The Marketing Operations Interface in Consumer Retail: Theory and Practical Approach

Wei Ke

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy
under the Executive Committee of
the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2009
ABSTRACT

The Marketing Operations Interface in Consumer Retail: Theory and Practical Approach

Wei Ke

This dissertation examines four consumer retail problems at the marketing operations interface. Interdisciplinary research between marketing science and operations research has gained traction in recent years. Operations research models can help fine-tune the managerial insights that arise from consumer behavioral research; the operations research field in turn makes new inroads as models begin to account for consumer needs, preferences and attitudes.

A variety of modeling techniques, including both theoretical models and empirical implementations, are used in the context of each unique problem. Of the four topics studied in this thesis, the first two are more theoretical in nature, while the latter two are based on applied work.

The first topic investigates the incentives for coordinating product positioning and pricing of horizontally differentiated products in the context of a vertical trading partnership between a retailer and multiple suppliers. We also discuss several benchmark systems and the implications on channel efficiency.

The second topic considers an extension on the work by Popescu and Wu (2006), who formulated a dynamic pricing problem of a monopolist firm facing product demand that was sensitive to a consumer’s memory and the firm’s pricing history. Our extension separates the effect of price communications from price changes, and analyzes the trade-offs that result from this separation.

The third topic examines the impact of physical locations on the sales of a product assortment, as part of an applied research project. We study the methods to optimize the location assignment of that product assortment and to forecast the impact of a location
change.

The last topic is an applied case study that focuses on the design and deployment of a causal forecasting model at a customer service network. We are interested in verifying the effectiveness of discrete choice models on forecasting real-world demand at a relatively granular level.
# Table of Contents

1 Introduction 1

2 Coordinating Vertical Partnerships for Horizontally Differentiated Products 5
   2.1 Decentralized System ........................................... 6
      2.1.1 Retailer’s Problem ........................................ 8
      2.1.2 Suppliers’ Game and Equilibrium Analysis ................. 14
   2.2 Benchmark Systems ............................................. 19
      2.2.1 Fully Centralized System .................................. 19
      2.2.2 Monopolist Supplier ....................................... 20
   2.3 Model Comparison and Concluding Remarks ..................... 21

3 Dynamic Strategies under Sticker Shock 25
   3.1 Model Formulation ............................................. 28
   3.2 Theoretical Results ............................................ 31
      3.2.1 Stationary Pricing and Price Communication Policies .... 31
      3.2.2 Steady States and Cycles .................................. 32
   3.3 Numerical Results ............................................. 36
      3.3.1 Monotonic Trajectories under a Flat Communication Policy 38
      3.3.2 Monotonic Trajectories under Communication, and Monotonic Trajectories in the Same Direction under No Communication 40
      3.3.3 Monotonic Trajectories under Communication, and Monotonic Trajectories in the Opposite Direction under No Communication 43
      3.3.4 Cyclic Trajectories under a Step-down Communication Policy 44
3.4 Conclusions and Future Research ............................................. 45

4 Shelf Placement Optimization .................................................. 48
  4.1 The Value of Placement ..................................................... 50
    4.1.1 Plan Modification Optimization (PMO) .......................... 51
    4.1.2 Empirical Results .................................................. 53
  4.2 Forecasting the Value of Placement .................................... 57
    4.2.1 Biclustering of Locational Impact on Demand ................. 58
    4.2.2 Empirical Results ................................................ 63
    4.2.3 A Note on Data Scarcity ......................................... 73
  4.3 Conclusions and Future Research ....................................... 75

5 Case Study in Choice Forecasting ............................................ 77
  5.1 Causal Forecasting Model Formulation ............................... 79
    5.1.1 Unit Choice Probability by Region ............................ 80
    5.1.2 Regional Volume Forecast ...................................... 83
    5.1.3 Unit-Level Demand Forecast ................................... 84
  5.2 Empirical Implementation ............................................... 85
    5.2.1 Estimated Choice Parameters .................................. 86
    5.2.2 Sample Volume Forecast Results ............................... 88
    5.2.3 Sample Demand Forecast Results .............................. 89
  5.3 Conclusions and Future Research ..................................... 89

6 Conclusions ............................................................................ 93

Bibliography ............................................................................. 96

A Technical Details of Coordinating Vertical Partnerships for Horizontally Differentiated Products ................................................. 101
  A.1 Technical Details of Decentralized System .......................... 101
    A.1.1 Retailer’s Problem .................................................. 101
    A.1.2 Suppliers’ Game ..................................................... 107
List of Figures

2.1 Neo-Hotelling Model ............................................. 7
2.2 Neo-Hotelling Setup with Two Products .......................... 8
2.3 Demand Scenarios for Products in Symmetric Positions ......... 9
2.4 Demand Scenarios for Products in Asymmetric Positions ...... 11
2.5 Retailer’s Optimal Solution under Symmetric Product Positions ... 15
2.6 Retailer’s Optimal Solution under Asymmetric Product Positions ... 15

3.1 Illustrative Policy That Leads to Cyclic Trajectories .......... 35
3.2 Cyclic State Transitions ......................................... 35
3.3 Outcome Type 1 \((\alpha_0 = 0.1, \alpha_1 = 0.2, K = 0.2, m = 0.1)\) ........ ........ 39
3.4 Outcome Type 2 \((\alpha_0 = 0.1, \alpha_1 = 0.2, K = 0.2, m = 0.6)\) ........ ........ 40
3.5 Outcome Type 5 \((\alpha_0 = 0.5, \alpha_1 = 0.6, K = 0.2, m = 0.6)\) ........ ........ 41
3.6 Outcome Type 6 \((\alpha_0 = 0.5, \alpha_1 = 0.6, K = 0.2, m = 0.5)\) ........ ........ 41
3.7 Outcome Type 9 \((\alpha_0 = 0.1, \alpha_1 = 0.4, K = 0.2, m = 0.6)\) ........ ........ 42
3.8 Outcome Type 10 \((\alpha_0 = 0.1, \alpha_1 = 0.4, K = 0.2, m = 0.4)\) ........ ........ 43
3.9 Outcome Type 11 \((\alpha_0 = 0.1, \alpha_1 = 0.5, K = 0.2, m = 0.4)\) ........ ........ 43
3.10 Outcome Type 11 Special Case \((\alpha_0 = 0.1, \alpha_1 = 0.7, K = 0.4, m = 0.5)\) ... 44
3.11 Outcome Type 12 \((\alpha_0 = 0.1, \alpha_1 = 0.5, K = 0.2, m = 0.5)\) ........ ........ 45

4.1 Revenue Zero-sum Game ......................................... 50
4.2 Submove Identification Algorithm ................................ 52
4.3 Placement Move as a Network ................................... 53
4.4 PMO Algorithm ................................................... 54
List of Tables

2.1 Summary of Results for the Retailer’s Problem in Decentralized System .... 16
2.2 Comparison of Channel Profits (C - centralized system, D - decentralized system, MS - monopolist supplier model, CD - cost difference, PD - positioning distortion, DM - double marginalization) ........................................ 22
2.3 Numerical Results ($v = 1$, $c_1 = 0.4$, $c_2 = 0.6$, $(#,#)=$(channel, retailer) except for the centralized system) ................................................ 23

3.1 Summary of All Numerical Outcomes ............................................. 37
3.2 Details of Pricing Policies for Each Outcome Type .............................. 38

4.1 Lagrangian System for Updating Standard Response .......................... 62
4.2 Ex. 1 - Fractional Response Matrix .............................................. 63
4.3 Ex. 1 Step 0a - Initial Standard Response Matrix ............................... 65
4.4 Ex. 1 Step 0b - Sorted Standard Response Matrix ............................. 65
4.5 Ex. 1 Step 1a - Standard Response Matrix After 1st Split .................. 65
4.6 Ex. 1 Step 1b - Updated Standard Response Matrix Given 1st Split ....... 66
4.7 Ex. 1 Step 2a - Standard Response Matrix After 2nd Split .................. 66
4.8 Ex. 1 Step 2b - Updated Standard Response Matrix Given 2nd Split ....... 66
4.9 Ex. 1 - Forecasted Fractional Response Matrix .................................... 67
4.10 Ex. 2 - Performance of Standard Response by Cluster ........................ 68
4.11 Ex. 2 - Fractional Response Matrix .............................................. 69
4.12 Ex. 2 - Final Standard Response Matrix ........................................ 70
4.13 Ex. 2 - Clustering Scheme .......................................................... 71
4.14 Ex. 2 - Forecasted Fractional Response Matrix ......................... 72

5.1 Estimated Choice Parameters ........................................... 87
5.2 Measures of Model Accuracy by Size Group .......................... 91

A.1 Location of Optimal Solutions for All Regions ...................... 111
A.2 Summary of Results for Retailer’s Problem in Fully Centralized System ... 115

C.1 Ex. 1 Step 0a - Lagrangian System to Initialize Standard Response ...... 128
C.2 Ex. 1 Step 1b - Lagrangian System to Update Standard Response Given 1st Split ........................................................................... 129
C.3 Ex. 1 Step 2b - Lagrangian System to Update Standard Response Given 2nd Split ........................................................................... 130
Acknowledgments

A doctoral dissertation holds far more than the culmination of years of graduate study and research—its completion reflects in many ways the generosity, encouragement and inspiration from those I have met since the beginning of my graduate work.

I am indebted to my advisor Garrett van Ryzin for his generosity, accessibility, flexibility and guidance over my research interests and career aspirations. I have sought to emulate his marvelous ability of developing intuition from an otherwise complex theoretical problem—a quality that will surely benefit my life after school. I am also thankful for the many doors that he has opened for me in these years, whether they be research or career opportunities.

I am grateful to committee member Sergei Savin for agreeing to serve on my proposal committee at moment's notice. I also appreciate the time and attention he has given me in our many chats about my future plans, as well as the opportunities to work on many of his enterprising projects.

I am thankful to committee member Oded Koenigsberg for showing a constant interest in and offering helpful hints to my research work over these years, and to committee members Guillermo Gallego and Soulaymane Kachani for their thoughtful criticism and time in their busy schedules.

I thank John Quillinan and Matthew Maron for making those chapters on applied research a tangible reality, and thank Elizabeth Chang and Daniel Spacher at the PhD office for keeping me on track of the logistical requirements of this dissertation.

To my parents Weihua and Guoqing, I am grateful for the love, help, patience and understanding that a son could ever ask for. And to my extended “American family,” I thank Rochelle, Eric, Anzia, Lincoln, Chris, Lina and Barbara for providing me a place that I can call home over the last 11 years while my parents are half a world away.

And finally to my dearest friend and companion Jia, I thank her for standing by my
side in this long and arduous process, and I shall endeavor to do the same as she starts her own graduate study at Columbia.
Chapter 1

Introduction

Marketing and operations are two important and closely related functions within a firm. Marketing enables the firm to study the behaviors of its customers in order to create demand for its products. The demand thus created is fulfilled through operations.

The importance of research on the marketing operations interface has long been understood and recognized by operations researchers and marketing scientists alike. Through the accounts of Anshen (1956), Montgomery (2001) and Morrison and Raju (2004), we can get a good sense of history in this area of research.

Academic journals, such as Management Science and Journal of Operations Management among others, have published special issues dedicated to research on the marketing operations interface. Berry et al. (1991) consider empirical research papers that link the strategic planning and alignment efforts of marketing and operations. Market analyses are shown to provide valuable insights into the firm’s operational decisions. Malhotra and Sharma (2002) revisit some of those issues in a more contemporary setting and broaden the scope to include issues “pertaining to organizational design, new product development, distribution, and supply chain management among others.” Ho and Tang (2004) survey latest approaches, methodologies and insights about the integration of marketing and supply chain to improve competitiveness and profit.

Technological progress in the digital age has also helped advance research in this area. The Internet, argues Chayet et al. (2004), speeds up production and marketing operations, allows better communications with the customers and enhances the data collection capabil-
CHAPTER 1. INTRODUCTION

ities. All of these technologies present opportunities that will enhance the overall customer experience in a “customer contact chain,” from product development, to its sales, to its delivery and finally to its post-sales service.

In practice, marketing principles and operational philosophies often find themselves in conflicts with each other (see, for example, Shapiro 1977, Porteus and Whang 1991). Typically speaking, revenue maximization is marketing’s responsibility, while cost minimization is operations’ concern. Thus, marketing sets the pricing and advertising policies, which incite market reactions in the form of product demand; operations then produce the quantities demanded at the minimum cost. This decoupling will likely lead to a suboptimal firm-wide result. A better solution is to develop joint marketing-operations policies, which can correct the undesirable suboptimality. These “coordinated decision making models” are discussed by Eliashberg and Steinberg (1993).

Continuing in the same spirit, Karmarkar (1996) argues that the integration between marketing and operations occur in three stages: 1) integration by recognizing cross functional interactions on time and quality, where marketing provides operations with constraints, requirements and prices and operations in turn provide marketing with marginal cost and average cost; 2) integration via joint decision making, so that marketing benefits and operational costs are aligned; 3) complete integration of marketing and operational functions.

Modeling techniques and philosophies have also been evolving since the inception of this area of research. In his letter to the Editor of Operations Research Quarterly, Lazer (1964) provides a short critical assessment on applying operations research models to marketing—his words are still of value today. He cautions against developing models that assume away such critical marketing factors as demand, consumer behavior, competitive strategies and reactions. He also points out that inappropriate scoping or abstraction of a real world marketing problem can lead to models that have only “potential applications” rather than “actual applications.”

Four decades later, Lazer (1964)’s views are still echoed in the development of the latest generation of models for the marketing operations interface. Chopra et al. (2004) suggest ample opportunities for models that are “more comprehensive and have greater fidelity
than the current state of the art." The behavioral aspect is key to many questions at this interface, thereby providing opportunities for more empirical research in this area.

Overview

Through a series of essays, this dissertation examines four topics on consumer retail, an area where ample opportunities exist for the close integration of marketing and operations. A variety of modeling techniques, ranging from stylized models to empirical implementations, are sampled and employed for each unique problem. The first two topics (Chapters 2 and 3) are more theoretical in nature, whereas the latter two (Chapters 4 and 5) are based on applied work. We introduce each topic in turn.

The first topic (Chapter 2) investigates the incentives for coordinating product positioning and pricing of horizontally differentiated products in the context of a vertical trading relationship between a retailer and multiple suppliers. The models allow us to analyze various trading relationships and characterize the channel efficiency of each one, leading to interesting insights about when such practices may be most effective.

The second topic (Chapter 3) considers an extension on the work by Popescu and Wu (2006), who formulated a dynamic pricing problem of a monopolist firm facing product demand that was sensitive to a consumer's memory and the firm's pricing history. They showed that consumers should perceive monotonic prices in the long run under the optimal pricing policy. Our extension separates the effect of price communications from price changes. Communicating a retail price that is lower than the current reference price may attract demand in the current period, but comes at the expense of reducing the reference price. Conversely, communicating a retail price that is higher than the current reference price reduces demand in the current period, but this creates a higher reference price that is beneficial in the long run. We investigate the conditions under which the pricing path converges to a steady state or ends up in cycles.

The third topic (Chapter 4) examines the impact of physical locations on the sales of a product assortment, as part of an applied research project. Specifically, we propose a heuristic algorithm, called "Plan Modification Optimization," to solve the $NP$-hard
placement optimization problem that is also difficult to fully specify due to the presence of arbitrary business rules. We also consider a biclustering methodology that correlates the locational sensitivity of an SKU with the inherent financial impact of a location, thereby providing the foundation for forecasting the financial consequence of a move that was not historically observed.

The last topic (Chapter 5) is an applied case study that deals with the design and deployment of a causal forecasting prototype at a customer service network, e.g. a hotel chain with different affiliated hotels that target different customer segments, or a restaurant group with different member restaurants that showcases different food styles. Here we use a time series model coupled with a discrete choice model to describe and forecast unit-level demand at a pilot facility. Through this experience, we seek to test the real-world effectiveness of consumer choice forecasting, and to provide a first step toward a pricing and inventory optimization system in the future.
Chapter 2

Coordinating Vertical Partnerships for Horizontally Differentiated Products

Category management, a common retail practice, involves a vertical partnership between a leading supplier (the category captain) and a retailer to increase profit of an entire product category by coordinating assortment, promotion and pricing decisions (see Federal Trade Commission 2001, Harris and McPartland 1993, McLaughlin 1994, Nielsen Marketing Research 1992, Verra 1998). While putting category decisions in the hands of a category captain can improve coordination, the market power it creates may lead to channel inefficiency, because the category captain has incentives to bias decisions in favor of his own brands (see Steiner 2001, Basuroy et al. 2001, Kurtuluş and Toktay 2004, Gruen and Shah 2000, Zenor 1994). How best to balance these two effects is an important problem in retailing practice.

Using stylized models, we investigate the incentives for coordinating product positioning and pricing of horizontally differentiated products\(^1\) in the context of a vertical trading relationship between a monopolist retailer and multiple suppliers. The models allow us to analyze various trading relationships and characterize the channel efficiency of each one, leading to interesting insights about when such practices may be most effective.

In particular, we first consider a two echelon decentralized system with two competing

---

\(^1\)Products are horizontally differentiated when they are different in features or attributes that cannot be ordered, such as color, style or taste. In contrast, vertically differentiated products can be ordered through a customer's perceived difference in quality.
suppliers and a monopolist retailer. The suppliers each offers a horizontally differentiated product defined by a price-attribute bundle. The retailer sets a retail price which directly affects the demand. Horizontal differentiation of product attributes is described through the neo-Hotelling model, introduced by Salop (1979). Readers are referred to Lancaster (1990) and Anderson et al. (1992, Ch. 8) for discussions on locational competition and discriminatory pricing models.

Due to the presence of a downstream retailer, we find that suppliers do not always have the incentive to maximally differentiate, in contrast to the direct selling model of Salop (1979). Instead, there are cost regimes where the suppliers have partial or complete indifference to differentiate horizontally.

We subsequently study the fully centralized system and the monopolist supplier model based on an appropriate modification of the decentralized system. When comparing channel profits of the three models, we observe that supply chain inefficiency has multiple causes. When production costs are sufficiently high, channel inefficiency results from double marginalization (see Spengler 1950 for a classic definition). When production costs are sufficiently low, the inefficiency is caused by positioning distortion, under which the retailer's profit is maximized when the two products maximally differentiate. Moreover, cost difference also plays a role in supply chain inefficiency.

In a typical category management partnership, the category captain often has some form of profit sharing agreement with the retailer in exchange for the gain in market power. We provide our take on this subject via a numerical study of a further twist on the monopolist supplier model where the retailer imposes a profit target, which protects some of the market power given away to the monopolist supplier.

2.1 Decentralized System

We first consider a two echelon decentralized system involving two competing suppliers that supply a monopolist retailer. Each supplier offers a single product. The suppliers first simultaneously set product positions and then wholesale prices. The retailer takes the product positions and wholesale prices as given and sets a retail price for each product.
Each player in the setting maximizes his own profit. The objective is to characterize an equilibrium for the optimal (product position, wholesale price, retail price) triple.

The retail prices directly affect the product demands via the neo-Hotelling model, typically represented with a circle of unit circumference. Consumers are assumed to be uniformly distributed on that circle. Product demands are realized wherever purchase utility, a function of both retail price and distance to the product position, is positive. An illustration of the neo-Hotelling model is provided in Figure 2.1.

![Figure 2.1: Neo-Hotelling Model](image)

We would like to note that the modeled results would probably be different if we used the finite line topology, known as the classic Hotelling model, instead of the circle, or infinite line topology, known as the neo-Hotelling model for this work. This is based on the observation that the equilibrium for the two-competitor finite line model is achieved with both competitors relocating to the mid-point, whereas the equilibrium for the two-competitor infinite line model, described in Salop (1979), is achieved with both competitors maximally differentiating their positions. The resulting difference between the classic Hotelling and the neo-Hotelling demand models is the subject of future investigation.

For product $i$, let $l_i$ be its position (see Figure 2.2). In a multi-product Hotelling setup, the distance between the two products, expressed as $|l_1 - l_2|$, is used to characterize the differentiation of product features. Let $w_i$ be its wholesale price, $c_i$ be its production cost,
CHAPTER 2. COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

Let $p_i$ be its retail price, and $D_i$ be its demand. In addition, let $v$ be consumer valuation$^2$, and $\theta$ be linear position cost to the consumer.

If we assume full rationality, the decentralized system can be solved via backward induction. First we analyze the retailer's problem by solving for $(p_1^*, p_2^*)$. Then we determine $(w_1^*, w_2^*)$ in the suppliers' game. And finally we find the optimal differentiation in product positions $|l_1 - l_2|$.

Figure 2.2: Neo-Hotelling Setup with Two Products

2.1.1 Retailer's Problem

Taking product positions and wholesale prices as given, the retailer solves a joint profit maximization problem to obtain the optimal retail prices. In order to determine the product demands in the objective function, we distinguish symmetric versus asymmetric positioning of the two products.

2.1.1.1 Products in Symmetric Positions

When the two products are positioned symmetrically ($|l_1 - l_2| = \frac{1}{2}$), it is possible for the demands to be (S1) separable, (S2) barely touching$^3$, or (S3) overlapping on both ends.

$^2$Horizontal differentiation implies that $v$ stays the same for both products.

$^3$Symmetrically positioned product demands barely touch when they extend the entire market without overlapping.
An illustration is provided in Figure 2.3. Let $l$ be the position of a consumer and the corresponding utility of purchasing product $i$ be $v - \theta |l_i - l| - p_i$.

![Figure 2.3: Demand Scenarios for Products in Symmetric Positions](image)

When product demands are separable, a consumer makes a purchase decision of product $i$ independent of all other products. Thus he buys product $i$ if

$$v - \theta |l_i - l| - p_i \geq 0.$$ 

Let $l^*$ be the borderline consumer position between buying and not buying. Then,

$$|l_i - l^*| = \frac{v - p_i}{\theta}.$$ 

We consider scenarios (S1) and (S2) together, i.e.

$$|l_1 - l^*| + |l_2 - l^*| = \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \leq \frac{1}{2},$$

where strict inequality describes (S1) and equality describes (S2). By symmetry of the circle, product demands under (S1) or (S2) can be written as

$$D_i = 2|l_i - l^*| = \frac{2(v - p_i)}{\theta}.$$ 

The retailer solves the following problem:

$$\text{(P1)} \quad \max_{p_1, p_2} \quad (p_1 - w_1) \frac{2(v - p_1)}{\theta} + (p_2 - w_2) \frac{2(v - p_2)}{\theta}$$

s.t. \quad $\frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \leq \frac{1}{2}$. 
CHAPTER 2. COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

Now we consider scenarios (S2) and (S3) together, i.e.
\[
\frac{v-p_1}{\theta} + \frac{v-p_2}{\theta} \geq \frac{1}{2},
\]
where strict inequality describes (S3) and equality again describes (S2). When product demands overlap, a consumer has to decide on purchasing which of the two products. He will buy product \( i \) if for \( j \neq i \),
\[
v - \theta|l_i - l| - p_i \geq v - \theta|l_j - l| - p_j.
\]

Let \( l^* \) denote the position where a consumer is indifferent between the two products (represented with a straight line that cuts across the overlapped demands of scenario (S3) in Figure 2.3). We can solve the following system to obtain the demands under (S2) or (S3):
\[
\begin{align*}
&v - \theta|l_1 - l'| - p_1 = v - \theta|l_2 - l'| - p_2 \\
&|l_1 - l'| + |l_2 - l'| = \frac{1}{2}
\end{align*}
\Rightarrow D_1 = 2|l_1 - l'| = \frac{1}{2} + \frac{p_2 - p_1}{\theta}; \\
D_2 = 2|l_2 - l'| = \frac{1}{2} + \frac{p_1 - p_2}{\theta}.
\]

The retailer solves the following problem:
\[
(P2) \quad \max_{p_1, p_2} (p_1 - w_1) \left( \frac{1}{2} + \frac{p_2 - p_1}{\theta} \right) + (p_2 - w_2) \left( \frac{1}{2} + \frac{p_1 - p_2}{\theta} \right)
\]
\[\text{s.t.} \quad \frac{v-p_1}{\theta} + \frac{v-p_2}{\theta} \geq \frac{1}{2}.
\]

2.1.1.2 Products in Asymmetric Positions

Assume \( |l_1 - l_2| \) refers to the shorter arc between \( l_1 \) and \( l_2 \), and \( 1 - |l_1 - l_2| \) refers to the longer arc. When the products are located asymmetrically \((0 < |l_1 - l_2| < \frac{1}{2})\), it is possible for product demands to be (AS1) separable, (AS2) barely touching on one end but not touching on the other end, (AS3) overlapping on one end but not touching on the other end, (AS4) overlapping on one end but barely touching on the other end, or (AS5) overlapping on both ends. An illustration is provided in Figure 2.4.

\[^4|l_1 - l_2| = 0 \text{ corresponds to cannibalization of one product by another. We will discuss later in the suppliers' game that this scenario in general does not lead to an equilibrium solution.}\]
CHAPTER 2. COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

Figure 2.4: Demand Scenarios for Products in Asymmetric Positions

We consider (AS1) and (AS2) together, i.e.

\[
\frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \leq |l_1 - l_2|,
\]

where inequality describes (AS1) and equality describes (AS2). The retailer solves the following problem, which has the same objective function as in (P1):

\[
(P3) \quad \max_{p_1, p_2} \left( p_1 - w_1 \right) \frac{2(v - p_1)}{\theta} + \left( p_2 - w_2 \right) \frac{2(v - p_2)}{\theta}
\]

s.t. \quad \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \leq |l_1 - l_2|.

(AS2), (AS3) and (AS4), considered together, can be described by the following:

\[
|l_1 - l_2| \leq \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \leq 1 - |l_1 - l_2|,
\]

where an active lower bound describes (AS2), an active upper bound describes (AS4), and all the interior points describe (AS3). The demand for product \( i \) on the longer arc between \( l_1 \) and \( l_2 \) is separable from that of product \( j \) (\( j \neq i \)), while on the shorter arc it overlaps with that of product \( j \). Let \( l^* \) denote the position where a consumer is indifferent between the two products on the shorter arc between \( l_1 \) and \( l_2 \) (represented with a straight line that cuts across the overlapped demands of scenarios (AS3), (AS4) and (AS5) in Figure 2.4). In order
to find the demand on the overlapping side, we solve the following system:

\[
\begin{align*}
\begin{cases}
v - \theta |l_1 - l'| - p_1 = v - \theta |l_2 - l'| - p_2 \\
|l_1 - l'| + |l_2 - l'| = |l_1 - l_2|
\end{cases}
\Rightarrow |l_1 - l'| = \frac{1}{2} |l_1 - l_2| + \frac{p_2 - p_1}{2\theta};
\end{align*}
\]

Demand for product \(i\) under (AS2), (AS3) or (AS4) consists of two pieces. The first piece measures from the product position \(l_i\) to the point where a consumer is indifferent between buying and not buying. The second piece measures from \(l_i\) to the point where a consumer is indifferent between buying product \(i\) and product \(j\). Therefore,

\[
D_1 = \frac{v - p_1}{\theta} + \left(\frac{1}{2} |l_1 - l_2| + \frac{p_2 - p_1}{2\theta}\right) = \frac{1}{2} |l_1 - l_2| + \frac{p_2 - 3p_1}{2\theta} + \frac{v}{\theta};
\]

\[
D_2 = \frac{v - p_2}{\theta} + \left(\frac{1}{2} |l_1 - l_2| + \frac{p_1 - p_2}{2\theta}\right) = \frac{1}{2} |l_1 - l_2| + \frac{p_1 - 3p_2}{2\theta} + \frac{v}{\theta}.
\]

The retailer solves the following problem:

(P4) \[\max_{p_1, p_2} (p_1 - w_1) \left(\frac{1}{2} |l_1 - l_2| + \frac{p_2 - 3p_1}{2\theta} + \frac{v}{\theta}\right) + (p_2 - w_2) \left(\frac{1}{2} |l_1 - l_2| + \frac{p_1 - 3p_2}{2\theta} + \frac{v}{\theta}\right)\]

s.t. \(|l_1 - l_2| \leq \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \leq 1 - |l_1 - l_2|\).

Now we consider (AS4) and (AS5) together, i.e.

\[
\frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \geq 1 - |l_1 - l_2|
\]

where strict inequality describes (AS5) and equality describes (AS4). Let \(l'_1\) be the position on the shorter arc between \(l_1\) and \(l_2\) where a customer is indifferent between buying either of the two products, and let \(l'_2\) be similarly defined as the product indifference point on the longer arc between \(l_1\) and \(l_2\). We can solve the following systems to obtain the product
demands:
\[
\begin{align*}
I_1 &= \left\{ \begin{array}{l}
v - \theta |l_1 - l_1^*| - p_1 = v - \theta |l_2 - l_2^*| - p_2 \\
|l_1 - l_1^*| + |l_2 - l_2^*| = |l_1 - l_2| 
\end{array} \right. \\
&\Rightarrow |l_1 - l_1^*| = \frac{1}{2}|l_1 - l_2| + \frac{p_2 - p_1}{2\theta}; |l_2 - l_2^*| = \frac{1}{2}|l_1 - l_2| + \frac{p_1 - p_2}{2\theta}; \\

I_2 &= \left\{ \begin{array}{l}
v - \theta |l_1 - l_2^*| - p_1 = v - \theta |l_2 - l_2^*| - p_2 \\
|l_1 - l_2^*| + |l_2 - l_2^*| = 1 - |l_1 - l_2| 
\end{array} \right. \\
&\Rightarrow |l_1 - l_2^*| = \frac{1}{2}(1 - |l_1 - l_2|) + \frac{p_2 - p_1}{2\theta}; |l_2 - l_2^*| = \frac{1}{2}(1 - |l_1 - l_2|) + \frac{p_1 - p_2}{2\theta}.
\end{align*}
\]

The demand for each product under (AS4) or (AS5) can be written as:
\[
D_1 = |l_1 - l_1^*| + |l_1 - l_2^*| = \frac{1}{2} + \frac{p_2 - p_1}{\theta}; D_2 = |l_2 - l_1^*| + |l_2 - l_2^*| = \frac{1}{2} + \frac{p_1 - p_2}{\theta}.
\]

The retailer solves the following problem:

\[
(P5) \max_{p_1, p_2} \left( p_1 - w_1 \left( \frac{1}{2} + \frac{p_2 - p_1}{\theta} \right) + (p_2 - w_2) \left( \frac{1}{2} + \frac{p_1 - p_2}{\theta} \right) \right) \\
\text{s.t. } \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} \geq 1 - |l_1 - l_2|.
\]

### 2.1.1.3 Characterization of Retailer’s Optimal Response

The solutions to (P1)-(P5) (see Appendix A.1.1) are used to construct the retailer’s optimal response to a given set of wholesale prices \((w_1, w_2)\) and relative product position \(|l_1 - l_2|\).

We first define the quantity \(v - \frac{w_1 + w_2}{2}\) as the value added over average wholesale cost. The results, as shown in the following proposition, are fully specified by comparing this quantity against a function of the linear position cost \(\theta\).

**Proposition 2.1** When product positions are symmetric, the retailer’s optimal solution can be segmented as the following:

- **Region 1** \((v - \frac{w_1 + w_2}{2} < \frac{\theta}{2})\) - Separable solution is optimal;
- **Region 2** \((v - \frac{w_1 + w_2}{2} > \frac{\theta}{2})\) - Barely touching solution is optimal.

When product positions are asymmetric, the retailer’s optimal solution can be segmented as the following:
• Region 3 \( (v - \frac{w_1 + w_2}{2} < \theta |l_1 - l_2|) \) - Separable solution is optimal;

• Region 4 \( (\theta |l_1 - l_2| < v - \frac{w_1 + w_2}{2} < \frac{3}{2} \theta |l_1 - l_2|) \) - The solution where one end barely touches and the other end does not touch is optimal;

• Region 5 \( (\frac{3}{2} \theta |l_1 - l_2| < v - \frac{w_1 + w_2}{2} < \theta - \frac{1}{2} \theta |l_1 - l_2|) \) - The solution where one end overlaps and the other end does not touch is optimal;

• Region 6 \( (v - \frac{w_1 + w_2}{2} > \theta - \frac{1}{2} \theta |l_1 - l_2|) \) - The solution where one end overlaps and the other end barely touches is optimal.

Proof: See Appendix A.1.1.

Below are illustrations of the two product position scenarios discussed in Proposition 2.1. Figure 2.5 corresponds to the symmetric case, whereas Figure 2.6 corresponds to the asymmetric case. In both figures, the axis \( v - \frac{w_1 + w_2}{2} \) characterizes the segments. The retailer’s optimal solution in each of those segments is illustrated with the corresponding neo-Hotelling demand plot.

Table 2.1 below summarizes the solution to each of the 6 regions discussed in Proposition 2.1.

### 2.1.2 Suppliers' Game and Equilibrium Analysis

We consider the “position-then-price” game for the suppliers. First, we note that cannibalization cannot be an equilibrium solution as long as it is not the only option, because the cannibalized firm can strictly increase its profit by complying with a range of wholesale prices that ensure both products are on the market.

We build a 2-stage game (product position then wholesale price) and find the equilibrium for the suppliers. Taking the optimal product positions and wholesale prices as given, the retailer will then be able to determine his optimal retail price. Via backward induction

---

5 Simultaneous selection of position and price does not lead to an equilibrium (see Anderson et al. 1992, Section 8.3).

6 For example, product \( i \) takes the same position as product \( j \) with a lower wholesale price, which induces the retailer to drop product \( j \) from the assortment.
CHAPTER 2. COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

Figure 2.5: Retailer’s Optimal Solution under Symmetric Product Positions

Figure 2.6: Retailer’s Optimal Solution under Asymmetric Product Positions
<table>
<thead>
<tr>
<th>Regions</th>
<th>$p_1^*$</th>
<th>$p_2^*$</th>
<th>$\pi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 3</td>
<td>$\frac{u_1 - \theta}{2}$</td>
<td>$\frac{u_2 - \theta}{2}$</td>
<td>$\frac{(\theta - \theta)^2}{2} + \frac{(\theta - \theta)^2}{2}$</td>
</tr>
<tr>
<td>2</td>
<td>$v + \frac{u_1 - \theta}{2} - \frac{\theta}{4}$</td>
<td>$v + \frac{u_2 - \theta}{2} - \frac{\theta}{4}$</td>
<td>$v + \frac{(\theta - \theta)^2}{4}$ - $\frac{\theta}{2} - \frac{\theta}{4}$</td>
</tr>
<tr>
<td>4</td>
<td>$v + \frac{u_1 - \theta}{4} - \frac{\theta}{2}</td>
<td>l_1 - l_2</td>
<td>$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{u_1 - \theta}{2} + \frac{\theta}{2}</td>
<td>l_1 - l_2</td>
<td>$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{u_1 - \theta}{4} + v + \frac{\theta}{2}</td>
<td>l_1 - l_2</td>
<td>- \frac{\theta}{2}$</td>
</tr>
</tbody>
</table>

Table 2.1: Summary of Results for the Retailer's Problem in Decentralized System
and the assumption of total rationality, the suppliers know that the monopolist retailer
will behave optimally in a manner prescribed by the 6 regions in Figure 2.6. Furthermore,
demand functions for the suppliers are symmetric in each region\(^7\). This implies that we can
first solve for suppliers’ equilibrium wholesale prices in each region, and then find the best
overall region, which is subsequently used to determine the optimal product positions.

Let \(\bar{w}_j, j \neq i\) be the wholesale price of the competing product \(j\) as given to supplier \(i\).
Suppliers’ demand functions are found by plugging in results from Table 2.1. We formulate
a set of unconstrained profit maximization problems for the suppliers, and use the derived
results to discuss equilibrium outcomes.

Regions 1 and 3 (refer to Figures 2.5 and 2.6) have the same demand functions, since
both correspond to separable demand solutions.

\[
D_1 = \frac{2(v - p^*_1)}{\theta} = \frac{v - w_1}{\theta}; \quad D_2 = \frac{2(v - p^*_2)}{\theta} = \frac{v - w_2}{\theta}.
\]

Each supplier maximizes his own profit, i.e.

\[
(SP13) \quad \max_{w_1} (w_1 - c_1) \frac{v - w_1}{\theta}; \quad \max_{w_2} (w_2 - c_2) \frac{v - w_2}{\theta}.
\]

The demand functions for region 2 are

\[
D_1 = \frac{2(v - p^*_1)}{\theta} = \frac{1}{2} - \frac{w_1}{2\theta} + \frac{w_2}{2\theta}; \quad D_2 = \frac{2(v - p^*_2)}{\theta} = \frac{1}{2} - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}.
\]

The demand functions for region 6 are

\[
D_1 = \frac{1}{2} |l_1 - l_2| + \frac{p^*_2 - 3p^*_1}{2\theta} + \frac{v}{\theta} = \frac{1}{2} - \frac{w_1}{2\theta} + \frac{w_2}{2\theta}; \quad D_2 = \frac{1}{2} |l_1 - l_2| + \frac{p^*_1 - 3p^*_2}{2\theta} + \frac{v}{\theta} = \frac{1}{2} - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}.
\]

Observe that regions 2 and 6 (refer to Figures 2.5 and 2.6) share the same demand functions.

Each supplier maximizes his own profit, i.e.

\[
(SP26) \quad \max_{w_1} (w_1 - c_1) \left(\frac{1}{2} - \frac{w_1}{2\theta} + \frac{w_2}{2\theta}\right); \quad \max_{w_2} (w_2 - c_2) \left(\frac{1}{2} - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}\right).
\]

The demand functions for region 4 (refer to Figure 2.6) are

\[
D_1 = \frac{1}{2} |l_1 - l_2| + \frac{p^*_2 - 3p^*_1}{2\theta} + \frac{v}{\theta} = |l_1 - l_2| - \frac{w_1}{2\theta} + \frac{w_2}{2\theta};
\]
\[
D_2 = \frac{1}{2} |l_1 - l_2| + \frac{p^*_1 - 3p^*_2}{2\theta} + \frac{v}{\theta} = |l_1 - l_2| - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}.
\]

\(^7\)In other words, interchanging subscripts \(i\) and \(j\) in product \(i\)’s demand gives product \(j\)’s demand.
Each supplier maximizes his own profit, i.e.

\[(SP4) \max_{w_1} (w_1 - c_1) \left( |l_1 - l_2| - \frac{w_1}{2\theta} + \frac{\bar{w}_1}{2\theta} \right); \max_{w_2} (w_2 - c_2) \left( |l_1 - l_2| - \frac{w_2}{2\theta} + \frac{\bar{w}_1}{2\theta} \right).\]

The demand functions for region 5 (refer to Figure 2.6) are

\[D_1 = \frac{1}{2} |l_1 - l_2| + \frac{p^*_1 - 3p^*_1}{2\theta} + \frac{v}{\theta} = \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_2}{4\theta} - \frac{3w_1}{4\theta};\]
\[D_2 = \frac{1}{2} |l_1 - l_2| + \frac{p^*_2 - 3p^*_2}{2\theta} + \frac{v}{\theta} = \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_1}{4\theta} - \frac{3w_2}{4\theta}.\]

Each supplier maximizes his own profit, i.e.

\[(SP5) \max_{w_1} (w_1 - c_1) \left( \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_2}{4\theta} - \frac{3w_1}{4\theta} \right); \max_{w_2} (w_2 - c_2) \left( \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_1}{4\theta} - \frac{3w_2}{4\theta} \right).\]

The solutions to the unconstrained maximization problems (detailed in Appendix A.1.2) have some immediate properties that are summarized in the following lemmas.

**Lemma 2.1** The transition of each supplier's demand is continuous between (i) regions 1 and 2 in the symmetric positioning case, and (ii) regions 3, 4, 5 and 6 in the asymmetric positioning case.

**Proof:** See Appendix A.1.2.

**Lemma 2.2** Both suppliers have incentive to move toward maximum product differentiation for the optimal solutions of regions 4 and 5.

**Proof:** See Appendix A.1.2.

The following theorem characterizes the suppliers' optimal responses to the downstream retailer in each production cost regime. We similarly define the quantity \(v - \frac{c_1 + c_2}{2}\) as the value added over average production cost.

**Theorem 2.1** When the value added over average production cost is high \((v - \frac{c_1 + c_2}{2} \geq \frac{3}{2} \theta)\),

1. The suppliers are completely indifferent to the locations of their products, when \(v - \frac{c_1 + c_2}{2} \geq 2\theta;\)
2. The suppliers switch from having incentive to differentiate to being indifferent to locations as \(|l_1 - l_2|\) increases past the threshold \(\frac{c_1 + c_2 - 2\theta + 4\theta}{6},\) when \(\frac{3}{4} \theta \leq v - \frac{c_1 + c_2}{2} \leq 2\theta;\)
3. The suppliers have incentive to maximally differentiate their products (i.e. $|l_1 - l_2| = \frac{1}{2}$), when 
\[ \frac{3}{2} \theta \leq v - \frac{\omega_1 + \omega_2}{2} \leq \frac{7}{4} \theta. \]
In the mean time, the retailer sets prices that let product demands extend the entire market.

When the value added over average production cost lies between $\theta$ and $\frac{3}{2} \theta$, the suppliers 
maximally differentiate their products (i.e. $|l_1 - l_2| = \frac{1}{2}$), and the retailer selects prices that let 
product demands extend the entire market.

When the value added over average production cost is low ($v - \frac{\omega_1 + \omega_2}{2} < \theta$), the suppliers 
have partial incentive to differentiate in order to maintain separable product demands. In other words, 
they switch from having incentive to differentiate to being indifferent as $|l_1 - l_2|$ increases past the 
threshold $\frac{2v - c_1 - c_2}{4 \theta}$.

Proof: See Appendix A.1.2.

2.2 Benchmark Systems

In this section, we discuss a few benchmark systems, based on the variation of the de­
centralized system in the previous section. Specifically, we consider the fully centralized 
system in Section 2.2.1, which allows us to compare channel profit loss later on. We also 
consider the monopolist supplier system in Section 2.2.2, where the two products are now 
supplied by a single vendor. The comparison of the monopolist supplier system with the 
decentralized system gives additional insights into the dynamics between between the 
upstream and downstream players in a supply chain.

2.2.1 Fully Centralized System

We consider a fully centralized system where the retailer selects (by picking $|l_1 - l_2|$) and 
then sells (by determining $(p_1, p_2)$) the two products all by herself. The cost of acquiring 
product $i$ is simply the production cost $c_i$. The fully centralized system can again be solved 
via backward induction. First we analyze the retailer's problem by solving for $(p_1^*, p_2^*)$.
We observe that the retailer's problem shares the same formulations as (P1) through (P5) 
in §2.1.1 for the decentralized system with $w_i$ replaced by $c_i$. Next we find the optimal 
differentiation in product positions $|l_1 - l_2|$, which is summarized in Theorem 2.2.
Theorem 2.2 When the value added over average production cost is high \((v - \frac{c_1 + c_2}{2} \geq \frac{\theta}{2})\), the retailer positions the two products symmetrically and sets prices such that the two demands extend the entire market without overlapping (barely touching). When the value added over average production cost is low \((v - \frac{c_1 + c_2}{2} < \frac{\theta}{2})\), the retailer has partial incentive to differentiate the products in order to maintain demand separability.

Proof: See Appendix A.2.

2.2.2 Monopolist Supplier

In this model, the two products are supplied by a single firm. The supplier first selects the product differentiation parameter \(|l_1 - l_2|\) and then picks the wholesale prices \((w_1, w_2)\). The retailer’s problem is the same as in §2.1.1 for the decentralized system.

In regions 1 and 3, demands are separable. The monopolist supplier solves the following problem:

\[
\text{(SP13')} \quad \max_{w_1, w_2} \quad (w_1 - c_1) \frac{v - w_1}{\theta} + (w_2 - c_2) \frac{v - w_2}{\theta}
\]
\[
\text{s.t.} \quad w_1 + w_2 \geq \alpha,
\]

where \(\alpha = 2v - \theta\) for region 1 and \(\alpha = 2v - 2\theta|l_1 - l_2|\) for region 3.

Regions 2 and 6 have the same demand functions due to location indifference. The monopolist supplier solves the following problem:

\[
\text{(SP26')} \quad \max_{w_1, w_2} \quad (w_1 - c_1) \left(\frac{1}{2} - \frac{w_1}{2\theta} + \frac{w_2}{2\theta}\right) + (w_2 - c_2) \left(\frac{1}{2} - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}\right)
\]
\[
\text{s.t.} \quad w_1 + w_2 \leq \alpha,
\]

where \(\alpha = 2v - \theta\) for region 2 and \(\alpha = 2v - 2\theta + \theta|l_1 - l_2|\) for region 6.

The monopolist supplier solves the following problem in region 4:

\[
\text{(SP4')} \quad \max_{w_1, w_2} \quad (w_1 - c_1) \left(|l_1 - l_2| - \frac{w_1}{2\theta} + \frac{w_2}{2\theta}\right) + (w_2 - c_2) \left(|l_1 - l_2| - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}\right)
\]
\[
\text{s.t.} \quad 2v - 3\theta|l_1 - l_2| \leq w_1 + w_2 \leq 2v - 2\theta|l_1 - l_2|.
\]
CHAPTER 2. COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

The monopolist supplier solves the following problem in region 5:

\[
\begin{align*}
(w_1 - c_1) \left( \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_1}{4\theta} - \frac{3w_1}{4\theta} \right) \\
+ (w_2 - c_2) \left( \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_2}{4\theta} - \frac{3w_2}{4\theta} \right)
\end{align*}
\]

s.t. \(2v - 2\theta + \theta |l_1 - l_2| \leq w_1 + w_2 \leq 2v - 3\theta |l_1 - l_2|.

The optimal solution is summarized in the following theorem.

**Theorem 2.3** When the value added over average production cost is high \((v - \frac{\zeta + \xi}{2} > \theta)\), the supplier positions the two products symmetrically, and the retailer sets prices that let product demands extend the entire market. When the value added over average production cost is low \((v - \frac{\zeta + \xi}{2} < \theta)\), the retailer has partial incentive to differentiate the products in order to maintain demand separability.

**Proof:** See Appendix A.2.

2.2.2.1 Monopolist Supplier with Retailer Profit Target

It is possible for the monopolist supplier to have too much market power, which encroaches upon the profit of the downstream retailer. To remedy that situation, we discuss a slight modification on the plain vanilla monopolist supplier system by introducing a guaranteed profit target for the retailer. Closed form solutions in this case become algebraically tedious. Instead, we present a numerical case in the next section.

2.3 Model Comparison and Concluding Remarks

We compare the channel profits of the centralized, decentralized and monopolist supplier systems in the following remark, verifiable through numerical experiments:

**Remark 2.1** Loss of channel profit occurs when the supply chain moves away from a centralized system.

- When the value added over average production cost is high \((v - \frac{\zeta + \xi}{2} \geq \theta)\), channel profit of the monopolist supplier system dominates that of the decentralized system.
CHAPTER 2. COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

- When the value added over average production cost is low \((v - \frac{\alpha + \omega}{2} < \frac{\theta}{2})\), the monopolist supplier system achieves the same channel profit as that of the decentralized system.

- When the value added over average production cost lies between \(\frac{\theta}{2}\) and \(\theta\), it is unclear whether the monopolist supplier system or the decentralized system incurs greater loss of channel profit.

Table 2.2 gives a quick summary of Remark 2.1 and attributes the causes of channel profit loss in each production cost regime. Between the monopolist supplier system and the decentralized system, we see that the former tends to dominate the latter when the value added over average production cost is high \((v - \frac{\alpha + \omega}{2} \geq \theta)\). The loss of channel profit in this case is mainly attributable to the difference in the production costs of the offered products. However, the benefit of a monopolist supplier system becomes questionable when the value added over average production cost falls below \(\theta\). In that case, double marginalization becomes the main cause of channel profit loss.

| \(v - \frac{\alpha + \omega}{2} \in \) | \((0, \frac{\theta}{2})\) | \([\frac{\theta}{2}, \theta]\) | \([\theta, \frac{3\theta}{2}]\) | \([\frac{3\theta}{2}, \infty)\) |
| Channel profit | \(C \geq MS > D\) | \(C \geq MS, C \geq D\) | \(C \geq MS > D\) | \(C \geq MS > D\) |
| Cause | DM | CD, DM | CD | PD, CD |

Table 2.2: Comparison of Channel Profits (C - centralized system, D - decentralized system, MS - monopolist supplier model, CD - cost difference, PD - positioning distortion, DM - double marginalization)

Table 2.3 summarizes the result of a numerical experiment \((v = 1, c_1 = 0.4, c_2 = 0.6)\). The product differentiation cost parameter \(\theta\) is varied to reflect the 4 cases discussed in Table 2.2, in the same order that was first presented. Except for the centralized system, where channel profit equals retailer profit, we present a pair of profit numbers in the parentheses for each system, with the first number corresponding to channel profit and second number corresponding to retailer profit. We also characterize the optimal regions (refer to Figures 2.5 and 2.6 for an illustration) that result in an equilibrium for all parties involved in each system.
<table>
<thead>
<tr>
<th>$\theta$</th>
<th>0.6</th>
<th>0.4</th>
<th>0.3</th>
<th>0.26</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v = \frac{1}{2}$ $\epsilon$</td>
<td>$(0, \frac{v}{2})$</td>
<td>$[\frac{v}{2}, \theta)$</td>
<td>$[\theta, \frac{v}{2})$</td>
<td>$[\frac{v}{2}, \infty)$</td>
</tr>
<tr>
<td>Centralized profit</td>
<td>0.367</td>
<td>0.425</td>
<td>0.458</td>
<td>0.473</td>
</tr>
<tr>
<td>Optimal region</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Decentralized profit (best)</td>
<td>(0.325,0.108)</td>
<td>(0.422,0.111)</td>
<td>(0.444,0.129)</td>
<td>(0.456,0.179)</td>
</tr>
<tr>
<td>As % of centralized</td>
<td>(88.6%,29.4%)</td>
<td>(99.3%,26.1%)</td>
<td>(96.9%,28.2%)</td>
<td>(96.4%,37.8%)</td>
</tr>
<tr>
<td>Optimal region</td>
<td>1 or 3</td>
<td>boundary of 1, 2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Decentralized profit (worst)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>(0.412,0.135)</td>
</tr>
<tr>
<td>As % of centralized</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>(87.1%,28.5%)</td>
</tr>
<tr>
<td>Optimal region</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Monopolist supplier profit</td>
<td>(0.325,0.108)</td>
<td>(0.419,0.106)</td>
<td>(0.45,0.083)</td>
<td>(0.464,0.075)</td>
</tr>
<tr>
<td>As % of centralized</td>
<td>(88.6%,29.4%)</td>
<td>(98.6%,24.9%)</td>
<td>(98.3%,18.1%)</td>
<td>(98.1%,15.9%)</td>
</tr>
<tr>
<td>Optimal region</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Monopolist supplier profit (retailer target)</td>
<td>(0.367,0.25)</td>
<td>(0.425,0.25)</td>
<td>(0.458,0.25)</td>
<td>(0.473,0.15)</td>
</tr>
<tr>
<td>As % of centralized</td>
<td>(100.0%,68.1%)</td>
<td>(100.0%,58.8%)</td>
<td>(100.0%,54.6%)</td>
<td>(100.0%,31.7%)</td>
</tr>
<tr>
<td>Optimal region</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.3: Numerical Results ($v = 1$, $c_1 = 0.4$, $c_2 = 0.6$, (#,#)=(channel, retailer) except for the centralized system)
Comparing retailer profit of the decentralized system with that of the monopolist supplier model, we see that the retailer is less profitable under the latter than the former, especially when the value added over average production cost is high \( (v - \frac{\alpha + \omega}{2} \geq \theta) \). In other words, the retailer suffers when market power is concentrated into the hands of the monopolist supplier.

However, the use of retailer profit target (as seen from the last row of Table 2.3) limits the market power of the monopolist supplier, thereby increasing channel efficiency. In this particular example, the channel achieves zero loss.
Chapter 3

Dynamic Strategies under Sticker Shock

The use of internal reference price is an important demand modeling concept. The internal reference price\(^1\) is defined as a price that is recallable from a customer’s memory and against which the offered price of a product or service is compared. Under the reference effect, a customer may buy the product when his reference price is greater than or equal to the product’s sticker price, otherwise he may walk away. Winer (1986) defined the “sticker shock” effect as a consumer’s reaction to the discrepancy between his internal reference price of a frequently purchased product and its retail price. This effect is typically captured as a component of the product demand function, which, in the simplest mathematical terms, is a function of the difference between the reference price and the current retail price. Setting a retail price that is lower than the current reference price may attract demand in the current period, but this comes at the expense of reducing the reference price. Conversely, setting a retail price that is higher than the current reference price reduces demand in the current period, but this creates a higher reference price that is beneficial in the long run. Popescu and Wu (2006) analyzed this effect in a recent paper by formulating the dynamic pricing problem of a monopolist firm facing product demand.

\(^1\)Cf. External reference price comes from the price of a competitor product, which is beyond the scope of this research.
that was sensitive to a consumer's memory and the firm's pricing history. They proved that consumers should perceive monotonic prices in the long run under the optimal pricing policy.

We consider an extension on Popescu and Wu (2006) by separating the effect of price communications from price changes. This relaxes the implicit assumption that consumers are instantly informed of any changes in retail prices. Communicating the current retail price quickly updates the reference price and potentially increases foot traffic to the store, but incurs a cost in return; conversely, not communicating the current retail price slows down reference price updating, but diminishes potential reference effects on demand. Our model analyzes this trade-off. Typical means of price communication includes catalogs, commercials and price stickers. Firms routinely communicate promotional prices as well as regular prices.

There are four relevant streams of literature for this area of research. The first stream, represented by Coase (1972), Stokey (1981), Conlisk et al. (1984), Sobel (1984), deals with inter-temporal price discrimination. Coase (1972) discusses a monopolist's exploitation of residual demand versus a customer's retaliation from delaying his purchase, and concludes that the selling strategy is "to keep the market saturated at all dates." Stokey (1981) extends Coase (1972)'s work and examines a monopolist's pricing strategy as affected by the length of the trading cycle. She finds that an equilibrium solution depends on the length of the trading cycle. Conlisk et al. (1984) discuss the cyclic high-low pricing strategy for a monopolist, and finds that the seller will "charge a price just low enough to sell immediately to consumer with a high willingness to pay," but periodically he will "drop the price far enough to sell to an accumulated group of consumers with a low reservation price." Sobel (1984) discusses a similar high-low pricing strategy for the competitive market, where "the sellers typically vary their prices over time, charging a high price in most periods, but occasionally cutting the price to sell to a large group of customers with a low reservation price."

CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

earlier. Tversky and Kahneman (1991) extend their Prospect Theory (See Kahneman and Tversky 1979) by including riskless choices, and introduces the much famed assumption that "losses and disadvantages have greater impact on preferences than gains and advantages." Kalyanaram and Winer (1995) propose three empirical generalizations of consumers' use of reference price in making brand choices. They find that "consumers rely on past prices as part of the reference price formation process" and verify Tversky and Kahneman (1991)'s finding that consumers are more sensitive to "losses" than "gains."

The third stream of research, represented by Vaidyanathan and Muehling (1999), Anderson and Simester (2004), examines the timing of pricing decisions. Vaidyanathan and Muehling (1999) discuss an empirical model on promotion management by focusing on "the internal reference price construct." Anderson and Simester (2004) examine consumers' perception of price fairness for durable goods sold through direct-mail catalog campaigns. They find that "deeper price discounts in the current period increased future purchases by first-time customers (a positive long-run effect) but reduced future purchases by established customers (a negative long-run effect)."

The last stream of literature uses a similar set of dual-control dynamic models, but tackles an entirely different issue, namely illicit drug enforcement. Most of the models are based on continuous time optimal control, instead of the discrete-time model that we will be describing. Nonetheless, the parallels between the issues facing a retail firm and those facing the drug enforcement agency are quite interesting. Behrens et al. (2000) investigate the optimal prevention and treatment decisions for drug use and its related problems. They find that prevention is most appropriate when the epidemic (i.e., the state function) involves mostly light users, while treatment is most preferred when the epidemic involves mostly heavy users; at the same time, there is a brief transition period where both prevention and treatment are required. Tragler et al. (2001) considers the optimal treatment and enforcement decisions for drug control. They find it optimal to use both treatment and enforcement, if the drug problem is at an early stage, otherwise it is optimal to moderate the growth of the epidemic by allocating more sources to treatment instead of enforcement.

In this work, we study the steady state properties of the infinite-horizon discrete-time dynamic pricing and price communication problem. There appear to be four types of price
communication policies:

1. Do not communicate the current retail price at any reference price;

2. Communicate the current retail price at all reference prices;

3. Do not communicate the current retail price at low reference prices, but communicate at high reference prices;

4. Communicate the current retail price at low reference prices, but do not communicate at high reference prices.

We have numerical evidence that suggests the existence of cyclic state and action trajectories—including retail prices, reference prices and price communication actions—especially under communication policy type 4. When the existing reference price falls below a threshold, the firm imposes and communicates a series of retail prices that are higher than the resulting reference prices, in order to quickly build up the reference price level. Once the reference price threshold is reached or surpassed, however, the firm should switch to but not communicate a series of retail prices that are lower than the resulting reference prices, thereby accelerating the demand. In the cases absent of trajectory cycles, the retail and reference prices will converge to a steady state price.

3.1 Model Formulation

The model formulation relies on a similar set of assumptions laid out by Popescu and Wu (2006)—those relevant assumptions will be re-stated verbatim or paraphrased here. The demand shall likewise be a function of both retail price $p$ and reference price $r$, i.e. $D(p, r)$.

**Assumption 3.1** (a) Demand $D(p, r)$ is non-negative bounded and continuous. Furthermore, it decreases in $p$ and increases in $r$. (b) $D(p, r) = D(p) + R(r - p, r)$, where $D(p)$ captures the intrinsic demand not affected by the reference effect and $R(r - p, r)$ captures the reference effect itself. (c) The reference effect $R(x, r)$ is increasing in $x$, and $R(r - p, r)$ is increasing in $r$. Furthermore, $R(x, r) \geq 0$ for $x > 0$, $R(x, r) \leq 0$ for $x < 0$, and $R(0, r) = 0$. 
Current period’s reference price $r_t$ is updated with the newly observed retail price $p_t$ to form next period’s reference price $r_{t+1}$.

**Assumption 3.2** At period $t$, reference prices are updated via the exponential smoothing formula:

$$r_{t+1} = a p_t + (1 - a) r_t,$$

where $a \in [0, 1]$.

Empirical evidence, as discussed in Winer (1986), supports the above reference price updating formula. We call the parameter $a$ in Assumption 3.2 the learning rate. The higher the learning rate (in other words, the more memorable the current retail price is), the faster the reference price is updated.

The firm controls both the retail price and the price communication action at each time period. The pricing action has the following regularity assumption:

**Assumption 3.3** For each time period $t$, the retail price $p_t$ is closed and bounded, i.e. $p_t \in [p, \bar{p}]$.

Assumptions 3.2 and 3.3 ensure that the state space (reference price) is also closed and bounded. With an initial reference price $r_0$, which may or may not be in $[p, \bar{p}]$, we observe that the state space is defined on the closed interval $\Omega = [r_0 \land p, r_0 \lor \bar{p}]$.

Price communication, as an additional control, brings about two immediate effects to the model dynamics:

1. It accelerates the learning rate: Consistent with empirical research on promotion and memory (see Vanhuele and Drèze 2002), we assume the learning rate under no communication, denoted $a_0$, is less than or equal to the learning rate under communication, denoted $a_1$.

2. It expands the market size: To reflect the impact of marketing activities on raised public awareness of a brand and increased foot traffic to retail outlets selling the brand, we denote $m \geq 0$ as the percent increase in demand due to price communication.

As a tradeoff, price communication incurs a fixed cost $K$. Fixed costs are typical in existing business practices, where firms usually decide on a fixed marketing budget for the entire fiscal year.

To denote the price communication action, we set $u_t = 1$ for periods where retail prices are being actively communicated, and set $u_t = 0$ otherwise. The binary price
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

communication action, together with the compact pricing space described in Assumption 3.3, ensures the compactness of the entire action space.

For an initial reference price $r_0$ and discount factor $\gamma \in (0,1)$, the firm’s discounted profit maximization problem can be described as:

$$-
\begin{align*}
(P) \sup_{p_t \in [p,\bar{p}], u_t \in \{0,1\}} \sum_{t=0}^{\infty} \gamma^t \Pi(p_t, u_t, r_t) \\
\text{s.t. } r_{t+1} = \alpha_u p_t + (1 - \alpha_u) r_t, \; t \geq 0
\end{align*}
$$

where

$$
\Pi(p_t, u_t, r_t) = p_t (1 + mu_t) D(p_t, r_t) - Ku_t.
$$

(3.1)

A few observations are immediate, as in Popescu and Wu (2006). By Assumptions 3.1(a) and 3.3, $\Pi(p_t, u_t, r_t)$ is bounded. And by Stokey et al. (1989), the value function is the unique bounded solution of the following Bellman equation:

$$
V(r) = \sup_{p \in [p,\bar{p}], u \in \{0,1\}} W(p, u, r) := \sup_{p \in [p,\bar{p}], u \in \{0,1\}} \Pi(p, u, r) + \gamma V(\alpha_u p + (1 - \alpha_u) r).
$$

(3.2)

Furthermore, the short term profit $\Pi(p, u, r)$ and the value function $V(r)$ are both increasing in the reference price $r$. The compactness of the action space $[p,\bar{p}] \times \{0,1\}$ and the continuity of all functions guarantee the existence of an optimal stationary pricing and price communication policy that solves Equation 3.2.

To conclude the formulation of (P), we make one additional regularity assumption on the components of the profit function. By Assumption 3.1(b) and Equation 3.1, we have the following relationship:

$$
\Pi(p, u, r) = (1 + mu)(\pi(p) + \pi^R(p, r)) - Ku,
$$

where $\pi(p) = pD(p)$ and $\pi^R(p, r) = pR(r - p, r)$.

Assumption 3.4 (a) $\pi(p)$ is non-monotonic and concave in $p$. (b) $\frac{\partial}{\partial p} \left( \pi(p) + \pi^R(p, r) \right) |_{p=r} = \pi'(r) - rR_x(0, r)$ is strictly decreasing in $r$. (c) $\pi^R(p, r)$ is concave in $p$, convex in $r$, and supermodular in $(p, r)$. (d) In addition, $\Pi(p, u, r)$ is supermodular in $(p, u, r)$. 
The equation in (b) holds as a result of the properties of $R(x, r)$ in Assumption 3.1(c). The intuition behind (b) and (c) is provided in Popescu and Wu (2006). In short, (b) states that the marginal profit diminishes as retail price approaches the reference price; (c) implies that the marginal short term profit increases with the reference price.

### 3.2 Theoretical Results

#### 3.2.1 Stationary Pricing and Price Communication Policies

Our goal in this section is to characterize the optimal stationary pricing and price communication policies, whose interactions result in various action and state trajectories. These trajectories in turn lend themselves to interesting strategies for the firm.

First we observe that the value function $V(r)$, as defined by the Bellman Equation 3.2, has the following properties:

**Lemma 3.1** $V(r)$ is increasing and convex in $r$.

**Proof:** The result holds by Assumptions 3.1(a), 3.4(c), and Smith and McCardle (2002, Proposition 5).

The properties of the optimal pricing policy are rather straightforward.

**Theorem 3.1** Under a given communication regime ($u \in \{0, 1\}$), the optimal stationary pricing policy $p_u^*(r)$ increases in the reference price $r$.

**Proof:** By Lemma 3.1 and Assumption 3.4(c), $W(p, u, r)$ in Equation 3.2 is supermodular in $(p, r)$, for a given $u$. Therefore, the result holds by Topkis (1998, Theorem 2.8.2).

Going forward, we will use $p^*(r)$ to denote the overall optimal pricing policy. Theorem 3.1, together with Assumption 3.2, implies the monotonicity of retail and reference price trajectories under each communication regime. A simple argument goes like this: If $p^*(r_i) \geq r_i$, we obtain immediately that $r_{i+1} = \alpha_u p^*(r_i) + (1 - \alpha_u) r_i \geq r_i$ and consequently $p_u^*(r_{i+1}) \geq p^*(r_i)$; the converse, when $p^*(r_i) \leq r_i$, also holds true.

The properties of the stationary price communication policy are directly related to the definition of a reference price threshold. Recall that the state space is defined on the closed interval $\Omega = [r_0 \wedge \underline{p}, r_0 \vee \overline{p}]$. 
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

Theorem 3.2 Under a given stationary pricing policy $p(r)$, if there exists an $\hat{r} \in \Omega$ that is the unique solution of

$$W(p(\hat{r}), 1, \hat{r}) = W(p(\hat{r}), 0, \hat{r}),$$

then for all $(r', r'') \in \{r' < \hat{r} < r'', r' \in \Omega, r'' \in \Omega\}$, the optimal stationary communication policy has the following relationship:

$$u^*(r'') = u^*(\hat{r}^+) 
eq u^*(\hat{r}^-) = u^*(r').$$

Proof: See Appendix B. ■

Equation 3.3, in essence, means the firm is indifferent between communicating and not communicating when the reference price is at $\hat{r}$. Theorem 3.2 suggests that there are four types of stationary price communication policies:

1. No communication policy: $u^*(r) = 0, \forall r$;
2. Constant communication policy: $u^*(r) = 1, \forall r$;
3. Step-up communication policy: $u^*(r) = 0, \forall r < \hat{r}$, and $u^*(r) = 1, \forall r \geq \hat{r}$;
4. Step-down communication policy: $u^*(r) = 1, \forall r < \hat{r}$, and $u^*(r) = 0, \forall r \geq \hat{r}$.

Here types 1 and 2 refer to the case where $\hat{r} \not\in \Omega$, and types 3 and 4 refer to the case where $\hat{r} \in \Omega$.

3.2.2 Steady States and Cycles

State and action trajectories are the direct result of the stationary policies characterized by Theorems 3.1 and 3.2. In this section, we will characterize the steady states and demonstrate the circumstances under which reference price cycles may occur. These cycles, as direct results of the threshold-type pricing and/or price communication policies, are particularly interesting, as they shed light on how firms can manage the retail prices and price communication actions of their products over fixed life cycles.

Theorem 3.3 (a) If Equation 3.2 admits a steady state $p^*$, it is the unique solution of

$$\frac{\pi'(p)}{1 - \gamma} = \frac{pR_x(0, p)}{1 - \gamma + \alpha u \gamma'}$$

(3.4)
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

where \( u^* \in \{0, 1\} \) is the stationary communication policy at \( p^* \). (b) \( u^* \) must satisfy the following condition:

\[
u^* = \begin{cases} 
0, & \text{if } \pi(p^*) < \frac{K}{m}; \\
1, & \text{if } \pi(p^*) \geq \frac{K}{m}.
\end{cases}
\]

(c) Furthermore, \( p^* \) is increasing in \( \alpha_u \) and in \( \gamma \), respectively.

**Proof:** See Appendix B.

With a small increase in price \( \Delta p \), the left hand side of Equation 3.4 implies that the long run profit without the reference effect is increased by

\[
\pi'(p)\Delta p + \gamma \pi'(p)\Delta p + \gamma^2 \pi'(p)\Delta p + \cdots = \frac{\pi'(p)}{1 - \gamma} \Delta p.
\]

At the same time, the small increase in price \( \Delta p \) results in a loss in profit due to the reference effect. This loss is increased by

\[
pRx(0, p)\Delta p + (1 - \alpha_u)\gamma pRx(0, p)\Delta p + (1 - \alpha_u)^2 \gamma^2 pRx(0, p)\Delta p + \cdots = \frac{pRx(0, p)}{1 - \gamma + \alpha_u \gamma}.
\]

At steady state, these two quantities must be equal.

A pricing and price communication scheme that converges to a steady state is generally more suitable for commodity products—e.g., consumer packaged goods, groceries, and so forth—that have constant supply and demand over a long life cycle and, at the same time, do not have to rely on frequent innovation to compete in the market place. For example, a gallon of milk has a relatively stable price and does not require much price communication or product innovation to renew its demand.

For durable products that benefit from rapid innovation (e.g., automobiles, consumer electronics, etc.), however, a cyclic pricing and price communication strategy that accompanies the introduction of a new product generation and the discontinuation of an old one may be beneficial. The following theorem illustrates the conditions under which cycles may occur and the properties that describe those cycles. We first define some notations:

\[
q(r) := \alpha_u r(p^*) + (1 - \alpha_u r) r,
\]

\[
q'(r) := q(q(...q(r)...)) = 
\]

\[
to \]
Theorem 3.4 (a) For the threshold \( \bar{r} \) defined in Theorem 3.2, action and state trajectories become cyclic under the following stationary pricing and price communication policies: \( u^*(r) = 1, p^*(r) > r, \forall r < \bar{r}; u^*(r) = 0, p^*(r) < r, \forall r \geq \bar{r} \). (b) The resulting reference price cycles are defined on \([r, \bar{r}]\).

The bounds \( r \) and \( \bar{r} \) are determined by the following system:

\[
\begin{align*}
\bar{r} &= q^m(r), \\
n &= \inf\{i \in \mathbb{Z}^+ : q^i(r) \geq \bar{r}\}, \\
m &= \inf\{i \in \mathbb{Z}^+ : q^i(r) > \bar{r}\}, \\
n &= \inf\{i \in \mathbb{Z}^+ : q^i(r) < \bar{r}\}.
\end{align*}
\]

This implies the following cycle states:

\[
\begin{align*}
\bar{r} &= q^n(r) < q^2(r) < \ldots < q^{n-1}(r) < q^n(r) < q^{n-2}(r) < \ldots < q^1(r) < \bar{r} = q^m(r)
\end{align*}
\]

Proof: See Appendix B.

Put simply, Theorem 3.4 says that the firm has incentives to set and communicate a retail price that is higher than the current reference price when the reference price is lower than \( \bar{r} \); conversely, it is optimal to set but not communicate a retail price that is lower than the current reference price when the reference price is higher than \( \bar{r} \). The stationary policies that produce cyclic trajectories are illustrated in Figure 3.1.

If it starts with an initial reference price \( r_0 = \bar{r} \), the system will gradually lower its reference price level to \( q^n(r) \) by communicating a retail price that is lower than the reference price at that time. The next action brings the reference price level down to \( \bar{r} = q^n(r) \) below the threshold \( \bar{r} \). At this point, it becomes necessary to charge and communicate a retail price that is higher than the current reference price, thereby building the reference price level up to \( q^{m-1}(r) \). The subsequent action then brings the reference price level all the way up to \( \bar{r} = q^m(r) \), thus completing the cycle. Similar arguments can be made with any initial reference price in \([r, \bar{r}]\). An illustration of the cyclic state transitions can be found in Figure 3.2.

As mentioned before, this cyclic pricing strategy is common for durable goods that rely on rapid innovation over a relatively short life cycle, e.g. automobiles, consumer electronics.
or personal computers. At the beginning of the cycle, a new product is introduced with a relatively high retail price and an advertising campaign. This retail price may sometimes be driven up by increased public interest, manifested as an increased reference price. As the product ages, advertising of the product is pulled and its retail price is gradually lowered to a point where the product is phased out and replaced by an upgraded product.

The conditions described in Theorem 3.4 are not the only ones leading to reference price cycles. We now describe the general conditions under which reference prices can form cycles:
Theorem 3.5 Cyclic reference price trajectories will occur if there exists a triplet \( r' < \bar{r} < r'' \), an \( m \in \mathbb{Z}^+ \) and an \( n \in \mathbb{Z}^+ \) such that \( p^*(\bar{r}) > \bar{r}, p^*(r') < \bar{r}, p^*(r') > r', p^*(r'') < r'', q^m(r') = r'', q^{m-1}(r') < \bar{r}, q^n(r'') = r', \) and \( q^{m-1}(r'') > \bar{r} \).

Proof: By construction, the reference price must cycle through at least these states \( r' = q^n(r'') < q^{m-1}(r') < q^{m-1}(r'') < r'' = q^{m}(r') \).

First we note that Theorem 3.5 applies to any of the four types of stationary communication policies specified by Theorem 3.2, as long as the corresponding pricing policies satisfy the general conditions laid out above. Unlike Inequality 3.6, the intermediate states here \( q(r'), q^2(r'), \ldots, q^{m-2}(r') \), and \( q(r''), q^2(r''), \ldots, q^{m-2}(r'') \) do not necessarily have to rank order. The conditions \( p^*(\bar{r}) > \bar{r} \) and \( p^*(\bar{r}) < \bar{r} \) represent a general form of the equation \( p^*(\bar{r}) = \bar{r} \) that allows the presence of a non-differentiable jump at \( \bar{r} \). As a special case, Theorem 3.4 suggests that \( \bar{r} = \bar{r} \). Interestingly, it is easiest to obtain cycles using the special conditions in Theorem 3.4, as opposed to the general conditions in Theorem 3.5. All cyclic cases in the numerical experiments described in the following section result from the conditions in Theorem 3.4.

3.3 Numerical Results

We illustrate the properties of various types of stationary policies using numerical experiments. The algorithm, coded in Matlab, calls the dpolve engine from the CompEcon toolbox developed by Miranda and Fackler (2002). The state space (reference price) is continuous, while the action space (retail price and price communication action) is discretized.

In all experiments, we use the following log-linear demand function

\[
D(p, r) = e^{-p} + \frac{e^{(r-p)/2} - 1}{2},
\]

together with these parameters:

- Retail price: \( p \in [0.5, 2] \) discretized to step-size 0.01;
- Discount factor: \( \gamma = 0.9; \)
- Initial reference price: \( r_0 \in [0, 0.5, 1, 1.5, 2, 2.5] \);
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

- Learning rates: $a_0 \in \{0.1, 0.2, \ldots, 0.8\}, a_1 \in \{a_0 + 0.1, a_0 + 0.2, \ldots, 0.09\}$;
- Fixed communication cost: $K \in \{0.2, 0.4, 0.6, 0.8, 1\}$;
- Market expansion factor: $m \in \{0.1, 0.2, \ldots, 0.6\}$.

A total of 6,480 cases are generated and summarized into 12 outcome types by their respective policy and trajectory characteristics, as presented in Table 3.1. Details of the

<table>
<thead>
<tr>
<th>Type</th>
<th>Policy of $u$</th>
<th>Policy of $p$</th>
<th>Trajectory of $u$</th>
<th>Trajectory of $p$</th>
<th>Trajectory of $r$</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$u'(r) = 0, V_r$</td>
<td>$p'(r) \uparrow r, V_r$</td>
<td>$u_t = 0, V_t$</td>
<td>$p_t$ monotonic, $V_t$</td>
<td>$r_t$ monotonic, $V_t$</td>
<td>5,555</td>
</tr>
<tr>
<td>2</td>
<td>$u'(r) = 1, V_r$</td>
<td>$p'(r) \uparrow r, V_r$</td>
<td>$u_t = 1, V_t$</td>
<td>$p_t$ monotonic, $V_t$</td>
<td>$r_t$ monotonic, $V_t$</td>
<td>74</td>
</tr>
<tr>
<td>3</td>
<td>$u'(r) = 0, V_r &lt; p$</td>
<td>$p'(r) \uparrow r, V_r$</td>
<td>$u_t = 0, V_t$</td>
<td>$p_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>$r_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>72</td>
</tr>
<tr>
<td>4</td>
<td>$1, V_r \geq p$</td>
<td>$p'(r) \uparrow r$ within each $u'(r)$ level</td>
<td>$u_t = 1, V_t$</td>
<td>$p_t$ monotonic, $V_t \geq \hat{r}$</td>
<td>$r_t$ monotonic, $V_t \geq \hat{r}$</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>$p'(r) \uparrow r$ within each $u'(r)$ level</td>
<td></td>
<td>$u_t = 0, V_t$</td>
<td>$p_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>$r_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>$u_t = 1, V_t &lt; \hat{r}; 0, V_t \geq \hat{r}$</td>
<td>same direction, $V_t \geq \hat{r}$</td>
<td>same direction, $V_t \geq \hat{r}$</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>$p'(r) \uparrow r$ within each $u'(r)$ level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>147</td>
</tr>
<tr>
<td>8</td>
<td>$u'(r) = 1, V_r &lt; p$</td>
<td>$p'(r) \uparrow r$ within each $u'(r)$ level</td>
<td>$u_t = 0, V_t$</td>
<td>$p_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>$r_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>$0, V_r \geq p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>10</td>
<td>$u'(r) = 1, V_r &lt; p$</td>
<td>$p'(r) \uparrow r$ within each $u'(r)$ level</td>
<td>$u_t = 0, V_t$</td>
<td>$p_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>$r_t$ monotonic, $V_t &lt; \hat{r}$</td>
<td>35</td>
</tr>
<tr>
<td>11</td>
<td>$0, V_r \geq p$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>71</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>cyclic</td>
<td>cyclic</td>
<td>cyclic</td>
<td>332</td>
</tr>
</tbody>
</table>

Table 3.1: Summary of All Numerical Outcomes

stationary pricing policy $p^*$ with respect to the reference price $r$ are summarized in Table 3.2. We use $p^* \times r$ to denote the cases where $p^*(r)$ crosses $r$, which is the 45° line, and use $p^*||r$ to denote the cases where $p^*(r)$ does not cross $r$. Note that if $p^* \times r$ and $p^*(r_0) \geq r_0$, then eventually $p^*(r_m) \leq r_m$ for some $m$; if $p^*||r$ and $p^*(r_0) \geq r_0$, then $p^*(r_m) \geq r_m$ for all $m$. Other scenarios are noted with a similar logic.

We will demonstrate each outcome in the subsections that follow. The first three subsections represent cases where action and state eventually converge to steady state. The last subsection, on the other hand, demonstrates the instances under which a cyclic pricing and price communication strategy may be optimal.
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

<table>
<thead>
<tr>
<th>Type</th>
<th>Pricing Policy under $u^* = 0$</th>
<th>Pricing Policy under $u^* = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
</tr>
<tr>
<td>3</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>4</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
</tr>
<tr>
<td>5</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>6</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>7</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>8</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
</tr>
<tr>
<td>9</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>10</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>11</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
<tr>
<td>12</td>
<td>$p^* \times r, p'(r_0) \geq r_0$</td>
<td>$p^* \times r, p'(r_0) \leq r_0$</td>
</tr>
</tbody>
</table>

Table 3.2: Details of Pricing Policies for Each Outcome Type

3.3.1 Monotonic Trajectories under a Flat Communication Policy

Outcome types 1 and 2 fall under this category. In outcome type 1, the optimal communication policy $u^*(r)$ is flat at no communication for any reference price, whereas the optimal pricing policy $p^*(r)$ is an increasing function of reference price. As a result, the communication trajectory $u_t$ is flat at no communication, while both the retail and reference price trajectories are monotonically increasing or decreasing for all $t$ and converging toward the same steady state. This is in essence the properties discussed by Popescu and Wu (2006). Figure 3.3 shows a specific instance of outcome type 1 with two different initial reference prices ($r_0 = 0, r_0 = 1$).

Outcome type 2 is similar to type 1, except that it is now optimal to always communicate the retail price at any reference price. The communication trajectory $u_t$ is correspondingly flat at always communicating. Figure 3.4 shows a specific instance of outcome type 2 with two different initial reference prices ($r_0 = 0.5, r_0 = 1$). A notable difference between types 1 and 2 is that $m = 0.6$ for outcome type 2 allows the market to be more expanded than does $m = 0.1$ for outcome type 1, as the retail price is being communicated to the customers.
Figure 3.3: Outcome Type 1 ($\alpha_0 = 0.1, \alpha_1 = 0.2, K = 0.2, m = 0.1$)
3.3.2 Monotonic Trajectories under Communication, and Monotonic Trajectories in the Same Direction under No Communication

Outcome types 5, 6, 9 and 10 fall under this category. In outcome type 5, the optimal communication policy \( u^*(r) \) is flat at no communication when the reference price is below a threshold \( \tilde{r} \); it steps up to communicating the retail prices when the reference price rises above \( \tilde{r} \). The optimal pricing policy \( p^*(r) \) is an increasing function of the reference price for each \( u^*(r) \) level. With a low initial reference price, the optimal communication trajectory starts out flat at no communication when \( t < \tilde{t} \), but switches to communicating when \( t \geq \tilde{t} \). The corresponding retail price trajectory increases during each \( u_t \) level, with the possible existence of a dip as time goes from \( \tilde{t}^- \) to \( \tilde{t}^+ \). In other words, it is possible, but not always the case, that \( p_{t^-} \leq p_{t^+} \). Figure 3.5 shows a specific instance of outcome type 5 with initial
reference price \( r_0 = 0 \).

Figure 3.5: Outcome Type 5 \((\alpha_0 = 0.5, \alpha_1 = 0.6, K = 0.2, m = 0.6)\)

With a high initial reference price and minor modifications on the input parameters, it is possible to obtain trajectories that are converse to those in outcome type 5. Outcome type 6 has a similar pricing and price communication policy to that of type 5, but it is optimal to communicate the retail prices when \( t < \bar{t} \) and stop when \( t \geq \bar{t} \). The corresponding retail price trajectory decreases during each \( u_t \) level, with the possible existence of a jump as time goes from \( \bar{t}^- \) to \( \bar{t}^+ \) (i.e. it is possible that \( p_{\bar{t}^+} \geq p_{\bar{t}^-} \)). Figure 3.6 shows a specific instance of outcome type 6 with initial reference price \( r_0 = 1.5 \) and a slightly lowered \( m = 0.5 \).

Figure 3.6: Outcome Type 6 \((\alpha_0 = 0.5, \alpha_1 = 0.6, K = 0.2, m = 0.5)\)
Outcome types 9 and 10 have a step-down communication policy, where it is optimal to communicate for reference prices below a threshold \( \hat{r} \) but stop when reference prices rise beyond \( \hat{r} \). As before, the pricing policy \( p^*(r) \) is an increasing function of reference price within each \( u^*(r) \) level. In outcome type 9, the communication trajectory starts out flat at no communication when \( t < \hat{r} \) but steps up to communicating when \( t \geq \hat{r} \). With a high initial reference price, the corresponding retail price trajectory decreases during each \( u_t \) level, with the possible existence of a jump as time goes from \( \hat{r}^- \) to \( \hat{r}^+ \) (\( p_{t^-} > p_{t^+} \)). Figure 3.7 shows a specific instance of outcome type 9 with initial reference price \( r_0 = 1.5 \).

![Optimal Trajectory (3306)](image)

Figure 3.7: Outcome Type 9 (\( \alpha_0 = 0.1, \alpha_1 = 0.4, K = 0.2, m = 0.6 \))

The trajectories for outcome type 10 are converse to those of type 9. This is achieved with a low starting reference price and slight modifications on the input parameters. For the resulting trajectories, it is optimal to communicate when \( t < \hat{r} \) but stop when \( t \geq \hat{r} \). The retail price trajectory increases during each \( u_t \) level, with the possible existence of a dip as time goes from \( \hat{r}^- \) to \( \hat{r}^+ \) (\( p_{t^-} < p_{t^+} \)). Figure 3.8 shows a specific instance of outcome type 10 with initial reference price \( r_0 = 0 \) and a slightly lowered \( m = 0.4 \).

A flat communication trajectory and a monotonic retail price trajectory may still be possible for a step-up or step-down communication policy. These special cases are illustrated in outcome types 3, 4, 7 and 8, but we will omit the details here.
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

3.3.3 Monotonic Trajectories under Communication, and Monotonic Trajectories in the Opposite Direction under No Communication

Outcome type 11 falls under this category. Only a step-down communication policy appears to lead to this case. For the resulting trajectories, it is optimal to communicate when $t < \hat{t}$ but stop when $t \geq \hat{t}$. The retail price trajectory increases when $t < \hat{t}$ but decreases when $t \geq \hat{t}$, with the possible existence of a dip as time goes from $\hat{t}^-$ to $\hat{t}^+$ ($p_{\hat{t}^-} \leq p_{\hat{t}^+}$). Figure 3.9 shows a specific instance of outcome type 11 with initial reference price $r_0 = 0$. 

Figure 3.8: Outcome Type 10 ($\alpha_0 = 0.1, \alpha_1 = 0.4, K = 0.2, m = 0.4$)

Figure 3.9: Outcome Type 11 ($\alpha_0 = 0.1, \alpha_1 = 0.5, K = 0.2, m = 0.4$)
CHAPTER 3. DYNAMIC STRATEGIES UNDER STICKER SHOCK

As a special case, it may be optimal to communicate the upper bound of the retail price \( \overline{p} \) for exactly one period at \( t = 1 \). Beyond the first period, it becomes optimal to stop communicating while simultaneously lowering the retail prices. Figure 3.10 shows a specific instance of this special case of outcome type 11 with initial reference price \( r_0 = 0 \).

![Optimal Trajectory (161)](image1)

![Optimal Stationary Policy (161)](image2)

Figure 3.10: Outcome Type 11 Special Case (\( \alpha_0 = 0.1, \alpha_1 = 0.7, K = 0.4, m = 0.5 \))

3.3.4 Cyclic Trajectories under a Step-down Communication Policy

Outcome type 12 falls under this category. Again, only a step-down communication policy appears to lead to this case. The resulting trajectories call for an alternating communication strategy, paired with a cyclic retail pricing strategy (see Theorem 3.4). It is optimal to communicate a retail price higher than the current reference price to drive up the reference price level, and harvest the higher reference price level by not communicating a retail price that is lower than the current reference price. With different initial reference prices, it may take different number of periods to for the system to stabilize into the same cyclic pattern. Figure 3.11 shows a specific instance of outcome type 12 with two different initial reference prices (\( r_0 = 0, r_0 = 1 \)). In this particular instance, it is optimal to communicate a high retail price for only one period before switching off, i.e. \( n = 1 \) in the terminology used in the system of equations (3.5).
3.4 Conclusions and Future Research

The extension on Popescu and Wu (2006) allowed us to separate the effect of price communications from price changes. Using an infinite horizon discounted model, we were able to analyze the cost and benefit of price communication. Communicating the retail price accelerates reference price updating and increases foot traffic to the store, but comes with a fixed advertising cost; on the other hand, not communicating the retail price slows down the reference price updating and saves advertising expense, but also reduces store traffic.

Our results show that in most cases, the price communication trade-off leads to a steady state, which admits a constant communication policy and a convergent pricing path, even though there may be, at the beginning of the product’s life cycle, a transient pricing...
and price communication strategy that differs from the steady state strategy. Products that benefit from a steady state pricing and price communication action generally have constant supply and demand over a long life cycle, and do not rely on rapid innovation to compete in the market place. Consumer packaged goods and groceries are good examples.

However, cycles in the pricing and communication trajectories can occur under the following circumstances: It becomes optimal for the firm to communicate a higher retail price at low reference price states, in order to shore up public interest of a newly introduced product that replaces the previous generation; conversely, the firm has incentives to set but not communicate a lower retail price at high reference price states, thereby harvesting the demand benefit of these high reference price states. These pricing and communication policies create cycles by bringing the reference price up when it is low and driving the reference price down when it is high. This cyclic tactic appears to be widely used by firms selling durable goods that rely on rapid innovation over a relatively short life cycle. Notable examples include automobile, consumer electronics or PC manufacturers.

The various action and state trajectories are the direct result of the pricing and price communication policies. We were able to summarize the communication policies into four distinct but simple types: 1) do not communicate at all; 2) communicate all the time; 3) do not communicate at low reference price states, but communicate at high reference price states; 4) communicate at low reference price states, but do not communicate at high reference price states. The cycles described previously were generated using the fourth communication policy type. The pricing policy has a very nice property: It is an increasing function of the reference price state within each communication regime. This property led directly to the steady states of retail and reference prices.

A few items for future research are immediately foreseeable:

1) We would like to explore how additional levels of communication would impact the model results. We believe that, in this case, threshold-type policies resulting in cyclic trajectories will still be found, and we expect to see a variety of communication levels that contribute to the cycle-generating thresholds.

2) It is also interesting to recast the formulation (P) as a continuous time optimal control problem. In other words, we would assume the firm has the ability to change the pricing
and price communication decisions at any given point in time, instead of the periodic
review structure adopted in the current formulation.

3) The problem (P) solved in this research is deterministic in nature. It would therefore
be interesting to see how our results will change when reference prices are not easily
observable, thereby forcing the firm to probe the reference prices modeled with a random
distribution.

4) The cyclic pricing and price communication tactics tend to backfire when the customers
start to anticipate the timing of price drops right before new product introductions. We
would like to explore how our model can accommodate this strategic customer behavior.
Chapter 4

Shelf Placement Optimization

In a typical retail environment, frequently encountered problems concerning the maximization of profit or margins of the entire assortment can usually be categorized as, but not limited to: 1) First Price Optimization, which determines what best initial price to offer for a product, 2) Markdown Price Optimization, which analyzes the timing and depth of price cuts from the first price, and 3) Assortment Optimization, which considers the best offered set of products subject to shelf space constraints. In contrast to these problems, for which extensive research literature exists, we focus our attention on placement optimization in an applied research project. The goal of this project is to create practical systems for placement optimization and to shed light on some under-researched issues in the process.

As its name suggests, placement optimization seeks to determine the profit-maximizing location—a shelf or a room—within the store where any subset of products should be positioned. In addition, it determines how much display space should be allocated to that product subset. Optimal placement decisions leverage manifested shopper behavior under different retail floor layouts.

Bryner (2005), Mendelsohn (2007), Abratt and Goodey (1990) summarize a few well-known retail shopping patterns that offer clues on where to place certain types of products, e.g. high margin items, in the store. For example, “most shoppers turn right,” suggesting that it may be advantageous to place high margin items along the right thoroughfare in the store; shoppers “stick to the perimeter of the store” and “do not weave up and down the aisle,” suggesting that products placed in the center of an aisle are less frequented; and
lastly, shoppers “have a natural tendency to focus and perceive at eye level.”

The impact of shelf space allocation on product sales has also been well studied in consumer behavior literature. Cox (1970) stated and verified that 1) there is no relationship between allocated shelf space and sales for a “staple product” brand; and that 2) there is relationship between allocated shelf space and sales for an “impulse product” brand. Later research by Kollat and Willett (1967, 1969), Curhan (1972), Bellenger et al. (1978), Wilkinson et al. (1982), Abratt and Goodey (1990) echoes Cox (1970)'s findings.

It appears that efficient store-level solutions do not yet exist in current literature on shelf placement. This is possibly because finding an optimal solution often requires exponential time. In addition, business rules and constraints are hard to fully specify. As a result, most placement decisions today rely on the experience of a planner.

Working with a sponsor that manages over 300 retail outlets at a major resort destination, we attempt to provide a system-level model that characterizes the aggregate assortment profit as a function of product locations, among other variables. This in turn provides the capability to experiment and determine optimal placement locations. The model also provides a strategic retail manager with an added benefit. Placement changes may achieve the same demand effect as price changes (e.g. relocating a product to a high traffic area, or conversely, the back of the store). This is particularly relevant, when the retail manager does not want a customer’s internal reference price (see Winer (1986) for a definition) adjusted inadvertently due to an otherwise unavoidable price change, thus helping to preserve the associated brand equity.

In the following sections, we first introduce the “plan modification optimization,” a heuristic algorithm that bypasses the computational complexity and modeling difficulties of the full placement optimization problem. The algorithm decomposes a set of moves proposed by store planners into feasible independent components and recommends the implementation of those components that contribute positively to revenue gain.

Next, we introduce a forecasting framework to complement the optimization algorithm. In reality, there are usually not enough natural experimentation that allows historically observed pairings between an SKU and any fixture location in a store—enumerations of all feasible pairings are astronomically prohibitive. Noting that 1) SKUs respond to their
placement locations in varying degrees and 2) prime real estate, such as high foot-traffic areas near the store entrance, generally lifts SKU sales, we propose a biclustering algorithm, operating on both the SKU- and the fixture-dimensions, to correlate the locational sensitivity of an SKU with the inherent financial impact of a location. This correlation provides a good estimate for forecasting the financial consequence of a move that was not observed historically.

4.1 The Value of Placement

The first question we must ask ourselves is whether placement is a zero-sum game from the revenue perspective. For space utilization, placement change appears to be a zero-sum game, because existing stock keeping units (SKUs) must be moved away before others can be moved onto a shelf or a fixture. However, will aggregate sales also remain unchanged when sales of individual SKUs are impacted by a move? We illustrate this question in Figure 4.1. In this example, we consider the inter-fixture moves of the shirts, the shoes and the jeans. Before the move, these items have achieved sales totalling $1,200; While individual sales have been impacted, the aggregate sales remain $1,200 after the move.

![Figure 4.1: Revenue Zero-sum Game](image)

There are well documented industry benefits from shelf space management, which determines how much space should be allocated for each product. Hansen and Heinsbroek (1979) finds "about 6%" in increased profit, when taking space elasticity into consideration;
Bultez and Naert (1988) reports “actual increases in profitability varying from a low of 6.9% to a high of 33.8%” due to implementation of the SH.A.R.P. method\(^1\); Drèze et al. (1994) finds “5-6% changes [in sales] due to shelf reorganization”; van Nierop et al. (2006) uses a simulated annealing algorithm to generate “10% to 15%” profit lift. To the best of our knowledge, however, there has not been any study on the financial benefits of placement optimization, which decides \textit{where} each product should be located on the retail floor. Using retail merchandise data provided by a multi-billion dollar entertainment company, we are able to verify revenue lifts averaging about 5% over existing practice, through what is called “Plan Modification Optimization” in a series of conservative what-if analyses. We explain the idea in the following section.

4.1.1 Plan Modification Optimization (PMO)

There are two immediate hurdles in solving the full placement optimization problem: 1) The problem formulation is likely to be combinatorial and thus mathematically complex; 2) Business rules (e.g., presentation/face-out constraints, promotions/special events, marketing aesthetics) are hard to fully specify. Therefore, we devise a suboptimal “semi-automatic” heuristic that allows us to bypass these two hurdles and to assess the value of placement from historical data. It is “semi-automatic,” because this procedure requires a store planner to propose a move before it can be analyzed and modified. The key is to intelligently improve proposed placement plans through the following steps:

\begin{itemize}
  \item \textbf{Step 1} Decompose proposed plan into feasible independent components, or submoves;
  \item \textbf{Step 2} Evaluate the revenue gain/loss of each submove;
  \item \textbf{Step 3} Implement only the submoves that produce positive revenue gains.
\end{itemize}

We can visualize SKU moves with the help of a network model. Let each fixture—for instance, a shelf—be represented by a capacitated node. Let stock room be an uncapacitated super-node, where we introduce new and retire old SKUs. A submove can be defined as a

\(^1\)SH.A.R.P., or Shelf Allocation for Retailer’s Profit, is a theoretical shelf space optimization/allocation model that accounts for demand interdependencies across and within product groups.
collection of nodes (excluding the stock room super-node) that are not connected to other nodes (i.e. a component or submove). The flow chart shown in Figure 4.2 contains the details of the submove identification algorithm. As a simple example, Figure 4.3 shows two submoves (marked by red and blue colors) identified in a move.

Figure 4.2: Submove Identification Algorithm

A proposed placement plan can be modified by deciding to execute or not execute any combinations of all submoves. Each submove already represents a collection of feasible moves, proposed by a store planner. Hence, this circumvents the difficulties that may arise when we model business rules. Key steps of the "Plan Modification Optimization" (PMO)
Chapter 4. Shelf Placement Optimization

Figure 4.3: Placement Move as a Network

are exhibited in Figure 4.4.

Using the example shown in Figure 4.3, we are able to obtain two modified plans. An improved plan can be determined by projecting the financial performance of each of the four plans (original, proposed, plus two modified plans). Figure 4.5 exhibits the case where an original proposal is improved by executing only the red-colored submove.

4.1.2 Empirical Results

A crucial PMO step involves projecting the financial impact of all proposed store plans, original or modified. The projection should account for the impact of placement, which is obtainable through a regression model. At the moment, the regression model has two limitations: 1) it only works for historical analysis, meaning we have executed the original proposal and are now looking back to see if it could have been improved; 2) estimated impact of location represents an average value and is not item-specific. Our commercial research partner is developing a set of regression models that can trace the impact of location for each item. As part of ongoing research, however, we wish to formulate a clustering/testing methodology that will address both limitations.

We use the following variables to explain the sales of an individual SKU: 1) fixture placement (dummy vector), 2) display quantity or face-out, 3) unit retail price, and 4) seasonality indices. The relative impact of placement with respect to a benchmark fixture
CHAPTER 4. SHELF PLACEMENT OPTIMIZATION

Start

Read in a pair of consecutive plans

Group all SKUs into 1) those present in both plans, 2) those retired in the next plan and 3) those introduced to the next plan

Initialize the n-by-n adjacency matrix (n = # of unique fixtures in both plans plus 1 for stockroom) with 0's

Enumerate all source-sink pairs and set the corresponding entry in the adjacency matrix to 1

Set all fixtures associated with current SKU in current plan as source; Set all fixtures associated with current SKU in the next plans as sink

Yes k

Move to next SKU in the list

Set all fixtures associated with current SKU in current plan as source; Set stockroom as sink

Enumerate all source-stockroom pairs and set the corresponding entry in the adjacency matrix to 1

Move to next SKU in the list

Yes k

Set stockroom as source; Set all fixtures associated with current SKU in the next plan as sink

Enumerate all stockroom-sink pairs and set the corresponding entry in the adjacency matrix to 1

Move to next SKU in the list

Yes

Remove all adjacency matrix rows and columns that are completely 0 (these rows and columns refer to fixtures that are not involved in any moves)

Temporarily remove the stockroom (last row and last column) from the adjacency matrix

Input the resulting matrix into the algorithm that finds all independent submoves and labels corresponding fixtures (see separate flowchart)

Project revenue of the original proposal (all submoves executed); Project revenue of the pre-move plan (no submoves executed)

Processed all submove groups?

Yes

A best plan is found by executing all submoves that help store performance; A worst plan is found by executing all submoves that hurt

Construct and save PMO output

End

Figure 4.4: PMO Algorithm
Figure 4.5: Sample Output of PMO Algorithm
can be inferred through the $\beta$'s associated with the placement dummy vector. Figure 4.6 shows average revenue contribution of 20 shelves at a sample store relative to the benchmark shelf, labeled 0.

![Graph showing average contribution of shelf placement to SKU sales.](image)

**Figure 4.6: Contribution of Shelf Placement to SKU Sales**

We applied PMO to 5 stores and a total of 29 pairs of consecutive placement plans, or 29 plan changes. Here, we do not consider consecutive moves across more than 2 consecutive plans. Suppose we have consecutive plans A, B, C and D. The analyses are conducted on 3 pairs of consecutive placement plans, i.e., plans A and B, plans B and C, and finally plans C and D. Even though an improvement can be found over plan B as we analyze plans A and B, we still use the original plan B as the starting point for the analyses carried out on plans B and C. This is reasonable because PMO is used to improve a proposed plan, rather than a series of moves.

The PMO procedure shows an average revenue improvement of 5% (range: 0-24%) over the original proposal. The example in Figure 4.7 shows revenue improvement for each of the 14 plan changes at a sample store. Certain information—specifically, store name and date range of each plan—is disguised due to a nondisclosure agreement with our commercial partner.

PMO is a conservative, fast and scalable procedure to obtain the value of placement. It considers only a subset of all possible moves, but enumerating plans using submoves is a polynomial-time operation. The procedure can be applied to multiple stores and inter-store moves. That said, the value of placement is likely to vary from store to store. Some stores
4.2 Forecasting the Value of Placement

In this section, we propose a method to forecast the impact of future moves through a biclustering algorithm that operates on both the SKU- and the fixture-dimensions. SKUs may get moved to a future location for the first time. In this case, we cannot rely directly
upon historical observations to analyze the impact of this move, because the models that work well on historical moves, e.g. regression and categorical analyses, are not sufficient to establish a correlation between an SKU and its correspondingly unobserved future location. Furthermore, we note that

1. SKUs do not all respond to their placement locations in the same manner: Some SKUs may be more sensitive to locations than others, resulting in a differing change in demand due to the move;

2. Prime real estate, such as high foot-traffic areas near the store entrance, generally lifts SKU sales (at varying degrees as discussed previously), but fixtures at the back usually provide no such lift, if not depressing the sales further.

A biclustering algorithm, which operates on both the SKU- and the fixture-dimensions and employs historically observed correlations between the locational sensitivity of an SKU and the inherent financial impact of a location, may provide a good guess of those unobserved future correlations that belong to the same cluster.

In addition, store managers are often faced with the daunting task of moving thousands of SKUs across hundreds of fixtures. It is unrealistic and almost impossible to test the financial impact of each product-location combination. A clustering methodology that groups "similar" products as well as "similar" fixtures will significantly reduce the size of the problem, and better ensure that a sufficient amount of data are collected for meaningful statistical analyses.

4.2.1 Biclustering of Locational Impact on Demand

As its name suggests, a biclustering algorithm performs simultaneous clustering of both the rows and the columns of an input matrix. The row and column values of an input matrix suited for biclustering are usually categorical variables instead of multi-dimensional attributes. A traditional clustering algorithm, e.g. k-means, relies on creating a Euclidean metric for all attribute dimensions and does not work well with categorical variables.

Biclustering as a term was first introduced by Cheng and Church (2000) in gene expression analysis, although one of the earliest biclustering algorithms was the direct clustering
Demand Normalization and the Fractional Response Matrix

First, we formalize a set of simple regression models, developed by our commercial partner and discussed in Section 4.1.2, to tease out the locational impact on demand for each SKU. Assume there are $m$ items and $n$ fixtures. Let $d_{it}$ be demand of SKU $i$ at time $t$. Let $z_{ijt} \in \{0, 1\}$ represent whether SKU $i$ was placed on fixture $j$ at time $t$. Let $H_i$ be the set of all fixtures on which SKU $i$ was placed historically. Let $h^0_i$ be the base location of SKU $i$. Let $y_{it}$ be the vector of non-location explanatory variables for SKU $i$ at time $t$, such as display quantity, unit retail price and seasonality indices.

The multi-linear regression to normalize demand of SKU $i$ to control for known sources of shocks can be written as:

$$d_{it} = \beta_{i0} + \sum_{j \in H_i \setminus h^0_i} \alpha_{ij} z_{ijt} + \beta_{it}' y_{it} + \epsilon_{it}. \quad (4.1)$$

Let $\bar{y}_i$ represent the average values of the vector of non-location dependent variables. We define the normalized demand of SKU $i$ sans the location effect as follows:

$$\bar{d}_i := \beta_{i0} + \beta_{i}' \bar{y}_i.$$ 

Furthermore, we define the fractional response of SKU $i$ when it is placed on fixture $j$ as follows:

$$\delta_{ij} := \frac{\alpha_{ij}}{\bar{d}_i}.$$ 

These $\delta_{ij}$'s can be interpreted as the percentage gain/loss in demand relative to a base location $h^0_i$, and can be organized into the fractional response matrix, to which we apply the biclustering algorithm. This suggests that the normalized demand of SKU $i$ on fixture $j$, which was historically observed, can be expressed as

$$d_{ij} := \bar{d}_i + \alpha_{ij} = \bar{d}_i(1 + \delta_{ij}). \quad (4.2)$$
Note that the fractional response matrix contains missing cells when the historical fixture set $H_i$ of an SKU $i \in \{1, \ldots, m\}$ is a subset of the full fixture set $S_i := \{1, \ldots, n\}$—This occurs quite frequently, as it takes an astronomical number of trials to cover all possible item-fixture combinations.

The Standard Response Matrix

Fractional responses cannot be compared from one row to another, with each row representing an SKU. This is due to the fact that the base location $h^0_i$, which is dropped from the regression (Equation 4.1) to prevent collinearity of the location vector, may be different from SKU to SKU. We wish to standardize these responses so that, for all SKUs, they represent a percentage gain/loss in demand relative to a standard location (e.g. the store-wide "average" location).

Let $x_{ij}$ be the standard response of SKU $i$ when it is placed on fixture $j$. Let $x_i$ be the standard response of SKU $i$ when it is placed on the base fixture $h^0_i$. For SKU $i$, we then have that

$$x_{ij} = x_i + \delta_{ij}, \forall j \in H_i. \quad (4.3)$$

Let $\bar{x}_{ij}$ be the average response of the bicluster to which SKU $i$ and fixture $j$ belong. Let $C_{ij}$ represent the ID of this bicluster. To update the standard response matrix given a clustering scheme, we solve the following quadratic program:

$$\min_x \sum_{(i,j) \text{ not missing}} (x_i + \delta_{ij} - \bar{x}_{ij})^2$$

$$\text{s.t.} \sum_{(k,j)C_{ij}=C_{ij}} (x_k + \delta_{kj} - \bar{x}_{ij}) = 0, \forall (i, j) \text{ not missing}$$

$$\sum_{(i,j) \text{ not missing}} (x_i + \delta_{ij}) = 0$$

The first equality condition can be interpreted as follows: The average response of a cluster is simply the simple average of all responses within that cluster.

The above quadratic problem can be solved by finding the solution to the linear La-
grangian system:

$$\mathcal{L}(x, \bar{x}, \lambda, \mu) = \sum_{(i, j) \text{ not missing}} (x_i + \delta_{ij} - \bar{x}_{ij})^2 + \sum_{(i, j) \text{ not missing}} \sum_{(k,l) : C_{kl} = C_{ij}} \lambda_{ij}(x_k + \delta_{kl} - \bar{x}_{ij}) + \mu \sum_{(i, j) \text{ not missing}} (x_i + \delta_{ij}).$$  \hfill (4.4)

We define the following auxiliary variables:

$$y_{ij} = \begin{cases} 1, & \text{if } (i, j) \text{ is not missing;} \\ 0, & \text{otherwise;} \end{cases}$$

$$z_{klij} = \begin{cases} 1, & \text{if item-fixture pair } (k, l) \text{ is in the same cluster as } (i, j) \text{ and is not missing;} \\ 0, & \text{otherwise.} \end{cases}$$

Then, Equation 4.4 can be arranged into a linear system described in Table 4.1. We also note that updating the standard response assumes a given clustering scheme. In other words, the standard response matrix should be updated after each biclustering iteration.

Using the final standard response matrix following the completion of the clustering steps, we can arrive at an estimate of the normalized demand for SKU $i$ when it is placed on a new fixture $j$ (i.e., it is not a historically observed move) by calculating:

$$d_{ij}^{\text{est}} := \bar{d}_i(1 + \delta_{ij}^{\text{est}}) = \bar{d}_i(1 + \bar{x}_{ij} - x_i),$$  \hfill (4.5)

where $\bar{x}_{ij}$ is the average standard response of the cluster, to which $(i, j)$ belongs. Equation 4.5 is obtained from the relationships described in Equations 4.2 and 4.3.

The Biclustering Algorithm

The biclustering algorithm applied to the standard response matrix is based on the direct clustering method introduced by Hartigan (1972), which uses a divide-and-conquer approach. The algorithm starts by sorting the rows (SKUs) by their variances and the columns (fixtures) by their means. We do this because it makes sense for SKUs to be grouped by their location sensitivities, measured through the variance of the standard response to different fixtures, and for fixtures to be grouped by the average level of sales lift, measured by the mean standard response across SKUs. At each iteration, the algorithm makes a
Table 4.1: Lagrangian System for Updating Standard Response
row or column split over one of the existing clusters to improve total square error. The iterative process continues until we have reached a pre-specified number of clusters or we are running out of feasible clusters to split.

We would like to remind the reader that the standard response updating problem \((P)\) at each iteration minimizes total square error, given the clustering scheme arrived at in the previous step. This means solving \((P)\) is compatible with the objective of each clustering step and as such does not affect the optimality of the clustering scheme that was input into \((P)\) in the first place.

The flow-chart of the algorithm is shown in Figure 4.9. A few numerical examples are discussed in the following section.

### 4.2.2 Empirical Results

**Example 1: 5 items / 3 fixtures / max 3 clusters**

We consider an example with the artificial fractional response matrix \((\delta_{ij}')s\) shown in Table 4.2, where blank entries indicate missing values.

<table>
<thead>
<tr>
<th>Item</th>
<th>Fix 1</th>
<th>Fix 2</th>
<th>Fix 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>0.10000</td>
<td>0.10000</td>
<td>0.20000</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.10000</td>
<td>0.20000</td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td>0.30000</td>
<td>0.30000</td>
<td>0.60000</td>
</tr>
<tr>
<td>Item 4</td>
<td>0.30000</td>
<td>0.60000</td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>0.30000</td>
<td>0.60000</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Ex. 1 - Fractional Response Matrix

**Step 0a:** Find the standard response matrix \((x_{ij}')s\) by solving the Lagrangian system in Table C.1 (in Appendix C) and applying Equation 4.3. We obtain Table 4.3.

The \(x_i\)'s (base locations) that solved the system are

\[ x_1 = -0.13333, x_2 = -0.15000, x_3 = -0.40000, x_4 = -0.45000, x_5 = -0.45000. \]

**Step 0b:** Sort the rows of the standard response matrix above by their variances and the columns by their means, and obtain Table 4.4.
CHAPTER 4. SHELF PLACEMENT OPTIMIZATION

Start

Read in the item-fixture response matrix (converted to fractions)

Clean up the fractional response matrix by removing rows or columns that are completely missing

Initialize the standard response matrix with a single cluster
Sort the rows by their variances and sort the columns by their means
Set min error to the total square error of the entire response matrix
Set total # of clusters to 1 and set cluster ID 1 to feasible

Total # < Max # of clusters and there exists a feasible cluster ID?

Output the final clustering scheme

End

Yes

Set clustering changed to false

No

Enumerated all existing clusters?

No

Current cluster feasible?

Yes

Current cluster contains at least 4 non-missing cells?

No

Set current cluster to infeasible

Yes

There are at least 2 rows in the current cluster?

No

Considered all row split scenarios?

Yes

Each of the two post-split parts contains at least 2 non-missing cells?

No

Tag the first post-split part with the current cluster ID and tag the second post-split part with a new cluster ID (= current total + 1)

Compute total square error of the new clustering scheme

Update incumbent clustering scheme and min error
Set clustering changed to true
Track which cluster was split

Yes

Considered all column split scenarios?

No

Each of the two post-split parts contains at least 2 non-missing cells?

Yes

Tag the first post-split part with the current cluster ID and tag the second post-split part with a new cluster ID (= current total + 1)

Compute total square error of the new clustering scheme

Update incumbent clustering scheme and min error
Set clustering changed to true
Track which cluster was split

No

There are at least 2 columns in the current cluster?

No

Considered all column split scenarios?

Yes

Each of the two post-split parts contains at least 2 non-missing cells?

No

Tag the first post-split part with the current cluster ID and tag the second post-split part with a new cluster ID (= current total + 1)

Compute total square error of the new clustering scheme

Update incumbent clustering scheme and min error
Set clustering changed to true
Track which cluster was split

Yes

There are at least 2 columns in the current cluster?

No

Considered all column split scenarios?

Yes

Each of the two post-split parts contains at least 2 non-missing cells?

No

Tag the first post-split part with the current cluster ID and tag the second post-split part with a new cluster ID (= current total + 1)

Compute total square error of the new clustering scheme

Update incumbent clustering scheme and min error
Set clustering changed to true
Track which cluster was split

No

There are at least 2 rows in the current cluster?

No

Considered all row split scenarios?

Yes

Each of the two post-split parts contains at least 2 non-missing cells?

No

Tag the first post-split part with the current cluster ID and tag the second post-split part with a new cluster ID (= current total + 1)

Compute total square error of the new clustering scheme

Update incumbent clustering scheme and min error
Set clustering changed to true
Track which cluster was split

Yes

Set both post-split parts to feasible
Increment total # of clusters by 1

Update the standard response matrix with the new clustering scheme

Figure 4.9: Biclustering Algorithm
Table 4.3: Ex. 1 Step 0a - Initial Standard Response Matrix

<table>
<thead>
<tr>
<th>( x_{ij} )</th>
<th>Fix 1</th>
<th>Fix 2</th>
<th>Fix 3</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.03333</td>
<td>-0.03333</td>
<td>0.06667</td>
<td>0.00333</td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.05000</td>
<td></td>
<td>0.05000</td>
<td>0.00500</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.10000</td>
<td>-0.10000</td>
<td>0.20000</td>
<td>0.03000</td>
</tr>
<tr>
<td>Item 4</td>
<td>-0.15000</td>
<td>0.15000</td>
<td></td>
<td>0.04500</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.15000</td>
<td></td>
<td>0.15000</td>
<td>0.04500</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.08333</td>
<td>-0.09444</td>
<td>0.12333</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Ex. 1 Step 0b - Sorted Standard Response Matrix

<table>
<thead>
<tr>
<th>( x_{ij} )</th>
<th>Fix 2</th>
<th>Fix 1</th>
<th>Fix 3</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.03333</td>
<td></td>
<td>0.06667</td>
<td>0.00333</td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.05000</td>
<td>0.05000</td>
<td></td>
<td>0.00500</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.10000</td>
<td>-0.10000</td>
<td>0.20000</td>
<td>0.03000</td>
</tr>
<tr>
<td>Item 4</td>
<td>-0.15000</td>
<td></td>
<td>0.15000</td>
<td>0.04500</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.15000</td>
<td>0.15000</td>
<td></td>
<td>0.04500</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.09444</td>
<td>-0.08333</td>
<td>0.12333</td>
<td></td>
</tr>
</tbody>
</table>

Step 1a: Find the first row or column split that minimizes the total square error (best error = 0.03129), and obtain Table 4.5.

Table 4.5: Ex. 1 Step 1a - Standard Response Matrix After 1st Split

<table>
<thead>
<tr>
<th>( x_{ij} )</th>
<th>Fix 2</th>
<th>Fix 1</th>
<th>Fix 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.03333</td>
<td>-0.03333</td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td></td>
<td>-0.05000</td>
<td>0.05000</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.10000</td>
<td>-0.10000</td>
<td>0.20000</td>
</tr>
<tr>
<td>Item 4</td>
<td></td>
<td>-0.15000</td>
<td>0.15000</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.15000</td>
<td></td>
<td>0.15000</td>
</tr>
</tbody>
</table>

Step 1b: Recompute the standard response matrix by solving the Lagrangian System in Table C.2, given the updated clustering scheme. We obtain Table 4.6.

The \( x_i \)'s that solved the system are

\[ x_1 = -0.15147, x_2 = -0.13186, x_3 = -0.41814, x_4 = -0.43186, x_5 = -0.43186. \]
Table 4.6: Ex. 1 Step 1b - Updated Standard Response Matrix Given 1st Split

<table>
<thead>
<tr>
<th>$x_{ij}$</th>
<th>Fix 2</th>
<th>Fix 1</th>
<th>Fix 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.05147</td>
<td>-0.05147</td>
<td>0.04853</td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.03186</td>
<td></td>
<td>0.06814</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.11814</td>
<td>-0.11814</td>
<td>0.18186</td>
</tr>
<tr>
<td>Item 4</td>
<td>-0.13186</td>
<td></td>
<td>0.16814</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.13186</td>
<td></td>
<td>0.16814</td>
</tr>
</tbody>
</table>

Table 4.7: Ex. 1 Step 2a - Standard Response Matrix After 2nd Split

<table>
<thead>
<tr>
<th>$x_{ij}$</th>
<th>Fix 2</th>
<th>Fix 1</th>
<th>Fix 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>-0.09167</td>
<td>-0.09167</td>
<td>0.00833</td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.09167</td>
<td></td>
<td>0.00833</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.09167</td>
<td>-0.09167</td>
<td>0.20833</td>
</tr>
<tr>
<td>Item 4</td>
<td>-0.09167</td>
<td></td>
<td>0.20833</td>
</tr>
<tr>
<td>Item 5</td>
<td>-0.09167</td>
<td></td>
<td>0.20833</td>
</tr>
</tbody>
</table>

Table 4.8: Ex. 1 Step 2b - Updated Standard Response Matrix Given 2nd Split

The $x_i$'s that solved the system are

$$x_1 = -0.19167, x_2 = -0.19167, x_3 = -0.39167, x_4 = -0.39167, x_5 = -0.39167.$$
We observe in this particular example that subtracting the final set of $x_i$'s from final standard response matrix in Table 4.8 recovers the original fractional response matrix Table 4.2 for the item-fixture combinations that were historically present. Furthermore, we are also able to say, for instance, that Item 2 will receive a fractional lift of 0.1 when it is placed on Fixture 2, i.e. $\delta_{22} = -0.09167 - 0.19176 = 0.1$. We arrive at the following fractional response table with a forecast for all previously empty cells, which represents historically unobserved item-fixture combinations:

\[
\begin{array}{c|ccc}
\delta_{ij} & \text{Fix 1} & \text{Fix 2} & \text{Fix 3} \\
\hline
\text{Item 1} & 0.10000 & 0.10000 & 0.20000 \\
\text{Item 2} & 0.10000 & 0.10000 & 0.20000 \\
\text{Item 3} & 0.30000 & 0.30000 & 0.60000 \\
\text{Item 4} & 0.30000 & 0.30000 & 0.60000 \\
\text{Item 5} & 0.30000 & 0.30000 & 0.60000 \\
\end{array}
\]

Table 4.9: Ex. 1 - Forecasted Fractional Response Matrix

Example 2: disguised store data with 45 items / 7 fixtures / max 20 clusters

There are a total of 161 items across 38 fixtures in the initial data that we received from our commercial partner. For simplicity, we will demonstrate the performance of the biclustering algorithm on a subset of that data containing 45 items and 7 fixtures. Item and fixture names are disguised. The exhibits are organized as follows:

- Table 4.11 shows the fractional response matrix;
- Table 4.12 shows the clustered standard response matrix;
- Table 4.13 shows the clustering scheme;
- Table 4.14 shows the forecasted fractional response matrix (see Equation 4.5).

We summarize the performance of the clustering scheme in Table 4.10, which is sorted by the standard deviation of each cluster. Most of the clusters appear to perform well with a relatively low standard deviation amongst the non-missing item-fixture combinations in
the cluster. Clusters with high standard deviations could be improved if we allow more clusters (greater than the designated maximum of 20 clusters) to be made.

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>-0.3497</td>
<td>0.0000</td>
<td>-0.3497</td>
<td>-0.3497</td>
</tr>
<tr>
<td>11</td>
<td>0.5957</td>
<td>0.0000</td>
<td>0.5957</td>
<td>0.5957</td>
</tr>
<tr>
<td>4</td>
<td>0.6729</td>
<td>0.0000</td>
<td>0.6729</td>
<td>0.6729</td>
</tr>
<tr>
<td>16</td>
<td>-1.2101</td>
<td>0.0000</td>
<td>-1.2101</td>
<td>-1.2101</td>
</tr>
<tr>
<td>10</td>
<td>0.0789</td>
<td>0.0058</td>
<td>0.0897</td>
<td>0.0681</td>
</tr>
<tr>
<td>1</td>
<td>0.3509</td>
<td>0.0115</td>
<td>0.3650</td>
<td>0.3369</td>
</tr>
<tr>
<td>6</td>
<td>0.7394</td>
<td>0.0273</td>
<td>0.7587</td>
<td>0.7201</td>
</tr>
<tr>
<td>17</td>
<td>-0.2298</td>
<td>0.0361</td>
<td>-0.2043</td>
<td>-0.2553</td>
</tr>
<tr>
<td>20</td>
<td>0.1179</td>
<td>0.0505</td>
<td>0.1536</td>
<td>0.0822</td>
</tr>
<tr>
<td>2</td>
<td>0.5007</td>
<td>0.0575</td>
<td>0.6606</td>
<td>0.4006</td>
</tr>
<tr>
<td>5</td>
<td>0.5907</td>
<td>0.0587</td>
<td>0.6908</td>
<td>0.4308</td>
</tr>
<tr>
<td>15</td>
<td>0.0030</td>
<td>0.0710</td>
<td>0.0532</td>
<td>-0.0472</td>
</tr>
<tr>
<td>13</td>
<td>0.3085</td>
<td>0.0862</td>
<td>0.3694</td>
<td>0.2475</td>
</tr>
<tr>
<td>19</td>
<td>-0.7636</td>
<td>0.0892</td>
<td>-0.6390</td>
<td>-0.8883</td>
</tr>
<tr>
<td>12</td>
<td>-1.6983</td>
<td>0.3532</td>
<td>-1.4486</td>
<td>-1.9480</td>
</tr>
<tr>
<td>18</td>
<td>-0.9623</td>
<td>0.3532</td>
<td>-0.7126</td>
<td>-1.2121</td>
</tr>
<tr>
<td>9</td>
<td>-0.2569</td>
<td>0.3709</td>
<td>0.1140</td>
<td>-0.6277</td>
</tr>
<tr>
<td>8</td>
<td>1.1372</td>
<td>0.5245</td>
<td>1.5080</td>
<td>0.7663</td>
</tr>
<tr>
<td>3</td>
<td>-2.4495</td>
<td>0.5245</td>
<td>-2.0786</td>
<td>-2.8203</td>
</tr>
<tr>
<td>7</td>
<td>-1.7982</td>
<td>0.6396</td>
<td>-1.1054</td>
<td>-2.3664</td>
</tr>
</tbody>
</table>

Table 4.10: Ex. 2 - Performance of Standard Response by Cluster
<table>
<thead>
<tr>
<th>( \delta_{ij} )</th>
<th>Fix 1</th>
<th>Fix 2</th>
<th>Fix 3</th>
<th>Fix 4</th>
<th>Fix 5</th>
<th>Fix 6</th>
<th>Fix 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>0.4341</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.7190</td>
</tr>
<tr>
<td>Item 2</td>
<td></td>
<td>-0.6278</td>
<td></td>
<td>-1.2879</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td></td>
<td></td>
<td>1.7331</td>
<td>0.9872</td>
<td></td>
<td>1.3962</td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td></td>
<td></td>
<td></td>
<td>-0.5763</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>2.5623</td>
<td>-2.2721</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td></td>
<td>0.6628</td>
<td></td>
<td></td>
<td>-0.8314</td>
<td></td>
<td>0.9544</td>
</tr>
<tr>
<td>Item 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.7683</td>
<td></td>
</tr>
<tr>
<td>Item 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.8725</td>
</tr>
<tr>
<td>Item 9</td>
<td></td>
<td>1.1867</td>
<td></td>
<td></td>
<td>0.8180</td>
<td>0.9224</td>
<td>1.2337</td>
</tr>
<tr>
<td>Item 10</td>
<td></td>
<td>0.3524</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.3274</td>
</tr>
<tr>
<td>Item 11</td>
<td></td>
<td>-0.6935</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 12</td>
<td></td>
<td></td>
<td></td>
<td>0.9275</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 13</td>
<td></td>
<td></td>
<td>1.6898</td>
<td></td>
<td></td>
<td></td>
<td>1.2268</td>
</tr>
<tr>
<td>Item 14</td>
<td></td>
<td></td>
<td></td>
<td>-1.5355</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.5563</td>
</tr>
<tr>
<td>Item 16</td>
<td></td>
<td>-0.4472</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.7859</td>
</tr>
<tr>
<td>Item 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.4966</td>
</tr>
<tr>
<td>Item 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.3522</td>
</tr>
<tr>
<td>Item 20</td>
<td></td>
<td>-0.6499</td>
<td></td>
<td>-0.5282</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 21</td>
<td></td>
<td></td>
<td>2.5386</td>
<td>0.9969</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 22</td>
<td></td>
<td></td>
<td>0.6348</td>
<td>0.4220</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 23</td>
<td></td>
<td></td>
<td></td>
<td>-0.3831</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 24</td>
<td></td>
<td></td>
<td>1.3273</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 25</td>
<td></td>
<td></td>
<td>1.3966</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 27</td>
<td></td>
<td>0.9596</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 28</td>
<td></td>
<td>0.6582</td>
<td></td>
<td>-0.7112</td>
<td>0.7397</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 29</td>
<td></td>
<td>0.2844</td>
<td></td>
<td>-0.8184</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 30</td>
<td></td>
<td></td>
<td></td>
<td>0.3757</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 31</td>
<td></td>
<td>1.6444</td>
<td>0.9650</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 32</td>
<td></td>
<td>1.4157</td>
<td>1.9132</td>
<td>0.6778</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 33</td>
<td></td>
<td>2.1597</td>
<td>0.8759</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 34</td>
<td></td>
<td>0.8764</td>
<td>0.6992</td>
<td>-0.6172</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 35</td>
<td></td>
<td>-0.9289</td>
<td></td>
<td></td>
<td>0.5355</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 36</td>
<td></td>
<td>-0.8472</td>
<td></td>
<td>-0.7529</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 37</td>
<td></td>
<td>-0.5649</td>
<td></td>
<td></td>
<td>0.3883</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 38</td>
<td></td>
<td>0.8712</td>
<td>0.4756</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 39</td>
<td></td>
<td>0.7895</td>
<td>1.1934</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 40</td>
<td></td>
<td>0.8749</td>
<td>0.7253</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 41</td>
<td></td>
<td></td>
<td></td>
<td>-0.9359</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.1254</td>
</tr>
<tr>
<td>Item 43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.2969</td>
</tr>
<tr>
<td>Item 44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.5685</td>
</tr>
<tr>
<td>Item 45</td>
<td></td>
<td>1.7574</td>
<td>1.5680</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.11: Ex. 2 - Fractional Response Matrix
### Table 4.12: Ex. 2 - Final Standard Response Matrix

<table>
<thead>
<tr>
<th>( x_{ij} )</th>
<th>Fix 2</th>
<th>Fix 6</th>
<th>Fix 5</th>
<th>Fix 4</th>
<th>Fix 3</th>
<th>Fix 7</th>
<th>Fix 1</th>
<th>Base (( x_i ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 4</td>
<td>0.5724</td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>1.0084</td>
</tr>
<tr>
<td>Item 7</td>
<td>0.7888</td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>0.8820</td>
</tr>
<tr>
<td>Item 8</td>
<td>0.7888</td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>0.8472</td>
</tr>
<tr>
<td>Item 12</td>
<td></td>
<td>0.6729</td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td>1.3659</td>
</tr>
<tr>
<td>Item 13</td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td>1.0064</td>
</tr>
<tr>
<td>Item 15</td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td>1.6144</td>
</tr>
<tr>
<td>Item 16</td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td>0.6352</td>
</tr>
<tr>
<td>Item 17</td>
<td>0.3594</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.7982</td>
</tr>
<tr>
<td>Item 18</td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>1.3766</td>
</tr>
<tr>
<td>Item 19</td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>1.0813</td>
</tr>
<tr>
<td>Item 20</td>
<td>0.5957</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.9460</td>
</tr>
<tr>
<td>Item 24</td>
<td>0.5724</td>
<td></td>
<td></td>
<td></td>
<td>0.5724</td>
<td></td>
<td></td>
<td>0.8837</td>
</tr>
<tr>
<td>Item 25</td>
<td>0.5966</td>
<td></td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>-0.4414</td>
</tr>
<tr>
<td>Item 26</td>
<td>0.5966</td>
<td></td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>-0.8060</td>
</tr>
<tr>
<td>Item 27</td>
<td>0.5724</td>
<td></td>
<td></td>
<td></td>
<td>0.5724</td>
<td></td>
<td></td>
<td>-0.4583</td>
</tr>
<tr>
<td>Item 30</td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td>0.2150</td>
<td>1.4358</td>
</tr>
<tr>
<td>Item 41</td>
<td>0.5724</td>
<td></td>
<td></td>
<td></td>
<td>0.5724</td>
<td></td>
<td></td>
<td>-1.0462</td>
</tr>
<tr>
<td>Item 42</td>
<td>0.7888</td>
<td></td>
<td></td>
<td></td>
<td>0.7888</td>
<td></td>
<td></td>
<td>-3.7063</td>
</tr>
<tr>
<td>Item 44</td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>-0.9778</td>
</tr>
<tr>
<td>Item 40</td>
<td>0.5966</td>
<td>0.6729</td>
<td></td>
<td></td>
<td>0.5966</td>
<td></td>
<td></td>
<td>-0.1346</td>
</tr>
<tr>
<td>Item 36</td>
<td>0.5957</td>
<td>0.5724</td>
<td></td>
<td></td>
<td>0.5724</td>
<td></td>
<td></td>
<td>1.3478</td>
</tr>
<tr>
<td>Item 21</td>
<td>0.3650</td>
<td>0.4867</td>
<td></td>
<td></td>
<td>0.4867</td>
<td></td>
<td></td>
<td>1.0149</td>
</tr>
<tr>
<td>Item 23</td>
<td>0.4550</td>
<td>0.6364</td>
<td></td>
<td></td>
<td>0.6364</td>
<td></td>
<td></td>
<td>0.0330</td>
</tr>
<tr>
<td>Item 39</td>
<td>0.6661</td>
<td>0.4377</td>
<td></td>
<td></td>
<td>0.4377</td>
<td></td>
<td></td>
<td>-0.3587</td>
</tr>
<tr>
<td>Item 1</td>
<td>0.3369</td>
<td></td>
<td>0.6470</td>
<td></td>
<td></td>
<td></td>
<td>0.6470</td>
<td>-0.0972</td>
</tr>
<tr>
<td>Item 6</td>
<td>0.4634</td>
<td>0.6975</td>
<td></td>
<td></td>
<td>0.6975</td>
<td></td>
<td></td>
<td>-0.2596</td>
</tr>
<tr>
<td>Item 9</td>
<td>0.5617</td>
<td>0.2475</td>
<td></td>
<td></td>
<td>0.2475</td>
<td></td>
<td></td>
<td>-0.6250</td>
</tr>
<tr>
<td>Item 10</td>
<td>0.6889</td>
<td>-0.4720</td>
<td>0.3694</td>
<td></td>
<td>0.3694</td>
<td></td>
<td></td>
<td>-0.8643</td>
</tr>
<tr>
<td>Item 38</td>
<td>0.3316</td>
<td>0.4555</td>
<td></td>
<td></td>
<td>0.4555</td>
<td></td>
<td></td>
<td>-0.4205</td>
</tr>
<tr>
<td>Item 2</td>
<td>0.8967</td>
<td>0.4699</td>
<td></td>
<td></td>
<td>0.4699</td>
<td></td>
<td></td>
<td>1.1177</td>
</tr>
<tr>
<td>Item 14</td>
<td>0.5365</td>
<td>0.8218</td>
<td></td>
<td></td>
<td>0.8218</td>
<td></td>
<td></td>
<td>-1.1445</td>
</tr>
<tr>
<td>Item 3</td>
<td>-0.2553</td>
<td>0.4954</td>
<td></td>
<td></td>
<td>0.4954</td>
<td></td>
<td></td>
<td>-1.2426</td>
</tr>
<tr>
<td>Item 11</td>
<td></td>
<td>0.1140</td>
<td>0.7663</td>
<td>0.1140</td>
<td>0.7663</td>
<td></td>
<td></td>
<td>0.4161</td>
</tr>
<tr>
<td>Item 31</td>
<td>-0.2425</td>
<td>0.4752</td>
<td></td>
<td></td>
<td>0.4752</td>
<td></td>
<td></td>
<td>-1.1692</td>
</tr>
<tr>
<td>Item 45</td>
<td>0.3594</td>
<td></td>
<td></td>
<td>-0.3497</td>
<td></td>
<td></td>
<td></td>
<td>-1.4065</td>
</tr>
<tr>
<td>Item 29</td>
<td>-1.4486</td>
<td>-1.2126</td>
<td>-0.3497</td>
<td></td>
<td>-0.3497</td>
<td></td>
<td></td>
<td>-0.6301</td>
</tr>
<tr>
<td>Item 37</td>
<td></td>
<td>-1.2166</td>
<td></td>
<td>-0.2569</td>
<td></td>
<td></td>
<td></td>
<td>-0.6451</td>
</tr>
<tr>
<td>Item 32</td>
<td>-1.9485</td>
<td>-0.7126</td>
<td>-1.2166</td>
<td></td>
<td>-1.2166</td>
<td></td>
<td></td>
<td>-2.6258</td>
</tr>
<tr>
<td>Item 33</td>
<td></td>
<td>-1.9329</td>
<td>-0.6390</td>
<td>-1.2166</td>
<td>-1.2166</td>
<td></td>
<td></td>
<td>-2.7987</td>
</tr>
<tr>
<td>Item 34</td>
<td>-2.3664</td>
<td>-1.1545</td>
<td>-0.8883</td>
<td></td>
<td>-0.8883</td>
<td></td>
<td></td>
<td>-1.7647</td>
</tr>
<tr>
<td>Item 28</td>
<td>-2.7862</td>
<td>-0.7636</td>
<td>-0.6277</td>
<td></td>
<td>-0.6277</td>
<td></td>
<td></td>
<td>-1.3674</td>
</tr>
<tr>
<td>Item 35</td>
<td></td>
<td>-0.7443</td>
<td>0.7256</td>
<td></td>
<td>0.7256</td>
<td></td>
<td></td>
<td>0.1846</td>
</tr>
<tr>
<td>Item 22</td>
<td></td>
<td>-0.7830</td>
<td>0.7587</td>
<td>1.5826</td>
<td>0.7587</td>
<td></td>
<td></td>
<td>-1.7798</td>
</tr>
<tr>
<td>Item 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.8233</td>
</tr>
</tbody>
</table>

Table 4.12: Ex. 2 - Final Standard Response Matrix
<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Fix 1</th>
<th>Fix 2</th>
<th>Fix 3</th>
<th>Fix 4</th>
<th>Fix 5</th>
<th>Fix 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 4</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 7</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 8</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 12</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 13</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 15</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 16</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 17</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 18</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 19</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 20</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 24</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 25</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 26</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 27</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 30</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 41</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 42</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 43</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 44</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 40</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 36</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 21</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 23</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 39</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 1</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 6</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Item 9</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Item 10</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Item 38</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>2</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Item 2</td>
<td>1</td>
<td>10</td>
<td>17</td>
<td>2</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Item 14</td>
<td>1</td>
<td>12</td>
<td>17</td>
<td>2</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Item 3</td>
<td>1</td>
<td>12</td>
<td>17</td>
<td>2</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Item 11</td>
<td>1</td>
<td>12</td>
<td>17</td>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 12</td>
<td>1</td>
<td>12</td>
<td>17</td>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 42</td>
<td>1</td>
<td>12</td>
<td>18</td>
<td>4</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 29</td>
<td>1</td>
<td>12</td>
<td>18</td>
<td>4</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 37</td>
<td>1</td>
<td>12</td>
<td>18</td>
<td>6</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 32</td>
<td>1</td>
<td>12</td>
<td>18</td>
<td>6</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 33</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>19</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 34</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>19</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 28</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Item 35</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Item 22</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Item 5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>19</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 4.13: Ex. 2 - Clustering Scheme
Table 4.14: Ex. 2 - Forecasted Fractional Response Matrix

<table>
<thead>
<tr>
<th>Item</th>
<th>( S_f )</th>
<th>Fix 2</th>
<th>Fix 6</th>
<th>Fix 5</th>
<th>Fix 4</th>
<th>Fix 3</th>
<th>Fix 7</th>
<th>Fix 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 4</td>
<td>-0.6574</td>
<td>-0.9295</td>
<td>-0.4126</td>
<td>-0.5076</td>
<td>-0.4177</td>
<td>-0.4177</td>
<td>-0.3355</td>
<td></td>
</tr>
<tr>
<td>Item 7</td>
<td>-0.5311</td>
<td>-0.8031</td>
<td>-0.2863</td>
<td>-0.3813</td>
<td>-0.2913</td>
<td>-0.2913</td>
<td>-0.2091</td>
<td></td>
</tr>
<tr>
<td>Item 8</td>
<td>-0.4962</td>
<td>-0.7683</td>
<td>-0.2514</td>
<td>-0.3645</td>
<td>-0.2565</td>
<td>-0.2565</td>
<td>-0.1743</td>
<td></td>
</tr>
<tr>
<td>Item 12</td>
<td>-1.0150</td>
<td>-1.2871</td>
<td>-0.7702</td>
<td>-0.8652</td>
<td>-0.7753</td>
<td>-0.7753</td>
<td>-0.6930</td>
<td></td>
</tr>
<tr>
<td>Item 13</td>
<td>-0.6555</td>
<td>-0.9275</td>
<td>-0.4107</td>
<td>-0.5057</td>
<td>-0.4157</td>
<td>-0.4157</td>
<td>-0.3335</td>
<td></td>
</tr>
<tr>
<td>Item 15</td>
<td>-1.2634</td>
<td>-1.5395</td>
<td>-1.0186</td>
<td>-1.1136</td>
<td>-1.0237</td>
<td>-1.0237</td>
<td>-0.9415</td>
<td></td>
</tr>
<tr>
<td>Item 16</td>
<td>-0.2842</td>
<td>-0.5563</td>
<td>-0.0394</td>
<td>-0.1345</td>
<td>-0.0445</td>
<td>-0.0445</td>
<td>0.0377</td>
<td></td>
</tr>
<tr>
<td>Item 17</td>
<td>-0.4472</td>
<td>-0.7193</td>
<td>-0.2024</td>
<td>-0.2974</td>
<td>-0.2075</td>
<td>-0.2075</td>
<td>-0.1253</td>
<td></td>
</tr>
<tr>
<td>Item 18</td>
<td>-1.0256</td>
<td>-1.2977</td>
<td>-0.7808</td>
<td>-0.8799</td>
<td>-0.7859</td>
<td>-0.7859</td>
<td>-0.7037</td>
<td></td>
</tr>
<tr>
<td>Item 19</td>
<td>-0.7304</td>
<td>-1.0024</td>
<td>-0.4856</td>
<td>-0.5806</td>
<td>-0.4907</td>
<td>-0.4907</td>
<td>-0.4084</td>
<td></td>
</tr>
<tr>
<td>Item 20</td>
<td>-0.5950</td>
<td>-0.8671</td>
<td>-0.3502</td>
<td>-0.4452</td>
<td>-0.3553</td>
<td>-0.3553</td>
<td>-0.2731</td>
<td></td>
</tr>
<tr>
<td>Item 24</td>
<td>-0.5328</td>
<td>-0.8049</td>
<td>-0.2880</td>
<td>-0.3830</td>
<td>-0.2931</td>
<td>-0.2931</td>
<td>-0.2109</td>
<td></td>
</tr>
<tr>
<td>Item 25</td>
<td>0.7924</td>
<td>0.5203</td>
<td>1.0372</td>
<td>0.9421</td>
<td>1.0321</td>
<td>1.0321</td>
<td>1.1143</td>
<td></td>
</tr>
<tr>
<td>Item 26</td>
<td>1.1569</td>
<td>0.8848</td>
<td>1.4017</td>
<td>1.3067</td>
<td>1.3966</td>
<td>1.3966</td>
<td>1.4789</td>
<td></td>
</tr>
<tr>
<td>Item 27</td>
<td>0.8092</td>
<td>0.9372</td>
<td>1.0540</td>
<td>0.9590</td>
<td>1.0489</td>
<td>1.0489</td>
<td>1.1312</td>
<td></td>
</tr>
<tr>
<td>Item 30</td>
<td>0.1360</td>
<td>-0.1361</td>
<td>0.3808</td>
<td>0.2887</td>
<td>0.3757</td>
<td>0.3757</td>
<td>0.4579</td>
<td></td>
</tr>
<tr>
<td>Item 41</td>
<td>-1.0849</td>
<td>-1.3569</td>
<td>-0.8401</td>
<td>-0.9351</td>
<td>-0.8452</td>
<td>-0.8452</td>
<td>-0.7629</td>
<td></td>
</tr>
<tr>
<td>Item 42</td>
<td>1.3971</td>
<td>1.1230</td>
<td>1.6419</td>
<td>1.5469</td>
<td>1.6368</td>
<td>1.6368</td>
<td>1.7190</td>
<td></td>
</tr>
<tr>
<td>Item 43</td>
<td>4.0572</td>
<td>3.7952</td>
<td>4.3020</td>
<td>4.2070</td>
<td>4.2969</td>
<td>4.2969</td>
<td>4.3792</td>
<td></td>
</tr>
<tr>
<td>Item 44</td>
<td>1.3288</td>
<td>1.0567</td>
<td>1.5736</td>
<td>1.4785</td>
<td>1.5685</td>
<td>1.5685</td>
<td>1.6507</td>
<td></td>
</tr>
<tr>
<td>Item 40</td>
<td>0.4855</td>
<td>0.2135</td>
<td>0.7304</td>
<td>0.6553</td>
<td>0.7253</td>
<td>0.7253</td>
<td>0.8075</td>
<td></td>
</tr>
<tr>
<td>Item 36</td>
<td>-0.9969</td>
<td>-1.2690</td>
<td>-0.7521</td>
<td>-0.8471</td>
<td>-0.7572</td>
<td>-0.7572</td>
<td>-0.2107</td>
<td></td>
</tr>
<tr>
<td>Item 21</td>
<td>-0.6499</td>
<td>-0.9360</td>
<td>-1.0119</td>
<td>-0.5282</td>
<td>-0.4242</td>
<td>-0.4242</td>
<td>0.1223</td>
<td></td>
</tr>
<tr>
<td>Item 23</td>
<td>0.3180</td>
<td>0.0459</td>
<td>-0.0300</td>
<td>0.4220</td>
<td>0.6035</td>
<td>0.5577</td>
<td>1.1042</td>
<td></td>
</tr>
<tr>
<td>Item 39</td>
<td>0.7097</td>
<td>0.4376</td>
<td>0.3617</td>
<td>1.0193</td>
<td>0.7895</td>
<td>0.9494</td>
<td>1.4959</td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>0.4341</td>
<td>0.1761</td>
<td>0.1002</td>
<td>0.5979</td>
<td>0.6879</td>
<td>0.7019</td>
<td>1.2344</td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td>0.6306</td>
<td>0.3385</td>
<td>0.2626</td>
<td>0.6603</td>
<td>0.8803</td>
<td>0.9504</td>
<td>1.3968</td>
<td></td>
</tr>
<tr>
<td>Item 9</td>
<td>0.9759</td>
<td>0.7038</td>
<td>0.6279</td>
<td>1.1687</td>
<td>0.9334</td>
<td>0.8725</td>
<td>1.7621</td>
<td></td>
</tr>
<tr>
<td>Item 10</td>
<td>1.2152</td>
<td>0.9934</td>
<td>0.8171</td>
<td>1.3605</td>
<td>1.1728</td>
<td>1.2337</td>
<td>2.0315</td>
<td></td>
</tr>
<tr>
<td>Item 38</td>
<td>0.7714</td>
<td>0.4993</td>
<td>0.4736</td>
<td>0.8710</td>
<td>0.5384</td>
<td>0.5384</td>
<td>1.5576</td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td>-0.7668</td>
<td>-1.0281</td>
<td>-1.3475</td>
<td>-0.6278</td>
<td>-0.9998</td>
<td>-0.9998</td>
<td>0.0194</td>
<td></td>
</tr>
<tr>
<td>Item 14</td>
<td>1.4955</td>
<td>-0.5538</td>
<td>0.9147</td>
<td>1.6810</td>
<td>1.2624</td>
<td>1.2267</td>
<td>2.2817</td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td>1.5935</td>
<td>-0.4557</td>
<td>0.9872</td>
<td>1.7331</td>
<td>1.3605</td>
<td>1.3962</td>
<td>2.3707</td>
<td></td>
</tr>
<tr>
<td>Item 11</td>
<td>-0.0651</td>
<td>-2.1143</td>
<td>-0.6459</td>
<td>-0.0847</td>
<td>-0.6729</td>
<td>-0.3021</td>
<td>0.3502</td>
<td></td>
</tr>
<tr>
<td>Item 31</td>
<td>1.5202</td>
<td>-0.5290</td>
<td>0.9650</td>
<td>1.6444</td>
<td>0.9124</td>
<td>0.9124</td>
<td>2.3064</td>
<td></td>
</tr>
<tr>
<td>Item 45</td>
<td>1.7574</td>
<td>-0.2918</td>
<td>0.4441</td>
<td>1.0568</td>
<td>1.1496</td>
<td>1.1496</td>
<td>2.5436</td>
<td></td>
</tr>
<tr>
<td>Item 29</td>
<td>0.9811</td>
<td>-0.8184</td>
<td>-0.5819</td>
<td>0.2804</td>
<td>0.3732</td>
<td>0.3732</td>
<td>1.7673</td>
<td></td>
</tr>
<tr>
<td>Item 37</td>
<td>0.9961</td>
<td>-1.0531</td>
<td>-0.3172</td>
<td>-0.5649</td>
<td>0.3883</td>
<td>0.3883</td>
<td>1.7823</td>
<td></td>
</tr>
<tr>
<td>Item 32</td>
<td>2.9767</td>
<td>0.6778</td>
<td>1.9132</td>
<td>1.4157</td>
<td>2.3689</td>
<td>2.3689</td>
<td>3.7630</td>
<td></td>
</tr>
<tr>
<td>Item 33</td>
<td>1.0005</td>
<td>1.0005</td>
<td>0.8759</td>
<td>2.1597</td>
<td>2.5419</td>
<td>2.5419</td>
<td>3.9359</td>
<td></td>
</tr>
<tr>
<td>Item 34</td>
<td>-0.0336</td>
<td>-0.6017</td>
<td>0.6592</td>
<td>0.8764</td>
<td>1.5078</td>
<td>1.5078</td>
<td>2.9018</td>
<td></td>
</tr>
<tr>
<td>Item 28</td>
<td>-1.0820</td>
<td>-0.7112</td>
<td>-1.0820</td>
<td>0.6038</td>
<td>1.1106</td>
<td>0.7397</td>
<td>2.5046</td>
<td></td>
</tr>
<tr>
<td>Item 35</td>
<td>-2.6341</td>
<td>-2.6341</td>
<td>-2.6341</td>
<td>-0.9289</td>
<td>0.3548</td>
<td>0.5355</td>
<td>0.9526</td>
<td></td>
</tr>
<tr>
<td>Item 22</td>
<td>-0.6696</td>
<td>-0.6696</td>
<td>-0.6696</td>
<td>0.9969</td>
<td>2.5386</td>
<td>2.5192</td>
<td>2.9170</td>
<td></td>
</tr>
<tr>
<td>Item 5</td>
<td>-2.2721</td>
<td>-1.9013</td>
<td>-1.9013</td>
<td>-0.2154</td>
<td>1.2876</td>
<td>1.2876</td>
<td>2.0562</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.14: Ex. 2 - Forecasted Fractional Response Matrix
4.2.3 A Note on Data Scarcity

The SKU-oriented regression model, first described in Section 4.1.2 and later formalized in 4.2.1, has an important restriction: an SKU needs to be present long enough before it can be fitted by the regression model. That restriction comes from the normality assumption of the error term in a multi-linear regression. In a regression with only one independent variable, it translates to a practical minimum of 30 weekly\(^2\) observations. Needless to say, the SKU-level regression model currently implemented by our commercial partner, has a lot more variables, plenty of which are used to capture locational impact on a particular SKU, and thus requires a lot more observations to ensure sufficient degrees of freedom in the error term.

Even with the adopted cutoff at 40 weeks, statistical implications aside, an SKU would have to be tracked for nearly 10 months before we can estimate the impact of location on its sales. A large majority of the SKUs sold by our commercial partner do have a shelf life much longer than 40 weeks, so eventually we have sufficient data to estimate locational impact. However at the inception of this project, locational data was just beginning to be properly tracked. As a result, a large number of SKUs did not have accurate location estimates.

While the biclustering method in Figure 4.9 is designed to accommodate a sparse SKU-fixture response matrix, arising from insufficient natural experimentation on historical moves, it still requires that some observations of the locational impact on SKU sales be present—i.e., the SKU still needs to be tracked for more than the adopted cutoff time. Until we allowed enough time for more SKUs to make the cut, implementing the biclustering algorithm was not a practical reality.

Our commercial partner has therefore proposed a fixture-oriented model. Instead of forecasting the SKU demand directly, the fixture-oriented model estimates the total aggregate demand of a fixture. It also relies on the consolidation of SKUs into item clusters ("ICs") as a first step, and therefore does not generally suffer from the same data scarcity problem as the SKU-oriented model.

\(^2\)Retailers generally track sales and inventory data on a weekly basis, possibly due to operational habit. Another reason is that one can smooth over the daily sales volatilities with weekly data.
CHAPTER 4. SHELF PLACEMENT OPTIMIZATION

The Fixture-Oriented Regression Model

SKUs are first grouped into $c$ distinct ICs by sales volume, price, and vintage\(^3\). At period $t$, let total aggregate sales volume of fixture $j \in \{1, \ldots, n\}$ be denoted $D_{jt}$, which is obtained by simply adding up the sales volume of each individual SKU on fixture $j$. $D_{jt}$ is then regressed against two sets of factors: 1) a total of $q$ sets of IC-specific attributes $u_{pjk}$ ($p \in \{1, \ldots, q\}$ and $k \in \{1, \ldots, c\}$), such as total number of different SKUs in each IC, average presentation quantity and average price, both weighted by historical sales of SKUs in each IC, and 2) store-level and seasonal indices $v_{jt}$, e.g. foot traffic to the store and weather. We then multiply the predicted total sales volume of fixture $j$ by the corresponding weighted average price to reach an estimate of the total dollar sales of the entire fixture, often referred to as the fixture productivity. The fixture-oriented regression model can be formally defined as the following:

$$D_{jt} = \psi_{0j} + \sum_{p=1}^{q} \sum_{k=1}^{c} \phi_{pjk} u_{pjk} + \psi_j' v_{jt} + \epsilon_j. \quad (4.6)$$

Tackling the Sparsity of Fixture-IC Coefficient Matrices with Biclustering

For each set $p \in \{1, \ldots, q\}$ of the IC-specific attributes from (4.6), we denote the fixture-IC coefficient matrix $\Phi_p := [\phi_{pjk}]_{j=1}^{n} k=1^{c}$. While this matrix does not generally suffer from data scarcity problem, i.e., every row $j$ contains a nonzero row vector, the matrix $\Phi_p$ is still not immune to the insufficiency of natural experimentation on historical IC moves. In other words, there may still be missing cells in $\Phi_p$, for example when a particular IC has never been placed on a particular fixture before. To forecast the missing coefficients, we propose applying the direct biclustering algorithm from Hartigan (1972) to $\Phi_p$ or the corresponding standardized coefficient matrix. We refer the reader to Bring (1994) for a discussion on standardization of regression coefficients.

\(^3\)That is, how many weeks since an SKU was first introduced.
4.3 Conclusions and Future Research

Through an applied project, we examined a number of issues surrounding shelf placement optimization in a retail environment. Placement optimization was a relatively new addition to the toolkit of retail revenue management, however we were able to establish its financial impact through the creation of a few practical systems.

First, we confirmed that placement was not a zero-sum game from a revenue perspective. In order to quantify the benefit of placement optimization, we developed the "Plan Modification Optimization" algorithm. The PMO algorithm constituted a lower bound on the benefit of placement, which we estimated to be roughly 5% using retail data from our commercial partner, and represented a practical (albeit partial) solution to the placement optimization problem. The PMO algorithm was designed to bypass the combinatorial nature of a full placement optimization formulation and the difficulty in adequately specifying business rules unique to each retail environment. The key of PMO was to intelligently improve a proposed placement plan by 1) decomposing it into feasible independent sub-moves, 2) by evaluating the revenue impact of each submove, and 3) by implementing only the submoves with positive revenue gains.

Next, we proposed a biclustering algorithm to forecast the impact of future moves. SKUs could move to a future location for the first time. In that case, traditional models that worked well on historical moves were no longer sufficient to establish a correlation between an SKU and its correspondingly unobserved location. A biclustering algorithm, which operated on both the SKU- and the fixture-dimensions, was the answer. The proposed algorithm relied on iterative error-minimizing cuts to cluster the sales responses due to location. Empirical experiments showed that it worked fairly well with sparse response matrices, which were characteristic of retail environments lacking sufficient natural experimentation of SKU placement changes. One interesting problem we hope to tackle in future research is the determination of an optimal number of clusters. Although this problem is likely to be \(NP\)-hard, we hope to find a fast heuristic that can also be easily incorporated into the biclustering framework.

We also plan to tackle the full placement optimization problem in future research. The
full placement optimization model, when properly specified, can help automate the decision support of shelf placement—the PMO algorithm, in comparison, is semi-automatic, because it requires a store manager to propose the next placement plan before it can start optimizing. We conjecture that product demand can be separated into two parts: 1) intrinsic demand, and 2) impulse demand. Intrinsic demand occurs wherever the SKU is placed, whereas impulse demand depends on traffic and location. The placement optimization model—a mixed integer program with a nonlinear objective—resembles, to some extent, a facility location problem, for which a greedy solution can be developed.
Chapter 5

Case Study in Choice Forecasting

We define a customer service network as one that consists of a portfolio of facilities, possibly of varying degrees of quality. These individual facilities in the portfolio are in turn made up of units, where customer demand ultimately originates.

A good example of a customer service network is a hotel chain. A typical hotel operator would position a variety of brands, each of which may have a different value proposition, in the same geographic area. Within each brand, there are rooms of varying features and prices that cater to different customer segments. For example, Starwood operates the Sheraton brand as its flagship luxury hotel and resort accommodation, caters the Westin brand to convention participants and business travelers, markets the W brand toward a younger, hipper clientele, and so forth. Other examples of a customer service network include a restaurant group, a retail store chain, the branch network of a retail bank, among others.

We conducted this project in conjunction with a commercial partner that operated a customer service network, similar to the examples described above. Due to the confidentiality agreement, however, we are not at liberty to disclose the specific industry with which our partner is associated.

The main purpose of the project was to design and implement a causal model that would forecast unit-level demand for a pilot facility within the customer service network. A facility network, operated by our commercial partner, serves regions of varying demographic compositions. Customers pay a fee to access units of various features.
A typical sequence of events that lead to demand origination can be described as the following: A customer calls in to request a unit of certain feature, e.g., room size in the case of a hotel business. Depending on inventory availability, a call center staff member will then provide several alternative choices from the requested facility as well as its neighboring facilities—this is increasingly common practice of a theme park operator that also owns a variety of on-property resorts. Based on these recommendations, the customer either picks a unit or walks away. The sequence of events may span several call-in sessions. Detailed accounts of these first contact conversations should be maintained, so that a customer's entire choice set and eventual decision can be recorded and traced.

We found this a good opportunity to test the real-world effectiveness of demand forecasting with consumer choice models. To forecast demand for various units, our partner currently employs for each group a time series model, which is fairly limiting. For one thing, a time series model is blind in that it does not concern itself with why changes in the forecasted quantity occur but only the fact that it has occurred. It tends to attribute all such changes to seasonality, which makes it difficult to study and distinguish whether these changes are caused by a managerial policy or consumer behavior. A causal forecasting model, by allowing what-if scenarios to be assessed through the adjustment of input explanatory variables, can adequately address this difficulty.

Starting with the seminal paper by Guadagni and Little (1983), it is common for consumer choice researchers to test their models with industry data. For example, Bucklin and Gupta (1999) provide an overview from both the industry and academic perspectives on one widely used data source retrieved from universal product code (UPC) scanners. However, we feel most of the efforts have focused on advancing the choice theory (see for example, Gupta 1988, Tellis 1988, Han et al. 2001, Bucklin and Lattin 1991), while there does not appear to be a substantial amount of work devoted to the topic of bottom-up causal forecasting using consumer choice models. As defined by Talluri and van Ryzin (2004, Section 9.1.4.3), such forecasting is first performed at a granular level, e.g. individual choices, after which demand forecast can be obtained by aggregating these granular forecasts. We refer the reader to Louviere and Hensher (1983) for an example of such work.

In recent years, improvements in data collection, e.g. deployment of enterprise resource
planning (ERP) systems and customer relationship management (CRM) systems, wide adoption of electronic point-of-purchase systems and internet shopper tracking systems, coupled with a renewed enthusiasm for "customer-centric" operations, have created an urgency among many businesses to push for causal forecasting models that can be used in what-if tests on consumer responses. Outputs of these systems often form the bases of better managerial decisions and tactical recommendations on the operations of the customer service network.

Our objectives at the conclusion of this case study include: 1) designing and deploying a causal demand forecasting model for all units at a pilot facility, and 2) conducting demand forecasting tests with hold-out samples.

5.1 Causal Forecasting Model Formulation

A causal model in this context is one that predicts or explains unit-level demand as a function of explanatory variables (e.g. price of the unit, prices of other competing units in the choice set, seasonality indicators, and so forth). Successful forecasting depends on two components: 1) a volume forecast that indicates the weekly number of first contacts from each geographic region serviceable by the pilot facility, and 2) a choice forecast that predicts what fraction of these first contacts will choose to pay for a particular unit in his choice set. In the bottom-up fashion, the first component allows us to aggregate the granular forecasts obtained in the second component into a unit-level demand forecast.

A pilot facility was chosen for the study. Using historical sales data, we also hand-picked a few facilities that competed directly with the pilot. From the first contact database, we are also able to identify the primary regions that are serviceable by the pilot facility. Specific to this data set, we learned that customers that chose a particular size group typically exhibited different purchase or usage patterns from those that picked the other size group, size being one of the important features of a unit. Our approach therefore relies on estimating the parameters of the region-specific models for these two size groups separately. Figure 5.1 shows a block diagram of the process. We discuss each step in turn.
5.1.1 Unit Choice Probability by Region

For the choice forecast, we conjecture that a typical customer has a size requirement with some flexibility for upward or downward adjustment, subject to inventory availability. The customer must decide which facility to rent from, or alternatively he may walk away without selecting any of the current facility options. We illustrate this scenario in Figure 5.2, where we assume that a customer prefers a unit with a specific size but is willing to go for a smaller or larger alternative. Due to inventory status, the preferred size is not available in facility 3, and similarly the smaller size is not available in facility 4.

The Nested Logit Model

We would like to predict the customer’s choice by estimating the probability of each available option in his choice set. Mathematically, we use the nested logit model (See
Ben-Akiva and Lerman (1985, chap. 10) and Train (2002, chap. 4) for an overview of this model). Let $C$ be the super set of all possible facility and unit size combinations, denoted $(f, s)$, where $f$ represents the facility choice and $s$ represents the size choice. Each individual, subscripted by $n$, may be faced with a subset $C_n \subseteq C$, reflecting his or her personal preference as well as inventory status at the time of inquiry or purchase. Let $F$ denote the set of all facilities and $S_{nf}, f \in F$, denote the set of all unit sizes in customer $n$'s choice set. Then,

$$C_n = F \times \bigcup_{f \in F} S_{nf}.$$

We conjecture that the dominant factors influencing a customer's choice include 1) the size of the considered unit relative to the size requirement, 2) the location of the unit within a given facility, 3) the price, and 4) the facility itself. In the case of a hotel, 1) corresponds to the room size/bed configuration, 2) corresponds to the floor-level of the room, smoker status, and scenic views, 3) corresponds to the accommodation fee for the room, and 4) corresponds to the type and general location of this specific hotel in the network. A

---

1We use the term "individual" or "customer" to loosely refer to each "observation" of the choice data set. Repeated rentals are allowed.
corresponding utility function has the following deterministic part:

\[
V = \beta_1 R + \beta_2 (S - R)^+ - \beta_3 (S - R)^- + \beta_4^T X - \beta_5 p + \beta_6^T Y. \tag{5.1}
\]

Here, we omit all subscripts for a clearer illustration of the modeling idea. \(S\) denotes the size of the considered unit; \(R\) represents the customer's size requirement; \(X\) is a dummy vector that keeps track of the positions of the considered unit; \(p\) is the price that the customer will have to pay for a particular unit; \(Y\) is a dummy vector that represents the facility to which the considered unit belongs. In the special case where \(\beta_1 = \beta_2 = \beta_3 = \beta \geq 0\), the first three terms of (5.1) simply becomes \(\beta S\), which implies that a larger unit has a positive impact on the utility.

Recognizing that factors 1 and 2 (size and position) are specific to the choice of unit size, that factor 3 (price) applies to both the unit and the facility, and that factor 4 is facility-specific (superscript \(f\)), we define the following components of the utility function:

\[
V = V_1(s) + V_2(f,s) + V_3(f),
\]

where

\[
V_1(s) = \beta_1 R + \beta_2 (S - R)^+ - \beta_3 (S - R)^- + \beta_4^T X;
\]
\[
V_2(f,s) = -\beta_5 p;
\]
\[
V_3(f) = \beta_6^T Y.
\]

Let \(P_n(f,s)\) denote the probability that customer \(n\) chooses facility-unit combination \((f,s)\). Let \(F\) denote the set of all available facilities, and \(S(f)\) denote the set of available unit-sizes within facility \(f \in F\). Then according to the nested logit model and defining no purchase utility\(^2\) as \(V_0\), we obtain the following:

\[
P_n(f,s) = \frac{e^{V(f,s)/\lambda_i} \left( \sum_{(j,k) \in S(f)} e^{V(j,k)/\lambda_i} \right)^{\lambda_i - 1}}{\sum_{j \in F} \left( \sum_{(j,k) \in S(j)} e^{V(j,k)/\lambda_j} \right)^{\lambda_j} + e^{V_0}}, \tag{5.2}
\]
\[
P_n(\text{no purchase}) = \frac{1}{\sum_{j \in F} \left( \sum_{(j,k) \in S(j)} e^{V(j,k)/\lambda_j} \right)^{\lambda_j} + e^{V_0}}. \tag{5.3}
\]

\(^2\)While no purchase utility is usually set to 0, the forecast error can be minimized by scaling the no purchase utility.
CHAPTER 5. CASE STUDY IN CHOICE FORECASTING

Here, the parameter $\lambda_f$ measures the degree of independence in unobserved utility among all available rooms in facility $f$. When $\lambda_f = 1$, the utility of choosing an available unit in facility $f$ becomes completely independent from each other.

Maximum Likelihood Estimation

The parameters specified by utility function (5.1) and choice probabilities (5.2) and (5.3) can be estimated via maximizing the log-likelihood function for each region. For those records with missing region and/or requested size information, we compute an empirical region distribution and a requested size distribution, and geometrically weigh each occurrence before including it in the log-likelihood function.

5.1.2 Regional Volume Forecast

The second task involves forecasting the volume of first contacts for each region that is in-scope for the pilot facility. Here, we characterize all primary regions serviceable by the pilot facility plus an additional super-segment for all non-primary regions. Without detailed demographic data, it is difficult to construct another causal model that can complete this task. We have determined that the commonly used time series Holt-Winters model, which captures the seasonality effect, works well here.

The Holt-Winters Model

A brief description of the additive Holt-Winters algorithm can be found in Brockwell and Davis (2002, Section 9.3). We use a multiplicative version of the model here to forecast regional volume. A Holt-Winters model, additive or multiplicative, consists of three terms indexed by time $t$: 1) the average level $A_t$, 2) the trend $T_t$, and 3) the seasonality $S_t$. Let $L$ be period length, which we set to 52 weeks. Let $N_t$ be historical volume at time $t$. Let $\alpha$, $\beta$, $\gamma$ be given. For $t = L + 2, \ldots, t_h$, where $t_h$ is the last historical date, we obtain the following
sequences:

\[ A_t = \alpha \frac{N_t}{S_t-L} + (1 - \alpha) (A_{t-1} + T_{t-1}); \]
\[ T_t = \beta (A_t - A_{t-1}) + (1 - \beta)T_{t-1}; \]
\[ S_t = \gamma \frac{N_t}{A_t} + (1 - \gamma)S_{t-L}. \]

The sequence can be initialized via the following:

\[ A_{L+1} = N_{L+1}; \]
\[ T_{L+1} = \frac{N_{L+1} - N_1}{L}; \]
\[ S_t = \frac{N_t}{N_1 + T_{L+1}(t-1)}, \quad t = 1 \ldots L + 1. \]

The forecast \( \hat{N}_{h+t} \) \( (t = 1, 2, \ldots) \) can be computed using the following equation:

\[ \hat{N}_{h+t} = \left( A_{h_t} + T_{h_t} \right) S_{h+t-L}. \]

**Least Square Error Estimation**

The parameters \( \alpha, \beta \) and \( \gamma \) can be estimated by minimizing the square error. We apply the least square error procedure to the historical data (100 weeks' worth) for each region serviceable by the pilot facility.

### 5.1.3 Unit-Level Demand Forecast

In this section, we put the two components together to obtain a regional forecast for unit-level demand (see Ben-Akiva and Lerman 1985, chap. 5). An aggregate unit-level demand forecast can be obtained by summing up the regional forecasts. This is done separately for the two size groups that exhibit distinct customer behaviors.

**The Unit-Level Demand Forecast Model**

Let \( N \) denote the forecasted weekly volume of potential customers, and let \( N(f,s) \) be the expected number of those potential customers that would choose facility-unit combination
(f, s) ∈ C. Then for each (f, s) ∈ C and a given vector (S, R, X, p, Y) from (5.1) as input to the utility function, we have

\[ N(f, s) = \sum_{n=1}^{N} P_n(f, s | S, R, X, p, Y). \]

**Mean Correction**

The data sets for this study contained a large number of no purchase records, which would have biased the choice parameters with a fixed, typically 0, no purchase utility \( V_0 \). In (5.2) and (5.3), we correct for this bias by adjusting and accepting a \( V_0 \) that minimizes the square difference between mean historical unit-level demand and the mean forecast over the prior 100 weeks for each region, serviceable by the pilot facility.

**5.2 Empirical Implementation**

The empirical implementation of the choice forecasting model follows these steps: first, we fit the choice model parameters with customer preference data at first contact; next, we forecast the arrival volume of first contact customers; and last, we forecast the unit-level demand of the pilot facility by leveraging results from both the choice model and the volume forecast.

Our commercial partner provided us with the following data sets for this study:

1. Size preferences and other basic demographic information at first contact,
2. Historical inventory and pricing records for all units, and
3. Customer purchase decisions (e.g., the unit that was ultimately chosen and paid for).

Due to the presence of a nondisclosure agreement, sensitive information will be disguised as we discuss the data and results here.

For the chosen pilot facility, we also selected 4 other competing facilities as potential alternatives in a customer’s nested choice set (please refer to Figure 5.2 for details). Out of 21 regions, we identified 3 primary regions, coded as Region 10, Region 15 and Region 16, for the pilot facility, and aggregated all other regions together into a super-region,
henceforth referred to simply as “other regions.” There are three unit positions, denoted Position 1–3, which are determined, based on business rules, to be relevant for Size Group 1 (SG1), but not so for Size Group 2 (SG2).

5.2.1 Estimated Choice Parameters

As discussed in the methodology section, we fitted a nested logit model to 100 weeks’ worth of customer preference data at first contact. While estimating the choice parameters for each region serviceable by the pilot facility, we ran into some numerical instabilities in the maximum likelihood procedure. As a result, we imposed upper and lower bounds for each choice parameter\(^3\). In some cases, these bounds may appear too limiting—however, their values were determined after trial and error using the available computing resources.

Table 5.1 exhibits the bounds along with the fitted parameters. The model results reveal demographic differences of various regions:

- For the positions of Size Group 1, Regions 10 and 15 favor Position 2 over either Position 1 or Position 3, whereas the rest of the regions favor Position 3 over Position 1 or Position 2;

- For facility choice, Region 10 most favors the competing facility #4 for Size Group 1 and the pilot facility for Size Group 2; Region 15 most favors the competing facility #4 for both size groups; Region 16 most favors the pilot facility for Size Group 1 and the competing facility #4 for Size Group 2; and all other regions favor the competing facility #3 for Size Group 1 and the competing facility #2 for Size Group 2.

Further adjustments are possible through a more fine-tuned set of bounds on the choice model parameters. For example, the price parameter shows up counter-intuitively as positive numbers a few times. We believe this may be due to the fact that price is confounded with size. For example, a larger room size in a hotel or a bigger party/table size at a restaurant, which is usually more preferred to a customer, almost always comes with a higher price. This can be corrected by forcing the price parameter to be negative. We

---

\(^3\)Due to the existing business rule, the position of a unit (Position 1–3) only applies to Size Group 1. Therefore, the upper and lower bounds for the dummy position parameters are set to 0 for Size Group 2.
### Table 5.1: Estimated Choice Parameters

<table>
<thead>
<tr>
<th>Choice Parameters</th>
<th>Bounds on SG1</th>
<th>Bounds on SG2</th>
<th>Region 10</th>
<th>Region 15</th>
<th>Region 16</th>
<th>Other Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Upper</td>
<td>Lower Upper</td>
<td>SG1</td>
<td>SG2</td>
<td>SG1</td>
<td>SG2</td>
</tr>
<tr>
<td>Neg. log like.</td>
<td>-15.0000 15.0000</td>
<td>-15.0000 15.0000</td>
<td>787</td>
<td>1,162</td>
<td>869</td>
<td>1,786</td>
</tr>
<tr>
<td>R</td>
<td>-15.0000 15.0000</td>
<td>-15.0000 15.0000</td>
<td>15.0000</td>
<td>7.1630</td>
<td>15.0000</td>
<td>9.2598</td>
</tr>
<tr>
<td>(S - R)*</td>
<td>-15.0000 15.0000</td>
<td>-15.0000 15.0000</td>
<td>-15.0000</td>
<td>-15.0000</td>
<td>14.9634</td>
<td>-15.0000</td>
</tr>
<tr>
<td>(S - R)*</td>
<td>-15.0000 15.0000</td>
<td>-15.0000 15.0000</td>
<td>-15.0000</td>
<td>-15.0000</td>
<td>-15.0000</td>
<td>-15.0000</td>
</tr>
<tr>
<td>(X_{position 1})</td>
<td>-5.0000 5.0000</td>
<td>-</td>
<td>-</td>
<td>0.0224</td>
<td>-</td>
<td>0.0836</td>
</tr>
<tr>
<td>(X_{position 2})</td>
<td>-5.0000 5.0000</td>
<td>-</td>
<td>-</td>
<td>0.4592</td>
<td>-</td>
<td>0.2316</td>
</tr>
<tr>
<td>(X_{position 3})</td>
<td>-5.0000 5.0000</td>
<td>-</td>
<td>-</td>
<td>0.1670</td>
<td>-</td>
<td>0.0945</td>
</tr>
<tr>
<td>(p)</td>
<td>-1.0000 1.0000</td>
<td>-1.0000 1.0000</td>
<td>1.0000</td>
<td>0.0780</td>
<td>0.2252</td>
<td>0.0310</td>
</tr>
<tr>
<td>(y_{pilot facility})</td>
<td>-5.0000 5.0000</td>
<td>-5.0000 5.0000</td>
<td>-3.5724</td>
<td>-2.5033</td>
<td>-3.5424</td>
<td>-3.5783</td>
</tr>
<tr>
<td>(y_{comp. facility 1})</td>
<td>-5.0000 5.0000</td>
<td>-5.0000 5.0000</td>
<td>-4.3962</td>
<td>-3.9580</td>
<td>-4.5591</td>
<td>-5.0000</td>
</tr>
<tr>
<td>(y_{comp. facility 2})</td>
<td>-5.0000 5.0000</td>
<td>-5.0000 5.0000</td>
<td>-4.9861</td>
<td>-3.7739</td>
<td>-5.0000</td>
<td>-4.4086</td>
</tr>
<tr>
<td>(y_{comp. facility 3})</td>
<td>-5.0000 5.0000</td>
<td>-5.0000 5.0000</td>
<td>-5.0000</td>
<td>-4.1680</td>
<td>-5.0000</td>
<td>-4.4549</td>
</tr>
<tr>
<td>(y_{comp. facility 4})</td>
<td>-5.0000 5.0000</td>
<td>-5.0000 5.0000</td>
<td>-3.0898</td>
<td>-2.7146</td>
<td>-1.0202</td>
<td>-1.4408</td>
</tr>
<tr>
<td>(\lambda_{pilot facility})</td>
<td>0.1000 1.0000</td>
<td>0.1000 1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>(\lambda_{comp. facility 1})</td>
<td>0.1000 1.0000</td>
<td>0.1000 1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.7266</td>
</tr>
<tr>
<td>(\lambda_{comp. facility 2})</td>
<td>0.1000 1.0000</td>
<td>0.1000 1.0000</td>
<td>1.0000</td>
<td>0.8207</td>
<td>0.7523</td>
<td>1.0000</td>
</tr>
<tr>
<td>(\lambda_{comp. facility 3})</td>
<td>0.1000 1.0000</td>
<td>0.1000 1.0000</td>
<td>0.2554</td>
<td>0.1000</td>
<td>0.9414</td>
<td>1.0000</td>
</tr>
<tr>
<td>(\lambda_{comp. facility 4})</td>
<td>0.1000 1.0000</td>
<td>0.1000 1.0000</td>
<td>0.1000</td>
<td>0.3547</td>
<td>0.1000</td>
<td>0.7888</td>
</tr>
<tr>
<td>Other Regions</td>
<td></td>
<td></td>
<td>18,208</td>
<td>25,264</td>
<td>15.0000</td>
<td>15.0000</td>
</tr>
</tbody>
</table>
decided not to implement these even more limiting bounds because they tended to de­
crease the model fit (by theory of constrained optimization) and disguise the confounding
interactions between some variables.

5.2.2 Sample Volume Forecast Results

First contact volumes were forecasted for each region using the Holt-Winters model. Hold­
ing out a 40-week sample, we fitted the model to the previous five years of first contact
history, and smoothing constants were chosen to minimize the error in the last 100 weeks
of data. The results are summarized below in Figure 5.3. Year numbers were removed to
protect data anonymity.

The volume forecast model predicts the volume and seasonal variations of back-tested
data fairly well—this is especially so for “other regions,” where a bigger volume helps
smooth out the forecast errors.
5.2.3 Sample Demand Forecast Results

We applied the causal forecasting model to a series of rolling 4-week holdout samples. The batch process starts by holding out a 40-week sample from the last available date and projecting the first 4 weeks of that sample using the remaining historical data. It then holds out a 36-week sample, of which it also projects the first 4 weeks, and so forth. The process ends with holding out and projecting a 4-week sample from the last available date. These 4-week projections are then pieced together into a 40-week forecast.

Figure 5.4 shows the unit-level demand forecast results aggregated under the two distinct size groups (SG1 & SG2) as well as various subgroups within each. The demand forecasts follow the trends of actual back-tested unit demand pretty well, although the forecasts are not as volatile as the actual demand. To quantify the accuracy of the causal forecasting model, we introduce a few measures in the following discussion.

Measures of Model Accuracy

The accuracy of the model can be measured through mean absolute deviation (MAD) or mean relative deviation (MRD) between projected and actual demand at different levels of aggregation, e.g. weekly, monthly, by facility or total demand. In the evaluation stage, we feed historical data into the model to generate would-be forecasts and associated errors; in the validation stage, we compare forecasts from the model with actual realizations of customer demand.

Measures of model accuracy over the 40-week horizon for the unit-level demand forecasts are summarized in Table 5.2. We note that each MAD measure is generally on the same order as the square root of the forecast mean—approximately an "ideal" forecast would have a MAD close to the square root of the forecast mean. This suggests that the model accuracy is well under control.

5.3 Conclusions and Future Research

We developed a prototype causal forecasting model for a customer service network. The model has two main components: 1) a first contact volume forecast, which uses a time-series
CHAPTER 5. CASE STUDY IN CHOICE FORECASTING

Figure 5.4: Demand Forecast by Size Group (4-Week Rolling Holdout)
CHAPTER 5. CASE STUDY IN CHOICE FORECASTING

<table>
<thead>
<tr>
<th>Size Group</th>
<th>Actual Mean</th>
<th>Forecast Mean</th>
<th>MAD</th>
<th>Sqrt(Fcst Mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG 1 (Aggregate)</td>
<td>13.38</td>
<td>12.62</td>
<td>4.09</td>
<td>3.55</td>
</tr>
<tr>
<td>SG 2 (Aggregate)</td>
<td>18.85</td>
<td>21.13</td>
<td>6.97</td>
<td>4.60</td>
</tr>
<tr>
<td>SG 1 (Position 3 or Null)</td>
<td>6.08</td>
<td>4.96</td>
<td>2.70</td>
<td>2.23</td>
</tr>
<tr>
<td>SG 1 (Position 1 or 2)</td>
<td>7.30</td>
<td>7.66</td>
<td>2.43</td>
<td>2.77</td>
</tr>
<tr>
<td>SG 2 (Subgroup A)</td>
<td>6.35</td>
<td>7.10</td>
<td>2.81</td>
<td>2.67</td>
</tr>
<tr>
<td>SG 2 (Subgroup B)</td>
<td>17.08</td>
<td>16.76</td>
<td>5.33</td>
<td>4.09</td>
</tr>
<tr>
<td>SG 2 (Subgroup C)</td>
<td>8.85</td>
<td>10.40</td>
<td>3.52</td>
<td>3.23</td>
</tr>
<tr>
<td>SG 2 (Subgroup D)</td>
<td>20.78</td>
<td>21.59</td>
<td>6.83</td>
<td>4.65</td>
</tr>
<tr>
<td>SG 2 (Subgroup E)</td>
<td>9.85</td>
<td>11.38</td>
<td>3.73</td>
<td>3.37</td>
</tr>
<tr>
<td>SG 2 (Subgroup F)</td>
<td>1.60</td>
<td>0.78</td>
<td>1.25</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Table 5.2: Measures of Model Accuracy by Size Group

model to predict the total number of regional first contacts, accounting for seasonality and trends in the contact volume, and 2) a choice model forecast, which predicts the probability that a contact will pay for a particular unit in a particular facility as a function of his size needs, the availability of units in each facility and the price of those units.

We tested the model on a series of rolling hold-out samples. The statistical errors produced by the model are generally within the expected accuracy range given the volume of demand for the pilot facility at various levels of aggregation. The pilot shows that reasonable forecasts can be constructed using a combination of volume and choice model forecasting methods over commercial data.

It is important to keep in mind that forecasting is a continuous improvement process in revenue management applications. Models are re-estimated, modified and fine-tuned on an on-going basis over time to correct for idiosyncratic features in the data and changes in market conditions. This test represents a first cut at this process—and one involving a quite intricate and sophisticated forecasting system with multiple components, each of which was untested on this particular set of commercial data. The fact that it works as well as it does “right out of the box” is, in our view, both promising and a validation of the overall modeling approach and architecture of the system.

The following refinements are likely to improve the forecasts further:
1. Segmenting the choice by different seasons to capture differences in price sensitivity throughout the year;

2. Segmenting the size requirement distribution by season to better reflect the changing mix of customer types (behavioral, attitudinal and demographic) throughout the year;

3. Incorporating CRM data into the choice model estimation, which will eliminate a considerable degree of guess work surrounding the choice sets presented to each customer;

4. Experimenting on a sample of transactions to better separate the price effect from preferences of facility/unit types or sizes.

Still, one must bear in mind that forecasts will never be perfect. The future is at some level inherently unpredictable. What a forecasting system can help at best is separating the predictable from the unpredictable. The residual uncertainty has to be managed in terms of pricing and capacity allocation decisions by using optimization techniques that account for uncertainty in outcomes (e.g. stochastic optimization). These techniques are used in most commercial revenue management systems and require both point forecasts and error estimates from a forecasting system.
Chapter 6

Conclusions

We used a variety of modeling techniques, ranging from new stylized models to empirical implementations of known methods, to examine each of the four unique topics on consumer retail in this dissertation. In the process, we hoped to explore and demonstrate a niche area where close integration of marketing and operations was possible.

For the first topic (Chapter 2), a series of stylized theoretical models were built to investigate the incentives for coordinating product positioning and pricing of horizontally differentiated products in the context of a vertical trading relationship between a retailer and multiple suppliers. By comparing the channel profits of the centralized, decentralized and monopolist supplier systems, we observed a loss in channel profit when the supply chain moved away from a centralized system. When total production costs were low, channel profit of the monopolist supplier system dominated that of the decentralized system. The loss of channel profit in this case could mainly be attributable to the difference in the production costs of the offered products. When production costs were extremely high, on the other hand, the monopolist supplier system achieved the same channel profit as that of the decentralized system. However, the benefit of a monopolist supplier system became questionable in this case, as double marginalization became the main cause of channel profit loss.

We used stylized theoretical models to examine the effect of pricing and price communication in a dynamic setting for the second topic (Chapter 3). Price communication, which incurs a fixed cost and serves as an additional control to pricing, accelerates the
CHAPTER 6. CONCLUSIONS

Learning rate of new retail prices and expands the market size of the product. In addition to those policies that lead to a steady state reference price, we discover cases where a cyclic pricing and price communication strategy may be optimal. In the cyclic cases, the firm has incentives to communicate a retail price higher than the current reference price to quickly build up the reference price level, and to alternate this tactic with the posting of a series of lower uncommunicated retail prices that harvest the benefits of a high reference price.

The third topic (Chapter 4) was a part of an applied research project, which sought to examine the impact of physical locations on retail sales. We designed and implemented a network-based Plan Modification Optimization (PMO) system for our commercial partner as a part of the project. This experience allowed us to understand various business constraints and requirements that were bypassed in the design of the PMO system. In our efforts to create a full merchandise placement optimization system, we hypothesized that the demand of every retail product consisted of two components—intrinsic and impulse. We proposed some heuristics to solve the $\mathcal{NP}$-hard placement optimization problem, whose formulation was likened to that of a facility location problem. Devising a clustering framework on both the product types and the location types, we are able to forecast the impact of a future move that was not yet observed historically, and effectively extend the application of PMO to all future moves.

For the last topic (Chapter 5), we took on a task to design and deploy a causal forecasting pilot for unit-level demand at a pilot facility in a customer service network. The forecasting model comprised a time series component for regional volume forecasts and a discrete choice component for individual preferences on the units in the pilot facility. This applied case study allowed us to observe the effectiveness of consumer choice forecasting when it was applied to real commercial data. We were also able to propose various ways to correct for biases and deviations that impacted the accuracy of the forecast model.

The four topics we have proposed for this dissertation are currently at various degrees of completion. However, we are already contemplating a few directions of future work. For the equilibria characterized in the first topic, we wonder whether the two-supplier decentralized system can be generalized to multiple suppliers and how this extension will affect the results. For placement optimization problem in the third topic, we ask whether
there are other, perhaps better, heuristics and clustering/testing designs. We are certain that more interesting problems will come up as we move toward their completion.
Bibliography


Appendix A

Technical Details of Coordinating Vertical Partnerships for Horizontally Differentiated Products

A.1 Technical Details of Decentralized System

A.1.1 Retailer’s Problem

Solution to (P1):

First observe that the objective function is concave, and the constraint is linear. Therefore, Karush-Kuhn-Tucker (KKT) conditions are both sufficient and necessary. The Lagrangian is

\[ L = (p_1 - w_1) \frac{2(v - p_1)}{\theta} + (p_2 - w_2) \frac{2(v - p_2)}{\theta} + \lambda \left( \frac{1}{2} - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} \right). \]

First order conditions yield the following:

\[ p_1 = \frac{v + w_1}{2} + \frac{\lambda}{4}; \quad p_2 = \frac{v + w_2}{2} + \frac{\lambda}{4}. \]  \hspace{1cm} (A.1)

Assume strict complementarity holds. We have the following 2 cases:

1. \( \lambda = 0 \) \( \iff \) The separable demand solution (S1) is optimal and dominates the barely
touching solution (S2).

\[
\Rightarrow \frac{1}{2} - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} > 0 \quad (A.2)
\]

\[
(A.1) \Rightarrow p_1^* = \frac{v + w_1}{2}; \quad p_2^* = \frac{v + w_2}{2} \quad (A.3)
\]

The corresponding retail profit is

\[
\pi^* = \frac{(v - w_1)^2}{2\theta} + \frac{(v - w_2)^2}{2\theta}. \quad (A.4)
\]

Plug (A.3) into (A.2) \( v - \frac{w_1 + w_2}{2} < \frac{\theta}{2}. \)

2. \( \lambda > 0 \iff \) The barely touching solution (S2) is optimal and dominates the separable demand solution (S1).

\[
\Rightarrow \frac{1}{2} - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} = 0 \quad (A.5)
\]

\[
(A.1), (A.5) \Rightarrow \lambda = 2v - \theta - (w_1 + w_2) > 0 \iff w_1 + w_2 < 2v - \theta;
\]

\[
p_1^* = v + \frac{w_1}{4} - \frac{w_2}{4} - \frac{\theta}{4}; \quad p_2^* = v + \frac{w_2}{4} - \frac{w_1}{4} - \frac{\theta}{4}. \quad (A.6)
\]

The corresponding retail profit is

\[
\pi^* = v + \frac{(w_1 - w_2)^2}{4\theta} - \frac{w_1 + w_2}{2} - \frac{\theta}{4}. \quad (A.7)
\]

Weak complementarity (\( \lambda = 0, \frac{1}{2} - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} = 0 \)) holds at the borderline of the two cases. In this case, \( v - \frac{w_1 + w_2}{2} = \frac{\theta}{2}. \)

Solution to (P2):

It can be verified that the objective function is concave, and the constraint is linear. Therefore, KKT conditions are both sufficient and necessary. The Lagrangian is

\[
\mathcal{L} = (p_1 - w_1)\left(\frac{1}{2} + \frac{p_2 - p_1}{\theta}\right) + (p_2 - w_2)\left(\frac{1}{2} + \frac{p_1 - p_2}{\theta}\right) + \lambda \left(\frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} - \frac{1}{2}\right).
\]

First order conditions yield

\[
p_1 - p_2 = \frac{\theta}{4} + \frac{w_1 - w_2}{2} - \frac{\lambda}{2}; \quad p_1 - p_2 = -\frac{\theta}{4} + \frac{w_1 - w_2}{2} + \frac{\lambda}{2}.
\]
Clearly we want

\[ \frac{\theta}{4} - \frac{\lambda}{2} = \frac{\theta}{4} + \frac{\lambda}{2} \Leftrightarrow \lambda = \frac{\theta}{2} > 0. \]

Since \( \lambda > 0 \), the inequality constraint in (P2) is always active, i.e. the barely touching solution (S2) will always dominate the overlapping demand solution (S3).

**Solution to (P3):**

First observe that the objective function is the same as in (P1) and is thus concave, and the constraint is linear. Therefore, KKT conditions are both sufficient and necessary. The Lagrangian is:

\[ \mathcal{L} = (p_1 - w_1) \frac{2(v - p_1)}{\theta} + (p_2 - w_2) \frac{2(v - p_2)}{\theta} + \lambda \left( |l_1 - l_2| - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} \right). \]

First order conditions yield the following, just as in (P1):

\[ p_1 = \frac{v + w_1}{2} + \frac{\lambda}{4}; \quad p_2 = \frac{v + w_2}{2} + \frac{\lambda}{4}. \]  \hspace{1cm} (A.8)

Assume strict complementarity holds. We have the following 2 cases:

1. \( \lambda = 0 \Leftrightarrow \) The separable demand solution (AS1) is optimal and dominates the solution where one end barely touches and the other end does not touch (AS2).

\[ \Rightarrow |l_1 - l_2| - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} > 0 \]  \hspace{1cm} (A.9)

\[ (A.8) \Rightarrow p_1^* = \frac{v + w_1}{2}; \quad p_2^* = \frac{v + w_2}{2}. \]  \hspace{1cm} (A.10)

The corresponding retail profit is

\[ \pi^* = \frac{(v - w_1)^2}{2\theta} + \frac{(v - w_2)^2}{2\theta}. \]

Plug (A.10) into (A.9) \( \Rightarrow \)

\[ v - \frac{w_1 + w_2}{2} < \theta |l_1 - l_2|. \]  \hspace{1cm} (A.11)
2. \( \lambda > 0 \Leftrightarrow \) The solution where one end barely touches and the other end does not touch (AS2) is optimal and dominates the separable demand solution (AS1).

\[
\Rightarrow |l_1 - l_2| - \frac{v - p_1}{\theta} = 0 \quad (A.12)
\]

\[
(A.8), (A.12) \Rightarrow \lambda = 2v - 2\theta |l_1 - l_2| - (w_1 + w_2) > 0 \iff w_1 + w_2 < 2v - 2\theta |l_1 - l_2| \quad (A.13)
\]

\[
p_1^* = v + \frac{w_1}{4} - \frac{w_2}{4} - \frac{\theta}{2} |l_1 - l_2|; \quad p_2^* = v + \frac{w_2}{4} - \frac{w_1}{4} - \frac{\theta}{2} |l_1 - l_2|.
\]

(A.14)

The corresponding retail profit is:

\[
\pi^* = 2v |l_1 - l_2| - (w_1 + w_2) |l_1 - l_2| - \theta |l_1 - l_2|^2 + \frac{(w_1 - w_2)^2}{4\theta}.
\]

Weak complementarity \((\lambda = 0, |l_1 - l_2| - \frac{v - p_1}{\theta} = \frac{v - p_2}{\theta} = 0)\) holds at the borderline of the two cases. In this case, \(v - \frac{w_1 + w_2}{2} = \theta |l_1 - l_2|\).

**Solution to (P4):**

It is easy to verify that the objective function is concave, and the constraints are linear. As a result, KKT conditions are both sufficient and necessary. Let \(\mu_1\) be the lagrange multiplier associated with the lower bound and \(\mu_2\) be lagrange multiplier associated with the upper bound. The Lagrangian is:

\[
L = (p_1 - w_1) \left( \frac{1}{2} |l_1 - l_2| + \frac{p_2 - 3p_1}{2\theta} + \frac{v}{\theta} \right) + (p_2 - w_2) \left( \frac{1}{2} |l_1 - l_2| + \frac{p_1 - 3p_2}{2\theta} + \frac{v}{\theta} \right) + \mu_1 \left( \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} - |l_1 - l_2| \right) + \mu_2 \left( 1 - |l_1 - l_2| - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} \right).
\]

First order conditions yield the following

\[
\begin{align*}
\theta |l_1 - l_2| + 2p_2 - 6p_1 + 2v + 3w_1 - w_2 - 2\mu_1 + 2\mu_2 &= 0 \\
\theta |l_1 - l_2| + 2p_1 - 6p_2 + 2v + 3w_2 - w_1 - 2\mu_1 + 2\mu_2 &= 0 \\
\Rightarrow p_1 - p_2 &= \frac{w_1 - w_2}{2}. \quad (A.16)
\end{align*}
\]
Assume strict complementarity holds. We have the following:

\[ \mu_1 > 0 \Rightarrow \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} - |l_1 - l_2| = 0; \] (A.17)

\[ \mu_1 = 0 \Rightarrow \frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} - |l_1 - l_2| > 0; \] (A.18)

\[ \mu_2 > 0 \Rightarrow 1 - |l_1 - l_2| - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} = 0; \] (A.19)

\[ \mu_2 = 0 \Rightarrow 1 - |l_1 - l_2| - \frac{v - p_1}{\theta} - \frac{v - p_2}{\theta} > 0. \] (A.20)

Note first that the upper and lower bounds cannot both be simultaneously active when product positions are asymmetric. This implies that \( \mu_1 \) and \( \mu_2 \) cannot both be \( > 0 \). We have the following 3 cases:

1. \( \mu_1 > 0, \mu_2 = 0 \Leftrightarrow \) The solution where one end of the product demands barely touches while the other end does not touch (AS2) is optimal and dominates (AS3) and (AS4).

\[ (A.16), (A.17) \Rightarrow p_1^* = v + \frac{w_1 - w_2}{4} - \frac{\theta}{2}l_1 - l_2; \quad p_2^* = v - \frac{w_1 - w_2}{4} - \frac{\theta}{2}l_1 - l_2. \] (A.21)

The corresponding retail profit is

\[ \pi^* = (2v - w_1 - w_2)|l_1 - l_2| - \theta|l_1 - l_2|^2 + \frac{(w_1 - w_2)^2}{4\theta}. \]

As a sanity check, observe that the results obtained above are identical to (A.14) when solving (P3).

Check (A.20) \( \Rightarrow 1 - |l_1 - l_2| - \frac{v - p_1^*}{\theta} - \frac{v - p_2^*}{\theta} = 1 - 2|l_1 - l_2| > 0, \) by asymmetry.

Plug (A.21) into (A.15) \( \Rightarrow 2\mu_1 = 3\theta|l_1 - l_2| + w_1 + w_2 - 2v > 0 \Leftrightarrow v - \frac{w_1 + w_2}{2} < \frac{3}{2}\theta|l_1 - l_2|.

Intersect above with (A.13) \( \Rightarrow \theta|l_1 - l_2| < v - \frac{w_1 + w_2}{2} < \frac{3}{2}\theta|l_1 - l_2|.

2. \( \mu_1 = 0, \mu_2 = 0 \Leftrightarrow \) The solution where one end overlaps while the other end does not touch (AS3) is optimal and dominates (AS2) and (AS4).

\[ (A.15) \Rightarrow p_1^* = \frac{v + w_1}{2} + \frac{\theta}{4}|l_1 - l_2|; \quad p_2^* = \frac{v + w_2}{2} + \frac{\theta}{4}|l_1 - l_2|. \] (A.22)

The corresponding optimal retail profit is

\[ \pi^* = \frac{1}{4\theta} \left(v - w_1 + \frac{\theta}{2}|l_1 - l_2|\right)^2 + \frac{1}{4\theta} \left(v - w_2 + \frac{\theta}{2}|l_1 - l_2|\right)^2 + \frac{1}{8\theta}(w_1 - w_2)^2. \]
Plug (A.22) into (A.18), (A.20) ⇒ $\frac{3}{2}\theta|l_1 - l_2| < v - \frac{w_1 + w_2}{2} < \theta - \frac{1}{2}\theta|l_1 - l_2|$. Observe that (A.13) is not violated.

3. $\mu_1 = 0, \mu_2 > 0$ ⇔ The solution where one end overlaps while the other end barely touches (AS4) is optimal and dominates (AS2) and (AS3).

\begin{align*}
(A.16), (A.19) \Rightarrow p_1^* &= \frac{w_1 - w_2}{4} + v + \frac{\theta}{2}|l_1 - l_2| - \frac{\theta}{2}; p_2^* = \frac{w_2 - w_1}{4} + v + \frac{\theta}{2}|l_1 - l_2| - \frac{\theta}{2}.
\end{align*}

The corresponding retail profit is
\begin{equation}
\pi^* = v - \frac{\theta}{2} + \frac{\theta}{2}|l_1 - l_2| - \frac{w_1 + w_2}{2} + \frac{(w_1 - w_2)^2}{4\theta}.
\end{equation}

Plug (A.23) into (A.15) ⇒ $2\mu_2 = 2v - 2\theta + \theta|l_1 - l_2| - w_1 - w_2 > 0$ ⇔ $v - \frac{w_1 + w_2}{2} > \theta - \frac{1}{2}\theta|l_1 - l_2|$. Observe that (A.13) is not violated.

Weak complementarity holds at the borderline of (AS2) and (AS3), i.e. $v - \frac{w_1 + w_2}{2} = \frac{3}{2}\theta|l_1 - l_2|$, and that of (AS3) and (AS4), i.e. $v - \frac{w_1 + w_2}{2} = \theta - \frac{1}{2}\theta|l_1 - l_2|$. 

Solution to (P5):

First observe that the objective function is the same as in (P2) and is thus concave, and the constraint is linear. Therefore, KKT conditions are both sufficient and necessary. The Lagrangian is
\begin{equation}
L = (p_1 - w_1)\left(\frac{1}{2} + \frac{p_2 - p_1}{\theta}\right) + (p_2 - w_2)\left(\frac{1}{2} + \frac{p_1 - p_2}{\theta}\right) + \lambda\left(\frac{v - p_1}{\theta} + \frac{v - p_2}{\theta} - 1 + |l_1 - l_2|\right).
\end{equation}

First order conditions yield
\begin{align*}
p_1 - p_2 &= \frac{\theta}{4} + \frac{w_1 - w_2}{2} - \frac{\lambda}{2}; p_1 - p_2 = -\frac{\theta}{4} + \frac{w_1 - w_2}{2} + \frac{\lambda}{2}.
\end{align*}

Clearly we want
\begin{equation}
\frac{\theta}{4} - \frac{\lambda}{2} = -\frac{\theta}{4} + \frac{\lambda}{2} \iff \lambda = \frac{\theta}{2} > 0.
\end{equation}

Since $\lambda > 0$, the inequality constraint in (P5) is always active, i.e. the solution where one end overlaps while the other end barely touches (AS4) will always dominate solution where both ends overlap (AS5).
APPENDIX A. TECHNICAL DETAILS OF COORDINATING VERTICAL PARTNERSHIPS  
FOR HORIZONTALLY DIFFERENTIATED PRODUCTS  

Proof of Proposition 2.1:

For symmetrically positioned products, solution to (P2) establishes that the overlapping demand solution (S3) is always dominated. The comparison is thus between separable demand (S1) and barely touching (S2). Solution to (P1) establishes the segments in the proposition statement and determines when (S1) dominates (S2) and vice versa.

For asymmetrically positioned products, solution to (P5) establishes that (AS5), the case where product demands overlap on both ends, is always dominated—we need only compare (AS1) through (AS4). Condition (A.13) links the optimal solutions of (P3) with those of (P4). When (A.11) is satisfied, (AS1) dominates (AS2) from results in (P3), and (AS2) dominates (AS3) and (AS4) from results in (P4) since (A.11) is a subset of the condition that ensures (AS2)'s domination over (AS3) and (AS4) in (P4). Conversely when (A.13) is satisfied, (AS1) is dominated by (AS2), which may itself be optimal or be dominated by (AS3) or (AS4) as described by the segments in (P4).

A.1.2 Suppliers' Game

Solution to (SP13):

The objective functions are concave. First order conditions yield the following:

\[ w_1^* = \frac{v + c_1}{2}; \quad w_2^* = \frac{v + c_2}{2}. \]  \hspace{1cm} (A.25)

And therefore,

\[ w_1^* + w_2^* = v + \frac{c_1 + c_2}{2}; \]  \hspace{1cm} (A.26)

\[ \pi_1^* = \frac{(v - c_1)^2}{4\theta}; \quad \pi_2^* = \frac{(v - c_2)^2}{4\theta}. \]  \hspace{1cm} (A.27)

Solution to (SP26):

The objective functions are concave. First order conditions yield the following:

\[ w_1^* = \frac{\theta + c_1 + \bar{w}_2}{2}; \quad w_2^* = \frac{\theta + c_2 + \bar{w}_1}{2}. \]
Solving jointly,

\[ w_1^* = \theta + \frac{2}{3} c_1 + \frac{1}{3} c_2; \quad w_2^* = \theta + \frac{1}{3} c_1 + \frac{2}{3} c_2; \]  
\[ \Rightarrow w_1^* + w_2^* = 2\theta + c_1 + c_2; \]  
\[ \pi_1^* = 2\theta \left( \frac{1}{2} + \frac{c_2 - c_1}{6\theta} \right)^2; \quad \pi_2^* = 2\theta \left( \frac{1}{2} + \frac{c_1 - c_2}{6\theta} \right)^2. \]  

(A.28)  
(A.29)  
(A.30)

Solution to (SP4):

The objective functions are concave. First order conditions yield the following:

\[ w_1^* = \frac{2\theta|l_1 - l_2| + c_1 + w_2}{2}; \quad w_2^* = \frac{2\theta|l_1 - l_2| + c_2 + w_1}{2}. \]

Solving jointly,

\[ w_1^* = 2\theta|l_1 - l_2| + \frac{2}{3} c_1 + \frac{1}{3} c_2; \quad w_2^* = 2\theta|l_1 - l_2| + \frac{1}{3} c_1 + \frac{2}{3} c_2; \]  
\[ \Rightarrow w_1^* + w_2^* = 4\theta|l_1 - l_2| + c_1 + c_2; \]  
\[ \pi_1^* = 2\theta \left( |l_1 - l_2| + \frac{c_2 - c_1}{6\theta} \right)^2; \quad \pi_2^* = 2\theta \left( |l_1 - l_2| + \frac{c_1 - c_2}{6\theta} \right)^2. \]  

(A.31)  
(A.32)  
(A.33)

Solution to (SP5):

The objective functions are concave. First order conditions yield the following:

\[ w_1^* = \frac{\theta|l_1 - l_2| + 2v + w_2 + 3c_1}{6}; \quad w_2^* = \frac{\theta|l_1 - l_2| + 2v + w_1 + 3c_2}{6}. \]

Solving jointly,

\[ w_1^* = \frac{1}{5} \theta|l_1 - l_2| + \frac{2}{5} v + \frac{18}{35} c_1 + \frac{3}{35} c_2; \quad w_2^* = \frac{1}{5} \theta|l_1 - l_2| + \frac{2}{5} v + \frac{18}{35} c_2 + \frac{3}{35} c_1; \]  
\[ \Rightarrow w_1^* + w_2^* = \frac{2}{5} \theta|l_1 - l_2| + \frac{4}{5} v + \frac{3}{5} (c_1 + c_2); \]  
\[ \pi_1^* = \frac{3}{4900\theta} (7\theta|l_1 - l_2| + 14v - 17c_1 + 3c_2)^2; \quad \pi_2^* = \frac{3}{4900\theta} (7\theta|l_1 - l_2| + 14v - 17c_2 + 3c_1)^2. \]  

(A.34)  
(A.35)  
(A.36)

Proof of Lemma 2.1:

Region 1 ↔ region 2. Consider supplier 1’s problem in region 2. Rewrite the right boundary of region 2 as \( w_2 = 2v - \theta - w_1 \). Plugging it into the objective function in region 2 (SP26)
yields the objective function in region 1 (SP13). Thus the transition from region 1 to region 2 is continuous. A similar argument holds for supplier 2.

Region 3 « region 4. Consider supplier 1’s problem in region 4. Rewrite the right boundary of region 4 as \( w_2 = 2v - 2\theta l_1 - l_2 - w_1 \). Plugging it into the objective function in region 4 (SP4) yields the objective function in region 3 (SP13). Thus the transition from region 3 to region 4 is continuous. A similar argument holds for supplier 2.

Region 4 « region 5 « region 6. Consider supplier 1’s problem in regions 4, 5 and 6. Rewrite the left boundary of region 4 or right boundary of region 5 as \( w_2 = 2v - 3\theta l_1 - l_2 - w_1 \). Plugging it into the objective function in region 4 (SP4) and the objective function in region 5 (SP5) yields the same objective value. Rewrite the right boundary of region 6 or left boundary of region 5 as \( w_2 = 2v - 2\theta + 2\theta l_1 - l_2 - w_1 \). Plugging it into the objective function in region 5 (SP5) and the objective function in region 6 (SP26) also yields the same objective value. Thus the transition from region 4 to region 5 and then to region 6 is continuous. A similar argument holds for supplier 2.

Proof of Lemma 2.2:

Region 4’s optimal objective functions for the suppliers are stated in (A.33). Both objective functions increase in \( |l_1 - l_2| \) if

\[
\begin{align*}
|l_1 - l_2| & \geq \frac{c_1 - c_2}{6\theta} \\
|l_1 - l_2| & \geq \frac{c_2 - c_1}{6\theta}
\end{align*}
\]

which is always satisfied because it is exactly suppliers’ participating constraints in region 4.

Region 5’s optimal objective function for each supplier is described respectively by (A.36). Both objective functions increase in \( |l_1 - l_2| \) if

\[
\begin{align*}
7\theta l_1 - l_2 & \geq -14v + 17c_1 - 3c_2 \\
7\theta l_1 - l_2 & \geq -14v + 17c_2 - 3c_1
\end{align*}
\]

\[
\begin{align*}
17c_1 - 3c_2 & \leq 7\theta |l_1 - l_2| + 14v; \\
17c_2 - 3c_1 & \leq 7\theta |l_1 - l_2| + 14v.
\end{align*}
\]

Suppliers will participate if they can make a positive profit by having the product on the market, i.e. \( w_i \geq c_i \) for each \( i \). The suppliers’ participating constraints in region 4 are obtained by plugging (A.31), respectively for each supplier, into the previously stated inequality.
The above is always satisfied because it is exactly suppliers’ participating constraints in region 5. Again there is incentive to increase $|l_1 - l_2|$ as much as possible.

**Proof of Theorem 2.1:**

First we establish a few inequalities to characterize the locations of the optimal solutions in all 6 regions (refer to Figures 2.5 and 2.6). These inequalities, summarized in Table A.1 below, are both sufficient and necessary. We observe that the only two boundary values in the absence of $|l_1 - l_2|$ on $c_1 + c_2$ are $2v - 2\theta$ and $2v - 3\theta$. This generates the following segments, namely $c_1 + c_2 \leq 2v - 3\theta$, $c_1 + c_2 \geq 2v - 2\theta$ and $2v - 3\theta \leq c_1 + c_2 \leq 2v - 2\theta$. These segments are equivalent, respectively, to the following segments on the value added over average production cost $v - \frac{c_1 + c_2}{2}$, i.e., $v - \frac{c_1 + c_2}{2} \geq \frac{3}{2}\theta$, $v - \frac{c_1 + c_2}{2} \leq \theta$ and $\theta \leq v - \frac{c_1 + c_2}{2} \leq \frac{3}{2}\theta$. This proof is based on deriving relationships around $c_1 + c_2$, which is algebraically more straightforward to follow.

When $c_1 + c_2 \leq 2v - 3\theta$, or equivalently, $v - \frac{c_1 + c_2}{2} \geq \frac{3}{2}\theta$, we compare the upper bound on $c_1 + c_2$ for this case with the inequalities in Table A.1 and observe that the optimal solution to region 2 lies in the interior and that the optimal solutions to regions 1 and 3 lie at the left boundary. Observe that the boundary solution to region 1 is dominated by the interior solution of region 2 due to Lemma 2.1 and the fact that the boundary solution is feasible to region 2. Recall that suppliers in regions 1 and 3 have the same objective functions. Therefore, region 3’s boundary solution is dominated by region 1’s boundary solution because it is feasible to region 1. In summary, region 2 dominates regions 1 and 3.

The optimal solutions to other regions can lie in a number of different places. Consider the following cases on region 6:

1. When $|l_1 - l_2| \in \left[\frac{c_1 + c_2 - 2v + 4\theta}{\theta}, \frac{1}{2}\right]$, region 6 has an interior solution, which is equivalent in objective value to region 2 because suppliers in both regions share the same objective functions. Both regions 4 and 5 have left boundary solutions. This is due to the following: We quote the appropriate results from Table A.1. When region 4 has a left boundary solution, the upper bound $2v - 7\theta|l_1 - l_2|$ is always greater than the upper

\[^{2}\text{The suppliers’ participating constraints in region 5 are obtained by plugging (A.34), respectively for each supplier, into the inequality } w_i \geq c_i.\]
### Table A.1: Location of Optimal Solutions for All Regions

<table>
<thead>
<tr>
<th>Regions</th>
<th>Locations of Optimal Solutions</th>
<th>Representation with Inequalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 3</td>
<td>Both at left boundary</td>
<td>$w_1^* + w_2^* \leq 2v - \theta$ (A.26) $c_1 + c_2 \leq 2v - 2\theta$</td>
</tr>
<tr>
<td></td>
<td>Both in the interior</td>
<td>$w_1^* + w_2^* \geq 2v - \theta$ (A.36) $c_1 + c_2 \geq 2v - 4\theta</td>
</tr>
<tr>
<td></td>
<td>In the interior for region 1;</td>
<td>$2v - 2\theta \leq c_1 + c_2 \leq 2v - 4\theta^2</td>
</tr>
<tr>
<td></td>
<td>At left boundary for region 3</td>
<td></td>
</tr>
<tr>
<td>2, 6</td>
<td>Both at right boundary</td>
<td>$w_1^* + w_2^* \geq 2v - \theta$ (A.39) $c_1 + c_2 \geq 2v - 2\theta</td>
</tr>
<tr>
<td></td>
<td>Both in the interior</td>
<td>$w_1^* + w_2^* \geq 2v - 2\theta + \theta</td>
</tr>
<tr>
<td></td>
<td>In the interior for region 2;</td>
<td>$2v - 4\theta + \theta</td>
</tr>
<tr>
<td></td>
<td>At right boundary for region 6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>At left boundary</td>
<td>$w_1^* + w_2^* \leq 2v - 3\theta</td>
</tr>
<tr>
<td></td>
<td>At right boundary</td>
<td>$w_1^* + w_2^* \geq 2v - 2\theta</td>
</tr>
<tr>
<td></td>
<td>In the interior</td>
<td>$2v - 7\theta</td>
</tr>
<tr>
<td>5</td>
<td>At left boundary</td>
<td>$w_1^* + w_2^* \leq 2v - 2\theta + \theta</td>
</tr>
<tr>
<td></td>
<td>At right boundary</td>
<td>$w_1^* + w_2^* \geq 2v - 3\theta</td>
</tr>
<tr>
<td></td>
<td>In the interior</td>
<td>$2v - \frac{10}{3} \theta + \theta</td>
</tr>
</tbody>
</table>
bound for region 6 to have an interior solution, i.e. \( 2v - 4\theta + \theta|l_1 - l_2| \). Similarly, the upper bound for region 5 to have a left boundary solution, i.e. \( 2v - \frac{10}{3}\theta + \theta|l_1 - l_2| \), is also always greater than \( 2v - 4\theta + \theta|l_1 - l_2| \) as well. By Lemma 2.1, region 6 then dominates regions 4 and 5. We can conclude that regions 2 and 6 dominate all other regions.

2. When \( |l_1 - l_2| \in (0, \frac{c_1 + c_2 - 2v + 4\theta}{\theta}] \), region 6 has a right boundary solution, which is dominated by region 2’s interior solution because suppliers in both regions share the same objective functions and the boundary solution is feasible to region 2. Region 4’s right boundary solution is the same as region 3’s left boundary solution, which is in turn dominated by region 2’s solution. Region 5’s left boundary solution is the same as region 6’s right boundary solution. When either region 4 or 5 has an interior solution, it degenerates into a form similar to region 2 by Lemma 2.2 and is therefore dominated by any optimal solution in region 2. When region 4 has a left boundary solution, it degenerates into region 5, which in turn degenerates into region 2 by Lemma 2.2, and is therefore dominated by any optimal solution to region 2. Similarly, when region 5 has a right boundary solution, it degenerates into region 4, which in turn degenerates into region 2 by Lemma 2.2, and is therefore dominated by any optimal solution to region 2.

Furthermore, the minimal set for region 6 to have interior solutions requires \( c_1 + c_2 \leq 2v - 4\theta \) (by setting \( |l_1 - l_2| = 0 \)), while the minimal set for region 6 to have a right boundary solution requires \( 2v - \frac{7\theta}{2} \leq c_1 + c_2 \leq 2v - 3\theta \) (by setting \( |l_1 - l_2| = \frac{1}{2} \) and upper-bounding by the low cost case). Observe that in these two sets, the choice of \( |l_1 - l_2| \) can be arbitrary without affecting the solution characteristics (interior or boundary) of region 6. Now for the segment \( 2v - 4\theta \leq c_1 + c_2 \leq 2v - \frac{7\theta}{2} \), region 6 can have either interior or right boundary solutions depending on \( |l_1 - l_2| \).

When \( c_1 + c_2 \geq 2v - 2\theta \), or equivalently, \( v - \frac{c_1 + c_2}{2} \leq \theta \), we compare the lower bound on \( c_1 + c_2 \) for this case with the inequalities in Table A.1 and observe that the optimal solution to region 1 lies in the interior and that the optimal solutions to regions 2 and 6 lie at the right boundary. Observe that the boundary solution to region 2 is dominated by the interior solution of region 1 due to Lemma 2.1 and the fact that the boundary solution
is feasible to region 1. Recall that suppliers in regions 2 and 6 have the same objective functions. Therefore, region 6's boundary solution is dominated by region 2's boundary solution because it is feasible to region 2. In summary, region 1 dominates regions 2 and 6.

The optimal solutions to other regions can lie in a number of different places. We consider the following cases on region 3:

1. When $|l_1 - l_2| \in \left[ \frac{2v - c_1 - c_2}{4\theta}, \frac{1}{2} \right]$, region 3 has an interior solution, which is equivalent in objective value to region 1 because suppliers in both regions share the same objective functions. Both regions 4 and 5 both have right boundary solutions. This is due to the following: We quote the appropriate results from Table A.1. When region 4 has a right boundary solution, the lower bound $2v - 6\theta|l_1 - l_2|$ is always lower than the lower bound for region 3 to have an interior solution, i.e. $2v - 4\theta|l_1 - l_2|$. Similarly, the lower bound for region 5 to have a right boundary solution, i.e. $2v - \frac{12}{3^2} \theta|l_1 - l_2|$ is also always lower than $2v - 4\theta|l_1 - l_2|$. By Lemma 2.1, region 3 then dominates regions 4 and 5. We can conclude that regions 1 and 3 dominate all other regions.

2. When $|l_1 - l_2| \in (0, \frac{2v - c_1 - c_2}{4\theta}]$, region 3 has a left boundary solution, which is dominated by region 1's interior solution because suppliers in both regions share the same objective functions and the boundary solution is feasible to region 1. In this case, region 4 cannot have an interior solution because the corresponding inequality in Table A.1 becomes $c_1 + c_2 = 2v$ regardless of any $|l_1 - l_2|$ in the range $(0, \frac{2v - c_1 - c_2}{4\theta}]$, leaving no profit to be gained for all parties involved. Region 4 may have a right boundary solution or a left boundary one. However we need not examine its right boundary solution case because it is the same as region 3's left boundary solution. Region 5 has an interior solution when $|l_1 - l_2| \in (0, \frac{12}{17} \cdot \frac{2v - c_1 - c_2}{4\theta}]$ and has a right boundary solution when $|l_1 - l_2| \in (\frac{12}{17} \cdot \frac{2v - c_1 - c_2}{4\theta}, \frac{2v - c_1 - c_2}{4\theta}]$. When it has an interior solution, region 5 dominates regions 3, 4 and 6 by Lemma 2.1. By Lemma 2.2, there is incentive for region 5's interior solution to maximally differentiate, which degenerates that solution into one similar to region 2, and is therefore dominated by any optimal solution in region 2—in this case, the only solution to region 2, or its right boundary solution. When region 5 has a right boundary solution, it degenerates into a solution similar to region 4, which in turn degenerates into one similar to region
2 by Lemma 2.2. Therefore, region 5's right boundary solution is dominated by any optimal solution in region 2, which again in this case is its right boundary solution, or the only solution.

When \( c_1 + c_2 \) lies between \( 2v - 3\theta \) and \( 2v - 2\theta \), or equivalently, when \( v - \frac{c_1 + c_2}{2} \) lies between \( \theta \) and \( \frac{3}{2}\theta \), both regions 2 and 6 have right boundary solutions, whereas regions 1 and 3 have left boundary solutions. Region 6's boundary solution is feasible to region 2 and is therefore dominated by its optimal solution—in this case, its right boundary solution. Similarly, region 3's left boundary solution is feasible to region 1 and is therefore dominated by its optimal solution—in this case, its left boundary solution. Note that region 1's left boundary solution is the same as region 2's right boundary solution, and we shall refer to it as the common boundary solution of regions 1 and 2 from now on. Region 4's right boundary solution is the same as region 3's left boundary solution, which is in turn dominated by the common boundary solution of regions 1 and 2. Region 5's left boundary solution is the same as region 6's right boundary solution, which is in turn dominated by the common boundary solution of regions 1 and 2. When either region 4 or 5 has an interior solution, it degenerates into a form similar to region 2 by Lemma 2.2 and is therefore dominated by any optimal solution in region 2—in this case, the only solution to region 2, or the common boundary solution to regions 1 and 2. When region 4 has a left boundary solution, it degenerates into region 5, which in turn degenerates into region 2 by Lemma 2.2, and is therefore dominated by any optimal solution to region 2—in this case, the common boundary solutions of regions 1 and 2. Similarly, when region 5 has a right boundary solution, it degenerates into region 4, which in turn degenerates into region 2 by Lemma 2.2, and is therefore dominated by any optimal solution to region 2—in this case, the common boundary solution of regions 1 and 2.

A.2 Technical Details of Benchmark Systems

Proof of Theorem 2.2:

First we establish a side proposition.
APPENDIX A. TECHNICAL DETAILS OF COORDINATING VERTICAL PARTNERSHIPS
FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

Proposition A.1 When product positions are symmetric, the retailer’s optimal solution can be segmented as the following:

- Region 1 \((v - \frac{\zeta(\ell_1 - \ell_2)}{4} < \frac{\theta}{2})\) - Separable solution is optimal;
- Region 2 \((v - \frac{\zeta(\ell_1 - \ell_2)}{4} > \frac{\theta}{2})\) - Barely touching solution is optimal.

When product positions are asymmetric, the retailer’s optimal solution can be segmented as the following:

- Region 3 \((v - \frac{\zeta(\ell_1 - \ell_2)}{4} < \theta|\ell_1 - \ell_2|\) - Separable solution is optimal;
- Region 4 \((\theta|\ell_1 - \ell_2| < v - \frac{\zeta(\ell_1 - \ell_2)}{4} < \frac{3}{2}\theta|\ell_1 - \ell_2|)\) - The solution where one end barely touches and the other end does not touch is optimal;
- Region 5 \((\frac{3}{2}\theta|\ell_1 - \ell_2| < v - \frac{\zeta(\ell_1 - \ell_2)}{4} < \theta - \frac{1}{2}\theta|\ell_1 - \ell_2|)\) - The solution where one end overlaps and the other end does not touch is optimal;
- Region 6 \((v - \frac{\zeta(\ell_1 - \ell_2)}{4} > \theta - \frac{1}{2}\theta|\ell_1 - \ell_2|)\) - The solution where one end overlaps and the other end barely touches is optimal.

The solution to each region is summarized in Table A.2.

Proof: We simply replace \(w_i\) in Proposition 2.1 with \(c_i\).
it is optimal for the centralized retailer to maximally differentiate the two products, i.e. \(|l_1 - l_2| = \frac{1}{2}\), leading to the two segments in the theorem statement.

Region 4’s optimal profit \(^3\) increases in \(|l_1 - l_2|\) if  
\[
\frac{\partial \pi^*}{\partial |l_1 - l_2|} > 0 \Rightarrow v - \frac{c_1 + c_2}{2} > \theta |l_1 - l_2|, 
\]
which turns out to be the right boundary of region 4 and is thus always satisfied. When \(|l_1 - l_2| \uparrow \frac{1}{2}\), region 4’s optimal profit becomes that of region 2. We can therefore conclude that region 4’s solution is always dominated by region 2.

Region 5’s optimal profit increases in \(|l_1 - l_2|\) if  
\[
\frac{\partial \pi^*}{\partial |l_1 - l_2|} > 0 \Rightarrow v - \frac{c_1 + c_2}{2} > -\frac{1}{2} \theta |l_1 - l_2|, 
\]
which is always satisfied because \(c_i \leq v\). Under \(|l_1 - l_2| \uparrow \frac{1}{2}\), region 5 becomes degenerate in the limit, i.e.
\[
c_2 = 2v - \frac{3\theta}{2} - c_1. \quad (A.37)
\]
Region 5’s optimal profit then becomes \((v-c_1)^2 - \frac{3}{2}(v-c_1) + \frac{17\theta}{16}\), which is the same as region 2’s optimal profit with (A.37) plugged in. Therefore region 5’s solution is always dominated by region 2.

Region 6’s optimal profit is linear in \(|l_1 - l_2|\). The coefficient in front of \(|l_1 - l_2|\) is \(\frac{\theta}{2} > 0\). Thus, region 6’s optimal profit always increases in \(|l_1 - l_2|\). When \(|l_1 - l_2| \uparrow \frac{1}{2}\), region 6’s optimal profit becomes that of region 2 in the limit. We can therefore conclude that region 6’s solution is always dominated by region 2.

Solution to (SP13’):

The objective function is concave, and the constraint is linear. Therefore, KKT conditions are both sufficient and necessary. The Lagrangian is:
\[
\mathcal{L} = (w_1 - c_1)\frac{v - w_1}{\theta} + (w_2 - c_2)\frac{v - w_2}{\theta} + \lambda(w_1 - w_2 - \alpha).
\]

KKT conditions yield the following:
\[
w_1 = \frac{v + c_1 + \lambda \theta}{2}; \quad w_2 = \frac{v + c_2 + \lambda \theta}{2}. \quad (A.38)
\]

\(^3\)Optimal profit for each region can be found under the column \(\pi^*\) in Table A.2.
Assume strict complementarity holds. There are two cases to consider:

1. \( \lambda = 0 \Leftrightarrow \) interior solution is optimal.

\[
\begin{align*}
(A.38) & \Rightarrow w_1^* = \frac{v + c_1}{2}; \quad w_2^* = \frac{v + c_2}{2}; \\
& \Rightarrow w_1^* + w_2^* = v + \frac{c_1 + c_2}{2} > \alpha; \\
& \Rightarrow \text{For region 1, } v - \frac{c_1 + c_2}{2} < \theta; \text{ for region 3, } v - \frac{c_1 + c_2}{2} < 2\theta|l_1 - l_2|.
\end{align*}
\]

2. \( \lambda > 0 \Leftrightarrow \) boundary solution is optimal.

\[
\begin{align*}
(A.38) & \Rightarrow w_1 + w_2 = v + \lambda \theta + \frac{c_1 + c_2}{2} = \alpha; \\
& \Rightarrow \lambda = \frac{\alpha - v}{\theta} - \frac{c_1 + c_2}{2\theta} > 0; \\
& \Rightarrow \text{For region 1, } v - \frac{c_1 + c_2}{2} > \theta; \text{ for region 3, } v - \frac{c_1 + c_2}{2} > 2\theta|l_1 - l_2|.
\end{align*}
\]

**Solution to (SP26'):**

The Lagrangian is:

\[
\mathcal{L} = (w_1 - c_1)\left(\frac{1}{2} - \frac{w_1}{2\theta} + \frac{w_2}{2\theta}\right) + (w_2 - c_2)\left(\frac{1}{2} - \frac{w_2}{2\theta} + \frac{w_1}{2\theta}\right) + \lambda(\alpha - w_1 - w_2).
\]

KKT conditions yield the following:

\[
w_1 - w_2 = \frac{c_1 - c_2}{2} + \frac{\theta}{2} - \theta \lambda; \quad w_1 - w_2 = \frac{c_1 - c_2}{2} - \frac{\theta}{2} + \theta \lambda.
\]

Clearly, we want

\[
\frac{\theta}{2} - \theta \lambda = -\frac{\theta}{2} + \theta \lambda \Rightarrow \lambda = \frac{1}{2} > 0.
\]

Therefore, the constraint is always active, and we have the following:

\[
\begin{align*}
& \left\{ \begin{array}{l}
w_1 - w_2 = \frac{c_1 - c_2}{2}; \\
w_1 + w_2 = \alpha;
\end{array} \right. \\
& \Rightarrow w_1^* = \frac{\alpha}{2} + \frac{c_1 - c_2}{4}; \quad w_2^* = \frac{\alpha}{2} + \frac{c_2 - c_1}{4}; \\
& \pi^* = \frac{\alpha}{2} - \frac{c_1 + c_2}{2} + \frac{(c_1 - c_2)^2}{8\theta}.
\end{align*}
\]
Note that \( \pi^* \) increases in \( a \). Since the \( a \) value associated with region 2 \( (2v - \theta) \) is greater than that of region 6 \( (2v - 2\theta + \theta|l_1 - l_2|) \), we can conclude that both region 2 and region 6 will achieve optimality at the right boundaries, and region 2’s solution dominates that of region 6.

Solution to (SP4'):

The Lagrangian is:

\[
\mathcal{L} = \left( w_1 - c_1 \right) \left( \frac{1}{4} |l_1 - l_2| - \frac{w_1}{2\theta} + \frac{w_2}{4\theta} \right) + \left( w_2 - c_2 \right) \left( \frac{1}{4} |l_1 - l_2| - \frac{w_2}{2\theta} + \frac{w_1}{4\theta} \right) \\
+ \mu \left( w_1 + w_2 - \frac{2v}{2\theta} + 3\theta|l_1 - l_2| \right) + \lambda \left( 2v - 2\theta + 6|l_1 - l_2| - w_1 - w_2 \right)
\]

KKT conditions yield the following:

\[
w_1 - w_2 = \frac{c_1 - c_2}{2} - \frac{1}{2} \mu \left( |l_1 - l_2| + \theta|l_1 - l_2| + \theta(\mu - \lambda) \right); \quad w_1 - w_2 = \frac{c_1 - c_2}{2} - \frac{1}{2} \mu \left( |l_1 - l_2| - \theta(\mu - \lambda) \right).
\]

Clearly, we want

\[
\theta|l_1 - l_2| + \theta(\mu - \lambda) = -\theta|l_1 - l_2| - \theta(\mu - \lambda) \Rightarrow \mu - \lambda = -|l_1 - l_2|.
\]

\( \mu \) and \( \lambda \) cannot both be positive, since the upper and lower bounds cannot both be active at the same time. \( \mu \) and \( \lambda \) cannot both be 0, since \( 0 \neq -|l_1 - l_2| \). If \( \mu > 0 \) and \( \lambda = 0 \), then we have \( \mu = -|l_1 - l_2| < 0 \), a contradiction to \( \mu \geq 0 \). The only possibility is therefore \( \mu = 0 \) and \( \lambda > 0 \), which leads to \( \lambda = |l_1 - l_2| > 0 \). This implies that region 4’s optimal solution always lies at the right bound.

Solution to (SP5'):

The objective function is concave, and the constraint is linear. Therefore, KKT conditions are both sufficient and necessary. The Lagrangian is:

\[
\mathcal{L} = \left( w_1 - c_1 \right) \left( \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_2}{4\theta} - \frac{3w_1}{4\theta} \right) + \left( w_2 - c_2 \right) \left( \frac{1}{4} |l_1 - l_2| + \frac{v}{2\theta} + \frac{w_1}{4\theta} - \frac{3w_2}{4\theta} \right) \\
+ \mu \left( w_1 + w_2 - \frac{2v}{2\theta} + 3\theta|l_1 - l_2| \right) + \lambda \left( 2v - 3\theta|l_1 - l_2| - w_1 - w_2 \right).
\]
KKT conditions yield the following:

\[ 3w_1 - w_2 = v + \frac{\theta}{2}|l_1 - l_2| + \frac{3c_1 - c_2}{2} + 2\theta(\mu - \lambda); \quad 3w_2 - w_1 = v + \frac{\theta}{2}|l_1 - l_2| + \frac{3c_2 - c_1}{2} + 2\theta(\mu - \lambda); \]

\[ \Rightarrow w_1 = \frac{v}{2} + \frac{\theta}{4}|l_1 - l_2| + \frac{c_1}{2} + 2\theta(\mu - \lambda); \quad w_2 = \frac{v}{2} + \frac{\theta}{4}|l_1 - l_2| + \frac{c_2}{2} + 2\theta(\mu - \lambda); \quad (A.39) \]

\[ w_1 + w_2 = v + \frac{\theta}{2}|l_1 - l_2| + \frac{c_1 + c_2}{2} + 2\theta(\mu - \lambda). \quad (A.40) \]

There are 3 cases to consider.

1. \( \mu = 0, \lambda > 0 \) \( \Leftrightarrow \) right boundary of region 5. Plug \( w_1 + w_2 = 2v - 3\theta|l_1 - l_2| \) and \( \mu = 0 \) into (A.40). By \( \lambda > 0 \), we have that

\[ v - \frac{c_1 + c_2}{2} < \frac{7}{2}\theta|l_1 - l_2|. \]

2. \( \mu > 0, \lambda = 0 \) \( \Leftrightarrow \) left boundary of region 5. Plug \( w_1 + w_2 = 2v - 2\theta + \theta|l_1 - l_2| \) and \( \lambda = 0 \) into (A.40). By \( \mu > 0 \), we have that

\[ v - \frac{c_1 + c_2}{2} > 2\theta - \frac{1}{2}\theta|l_1 - l_2|. \]

3. \( \mu = 0, \lambda = 0 \) \( \Leftrightarrow \) region 5 has interior solution.

\[ (A.39) \Rightarrow w_1^* = \frac{v}{2} + \frac{\theta}{4}|l_1 - l_2| + \frac{c_1}{2}; \quad w_2^* = \frac{v}{2} + \frac{\theta}{4}|l_1 - l_2| + \frac{c_2}{2}. \]

Supplier's optimal category profit is therefore

\[ \pi^* = \frac{1}{32\theta}(2v + \theta|l_1 - l_2| - 2c_1)(2v + \theta|l_1 - l_2| + c_2 - 3c_1) \]

\[ + \frac{1}{32\theta}(2v + \theta|l_1 - l_2| - 2c_2)(2v + \theta|l_1 - l_2| + c_1 - 3c_2). \quad (A.41) \]

Plugging \( w_1^* + w_2^* \) into the constraints of region 5's optimization problem yields

\[ \frac{7}{2}\theta|l_1 - l_2| < v - \frac{c_1 + c_2}{2} < 2\theta - \frac{1}{2}\theta|l_1 - l_2|. \]

Proof of Theorem 2.3:

First we show the following side lemma.

**Lemma A.1** There is incentive to move toward maximum product differentiation in region 5 if the solution lies in the interior. The interior solution of region 5 is dominated by the optimal solution of region 2.
APPENDIX A. TECHNICAL DETAILS OF COORDINATING VERTICAL PARTNERSHIPS FOR HORIZONTALLY DIFFERENTIATED PRODUCTS

Proof: Region 5's optimal profit (A.41) increases in \(|l_1 - l_2|\) if

\[
\frac{\partial \pi^*}{\partial |l_1 - l_2|} > 0 \quad \Rightarrow \quad v - \frac{c_1 + c_2}{2} > -\frac{1}{2} \theta |l_1 - l_2|,
\]

which is always satisfied since \(c_i \leq v\). As \(|l_1 - l_2| \uparrow \frac{1}{2}\), region 5 becomes degenerate in the limit and has an interior solution if

\[
c_2 = 2v - \frac{7\theta}{2} - c_1. \tag{A.42}
\]

Region 5's solution then becomes \((v - c_3)^2 + 8\frac{\theta}{2} - \frac{7(v - c_3)}{4}\), which is less than \((v - c_3)^2 + \frac{8\theta}{2} - \frac{7(v - c_3)}{4}\), region 2's optimal solution with (A.42) plugged in. Therefore, region 5's solution is always dominated by region 2.

It is easy to verify in the same manner as Lemma 2.1 that the transition from a region to its neighbor is continuous for the supplier's problem. Between regions 1 and 3, which share the same objective function, we know that region 1's solution will always dominate (if not the same as) region 3's solution. This can be seen by the following argument. If both region 1 and region 3 have interior solution, then they have the same value. If region 1 has interior solution but region 3 has left boundary solution, then region 3's solution is feasible to region 1 and is thus dominated by region 1's optimal interior solution. If both regions 1 and 3 have left boundary solution, then region 1's larger feasibility set ensures that region 1 will again dominate region 3. Therefore, the attack plan for an optimal solution to the monopolist supplier model is by considering when region 1 has interior or boundary solutions.

Optimal solution to region 1 lies in the interior \((v - \frac{c_1 + c_2}{2} < \theta)\): Region 2 has right boundary solution, which is dominated by region 1's interior solution.

Consider region 5 at the right boundary. Due to continuity of objective values across regions, region 6's right boundary solution is dominated by region 5's right boundary solution, which is in turn dominated by region 4's right boundary solution. If region 3's solution is at the interior and is thus equal to region 1's interior solution, then region 4's right boundary solution is dominated by it. If region 3's solution is at the left boundary and is thus dominated by region 1's interior solution (as shown in the opening of this section), then it is equal to region 4's right boundary solution, which makes it also dominated by region 1's interior solution.
Consider region 5 at the left boundary. Then region 5 has the same value as region 6, which is dominated by region 2 because region 2 has a larger feasibility set and both regions 2 and 6 have the same objective function. Relationship across all other regions remain the same. Therefore region 1 still dominates all.

Consider region 5 at the interior. Then by Lemma A.1, region 5 is a feasible point in region 2 and is thus dominated by region 2’s optimal solution at the right boundary, which is in turn dominated by region 1’s interior solution. All other relationships remain the same. Therefore region 1 still dominates all.

Optimal solution to region 1 lies at the boundary ($v - \frac{\xi_1 + \xi_2}{2} > \theta$): Region 2’s right boundary solution is the same as region 1’s left boundary solution.

Consider region 5 at the right boundary. Then region 3 dominates (or is equal to) region 4, which dominates region 5, which dominates region 6. By argument at the beginning of this section, region 1 will dominate region 3.

Consider region 5 at the left boundary. Then it is the same as region 6’s right boundary solution, which is dominated by region 1 and 2’s common boundary solution. All other relationships remain the same. Region 1 still dominates all.

Consider region 5 at the interior. Then by Lemma A.1, region 5 is feasible to region 2 and is thus dominated by region 2’s optimal right boundary solution (same as region 1’s left boundary solution). All other relationships remain the same. Region 1 still dominates all.
Appendix B

Technical Details of Dynamic Strategies under Sticker Shock

Proof of Theorem 3.2:

By Assumption 3.4 and Smith and McCardle (2002, Proposition 5), it is easy to obtain that $W(p, u, r)$ is supermodular in $(p, u, r)$. There are two cases to consider:

(i) $u'(p^r) = 0, u'(p^f) = 1 \Rightarrow u'(r') = 0, u'(r'') = 1$;

(ii) $u'(p^r) = 1, u'(p^f) = 0 \Rightarrow u'(r') = 1, u'(r'') = 0$.

Let $r' < p < r''$. By the supermodularity of $W(p, u, r)$ and Theorem 3.1, we have for Case (i) that

$$W((p(r'), 1, r') \lor (p(r''), 0, r'')) + W((p(\bar{r}), 1, \bar{r}) \land (p(r''), 0, r'')) \geq W(p(\bar{r}), 1, \bar{r}) + W(p(r''), 0, r'')),$$

$$W(p(r'), 1, r') + W(p(\bar{r}), 0, \bar{r}) \geq W(p(\bar{r}), 1, \bar{r}) + W(p(r''), 0, r'')),$$

$$W(p(r'), 1, r') - W(p(r''), 0, r'')) \geq W(p(\bar{r}), 1, \bar{r}) - W(p(\bar{r}), 0, \bar{r}) = 0,$$ by Equation 3.3.

Therefore,

$$W(p(r'), 1, r') > W(p(r''), 0, r'')) \Rightarrow u'(r'') = 1.$$
APPENDIX B. TECHNICAL DETAILS OF DYNAMIC STRATEGIES UNDER STICKER SHOCK

The other side can be argued similarly.

\[ W((p(\tilde{r}), 0, \tilde{r}) \lor (p(r'), 1, r')) + W((p(\tilde{r}), 0, \tilde{r}) \land (p(r'), 1, r')) \geq W(p(\tilde{r}), 0, \tilde{r}) + W(p(r'), 1, r')) , \]

\[ \Rightarrow W(p(\tilde{r}), 1, r') + W(p(r'), 0, r') \geq W(p(\tilde{r}), 0, \tilde{r}) + W(p(r'), 1, r')) , \]

\[ \Rightarrow W(p(r'), 1, r') - W(p(r'), 0, r')) \leq W(p(\tilde{r}), 1, r') - W(p(\tilde{r}), 0, \tilde{r}) = 0, \text{ by Equation 3.3.} \]

In other words,

\[ W(p(r'), 1, r') < W(p(r'), 0, r')) \Rightarrow u^*(r') = 0. \]

The argument for Case (ii) follows a similar logic.

**Proof of Theorem 3.3:**

To prove part (a), we use a variational approach similar to the one laid out in Popescu and Wu (2006). Suppose there exists a steady state \( r \), such that \( p^*(r) = r \). Here, we can drop the subscript \( u \) in the optimal pricing policy \( p^*_u(r) \) because the optimal communication policy \( u^*(r) \) applied to the steady state \( r \) is by definition and by construction a constant. For a small enough \( \delta > 0 \), let a feasible pricing policy be

\[ \phi(\delta) := (p_0 = r - \delta; \ p_1 = r + (1 - \alpha_u)\delta; \ p_t = r, \forall t \geq 2), \]

where \( u \) is the same communication action applied to all \( t \). The corresponding reference price trajectory is, therefore,

\[ r_0 = r; \]

\[ r_1 = \alpha_u p_0 + (1 - \alpha_u) r_0 = r - \alpha_u \delta; \]

\[ r_2 = \alpha_u p_1 + (1 - \alpha_u) r_1 = r; \]

\[ \ldots \]

We obtain the value function under policy \( \phi(\delta) \) as

\[ V_\delta(r) = (1 + mu)(\pi(p_0) + p_0 R(r_0 - p_0, r_0)) - Ku \]

\[ + \gamma ((1 + mu)(\pi(p_1) + p_1 R(r_1 - p_1, r_1)) - Ku) + \gamma^2 V(r) \]

\[ = (1 + mu)(\pi(r - \delta) + (r - \delta) R(\delta, r)) - Ku \]

\[ + \gamma ((1 + mu)(\pi(r + (1 - \alpha_u)\delta) + (r + (1 - \alpha_u)\delta) R(-\delta, r - \alpha_u \delta)) - Ku) + \gamma^2 V(r). \]
Since $V(r)$ is the optimal value function obtained under the constant pricing policy $p^*(r) = r$, we have therefore

$$V(r) = (1 + \gamma)((1 + m\mu)(\pi(r) + rR(0, r)) - Ku) + \gamma^2 V(r)$$

$$= (1 + \gamma)((1 + m\mu)(\pi(r) - Ku) + \gamma^2 V(r)$$

$$\geq V_\delta(r).$$

The second equality is due to Assumption 3.1. The inequality is by the optimality of $V(r)$. Combining and rearranging the terms, we obtain the following relationship:

$$(\pi(r - \delta) - \pi(r + (r - \delta)R(\delta, r)) + \gamma (\pi(r + (1 - \alpha_u\delta)) - \pi(r) + (r + (1 - \alpha_u\delta))R(-\delta, r - \alpha_u\delta)) \leq 0.$$  

Observing that $\lim_{\delta \to 0} \frac{R(\delta, r)}{\delta} = R_x(0, r)$, $R(0, r) = 0$, and $\lim_{\delta \to 0} \frac{R(-\delta, -\alpha_u\delta)}{\delta} = -R_x(0, r)$, we divide both sides of above inequality by $\delta$, take $\delta$ to 0 and obtain

$$-\pi'(r) + \gamma \pi'(r)(1 - \alpha_u) + rR_x(0, r) - \gamma rR_x(0, r) \leq 0.$$  

In other words,

$$\frac{\pi'(p)}{1 - \gamma} \geq \frac{pR_x(0, p)}{1 - \gamma + \alpha_u\gamma}.$$  

A similar argument can be made with the application of the feasible pricing policy $p(-\delta)$. Therefore, the equality holds for any constant communication action $u$.

To prove part (b), we plug $p^*(r) = r = p^{**}$ into Equation 3.2. By Assumption 3.1, we obtain:

$$V(p^{**}) = \max\{\pi(p^{**}) + \gamma V(p^{**}), (1 + m)\pi(p^{**}) - K + \gamma V(p^{**})\}.$$  

Clearly, $u^{**} = 1$ if

$$(1 + m)\pi(p^{**}) - K + \gamma V(p^{**}) \geq \pi(p^{**}) + \gamma V(p^{**}) \Rightarrow \pi(p^{**}) \geq \frac{K}{m};$$

and $u^{**} = 0$ if

$$(1 + m)\pi(p^{**}) - K + \gamma V(p^{**}) < \pi(p^{**}) + \gamma V(p^{**}) \Rightarrow \pi(p^{**}) < \frac{K}{m}. $$
To prove part (c), i.e. the monotonicity of \( p^* \) in both \( a_u \) and \( y \), we denote the following function

\[
G(p, a_u, \beta) = \frac{1 - \gamma + a_u \gamma}{1 - \gamma} \pi'(p) - pR_x(0, p) \\
= (\pi'(p) - pR_x(0, p)) + \frac{a_u \gamma}{1 - \gamma} \pi'(p).
\]

By Assumption 3.4(a) and (b), therefore, \( G \) is decreasing in \( p \). Because there exists a unique \( p^* \) that solves

\[
G(p^*, a_u, \gamma) = 0,
\]

We have therefore \( G \uparrow a_u \) and \( G \uparrow \gamma \). As a result, \( p^* \uparrow a_u \) and \( p^* \uparrow \gamma \).

**Proof of Theorem 3.4:**

We will prove the theorem via construction of two series. For an initial reference price \( r_0 < q \), there exists an \( m \in \mathbb{Z}^+ \) such that the following sequence applies:

\[
\hat{p} > r_1 = q(r_0) = a_1 p^*(r_0) + (1 - a_1) r_0 > r_0,
\]

\[
\hat{p} > r_2 = q(r_1) = a_1 p^*(r_1) + (1 - a_1) r_1 > r_1,
\]

\[ \vdots \]

\[
\hat{p} > r_{m-1} = q(r_{m-2}) = a_1 p^*(r_{m-2}) + (1 - a_1) r_{m-2} > r_{m-2},
\]

but

\[
\hat{p} \leq r_m = q(r_{m-1}) = a_1 p^*(r_{m-1}) + (1 - a_1) r_{m-1} > r_{m-1}.
\]

Conversely for an \( r'_0 \geq \hat{p} \), there exists an \( n \in \mathbb{Z}^+ \) such that

\[
\hat{p} \leq r'_1 = q(r'_0) = a_0 p^*(r'_0) + (1 - a_0) r'_0 < r'_0,
\]

\[
\hat{p} \leq r'_2 = q(r'_1) = a_0 p^*(r'_1) + (1 - a_0) r'_1 < r'_1,
\]

\[ \vdots \]

\[
\hat{p} \leq r'_{n-1} = q(r'_{n-2}) = a_0 p^*(r'_{n-2}) + (1 - a_0) r'_{n-2} < r'_{n-2},
\]

but

\[
\hat{p} > r'_n = q(r'_{n-1}) = a_0 p^*(r'_{n-1}) + (1 - a_0) r'_{n-1} < r'_{n-1}.
\]
A cycle can be constructed if we let $r_m = r'_0$ and $r'_n = r_0$, where $m = \inf\{i \in \mathbb{Z}^+: q(r_{i-1}) \geq \delta\}$ and $n = \inf\{i \in \mathbb{Z}^+: q(r'_{i-1}) \leq \delta\}$. This is equivalent to the system of equations (3.5). The ranking in (3.6) is also easily obtained from this construction.
Appendix C

Technical Details of Shelf Placement Optimization

The Lagrangian systems used in the clustering steps of Example 1 in Section 4.2.2 are shown here. Solutions to these systems are used to advance the clustering iterations.
<table>
<thead>
<tr>
<th></th>
<th>0.6</th>
<th>0.8</th>
<th>1.0</th>
<th>1.2</th>
<th>1.4</th>
<th>1.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C.1: Ex. 1 Step 0a - Lagrangian System to Initialize Standard Response
Table C.2: Ex. 1 Step 1b - Lagrangian System to Update Standard Response Given 1st Split
Table C.3: Ex. 1 Step 2b - Lagrangian System to Update Standard Response Given 2nd Split