta = p_j f_{D_j}(q_j)$ .

We have

$$\begin{vmatrix} \frac{\partial q_i}{\partial w_i} \\ \frac{\partial q_j}{\partial w_i} \end{vmatrix} = -\frac{1}{\gamma\delta(1-b^2)} \begin{vmatrix} -\delta & b\gamma \\ b\delta & -\gamma \end{vmatrix} \begin{vmatrix} -1 \\ 0 \end{vmatrix} = \frac{1}{\gamma(1-b^2)} \begin{vmatrix} -1 \\ b \end{vmatrix}$$

From (6.3) we have

$$p_i = w_i / \Pr\{q_i < D_i\}$$

Substitute  $p_i$  and  $\partial q_i/\partial w_i$  into (6.6). We have

$$\left( -(1-b^2) \frac{(w_i - c)}{w_i} \frac{\Pr\{q_i(w_i) < D_i\}}{f(q_i(w_i))q_i(w_i)} + 1 \right) q_i(w_i) = 0 \quad (6.7)$$

Clearly  $q_i(w_i)$  and  $(w_i - c)/w_i$  are decreasing in  $w_i$ . Since demand is IGFR,

$$\frac{\Pr\{q_i(w_i) < D_i\}}{f(q_i(w_i))q_i(w_i)}$$

is decreasing in  $q_i(w_i)$  and  $q_i(w_i)$  is decreasing in  $w_i$ . Hence the negative term in (6.7) is decreasing in  $w_i$ . We also have  $q_i(w) \geq 0$ . Hence the second derivative is zero at the points where the first order condition holds. This guarantees the unimodality of each supplier's profit in her decision. Note also that the action space  $w_i \in [c_i, p_i]$  is compact. Therefore there exists an equilibrium (Theorem 1.2. in Fudenberg and Tirole, 2000). ■

**Proposition 6.6.2** *Consider the case of two symmetric products with IGFR demand distribution .*

*i) There exists a unique equilibrium  $\mathbf{w}^{DD}$  that is symmetric (i.e.,  $w_1^{DD} = w_2^{DD}$ ) and characterized by*

$$(w_i - c) \frac{\partial q_i^{DD}(\mathbf{w})}{\partial w_i} + q_i^{DD}(\mathbf{w}) = 0, \quad i = 1, 2$$

*ii)  $\mathbf{w}^{DD} \leq \mathbf{w}^{CD}$ .*

*iii)  $\mathbf{q}^D \geq \mathbf{q}^{DD} \geq \mathbf{q}^{CD}$ .*

**Proof.** *i)* We first establish that the reaction functions are increasing.

$$\frac{\partial w_i(w_{-i})}{\partial w_j} = - \frac{(w_i - c_i) \frac{\partial^2 q_i^{DD}(\mathbf{w})}{\partial w_j \partial w_i} + \frac{\partial q_i^{DD}(\mathbf{w})}{\partial w_j}}{(w_i - c_i) \frac{\partial^2 q_i^{DD}(\mathbf{w})}{\partial w_i^2} + 2 \frac{\partial q_i^{DD}(\mathbf{w})}{\partial w_i}}$$

The denominator is negative for IGFR demand distributions as established in Proposition (6.6.1). We proceed to evaluate

$$\begin{aligned} \frac{\partial^2 q_i^{DD}(\mathbf{w})}{\partial w_j \partial w_i} &= \frac{\partial}{\partial w_j} \frac{-1}{\gamma(1-b^2)} = -(1-b^2)^{-1} \left( \frac{\partial \gamma^{-1}}{\partial q_i} \frac{\partial q_i}{\partial w_j} + \frac{\partial \gamma^{-1}}{\partial q_j} \frac{\partial q_j}{\partial w_i} \right) \\ &= -(1-b^2)^{-2} \delta^{-1} \left( b \frac{\partial \gamma^{-1}}{\partial q_i} - \frac{\partial \gamma^{-1}}{\partial q_j} \right) \\ &= (1-b^2)^{-2} \delta^{-1} \gamma^{-2} (b\gamma' - b\gamma') = 0. \end{aligned}$$

Hence  $\frac{\partial w_i(w_{-i})}{\partial w_j} \geq 0$ . In a symmetric game with increasing reaction functions, all equilibria are symmetric, i.e., there does not exist asymmetric equilibria.

Uniqueness comes from the observation that at each symmetric equilibrium, the slope of the reaction function  $w_i(w_{-i})$  defined by (6.6) is less than 1:  $\frac{\partial w_i(w_{-i})}{\partial w_j} < 1$  if  $\frac{\partial q_i^{DD}(\mathbf{w})}{\partial w_j} < -2 \frac{\partial q_i^{DD}(\mathbf{w})}{\partial w_i}$ , which means  $b\delta^{-1} < 2\gamma^{-1}$ . At a symmetric equilibrium,  $w_i = w_j$ , which implies  $q_i = q_j$ , and  $\gamma = \delta$ . By definition,  $b < 1$ . Therefore at each symmetric equilibrium, the

slope of the reaction function is less than 1, which implies that there can be at most one symmetric equilibrium, as the reaction function cannot cross the  $w_i = w_j$  line more than once.

ii) Take the optimality conditions for the supplier's optimization problem (6.4) at  $\mathbf{w}^{CD}, \mathbf{q}^{CD}$ .

$$(w - c) \frac{\partial q_i^{CD}(\mathbf{w})}{\partial w_i} + q_i^{CD}(\mathbf{w}) + \sum_{j \neq i} (w_j - c_j) \frac{\partial q_j^{CD}(\mathbf{w})}{\partial w_i} = 0 \text{ for all } i$$

We can show that  $\frac{\partial q_i^{CD}(\mathbf{w})}{\partial w_i} \geq 0$  by using the Implicit Function Theorem on the equilibrium conditions (6.3) for the retailer's stocking quantity game. We have

$$(w - c) \frac{\partial q_i^{CD}(\mathbf{w})}{\partial w_i} + q_i^{CD}(\mathbf{w}) \leq 0 \text{ for all } i.$$

The left-hand side of the inequality is the equilibrium conditions (6.6) for  $\mathbf{w}^{DD}$ . The first derivative of the supplier's profit is positive at  $w_i = c^+$ , because the supplier's profit is increasing from zero at  $w_i = c$  to a positive profit. Therefore  $\exists w_i < w^{CD}$  such that  $(w - c) \frac{\partial q_i(\mathbf{w})}{\partial w_i} + q_i(\mathbf{w}) = 0$ . This means that the reaction function in the game  $DD$  satisfy  $w_i^{DD}(w_{-i}) \leq w_i^{CD}(w_{-i})$ . If there exists a symmetric equilibrium  $\mathbf{w}^{DD}$ , the reaction function of the game  $DD$  must cross the diagonal of the unit cube before  $\mathbf{w}^{CD}$ , i.e.,  $\mathbf{w}^{DD} \leq \mathbf{w}^{CD}$ . This implies  $\mathbf{q}^{DD} \geq \mathbf{q}^{CD}$ .

iii) Compare  $D$  and  $DD$ . The retailer's optimality equations are the same except that  $DD$  has  $w_i$  instead of  $c_i$ . Since the equilibrium wholesale prices (of the game between suppliers in  $DD$ ) cannot be lower than  $c_i$ ,  $\mathbf{q}^D \geq \mathbf{q}^{DD}$ . ■

In this model, there are three effects: double marginalization, competition at the retail level, and (implicit) competition at the supplier level. Suppliers compete with each other

through the competing retailers and they set lower prices than they would in the centralized suppliers case, resulting in more inventory in the system.

### 6.6.2 Revenue sharing contracts

The suppliers offer a revenue sharing contract, where the retailer shares  $(1 - \phi_i)p_i$  for each unit sold with its supplier. Similar to the revenue sharing contracts in one-to-one settings, the supplier can get the retailer to order higher than they would with a wholesale price contract, by setting  $w_i = \phi_i c_i$ . To see this consider the game between the retailers

$$\max_{q_i} \phi_i p_i E[\min\{q_i, D_i\}] - \phi_i c_i q_i \text{ for all } i$$

The equilibrium conditions are

$$\phi_i(p_i \Pr\{q_i \leq D_i\} - c_i) = 0 \text{ for all } i \quad (6.8)$$

This implies that for any  $\phi = (\phi_1, \dots, \phi_n)$ , the retailers stocking quantities are  $\mathbf{q}^{DD} = \mathbf{q}^D \geq \mathbf{q}^C$ . The revenue sharing contracts eliminate the effect of double marginalization. In essence, it transforms the system to competing vertically integrated chains, which clearly carries more inventory than system optimal.

Also note that we are confined to revenue sharing contracts of the type  $(\phi_i, \phi_i c_i)$  rather than  $(\phi_i, w_i)$ . Analyzing the game between suppliers offering revenue sharing contracts with two parameters is left as future research.

## 6.7 Model CC: Centralized suppliers, centralized retailers

### 6.7.1 Wholesale price contracts

Given wholesale prices  $\mathbf{w} = (w_1, \dots, w_n)$ , the retailer maximize its total profit. That is

$$\max_{\mathbf{q}} \sum_i \{p_i E[\min\{q_i, D_i\}] - w_i q_i\} \text{ for all } i$$

The retailer's optimal stocking quantities  $\mathbf{q}^{CC}(\mathbf{w})$  is the unique solution to

$$p_i \Pr\{q_i \leq D_i\} - \sum_{j \neq i} p_j \Pr\{D_j \leq q_j\} b - w_i = 0 \text{ for all } i \quad (6.9)$$

The supplier maximize its profits given the retailer's optimal solution by (6.9).

$$\max_{\mathbf{w}} \sum_i (w_i - c_i) q_i^{CC}(\mathbf{w}) \text{ for all } i$$

**Proposition 6.7.1** *Consider the case with symmetric products with symmetric wholesale prices. There exists a unique solution  $w^{CC}$  to the supplier's optimization problem.  $\mathbf{q}^{CC}(\mathbf{w}) \leq \mathbf{q}^C$ .*

**Proof.** The uniqueness of  $w$  is similar to our generalization of NZ's result for model CD.

A comparison of (6.9) and (6.1) for any  $w > c$  shows that (6.9) yields stocking quantities  $\mathbf{q}^{CC}(\mathbf{w}) \leq \mathbf{q}^C$ . ■

The double marginalization effect is the only one at work in this case. Hence the system carries less inventory than system optimal.

### 6.7.2 Revenue sharing contracts

The supplier offers the retailer a revenue sharing contract with  $(\phi_1, \dots, \phi_n)$ , and  $w_i = \phi_i c_i$ .

The retailer's optimization problem is

$$\max_{\mathbf{q}} \sum_i \{\phi_i p_i E[\min\{q_i, D_i\}] - \phi_i c_i q_i\} \text{ for all } i$$

The optimal stocking quantities  $\mathbf{q}^{CC}(\phi)$  is the unique solution to

$$\phi_i p_i \Pr\{q_i \leq D_i\} - \sum_{j \neq i} \phi_j p_j \Pr\{D_j \leq q_j\} b - \phi_i c_i = 0 \text{ for all } i$$

**Proposition 6.7.2** *With  $\phi_i = \phi$  for all  $i$ , arbitrary  $\phi \in [0, 1]$ ,  $\mathbf{q}^{CC}(\phi) = \mathbf{q}^C$ .*

In the CD model the supplier has the flexibility to offer different  $\phi_i$ 's to the retailers. In this case however with different  $\phi_i$ 's to the retailer will deviate from  $\mathbf{q}^C$ . Note that the supplier has no incentive to choose nonidentical  $\phi_i$ , because with identical  $\phi_i$ , the system optimal is reached and supplier can keep all the profits.

## 6.8 Model DC: Supplier competition through centralized retailer

### 6.8.1 Wholesale price contracts

The retailer's optimization problem is the same as Model CC. The optimal stocking quantity for the retailer satisfy the optimality conditions (6.9). Based on the retailer's optimality equations, each supplier maximizes its own profits given the other's wholesale price.

$$\max_{w_i} \pi_i^s = (w_i - c_i) q_i^{DC}(\mathbf{w}) \text{ for all } i$$

Similar to the game between suppliers in Model *DD*, to guarantee the existence of an equilibrium, it is sufficient to show that the second order condition holds. Let's evaluate  $\partial q_i(\mathbf{w})/\partial w_i$  and  $\partial^2 q_i(\mathbf{w})/\partial w_i^2$ . Consider the case with two products. By the Implicit Function Theorem,

$$\begin{vmatrix} \frac{\partial q_i}{\partial w_i} \\ \frac{\partial q_j}{\partial w_i} \end{vmatrix} = - \begin{vmatrix} -\gamma - \delta b^2 & -\gamma b - \delta b \\ -\gamma b - \delta b & -\gamma b^2 - \delta \end{vmatrix}^{-1} \begin{vmatrix} \frac{\partial R_i}{\partial w_i} \\ \frac{\partial R_j}{\partial w_i} \end{vmatrix}$$

where  $\gamma = p_i f_{D_i}(q_i)$  and  $\delta = p_j f_{D_j}(q_j)$ .

We have

$$\begin{aligned} \begin{vmatrix} \frac{\partial q_i}{\partial w_i} \\ \frac{\partial q_j}{\partial w_i} \end{vmatrix} &= -\frac{1}{\gamma\delta(1-b^2)^2} \begin{vmatrix} -\gamma b^2 - \delta & (\gamma + \delta)b \\ (\gamma + \delta)b & -\gamma - \delta b^2 \end{vmatrix} \begin{vmatrix} -1 \\ 0 \end{vmatrix} \\ &= \frac{1}{\gamma\delta(1-b^2)^2} \begin{vmatrix} -\gamma b^2 - \delta \\ (\gamma + \delta)b \end{vmatrix} \end{aligned}$$

In Appendix C, we show that  $\frac{\partial^2 q_i}{\partial w_i^2} = 0$  for uniform demand distribution and that makes the supplier  $i$ 's profit concave in  $w_i$ . This ensures that each supplier's reaction function is a continuous function of the other supplier's wholesale price, which guarantees the existence of an equilibrium.

**Proposition 6.8.1** *In model CD, there exists an equilibrium in the suppliers' pricing game with two products and uniform demand distribution.*

**Proposition 6.8.2** *Consider model CD with two symmetric products with uniform demand distributions.*

i) There exists a unique equilibrium  $\mathbf{w}^{DC}$  that is symmetric and characterized by

$$(w_i - c) \frac{\partial q_i^{DC}(\mathbf{w})}{\partial w_i} + q_i^{DC}(\mathbf{w}) = 0 \text{ for all } i \text{ where } w_i = w^{DC} \text{ for all } i$$

ii)  $\mathbf{w}^{DC} \leq \mathbf{w}^{CC}$

iii)  $\mathbf{q}^C \geq \mathbf{q}^{DC} \geq \mathbf{q}^{CC}$ .

**Proof.** The proof is similar to that of Proposition 6.6.2., with the difference that we can establish that the reaction functions have positive slope only for uniform demand distribution. Specifically, In Appendix C, we show that  $\frac{\partial^2 q_i}{\partial w_j \partial w_i} = 0$  for uniform demand distribution and that  $\frac{\partial w_i(w_{-i})}{\partial w_j} < 1$  for symmetric solutions. ■

$\mathbf{q}^C \geq \mathbf{q}^{DC}$  is an interesting result. Competition between the suppliers is not enough to offset the double marginalization effect.

### 6.8.2 Revenue sharing contracts

With revenue sharing contracts  $(\phi_i, \phi_i c_i)$  offered by the suppliers, the retailer's order quantities  $\mathbf{q}^{DC}(\phi)$  is the unique solution to

$$\phi_i p_i \Pr\{q_i \leq D_i\} - \sum_{j \neq i} \phi_j p_j \Pr\{D_j \leq q_j\} b - \phi_i c_i = 0 \text{ for all } i \quad (6.10)$$

Clearly, for  $\phi_i = \phi$  for all  $i$ , the retailer orders the system optimal,  $\mathbf{q}^C$ . For nonidentical  $\phi_i$ , however, the retailer will tend to order more from the suppliers who leaves a higher share of the revenues to the retailer. Each supplier optimizes her profits given the other suppliers' contract parameters and retailer's ordering function  $\mathbf{q}^{DC}(\phi)$ .

$$\max_{\phi_i} (1 - \phi_i)(p_i - c_i) q_i^{DC}(\phi) \text{ for all } i$$

The equilibrium revenue sharing parameters  $\phi^{DC}$  satisfies the following conditions.

$$-q_i^{DC}(\phi) + (1 - \phi_i) \frac{\partial q_i^{DC}(\phi)}{\partial \phi_i} = 0 \text{ for all } i \quad (6.11)$$

Consider the case with two symmetric products  $i$  and  $j$ . From the Implicit Function Theorem,

$$\begin{vmatrix} \frac{\partial q_i}{\partial \phi_i} \\ \frac{\partial q_j}{\partial \phi_j} \end{vmatrix} = - \begin{vmatrix} -\phi_i p_i f_{D_i}(q_i) - \phi_j p_j f_{D_j}(q_j) b^2 & -\phi_i p_i f_{D_i}(q_i) b - \phi_j p_j f_{D_j}(q_j) b \\ -\phi_i p_i f_{D_i}(q_i) b - \phi_j p_j f_{D_j}(q_j) b & -\phi_i p_i f_{D_i}(q_i) b^2 - \phi_j p_j f_{D_j}(q_j) \end{vmatrix}^{-1} \begin{vmatrix} \frac{\partial R_i}{\partial \phi_i} \\ \frac{\partial R_j}{\partial \phi_j} \end{vmatrix}$$

$\partial R_i / \partial \phi_i = p_i \Pr\{D_i \geq q_i\} - c_i$  and  $\partial R_i / \partial \phi_i = -p_i \Pr\{D_i \leq q_i\}$ .

For uniform demand distribution, we have  $\partial q_i / \partial \phi_i > 0$ ,  $\partial q_j / \partial \phi_i < 0$ , and  $\partial^2 q_i / \partial \phi_i^2 < 0$  if  $(p_i - c_i) > p_i \Pr\{D_i \leq q_i\}(1 - b^2)$  (See Appendix C for proofs of these results). This again ensures the existence of equilibrium for uniform demand distribution.

**Proposition 6.8.3** *In model DC, there exists an equilibrium in the suppliers' revenue sharing game with two products and uniform demand distribution if  $(p_i - c_i) > p_i \Pr\{D_i \leq q_i\}(1 - b^2)$  for all  $i$ .*

**Proposition 6.8.4** *Consider model DC with two symmetric products and uniform demand distribution.*

i) *If  $(p_i - c_i) > p_i \Pr\{D_i \leq q_i\}(1 - b^2)$  for all  $i$ , there exists a symmetric equilibrium that is characterized by (6.11) where  $\phi_i = \phi^{DC}$  for all  $i$*

ii) *If  $b > 1/2$ ,  $\phi^{DC}$  is the unique equilibrium of the game between suppliers..*

iii)  *$\mathbf{q}^{DC}(\phi^{DC}) = \mathbf{q}^C$ .*

**Proof.** *i)* In any symmetric game with continuous reaction functions, there exists a symmetric equilibrium.

*ii)* We find in Appendix C that  $\frac{\partial^2 q_i}{\partial \phi_j \partial \phi_i} \geq 0$  for uniform demand distribution if  $\frac{\phi_i^2}{\phi_j^2} < \frac{(1+b)(p_i-c_i)}{b(p_i-c_i)}$ . That establishes that reaction functions are increasing, i.e.,  $\frac{\partial w_i(w_{-i})}{\partial w_j} > 0$  in a cone around the diagonal. Since the reaction functions are increasing in the cone, there is no asymmetric equilibria. Moreover, we show that  $\frac{\partial w_i(w_{-i})}{\partial w_j} < 1$  if  $b > 1/2$ . Similar to the proof of Proposition (6.6.2), this establishes that under uniform demand distribution there exists a unique symmetric equilibrium.

*iii)*  $\phi_i = \phi$  for all  $i$  implies that the retailer's first order conditions are equivalent to (6.1) the optimality conditions in model  $J$ . ■

Consider the symmetric equilibria. At  $\phi_i = \phi_j = \phi$ ,  $q_i(\phi) = q$  and  $\gamma = \delta$ . At  $\mathbf{q}(\phi)$ ,  $\partial q_i / \partial \phi_i = \frac{1}{\phi} \frac{b}{(1-b)^2} \frac{F_D(q)}{f_D(q)}$  which decreases with  $\phi$  and increases with  $b$  (coefficient of negative externality).

An examination of retailer's optimality conditions (6.9) reveals that the equilibria between the suppliers are always interior solutions under both contract types. That is  $\phi = (0, 0)$  or  $\phi = (1, 1)$  cannot be an equilibrium of the revenue sharing game and  $\mathbf{w} = (c_i, c_j)$  or  $\mathbf{w} = (p_i, p_j)$  cannot be an equilibrium of the wholesale pricing game. At  $\phi = (0, 0)$  for example, each supplier unilaterally has an incentive to increase the share of the retailer, because the retailer will drop the other supplier entirely. Hence  $\phi = (0, 0)$  cannot be an equilibrium of the game between suppliers. Now consider the other extreme,  $\phi = (1, 1)$ , left-hand side of (6.11) is negative, meaning that each supplier would like to decrease the share of the retailer. Hence  $\phi = (1, 1)$  cannot be an equilibrium either.

## 6.9 Numerical Examples

In this section, we present two numerical examples for the multiple suppliers selling through a common retailer model (Model *DC*). Namely, we investigate the impact of the two contract types on suppliers' and retailer's profits.

### **Symmetric independent suppliers selling to a centralized retailer**

The equilibria under wholesale price contracts and revenue sharing contracts for symmetric products (i.e., identical selling price, cost and demand data) are summarized in Table 6.1. At the equilibrium under revenue sharing contracts, the retailer orders the system optimal quantities, and suppliers take 98.3% of the profits. The suppliers are better off in two ways with the revenue sharing contracts. They sell more and their margins are much higher, resulting in 66% increase in their profits. The revenue sharing contracts hit the retailer hard: she makes 89% less than the wholesale price case.

### **Asymmetric independent suppliers selling to a centralized retailer**

The two products have the same price and cost, but the mean demand for product 1 is double the mean demand for product 2. We call supplier 1 the large supplier and supplier 2 the small supplier. The equilibria under wholesale price contracts and revenue sharing contracts (i.e., identical selling price, cost and demand data) are summarized in Table 6.2. In both contract types, the small supplier offers a higher margin (i.e., lower wholesale price or higher share) to the retailer than the large supplier. Similar to the asymmetric case, with revenue sharing contracts, both suppliers' and total profits increase significantly and the retailer makes significantly less.

Suppose that the large supplier is the leader and announces her revenue sharing contract parameters. Set  $\phi_1 = 0.2$ , the small supplier chooses her best response  $\phi_2(\phi_1) = 0.05$ . The retailer orders (231, 107) units in comparison to the (223, 123) units in the Nash equilibrium. Both suppliers are still better off than they are under wholesale price contracts. However in this case, the retailer makes more money than the wholesale price case. Reflecting on the Anheuser-Busch example, the retailer would be likely to accept to switch to the revenue sharing contracts. However, it must be noted the small supplier also benefits significantly from switching to revenue sharing contracts. The industry leader should be cautious in such an initiative if her goal is the elimination of small competitors.

## 6.10 Summary and Discussion of Results

|       | $n = 2, D_i = d_i - bq_j, b \in (0, 1).$                                                                                                                                                                            |                                                                                                                                                                                                                                        |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Model | Wholesale price                                                                                                                                                                                                     | Revenue Sharing                                                                                                                                                                                                                        |
| $CD$  | $\pi^s$ -concave in $w$ for $w_i = w, \forall i$ ,<br>for IGFR distributions.                                                                                                                                       | $q^{CD} = q^D \geq q^C, \forall (\phi_i, \phi_j)$                                                                                                                                                                                      |
| $DD$  | $\pi_i^s$ -concave in $w_i$ for IGFR. dist.<br>existence<br>$w_i(w_j)$ - increasing<br>(slope) $< 1$ if $w_i = w_j$<br>unique, symmetric equilibrium<br>$w^{DD} \leq w^{CD}$<br>$q^D \geq q^{DD} \geq q^{CD}$       | $q^{DD} = q^D, \forall (\phi_i, \phi_j)$                                                                                                                                                                                               |
| $CC$  | $\pi^s$ -concave in $w$ for $w_i = w, \forall i$ ,<br>for IGFR distributions.<br>$q^{CC} \leq q^C$                                                                                                                  | $q^{CC} = q^C, \forall (\phi_i, \phi_j)$ if $\phi_i = \phi_j$                                                                                                                                                                          |
| $DC$  | $\pi_i^s$ -concave in $w_i$ (unif).<br>existence<br>$w_i(w_j)$ - increasing (unif)<br>(slope) $< 1$ if $w_i = w_j$ (unif)<br>unique, symmetric equilibrium<br>$w^{DC} \leq w^{CC}$<br>$q^C \geq q^{DC} \geq q^{CC}$ | $\pi_i^s$ -concave in $w_i$ for unif. dist.<br>existence<br>$\phi_i(\phi_j)$ - increasing (unif.)<br>(slope) $< 1$ if $\phi_i = \phi_j, b \geq 1/2$ (unif.)<br>unique, symmetric equilibrium<br>$q^{DC} = q^C$ at the symmetric equil. |

### **Wholesale price contracts**

The suppliers' optimization problem is well behaved for IGFR distributions when the supplier is centralized and sets symmetric wholesale prices and for uniform distributions when the suppliers are decentralized. The reaction functions are increasing in the other supplier's wholesale price and there exists a symmetric equilibrium when the products are symmetric and have uniform demand distributions.

We find that horizontal integration of either echelon (i.e., centralization at the retailer level or the supplier level) increases the wholesale prices and reduces the total stock level in the system regardless of the other echelon being centralized or decentralized. Vertical integration of supplier-retailer pairs on the other hand increases the total stock in the system. Horizontal integration eliminates the competition effect whereas vertical integration eliminates the double marginalization effect. Hence their effect on total inventory is in different directions.

Introducing competition always leads to higher inventories under wholesale price contracts. Competition leads to higher inventory than system optimal when supplier-retailer pairs are vertically integrated, resulting in less profits. However, it may be desirable to introduce competition and increase the inventory in the system. For example when there is a centralized retailer who buys from suppliers with a wholesale price contract, having decentralized suppliers leads to an inventory level that is closer to (but still smaller than) the system optimal in comparison to having centralized suppliers. The retailer wins in two ways from competition between suppliers: lower wholesale prices and increased sales.

### **Revenue sharing contracts**

Revenue sharing contracts eliminate the double marginalization effect always in the decentralized retailers case and when revenue share of all products are identical in the centralized retailer case. Surprisingly, it also eliminates the competition effect at the supplier level. Competition leads to more inventories when it is at the retail level, but has no effect when it is at the supplier level. Hence, the system optimal is reached with revenue sharing contracts when there is a centralized retailer and more than system optimal inventory is held when the retailers are decentralized.

We show that revenue sharing contracts leads to higher inventory than wholesale price contracts in any of the configurations. We also note that the equilibria between the suppliers are always interior solutions under both contract types. This means that under models DD and DC,  $\phi = (0, 0)$  or  $\phi = (1, 1)$  cannot be an equilibrium of the revenue sharing game and  $\mathbf{w} = (c_i, c_j)$  or  $\mathbf{w} = (p_i, p_j)$  cannot be an equilibrium of the wholesale pricing game.

## 6.11 Future research

This chapter is a preliminary attempt to analyze the issues in competition between suppliers selling to the same retailer. Hence it is not mature enough and far from being a final product. The models presented here may not be the most appropriate model. Our results are not general enough since they require conditions on the demand distributions and apply to symmetric products only. Below is a list of potential issues and alternative models that needs to be explored.

Introducing pricing decisions at the retail level can simplify the analysis. It is known from van Mieghem and Dada (1999) that inventory decisions under uncertainty with post-

poned pricing (i.e., retailer's make the pricing decisions after the uncertainty is revealed) simplifies to a capacity game where there is no uncertainty. Rewriting the retailer's optimality equations for this simplified form may simplify the analysis of the game pricing game between suppliers.

Other forms of competition at the supplier level can be considered. For example the number of suppliers in the market can be explored where retailer's choice have an impact on the level of consolidation of the suppliers. Alternatively, suppliers can have fixed prices fixed but commit to increased advertising or sales effort to help out the retailer. In this case the retailer chooses the allocation of her sales between suppliers based on factors other than prices (i.e., effort level). Suppliers can also compete on delivery leadtimes (lower leadtimes reduce the stockout rate at the retailer), transportation costs (charge smaller fraction of the transportation costs to the retailer), or payment to the retailer for unit shelf space her products use. All these alternative forms of competition requires modeling the costs and benefits of these variables at the retailer and the supplier carefully.

Manufacturers leverage their main product line to make their secondary product lines more attractive to the retailers. Anheuser-Busch, a national beer company, is a good example. In an effort to block out the local brands that compete mostly with its premium beers, Anheuser-Busch provides discounts and other incentives for distributors who don't carry local brands. As a result of this campaign, many Anheuser-Busch distributors are dropping independent brews in favor of smaller Anheuser-Busch brands such as Red Wolf and Black & Tan Porter (Wall Street Journal, 1998). We suspect that under Anheuser-Busch's discount scheme may not be effective: the retailer buys more of the main product

rather than increasing the sales of the secondary product. We can investigate whether it is a good idea for Anheuser-Busch to link the pricing of the main and secondary products or whether there exists a contract that will help Anheuser-Busch achieve its goal.

Going back to the model presented in this paper, throughout the analysis we assumed that the centralized retailer's optimal quantity is an interior solution, i.e., the solution is characterized by the first order conditions. In some numerical examples, we have found that when the large supplier gives a very small share to the retailer (e.g., 0.01), the small retailer can increase the retailers share significantly (e.g. 0.2) and the retailer drops the large supplier entirely. This leads to reaction functions that are discontinuous (with a downward jump) and the existence of equilibrium is in jeopardy. This is one area we can explore in depth.

It is possible that modeling the retailer's operational costs in an EOQ framework rather than using newsvendor model can simplify the analysis and further results and comparisons can be made. This option certainly needs to be investigated. The supplier's operational costs (due to setups or holding inventory) can be also incorporated into this model. Role of vendor managed inventories would be an interesting topic of research when multiple suppliers are selling through a common retailer with the EOQ framework. This has been studied for distribution systems by Chen et. al. (2001) and Bernstein et. al. (2003) among others.

## Chapter 7

# Summary of Contributions

In the first part of the dissertation, we study the joint assortment selection and inventory planning problem, defined as the selection of the products in a category and setting the inventory levels for each product. Limited shelf space creates a tradeoff between breadth of assortment and in-stock rates of products in the assortment. The degree of importance of the tradeoff between breadth of assortment and in-stock rates of products depends on the customers' willingness to accept another product as a substitute if their favorite product is not available. Assessing how likely consumers are to accept one of the available variants as a substitute and incorporating this into assortment planning is therefore essential for retailers.

The contributions of the first part are as follows.

1. We discuss the applicability of existing assortment selection models and propose a methodology that is applicable even for problems with large number of products.

2. We develop an estimation scheme for the substitution rate, which can be used in understanding the roles of categories better and as an input to assortment planning algorithms.
3. We develop an iterative optimization heuristic for a nonlinear, nonseparable, discrete resource allocation problem and assess the performance of the heuristic and establish new structural properties of the optimal assortment based on its solution. .
4. We demonstrate the value of taking consumer behavior into account and the value of an operational optimization model in a grand scheme. We applied our method at a large retail chain in the grocery industry and found the optimal assortment for each store. Comparing results with the existing assortments suggest more than 50% increase in profits.

The second part of the dissertation studies the assortment planning problem with multiple merchandise categories. The categories interact with each other due to limited shelf space and the effects of basket shoppers, consumers who purchase from more than one category in a shopping trip. We present a model where the retailer chooses the number of variants to offer in each category and the consumers of multiple categories make their store choice between the retail store and an exogenous outside alternative to maximize their total shopping utility. We investigate the interaction of the category variety decisions at the retail store under centralized and decentralized management regimes. The common practice of category management is an example of a decentralized regime for controlling assortment because each category manager is charged with maximizing profit of his or her assigned category.

Major contributions of the second part can be summarized as follows.

1. Variety levels of categories can be both strategic complements and strategic substitutes depending on their levels in comparison to competition.
2. Category management never achieves the optimal solution. It leads to a significant profit loss and mostly provides less variety than the optimal.
3. We propose a decentralized heuristic that is simple to implement and recovers most of the loss due to category management.

The third part of the dissertation focuses on supplier competition in a retail supply chain. We extend the retailer's assortment planning problem to include the price competition between suppliers, i.e. each supplier use price or discounts as a competitive tool to effect retailer's assortment planning decisions in her favor. Suppliers offer contracts to the retailer(s), whose terms are the wholesale prices per unit under a price-only contract and supplier's share of revenues under revenue sharing contracts. Given a set of contracts, retailer(s) choose the stocking quantities of the products to meet the single period stochastic demand. The products are strategic substitutes, that is demand of a product is negatively effected by other products' stocking quantities. Under different centralization/decentralization configurations we study two contract types: wholesale price contracts and revenue sharing contract types. Following is a brief summary of the differentiating features of our paper.

1. The fundamental difference of our model is that the products are supplied by different suppliers, not by the same supplier. Suppliers compete with each other through a

centralized retailer's optimization decisions or equilibria of a noncooperative game between retailers.

2. Contracts serve as a tool for competition rather than achieving system optimality.

There isn't much research on why some contracts are more prominent in practice.

Ours is a first attempt in understanding the impact of different contract types on the profits of the supply chain players.

3. We show that introducing competition at the supplier level can bring the total inventory level closer to the system optimal solution under wholesale price contracts and

benefits the retailer in two ways: higher sales and lower wholesale prices. We also show that revenue sharing contracts coordinate the system with symmetric products

when the retailer is centrally managed. Finally, revenue sharing contracts yield higher inventory than the wholesale price contracts.

4. Although at a preliminary stage, our research is the first to employ supply chain

contracts in a multiple competing products setting. In supply chain coordination with contracts, usually contract parameters can be arbitrarily chosen where the realistic

outcomes are the result of bargaining power or negotiations between the supplier and the retailer. In our model, the contract parameters emerge as the result of the competition between suppliers.

## Chapter 8

# Appendices

### 8.1 Appendix A: Sum of two Gumbel distributed random variables

The sum of two gumbel distributions does not have a closed form distribution function. We approximated in Chapter 5, the sum of gumbel distributed random variables with a gumbel distribution. Here we discuss the quality of this approximation. One measure of similarity of two probability distributions is the Kolmogorov-Smirnov (K-S) distance, which is the supremum of the distance between the cumulative density functions. Consider the case with two random variables,  $X, Y, Z$  be Gumbel with parameters  $(\eta_X, \mu_X), (\eta_Y, \mu_Y)$ , and  $(\eta_Z, \mu_Z)$ . The mean of  $X + Y$  is  $\eta_X + \eta_Y + \gamma(\mu_X + \mu_Y)$  and variance is  $\pi^2 \mu^2 / 6$ . We set the parameters of  $Z$  to have the same mean and variance with  $X + Y$  as follows.

$$\mu_Z = \sqrt{\mu_X^2 + \mu_Y^2} \text{ and } \eta_Z = \eta_X + \eta_Y + \gamma(\mu_X + \mu_Y - \mu_Z)$$

K-S distance is

$$\sup_u |F_{X+Y}(u) - F_Z(u)| \approx \max_{p \in \{0.1, \dots, 0.99\}} |F_{X+Y}(F_Z^{-1}(p)) - p|$$

We computed  $F_{X+Y}$  using numerical integration and searched for the maximum over percentiles for the combination of the following parameters.  $\eta_X = \{10, 40\}, \eta_Y = \{10, 40\}, \mu_X = \{5, 20\}, \mu_Y = \{5, 20\}$ . K-S distance never exceeded 0.019 in these examples, which we believe is fairly small. It is also of interest to note that the K-S distance seems to be effected by only the ratio of  $\mu_X/\mu_Y$  but none of the other parameters, reaching the maximum at  $\mu_X/\mu_Y = 1$ . Figure 8.1. provides an example of the cumulative density functions of  $X+Y$  and  $Z$ .

By the central limit theorem, the sum of  $n$  gumbel distributions asymptotically converges to a normal distribution. Since we approximate sum of any number of Gumbel distributions with a Gumbel distribution, measuring the distance of a Gumbel to a normal distribution can give us an idea of the quality of the approximation in the limit. Regardless of the parameters of the normal distribution, the K-S distance is 0.07 provided at the median of the normal distribution. This may be considered as a reasonably small asymptotic upper bound.

Finally, from the perspective of our model, we are interested more in how the resulting choice probabilities are affected with the approximation. Consider that  $X, Y, A, B$  are all Gumbel distributions,  $X+Y$  and  $A+B$  are approximated by Gumbels  $Z$  and  $C$  respectively. We compare  $\Pr\{X + Y < A + B\}$  and  $\Pr\{Z < C\}$ . Notice that the latter one can be computed using the MNL formula. For the combination of the set of parameters above, the maximum difference in probabilities is 0.010, which is also fairly small and in the same

order of magnitude with the K-S distance. Table 8.1. presents a set of resulting choice probabilities with both the true distributions and the Gumbel approximations.

## 8.2 Appendix B: Example of a replenishment system with convex operating cost

Suppose there are  $x$  products in the category. Category demand is  $\lambda(x)$ . Recall that in our model  $\lambda(x) = \frac{A(x)}{A(x) + A(z)}$  where  $A'(x) > 0$  and  $A''(x) \geq 0$ . Each item's demand rate is  $\lambda_i(x)$ , where  $\sum_i \lambda_i(x) = \lambda(x)$ . Total shelf space capacity is  $S$ , assume that each unit has a maximum inventory level proportional to its' demand rate, i.e.,  $S_i = S\lambda_i(x)/\lambda(x)$ . Consider a continuous review system with zero lead time for each product: an order is placed whenever the inventory level drops to zero. Average inventory level for a single product is  $S_i/2$ . Total average inventory is  $S/2$ , hence inventory holding cost is independent of  $x$ . Number of orders per unit time for each product is  $\frac{\lambda_i(x)}{S_i(x)}$  which is equal to  $\frac{\lambda(x)}{S(x)}$ . That is the frequency of orders is the same for all products. Total number of orders in the category is  $x\lambda(x)/S$ , whose second derivative is

$$\begin{aligned} \frac{2\lambda'(x) + x\lambda''(x)}{S} &= \frac{A(z)}{(A(x) + A(z))^3} (2A'(x)((A(x) + A(z) - xA'(x) + A''(x)(A(x) + A(z)))) \\ &\geq \frac{A(z)}{(A(x) + A(z))^3} (2A'(x)A(z) + A''(x)(A(x) + A(z))) \end{aligned}$$

The inequality is due to the concavity of  $A(x)$ , i.e.,  $A''(x) \leq 0$  and  $A(0) = 0$  implies  $A(x) \geq xA'(x)$ . If the  $x$  products have similar demand rates,  $A(x)$  is linear.  $A''(x) = 0$  and the cost function is convex in  $x$ . With strictly concave  $A(x)$ , the cost function is still convex if  $A(x)$  is not too concave. For example, if  $A(x) = \ln x$ , the sign of the second

derivative is determined by  $A(z)(-1 + 2x) - \ln x$ , which is strictly positive for  $z \geq e$  and  $x > 1$ .

### 8.3 Appendix C: Proof of second order properties of the contract parameters in Model DC :

#### 8.3.1 Wholesale price contracts

$$\begin{aligned}
\frac{\partial^2 q_i}{\partial w_i^2} &= \frac{\partial}{\partial w_i} \frac{-\gamma b^2 - \delta}{\gamma \delta (1 - b^2)^2} = -(1 - b^2)^{-2} \left( \frac{\partial}{\partial q_i} \frac{\gamma b^2 + \delta}{\gamma \delta} \frac{\partial q_i}{\partial w_i} + \frac{\partial}{\partial q_j} \frac{\gamma b^2 + \delta}{\gamma \delta} \frac{\partial q_j}{\partial w_i} \right) \\
&= -(1 - b^2)^{-4} (\gamma \delta)^{-3} \left( \begin{aligned} &(\gamma \delta (b^2 \gamma' + \delta' b)) - (\gamma' \delta + \gamma \delta' b)(\gamma b^2 + \delta)(-\gamma b^2 - \delta) \\ &+ (\gamma \delta (b^3 \gamma' + \delta') - (b \gamma' \delta + \gamma \delta')(\gamma b^2 + \delta))(\gamma + \delta) b \end{aligned} \right) \\
&= -(1 - b^2)^{-4} (\gamma \delta)^{-3} \left( \begin{aligned} &\gamma \delta ((b^2 \gamma' + \delta' b)(-\gamma b^2 - \delta) + (b^3 \gamma' + \delta')(\gamma + \delta) b) \\ &-(\gamma b^2 + \delta)((\gamma' \delta + \gamma \delta' b)(-\gamma b^2 - \delta) + (b \gamma' \delta + \gamma \delta')(\gamma + \delta) b) \end{aligned} \right) \\
&= -(1 - b^2)^{-4} (\gamma \delta)^{-3} \left( \begin{aligned} &\gamma \delta (1 - b^2)(b \delta' \gamma - b^2 \gamma' \delta) \\ &-(\gamma b^2 + \delta)(1 - b^2)(b \gamma^2 \delta' - \gamma' \delta^2) \end{aligned} \right) \\
&= -(1 - b^2)^{-3} (\gamma \delta)^{-3} \left( \begin{aligned} &\gamma \delta (b \delta' \gamma - b^2 \gamma' \delta) \\ &-(\gamma b^2 + \delta)(b \gamma^2 \delta' - \gamma' \delta^2) \end{aligned} \right) \\
&= -(1 - b^2)^{-3} (\gamma \delta)^{-3} (\gamma' \delta^3 - \delta' \gamma^3 b^3) \\
&= 0 \text{ for uniform distribution.}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 q_i}{\partial w_j \partial w_i} &= \frac{\partial}{\partial w_j} \frac{-\gamma b^2 - \delta}{\gamma \delta (1 - b^2)^2} = -(1 - b^2)^{-2} \left( \frac{\partial}{\partial q_i} \frac{\gamma b^2 + \delta}{\gamma \delta} \frac{\partial q_i}{\partial w_j} + \frac{\partial}{\partial q_j} \frac{\gamma b^2 + \delta}{\gamma \delta} \frac{\partial q_j}{\partial w_i} \right) \\
&= -(1 - b^2)^{-4} (\gamma \delta)^{-3} \left( \begin{aligned} &(\gamma \delta (b^2 \gamma' + \delta' b)) - (\gamma' \delta + \gamma \delta' b)(\gamma b^2 + \delta)(\gamma + \delta) b \\ &+ (\gamma \delta (b^3 \gamma' + \delta') - (b \gamma' \delta + \gamma \delta')(\gamma b^2 + \delta))(-\delta b^2 - \gamma) \end{aligned} \right) \\
&= -(1 - b^2)^{-4} (\gamma \delta)^{-3} \left( \begin{aligned} &\gamma \delta ((b^2 \gamma' + \delta' b)(\gamma + \delta) b + \\ &(b^3 \gamma' + \delta')(-\delta b^2 - \gamma)) \\ &-(\gamma b^2 + \delta)((\gamma' \delta + \gamma \delta' b)(\gamma + \delta) b + \\ &(b \gamma' \delta + \gamma \delta')(-\delta b^2 - \gamma)) \end{aligned} \right) \\
&= -(1 - b^2)^{-4} (\gamma \delta)^{-3} \left( \begin{aligned} &\gamma \delta (1 - b^2)(b \delta' \gamma - b^2 \gamma' \delta) \\ &-(\gamma b^2 + \delta)(1 - b^2)(b \gamma^2 \delta' - \gamma' \delta^2) \end{aligned} \right) \\
&= -(1 - b^2)^{-3} (\gamma \delta)^{-3} \left( \begin{aligned} &\gamma \delta (b^3 \gamma' \delta - \gamma \delta') \\ &-(\gamma b^2 + \delta)(b \gamma' \delta^2 - \gamma^2 \delta') \end{aligned} \right) \\
&= -(1 - b^2)^{-3} (\gamma \delta)^{-3} (\gamma' \delta^3 b - \delta' \gamma b^2) \\
&= 0 \text{ for uniform distribution.}
\end{aligned}$$

### 8.3.2 Revenue sharing contracts

The application of Implicit Function Theorem to the retailer's optimality equations (6.9)

yields

$$\begin{vmatrix} -\phi_i p_i f_{D_i}(q_i) - \phi_j p_j f_{D_j}(q_j) b^2 & -\phi_i p_i f_{D_i}(q_i) b - \phi_j p_j f_{D_j}(q_j) b \\ -\phi_i p_i f_{D_i}(q_i) b - \phi_j p_j f_{D_j}(q_j) b & -\phi_i p_i f_{D_i}(q_i) b^2 - \phi_j p_j f_{D_j}(q_j) \end{vmatrix} \begin{vmatrix} \frac{\partial q_i}{\partial \phi_i} \\ \frac{\partial q_j}{\partial \phi_i} \end{vmatrix} = - \begin{vmatrix} \frac{\partial R_i}{\partial \phi_i} \\ \frac{\partial R_j}{\partial \phi_i} \end{vmatrix}$$

$$\gamma = \phi_i p_i f(q_i + b q_j), \partial \gamma / \partial q_i = \gamma', \partial \gamma / \partial q_j = b \gamma'$$

$$\delta = \phi_j p_j f_{D_j}(bq_i + q_j), \partial\delta/\partial q_i = b\delta', \partial\delta/\partial q_j = \delta'$$

$$\eta 1 = \partial R_i / \partial \phi_i = p_i \Pr\{q_i \leq D_i\} - c = p_i - p_i \Pr\{D_i \leq q_i\} - c$$

$$\eta 2 = -p_i \Pr\{D_i \leq q_i\} b$$

$$\begin{vmatrix} -\gamma - \delta b^2 & -\gamma b - \delta b \\ -\gamma b - \delta b & -\gamma b^2 - \delta \end{vmatrix} \begin{vmatrix} \frac{\partial q_i}{\partial \phi_i} \\ \frac{\partial q_j}{\partial \phi_i} \end{vmatrix} = - \begin{vmatrix} p_i - \eta - c_i \\ -\eta b \end{vmatrix}$$

$$\begin{aligned} \begin{vmatrix} \frac{\partial q_i}{\partial \phi_i} \\ \frac{\partial q_j}{\partial \phi_i} \end{vmatrix} &= \frac{1}{\gamma \delta (1 - b^2)^2} \begin{vmatrix} -\gamma b^2 - \delta & (\gamma + \delta)b \\ (\gamma + \delta)b & -\gamma - \delta b^2 \end{vmatrix} \begin{vmatrix} \eta - (p_i - c_i) \\ \eta b \end{vmatrix} \\ \begin{vmatrix} \frac{\partial q_i}{\partial \phi_i} \\ \frac{\partial q_j}{\partial \phi_i} \end{vmatrix} &= \frac{(1 + b)}{\gamma \delta (1 - b^2)^2} \begin{vmatrix} (\gamma b^2 + \delta)(p_i - c_i) - \eta \delta (1 - b^2) \\ -(\gamma + \delta)b(p_i - c_i) + \eta \delta b(1 - b^2) \end{vmatrix} \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 q_i}{\partial \phi_i^2} &= \frac{\partial}{\partial \phi_i} \left( \frac{(1 + b)}{(1 - b^2)^2} \left| \frac{(\gamma b^2 + \delta)(p_i - c_i) - \eta \delta (1 - b^2)}{\gamma \delta} \right| \right) \\ &= \frac{(1 + b)}{(1 - b^2)^2} \left( \frac{\partial}{\partial \phi_i} () + \frac{\partial}{\partial q_i} () \frac{\partial q_i}{\partial \phi_i} + \frac{\partial}{\partial q_j} () \frac{\partial q_j}{\partial \phi_i} \right) \\ &= \frac{(1 + b)}{(1 - b^2)^2} \left( \begin{aligned} & -\frac{(p_i - c_i) - \eta(1 - b^2)}{\gamma^2} \frac{\gamma}{\phi_i} + \\ & \gamma^{-2} \delta^{-2} \begin{pmatrix} \gamma \delta \begin{pmatrix} (\gamma' b^2 + \delta' b)(p_i - c_i) \\ -(1 - b^2)(\delta \gamma / \phi_i + \eta \delta' b) \end{pmatrix} \\ ((\gamma b^2 + \delta)(p_i - c_i) - \eta \delta (1 - b^2))(\gamma' \delta + \gamma \delta' b) \end{pmatrix} \frac{\partial q_i}{\partial \phi_i} + \\ & \gamma^{-2} \delta^{-2} \begin{pmatrix} \gamma \delta (\gamma' b^3 + \delta') (p_i - c_i) \\ -(1 - b^2)(\delta \gamma b / \phi_i + \eta \delta') \\ ((\gamma b^2 + \delta)(p_i - c_i) - \eta \delta (1 - b^2))(\gamma' b \delta + \gamma \delta') \end{pmatrix} \frac{\partial q_j}{\partial \phi_i} \end{aligned} \right) \end{aligned}$$

for  $\gamma'\delta' = 0$

$$\begin{aligned}
&= \frac{(1+b)}{(1-b^2)^2} \begin{pmatrix} -\frac{(p_i-c_i)-\eta(1-b^2)}{\gamma^2} \frac{\gamma}{\phi_i} \\ +\gamma^{-2}\delta^{-2} \left( \gamma\delta \left( -(1-b^2)(\delta\gamma/\phi_i) \right) \right) \frac{\partial q_i}{\partial \phi_i} + \\ \gamma^{-2}\delta^{-2} \left( \gamma\delta \left( -(1-b^2)(\delta\gamma b/\phi_i) \right) \right) \frac{\partial q_i}{\partial \phi_i} \end{pmatrix} \\
&= \frac{(1+b)}{(1-b^2)^2} \frac{\gamma}{\phi_i} \gamma^{-2} \begin{pmatrix} -((p_i-c_i)-\eta(1-b^2)) \\ +(1-b^2)\gamma \left( -\frac{\partial q_i}{\partial \phi_i} - b\frac{\partial q_i}{\partial \phi_i} \right) \end{pmatrix} \\
&= \frac{(1+b)}{(1-b^2)^2} \frac{\gamma}{\phi_i} \gamma^{-2} \begin{pmatrix} -((p_i-c_i)-\eta(1-b^2)) \\ +\frac{(1-b^2)\gamma(1+b)}{\gamma\delta(1-b^2)^2} -((\gamma b^2+\delta)(p_i-c_i)-\eta\delta(1-b^2)) \\ -b \left( -(\gamma+\delta)b(p_i-c_i) + \eta\delta b(1-b^2) \right) \end{pmatrix} \\
&= \frac{(1+b)}{(1-b^2)^2} \frac{\gamma}{\phi_i} \gamma^{-2} \begin{pmatrix} -((p_i-c_i)-\eta(1-b^2)) \\ -(1+b) \left( (p_i-c_i)-\eta(1-b^2) \right) \end{pmatrix} \\
&= \frac{(1+b)}{(1-b^2)^2} \frac{\gamma}{\phi_i} \gamma^{-2} (-1-(1+b)) \left( (p_i-c_i)-\eta(1-b^2) \right) \\
&\leq 0 \text{ if } (p_i-c_i) > \eta(1-b^2). \begin{pmatrix} \text{same condition is sufficient to} \\ \text{have } \partial q_i/\partial \phi_i > 0 \text{ and } \partial q_j/\partial \phi_i < 0 \end{pmatrix}
\end{aligned}$$

The cross partial is

$$\begin{aligned}
\frac{\partial^2 q_i}{\partial \phi_j \partial \phi_i} &= \frac{\partial}{\partial \phi_j} \left( \frac{(1+b)}{(1-b^2)^2} \left| \frac{(\gamma b^2 + \delta)(p_i - c_i) - \eta \delta(1-b^2)}{\gamma \delta} \right| \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( \frac{\partial}{\partial \phi_j} () + \frac{\partial}{\partial q_i} () \frac{\partial q_i}{\partial \phi_j} + \frac{\partial}{\partial q_j} () \frac{\partial q_j}{\partial \phi_j} \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( \begin{aligned} & -\frac{(p_i - c_i)b^2}{\delta^{-2}} \frac{\delta}{\phi_j} + \\ & \gamma^{-2} \delta^{-2} \left( \gamma \delta \begin{pmatrix} (\gamma' b^2 + \delta' b)(p_i - c_i) \\ -(1-b^2)(\delta \gamma / \phi_i + \eta \delta' b) \end{pmatrix} \right) \frac{\partial q_i}{\partial \phi_j} + \\ & ((\gamma b^2 + \delta)(p_i - c_i) \\ & -\eta \delta(1-b^2))(\gamma' \delta + \gamma \delta' b) \end{aligned} \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( \begin{aligned} & \gamma^{-2} \delta^{-2} \begin{pmatrix} \gamma \delta (\gamma' b^3 + \delta')(p_i - c_i) \\ -(1-b^2)(\delta \gamma b / \phi_i + \eta \delta') \\ ((\gamma b^2 + \delta)(p_i - c_i) \\ -\eta \delta(1-b^2))(\gamma' b \delta + \gamma \delta') \end{pmatrix} \frac{\partial q_j}{\partial \phi_j} \end{aligned} \right)
\end{aligned}$$

$$\text{for } \gamma' = \delta' = 0$$

$$\begin{aligned}
&= \frac{(1+b)}{(1-b^2)^2} \left( -\frac{(p_i-c_i)b^2}{\delta^{-2}} \frac{\delta}{\phi_j} + \gamma^{-2} \delta^{-2} (\gamma \delta (-(1-b^2)(\delta \gamma / \phi_i))) \frac{\partial q_i}{\partial \phi_j} + \right. \\
&\quad \left. \gamma^{-2} \delta^{-2} (\gamma \delta (-(1-b^2)(\delta \gamma b / \phi_i))) \frac{\partial q_i}{\partial \phi_j} \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( -(p_i-c_i)b^2/(\delta \phi_j) - \frac{(1-b^2)}{\phi_i} \left( \frac{\partial q_i}{\partial \phi_j} + b \frac{\partial q_j}{\partial \phi_j} \right) \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( -(p_i-c_i)b^2/(\delta \phi_j) - \frac{(1-b^2)}{\phi_i} \left( \frac{\partial q_i}{\partial \phi_j} + b \frac{\partial q_j}{\partial \phi_j} \right) \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( \begin{array}{c} -(p_i-c_i)b^2/(\delta \phi_j) \\ -\frac{(1-b^2)}{\phi_i} \frac{(1+b)}{\gamma \delta (1-b^2)^2} \left( \begin{array}{c} -(\gamma + \delta)b(p_j - c_j) + \eta_j \gamma b(1-b^2) \\ b(\delta b^2 + \gamma)(p_j - c_j) - \eta_j \gamma b(1-b^2) \end{array} \right) \end{array} \right) \\
&= \frac{(1+b)}{(1-b^2)^2} \left( -(p_i-c_i)b^2/(\delta \phi_j) + \frac{(1-b^2)}{\phi_i} \frac{(1+b)}{\gamma \delta (1-b^2)^2} b(p_j - c_j) \delta (1-b^2) \right) \\
&= \frac{(1+b)b}{(1-b^2)^2} \left( -\frac{(p_i-c_i)b}{\delta \phi_j} + \frac{(1+b)}{\phi_i \gamma} (p_j - c_j) \right) \\
&\geq 0 \text{ if } \frac{\phi_i \gamma}{\delta \phi_j} < \frac{(1+b)(p_j - c_j)}{b(p_i - c_i)}
\end{aligned}$$

which holds for symmetric data and symmetric solution.

$$-q_i(\phi) + (1 - \phi_i) \frac{\partial q_i(\phi)}{\partial \phi_i} = 0$$

$$\frac{\partial \phi_i^*}{\partial \phi_j} = -\frac{-\frac{\partial q_i(\phi)}{\partial \phi_i} + (1 - \phi_i) \frac{\partial^2 q_i(\phi)}{\partial \phi_i \partial \phi_i}}{-2 \frac{\partial q_i(\phi)}{\partial \phi_i} + (1 - \phi_i) \frac{\partial^2 q_i(\phi)}{\partial \phi_i^2}}$$

For uniqueness we need

$$-\frac{\partial q_i(\phi)}{\partial \phi_j} + (1 - \phi_i) \frac{\partial^2 q_i(\phi)}{\partial \phi_j \partial \phi_i} \leq 2 \frac{\partial q_i(\phi)}{\partial \phi_i} - (1 - \phi_i) \frac{\partial^2 q_i(\phi)}{\partial \phi_i^2}$$

$$\left( \begin{array}{c} -\frac{(1+b)}{\gamma\delta(1-b^2)^2} \left( \begin{array}{c} -(\gamma+\delta)b(p_j - c_j) \\ +\eta_j\gamma b(1-b^2) \end{array} \right) \\ (1-\phi_i)\frac{(1+b)b}{(1-b^2)^2} \left( \begin{array}{c} -\frac{(p_i-c_i)b}{\delta\phi_j} \\ +\frac{(1+b)}{\phi_i\gamma}(p_j - c_j) \end{array} \right) \end{array} \right) < \left( \begin{array}{c} \frac{2(1+b)}{\gamma\delta(1-b^2)^2} \left( \begin{array}{c} (\gamma b^2 + \delta)(p_i - c_i) \\ -\eta\delta(1-b^2) \end{array} \right) \\ - \left( \begin{array}{c} (1-\phi_i)\frac{(1+b)}{(1-b^2)^2}\frac{\gamma}{\phi_i} \\ \times \gamma^{-2}(-1-(1+b)) \end{array} \right) \\ \times \left( \begin{array}{c} (p_i - c_i) \\ -\eta(1-b^2) \end{array} \right) \end{array} \right)$$

at symmetric data & solution,  $\phi_i = \phi_j, q_i = q_j$ .

$$\left( \begin{array}{c} -\frac{1}{\gamma\gamma} \left( \begin{array}{c} -2\gamma b(p-c) \\ +\eta\gamma b(1-b^2) \end{array} \right) \\ (1-\phi)b \left( \begin{array}{c} -\frac{(p-c)b}{\gamma\phi} \\ +\frac{(1+b)}{\phi\gamma}(p-c) \end{array} \right) \end{array} \right) < \left( \begin{array}{c} \frac{2}{\gamma\gamma} \left( \begin{array}{c} (\gamma(b^2+1)(p-c) \\ -\eta\gamma(1-b^2) \end{array} \right) \\ - \left( \begin{array}{c} (1-\phi)\frac{1}{\phi} \\ \times \gamma^{-1}(-2-b) \end{array} \right) \\ \times \left( \begin{array}{c} (p-c) \\ -\eta(1-b^2) \end{array} \right) \end{array} \right)$$

$$\left( \begin{array}{c} -(-2b(p-c) + \eta b(1-b^2)) \\ \frac{(1-\phi)}{\phi} b(-(p-c)b + (1+b)(p-c)) \end{array} \right) < \left( \begin{array}{c} 2((b^2+1)(p-c) - \eta(1-b^2)) \\ -\frac{(1-\phi)}{\phi}(-2-b)((p-c) - \eta(1-b^2)) \end{array} \right)$$

$$\left( \begin{array}{c} b(2(p-c) - \eta(1-b^2)) \\ \frac{(1-\phi)}{\phi} b(p-c) \end{array} \right) < \left( \begin{array}{c} 2(b^2+1)(p-c) - 2\eta(1-b^2) \\ \frac{(1-\phi)}{\phi}(2+b)((p-c) - \eta(1-b^2)) \end{array} \right)$$

$$\eta(1-b^2)(-b+2 + \frac{(1-\phi)}{\phi}(2+b)) < (p-c) \left( \begin{array}{c} \frac{(1-\phi)}{\phi}(2+b-b) \\ +2(b^2+1) - 2b \end{array} \right)$$

$$\eta(1-b^2)(-b) < (p-c) \left( \frac{(1-\phi)}{\phi}(-b) + b(2b-2) \right)$$

$$-\eta(1-b^2) < (p-c)(2b-2)$$

$$1 < 2b$$

# Bibliography

- [1] Anderson, S.P., A. de Palma., J.F. Thisse. 1992. *Discrete choice theory of product differentiation*. Cambridge, MA. The MIT Press.
- [2] Anupindi R., M. Dada, S. Gupta. 1998. Estimation of consumer demand with stockout based substitution: An application to vending machine products. *Marketing Science*. **17** 406-423.
- [3] Anupindi, R., Y. Bassok. 1999. Centralization of stocks: retailers vs. manufacturers. *Management Science*. **45** 178-191.
- [4] Avsar, Z.M., M. Baykal-Gursoy. 2002. Inventory control under substitutable demand: A stochastic game application. *Naval Research Logistics*. **49**. 359-375
- [5] Bell D.R., G.J. Fitzsimons. 1999. An experimental and empirical analysis of consumer response to stockouts. Working paper. Marketing department. Wharton School.
- [6] Bell, D.E., R.L. Keeney, J.D.C. Little. 1975. Market share theorem. *Journal of Marketing Research*. **12** 136-41.

- [7] Bell, D. R., T-H Ho, C.S. Tang 1998. Determining where to shop: Fixed and variable costs of shopping. *Journal of Marketing Research*. **35** 352-369.
- [8] Bell, D.R., J.M. Lattin. 1998. Shopping behavior and consumer preference for store price format: Why large basket shoppers prefer EDLP. *Marketing Science*. **17** 66-88.
- [9] Ben-Akiva, M. S.R. Lerman. 1985. *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA. The MIT Press.
- [10] Benson, H.P., S.S. Erenguc, R. Horst. 1990. A note on adapting methods for continuous global optimization to the discrete case. *Annals of Operations Research*. **25** 243-252.
- [11] Bernstein, F., F. Chen, A. Federgruen. 2003. Coordinating supply chains with simple pricing schemes: the role of vendor managed inventories. Working paper, Fuqua School of Business, Duke University.
- [12] Bernstein, F., F. de Vericourt. 2002. Allocation of supply contracts with service guarantees. Working paper, Fuqua School of Business, Duke University.
- [13] Bernstein, F., A. Federgruen. 2001. Decentralized supply chains with competing retailers under demand uncertainty. Forthcoming in *Management Science*.
- [14] Bernstein, F., A. Federgruen. 2003. Pricing and replenishment strategies in a distribution system with competing retailers. *Operations Research*. **51** 409-426.
- [15] Boatwright, P., J.C. Nunes. 2001. Reducing assortment: An attribute-based approach. *Journal of Marketing* **65** 50-63.

- [16] Bodapati, A.V., V. Srinivasan. 1999. The impact of out-of-store advertising on store sales. Working paper. Northwestern University.
- [17] Borin, N., P. Farris. 1995. A sensitivity analysis of retailer shelf management models. *J Retailing*. **71** 153-171.
- [18] Bretthauer, K.M., B. Shetty. 2002. The nonlinear knapsack problem - algorithms and applications. *European J Operations Research*. **138** 459-472.
- [19] Broniarczyk, S.M., W.D. Hoyer, L. McAlister. 1998. Consumers' perception of the assortment offered in a grocery category: The impact of item reduction. *Journal of Marketing Research*. **35** 166-76.
- [20] Bucklin, R.E., S. Gupta. 1992. Brand choice, purchase incidence, and segmentation: An integrated modeling approach. *Journal of Marketing Research*. **29** 201-215.
- [21] Bultez, A., P. Naert. 1988. SHARP: Shelf allocation for retailers profit. *Marketing Science*. **7** 211-231.
- [22] Cachon, G.P. 2002. Supply chain coordination with contracts. Forthcoming in *Handbooks in Operations Research and Management Science: Supply Chain Management*. S. Graves and T. de Kok, editors.
- [23] Cachon, G.P., P.T. Harker. 2002. Competition and outsourcing with scale economies. *Management Science*. **48** 1314-33
- [24] Cachon, G.P., A.G. Kok. 2002. Heuristic equilibrium in the newsvendor model with clearance pricing. Working paper, Wharton School, University of Pennsylvania.

- [25] Cachon G. P., M. Lariviere 2000. Supply chain coordination with revenue sharing contracts: strengths and limitations. Forthcoming in *Management Science*.
- [26] Cachon G. P., M. Lariviere 2001. Turning the supply chain into a revenue chain. *Harvard Business Review*. March.
- [27] Caprara, A., D. Pisinger, P. Toth. 1999. Exact solution of the quadratic knapsack problem. *INFORMS Journal on Computing*. **11** 125-137.
- [28] Chen, F., 2001. Auctioning supply contracts. Working paper, Columbia University.
- [29] Chen, F., A. Federgruen, Y-S. Zheng. 2001. Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Management Science*. **47** 693-708.
- [30] Chiang, J. 1991. A simultaneous approach to the whether, what and how much to buy questions. *Marketing Science*. **10** 297-315.
- [31] Chintagunta, P.K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*. **12** 184-208.
- [32] Choi, S.C. 1991. Price competition in a channel structure with a common retailer. *Marketing Science*. **10** 271-296.
- [33] Corstjens, M., P. Doyle. 1981. A model for optimizing retail space allocations. *Management Science*. **27** 822-833.
- [34] Dana Jr., J., and N. Petruzzi. 2001. Note: The newsvendor model with endogenous demand. *Management Science* **47** 1488-1497.

- [35] Downs B., R. Metters, J. Semple. 2002. Managing inventory with multiple products, lags in delivery, resource constraints, and lost sales: A mathematical programming approach. *Management Science*. **47** 464-479.
- [36] Dreze, X., S.J. Hoch and M.E. Purk. 1994. Shelf Management and Space Elasticity. *Journal of Retailing* **70** 301-26.
- [37] Dussault, J.P., J.A. Ferland, B. Lemaire.1986. Convex quadratic programming with one constraint and bounded variables. *Mathematical Programming* **36** 90-104.
- [38] Eppen, G.D. 1979. Effects of centralization on expected costs in multi-location news-boy problems. *Management Science*. **25**. 498-501.
- [39] Fisher, M.L., A. Raman, A.S. McClelland. 2000. Rocket science retailing is almost here - Are you ready? *Harvard Business Review*. July/August 115-124.
- [40] Food marketing Institute Report. 1993. Variety or duplication: A process to know where you stand. Prepared by Willard Bishop Consulting and Information Resources in cooperation with Frito Lay.
- [41] Fudenberg,D., J. Tirole 2000. *Game Theory*. Cambridge, MA. The MIT Press.
- [42] Gallo, G., P.L. Hammer, B. Simeone. 1980. Quadratic knapsack problems. *Mathematical Programming*. **12** 132-149.
- [43] Gans, N. 2002. Customer Loyalty and Supplier Quality Competition. *Management Science*. **48**. 207-21.

- [44] Gaur, V. 2001. Essays in Retail Operations Management. Ph.D. Dissertation. The Wharton School, University of Pennsylvania.
- [45] Guadagni, P.M., J.D.C. Little. 1983. A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science*. **2** 203-238.
- [46] Hadley, G; T.M. Whitin. 1963. *Analysis of Inventory Systems*. Prentice Hall.
- [47] Hall, J., E. Porteus. 2000. Customer Service Competition in Capacitated Systems. *Manufacturing & Service Operations Mgmt.* **2** 144-65.
- [48] Hoch, S.J., E.T. Bradlow and B. Wansink. 1999. The variety of an assortment. *Marketing Science*. **21** 342-55.
- [49] Hochbaum, D.S. 1995. A nonlinear knpsack problem. *Operations Research Letters*. **17** 103-110.
- [50] Horn, R.A., C.R. Johnson. 1996. *Matrix Analysis*. New York, NY. Cambridge University Press.
- [51] Hotelling, H. 1929. Stability in competition. *Economic Journal*. **39** 41-57
- [52] Horst, R., H. Tuy. 1990. *Global Optimization: Deterministic Approaches*. Berlin. Springer.
- [53] Huffman, C., B.E. Kahn. 1998. Variety for sale: Mass customization or mass confusion? *Journal of Retailing*. **74** 491-513.
- [54] Jayaraman V., R. Srivastava, W.C. Benton. 1998. A joint optimization of product variety and ordering approach. *Comput Oper. Res.* **25** 557-566.

- [55] Jeuland A. P. and S. P. Shugan 1983. Managing channle profits. *Marketing Science*. **2** 239-272.
- [56] Kahn, B.E. 1995. Consumer variety-seeking in goods and services: an integrative review. *Journal of Retailing and Consumer Services*. **2** 139-48.
- [57] Klemperer, P., 1992. Equilibrium product lines: Competing head-to-head may be less competitive. *American Economic Review*. **82** 740-55.
- [58] Kurt Salmon Associates. 1993. Efficient consumer response: Enhancing consumer value in the grocery industry. *Food Marketing Institute Report #* 9-526. Washington, DC: Food Marketing Institute.
- [59] Lancaster, K. 1966. A new approach to consumer theory. *Journal of Political Economy*. **74** 132-57.
- [60] Lariviere, M., E. Porteus. 2001. Selling to the newsvendor: an analysis of price-only contracts. *Manufacturing and Service Operations Management*. **3** 293-305.
- [61] Lederer, P.J., L.D. Li. 1997. Pricing, production, scheduling, and delivery-time competition. *Operations Research*. **45** 407-420.
- [62] Cooper, L.G., M. Nakanishi. 1988. *Market-Share Analysis: Evaluating Competitive Marketing Effectiveness*. Kluwer Academic Publishers.
- [63] Lee, H.L., M.J. Rosenblatt. 1986. A generalized quantity discount pricing model to increase supplier's profits. *Management Science*. **32** 1177-1185.

- [64] Li L.D., Y.S. Lee. 1994. Pricing and delivery-time performance in a competitive environment. *Management Science*. **40** 633-646.
- [65] Lippman S.A., K.F. McCardle. 1997. The competitive newsboy. *Operations Research*. **45** 54-65.
- [66] Mahajan, S., G. van Ryzin. 1999. Supply chain coordination under horizontal competition. Working paper. Columbia University.
- [67] Mahajan, S., G. van Ryzin. 2001. Stocking retail assortments under dynamic consumer substitution. *Operations Research*. **49** 334-351.
- [68] Mahajan, S., G. van Ryzin. 2001b. Inventory competition under dynamic consumer choice. *Operations Research*. **49** 646-657.
- [69] Martello, S., P. Toth. 1990. *Knapsack Problems: Algorithms and Computer Implementations*. Wiley, New York.
- [70] McDowell, S.D., J. Lee. 2003. India's experiments in mobile licensing. *Telecommunications Policy*. **27** 371-382.
- [71] McGuire T.W., R. Staelin. 1983. An industry equilibrium analysis of downstream vertical integration. *Marketing Science*. **2** 161-191.
- [72] Monahan, J.P., 1984. A quantity discount pricing model to increase vendor profits. *Management Science*. **30** 720-726.
- [73] Moorthy, S. 1984. Market segmentation, self-selection, and product line design. *Marketing Science*. **3** 288-307.

- [74] Moorthy, S. 1987. Product line competition. *Annales des Telecommunications*. **42** 655-663.
- [75] Nahmias, S., C.P. Schmidt. 1984. An efficient heuristic for the multi-item newsboy model with a single resource constraint. *Naval Research Logistics*. **31**. 463-474.
- [76] Netessine, S., N. Rudi. 2003. Centralized and competitive inventory models with demand substitution. *Operations Research*. **51**. 329-335.
- [77] Netessine S. F. Zhang. 2003. The impact of supply-side externalities among downstream firms on supply chain efficiency. Working paper. The Wharton School. University of Pennsylvania.
- [78] Nijs, V.R., M.G. Dekimpe, J.E.M. Steenkamp, D.M. Hanssens. 2001. The category-demand effects of price promotions. *Marketing Science*. **20**. 1-22
- [79] Noble, P.M., T.S. Gruca. 1999. Industrial pricing: Theory and managerial practice. *Marketing Science* **18** 435-454.
- [80] Parlar, M. 1985. Optimal ordering policies for a perishable and substitutable product: A Markov decision model. *Infor*. **23**. 182-195.
- [81] Pauwels, K., D.M. Hanssens, S. Siddharth. 2002. The long-term effects of price promotions on category incidence, brand choice and purchase quantity. *Journal of Marketing research*. **39**. 421-439.
- [82] Quelch, J.A., D. Kenny. 1994. Extend profits, not product lines. *Harvard Business Review*. **72** 153-160.

- [83] Raman, A., N. DeHoratius, Z. Ton., 2001. The Achilles Heel of Supply Chain Management. *Harvard Business Review* **79**136-152.
- [84] Rajaram K. 2001. Assortment planning in fashion retailing: methodology, application and analysis. *Eur J of Oper Res.* **129** 186-208.
- [85] Russell, G.J., S. Bell, D.R. et.al. 1997. Perspectives on multiple category choice. *Marketting Letters.* **8** 297-305.
- [86] Schary, P.B., M. Christopher. 1979. Anatomy of a stock out. *Journal of Retailing.* **55** 59-70.
- [87] Smith, S.A., D.D. Achabal. 1998. Clearance pricing and inventory policies for retail chains. *Management Science* **44** 285-300.
- [88] Smith, S.A., N. Agrawal. 2000. Management of multi-item retail inventory systems with demand substitution. *Operations Research.* **48** 50-64.
- [89] Simonson, I. 1999. The effect of product assortment on buyer preferences. *Journal of Retailing.* **75** 347-70.
- [90] Thonemann, U.W., J.R. Bradley. 2002. The effect of product variety on supply-chain performance. *European Journal of Operational Research.* **143** 548-69.
- [91] Topkis D.M. 1998. *Supermodularity and Complementarity*. Princeton, NJ. Princeton University Press..
- [92] Urban, T.L. 1998. An inventory-theoretic approach to product assortment and shelf space allocation. *Journal of Retailing.* **74** 15-35.

- [93] van Mieghem, J.A., M. Dada. 1999. Price versus production postponement: Capacity and competition. *Management Science*. **45** 1631-1649.
- [94] van Ryzin, G., S. Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*. **45** 1496-1509.
- [95] Vives, X. 1999. *Oligopoly pricing: old ideas and new tools*. Cambridge, MA. The MIT Press.
- [96] Wall Street Journal. 1998. The Wall Street Journal Interactive Edition – March 9, 1998.
- [97] Weng, Z.K. 1995. Channel coordination and quantity discounts. *Management Science*. **41** 1509-1522.
- [98] Yuskavage and Strassner 2003. Gross domestic product by industry for 2002. *Survey of Current Business*. <http://www.bea.doc.gov/bea/dn2/iedguide.htm>.
- [99] Zheng, Y-S., 1992. On properties of stochastic inventory systems. *Management Science*. **38** 87-103.
- [100] Zipkin, P.H. 1995. Performance analysis of a multi-item production-inventory system under alternative policies. *Management Science*. **41** 690-703.

## Tables & Figures

| Category name | # Subcategories | # SKUs     |
|---------------|-----------------|------------|
| 1             | 24              | 219        |
| 2             | 18              | 58         |
| 3             | 25              | 177        |
| 17            | 16              | 260        |
| 53            | 14              | 55         |
| 55            | 9               | 55         |
| 127           | 8               | 56         |
| <b>Total</b>  | <b>114</b>      | <b>880</b> |

**Table 3.1. Summary of merchandise categories in data**

| Subcategory | Sample | % Bias | MAD |
|-------------|--------|--------|-----|
| 10          | Fit    | -3%    | 22% |
|             | Test   | -4%    | 22% |
| 30          | Fit    | -7%    | 26% |
|             | Test   | -6%    | 27% |
| 73          | Fit    | -1%    | 16% |
|             | Test   | 4%     | 17% |

**Table 3.2. Estimation quality for purchase incidence models for all stores**

| SKU    | Share within subcategory ( $p_j$ ) | MAD  | Bias |
|--------|------------------------------------|------|------|
| 527487 | 3%                                 | 101% | -67% |
| 8420   | 6%                                 | 53%  | -25% |
| 38011  | 6%                                 | 36%  | -7%  |
| 8448   | 7%                                 | 38%  | -9%  |
| 521989 | 8%                                 | 43%  | -18% |
| 35344  | 10%                                | 35%  | -11% |
| 8445   | 12%                                | 24%  | 1%   |
| 36315  | 14%                                | 19%  | 4%   |
| 35343  | 33%                                | 14%  | 6%   |

**Table 3.3. Estimation quality for the choice model for subcategory 10**

|      | Sample | Choice | logitPQ | DirectPQ |
|------|--------|--------|---------|----------|
| MAD  | Fit    | 72%    | 67%     | 72%      |
|      | Test   | 74%    | 94%     | 70%      |
| Bias | Fit    | 0%     | -43%    | 1%       |
|      | Test   | -9%    | -30%    | 2%       |

Table 3.4. Comparison of the *choice* model with other alternative models

| segments | Week = 91  |                 | Week = 38  |                 |
|----------|------------|-----------------|------------|-----------------|
|          | SR         | Error reduction | SR         | Error reduction |
| 1        | 0.3        | 0%              | 0          | 0%              |
| 2        | 0.1        | 0%              | 0.3        | 0%              |
| 3        | 1          | 1%              | 1          | 1%              |
| 4        | <b>0.9</b> | <b>34%</b>      | <b>0.8</b> | <b>30%</b>      |
| 5        | <b>0.6</b> | <b>38%</b>      | <b>0.5</b> | <b>31%</b>      |
| 6        | <b>0.9</b> | <b>33%</b>      | <b>0.7</b> | <b>24%</b>      |
| 7        | <b>1</b>   | <b>20%</b>      | <b>1</b>   | <b>19%</b>      |
| 8        | 0.1        | 0%              | 0          | 0%              |
| 9        | 0.8        | 0%              | 0.8        | 0%              |
| 10       | <b>1</b>   | <b>29%</b>      | <b>1</b>   | <b>30%</b>      |
| 11       | 1          | 5%              | 1          | 7%              |
| 12       | <b>1</b>   | <b>36%</b>      | <b>1</b>   | <b>35%</b>      |

Table 4.1. A sample of substitution estimates from different weeks in the planning period

| Iteration | Products                   | Subcategory 1 |             | Subcategory 2 |             |          |
|-----------|----------------------------|---------------|-------------|---------------|-------------|----------|
|           |                            | A             | B           | C             | D           | E        |
| 1         | Demand rate                | 10            | 5           | 9             | 4           | 4        |
|           | GP(facing1)                | <b>4.95</b>   | <b>4.12</b> | <b>4.91</b>   | <b>3.59</b> | 3.59     |
|           | GP(facing 2) – GP(facing1) | <b>3.8</b>    | 0.85        | 3.31          | 0.5         | 0.5      |
|           | <b>Facings allocation</b>  | <b>2</b>      | <b>1</b>    | <b>1</b>      | <b>1</b>    | <b>0</b> |
| 2         | Updated demand rates       | 10            | 5           | 11.8          | 5.2         | -        |
|           | GP(facing1)                | <b>4.95</b>   | <b>4.12</b> | <b>4.97</b>   | <b>4.2</b>  | -        |
|           | GP(facing 2) – GP(facing1) | 3.8           | 0.85        | <b>4.15</b>   | 0.97        | -        |
|           | <b>Facings allocation</b>  | <b>1</b>      | <b>1</b>    | <b>2</b>      | <b>1</b>    | <b>0</b> |

Table 4.2. Iterative optimization example

|                                                  | No substitution | With substitution |           |
|--------------------------------------------------|-----------------|-------------------|-----------|
|                                                  | Greedy          | Greedy            | Iterative |
| Average % gap                                    | 0.1%            | 3.0%              | 0.5%      |
| Maximum % gap                                    | 3.2%            | 24.1%             | 12.1%     |
| # instances with zero-gap<br>out of 93 instances | 85              | 14                | 40        |

**Table 4.3. Performance of the optimization method in comparison to the optimal solution**

|           |               | Shelf space |        |      |
|-----------|---------------|-------------|--------|------|
|           |               | low         | medium | high |
| Greedy    | Average % gap | 3.2%        | 1.5%   | 0.0% |
|           | Maximum % gap | 24.1%       | 12.1%  | 0.4% |
| Iterative | Average % gap | 0.5%        | 0.4%   | 0.0% |
| Heuristic | Maximum % gap | 12.1%       | 6.5%   | 0.1% |

**Table 4.4. Performance of the optimization method in comparison to the optimal solution for different levels of shelf space**

| store        | Space | Peakload Gross Profit |               |              | Weekly Gross Profit |                |             |
|--------------|-------|-----------------------|---------------|--------------|---------------------|----------------|-------------|
|              |       | Current               | Suggested     | % increase   | Current             | Suggested      | % increase  |
| 1            | 19859 | 109.6                 | 115.4         | 5.3%         | 532.0               | 550.3          | 3.4%        |
| 2            | 29617 | 393.5                 | 403.5         | 2.6%         | 1590.8              | 1611.4         | 1.3%        |
| 3            | 17795 | 83.3                  | 95.5          | 14.7%        | 353.0               | 359.8          | 1.9%        |
| 4            | 23696 | 170.5                 | 180.8         | 6.0%         | 919.6               | 940.2          | 2.2%        |
| 5            | 20847 | 121.7                 | 130.0         | 6.8%         | 437.6               | 459.6          | 5.0%        |
| 6            | 29665 | 127.7                 | 145.4         | 13.9%        | 750.4               | 801.3          | 6.8%        |
| 7            | 23709 | 108.3                 | 122.4         | 13.0%        | 490.4               | 518.7          | 5.8%        |
| 8            | 17810 | 74.1                  | 86.7          | 17.1%        | 186.2               | 243.5          | 30.8%       |
| 9            | 29701 | 114.9                 | 124.5         | 8.4%         | 728.6               | 750.4          | 3.0%        |
| 10           | 23738 | 96.0                  | 132.4         | 37.9%        | 626.8               | 724.1          | 15.5%       |
| 1            | 13290 | 151.4                 | 171.9         | 13.5%        | 734.2               | 781.1          | 6.4%        |
| 2            | 26869 | 315.5                 | 332.7         | 5.5%         | 2135.3              | 2231.1         | 4.5%        |
| 3            | 13363 | 104.8                 | 142.9         | 36.3%        | 487.3               | 588.8          | 20.8%       |
| 4            | 20099 | 244.6                 | 260.7         | 6.6%         | 1204.5              | 1257.4         | 4.4%        |
| 5            | 13402 | 210.8                 | 240.9         | 14.3%        | 734.8               | 780.4          | 6.2%        |
| 6            | 20226 | 243.9                 | 273.2         | 12.0%        | 1197.1              | 1276.5         | 6.6%        |
| 7            | 13365 | 152.3                 | 160.6         | 5.4%         | 766.3               | 785.1          | 2.5%        |
| 8            | 6759  | 50.9                  | 61.8          | 21.4%        | 223.4               | 250.4          | 12.1%       |
| 9            | 26757 | 168.8                 | 239.7         | 42.0%        | 1087.6              | 1191.0         | 9.5%        |
| 10           | 13355 | 144.7                 | 159.3         | 10.1%        | 953.4               | 995.8          | 4.4%        |
| <b>TOTAL</b> |       | <b>3187.3</b>         | <b>3580.4</b> | <b>12.3%</b> | <b>16139.4</b>      | <b>17096.7</b> | <b>5.9%</b> |

**Table 4.5. Comparison of existing and suggested assortments**

| Proportion of basket shoppers | N = 2    | N = 3        | N = 5                |
|-------------------------------|----------|--------------|----------------------|
| low                           | {.8, .2} |              | {.8, .2, 0, 0, 0}    |
|                               |          | {.8, .2, 0}  | {.5, .5, 0, 0, 0}    |
|                               |          | {.5, .5, 0}  | {.2, .8, 0, 0, 0}    |
|                               |          |              | {.6, .2, .2, 0, 0}   |
| medium                        | {.5, .5} |              | {.2, .6, .2, 0, 0}   |
|                               |          | {.2, .8, 0}  | {.2, .2, .6, 0, 0}   |
|                               |          | {.6, .2, .2} | {.2, .2, .2, .2, .2} |
| high                          | {.2, .8} | {.2, .6, .2} | {.2, 0, .6, 0, .2}   |
|                               |          | {.2, .2, .6} | {.2, 0, .2, 0, .6}   |
|                               |          | {0, .2, .8}  | {0, 0, .2, .2, .6}   |

**Table 5.1. Allocation of basket shoppers in numerical experiments**

|                                |        | % Profit Loss | % less variety |
|--------------------------------|--------|---------------|----------------|
| N                              | 2      | 13%           | 31%            |
|                                | 3      | 24%           | 43%            |
|                                | 5      | 33%           | 50%            |
| z                              | low    | 13%           | 34%            |
|                                | high   | 42%           | 56%            |
| c                              | low    | 13%           | 34%            |
|                                | high   | 42%           | 56%            |
| Ratio of<br>basket<br>shoppers | low    | 13%           | 31%            |
|                                | medium | 27%           | 47%            |
|                                | high   | 44%           | 58%            |

**Table 5.2. Summary of numerical results**

| Category Management |                   | Aggregation Heuristic |                   |
|---------------------|-------------------|-----------------------|-------------------|
| $1 - x^d/x^c$       | $1 - \Pi^d/\Pi^c$ | $1 - x^h/x^c$         | $1 - \Pi^h/\Pi^c$ |
| 29%                 | 4%                | 1%                    | 0.0%              |
| 36%                 | 8%                | 2%                    | 0.0%              |
| 60%                 | 42.9%             | -18%                  | 2.3%              |
| 38%                 | 11.4%             | -6%                   | 0.2%              |
| 48%                 | 23.7%             | 4%                    | 0.1%              |
| 31%                 | 5%                | 7%                    | 0.2%              |
| 19%                 | 2%                | 8%                    | 0.3%              |
| 39%                 | 10%               | 0%                    | 0.0%              |
| 54%                 | 34.4%             | 6%                    | 0.3%              |
| 66%                 | 59.2%             | 12%                   | 1.4%              |
| 56%                 | 36.8%             | 6%                    | 0.3%              |

**Table 5.3. Performance of the category management and aggregation heuristic in comparison to the optimal solution for a set of examples**

| Contract type   | $w_1^*$ or $\phi_1^*$ | $w_2^*$ or $\phi_2^*$ | $q_1^*$ | $q_2^*$ | Supplier Profits |           | Retailer's Profit | Total Profit |
|-----------------|-----------------------|-----------------------|---------|---------|------------------|-----------|-------------------|--------------|
|                 |                       |                       |         |         | $\pi_1^*$        | $\pi_2^*$ |                   |              |
| Wholesale price | 78.2                  | 78.2                  | 72.0    | 72.0    | 4667.4           | 4667.4    | 2440.5            | 11775.2      |
| Revenue sharing | 0.0171                | 0.0171                | 125.7   | 125.7   | 7759.7           | 7759.5    | 269.8             | 15789.0      |

**Table 6.1. Comparison of equilibrium under wholesale price and revenue sharing contracts for two symmetric suppliers selling through a centralized retailer.  $d_i \sim \text{Normal}(100,30)$ ,  $b=0.05$ ,  $p_i=100$ ,  $c_i=10$  for  $i=1,2$ .**

| Contract type   | $w_1^*$ or $\phi_1^*$ | $w_2^*$ or $\phi_2^*$ | $q_1^*$ | $q_2^*$ | Supplier Profits |           | Retailer's Profit | Total Profit |
|-----------------|-----------------------|-----------------------|---------|---------|------------------|-----------|-------------------|--------------|
|                 |                       |                       |         |         | $\pi_1^*$        | $\pi_2^*$ |                   |              |
| Wholesale price | 90.3                  | 77.5                  | 155.6   | 69.2    | 12395.9          | 4408.9    | 2570.8            | 19375.6      |
| Revenue sharing | 0.0109                | 0.0162                | 223.1   | 122.9   | 16708.5          | 7342.9    | 304.4             | 24355.8      |

**Table 6.2. Comparison of equilibrium under wholesale price and revenue sharing contracts for asymmetric suppliers selling through a centralized retailer.  $d_1 \sim \text{Normal}(200,30)$ ,  $d_2 \sim \text{Normal}(100,30)$ ,  $b=0.05$ ,  $p_i=100$ ,  $c_i=10$  for  $i=1,2$ .**

| Mean<br>(X,Y), (A,B) | Pr {X+Y>A+B} |       |
|----------------------|--------------|-------|
|                      | True         | MNL   |
| (5,5 ), (10,10)      | 0.401        | 0.395 |
| (5,10 ), (10,10)     | 0.458        | 0.456 |
| (10,10 ), (10,10)    | 0.500        | 0.500 |
| (5,20 ), (10,10)     | 0.542        | 0.544 |
| (20,20 ), (10,10)    | 0.661        | 0.669 |

**Table 8.1. Resulting choice probabilities using the sum of Gumbel random variables and its multinomial logit approximation**

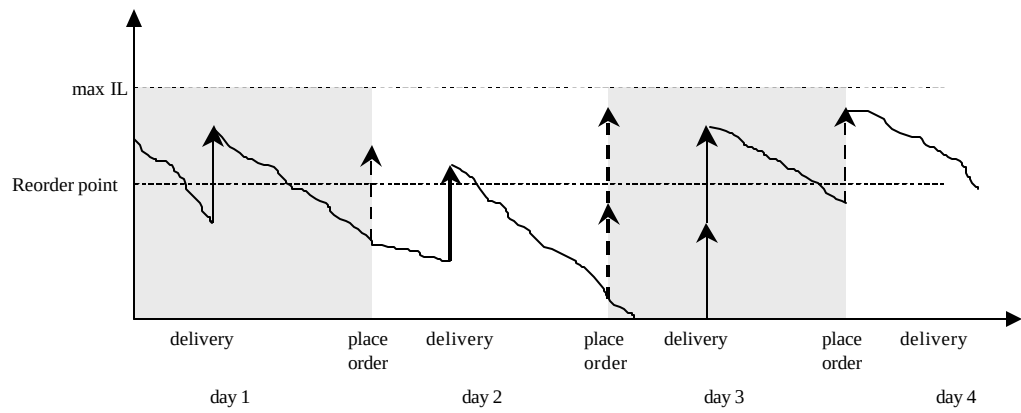


**Figure 3.1. A sample realization of inventory trajectory over time**

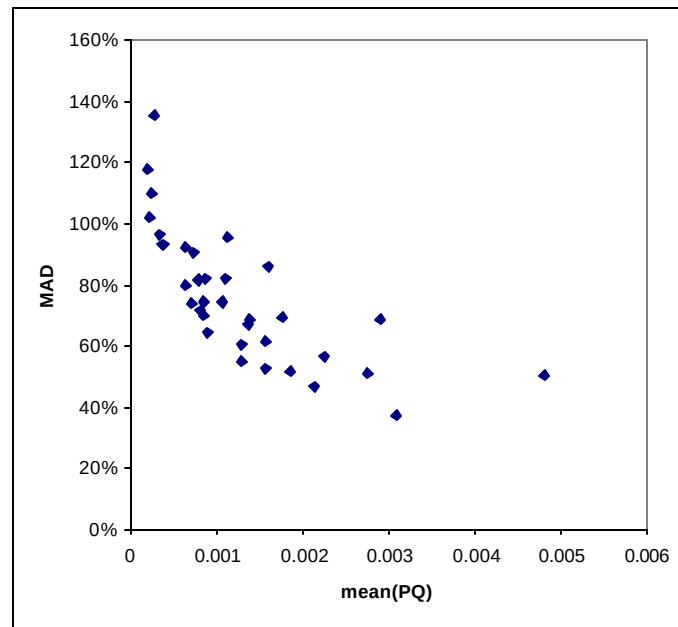


**Figure 3.2. Average of mean absolute deviation (MAD) values of day-store-SKU level PQ estimates.**

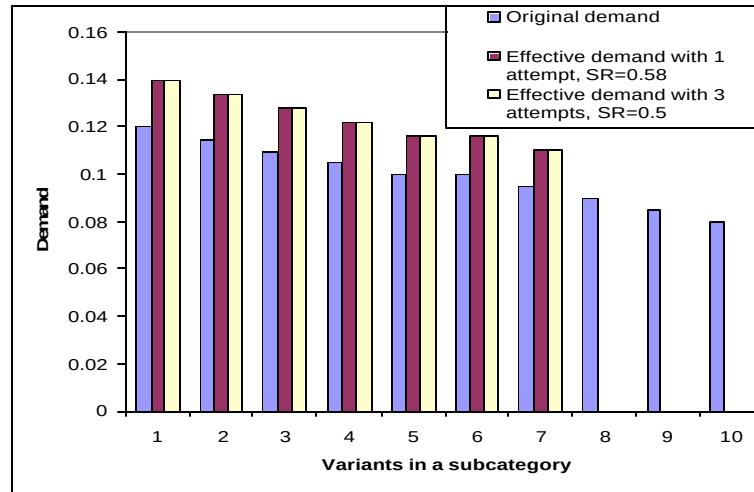


**Figure 4.1. Demand of variants under single and multiple substitution attempts when variants 8,9 and 10 are left out of the assortment**

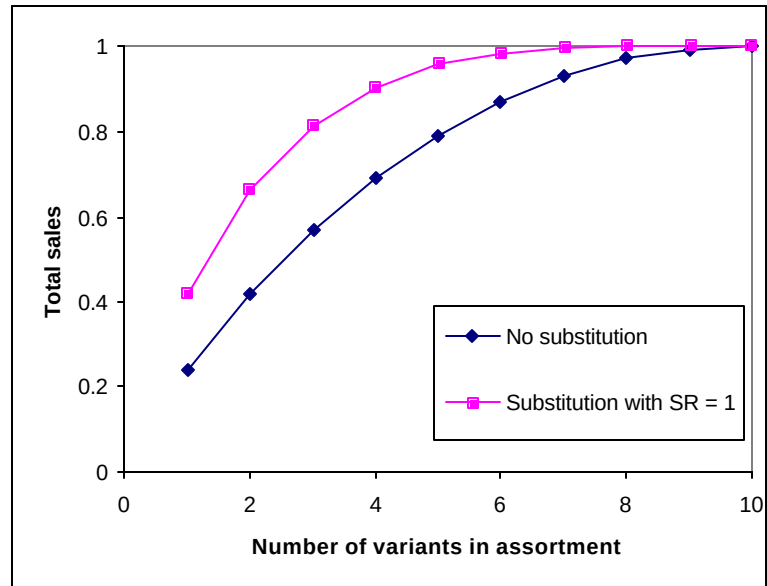


Figure 4.2. Sales vs. Number of variants

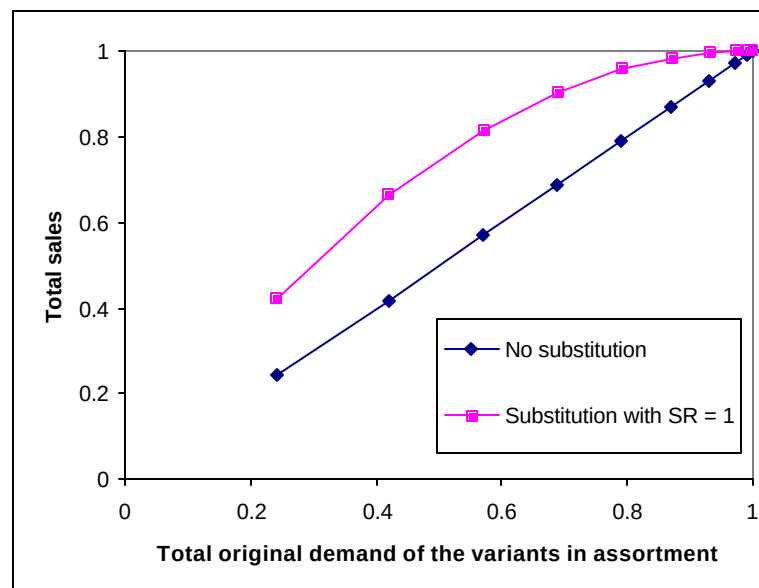


Figure 4.3. Sales vs. Total original demand

## Estimating substitution

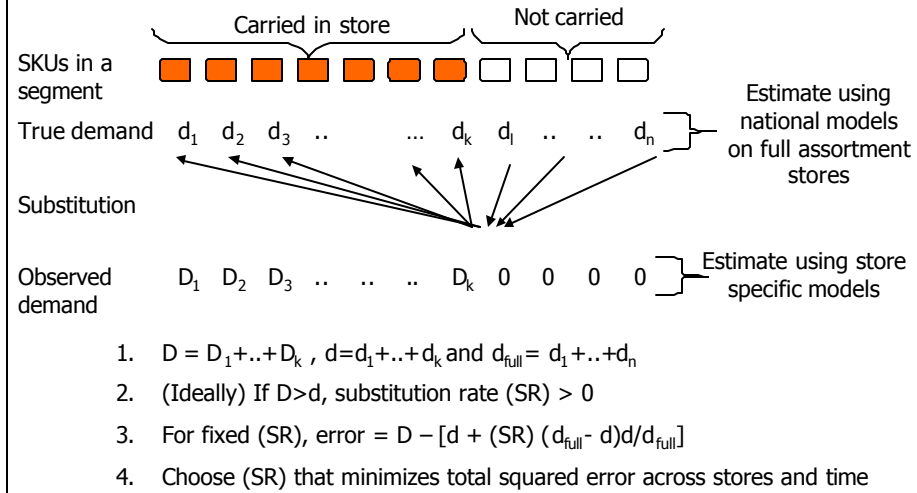


Figure 4.4. Substitution estimation scheme

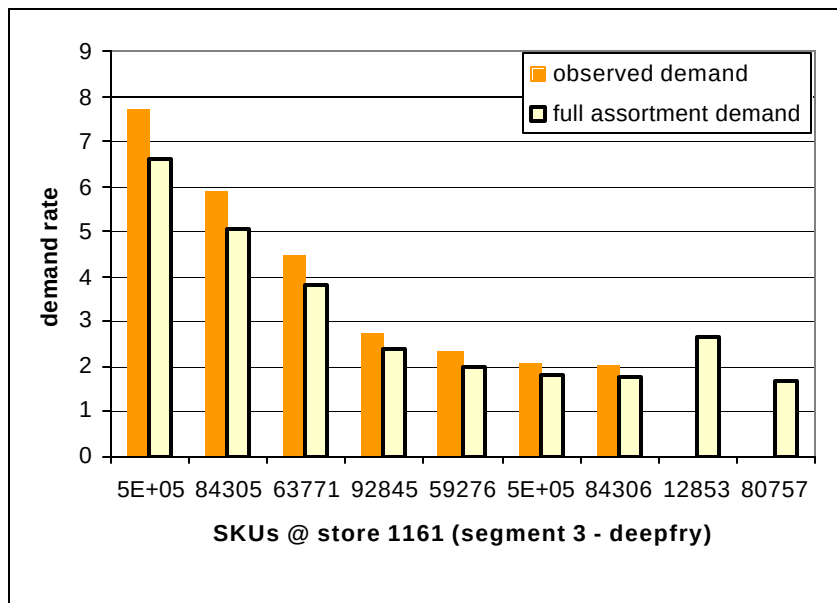


Figure 4.5. Estimates of observed demand and original demand for a subcategory

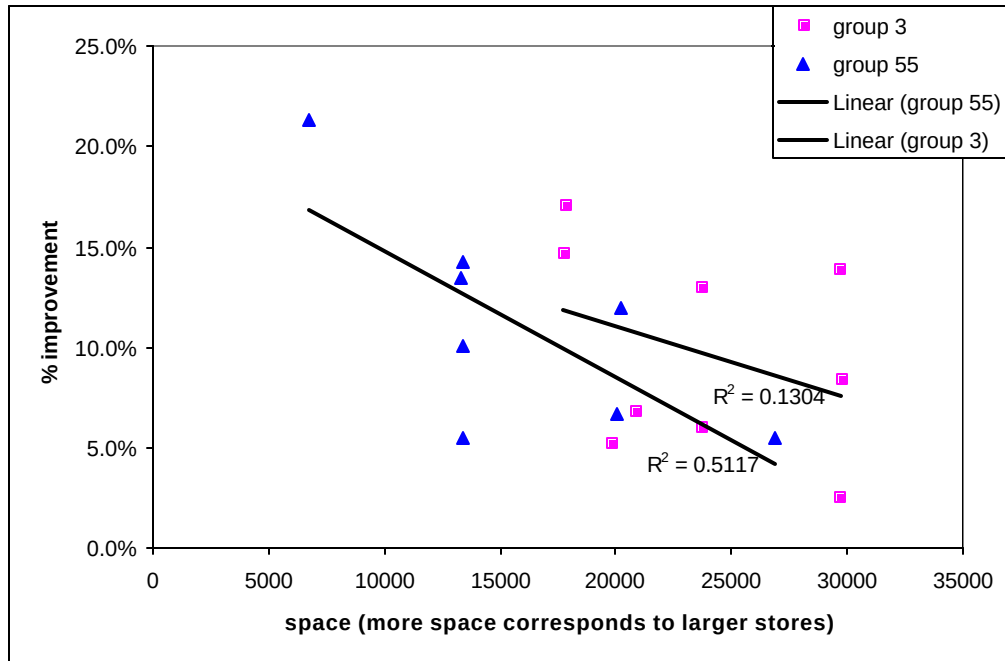


**Figure 4.6. Percentage improvement in category gross profits by the suggested assortment over the current assortments**

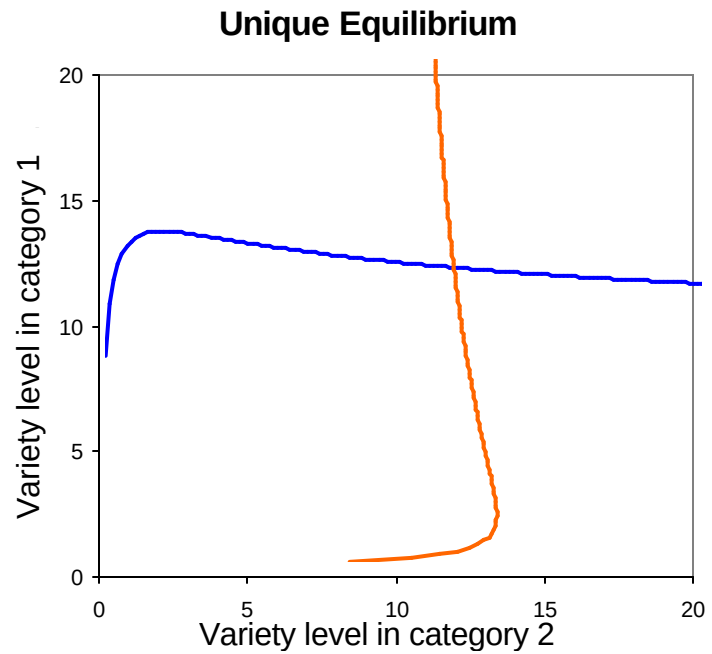


Figure 5.1. A two category example with unique category management equilibria

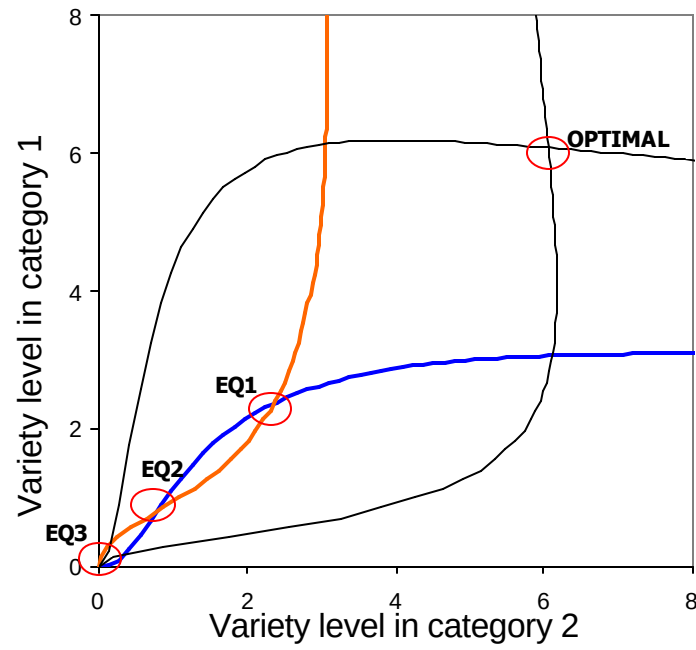


Figure 5.2. A two category example with three category management equilibria.

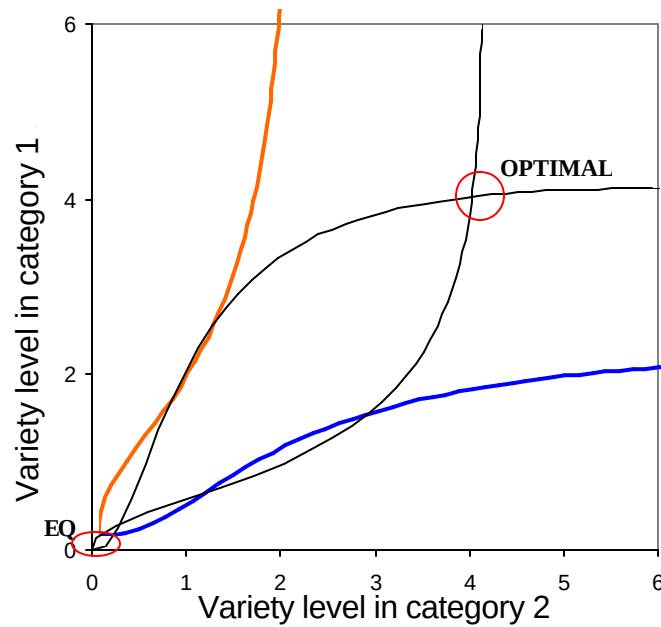


Figure 5.3. A two category example where the unique category management equilibrium is at zero

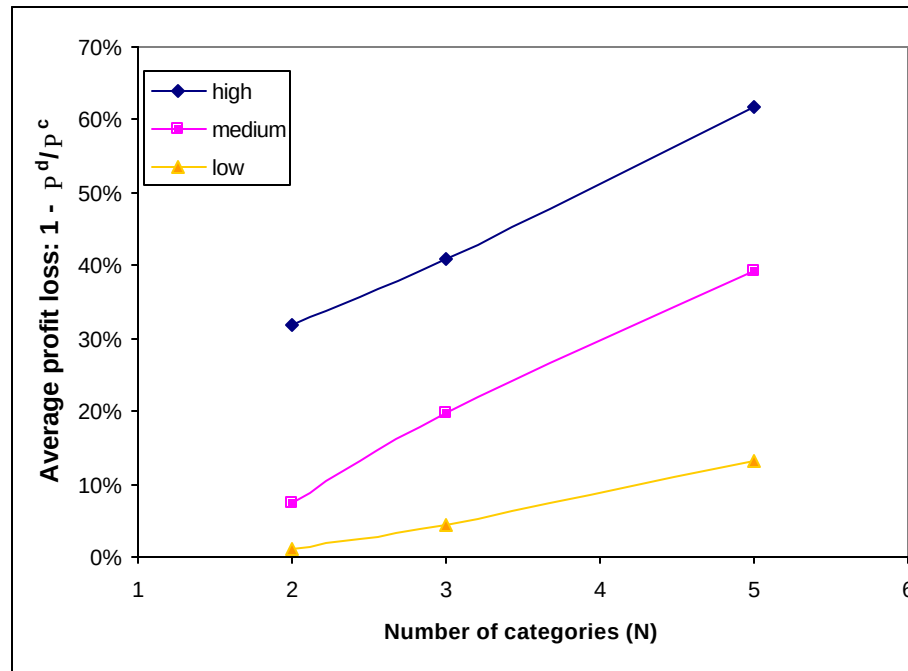


Figure 5.4. Average profit loss due to decentralization

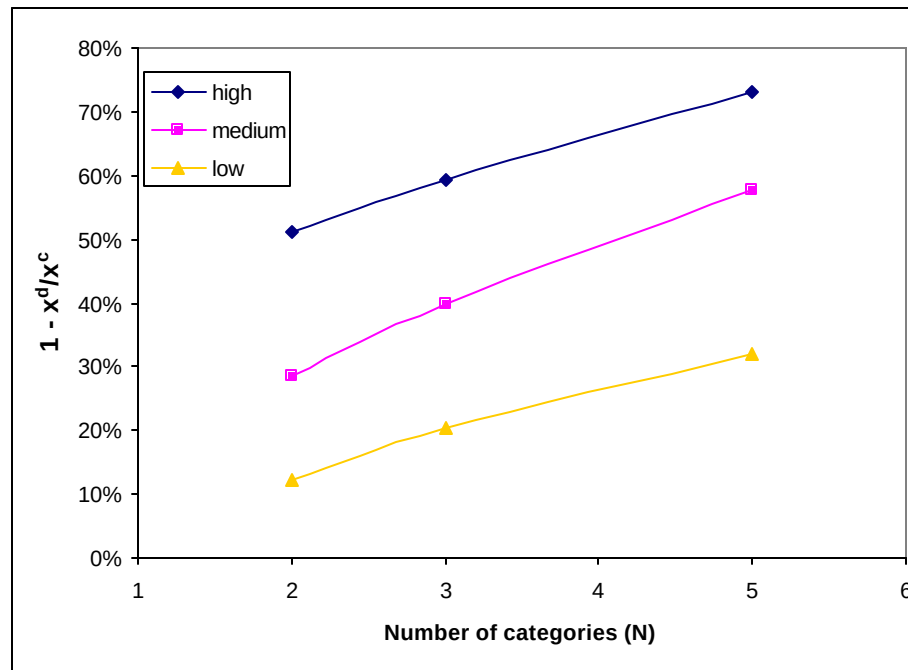


Figure 5.5. Average decrease in variety level due to decentralization

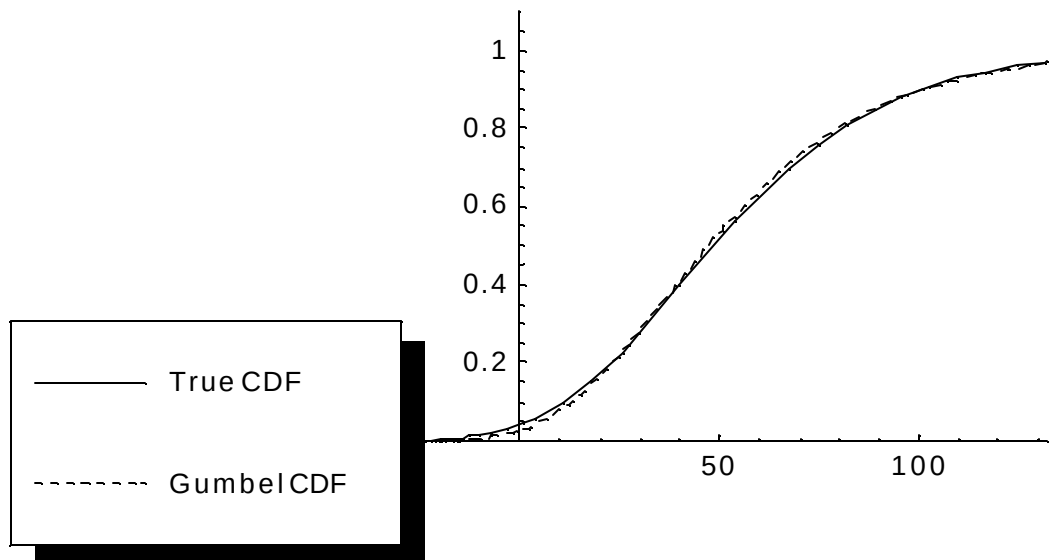


**Figure 8.1.** Cumulative density functions of the sum of Gumbel random variables and its Gumbel approximation