Multiperiod Optimization Models in Operations Management

by

Kevin Bozhe Li

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Industrial Engineering and Operations Research

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Candace Arai Yano, Chair
Professor Philip Kaminsky
Professor Miguel Villas-Boas
Assistant Professor Anil Aswani

Fall 2018

Multiperiod Optimization Models in Operations Management

by

Kevin Bozhe Li

Abstract

Multiperiod Optimization Models in Operations Management

by

Kevin Bozhe Li

Doctor of Philosophy in Engineering - Industrial Engineering and Operations Research

University of California, Berkeley

Professor Candace Arai Yano, Chair

In the past two decades, retailers have witnessed rapid changes in markets due to an increase in competition, the rise of e-commerce, and ever-changing consumer behavior. As a result, retailers have become increasingly aware of the need to better coordinate inventory control with pricing in order to maximize their profitability. This dissertation was motivated by two of such problems facing retailers at the interface between pricing and inventory control. One considers inventory control decisions for settings in which planned prices fluctuate over time, and the other considers pricing of multiple substitutable products for settings in which customers hold inventory as a consequence of stockpiling when promotional prices are offered.

In Chapter 1, we provide a brief motivation for each problem. In Chapter 2, we consider optimization of procurement and inventory allocation decisions by a retailer that sells a product with a long production lead time and a short selling season. The retailer orders most products months before the selling season, and places only one order for each product due to short product life cycles and long delivery lead times. Goods are initially stored at the warehouse and then sent to stores over the course of the season. The stores are in high-rent locations, necessitating efficient use of space, so there is no backroom space and it is uneconomical to send goods back to the warehouse; thus, all inventory at each store is available for sale. Due to marketing and logistics considerations, the planned trajectory of prices is determined in advance and may be non-monotonic. Demand is stochastic and price-dependent, and independent across time periods. We begin our analysis with the case of a single store. We first formulate the inventory allocation problem given a fixed initial order quantity with the objective of maximizing expected profit as a dynamic program and

explain both technical and computational challenges in identifying the optimal policy. We then present two variants of a heuristic based on the notion of equalizing the marginal value of inventory across the time periods. Results from a numerical study indicate that the more sophisticated variant of the heuristic performs well when compared with both an upper bound and an industry benchmark, and even the simpler variant performs fairly well for realistic settings. We then generalize our approaches to the case of multiple stores, where we allow the stores to have different price trajectories. Our numerical results suggest that the performance of both heuristics is still robust in the multiple store setting, and does not suffer from the same performance deterioration observed for the industry benchmark as the number of stores increases or as price differences increase across stores and time periods. For the pre-season procurement problem, we develop a heuristic based on a generalization of the newsvendor problem that accounts for the two-tiered salvage values in our setting, specifically, a low price during end-of-season markdown periods and a very low or zero salvage value after the season has concluded. Results for numerical examples indicate that our modified newsvendor heuristic provides solutions that are as good as those obtained via grid search.

In Chapter 3, we address a retailer's problem of setting prices, including promotional prices, over a multi-period horizon for multiple substitutable products in the same product category. We consider the problem in a setting in which customers anticipate the retailer's pricing strategy and the retailer anticipates the customers' purchasing decisions. We formulate the problem as a two-stage game in which the profit maximizing retailer chooses prices and the utility maximizing customers respond by making explicit decisions regarding purchasing and consumption, and thus also implicit decisions regarding stockpiling. We incorporate a fairly general reference price formation process that allows for cross-product effects of prices on reference prices. We initially focus on a single customer segment. The representative customer's utility function accounts for the value of consumption of the products, psychological benefit (for deal-seekers) from purchasing at a price below his/her reference price but with diminishing marginal returns, costs of purchases, penalties for both shortages and holding inventory, and disutility for deviating from a consumption target in each period (where applicable). We are the first to develop a model that simultaneously accounts for this combination of realistic factors for the customer, and we also separate the customer's purchasing and consumption decisions. We develop a methodology for solving the customer's problem for arbitrary price trajectories based on a linear quadratic control formulation of an approximation of the customer's utility maximization problem. We derive analytical representations for the customer's optimal decisions as simple linear functions of prices, reference prices, inventory levels (as state variables), and the cumulative aggregate consumption level (as a state variable). We then embed the consumer's optimal policy (in analytic form) in the retailer's profit maximization problem and show that the retailer's problem can be reformulated as a quadratic program that can be solved numerically. Despite the additional generality of our problem context, our solution methodology is computationally tractable. Results for numerical examples indicate that the separation of the customer's two types of decisions turns out to be important in inducing realistic purchasing patterns, including demand spikes like those observed in practice and the associated stockpiling, as well as pre- and post-promotion dips, which are also observed in practice. We discuss how our approach can be extended to multiple customer segments, limitations of our approach, and future research directions.

To my family.

Contents

List of Figures

List of Tables

Acknowledgments

Throughout my PhD, I have received a lot of encouragement and support from many people. My words here cannot fully express my gratitude and appreciation to all of those who have helped me in the last five years.

First and foremost, I would like to thank my advisor, Candace A. Yano, for her support, encouragement, and guidance. She gave me the freedom to choose my research topic, structure my time, and explore other interests and ventures. She always made herself available whenever I needed guidance and supported me through the various phases of my PhD journey and my job search. Since the day I arrived at Berkeley, she has been my mentor and has always had my interests in mind.

I would like to thank Professor Philip Kaminsky, Professor Miguel Villas-Boas, and Professor Anil Aswani for their academic guidance and for serving on my thesis committee and qualifying exam committee. They have provided invaluable feedback that has helped shape my work detailed in this dissertation. Professor Kaminsky and Professor Aswani also taught courses in supply chain management, learning and optimization which helped build key fundamental technical knowledge that I needed to approach my research.

I would also like to thank the many members of the department staff, especially Keith McAleer, Yeri Caesar-Kaptoech, Sonia Chahal, Anayancy Paz, Rebecca R. Pauling, and Diana Salazar. I appreciate all the timely help they provided when I needed it the most in the last five years.

I am incredibly grateful to the many friends, officemates, and colleagues who have supported me through the journey and made my Berkeley experience truly unforgettable. I am especially grateful to many of them which I will list only a few here: Jiung Lee, Erik Bertelli, Raj Agrawal, Joe Borja, Tugce Gurek, Guang Yang, Clay Campaigne, Alexander Fred-Ojala, Quico Spaen, Amber Richter, Cheng Lyu, Stewart Liu, Angel Yang, Wen Gao, Morteza Ashraphijuo, Jared Bauman, Carlos Deck, Brent Eldridge, Auyon Siddiq, Alfonso Andres Lobos Ruiz, Mo Zhou, Jiaying Shi, Shiman Ding, Rebecca Sarto, Titouan Jehl, Rebecca Sarto, Matt Olfat, Salar Fattahi, Dean Grosbard, Tianyi Lin, Renyuan Xu, and Mark Velednitsky.

I would like to thank my family members who have given me all the love throughout my PhD journey and have been my biggest cheerleader. It is all of your support that has kept me motivated and strong. Each and every one of you continues to inspire me on a daily basis and remind me that "We are all stumblers, and the beauty and meaning of life are in the stumbling" (David Brooks, 2016). I am forever grateful to all of you. Last but not least, I would like to thank my girlfriend and love of my life, Gina Nieh, who has given me all the support, love and care in the last 5 years. You supported me through all the highs and lows

since our undergraduate days at the University of Virginia, and I hope to be your pillar the same way moving forward.

Chapter 1

Motivation and Overview

In recent decades, retailers have witnessed rapid changes stemming from an increase in competition, the rise of e-commerce, and ever-changing consumer behavior. Retailers have become increasingly aware of the need to better coordinate inventory control with pricing in order to maximize their profitability in view of these phenomena. In this dissertation, we address two problems, both motivated by discussions with large retailers. One is concerned with inventory control decisions for settings in which planned prices fluctuate over time, and the other is concerned with pricing of multiple substitutable products for settings in which customers hold inventory as a consequence of stockpiling when promotional prices are offered. Below, we briefly summarize the motivation for each problem.

In Chapter 2, we consider inventory management for retailers that sell seasonal products and offer multiple promotional prices during the season. Our work was motivated by a problem facing a Fortune-500 clothing retailer that sells almost exclusively its own-branded products. As is common in many retail sectors, demand uncertainty is high. The retailer orders most products four to six months before the selling season, and typically places only one order for each product due to short product life cycles and long delivery lead times. The order is initially stored at the warehouse and periodically allocated to store(s) during the season. Furthermore, the retailer often plans prices and schedules price promotions well in advance of the selling season due to the need to plan for advertising and operations, including logistics. The stores are in high-rent locations, necessitating efficient use of space, so there is no backroom storage. In addition, shipping excess inventory from the store back to the warehouse is uneconomical. For these two reasons, all inventory at each retail store is available for sale at the current price. This poses a complication that does not exist in settings with constant or declining prices, namely, that it may not be optimal to provision a store with a high inventory level in a high-price period if the chance of selling it is low, thus leaving too large a supply available for sale at a lower price in the following period. Although our work is motivated by a particular retailer, the problem setting is prevalent among retailers that sell seasonal goods. We address both the inventory allocation problem, initially for a single store and then for multiple stores, and the pre-season procurement problem in such a context, with the goal of maximizing the retailer's expected profit.

In Chapter 3, we consider pricing of multiple non-perishable substitutable products in the same category over multiple time periods. Our research was motivated by discussions with major retailers whose customers visit their store(s) regularly and not only can observe the prices of all products in the category when making a purchase but also have knowledge of the pricing patterns of the retailer due to their prior shopping experiences. The retailers recognize that many customers choose among substitutable products on the basis of their relative valuations for the products and the offered prices, that some customers are deal-seekers who gain psychological benefit from purchasing at a discount, that frequent and/or deep discounting of products within the category can influence customer's willingness to pay for the products, and that customers often stockpile when products are offered at a discount. We account for these factors in our formulation of the customer's problem of choosing purchase and consumption quantities over time, whose approximate solution enables us to devise a method to solve the retailer's pricing problem.

Chapter 2

Promotion Driven Inventory Control

2.1 Introduction

Many apparel retailers intentionally change prices during a product's selling season as part of a broader promotional strategy, often offering early-season discounts followed by in-season full pricing with occasional in-season discounts, and clearance prices at the end of the season. Both early-season and in-season discounts may be synchronized with store-wide sales designed to enhance customer traffic. In such contexts, retailers face challenges in deciding seasonal ordering quantities as well as weekly allocations from the warehouse to stores.

Our work was motivated by a problem facing a Fortune-500 clothing retailer that sells almost exclusively its own-branded products. The retailer orders most products four to six months before the selling season, and typically places only one order for each product due to short product life cycles and long delivery lead times. Furthermore, the retailer often plans prices and schedules price promotions well in advance of the selling season due to the need to plan for advertising and operations, including logistics. In practice, prices would be adjusted if demand turns out to be much higher or lower than originally expected, but the methodology that we develop can be re-applied if pricing plans change. The stores are in high-rent locations, necessitating efficient use of space, so there is no backroom storage and therefore all inventory at each retail store is available for sale at the current price; it is not possible to hold back inventory, nor is it possible to send excess inventory back to the warehouse or to other retail locations due to the high cost of doing so. The lack of backroom storage also means that there is no practical option but to store most of the goods in warehouses and ship them to the stores gradually during the season. In our motivating context, each warehouse ships inventory to about 100 stores, typically once a week.

Although our work is motivated by a particular retailer, the problem setting is prevalent in both apparel retail and seasonal goods industries where competition is intensifying. Many retailers, as a result, are offering discounts more frequently, leading to non-monotonic price trajectories that are unlike those typically considered in the literature. Non-monotonic price trajectories add a new dimension to managing the risks of over- or under-ordering. Because prices fluctuate during the season, it may be optimal to temper current-period inventory allocations to take advantage of the opportunity to sell the units at a higher price later. Empty racks are not necessarily bad nor do they necessarily drive customers to competitors either immediately or in the longer term because shoppers have other items to choose from and may not expect abundant supply of deeply discounted goods. (We do not incorporate the effects of shortage-based substitution nor the effect of shortages on customer loyalty in this paper.) On the other hand, full racks during high-price periods are not necessarily good nor do they necessarily improve margins: if a high-price period is followed by a low-price period and if, when the price drops, there is excess inventory in the store that cannot be hidden in a back room, more units than desired may be sold at the low price, thus decreasing overall profit margins. Consequently, inventory allocation policies must take into account the trajectory of future prices and the likelihood of having too much inventory at the "wrong" time, as well as the usual concern of not having enough at the "right" time. Finally, optimizing the seasonal order quantity in view of the subsequent dynamic inventory allocation is also a challenge. Although there is significant research on inventory control problems for seasonal goods, to the best of our knowledge, inventory management for settings with pre-planned non-monotonic pricing has received little or no attention.

2.2 Literature Review

In this section, we provide an overview of the three streams of literature that are most closely related to our work: (i) revenue management, (ii) dynamic pricing and inventory models for seasonal products with fixed initial inventory or one or two orders whose quantities can be decided, and (iii) allocation of limited inventory to multiple stores over multiple time periods. The three streams of literature concern similar yet distinct problems. Revenue management is primarily concerned with problems in which the seller is constrained by a fixed initial inventory (e.g., number of seats on a flight or number of rooms in a hotel). All articles in the second stream consider settings with stochastic, price-sensitive demand and although quantities for one or two orders may be decisions, the emphasis is on the joint price-inventory decisions over multiple periods. The papers in the third stream emphasize "balance" in the inventory allocation process to reduce expected total shortages. Our emphasis is on articles involving finite horizon settings, as in our motivating scenario, and those involving infinite horizon settings in which allocation over a short time frame is a key factor.

2.2.1 Revenue Management

The revenue management literature is very rich, and we focus on the most relevant substream concerning how to dynamically allocate available inventory of a single resource to each demand class (generally defined by time of arrival and willingness to pay) over multiple time periods, where the price paid by each class is prespecified. The objective is to maximize the total expected profit (or revenue if the costs are fixed) over a finite horizon. This class of problems is known as *quantity-based revenue management*, and the allocation of the single resource is typically controlled via booking limits that restrict the amount of capacity/inventory that can be sold to any particular class during a given time period, and protection levels that specify the amount of capacity/inventory to reserve for a particular class or set of classes in subsequent periods.

The earliest article on quantity-based revenue management is due to Littlewood (1972). He assumes there are two customer classes, with class 1 customers arriving later but willing to pay a higher price. The problem is how much class 2 demand to accept before observing the class 1 demand realization. The well-known Littlewood rule states that it is optimal for the seller to accept class 2 requests as long as the corresponding price exceeds the expected marginal revenue of class 1 demand. During the past four decades, many researchers have extended Littlewood's results to consider more general and realistic settings. Papers by Brumelle and McGill (1993), Curry (1990), Robinson (1995), and Wollmer (1992) generalize his result to settings with an arbitrary number of demand classes. The optimal policy is typically derived by ensuring that the expected marginal value of capacity for the future periods is equal to the (known) marginal value of capacity in the current period under the allocation policy. The problem addressed by Robinson is most closely related to our work as it allows for customer classes arriving in non-overlapping intervals but not necessarily in order of increasing revenue. Robinson (1995) models the problem as a finite-horizon discrete-time dynamic program and shows that the optimal policy has a booking limit structure, i.e., accept reservations as long as the cumulative bookings do not exceed the booking limit in each period.

More recently, researchers have relaxed the assumption that different classes of customer arrive sequentially in non-overlapping intervals. Lee and Hersh (1993) and Lautenbacher and Stidham (1999) develop models that allow more than one class to arrive in the same interval, which necessitates dynamic allocation among classes, not just prescribing booking limits for each interval. Differently from the models discussed above, Lee and Hersh (1993) model demand as a stochastic process, rather than as a probability distribution of the aggregate demand for each booking class, and consider group arrivals. The authors divide the planning horizon into decision periods and develop an optimal discrete-time dynamic programming algorithm. For multiple classes without group arrivals, the authors show that the optimal policy consists of a booking limit for each customer class in each period, and for any booking capacity, there is a threshold time period after which customers of a given class will not be accepted. Finally, they show that for any period and booking capacity, there is a critical booking class that partitions customers who should be accepted and those who should not. The authors also characterize properties of the optimal policy for the case of multiple classes with group arrivals. Lautenbacher and Stidham (1999) develop a unified framework for analyzing both static and dynamic single-resource capacity control models. In their static models, demand classes arrive sequentially in non-overlapping intervals while in their dynamic models, demand from each class may arrive at any time during the horizon. They establish the optimality of a booking-limit policy for both overlapping and non-overlapping arrivals of demand classes in the absence of group arrivals.

Heuristics are often used in practice as they are easier to implement and may achieve good performance. EMSR-a (Belobaba, 1992) and EMSR-b (Belobaba and Weatherford, 1996) are two of the most commonly-used heuristics developed for the quantity-based revenue management problem in which the set of possible prices is prespecified and the decisions pertain to the quantities to make available at each price. Both heuristics are based on approximations that reduce the pertinent problem at each stage to a two-class problem to determine protection levels for future demand classes with a higher revenue per unit, but they differ with respect to the approximation scheme. EMSR-a is based on the idea of determining a protection level for each class by applying Littlewood's rule to successive pairs of classes, and then adding them to obtain the protection level for a combination of classes. EMSR-b improves upon EMSR-a by pooling demand across classes and optimizing protection levels (or analogously, booking limits) for the corresponding sets of customer classes rather than adding protection levels determined for individual classes. It typically provides better revenue performance in practice. Robinson (1995) discusses the extensive computational effort required to identify the exact solution, which requires the maximization of a quasiconcave function for each period and numerical calculations of multiple integrals. He develops a heuristic that relies on Monte Carlo integration that can be utilized to identify policies that are arbitrarily close to the optimum.

The problem that we investigate differs from the traditional quantity-based revenue management problem in several ways. First, the firm holds inventory at both a central warehouse and at retail store(s), but inventory held at the warehouse cannot be sold. Second, excess inventory is not returned to the warehouse for future reallocation. Third, there is no backroom storage at the retail locations. For these reasons, the firm faces constraints and risks that do not exist in the traditional quantity-based revenue management problem. In particular, the combination of the second and third factors means that retail locations cannot prevent excess leftover inventory at the end of the previous period from being sold at low prices, and, as a result, protection levels cannot play the same role as when a backroom is available.

2.2.2 Multi-Period Dynamic Pricing and Inventory Control for Seasonal Products with Stochastic, Price-Sensitive Demand

In this subsection, we discuss articles on pricing and inventory control in finite-horizon settings with price-sensitive stochastic demand for a single product in which a one or two orders, which limit supply for the season, are placed before the beginning of the horizon. There is a substantial literature within this stream for situations in which demand is modeled via a Poisson (or similar) process whose mean demand arrival rate is a decreasing function of price. We require a more general and flexible demand model, and therefore limit our survey to articles involving such demand models. These articles differ with respect to whether the initial inventory is exogenously determined or a decision variable, the number of price changes allowed, the specific forms of the demand functions, and whether the seller operates a two-echelon distribution system.

Monahan et al. (2004) studies a dynamic pricing problem in which demand in each period has a mean which is isoelastic in price and the demand random variable is represented as the product of this mean and a random multiplicative factor with a mean of 1. The authors formulate the problem as a dynamic program and show the equivalence between their dynamic pricing problem and a price-setting newsvendor problem with recourse. The authors demonstrate that the pricing problem can be solved as a series of static single-variable optimization problems after state space reduction. They also show that the optimal stocking factors, i.e., the unit-normalized representations of safety stock in the various periods, only depend on the number of periods remaining in the season, but are independent of the inventory available for sale. Given the inventory-level-independent sequence of optimal stocking factors, the optimal price in each period can then be recovered from the definition of the stocking factor once the leftover inventory from the previous period is observed. The authors establish the structure of the optimal pricing policy and discuss implications for the optimal stocking levels. They prove that the optimal stocking level is at least as large as the optimal stocking level for the price-setting newsvendor problem and is increasing in the number of periods remaining in the selling season. For the special case of deterministic demand, the authors demonstrate that the optimal pricing policy is a single-price policy.

Choi (2007) studies joint stocking and pricing decisions with two demand updates. The retailer can place two orders before the selling season: the first order must be placed before any demand update takes place and takes longer to arrive but is cheaper, while the second order is placed after observing demand for a different but similar product whose demand is correlated with the in-season product and whose demand depends upon its own price. Both orders of the in-season product arrive at the beginning of the season. The author formulates the two-stage stocking problem as a dynamic program and derives the optimal policy. He then considers the pricing problem and derives the optimal price during the selling season under various objectives: expected profit, probability of meeting a profit target, Value-at-Risk, and target service levels.

Tian (2010) also studies a discrete-time dynamic pricing problem for a seasonal product with fixed initial inventory and no replenishment. He models customer arrivals as a Poisson process and approximates the aggregate demand process as a Binomial process. The author uses marginal analysis and backward dynamic programming to derive the optimal pricing policy as a function of the purchase probability and remaining inventory. He shows that the discrete-time dynamic programming approach is also applicable to other common demand models such as linear demand, multinomial logit demand models, and the multiplicative competitive interaction model. His numerical results show that dynamic pricing policies dominate static pricing policies when inventory is scarce.

Chung et al. (2009) consider a variant of the newsvendor problem in which the retailer makes an in-season price adjustment at a specified review time after collecting information on demand (which depends upon the early-season price) and updating the estimate of the remaining demand during the season. The authors model (mean) demand after the review time using a linear approximation of the demand-price curve and solve for the optimal percentage change in price. The authors develop both optimal and heuristic algorithms to compute the optimal order quantity.

Chung and Flynn (2011) extend the model of Monahan et al. (2004) to incorporate inventory holding costs, and initial inventory as a decision variable. The authors show similar results to those of Monahan et al. (2004) and derive optimal solutions for two special cases: deterministic demand and nonnegative holding costs that are proportional to selling prices. For the general problem, the structural properties derived by Monahan et al. (2004) are not preserved when incorporating inventory holding costs. The authors show that it is possible to reduce the dynamic problem to a series of single-variable optimization problems. They then devise two computationally tractable heuristics that generate good solutions.

Chew et al. (2009) study the joint pricing and inventory allocation problem for a perishable product. They assume that prices increase over time and demand is price-sensitive. The authors develop a discrete-time dynamic programming model to obtain the optimal prices and inventory allocation policy for two-period problems. They were unable to derive analytical solutions for problems with three or more periods. The authors develop three heuristics and utilize an iterative procedure to update the policy in each period during the selling horizon. The heuristics achieve near-optimal results in their computational studies. The authors also study two special cases: monotonically decreasing prices, and a price trajectory that first increases and then decreases. For the first special case, the authors derive the optimal markdown pricing policy. For the second special case, the authors propose different solution procedures depending on how long prices initially increase. For the case with two periods of increasing prices, they develop an efficient search algorithm to find the optimal pricing and inventory allocation. For cases with more than two periods of increasing prices, the authors propose a hybrid procedure that combines any of their three heuristics to first determine the inventory allocation and pricing solution for the interval of increasing prices with the optimal markdown pricing policy for the remainder of the horizon.

Cachon and Kök (2007) extend the classical newsvendor model to settings with non-constant salvage values. In their problem setting, the clearance price is selected in response to the observed demand during the regular season: a deep discount is offered during the markdown period when the seller has large quantities of leftover inventory, and a much smaller discount is offered when leftover inventory is limited. The authors show that the classical newsvendor model would yield suboptimal solutions in such settings. The authors derive both the optimal clearance pricing policy and the corresponding optimal order quantity. They show that the inclusion of a nonlinear salvage value function is insufficient to correct for the classical newvendor model's deficiency, and that the correct estimate of the salvage value is critical in determining a near optimal solution. The authors then show that several intuitive estimation methods can lead to excessively large order quantities, and develop an estimation method that accounts for the interaction between decisions and demand data, and yields the optimal profit.

Our research differs from that in the articles discussed in this subsection in the following ways. First, in our research, the price trajectory is specified prior to the selling season and the firm's main levers are the procurement quantity and inventory allocation. Second, we address the problem for a two-echelon distribution system, a setting treated only by Chew et al. (2009). Our model, however, differs from his as we allow arbitrary price trajectories, including trajectories that are not monotonic nor follow an increasing then decreasing pattern, and the prices are decided in advance. Furthermore, we also consider the single-warehouse, multiple-retailer setting, allowing for nonidentical price trajectories.

2.2.3 Allocation of Limited Inventory to Multiple Stores over Multiple Periods

Our problem concerns allocation of inventory over multiple time periods from a warehouse to one or more retail locations over a finite horizon. Although our problem involves only one such external order and prices are time-varying, it does share features of inventory allocation problems in which the warehouse receives periodic shipments from a supplier and fulfills orders from retailers multiple times between supplier shipments. Jackson (1988) considers such a setting with multiple non-identical retailers and develops an approximate model of the expected shortages after the warehouse runs out of stock. With this, he determines the best (constant) order-up-to levels, one for each retailer. Jackson and Muckstadt (1989) later develop a methodology to compute optimal policies for this problem when there are *two* opportunities to allocate inventory to retailers between supplier shipments. McGavin et al. (1993) examine a case with identical retailers and two retailer replenishments per warehouse replenishment. They show that a balanced replenishment policy is optimal when there is short supply. Graves (1996) considers a related problem in which both the warehouse and retailers order at fixed intervals. He focuses on the problem of allocating stock in short supply and develops a virtual allocation heuristic that performs near-optimally in numerical tests. Considerable follow-on work has been done on this topic. In the interest of brevity, we refer the reader to Agrawal and Smith (2009), Sections 5 and 6 in Wang (2011), Marklund and Rosling (2012), and the references therein for more recent research on this topic. We emphasize that all of these papers are based on the assumption of constant prices.

There is a very substantial literature on methods to reduce a retailer's risk and costs for seasonal products with long lead times. Articles in this stream consider supply contracts, multiple orders, multiple suppliers and demand updating, among other things, but we are aware of only one that addresses inventory allocation over multiple periods. Agrawal and Smith (2013) develop a two-period inventory allocation model for a seasonal item with a constant selling price at a retail chain with nonidentical stores whose demands are correlated across the stores and periods. At the beginning of the second period, demand forecasts and inventory allocation can be updated based on the observed demands in the first period. Assuming that the store demand distributions share an unknown parameter, the authors develop a Bayesian inference model and a two-stage optimization methodology to determine the initial and revised inventory allocation policies. The authors also extend their approach to jointly determine the initial order quantity and inventory allocation policy.

Our research differs from the articles discussed above because we study the inventory allocation problem under non-monotonic price trajectories that may also differ among retailers; all of the articles mentioned in this subsection are based on the assumption of constant prices. We also consider a finite horizon problem, whereas only the Agrawal and Smith article does so. In the other articles, backlogging is the standard assumption, which is reasonable in a repeated procurement environment with non-seasonal products, but it is not realistic in our setting. For an extensive overview of procurement problems in apparel retailing, we refer the reader to Caro and Martínez-de-Albéniz (2013).

2.2.4 Our Contributions

To the best of our knowledge, our work is the first to address a retailer's inventory management problem for a seasonal product when the inventory is gradually shipped to retail locations during the season, prices are planned in advance and non-monotonic over time and consequently, the demand distributions are also non-stationary. Two additional realistic but complicating features in our problem setting are that excess inventory is not returned to the warehouse and there is no backroom storage at the retail locations. In our problem context, there are two types of operational decisions: (i) how much to procure prior to the selling horizon, and (ii) how much inventory should be transferred from the central warehouse(s) to each store in each period.

We first devise a heuristic for setting order-up-to levels based on equating the marginal value of an incremental unit of inventory across all periods for the single-store setting that is intuitive and fairly easy to implement. In a numerical study, we compare the expected profit from our heuristic solution against a theoretical upper bound which is based on a strict relaxation of our problem, as well as an industry benchmark. The results show that our heuristic performs well under most conditions. We explain how our heuristic can be generalized to consider multiple retail locations. We then address the problem of determining the initial order quantity in view of the subsequent allocation of inventory over the selling horizon. Because the optimal solution is elusive, we develop a heuristic based on a modified newsvendor model that explicitly accounts for the two-tiered nature of salvage values—a positive but low markdown price late in the selling horizon and zero or minimal salvage value for any units remaining at the end of the horizon. Our modified newsvendor approach performs surprisingly well in many settings despite its simplicity.

2.3 Problem Statement

We consider a retailer, facing a long production lead time and a short product life cycle, who procures and sells a single seasonal good over periods $t = 1, \ldots, T$ at possibly non-identical stores indexed $n = 1, \ldots, N$. For each store n, the prices during the selling horizon, $\{p_t^n\}_{t=1}^T$, are given and fixed. We assume that there is at least one price increase during the selling horizon, and there is a markdown regime for one or more periods at the end of the horizon. Demand at store n in period t is random and denoted by the random variable D_t^n and its realization is denoted by d_t^n. D_t^n has a probability density function $f_t^n(\cdot)$ and cumulative distribution function $F_t^n(\cdot)$. We assume demands are independent across periods. Although there is likely to be mild negative correlation between proximate time periods, few retail firms that sell goods such as apparel can estimate correlation effects due to the presence of an array of confounding factors, and consequently, they consider assuming independence to be a practical choice. For similar reasons, we assume demand is independent across stores. As is typical in settings like our motivating example, we assume excess demand is lost as consumers can find comparable options at competitors. We also assume that the effects of any advertising or other forms of promotion are accounted for in the parameters of the demand distributions in the various periods.

Prior to the selling horizon, the retailer places an order, whose quantity is denoted by x_1^0, with an external supplier. The unit ordering cost is c. The order is shipped and stored at the warehouse before the season begins. Inventory in the warehouse and at store n at the beginning of period t are denoted by x_t^0 and x_t^n, respectively. We assume that a shipment from the warehouse to any store n requires zero lead time and occurs at the beginning of each period t prior to the demand realization; the shipment quantity is denoted by $q_t^n \geq 0$. For most apparel retailers, store replenishment from a local or regional warehouse is relatively rapid, so the zero lead time assumption is a reasonable approximation. We assume shelf space at the retailer does not restrict the inventory, although our heuristic can be modified to account for such a constraint.

We assume that the retailer incurs per unit per period inventory holding costs at both the warehouse (h_w) and the retail locations (h_s), and that the latter is larger than the former because shrinkage and damage tend to be higher at retail locations. The unit price in period t, p_t, may vary from period to period. If there is a lost sale penalty beyond the loss of gross margin, it can be included in p_t. We assume that the salvage value of any units remaining at the end of the horizon is zero, but our model can accommodate any positive constant provided that all remaining units can be sold at that price. The retailer seeks to maximize expected profit considering revenue, purchase costs, and inventory holding costs at both the warehouse and the retail locations.

The most general version of our problem is quite complicated, so we first focus on the simpler single-warehouse single-store setting to develop some insights before discussing the extension to multiple retail stores. We formulate our problem for the single-store setting with fixed initial inventory in Section 2.3.1 and present our heuristic for this problem in Section 2.3.2. A numerical study to evaluate our heuristic in the single-store case appears in Section 2.4. We analyze the one-warehouse multi-store setting in Section 2.5 and the procurement problem in Section 2.6. We conclude the paper in Section 2.7.

2.3.1 Single Warehouse Single Store Multi-Period Inventory Allocation Problem with Fixed Initial Order Quantity

In this section, we discuss the inventory allocation problem with a fixed initial order quantity in the single-warehouse, single-store setting. Our problem is a multi-period stochastic optimization problem that can, in principle, be solved in many ways, including stochastic programming, (approximate) dynamic programming, simulation, and stochastic search (Powell, 2014). To make the problem concrete, we first present a dynamic programming (DP) formulation and discuss why, even beyond the issue of state space explosion, the problem is difficult to solve as a DP. We then present a heuristic procedure that builds upon ideas from quantity-based revenue management but accounts for the special features of our problem. We also present results of numerical studies in which we compare the performance of solutions from our heuristic against those of two benchmark solutions.

We can formulate the problem as a dynamic program in which the state of the system in each period is defined by the levels of inventory at the warehouse and at the store at the beginning of the period. Let x_t^0 and x_t^1 be the inventory at the warehouse and store 1, respectively, at the beginning of period t. We assume the initial inventory at the warehouse, x_1^0, is a constant, C, and the initial inventory at the store, x_1^1, is zero. Let $V_t(x_t^0, x_t^1, q_t)$ be the expected gross margin starting with x_t^0 (x_t^1) units of inventory at the warehouse (store) at the beginning of period t if q_t is sent from the warehouse to the store. We use the term gross margin throughout the remainder of the paper when referring to the retailer's objective considering the cost of holding inventory, so as to distinguish it from the usual definition of profit in inventory-pricing models. The dynamic program can be expressed as follows:

$$\begin{aligned} V_t(x_t^0, x_t^1, q_t) &= \mathbb{E}_{D_t}[-h_w(x_t^0 - q_t) - h_s(x_t^1 + q_t - D_t)^+ + p_t \min(x_t^1 + q_t, D_t) \\ &\quad + V_{t+1}^*(x_t^0 - q_t, (x_t^1 + q_t - D_t)^+)] \quad \forall t \in \{1, \ldots, T\}, && (2.1) \\ V_{T+1}(x_{T+1}^0, x_{T+1}^1) &= 0, && (2.2) \\ V_t^*(x_t^0, x_t^1) &= \max_{q_t | 0 \le q_t \le x_t^0} V_t(x_t^0, x_t^1, q_t) \quad \forall t \in \{1, \ldots, T\}, && (2.3) \end{aligned}$$

where the inventory state transitions at the warehouse and store are governed by $x_{t+1}^1 = (x_t^1 + q_t - d_t)^+ \ge 0$ where d_t is the observed demand in period t, and $x_{t+1}^0 = x_t^0 - q_t \ge 0$. In (2.1), the first two terms are the cost of holding inventory at the warehouse and store, respectively, and the third term is the expected revenue, while the last term is the optimal value to go from period $t+1$ onward. Expression (2.2) is the end-of-horizon boundary condition, and (2.3) states the optimal value function, where the optimization is with respect to the order quantity, which is constrained by the inventory at the warehouse. There is a non-negativity constraint on the order quantity, q_t, in each period. Thus, the store inventory level after replenishment in each period is the maximum of the leftovers and the achieved order-up-to level, where the latter may be constrained by available inventory at the warehouse.

Deriving the optimal solution exactly is extremely difficult. Apart from the usual challenges due to the size of the state space, the DP presented above poses additional difficulties because the single-period value functions may have flat sections when the demand distributions do not have infinite right tails; this, in turn, may lead to discontinuous derivatives. Furthermore, even if the single-period value functions do not have such flat sections, without restrictive assumptions on the form of the demand distributions, it is not possible to guarantee that the value function is sufficiently well behaved to enable identification of optimal policies from first-order necessary conditions alone.

2.3.2 Marginal Value Analysis (MVA) Based Heuristic Algorithm

Although computing the optimal policy is difficult, we can construct a good heuristic by taking advantage of the specific features of our problem. We assume a time-dependent order-up-to policy is implemented and develop a heuristic, which we call the Marginal Value Analysis (MVA) heuristic, to determine good order-up-to levels. The heuristic is based on the principle that available inventory should be allocated across the periods such that the incremental benefit from the marginal unit of inventory is the same in all periods; otherwise, a reallocation would increase expected gross margin.

Under an order-up-to policy, due to demand uncertainty, even if the order-up-to level is less than the mean demand, the expected end-of-period inventory is positive, which in turn implies that the expected allocation of inventory in the subsequent period is less than that period's order-up-to level. For this reason, one cannot simply allocate the initial inventory among order-up-to levels. Furthermore, the presence of time-dependent demand distributions and prices can lead to situations in which the commonly-assumed outcome for a leftover unit, i.e., being sold in the subsequent period, does not occur. This can arise, for example, if the order-up-to level in the current period is much higher than that of the subsequent period and demand in the current period turns out to be very low. In such a case, an incremental unit of inventory in the current period could become an excess unit—above and beyond the order-up-to level—in the subsequent period, and may have a very low chance of being sold at that time. When computing an estimate of the incremental value of allocating inventory to a period, we account for this in an approximate fashion.

The most sophisticated (dynamic) version of our heuristic is a telescoping-horizon procedure in which iteration i, $i = 1, \ldots, T$, generates a solution for periods i through T. In each iteration, the first phase of the heuristic involves constructing an initial solution (set of order-up-to levels) essentially by iteratively assigning inventory to maximize expected marginal gross margin in a myopic fashion, temporarily ignoring the impact of leftover inventory. The resulting solution is the same as what would be obtained by solving an asymmetric multi-market newsvendor problem with an aggregate inventory constraint, where prices differ across the markets. The analogy is as follows: The time periods in our model correspond to markets, which are independent and have asymmetric prices. The retailer has a fixed supply of the product to allocate across the markets.

Because demand is uncertain, we cannot know the levels of leftover inventory *a priori*, so we incorporate the effects of *expected* leftover inventory in an adjustment process that accounts for the fact that the inventory constraint applies to the sum of the shipment quantities, not to the order-up-to levels. The adjustment procedure entails computing and totaling the expected leftover inventory in all periods and iteratively allocating this inventory, one unit at a time, to the period with the highest incremental benefit. The second phase entails another series of adjustments to the order-up-to levels: in each adjustment step, the order-up-to level in the period with the highest expected marginal value of inventory is increased by one unit (or a managerially-determined increment) and the one in the period with the lowest is decreased, where the estimate of the expected marginal value of inventory accounts for the phenomena that we described earlier. (We present mathematical details later in this section.) Adjustments are made until the variance among the expected marginal values cannot be reduced further. The set of order-up-to levels after this series of adjustments forms the *simple static* heuristic solution for the selling horizon under consideration. The retailer utilizes these order-up-to levels as the basis for shipping inventory to the retailer in the period so long as there is sufficient inventory at the warehouse. If there is insufficient inventory at the warehouse, the policy is to ship as much as possible. We later explain an improved version of the simple static heuristic that requires very little additional effort in implementation.

In the dynamic version of the heuristic, the retailer implements the heuristic solution only for the current period. After observing demand in the period and fulfilling it as much as possible, the retailer recomputes the heuristic solution. Starting with the updated inventory level, the retailer implements the initialization and adjustment phases of the heuristic to update the allocation plan for the remainder of the horizon. After the order-up-to level for period i is finalized following the second phase adjustment, the retailer uses the order-up-to level in period i to determine how much to ship, observes demand in period i and then moves on to the next period and repeats this process in each successive period.

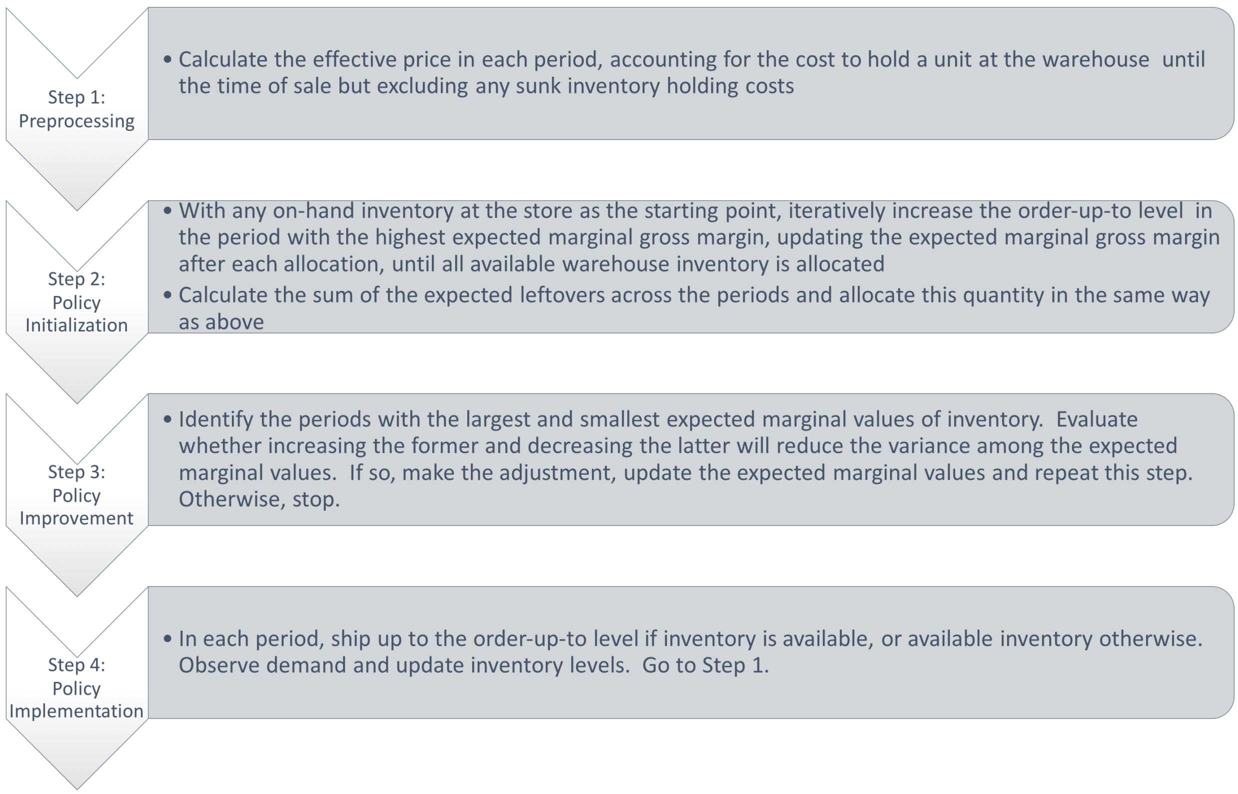


Figure 2.1 High-Level Flowchart of Dynamic MVA

For computational convenience, we assume that quantities are integer-valued, which is also true in our motivating scenario. To calculate the expected marginal value of inventory in each period, we need to approximate the DP value function at the current order-up-to level, S_t, as well as at an order-up-to level one unit (or one increment, although we use one unit for ease of exposition) larger. In our heuristic, which we describe in more detail in the next subsection, at each iteration, we utilize approximations of the value function that are based on the assumption that the current vector of order-up-to values, which we denote by $\boldsymbol{S}$, is valid, but because one order-up-to value is increased and another is decreased at the end of each iteration, the computed marginal value in each period is based on other order-up-to levels that may change in both the current and in future iterations. In particular, the marginal value of inventory in any period t depends heavily on the marginal value in the few succeeding periods and our algorithm does not account for the fact that the order-up-to levels in those periods may change due to the change in the order-up-to level in period t. A high-level flowchart of the dynamic version of our heuristic appears in Figure 2.1.

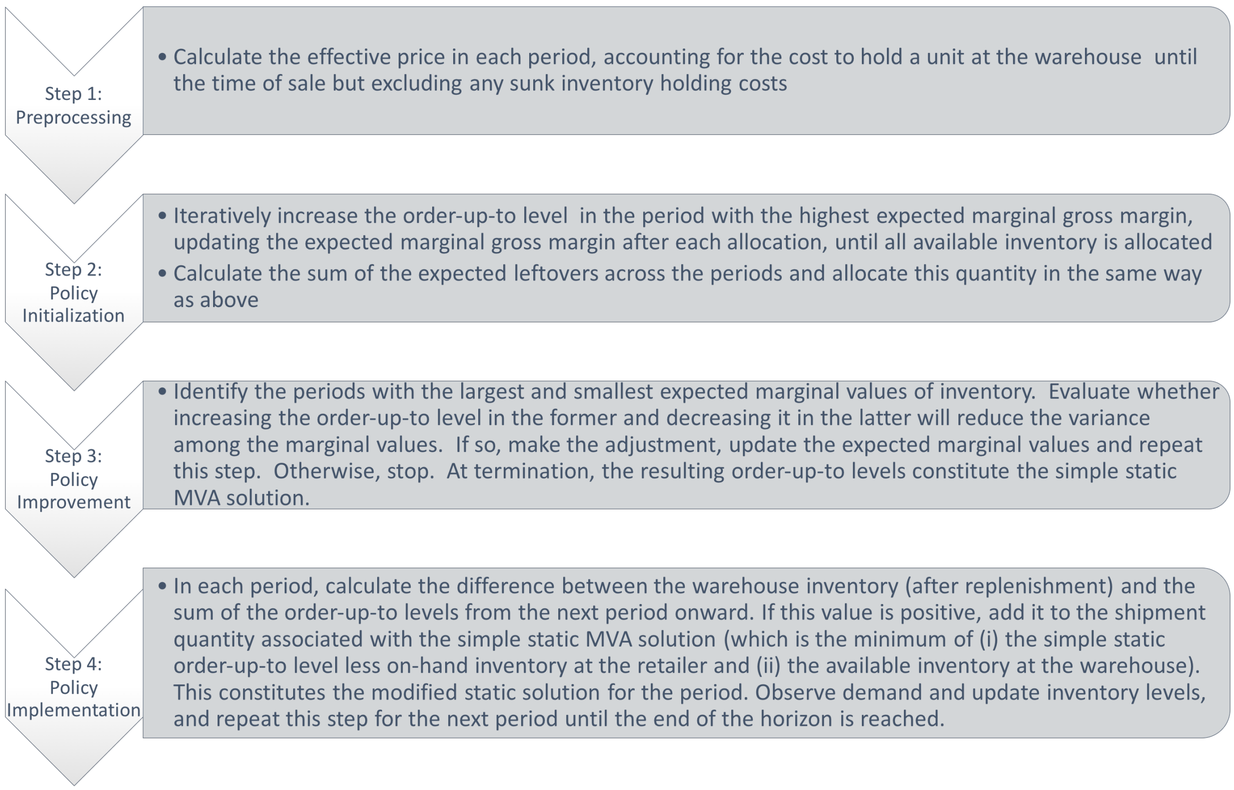


Figure 2.2 High-level Flowchart of *Modified* Static MVA

Retailers may not be able to dynamically update the order-up-to levels in response to actual demands. If updating is too difficult, the retailer could simply apply the simple static heuristic solution by utilizing the order-up-to levels as the basis for shipping inventory to the retailer, so long as there is sufficient inventory at the warehouse. However, when initial inventory is ample and/or demand in the initial periods is lower than expected, a completely static policy can cause the retailer to forgo considerable gross margin, as we observed in preliminary tests. To ameliorate this problem, we propose an adjustment step that would be executed only if the sum of order-up-to levels in the remaining $T - t$ periods is smaller than the warehouse inventory level after replenishment. If this occurs, the adjustment entails sending the store an additional $y_t - \sum_{i \in \{t+1,\ldots,T\}} S_i$ units of inventory, where y_t is the number of units available at the warehouse after replenishment. Observe that this is a conservative allocation because it protects enough inventory at the warehouse to cover the sum of the order-up-to levels in subsequent periods. As noted earlier, the expected leftover inventory is always positive due to our assumption of lost sales, so, in expectation, there will be more inventory available to allocate, even after we implement this adjustment. We note that the need for these adjustments arises only if initial inventory is far above total expected demand or if demand turns out to be much lower than expected. Also, the calculations required for these adjustments are extremely simple and there is no need to perform any MVA-like calculations. A high-level flowchart of the *modified static* version of our heuristic appears in Figure 2.2.

2.3.2.1 Marginal Value Analysis Policy Improvement Step

In this subsection, we describe the policy improvement step which is implemented in both the second phase of the first iteration (which culminates in the simple static solution) as well as in the dynamic version of our heuristic in which order-up-to levels are adjusted after each demand is observed. We seek to reallocate available inventory across the periods such that marginal values of inventory in the various periods are as equal as possible. In order to do so, we must calculate the expected marginal value of inventory in each period for the current order-up-to vector, $\boldsymbol{S}$. The marginal value of inventory in a period is difficult to express exactly as it depends on not only the expected marginal gross margin if the incremental unit is sold in that period, but also the revenue and cost consequences if it is not sold and is carried forward to the next period.

We now present an approximation of the expected marginal value of inventory for any given order-up-to vector, $\boldsymbol{S}$. Let the expected marginal value of the S_t^{th} unit of inventory in period t be denoted by $\Delta_t(S_t)$. For ease of exposition, we omit the dependence on the other elements of $\boldsymbol{S}$. Let $\tilde{p}_t = p_t - h_w(t-1)$ represent the effective price per unit in period t, which is the selling price in period t less the cost of holding the unit in the warehouse until period t; given that the initial inventory is fixed and thus, the variable purchase costs are sunk, this is the maximum gross margin one can obtain from selling a unit in period t. We also define $h_d = h_s - h_w$ as the difference between holding a unit for one period at the retail store and at the warehouse.

Our formula for $\Delta_t(S_t)$ accounts for all of the possible revenue and cost outcomes associated with the S_t-th unit of inventory, some of them approximately. We first present the formula and then explain each term.

$$\begin{aligned}\Delta_t(S_t) := {} & \tilde{p}_t Pr(D_t \geq S_t) - h_d Pr(D_t < S_t) + \Delta_{t+1}(S_{t+1}+1)) Pr(D_t < S_t - S_{t+1}) \\ & + \max_i \{\Delta_i(S_i + 1) \mid i > t\} Pr(S_t - S_{t+1} \leq D_t < S_t)\end{aligned} \tag{2.4}$$

The first term is the expected marginal gross margin from the incremental unit, accounting for cases in which the observed demand exceeds the current S_t. The second term is the expected holding cost incurred due to holding the incremental unit at the store rather than the warehouse, accounting for situations in which observed demand is less than S_t (so the incremental unit would not be sold). The third term accounts for situations in which $S_t >$

S_{t+1} and observed demand is less than $S_t - S_{t+1}$. In such cases, leftover inventory in period t would exceed the order-up-to level in period $t+1$ and an incremental unit of inventory beyond $S_{t+1}+1$ units would be expected to generate no more than $\Delta_{t+1}(S_{t+1}+1)$ in value, because successive units allocated to period $t+1$ provide weakly decreasing benefit. The third term is the probability of the aforementioned type of demand occurring multiplied by the upper bound on the incremental value, $\Delta_{t+1}(S_{t+1}+1)$. The last term accounts for situations in which a leftover unit is simply carried forward to the next period without being a unit in excess of S_{t+1}, which would occur when demand is large enough that the leftover inventory does not exceed S_{t+1}. One may think that the value of the incremental unit in period t in such situations would be simply $\Delta_{t+1}(S_{t+1}+1)$, but it is not. The reason is that the availability of the incremental unit in period $t+1$ enables one to "save" one of the units at the warehouse for the best future period. This is why the fourth term contains the expression $\max_i\{\Delta_i(S_i+1) \mid i > t\}$ to capture the incremental value associated with the best future period. This value is multiplied by the probability of demand being in the proper interval for this economic outcome to be accurate. One implication of the above analysis is that the consequences of sending an "extra" unit to the store are minimal so long as the incremental unit does not cause an overshoot of the order-up-to levels in subsequent periods and does not cause depletion of warehouse inventory too early (which would make it impossible to ship enough to match the order-up-to levels in future periods, whose prices may be higher).

If the objective function is jointly unimodal in the S_i values, the swaps in the policy improvement routine would guarantee an optimal allocation under our approximation of the marginal values. In reality, notable discontinuities (not just those due to the integrality of the S value) in the approximate marginal values can arise when the swaps occur between consecutive periods if the swap reverses the relative magnitudes of the S_t values. We have found that despite this limitation, our approximation is adequate for making the necessary comparisons.

During the policy improvement phase, we first compute the variance among the initial $\Delta_t(S_t)$ values given the order-up-to vector, $\boldsymbol{S}$, after the policy initialization phase. For each iteration in the policy improvement phase, we identify improvement opportunities as follows. Let

$$j = \max_t\{\Delta_t(S_t+1),\ \forall t\} \text{ and } k = \min_t\{\Delta_t(S_t),\ \forall t \in \{i|S_i \geq 1, \forall i\}\}.$$

Then j indexes the period with the largest expected marginal value and k indexes the period with the smallest expected marginal value among the periods whose order-up-to levels are strictly positive. If $\Delta_j(S_j+1) > \Delta_k(S_k) \geq 0$, an improvement opportunity exists and we increase S_j by one unit and decrease S_k by one unit. We then recompute the $\Delta_t(S_t)$ values and the variance among these values. This process continues and we search for improvements as long as the variance is decreasing and/or an improvement opportunity exists; otherwise the process terminates.

2.3.2.2 An Illustrative Example

Before presenting our numerical study, we utilize an example to illustrate the four steps of the dynamic version of our MVA heuristic. The problem parameters are: $T = 3$, $N = 1$ $\mathcal{C} = 4$, $x_1^0 = 4$, $x_1^1 = 0$, $h_w = 1$, $h_s = 2$, $p_1 = 24$, $p_2 = 31$, $p_3 = 12$, and $Pr(D_1(p_1) = i) = 0.25$, $i \in \{0, \ldots, 3\}$, $Pr(D_2(p_2) = i) = 0.25$, $i \in \{0, \ldots, 3\}$ and $Pr(D_3(p_3) = i) = 0.20$, $i \in \{0, \ldots, 4\}$. We show how to implement the dynamic version of the MVA heuristic for demand path $\{2, 2, 3\}$, which is initially unknown to the retailer.

For $t = 1$:

Step 1: Preprocess

We first determine initial systemwide inventory and effective prices as follows:

$$\mathcal{C} = x_1^0 + x_1^1 = 4;$$

$$\tilde{p}_1 = 24; \tilde{p}_2 = 31 - 1 = 30; \tilde{p}_3 = 12 - 1(2) = 10.$$

Step 2: Policy Initialization

We calculate the initial S_t values by iteratively allocating inventory to the period with the highest expected marginal gross margin, $\tilde{p}_i Pr(D_i > S_i)$, using a greedy procedure. Table 2.1 shows the necessary calculations.

Table 2.1 Step 2: Calculations for Initializing $\boldsymbol{S}$ in Period 1

Iteration	$\mathcal{C}$	S_1	$\tilde{p}_1 Pr(D_1 > S_1)$	S_2	$\tilde{p}_2 Pr(D_2 > S_2)$	S_3	$\tilde{p}_3 Pr(D_3 > S_3)$
1	4	0	18	0	22.5	0	8
2	3	0	18	1	15	0	8
3	2	1	12	1	15	0	8
4	1	1	12	2	7.5	0	8
5	0	2	6	2	7.5	0	8

At the beginning of iteration 1, there are 4 units of inventory available. We calculate the expected marginal gross margin for each period and increase S_2, the order-up-to level in the period with the highest such value, by one unit. We repeat the process iteratively, calculating $\tilde{p}_i Pr(D_i > S_i)$ for all i and allocating inventory to the period with the highest expected marginal gross margin, until all units of inventory are assigned.

We then take into account the positive expected end-of-period inventory. Under the current order-up-to level policy, $\{S_t\}_{t=1}^3$, expected sales is: $\sum_{t=1}^3 \mathbb{E}[\min(D_t, S_t)] = 2.5 < \mathcal{C}$. We compute the total expected leftover inventory across all periods, and iteratively allocate the aggregate leftovers, rounding up to the nearest integer, in the same way as described above until total expected sales reaches the initial inventory level. This step ensures that, in expectation, all of the initial inventory will be sent to the store during the selling season. We show the necessary calculations in Table 2.2.

Table 2.2 Step 2: Calculations for Initializing $\boldsymbol{S}$ (Adjustment) in Period 1

Iter.	S_1	$\tilde{p}_1 Pr(D_1 > S_1)$	S_2	$\tilde{p}_2 Pr(D_2 > S_2)$	S_3	$\tilde{p}_3 Pr(D_3 > S_3)$	Expected Sales
5	2	6	2	7.5	0	8	2.5
6	2	6	2	7.5	1	6	3.3
7	2	6	3	0	1	6	3.55
8	3	0	3	0	1	6	3.8
9	3	0	3	0	2	4	4.4

After the adjustment procedure, the order-up-to vector, $\boldsymbol{S}$, is $\{3, 3, 2\}$.

Step 3: Policy Improvement

In the policy improvement step, we refine the order-up-to policy, taking into account all possible revenue and cost outcomes and their associated probabilities of occurring. We calculate the expected marginal value of inventory in each period under the order-up-to levels, $\{\Delta_t(S_t)\}_{t=1}^3$, and the expected marginal value of inventory in each period t if an additional unit of inventory is assigned, $\{\Delta_t(S_t+1)\}_{t=1}^3$. With these values, we can determine whether increasing the order-up-to level by one unit in one period and decreasing it by one unit in another will reduce the variance among the resulting $\Delta_t(S_t)$ values, in which case we would expect an improvement in the total expected gross margin. The improvement procedure terminates when no further opportunities exist. Table 2.3 shows the necessary calculations.

Table 2.3 Step 3: Calculations for Improvement Procedure in Period 1

Iteration	S_1	$\Delta_1(S_1), \Delta_1(S_1+1)$	S_2	$\Delta_2(S_2), \Delta_2(S_2+1)$	S_3	$\Delta_3(S_3), \Delta_3(S_3+1)$
1	3	7.05, 1.4	3	9.3, 2.4	2	5.6, 3.4

At this point, there are no improvement opportunities as the marginal values for assigning an additional unit, $\Delta_t(S_t+1), \forall\, t$, are strictly less than the marginal value of inventory under the current order-up-to levels, so increasing one order-up-to level and decreasing another cannot improve the solution.

Step 4: Inventory Allocation

$S_1 = 3$, $\mathcal{C} = 4$, $x_1^0 = 4$ and $x_1^1 = 0$ so the retailer sends three units of inventory to the store, i.e., $q_1 = 3$.

For $t = 2$:

At the beginning of period 2, the retailer determines an updated heuristic solution by implementing Steps 1 through 4 described above. For brevity, we present an abbreviated discussion below.

Step 1: Preprocess
We have $d_1 = 2$, so

$$x_2^0 = 1 \text{ and } x_2^1 = 1 \text{ so } \mathcal{C} = 2$$

$$\tilde{p}_2 = 31; \tilde{p}_3 = 12 - 1 = 11$$

Note that $\tilde{p}_2$ and $\tilde{p}_3$ differ from those utilized in period 1 as the cost of holding inventory during period 1 is now sunk.

Step 2: Policy Initialization
Table 2.4 presents the necessary calculations.

Table 2.4 Step 2: Calculations for Initializing $\boldsymbol{S}$ in Period 2

Iteration	$\mathcal{C}$	S_2	$\tilde{p}_2 Pr(D_2 > S_2)$	S_3	$\tilde{p}_3 Pr(D_3 > S_3)$
1	2	0	23.25	0	8.8
2	1	1	15.5	0	8.8
3	0	2	7.75	0	8.8

In iteration 1, we assign a unit of inventory to period 2 because the expected marginal value of inventory is higher than that of period 3. We repeat the calculation and assignment until all available units of inventory are assigned. We must once again take into account the positive expected end-of-period inventory effects. Under the current order-up-to level policy, $\{S_t\}_{t=2}^{3}$, expected sales is: $\sum_{t=2}^{3} \mathbb{E}[\min(D_t, S_t)] = 1.25 < \mathcal{C}$. We compute the total expected leftover inventory in all periods, and iteratively allocate inventory until expected sales is greater or equal to the initial inventory level. We show the necessary calculations in Table 2.5.

Table 2.5 Step 2: Calculations for Initializing $\boldsymbol{S}$ (Adjustment) in Period 2

Iteration	S_2	$\tilde{p}_2 Pr(D_2 > S_2)$	S_3	$\tilde{p}_3 Pr(D_3 > S_3)$	Expected Sales
3	2	7.75	0	8.8	1.25
4	2	7.75	1	6.6	2.05

After the adjustment procedure, the order-up-to vector for the remaining periods in the horizon, $\boldsymbol{S}$, is $\{2, 1\}$.

Step 3: Policy Improvement

Table 2.6 presents the necessary calculations.

Table 2.6 Step 3: Calculations for Improvement Procedure in Period 2

Iteration	S_2	$\Delta_2(S_2), \Delta_2(S_2+1)$	S_3	$\Delta_3(S_3), \Delta_3(S_3+1)$
1	2	18.6, 11.65	1	8.6, 6.2
2	3	13.45, 7.6	0	11, 8.6

Because $\Delta_2(S_2+1) > \Delta_3(S_3)$ under the order-up-to policy obtained in Step 2, an improvement opportunity exists; we increase S_2 by 1 and decrease S_3 by 1 to improve the order-up-to policy. MVA terminates as there are no further improvement opportunities.

Step 4: Inventory Allocation

$S_2 = 3$, $\mathcal{C} = 2$, $x_2^1 = 1$ and $x_2^0 = 1$ so the retailer sends 1 unit of inventory to the store, i.e., $q_t = 1$. We have $d = 2$, so all inventory is depleted and the MVA heuristic terminates. The inventory in period 3 is zero.

2.4 Numerical Study for the Single-Store Setting

In this section, we report on the performance of the MVA heuristic for a variety of realistic numerical examples and explore the impact of the price trajectory and the pattern of demand uncertainty over the course of the horizon. We first consider MVA's performance when the retailer faces different price trajectories with the same coefficient of variation of demand uncertainty throughout the horizon. We then explore the robustness of the heuristic to different patterns of demand uncertainty over the horizon. We compare MVA to two benchmark policies, one based on a theoretical upper bound and one industry benchmark.

The theoretical upper bound is based on a relaxation of our problem in which all of the inventory is held at a single location from which it can be sold and there is infinite backroom storage. This relaxation is equivalent to the revenue management problem addressed by Robinson (1995), which allows time-varying prices and for which the optimal solution can be expressed analytically and computed efficiently. In the remainder of the chapter, we refer to the upper bound obtained from Robinson's algorithm as the Robinson upper bound.

The industry benchmark that we use is the best constant service level policy (BCSLP), where service level is defined as the probability of not stocking out (at the retailer) in a period. The firm that motivated our research has been using a constant service level policy with a managerially-selected service level. The BCSLP is the best of such policies obtained via a grid search over a wide range of service levels. As such, it provides a practical upper bound on what the firm could hope to achieve. In all of our numerical examples, we consider service levels from 60% to 97.5% in 2.5% increments.

In all of our numerical studies, we perform 100 simulation runs for each combination of problem parameters from which we estimate the expected gross margin of any policy, accounting for revenue, inventory holding costs and unit costs, by calculating the average across the simulation runs.

2.4.1 Impact of Price Trajectories on MVA Performance

We now report on the performance of MVA for four price trajectories: (1) high-low pricing, (2) regular price with pre-season and in-season promotions, (3) markdown pricing and (4) high-amplitude non-monotonic price trajectory. The first three price trajectories are similar to those seen in practice, while the last price trajectory poses a challenging combination of conditions.

The four examples differ in their price trajectories well as the horizon lengths, but share the following parameters: $N = 1$, $c = 25$, $h_w = 0.25$, $h_s = 0.5$, $D_t = Ap_t^{-b}\xi_t$, where $A = 100000, b = -1.5$, $\xi_t \sim \mathcal{N}(\mu, \sigma^2)$, where $\mu = 1, \sigma/\mu = 0.3$ and $\alpha = 0.5, 0.6, \ldots, 1.5$. The demand random noise terms, ξ_t, are assumed to be independently and identically distributed. For each α, we set the order quantity, C, to α times the sum of expected demands in profitable periods. We note that if the price in any period is equal to the purchase cost, then, because retailer incurs inventory holding costs, any inventory sold in that period incurs a net loss. The parameter α can be interpreted as the fraction of profitable demand covered by the order from the supplier, and we henceforth refer to it as the stocking factor. The prices and demand parameters for the four examples appear in Tables 2.7 through 2.10.

Table 2.7 Prices and Demand Parameters for Example 1

Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.9
2	40	395.3	118.0
3	50	282.8	84.9
4	40	395.3	118.0
5	30	608.6	181.6

Table 2.8 Prices and Demand Parameters for Example 2

Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	40	395.3	118.0
2	50	282.8	84.9
3	50	282.8	84.9
4	40	395.3	118.0
5	50	282.8	84.9
6	50	282.8	84.9
7	50	282.8	84.9
8	20	1118.0	335.4

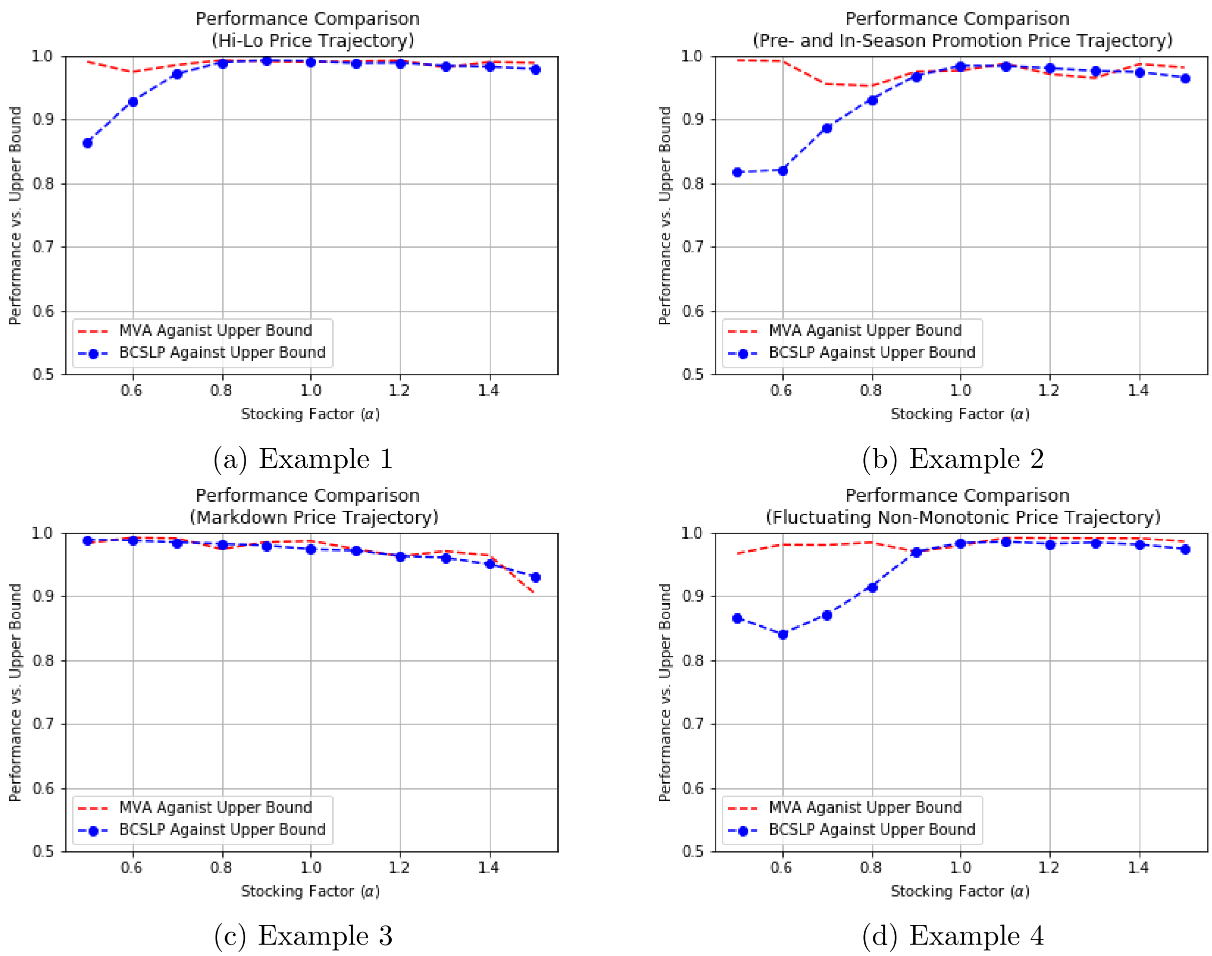
Table 2.9 Prices and Demand Parameters for Example 3

Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.4
2	50	282.8	84.4
3	45	331.3	99.4
4	45	331.3	99.4
5	40	395.3	118.6
6	40	395.3	118.6
7	30	608.6	182.6
8	30	608.6	182.6
9	20	1118.0	335.4

Table 2.10 Prices and Demand Parameters for Example 4

Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.9
2	60	215.2	64.5
3	60	215.2	64.5
4	40	395.3	118.6
5	50	282.8	84.9
6	60	215.2	64.5
7	50	282.8	84.9
8	20	1118.0	335.4

Results for the four examples are shown in Figures 2.3(a), 2.3(b), 2.3(c), and 2.3(d), respectively. In each subfigure in Figure 2.3, we plot the ratio of the expected gross margin from MVA (red dashed line) and BCSLP (blue dashed line with dots as markers) to that from the Robinson upper bound at each stocking factor level. The results suggest that the performance of BCSLP is comparable to that of MVA for settings in which it is not optimal to hold back inventory due to price changes, such as in the case of monotonically non-increasing prices, but the performance of MVA is superior for settings with non-monotonic trajectories. In the third example, which has monotonically non-increasing prices, unless inventory holding costs are extremely high, there is no incentive for the retailer to withhold inventory for future periods, as prices weakly decline over the horizon. In other words, the retailer can do no better than if he sets high service levels throughout the selling season in order to sell the units as soon as possible.



(a) Example 1

(b) Example 2

(c) Example 3

(d) Example 4

Figure 2.3 Gross Margin Performance of MVA and Best Constant Service Level Policy versus an Upper Bound

Under non-monotonic price trajectories, however, the retailer can reap significant benefits by intelligently holding back inventory for future high-price periods. Using MVA, the retailer implicitly sets varying service levels in alignment with the selling price in the current and immediately succeeding period, while under BCSLP, the retailer sets a "one-size-fits-all" service level for the entire horizon, thus making too much inventory available for purchase in low-price periods, thereby putting at risk availability during later high-price periods. When the stocking factor is intermediate to high, MVA and BCSLP perform similarly but the difference in performance is particularly significant when the retailer starts with low to intermediate levels of the stocking factor. In practice, retailers offer thousands of distinct products and face a budget constraint when deciding initial order quantities, and as such, low to intermediate stocking factor levels are common—exactly the conditions under which careful inventory allocation is most valuable.

To evaluate the statistical significance of the difference between the expected gross margin from the MVA solution and that from the benchmark(s), we perform the sign test (Howell, 2007), which is a non-parametric test that requires very few assumptions about the distribution of the data. In our study, the null hypothesis is that MVA and a benchmark achieve the same gross margin. If the null hypothesis is true, one would expect MVA to outperform a benchmark in 50% of the simulation runs and we would expect the number of simulation runs in which MVA outperforms a benchmark in a total of m simulation runs to follow the binomial distribution $Bin(m, 0.5)$. We can therefore use the standard binomial test to calculate statistical significance. In the interest of brevity, we do not report details. However, when examining our results, it is safe to infer that when MVA outperforms a benchmark by 5% on the average, the null hypothesis can be rejected at the 5% significance level.

To evaluate MVA against a practical approximation of the optimal solution, for six three-period problems (two price trajectories and three stocking factor levels), we compare solutions obtained using MVA with the best solution that can be found via a grid search over a ranges of order-up-to levels in each period. The results show that the expected gross margin for solutions from MVA are within 0.5% of those obtained from the grid search. Further details can be found in Appendix A.

2.4.2 Impact of Degree and Pattern of Demand Uncertainty on MVA Performance

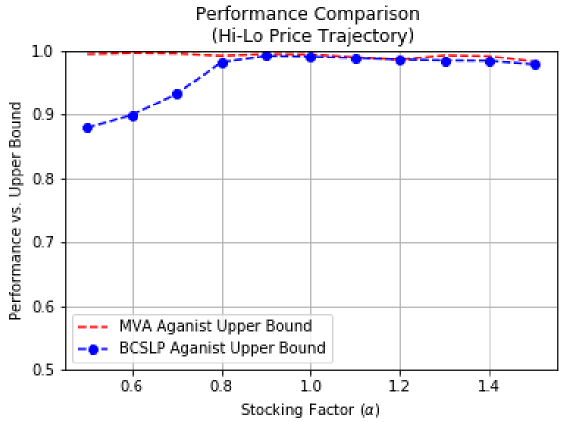
Results from the numerical examples thus far suggest that the MVA heuristic performs well even for non-monotonic price trajectories. We now present additional numerical studies to demonstrate its robustness to different degrees and patterns of demand uncertainty during the selling horizon. In the earlier examples, the multiplicative random noise random variables were i.i.d. and with a high coefficient of variation ($\sigma/\mu = 0.3$). In this subsection, we first vary the magnitude of the coefficient of variation, still holding it constant across periods, and then consider non-i.i.d. multiplicative random noise terms.

Before doing so, we first present baseline results. All problems in this section share the following parameters: $N = 1$, $c = 25$, $h_w = 0.25$, $h_s = 0.5$, $D_t = Ap_t^{-b}\xi_t$, where $A = 100000, b = -1.5,\ \xi_t \sim \mathcal{N}(\mu, \sigma_t^2), \mu = 1$ and $\alpha = 0.5, 0.6, \ldots, 1.5$. Note that σ_t is not fixed and the pertinent value(s) will be specified for each problem instance. For each stocking level α, we set the order quantity, C, to α times the sum of expected demands in profitable periods. We present the prices and demand parameters for the three price trajectories in Table 2.11. In the remainder of the paper, we refer to the price trajectories for Examples 5, 6 and 7 as price trajectories 5, 6 and 7, respectively. For the baseline results, we use $\sigma_t = 0.2$ for all t with μ_t held constant at 1.0. Results are shown in Figure 2.4. Results obtained when varying the coefficient of variation appear in Figures 2.10 and 2.11 in Appendix B.

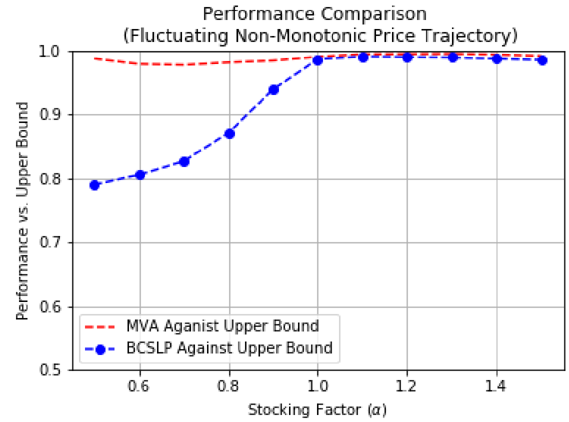
Table 2.11 Prices and Demand Parameters for Examples 5, 6, and 7

	Example 5		Example 6		Example 7	
t	Price (p_t)	Expected Demand (μ_t^D)	Price (p_t)	Expected Demand (μ_t^D)	Price (p_t)	Expected Demand (μ_t^D)
1	50	282.8	40	395.3	35	483.0
2	40	395.3	60	215.2	50	282.8
3	50	282.8	60	215.2	50	282.8
4	40	395.3	40	395.3	40	395.3
5	50	282.8	50	282.8	50	282.8
6	40	395.3	60	215.2	50	282.8
7	25	800.0	25	800.0	25	800.0

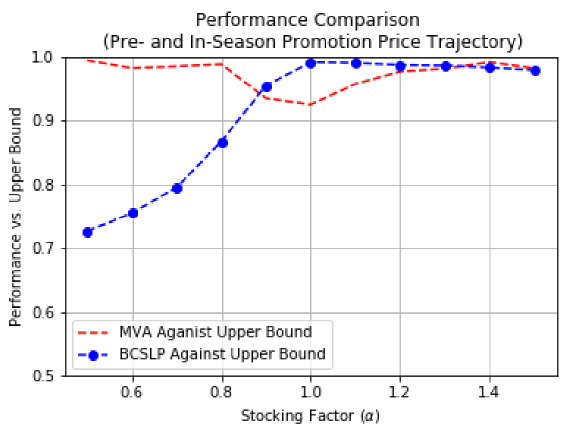
The results for Examples 5 and 6 are consistent with those of the first four examples: for low to intermediate levels of the stocking factor, MVA's performance is far superior to that of BCSLP; for intermediate to high levels of stocking factor, MVA's performance is still marginally better but BCSLP also performs quite well, as the retailer can simply set high service levels to maximize sell-through during the selling season. Example 7, however shows that BCSLP may outperform MVA when the retailer starts the season with intermediate levels of the stocking factor. In this particular case, the weaker performance of MVA can be attributed to the following reason. With the price in the first period being lower than in any other period but the markdown period, the MVA heuristic generates a solution in which the order-up-to level in the first period is only 458, which is less than the mean demand. This turns out to be too low considering that any leftovers incur low holding costs and could be sold at a much higher price in the subsequent period. The approximation of the expected marginal value of inventory in the MVA heuristic works well in most settings, but doesn't explicitly account for the value of time. It is intuitive that inventory allocations made earlier in the horizon have a better chance of being sold and being sold at a high price than in the final-period markdown price or a price of zero if unsold, but the MVA heuristic does not account for the number of periods remaining in the horizon except near the end of the horizon. Further research is needed to modify MVA to account for the number of periods remaining.



(a) Example 5



(b) Example 6

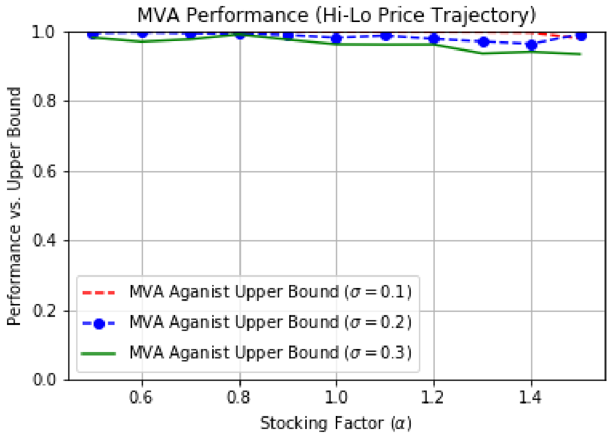


(c) Example 7

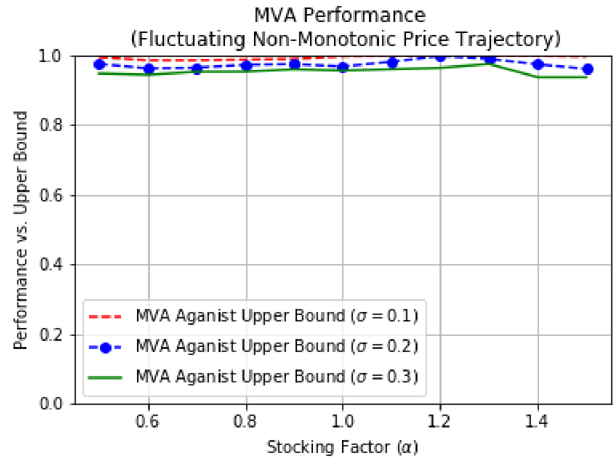
Figure 2.4 Gross Margin Performance of MVA and Best Constant Service Level Policy versus an Upper Bound for $\sigma = 0.2$

2.4.2.1 Impact of Coefficient of Variation

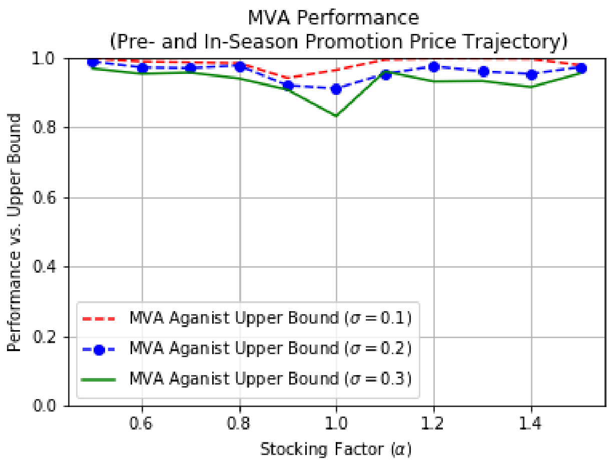
We now investigate how the degree of demand uncertainty affects the performance of MVA versus the other benchmarks. In apparel retailing, demand uncertainty is known to be high (Agrawal and Smith (1996)). For this reason, we focus on numerical examples with this characteristic, and study the performance of the MVA heuristic for a range of levels of the coefficient of variation, where $\frac{\sigma}{\mu} = 0.1, 0.2, 0.3$, with the coefficient of variation held constant across all periods. Figures 2.5(a), 2.5(b), and 2.5(c), show plots of the performance of MVA for different levels of coefficient of variation of demand for Examples 5, 6 and 7, respectively. The results clearly indicate that the performance of MVA declines steadily but slowly as the coefficient of variation increases. This is not surprising, as the problem becomes more challenging as demand uncertainty grows. More importantly, our results suggests that MVA's performance is fairly robust to varying magnitudes of demand uncertainty.



(a) Example 5



(b) Example 6

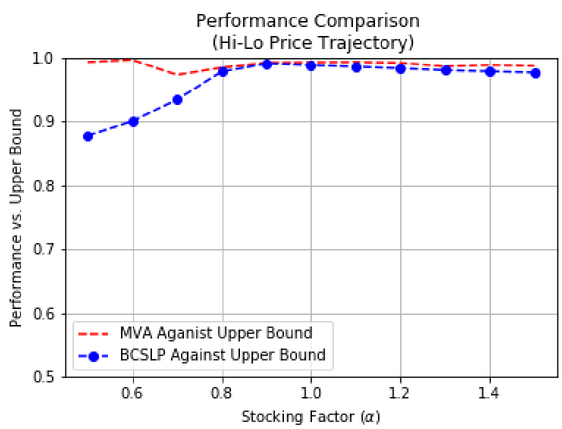


(c) Example 7

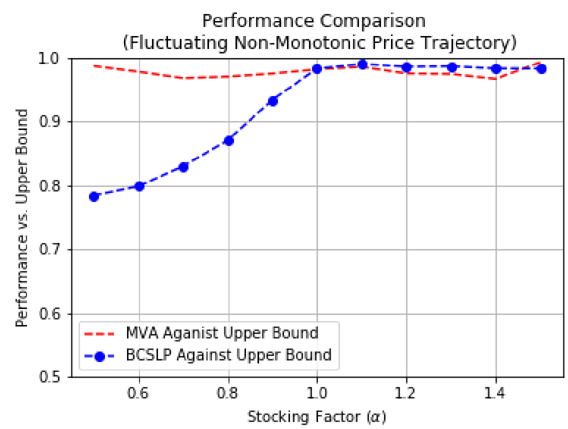
Figure 2.5 Gross Margin Performance of MVA versus an Upper Bound for Different Levels of the Coefficient of Variation

2.4.2.2 Impact of Non-I.I.D. Multiplicative Random Noise Random Variables

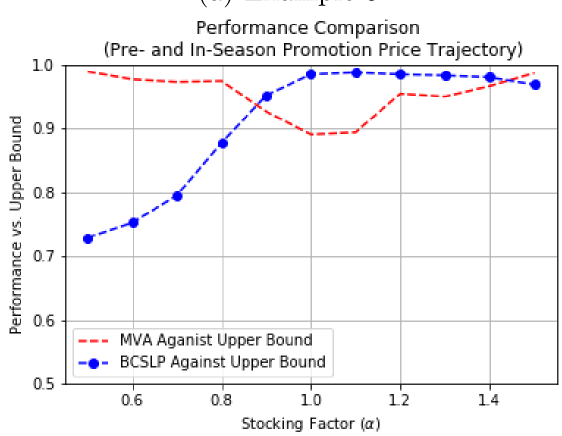
In practice, the magnitude of demand uncertainty is often time-varying. For example, demand uncertainty for a new product is typically larger immediately after launch but gradually decreases during the selling season. In this subsection, we analyze MVA's performance in the presence of non-i.i.d. random noise. We consider patterns with decreasing demand uncertainty and fluctuating demand uncertainty during the selling season. For the former case, the standard deviation of the random noise term declines linearly with time, specifically, $\{\sigma_t\}_{t=1}^{7} = \{0.3, 0.267, 0.233, 0.2, 0.167, 0.133, 0.1\}$. For the latter case, the standard deviation of the random noise term in full-price periods is 0.3, while the standard deviation in promotional periods is 0.1. Under the assumption of lost sales, the latter case is quite challenging because periods with a high price (and thus low mean demand) and high demand variance precede periods with a low price (and thus high mean demand) and low demand variance. In such a case, it is usually desirable to choose an order-up-to level much higher than the mean in the high-price periods, but observed demand may be very low, thus leaving large amounts of leftover inventory available for sale in low-price periods. The results for the case of decreasing demand uncertainty are presented in Figure 2.6, and those for the case of fluctuating demand uncertainty are presented in Figure 2.7.



(a) Example 5

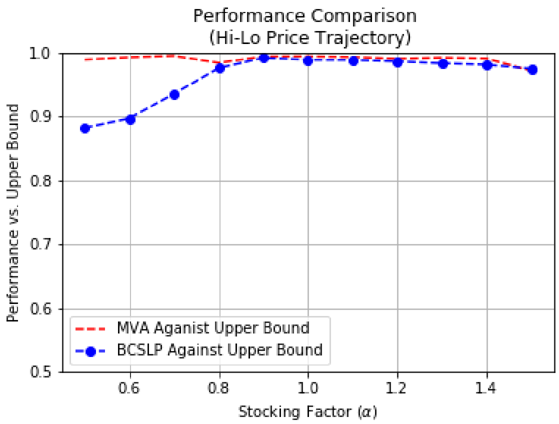


(b) Example 6

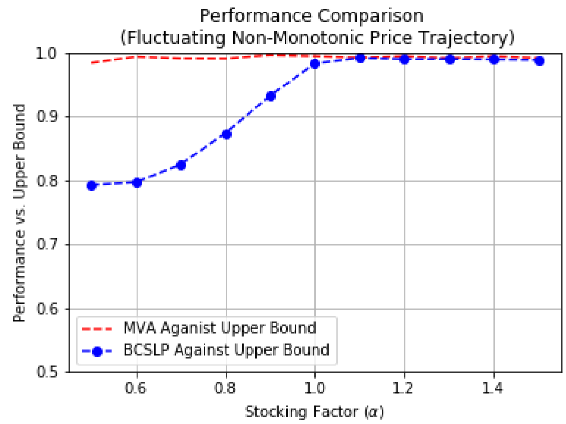


(c) Example 7

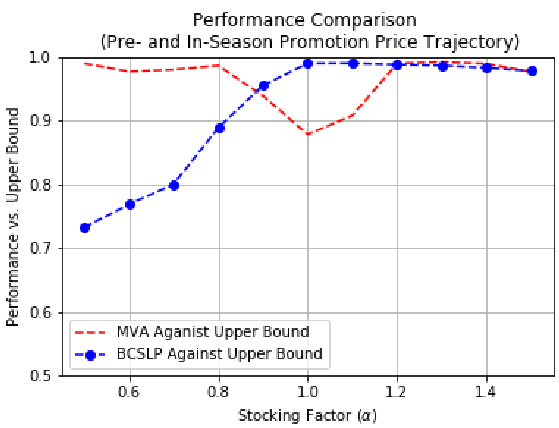
Figure 2.6 Gross Margin Performance of MVA and Best Constant Service Level Policy versus an Upper Bound with Decreasing Demand Uncertainty During the Selling Season



(a) Example 5



(b) Example 6



(c) Example 7

Figure 2.7 Gross Margin Performance of MVA and Best Constant Service Level Policy versus an Upper Bound with Non-Stationary Demand Uncertainty During the Selling Season

MVA's performance is similar to that under the base case with constant demand uncertainty. For Examples 5 and 6, neither pattern of demand uncertainty leads to a major deterioration in MVA's performance. We do observe a notable decline in performance for Example 7 at intermediate stocking factor levels. Interestingly, this pattern does not arise when the price in period 1 is set to 40. Therefore, the dip in the performance curve for MVA for Example 7 can be attributed to the low price in the first period. The deterioration in MVA's performance is due to the same reasons as we explained earlier: too conservative an allocation in period 1. Nevertheless, we observed a performance gap no worse than 12% when compared to that of the Robinson upper bound. These results suggest that MVA's performance is robust to time-varying demand uncertainty, but care should be taken to avoid making too conservative an allocation to the store in early, low-price periods, even when the demand uncertainty in those periods is high. An adjustment to the MVA algorithm for situations of this type may be warranted.

2.5 Single-Warehouse Multiple-Store Inventory Allocation Problem

In this subsection, we discuss the inventory allocation problem for the single-warehouse multi-store setting and extend the MVA heuristic to handle multiple stores. The main idea underlying the algorithm remains unchanged and is based on the principle that available inventory should be allocated such that the incremental benefit from a marginal unit of inventory is the same across all stores and periods; otherwise, a reallocation would increase expected gross margin.

We consider a firm that procures and sells a single good over T periods at stores indexed by $n = 1, \dots, N$. Our first task is to extend our approximation of the marginal value of inventory for any given set of order-up-to trajectories, $S_t^n, t = 1, \dots T$ for $n = 1, \dots, N$. Let the marginal expected value of the S_t^n-th unit of inventory in period t and store n be $\Delta_t^n(S_t^n)$. Let $\tilde{p}_t^n = p_t^n - h_w(t-1)$ represent the effective price from a unit sold at store n during period t, which is the selling price at store n during period t less the cost of holding the unit in the warehouse until period t; given a fixed initial inventory and therefore sunk variable purchasing costs, this is the maximum gross margin one can obtain from selling a unit in store n during period t.

The formula for $\Delta_t^n(S_t^n)$ must now account for all of the possible revenue and cost outcomes—at all stores in all periods—associated with the S_t^n-th unit of inventory for store n in the t-th period. We present the formulation and then explain each term.

$$\begin{aligned}\Delta_t^n(S_t^n) := \tilde{p}_t^n Pr(D_t^n \geq S_t^n) - h_d Pr(D_t^n < S_t^n) + \Delta_{t+1}^n(S_{t+1}^n + 1))Pr(D_t^n < S_t^n - S_{t+1}^n) \\ + \max_{i,j}\{\Delta_i^j(S_i^j + 1) \mid i > t,\ j \in \{1, \dots, N\}\} Pr(S_t^n - S_{t+1}^n \leq D_t^n < S_t) \quad (2.5)\end{aligned}$$

The first term is the expected marginal gross margin from the incremental unit, accounting for cases in which the observed demand exceeds the current S_t^n. The second term is the expected holding cost incurred due to holding the incremental unit at the store rather than at the warehouse, accounting for situations in which observed demand is less than S_t^n (so the incremental unit would not be sold). The third term accounts for situations in which $S_t^n > S_{t+1}^n$ and observed demand is less than $S_t^n - S_{t+1}^n$. In such cases, leftover inventory in period t would exceed the order-up-to level in period $t+1$ and the unit of inventory would be expected to generate *at most* $\Delta_{t+1}^n(S_{t+1}^n + 1)$. (We explained earlier why this is an upper bound.) The third term is the probability of such an event occurring multiplied by the incremental value $\Delta_{t+1}(S_{t+1} + 1)$. The last term accounts for situations in which the incremental unit in period t is left over (i.e., $D_t^n < S_t^n$) but is simply carried forward to the next period without being a unit in excess of S_{t+1}^n. Such a unit of inventory can be utilized indirectly for the best future opportunity at all stores in all future time periods because it substitutes for a unit that would otherwise have to be sent from the warehouse to the store;

this explains the *max* expression. The last term in the approximation is, in fact, the main difference between single-store and multiple-store settings, i.e., in the latter, the retailer can capture the incremental value associated with the best future period at the best store. In the remainder of this subsection, we examine the performance of MVA in the multiple-store setting.

2.5.1 Numerical Study for the Two-Store Setting

Similarly to the numerical study presented in Section 2.4, we compare MVA's performance against an upper bound which is obtained by aggregating the locations and relaxing the no backroom assumption, which allows us to reduce the original problem to an equivalent revenue management problem studied by Robinson (1995) with a modification to account for the two-store setting, and to the best constant service level policy. For the Robinson upper bound analysis, we incorporate another relaxation that allows the retailer to make allocations to stores in decreasing order of their offered prices in each period, with ties broken arbitrarily. Because it may be difficult for a retailer to implement the dynamic version of MVA in a multi-store setting, we also report on the performance of the modified static MVA to examine its viability as an inventory allocation policy when the dynamic version cannot be implemented.

For simplicity, we focus our numerical study on problems involving two stores. We use the following four types of price trajectories in our study: (i) identical in-phase Hi-Lo pricing, (ii) out-of-phase Hi-Lo pricing, (iii) Hi-Lo pricing at one store and constant pricing at the other, and (iv) non-identical in-phase Hi-Lo pricing. The numerical examples differ with respect to the price trajectories but share the following parameters: $T = 7$, $h_w = 0.25$, $h_s = 0.50$ and $c = 25$. Also, as before, C is equal to α times the total expected demand in profitable periods for $\alpha \in \{0.5, 0.6, \ldots, 1.5\}$. For each store, $D_t^i = A(p_t^i)^{-b}\xi_t$, where $A = 100000, b = -1.5,\ \xi_t \sim \mathcal{N}(\mu, \sigma^2),\ \forall t \in \{1, 2, 3, 4, 5, 6, 7\}$, where $\mu = 1$ and $\sigma = 0.3$. The prices and demand parameters for the examples are presented in Tables 2.12 through 2.18.

1. *Example 8: In-Phase Hi-Lo Pricing with Identical Price Trajectories*

Table 2.12 Prices and Demand Parameters for Stores 1 and 2 (Example 8)

		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.9
2	40	395.3	118.0
3	50	282.8	84.9
4	40	395.3	118.0
5	50	282.8	84.9
6	40	395.3	118.0
7	25	800.0	240.0

2. *Example 9: Out-of-Phase Hi-Lo Pricing (Hi Price and Low Price are the Same at Both Stores*

Table 2.13 Prices and Demand Parameters for Store 1 (Example 9)

		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	40	395.3	118.0
2	50	282.8	84.9
3	40	395.3	118.0
4	50	282.8	84.9
5	40	395.3	118.0
6	50	282.8	84.9
7	25	800.0	240.0

Table 2.14 Prices and Demand Parameters for Store 2 (Example 9)

		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.9
2	40	395.3	118.0
3	50	282.8	84.9
4	40	395.3	118.0
5	50	282.8	84.9
6	40.	395.3	118.0
7	25	800.0	240.0

3. *Example 10: Hi-Lo vs Constant Pricing*

Table 2.15 Prices and Demand Parameters for Store 1 (Example 10)

		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.9
2	40	395.3	118.0
3	50	282.8	84.9
4	40	395.3	118.0
5	50	282.8	84.9
6	40	395.3	118.0
7	25	800.0	240.0

Table 2.16 Prices and Demand Parameters for Store 2 (Example 10)

		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	45	331.3	99.4
2	45	331.1	99.4
3	45	331.1	99.4
4	45	331.1	99.4
5	45.	331.1	99.4
6	45	331.1	99.4
7	25	800.0	240.0

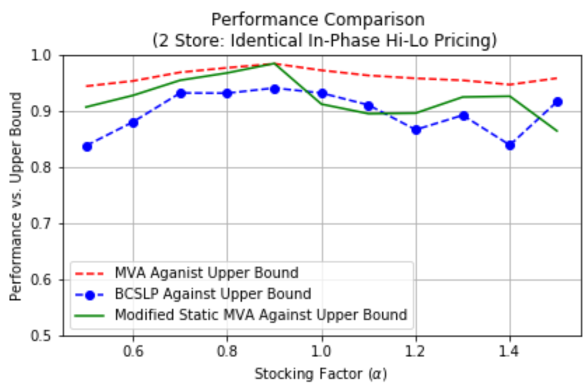
4. *Example 11: In-Phase Hi-Lo Pricing with Non-Identical Price Trajectories*

Table 2.17 Prices and Demand Parameters for Store 1 (Example 11)

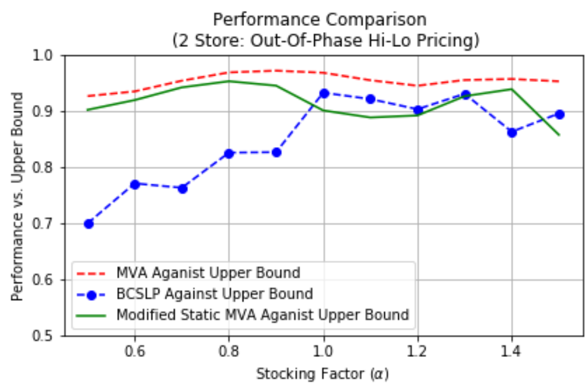
		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	60	215.2	64.5
2	30	608.6	182.6
3	60	215.2	64.5
4	30	608.6	182.6
5	60	215.2	64.5
6	30	608.6	182.6
7	25	800.0	240.0

Table 2.18 Prices and Demand Parameters for Store 2 (Example 11)

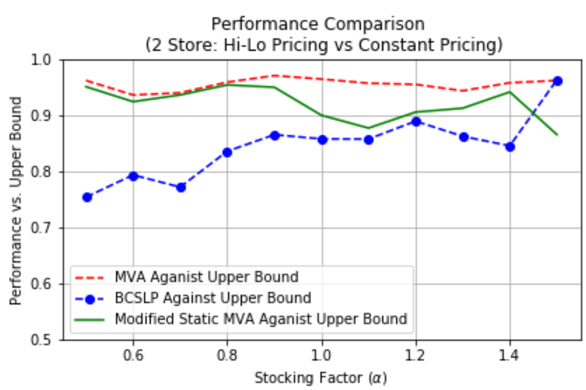
		Demand Statistics	
Period (t)	Price (p_t)	Expected Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	50	282.8	84.9
2	40	395.3	118.0
3	50	282.8	84.9
4	40	395.3	118.0
5	50	282.8	84.9
6	40	395.3	118.0
7	25	800.0	240.0



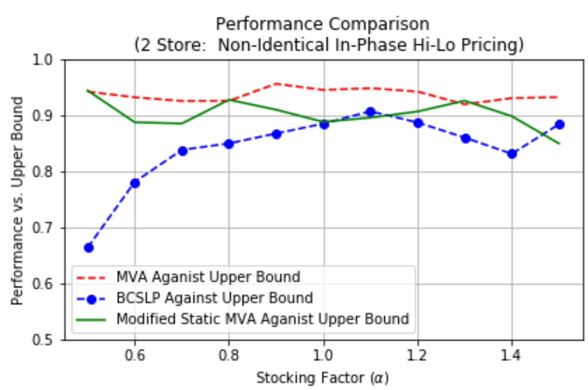
(a) Example 8



(b) Example 9



(c) Example 10



(d) Example 11

Figure 2.8 Performance Comparison: (a) In-Phase Hi-Lo Pricing with Identical Price Trajectories, (b) Out-Of-Phase Hi-Lo Pricing, (c) Hi-Lo vs Constant Pricing, (d) Hi-Lo Pricing with Non-Identical Price Trajectories

Performance results for the MVA heuristic in the two-store problems are shown in Figure 2.8. In the figure, each subfigure provides a graphical comparison of the performance of the dynamic MVA (dashed red line) and modified static MVA (solid green line) heuristics, and BCSLP (dashed blue line with dots as markers). As in the single-store problems, the dynamic MVA heuristic performs well, producing expected gross margins close to the respective Robinson upper bounds. Not surprisingly, the modified static MVA heuristic does not perform as well as the dynamic version, but still significantly outperforms BCSLP a great majority of the time.

As Figure 2.8(d) shows, the differences among the heuristic solutions are most prominent when the initial inventory level is low and/or the price differences across stores and periods are large. For these numerical examples, if the retailer sets the same service level at both stores even if they offer different prices, he would order up to a larger quantity when the price is 30 than when the price is 60 (due to the fact that both the mean and variance of demand are high when the price is 30). This would increase the risk of stocking out well before the end of the selling season, which would cause the retailer to forgo the opportunity to sell the inventory at a higher price in a later period and/or different location. Interestingly, the modified static version of MVA performs nearly as well as the dynamic MVA for stocking factors of 0.9 or smaller. The order-up-to levels for the simple static MVA solution can be computed off-line in advance, and the additional calculation for the modified static MVA are simple additions and subtractions, and thus, could be implemented without difficulty.

To summarize, the numerical examples collectively show that even the modified static MVA outperforms BCSLP when the stocking factor is low to intermediate, and the dynamic MVA performs quite well over the full range of stocking factor levels. BCSLP works well for monotonically decreasing price trajectories because, in such settings, a "push" strategy is optimal, which can be implemented via a high constant service level. However, it does not perform well for non-monotonic trajectories.

2.6 Procurement Problem

We now address the retailer's pre-season procurement problem under the assumption that the MVA heuristic will be used for inventory allocation. The main challenge in choosing the procurement quantity is that the recourse actions are complex—spanning multiple periods and, where appropriate, across multiple locations. Due to the two-echelon supply chain in our problem context, and the fact that inventory is allocated to the store(s) over multiple time periods contingent on demand outcomes, writing an exact formulation for the retailer's procurement problem is essentially impossible. For this reason, we develop a heuristic for a generalization of the newsvendor problem that captures most of the key high-level tradeoffs in our model. Our heuristic is in the spirit of the ideas in Cachon and Kök (2007), but we explicitly account for the known price in the markdown periods(s) and zero salvage value for unsold inventory, whereas their aim is to find salvage values that lead to the optimal order quantity. We present a model for the case of a single retail store but an extension to consider multiple stores is conceptually straightforward.

One salient feature of real-world multi-period newsvendor-like problems is that there is no single salvage value. The product may be sold at a low price during the last (or last few) period(s) in the selling season, but any leftovers at the end of the selling season must be salvaged at a much lower price, possibly zero. The number of units sold at either of these salvage values depends on the order quantity and the demand distribution. We present a modified newsvendor model that involves an approximation in which we collapse the supply chain to a single location and partition the selling season into two stages: the regular season and the final markdown period (we use period T alone here, but the final markdown period could span multiple time periods, in which case the season would be partitioned differently).

We assume, as an approximation, that the inventory is sold at a fixed constant price, $\hat{p}$, during the regular season and at a price of p_T during the final markdown period, where $\hat{p}$ can be approximated by the demand-weighted average of prices in the profitable periods during the regular season. (For ease of exposition, we assume that the product is profitable during all periods during the regular season.) The demand random variable for the first stage is approximated as $D_{Reg} = \sum_{t=1}^{T-1} D_t$ and the demand random variable for the final markdown period is represented as D_T, where we have suppressed the dependence on the fixed prices. D_{Reg} and D_T are assumed to be independent and have cumulative distributions $F_{Reg}(\cdot)$ and $F_T(\cdot)$ respectively, which we assume are continuous and twice differentiable. (Accordingly, we denote the probability density functions of D_{Reg} and D_T as $f_{Reg}(\cdot)$ and $f_T(\cdot)$, respectively.) The observed values of D_{Reg} and D_T are denoted by d_{reg} and d_T, respectively. Three types of demand outcomes may occur during the selling season. First, demand during the regular season may exceed the initial order quantity. i.e., $d_{Reg} \geq x_0^0$. Second, the regular season demand may be less than the initial order quantity, but the total demand during the regular season and the final markdown period may exceed the initial order quantity, i.e., $d_{Reg} < x_0^0$ and $d_{Reg} + d_T \geq x_0^0$. Third, the total demand during the regular season and the final markdown period may be less than the initial order quantity and if so, demand during the regular season would necessarily be less than the initial order quantity, in which case we would have $d_{Reg} < x_0^0$ and $d_{Reg} + d_T < x_0^0$. We account for these three types of outcomes in our formulation:

$$\max_{x_0^0 \geq 0} \text{profit} = \hat{p}\int_{x_0^0}^{\infty} x_0^0 f_{Reg}(x)dx + \int_0^{x_0^0}\int_0^{\infty}(\hat{p}x + p_T\min(y, x_0^0 - x))f_{Reg}(x)f_T(y)dydx - cx_0^0 \tag{2.6}$$

$$= \hat{p}\int_{x_0^0}^{\infty} x_0^0 f_{Reg}(x)dx + \int_0^{x_0^0}\int_0^{x_0^0-x}(\hat{p}x + p_T y)f_{Reg}(x)f_T(y)dydx$$
$$+ \int_0^{x_0^0}\int_{x_0^0-x}^{\infty}(\hat{p}x + p_T(x_0^0 - x))f_{Reg}(x)f_T(y)dydx - cx_0^0 \tag{2.7}$$

In (2.7), the first term in the profit function is the expected revenue, accounting for situations in which demand during the regular season exceeds initial inventory. The second term is the expected revenue, accounting for situations in which demand during the regular season does not exceed initial inventory, nor does total demand during the regular season and markdown period. The third term is the expected revenue, accounting for situation in which demand during the regular season does not exceed initial inventory, but total demand during the regular season and the markdown period do. The fourth term is the variable purchase cost. Note that we have not accounted for any inventory holding costs in this model. The profit function is well behaved as we next show.

Theorem 1:
The profit function of the modified newsvendor model is unimodal in the initial order quantity, x_0^0, and a unique stationary point, x^*, of the profit function exists.

Proof: Let us consider the derivative of the left hand side of (2.7):

1. Derivative of the first term:

$$\begin{aligned}
&\frac{d[\hat{p}\int_{x_0^0}^{\infty} x_0^0 f_{Reg}(x)dx]}{dx_o^0} \\
&= -\hat{p}x_0^0 f_{Reg}(x_0^0) + \hat{p}\int_{x_0^0}^{\infty} f_{Reg}(x)dx \\
&= -\hat{p}x_0^0 f_{Reg}(x_0^0) + \hat{p}[1 - F_{Reg}(x_o^0)] - (i)
\end{aligned}$$

2. Derivative of the second term:

$$\frac{d[\int_0^{x_0^0}\int_0^{x_0^0-x}(\hat{p}x + p_T y)f_{Reg}(x)f_T(y)dydx]}{dx_o^0}$$
$$= \int_0^{x_0^0}(\hat{p}x + p_T(x_0^0 - x))f_{Reg}(x)f_T(x_0^0 - x)dx - (ii)$$

3. Derivative of the third term:

$$\begin{aligned}
&\frac{d[\int_0^{x_0^0}\int_{x_0^0-x}^{\infty}(\hat{p}x+p_T(x_0^0-x))f_{Reg}(x)f_T(y)dydx]}{dx_o^0}\\
=\ &\int_0^{\infty}\hat{p}x_0^0 f_{Reg}(x_0^0)f_T(y)dy-\int_0^{x_0^0}(\hat{p}x+p_T(x_0^0-x))f_{Reg}(x)f_T(x_0^0-x)dx\\
&+\int_0^{x_0^0}\int_{x_0^0-x}^{\infty}p_T f_{Reg}(x)f_T(y)dydx\\
=\ &\hat{p}x_0^0 f_{Reg}(x_0^0)-\int_0^{x_0^0}(\hat{p}x+p_T(x_0^0-x))f_{Reg}(x)f_T(x_0^0-x)dx\\
&+p_T\int_0^{x_0^0}[1-F_T(x_0^0-x)]f_{Reg}(x)dx-(iii)
\end{aligned}$$

4. Consolidating terms, the derivative of the profit function is:

$$\hat{p}[1-F_{Reg}(x_0^0)]+p_T\int_0^{x_0^0}[1-F_T(x_0^0-x)]f_{Reg}(x)dx-c$$

First, we show the existence of a stationary point, x^*, satisfying:

$$\hat{p}[1-F_{Reg}(x^*)]+p_T\int_0^{x^*}[1-F_T(x^*-x)]f_{Reg}(x)dx-c=0.$$

Let

$$\gamma(x_0^0):=\hat{p}[1-F_{Reg}(x_0^0)]+p_T\int_0^{x_0^0}[1-F_T(x_0^0-x)]f_{Reg}(x)dx-c,$$

and consider the following upper bound on $\gamma(x_0^0)$:

$$g(x_0^0):=\hat{p}[1-F_{Reg}(x_0^0)]+p_T\int_0^{x_0^0}f_{Reg}(x)dx-c.$$

Then, for $\hat{p}>p_T$:

$$g'(x_0^0)=-\hat{p}f_{Reg}(x_0^0)+p_T f_{Reg}(x_0^0)<0,$$

and

$$\gamma'(x_0^0)=-\hat{p}f_{Reg}(x_0^0)+p_T f_{Reg}(x_0^0)-p_T\int_0^{x_0^0}f_T(x_0^0-x)f_{Reg}(x)dx<0.$$

If $\hat{p} > p_T$, then g and γ are monotonically and strictly decreasing functions of x_0^0. If we further assume that there exist some M and $\delta > 0$ such that:

$$f_{Reg}(t) \geq \delta, \forall t \leq M,$$

and

$$M\delta > \frac{\hat{p} - c}{\hat{p} - p_T},$$

then,

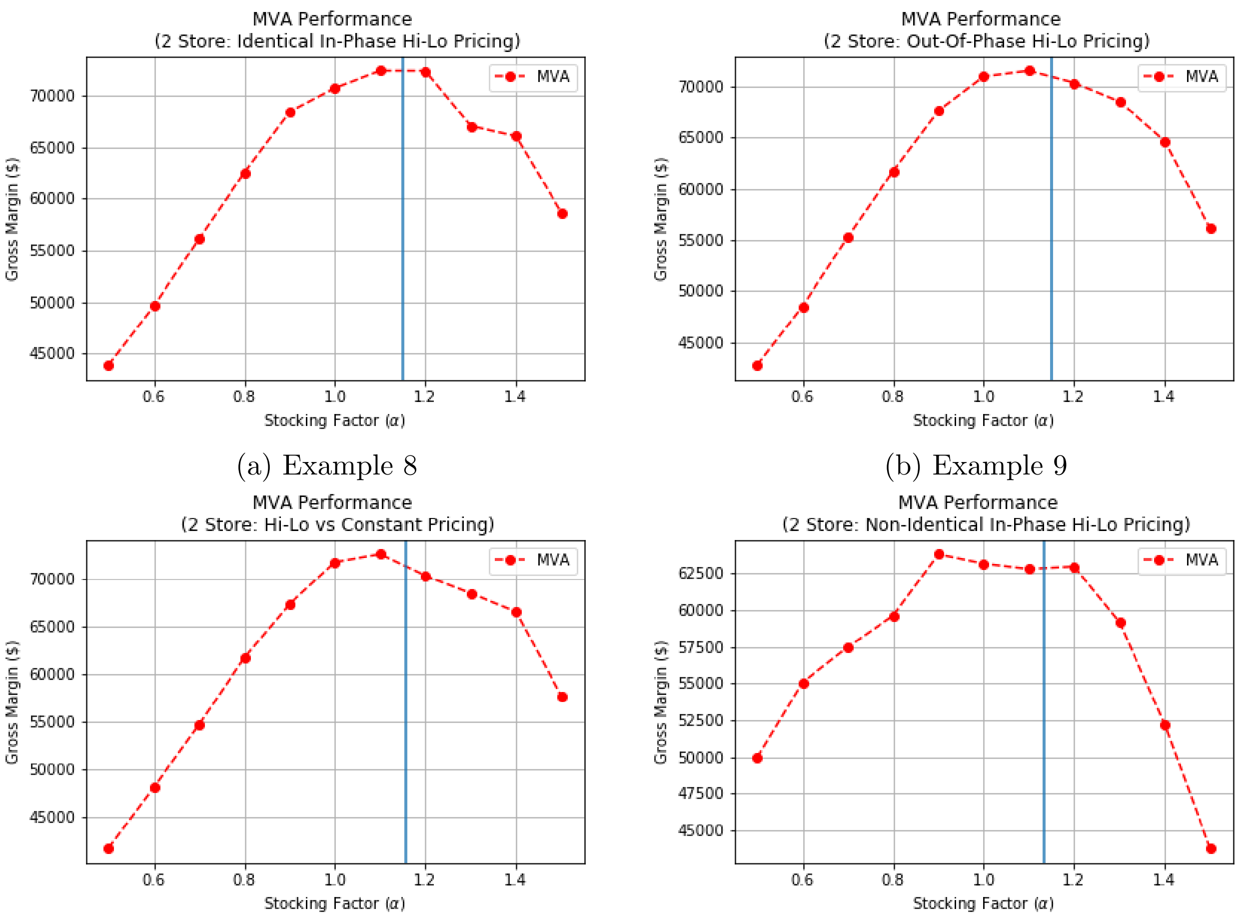
$$\begin{aligned} g(M) &= g(0) + \int_0^M g'(t)dt \\ &= \hat{p} - c - (\hat{p} - p_T)\int_0^M f_{Reg}(t)dt \\ &\leq \hat{p} - c - (\hat{p} - p_T)M\delta \\ &< 0. \end{aligned}$$

This implies $\gamma(M) \leq g(M) < 0$. By the Intermediate Value Theorem, there exists a stationary point, x^*, with $\gamma(x^*) = 0$. Furthermore, because γ is also monotonically decreasing in x_0^0, the stationary point of the profit function is also unique. (Q.E.D)

Theorem 1 indicates that the stationary point of the profit function is the global maximizer of our approximate profit function. As a result, the retailer only needs to solve the first order necessary condition for x_0^0:

$$\hat{p}[1 - F_{Reg}(x_0^0)] + p_T \int_0^{x_0^0} [1 - F_T(x_0^0 - x)]f_{Reg}(x)dx - c = 0,$$

which can be accomplished via a line search. If we consider only the first and last term on the left hand side of the above expression, we would have the optimality condition of the classical newsvendor model. The second term is a "correction term" that accounts for the salvage value of inventory sold in the markdown period.



(a) Example 8 (b) Example 9

(c) Example 10 (d) Example 11

Figure 2.9 Gross Margin Performance of MVA Starting with a Range of Initial Inventory Levels. The dark vertical bar shows the solution from the Modified Newsvendor Model.

We now apply the modified newsvendor approach to Examples 8 through 11. To do so, we aggregate demands across the two stores in each period and then proceed in the same way as for the single-store setting. More specifically, we approximate the demand random variable for the first stage as $D_{Reg} = \sum_{t=1}^{T-1} \sum_{n=1}^{2} D_t^n$ and that for the final markdown period as $D_T = \sum_{n=1}^{2} D_T^n$, and set $\hat{p}$ to the demand-weighted average of prices in the profitable periods. The results are shown in Figure 2.9, where we utilize the dynamic MVA heuristic to determine the inventory allocation. For Examples 8 through 10, the modified newsvendor solutions are near optimal, not only in terms of order quantity, but also in terms of the expected profit. For all three examples, the best stocking factor obtained via grid search is 1.1, and the corresponding modified newsvendor solutions are 1.15, 1.15, and 1.16, respectively. The gross margin gaps are 3.7%, 3.0%, 1.3%, respectively. For Example 11, the modified newsvendor does not perform as well in identifying an order quantity of the right magnitude:

the best stocking factor identified via grid search is 0.9 and the modified newsvendor solution is 1.135; however, the resulting gross margin gap is only 1.0%. As noted earlier, Example 11 is challenging due to the combination of parameters. In view of this, the performance of the modified newsvendor heuristic is good, considering its simplicity.

2.7 Conclusions

We have considered the procurement and inventory allocation problems of a retailer that sells a product with a long production lead time and short product life cycle, and utilizes a traditional two-echelon distribution system in which goods are initially stored at the warehouse and then sent to stores over the course of the season. We consider settings in which, due to marketing and logistics considerations, the planned trajectory of prices is fixed in advance and may be non-monotonic. We develop a heuristic, which we call the Marginal Value Analysis (MVA) heuristic, for inventory allocation based on the principle that the marginal value of an incremental unit of inventory should be the equal at all retail locations and all time periods. Results from numerical studies demonstrate that MVA achieves good performance versus both an upper bound and an industry benchmark for a setting with a single warehouse supplying either a single store or multiple (possibly non-identical) stores. Our results suggest that the performance of MVA is robust, and does not suffer from the same performance deterioration observed for the industry benchmark as the number of stores increases or as price differences increase across stores and time periods.

We also present a modification of the standard newsvendor model to account for the two-tiered salvage values that exist in our motivating scenario: positive markdown prices at the end of the season and zero salvage values for leftovers at the end of the season. Numerical examples suggest that our approach provides solutions that are nearly as good as those obtained via a grid search, despite the simplicity of the model. In our comparison, we use the MVA heuristic to determine demand-contingent inventory allocations during the selling season, and as such, our profit calculations (we actually calculate gross margin) reflect what would occur in actual implementation. Although the MVA heuristic is conceptually simple, it is too complicated to be applied in practice. Further research is needed to identify computationally less expensive methods that perform well; the MVA heuristic could serve as a more realistic benchmark than the Robinson upper bound.

Appendices

Appendix A: MVA versus "Best" Order-Up-To Policy

In this section, we compare solutions obtained using MVA with the best solution that can be found via a grid search over a range of order-up-to levels for each period. We focus on three-period problems, which are the simplest problems with non-monotonic price trajectories and are possible to solve via a relatively fine grid search. To implement the grid search, rather than using order-up-to levels, we search over service levels from 60% to 97.5% in 2.5% increments for each period and calculate the expected gross margin for all service level combinations. We consider two three-period examples that we refer to as Examples A and B, respectively. Both examples share the following parameters: $N = 1$, $c = 25$, $h_w = 0.25$, $h_s = 0.5$, $D_t = Ap_t^{-b}\xi_t$, where $A = 10000$, $b = -1.5$, $\xi_t \sim \mathcal{N}(\mu, \sigma^2)$, $\mu = 1$, $\sigma = 0.2$ and $\alpha = \{0.7, 0.8, 0.9\}$. For each stocking level α, we set the order quantity, C, to α times the sum of expected demands in profitable periods. The prices and demand parameters for the two examples are presented in Table 2.19. The MVA and "best" solutions are presented in Table 2.20. For all numerical examples, the expected gross margin using the MVA heuristic is within 0.5% of the best solution from the grid search, and in four of the six cases, the gap is less than 0.1%. These examples provide some evidence that MVA performs well in an absolute sense.

Table 2.19 Prices and Demand Parameters for Examples A and B

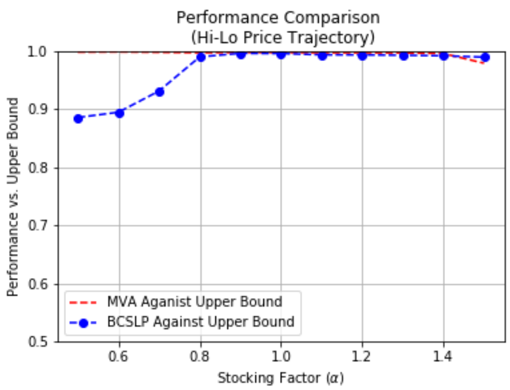
	Price Trajectory A			Price Trajectory B		
t	Price	Avg. Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)	Price	Avg. Demand (μ_t^D)	Std. Dev of Demand (σ_t^D)
1	40	39.5	7.9	50	28.5	5.7
2	60	21.5	4.3	60	21.5	4.3
3	25	80	16	25	80	16

Table 2.20 MVA and Best Solution from Grid Search for Examples A and B

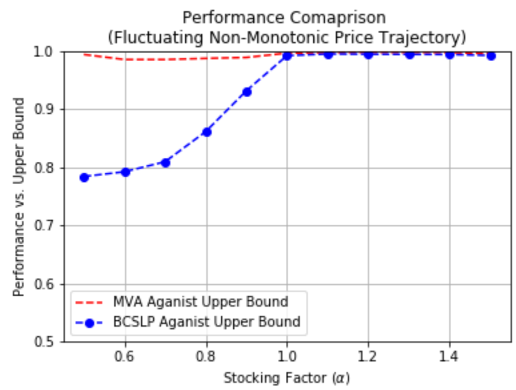
Example	α	C	MVA Solution		Best Solution		Profit Gap (%)
			s_1	s_2	s_1	s_2	
A	0.7	42	22	22	21.3	20.8	0.1%
	0.8	48	27	23	26.6	23.4	0.0%
	0.9	54	34	23	32.0	25.0	0.0%
B	0.7	34	17	18	17.2	17.1	0.1%
	0.8	39	22	18	20.2	19.6	0.5%
	0.9	44	27	20	24.0	22.9	0.3%

Appendix B: MVA Performance for Examples 5, 6, and 7

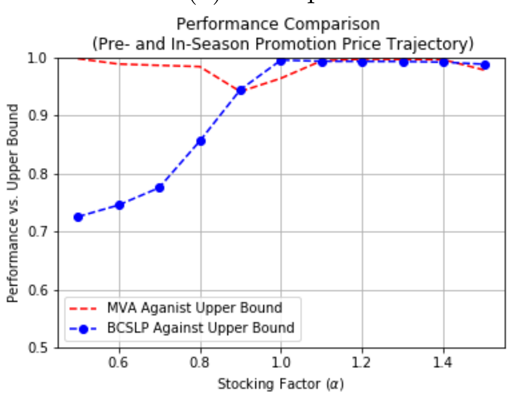
In this Appendix, we present results for Examples 5, 6, and 7 (discussed in Section 2.4.2) for $\frac{\sigma}{\mu} = 0.1$ and 0.3.



(a) Exampe 5

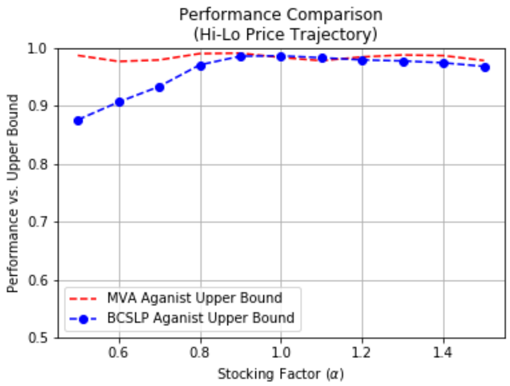


(b) Example 6

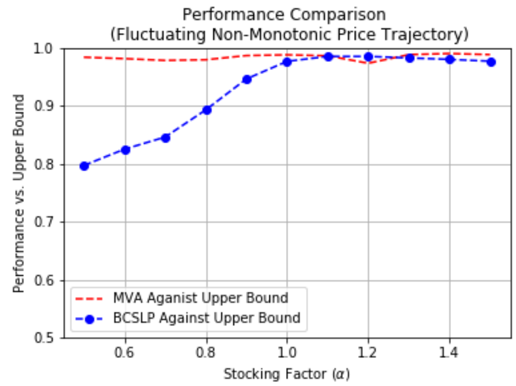


(c) Example 7

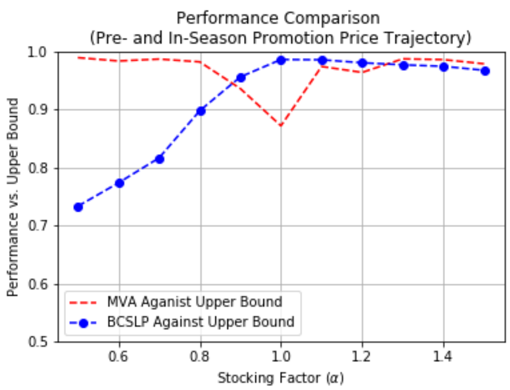
Figure 2.10 Gross Margin Performance of MVA and Best Constant Service Level Policy versus an Upper Bound for $\sigma = 0.1$



(a) Example 5



(b) Example 6



(c) Example 7

Figure 2.11 Gross Margin Performance of MVA and Best Constant Service Level Policy versus an Upper Bound for $\sigma = 0.3$

Chapter 3

Multi-Product Pricing Over Multiple Time Periods

3.1 Introduction

An important question that frequently arises in today's highly competitive retail sector is how to price multiple substitutable products over multiple time periods when customers' decisions in a given time period depend not only on the current prices of the products but also on the past trajectories of prices and anticipated future prices. We consider a category of consumable non-perishable products sold by a brick-and-mortar monopolist retailer whose customers visit its store(s) regularly and not only can observe the prices of all products in the category when making a purchase but also have knowledge of the pricing patterns of the retailer due to their prior shopping experience. A prototypical setting would be a grocer whose customers shop approximately weekly, and some prototypical products would be staple goods such as breakfast cereals, laundry detergent and non-seasonal condiments, as well as more discretionary items such as bottled beverages.

The growth of e-tailers and discount retailers has forced brick-and-mortar retailers to resort to offering discounts in order to lure customers, leading to the difficult problem of determining the best price trajectories when customers' buying behavior is a complex function of these trajectories. Empirical studies (e.g., Hinterhuber and Bertini 2011) have found that small changes in prices can increase or decrease profitability by as much as 50%. Ailawadi et al. (2007) found that more than 50% of promotions at the drug store chain CVS were not profitable, as incremental sales did not sufficiently compensate for lower profit margins. They estimate that CVS could improve annual profit by approximately $52.6 million by eliminating chain-wide promotions in the 15 worst-performing categories. One key reason why price promotions are not profitable is that they lead to cannibalization of demand for other substitute products such as different pack sizes of the same brand or different products within a broad category, or lead to forward buying. van Heerde et al. (2004) report on the effects of sales promotions using store-level data and conclude that expansion of category sales, cross-period substitution (forward buying) and cross-brand substitution each account for about one third of the overall sales bump.

We seek to develop a near-optimal approach for choosing prices for multiple products over multiple time periods, accounting for reference price effects (including deal-seeking behavior), customers who are forward-looking and make purchasing decisions strategically in view of anticipated future prices, and price-based substitution by customers who are not loyal to a specific product. Although researchers have incorporated some combinations of these factors into multi-period pricing models, the vast majority of this research has been done in single-item contexts. The literature on pricing multiple products over multiple periods is sparse and does not explicitly account for all of the factors mentioned above. Our work was motivated by conversations with category managers and pricing managers who are eager to develop multi-product, multi-period pricing strategies that account for realistic effects of prices on customer purchases.

The remainder of the paper is organized as follows. In Section 3.2 we provide a literature review. A problem statement and related problem formulations appear in Section 3.3. We present structural results and describe our solution methodology in Section 3.4. Section 3.5 contains numerical examples that illustrate the effects of separating the customer's purchasing and consumption decisions (which is an important distinctive feature of our approach). Finally, we present conclusions, discuss limitations of our approach and suggest directions for future research in Section 3.6.

3.2 Literature Review

In this section, we provide a review of the pertinent literature, focusing on settings involving non-seasonal products that are consumed on an ongoing basis and whose supply is unconstrained for practical purposes. This differs from traditional revenue management settings in which each customer is interested in buying one unit of a product (e.g., an airline ticket for a specific date and destination) and needs to decide when, but not how much, to purchase. We first provide a review of the literature on customer response to promotions, including both conceptual models of, and empirical studies on, reference price effects, as well as strategic customer behavior in response to discounts. We then discuss the literature on multi-period pricing to highlight existing work on single-product pricing with reference price effects, single-product pricing with strategic customer stockpiling, and finally, multi-period multi-product pricing.

3.2.1 Customer Response to Promotions

There is a large empirical marketing literature on customer response to price promotions. In this subsection, we discuss key empirical findings regarding how strategic customers respond to discounts and the effect of discounts on reference prices. We refer the interested reader to Neslin and van Heerde (2008) for a very comprehensive review of the literature on the intertemporal effects of price promotions, including associated phenomena such as stockpiling, accelerated purchases of infrequently-purchased products, impact on customers' reference prices and price sensitivity, and permanent effects.

3.2.1.1 Empirical Studies of Strategic Customer Behavior in Response to Discounts

When customers are strategic, their knowledge or forecasts of future prices affect how they time their purchases and how much they purchase at each buying epoch. Many researchers have reported empirical evidence that discounts lead customers to increase their purchase quantities, stockpiling a portion of them, and thus also increase the duration until the next purchase. As one example of purchase acceleration, Psendorfer (2002), in a study involving ketchup products, finds that customers purchase about seven times as much at low prices than they do at high prices, and presumably stockpile the excess for use during periods of high prices. As another example, Hendel and Nevo (2006b) study two years of household- and store-level demand data and observe post-promotion dips for many products, which suggests that customers are stockpiling during promotions. See Shoemaker (1979), Ward and Davis (1978), Blattberg et al. (1981), and Neslin et al. (1985) for other empirical studies on purchase acceleration.

Many empirical studies (e.g., Gupta 1988, Chiang 1991, Chintagunta 1993, and Bell et al. 1999) are concerned with understanding the effects of promotional discounts on primary (brand switching) and secondary demand effects (stockpiling and purchase acceleration), and generally conclude that the secondary effects are larger. There is also a substantial amount of research on how forward-looking customers form expectations of future prices. Gönül and Srinivasan (1996), Erdem et al. (2003), and Hendel and Nevo (2006a) develop dynamic structural models of customer price expectations and show that these expectations have a significant impact on customer purchases. Their results indicate that it is important to incorporate price expectations in demand models because of their effects on price elasticity. Our approach to the problem accounts for these price expectations.

3.2.1.2 Empirical Studies of Reference Price Effects

In the past four decades, behavioral economics, which is concerned with the effects of psychological, cognitive, emotional, cultural and social factors on the economic decisions of individuals and institutions, has attracted growing interest across multiple fields (Thaler, 2017). In this subsection, we discuss both conceptual models of, and empirical studies on, reference price effects. The concept of reference points was introduced by Kahneman and Tversky (1979) as an element of their proposed prospect theory, which is an alternative to expected utility theory (von Neumann and Morgenstern 1944). The authors observe that people make decisions based on perceived gains and losses relative to a reference point rather than on the basis of the final outcome alone. Monroe (1973) defines a reference point as a standard against which the outcomes are judged: worse outcomes are viewed as losses while better outcomes are viewed as gains (see also Kumar et al. 1998). In our context, the reference points are reference prices, which are usually defined as a weighted average of past prices, typically based on exponential smoothing, or the previous period's price (Winer 1986). However, researchers have also reported empirical evidence that reference prices may be shaped not only by internal references prices that are based on customers' memory of historical prices and but also by external reference prices, i.e., observed prices at the time of purchase (see Mazumdar et al. 2005 and the references therein.).

Researchers have studied how discounts affect reference prices and reservation prices. The distinction between a reservation price (the highest price that a customer is willing to pay) and a reference price has been hotly debated among academics. Garbarino and Slonim (2003) suggest that reference prices and reservation prices are different concepts and are formed based on different inputs and human perceptions. However, Blattberg et al. (1981) argue that, from a practical standpoint, retail discounts shape customers' reference and reservation prices in the same way. Some researchers (e.g., Briesch et al. 1997) have even used estimates of reference prices in choice models as if they were reservation prices.

There is a significant stream of literature on the impact of reference prices on customer choices. We discuss key results and refer the interested reader to Mazumdar et al. (2005) and Arslan and Kachani (2010) for surveys of research on reference prices. Winer (1986) develops a brand choice model for frequently-purchased products that accounts for both probability-of-purchase and reference price effects, and finds that his empirical model outperforms standard demand models that are based solely on price in predicting customer choices. Briesch et al. (1997) compare several reference price models, and conclude that memory-based (i.e., based on price history) internal reference price models outperform stimulus-based (i.e., based on information available at the point-of-purchase) external reference price models in modeling customer brand choice. Hardie et al. (1993) and Greenleaf (1995) also find that memory-based internal reference price models outperform external reference price models. Kumar et al. (1998) also study the impact of internal and external reference price and find that the latter has a stronger effect on product choice. Putler (1992) develops a brand choice model with reference price effects and empirically validates his model on weekly retail egg sales data. Considering an aggregate demand model and assuming that the initial price is the initial reference price, he finds that reference prices do significantly impact brand choice in an asymmetric fashion: in particular, demand fell in response to a price increase by two and a half times as much as it increased in response to a price decrease of equal magnitude.

3.2.2 Multi-Period Pricing

In this subsection, we describe both empirical and analytical research on multi-period pricing. We use the term multi-period pricing to indicate that the retailer's selected price may differ from one period to the next. In a portion of the literature, and especially the articles in which demand is assumed to be a deterministic function of the price trajectory (although the prices themselves may change from period to period), the prices are determined at the beginning of the horizon and are not modified, and as such, are static. A good portion of the so-called dynamic pricing literature fits into this category. In the literature on truly dynamic pricing, the dynamics are typically due to demand uncertainty, and prices are updated in response to demand observations. In our models, the demands are deterministic functions of the price trajectories of the products, so the prices can be determined at the beginning of the horizon. This is why we use the phrase "multi-period pricing" rather than dynamic pricing. In the next two subsections, we discuss research on normative models for multi-period pricing, first considering reference price effects and then strategic customer behavior.

3.2.2.1 Single-Product Pricing with Reference Price Effects

In this subsection, we discuss the literature on single-product pricing with reference price effects. Most of these articles address the pricing problem of a monopolist selling a single non-seasonal product to customers whose reference price is a function of past prices, in most cases with an exponential weighting. The authors model demand in the aggregate, not for each customer separately, considering one or multiple customer segments. If customers are heterogeneous, they may differ in the reference effects associated with gains (purchasing at a price below their reference price) or losses (purchasing at a price above their reference price).

Greenleaf (1995) presents analysis for two different settings in which the total expected demand in each period is a linear function of the price in that period as well as both the positive and negative deviations of the offered price from the customer's reference price. The customer's reference price is updated via exponential smoothing. The author first considers an infinite-horizon setting in which, in each period, the retailer can choose either the regular price or a pre-specified discount price; the other is a finite-horizon setting in which the retailer can choose any price within a range in each period. The retailer seeks to maximize total discounted profit. For the first setting, the author's focus is on deriving the effects of the references prices on the retailer's profit. For the second setting, the author presents a dynamic programming formulation that he utilizes for solving numerical examples based on real data for a national brand of peanut butter. Among other things, Greenleaf's numerical results suggest that a high-low pricing strategy is optimal, with the high price contributing to increasing the customers' reference price, and that promotions become more profitable as the sales response due to gain-seeking increases relative to that of loss aversion. He also observes that both regular and irregular cycles of promotion may be optimal.

Kopalle et al. (1996) generalize Greenleaf's model, first to consider the effects of different relative weights associated with gains and losses for a single customer segment, and then to consider a two-segment model in which one segment weights losses more heavily than gains, and the other has the opposite relative weighting. The retailer chooses one price (within a range) in each period with the goal or maximizing total discounted profit. Like Greenleaf, the authors formulate the problems as dynamic programs, from which they derive properties of the optimal solution. The authors find that the relative weighting of gains and losses plays a key role in the single-segment model. In particular, if gains have a higher weight than losses, the optimal pricing policy is cyclic, whereas when the weights are equal, the optimal price converges monotonically to a steady-state optimal price which depends on the initial reference price. If losses have a higher weight than gains, under certain conditions on the smoothing constant for reference price updating and/or the form of the demand function, the optimal price monotonically converges to a constant price. Kopalle et al. also generalize their model to consider multiple products offered by one or more retailers. They assume that the demand for each product is a function of the product's own price, the customer's reference price for that product, and the prices of all other products. The authors consider two objectives for the retailer, both maximizing discounted profit and maximizing the average profit per period. They derive sufficient conditions for constant prices to be suboptimal.

Fibich et al. (2003) address a continuous-time, infinite-horizon version of the single-product pricing problem in which the reference price is the (continuous) exponentially weighted average of past prices. The aggregate demand function is linearly decreasing in price. In the authors' first model, the (linear) effects of gains and losses are symmetric. They derive the optimal policy in closed form and then utilize the results for the symmetric case to derive a closed-form optimal policy for the asymmetric case, although the problem is a non-smooth optimal control problem. The authors show that there is an optimal steady-state price which is lower than that in the absence of reference price effects and the trajectory is initially monotonically increasing (decreasing) if the initial reference price is lower (higher) than the steady-state optimal price. They also show that the *form* of the optimal policy is the same for the asymmetric case as it is for the symmetric case. Furthermore, they show that for the finite horizon case, the optimal price is monotonically decreasing at the end of the horizon after an interval of constant prices. The authors illustrate the benefits of using the optimal policy by comparing it with several benchmarks, including the best constant pricing strategy. They also show how to generalize the model and results to incorporate demand functions that are nonlinear in price.

Chen et al. (2017) consider a generalization of the discrete-time, finite-horizon version of the problem studied by Fibich et al. (2003) that allows for time-varying demand parameters, and establish some structural properties of the optimal policy, including the discontinuity of the optimal value function and the maximum number of kink points in it. They also develop an approximation heuristic with polynomial time complexity. Via numerical examples, they illustrate the impact of seasonality of the demand parameters, which tends to cause the optimal prices to be cyclic, as well as loss aversion, which tends to dampen price fluctuations. They also present results of a numerical study which indicates that their heuristic performs well.

Popescu and Wu (2007) consider infinite-horizon, discrete-time settings under assumptions like those in Kopalle et al. (1996) and Fibich et al (2003), with the aim of incorporating more general demand models. Among other things, their demand models allow for the marginal effect of gains and losses relative to the reference price to decrease as the reference price increases. The authors first study a setting with loss-neutral customers and show that the steady-state optimal price is increasing in the reference price, and then derive conditions under which convergence of the optimal price is monotonic. They then show that the optimal price is cyclic if customers place more weight on gains than on losses. For the case of loss averse customers, they show that the optimal pricing policy depends on the initial reference price. For initial reference prices in an intermediate range, the optimal price is constant. For either higher or lower initial reference prices, each of the optimal price trajectories converges monotonically to an optimal price, but the optimal steady state prices differ for the three ranges of the initial reference price. Qualitatively, their results confirm that the conclusions of Kopalle et al. and Fibich et al. apply to a broader range of demand functions, including nonlinear representations, that depend upon reference prices.

Nasiry and Popescu (2011) consider a different memory-based reference price model based on the peak-end rule (see Fredrickson and Kahneman, 1993), which captures the notion that customers remember the peak (here, actually the trough thus far) and end (most recent) price; the reference price is a weighted average of the two. The authors study an infinite-horizon, discrete-time, single-product setting with loss averse customers. They show that the optimal policy has three regions, with thresholds on the trough at the beginning of the horizon (given as an initial condition) defining boundaries between regions. If the initial trough is in a (characterizable) intermediate range, the optimal steady-state price is equal to the initial trough. If the initial trough is low, the optimal price trajectory converges to a constant price that depends on the initial trough, while if the initial trough is high, the optimal steady-state price is equal to the price at the upper limit of the intermediate range. From numerical examples, they infer the intuitive result that the value of manipulating prices when customers' reference prices are based on the peak-end rule is less than that under exponential smoothing, and consequently, the optimal prices fluctuate less under the peak-end rule.

The articles discussed above establish the optimality of constant or converging prices under most realistic settings in which customers are homogeneous and loss averse. Two recent papers have explored a variant of the problem in which customers are heterogeneous, and demonstrate the optimality of non-stationary pricing strategies. Wang (2016) addresses a setting involving different customer segments who arrive in heterogeneous sets of time periods and each customer segment's reference price is updated only when the segment visits the store and observes the price. He shows that even when customers are loss-averse, the optimal price trajectory is cyclic and enables the retailer to price discriminate among segments. Chen and Nasiry (2018) show that even in a simpler setting than that in Wang (2016), namely one with customer segments that differ in their sensitivity to losses (and may also be gain-seeking), a cyclic pricing policy may be optimal. The authors show that previous findings that support optimality of constant pricing depend heavily on the assumption of homogeneous customer behavior vis-a-vis reference price effects. They show that even if a retailer cannot price-discriminate among multiple customer segments, a periodic markdown pricing policy allows the retailer to reduce demand from a customer segment with lower willingness to pay, boost demand from gain-seeking customers, and improve profits.

The articles discussed above are based on models involving loss aversion alone or linear reference price effects, but empirical research (e.g. Mazumdar et al. 2005, Krishnamurthi et al. 1992, Bell and Lattin 2000) suggests that loss aversion is not a universal phenomenon. Building on these findings, Hu et al. (2016) study a single-product pricing problem of a firm facing gain-seeking customers. They demonstrate that even if the retailer sets prices myopically (maximizing profit in the current period), complicated reference price dynamics may transpire. Unable to derive the optimal pricing strategy under more general assumptions, the authors present numerical results showing that a high-low pricing strategy is suboptimal, and that both the optimal price and reference price trajectories are so complex that there is a connection to one-dimensional discontinuous maps. The authors show that in the special case in which customers remember only the most recent price and demand is insensitive to a surcharge above the reference price, a cyclic skimming pricing strategy is optimal. They also provide conditions that guarantee the optimality of high-low pricing strategies.

To summarize, existing research has demonstrated the optimality of constant or converging prices (although not always monotonically) under many realistic conditions when customers are homogeneous and loss averse, and have established the optimality of non-stationary pricing strategies when customers are gain-seeking and in other settings involving heterogeneous customers. The work of Chen et al. (2017) provides a methodology that performs well for the case of homogeneous customers, incorporating the effects of both loss aversion and gain-seeking as well as non-stationary demands. However, none of the single-product models explicitly separates the customer's purchase and consumption decisions, and therefore they all rely on the implicit assumption that customers immediately consume the purchased product, which is not always true in practice. Furthermore, none of models in these articles on pricing in view of reference price effects can be generalized to handle multiple substitutable products, as the underlying demands in the models are functions of only the product's own price and its reference price and not of the prices of substitute products. Finally, Chen and Nasiry (2018) point to the need to incorporate forward-looking customer behavior into pricing models with reference effects. None of the articles discussed in this section combines a memory-based reference price effect and strategic forward-looking customer behavior, but customers, in reality, may be both strategic and influenced by reference price effects, as well. For example, a customer may stockpile, quite rationally, to take advantage of a price discount, but may purchase even more than economic tradeoffs would dictate due to utility accrued from purchasing at a price below her reference price (as would be applicable for gain-seeking customers).

3.2.2.2 Single-Product, Multi-Period Pricing with Strategic Stockpiling by Customers

Recent research has begun to incorporate the effects of strategic customer behavior on aggregate demand and the optimal pricing strategy. In this subsection, we discuss the literature on a retailer's single-product multi-period pricing problem when facing strategic customers who may stockpile.

The problem of how to optimally price in settings in which customers make purchases to (try to) satisfy their known, constant demand and may stockpile was first studied by Blattberg et al. (1981) and subsequently by Jeuland and Narasimhan (1985). In both papers, the authors assume that the firm has a regular price but offers periodic promotions to two customer segments that have different inventory holding cost rates. In Blattberg et al. (1981), customers seek to minimize their long-run average cost, which consists of purchase and inventory holding costs, to satisfy their known demands. The retailer seeks to maximize long-run average cost, which consists of a fixed cost per order from the supplier, loss of revenue due to discounting vis-a-vis the undiscounted price, and inventory holding costs. The authors assume that a discount is offered each time an order is placed with the supplier, which is shown to be optimal in Eppen and Liebermann (1984). They derive the retailer's optimal policy analytically and propose two potential explanations for manufacturer-incentivized promotions, as we next describe. The information acquisition explanation is based on the premise that customers seek information via trials and retailers offer discounts to reduce the cost of a trial. The inventory reduction explanation suggests that retailers offer discounts to reduce inventory costs. Examining four product categories in a panel dataset, the authors find strong alignment between the inventory reduction explanation and the retailer's pricing trajectories, finding retailers often promote high market share brands more frequently than low market share brands in all four product categories, which, in turn, leads to customer stockpiling.

Jeuland and Narasimhan (1985) consider a setting with two customer segments that have different price-sensitive demand rates. The segment with the lower demand rate also has a lower inventory holding cost rate, and will purchase to satisfy their desired consumption rate only if the purchase price plus the cost of holding inventory until consumption is below their (implied) willingness to pay. The authors develop an analytical model to study what happens if the retailer offers a fixed discount every T periods. The authors show that because the segment with the higher demand rate has a higher inventory holding cost rate, the retailer can utilize periodic discounting as a mechanism to profitably price discriminate if the segment with the higher inventory holding cost does not stockpile when the discount is offered and the other segment stockpiles, purchasing more than they would in the absence of a discount.

Su (2010) studies a discrete time, infinite-horizon pricing problem in a setting with n customer segments that differ in their valuation of the product, inventory holding cost rate, maximum consumption rate and fixed cost for making a purchase. In each period, the retailer sets a price for the product, and each customer chooses the number of units to purchase. Customers seek to maximize long-run average utility, which is equal to valuation from consumption less fixed and variable procurement costs and the cost of holding inventory. (To avoid unnecessary inventory holding costs, customers always consume at their maximum consumption rate if they have sufficient inventory, or their available inventory otherwise.) The retailer seeks to maximize long-run average revenue. The author derives the rational expectations equilibrium and shows that the retailer should use recurring promotions when frequent shoppers have higher willingness to pay than occasional shoppers. He also presents analytical results that provide insights into conditions under which high-low pricing or an everyday-low-price strategy is optimal, and the effect of various parameters on the optimal frequency and depth of promotions.

In addition to the articles that discussed above, our work is also related to the research of Besbes and Lobel (2015), who consider a pricing problem under price commitment when there is strategic customer stockpiling. The authors consider a monopolist facing a continuous-time, finite-horizon, single-product pricing problem. The customers are assumed to be infinitesimally small and to be present during the entire horizon. The authors first study the pricing problem under the unit-demand setting before extending their results to settings in which customers may purchase more than one unit at time. The customer population is assumed to be heterogeneous with respect to their valuation of the product, waiting cost, and storage capacity. At the start of the selling horizon, the retailer has a fixed amount of inventory which has zero value at the end of the selling horizon. The retailer optimizes both pricing and rationing policies, and commits to them at the beginning of the horizon, seeking to maximize revenue. During the selling horizon, customers decide whether or not to purchase the product and whether or not to leave the market based on the announced pricing and rationing policies: staying in the market results in waiting cost while leaving the market yields zero payoff. The customers are strategic in timing their purchases and maximize their utility by postponing purchases until a low price period. For the unit-demand setting, using a dynamic programming approach, the authors prove that the optimal pricing policy is cyclical, with a cycle length that is at most twice as long as the maximum duration the customer population is willing to wait, and typically consists of moderate discounts throughout each cycle, a larger discount at the midpoint of the cycle, and the largest discount at the end of the cycle. The authors also show that monotonic price trajectories are suboptimal. The authors extend their model to allow customers to purchase more than one unit at a time and to stockpile. They assume that the customer population is present throughout the horizon. Each individual has unit demand in each period and customers are heterogeneous with respect to their storage capacities. The authors show that the optimal pricing policy is also cyclic and the seller should first sell to impatient customers with high-valuations before selling to more patient customers later in each cycle.

This substream of the multi-period pricing literature elucidates how different facets of strategic customer behavior affect the optimal pricing policy. Our work, similarly to the articles discussed in this subsection, incorporates strategic customer behavior, but also simultaneously considers the effect of reference prices.

3.2.3 Multi-Period, Multi-Product Pricing

In the final subsection of our literature review, we discuss the literature on optimizing multi-period multi-product pricing for frequently-purchased goods. There is a large literature on empirical models involving many different aspects of consumer response to prices in multi-product, multi-period contexts. We refer interested readers to Chintagunta (1993), Kim et al. (1995), Foekens et al. (1998), Cooper et al. (1999), Anderson and Vilcassim (2001), Chintagunta (2002), van Heerde et al. (2004), Baltas (2005), Liu et al. (2009), and references therein, and the articles on brand switching discussed in Section 3.2.1.1 for an overview of the subject. There is also a significant literature on multi-period multi-product dynamic pricing in revenue management contexts, i.e., where customers wish to purchase at most one unit within the product category before a deadline (e.g., a seat in one of the fare classes on an airline flight) and need to decide when to buy in view of the seller's anticipated price trajectory. See Gallego and van Ryzin (1997), Bertsimas and De Boer (2005), and Maglaras and Meissner (2006) for representative work on the topic, and McGill and van Ryzin (1999) for an overview.

There are very few articles on optimizing multi-period multi-product pricing decisions in which there is an interaction among the products, and we are aware of only one paper that pertains to frequently-purchased goods. Cohen et al. (2017a) study the multi-product promotion optimization problem (POP) for fast-moving consumer goods. The authors assume that the retailer commits to the price trajectories at the beginning of the horizon and carries sufficient inventory to meet all demand for each item during the planning horizon. They also assume that demand can be expressed as a deterministic, time-dependent, nonlinear function of prices, but limit the demand function to depend on only self-past and current prices and on cross-current prices. Demand may exhibit seasonality. The retailer aims to maximize the total profit from all items during the selling horizon. The authors first model the multi-product POP problem as a nonlinear integer program that is computationally intractable. An extension of their methodology for the single-item POP problem (Cohen et al., 2017b) does not perform well for the multi-item POP problem, so they propose a more general class of integer programming approximation methods. The authors find that App(2), in which the objective function is approximated by the sum of marginal effects on the promoted item and cross-item effects of promotions, provides near-optimal performance in most of the instances in their numerical study. We refer the interested reader to Cohen and Perakis (2018) and the references therein for an overview of issues that arise in multi-item POP decision contexts.

To the best of our knowledge, our work is the first to study the multi-period, multi-product pricing problem with both strategic customer behavior and reference price effects. We first develop a unified approach in which customer purchasing and consumption decisions are explicitly modeled, which enables us to capture implicit stockpiling, and derive the customer's optimal purchase and consumption policy as a function of the price trajectories (expressed analytically). The customer's response can then be incorporated into the retailer's profit maximization problem to determine optimal price trajectories. Later in the paper, we show that our approach can be generalized to multiple customer segments.

3.3 Problem Statement

We address a multi-period pricing problem faced by a retailer who sells multiple substitutable products in the same category to a class of homogeneous customers. Although we treat a finite-horizon problem, our solution approach allows us to solve problems with horizons that are long enough to keep the end-of-horizon effects small. End-of-horizon effects can be further mitigated via appropriate modeling constructs that we describe later.

Both the retailer and customers are strategic agents and anticipate the other party's optimal choices. We assume that the retailer commits to the price trajectories that it selects for the various products and that customers are able to learn the retailer's pricing patterns over time. In practice, retailers often need to commit to prices over some time horizon due to the need to arrange for advertising and for purchases of products to meet demands that fluctuate due to the changing prices. For frequently-purchased goods, customers do learn about retailer's pricing patterns and they are further aided by web sites such as camelcamel.com that provide extensive price histories. Early in our research on this topic, we analyzed a publicly available data set that includes weekly prices and demands at a grocery chain for a variety of product categories. We observed large demand spikes for some products when discounts were offered. We then attempted to fit parameters (coefficients) for existing models of demand for non-seasonal products that incorporate reference price effects, including loss aversion and gain-seeking behavior. We found that it was insufficient to include reference price effects to generate the magnitude of demand spikes that we observed in the data. Rather, it was necessary to explicitly separate customers' purchasing and consumption decisions, and thus also implicit decisions regarding stockpiling, to induce the observed demand spikes. This is the main motivation behind our approach to the problem.

We formulate the problem as a two-stage game, with the retailer as the Stackelberg leader, setting the price of each product in each period, and customers as followers, making purchase and consumption decisions in each period. The retailer maximizes profit, which is equal to revenue less variable purchase costs. For ease of exposition, we assume variable purchasing costs for any given product are constant over the horizon, although it is straightforward to relax this assumption. We assume, for ease of exposition, that one unit of any product offers the same functional value, e.g., same number of servings of a food product. Customers obtain utility and disutility from various sources: (i) utility from consumption of the various products, which we assume is linear in the total consumption, but may vary from one product to another; (ii) disutility from deviation of consumption from a (constant) aggregate target for the product category in each period (if constant consumption is desired for the category), which we assume is quadratic, reflecting increasing disutility for greater deviations as is commonly observed in practice; (iii) utility from purchasing products below the respective reference prices but with decreasing marginal returns, reflecting that even deal-seekers obtain decreasing marginal pleasure from each additional unit purchased at a bargain; (iv) disutility for expenditures for purchases; and (v) disutility from incurring inventory holding costs; (vi) disutility from shortages due to insufficient purchases. We represent the combination of the last two factors in the form of a convex penalty term associated with each product in each time period that increases as either inventory or shortages increase. The customer's total expenditures on products within the category during the horizon are limited by a budget constraint.

We assume that the customer's reference price for each product is modeled as an exponentially weighted sum of past prices for the product itself as well as those of potentially all other products in the category. Exponential weighting is the most commonly used approach in the literature; our approach can be generalized to other weighting schemes. Our inclusion of prices for potentially all products in the category reflects the fact that a customer's willingness to pay for a product may depend, in part, on the prices of available substitutes. We also assume that the time horizon under consideration is short (weeks over the course of several months, not over years), so for ease of exposition, we do not incorporate discounting, but it is straightforward to do so. For simplicity, we also assume that all products have the same unit size and that customers may purchase fractional quantities. To avoid aberrations due to end-of-horizon effects, we use a standard technique employed in the inventory control literature, which is to associate a value with each unit of inventory remaining at the end of the horizon. If this value accurately reflects its true marginal value in subsequent periods, the decisions within the horizon under consideration will not be distorted by the finiteness of the horizon.

We next present notation used in the customer's and retailer's optimization problems.

n: number of customers;

t: time period index, $t = 1, \ldots, T$;

i: product index, $i = 1, \ldots, N$;

p_{it}: price of product i in period t (constant in the customer's problem and decision variable in the retailer's problem);

v_i: unit purchase cost of product i for the retailer;

α_i: per unit utility from consumption of one unit of product i;

ν: coefficient for disutility associated with non-smooth consumption;

β: coefficient for utility associated with reference-effect-related gains and losses;

τ: coefficient for disutility that captures the effect of diminishing marginal returns of purchases at a single epoch (for the customer);

h: per period coefficient on the inventory and shortage penalty function;

μ: coefficient for disutility of each dollar of expenditure;

$\bar{c}$: consumption target for the product category in each period;

x_{it}: purchase quantity of item i in period t;

c_{it}: consumption of item i in period t;

C_{it}: cumulative consumption of item i as of the beginning of period t;

m_t: cumulative expenditure as of the beginning of period t;

y_{it}: customer's inventory of product i at the beginning of period t (with $y_{i0} = 0$ for all i);

r_{it}: reference price for product i at the beginning of period t (with r_{i0} given for all i)

γ_{ij}: coefficient representing the effect of product i's reference price in the prior period on product j's reference price in the current period;

ψ_{ij}: coefficient representing the effect of product i's price in the prior period on product j's reference price in the current period;

Each customer's reference price is modeled via exponential smoothing and evolves as follows:

$$r_{i,t+1} = \sum_j \gamma_{ji} r_{it} + \psi_{ji} p_{it} \quad \forall i, t,$$

where $\sum_{j=1}^{n} (\gamma_{ij} + \psi_{ij}) = 1, \forall i$. As such, the reference price in the next period is a weighted average of the current period's reference price and prices of the substitutable products in the category.

We can now express the customer's optimization problem as:

$$
\begin{aligned}
\text{(P1)}\quad & \max_{x_{it},c_{it}\forall i,t} \textstyle\sum_{t=1}^{T}\sum_{i=1}^{N}\alpha_i c_{it} - \nu\sum_{t=1}^{T}(\sum_{i=1}^{N} c_{it} - \bar{c})^2 \\
& +[\beta \textstyle\sum_{t=1}^{T}\sum_{i=1}^{N}(R_{it}-p_{it})x_{it} - \tau\sum_{t=1}^{T}\sum_{i=1}^{N} x_{it}^2] \\
& -\textstyle\sum_{i=1}^{N} h_i \sum_{t=1}^{T} y_{it}^2 - \eta\sum_i y_{i,T+1}^2 - \mu m_{T+1} \\
\text{s.t.}\quad & r_{i,t+1} = \textstyle\sum_j \gamma_{ji} r_{jt} + \sum_j \psi_{ji} p_{jt}, \quad \forall i,t \\
& m_{t+1} = m_t + \textstyle\sum_i p_{it} x_{it} \quad \forall t \\
& m_{T+1} \le \bar{M} \\
& C_{i,t+1} = C_{it} + c_{it} \quad \forall i,t \\
& y_{i,t+1} = y_{it} + x_{it} - c_{it} \quad \forall i,t \\
& m_1 = 0; \quad y_{i1} = 0 \quad \forall i; \quad C_{i1} = 0 \quad \forall i \\
& x_{it} \ge 0, c_{it} \ge 0, y_{it} \ge 0 \quad \forall i,t
\end{aligned}
$$

The terms in the objective represent (i) utility from direct consumption; (ii) disutility from non-smooth consumption; (iii) purchase utility from "gains" ("losses") associated with quantities purchased at below (above) the reference price, which exhibits decreasing marginal returns; (iv) inventory holding and shortage costs during the horizon; (v) penalties for deviations of inventory from zero at the end of the horizon; and (vi) disutility from expenditures for the products. We note that the terms associated with consumer inventory and shortage are not in the traditional linear form. We are using the expression $\sum_i h_i \sum_t y_{it}^2$ to approximate $\sum_i \sum_t \omega_i y_{it} + M\sum_t[(\sum_i y_{it})^-]^2$, where ω_i is the per unit per period holding cost for product i and M is a large number. The two terms in the foregoing expression are the usual expression for inventory holding costs and a term to penalize the square of the total shortages in each period. The expression that we actually utilize allows us to retain the quadratic form essential for a linear quadratic control formulation, but it has two important limitations. First, we should be imposing an infinite penalty on shortages (to play the role of a constraint) and a modest penalty on inventory, but the expression that we use forces us to use the same weight for both. Second, penalties are charged on the square of shortages of each product individually in each period rather than the sum of shortages in each period. This has the effect of inducing the customer to equalize shortages among the products, and thus each shortage quantity tends to remain small. This, in turn, necessitates the use of large values of h_i to ensure that inventory remains non-negative, but large values of h_i do not accurately reflect inventory holding costs. A common approach to model an asymmetric penalty using a quadratic penalty function is to center the function at a non-zero value, which we refer to as ξ. We later discuss challenges in choosing appropriate values of h_i and ξ as well as the limitations associated with a quadratic penalty function when the actual penalties for positive and negative deviations (inventory and shortages, respectively) are asymmetric.

Problem (P1), in the form stated above, is a nonlinear optimization problem with a concave objective and linear constraints. It is straightforward to reformulate it as a dynamic program with the same decision variables and with the state variables $y_{i,t}$, $C_{i,t}$, and m_t (for all i and t). Neither version of the problem is difficult to solve numerically. However, we would like to obtain solutions in explicit algebraic form as a function of the problem parameters to enable us to embed the algebraic form into the retailer's optimization problem. Fortunately, for any given set of price trajectories for the products, if we relax the non-negativity constraints on the x_{it}, c_{it}, and y_{it} variables (for all i and t), the problem is messy but otherwise straightforward to solve as a linear quadratic control problem. In practice, small violations of this constraint will not significantly distort the customer's purchasing patterns in response to the retailer's prices, and the deviations can be limited by the quadratic penalty in the third term if the values of h_i are large enough.

Let $x^*_{it}, i = 1, \ldots, N, t = 1, \ldots, T$ be the customer's optimal purchase decisions for (P1) given a set of price trajectories for the products, $p_{it}, i = 1, \ldots, N, t = 1, \ldots, T$, and let n be the number of customers who visit the retailer. The retailer's profit maximization problem can then be expressed as:

$$\text{(P2)} \qquad \max_{p_{it} \forall i,t} n \sum_{t=1}^{T} \sum_{i=1}^{N} (p_{it} - v_i) x^*_{it}(\mathbf{P})$$

where $\mathbf{P}$ is the matrix of prices.

3.4 Solution Methodology and Structural Results

We now proceed to present methodologies for solving the customer's utility maximization and retailer's profit maximization problem. At the end of this section, we discuss the extension to multiple customer segments.

3.4.1 Customer's Utility Maximization Problem

Before presenting our solution methodology for the customer's problem, we first reformulate (P1) as a finite-horizon linear-quadratic (LQ) control problem without non-negativity constraints on consumption and purchase quantities, and on on-hand inventory levels. Problems of this form are well studied in the literature on optimal control (see Bertsekas (1995)) and yield solutions in closed form in the absence of non-negativity constraints. Provided that the objective function is well behaved and $\boldsymbol{x_t}$, $\boldsymbol{c_t}$, and $\boldsymbol{y_t}$ do not become too negative, we can drop these non-negativity constraints because small negative quantities in the *customer's* solution will not have a meaningful effect on the retailer's decisions.

Once the non-negativity constraints have been dropped, for any specific retailer price trajectory, the customer's optimal purchase and consumption policy can be derived as the solution of generalized Riccati equations. To express the customer's purchase and consumption decisions *analytically* as a function of the retailer's prices, we incorporate two additional terms pertaining to the customer's aggregate consumption into her objective function, specifically $\lambda(\mathbf{1})'[\boldsymbol{\mathcal{C}}_{T+1}]-\pi\|\boldsymbol{\mathcal{C}}_{T+1}\|_2^2$, where λ is the coefficient of utility for each unit of consumption in the product category and π is the coefficient associated with a term that captures the diminishing marginal returns for aggregate consumption over the horizon. The addition of the two terms allows us not only to model the customer's desire for higher aggregate consumption at a diminishing rate, but also to ensure that the objective function is sufficiently well behaved that the problem can be solved tractably as a LQ-control problem. Finally, the budget constraint also may be safely removed as the parameter μ can always be calibrated so that the budget constraint is satisfied.

We can now express the simplified formulation as follows:

$$
\begin{aligned}
\text{(P1B)}\quad & \max_{x_{it},c_{it}\forall i,t} \textstyle\sum_{t=1}^{T}\sum_{i=1}^{N}\alpha_i c_{it} - \nu\sum_{t=1}^{T}(\sum_{i=1}^{N}c_{it}-\bar{c})^2 \\
& +[\beta\textstyle\sum_{t=1}^{T}\sum_{i=1}^{N}(R_{it}-p_{it})x_{it}) - \tau\sum_{t=1}^{T}\sum_{i=1}^{N}x_{it}^2] \\
& -\textstyle\sum_{i=1}^{N}h_i\sum_{t=1}^{T}(y_{it}-\xi)^2 - \eta\sum_i y_{i,T+1}^2 - \mu m_{T+1} \\
& +\lambda(\mathbf{1})'\boldsymbol{\mathcal{C}}_{T+1} - \pi\|\boldsymbol{\mathcal{C}}_{T+1}\|_2^2 \\
\text{s.t.}\quad & r_{i,t+1} = \textstyle\sum_j \gamma_{ji}r_{jt} + \sum_j \psi_{ji}p_{jt}, \quad \forall i,t \\
& m_{t+1} = m_t + \textstyle\sum_i p_{it}x_{it} \quad \forall t \\
& C_{i,t+1} = C_{it} + c_{it} \quad \forall i,t \\
& y_{i,t+1} = y_{it} + x_{it} - c_{it} \quad \forall i,t \\
& m_1 = 0; \quad y_{i1} = 0 \quad \forall i; \quad C_{i1} = 0 \quad \forall i
\end{aligned}
$$

It is straightforward to rewrite (P1B) as a dynamic program (DP) in which the periods are the stages. The DP value function at each stage is jointly concave in the pertinent decision variables. As a result, solving the first order optimality conditions allows us to derive the full solution in the terminal stage, and by backward induction, we can derive an analytical solution for the customer's utility maximization problem and the utility-to-go function at each stage. This result is stated in Theorem 2.

Theorem 2
The customer's optimal purchase and consumption policies can be expressed as linear functions of the elements in the retailer's price trajectories, customer inventory levels, and the customer's cumulative aggregate consumption level.

A proof of Theorem 2 appears in Appendix C. There, we present a proof under the simplifying assumption that $h_i = h$ for all i but the result generalizes to the case of unequal h_i values.

The key result is that the optimal purchase and consumption decisions in each period t can be expressed as:

$$\begin{bmatrix} \mathbf{x}_t \\ \mathbf{c}_t \end{bmatrix} = K_t \begin{bmatrix} \mathbf{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix} + G_t \begin{bmatrix} \mathbf{p}_T \\ \mathbf{r}_T \\ \mathbf{p}_{T-1} \\ \mathbf{r}_{T-1} \\ \dots \\ \mathbf{p}_{t+1} \\ \mathbf{r}_{t+1} \\ \mathbf{p}_t \\ \mathbf{r}_t \end{bmatrix} + F_t,$$

where K_t, G_t, and F_t are parameter matrices (see Section 3.4.2 for detailed mathematical expressions) that only depend on exogenous model parameters and are independent of prices, reference prices, the customer's inventory level and the cumulative aggregate consumption. The purchase and consumption decisions are linear functions of the reference price, current prices, and future prices. The customer's purchasing decisions reflect his (approximately) optimal response to the complex tradeoffs among the factors that we have discussed above. The terms in the parameter matrices are complicated but are easy to compute.

3.4.2 Algorithm Implementation

In this subsection, we provide a summary of the steps required to compute the customer's optimal purchase and consumption policy as a function of the customer's state variables and the retailer's price trajectories under the simplifying assumption that $h_i = h$ for all i. The optimal purchasing-consumption policy takes the form:

$$\begin{bmatrix} \mathbf{x}_t \\ \mathbf{c}_t \end{bmatrix} = K_t \begin{bmatrix} \mathbf{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix} + G_t \begin{bmatrix} \mathbf{p}_T \\ \mathbf{r}_T \\ \mathbf{p}_{T-1} \\ \mathbf{r}_{T-1} \\ \dots \\ \mathbf{p}_{t+1} \\ \mathbf{r}_{t+1} \\ \mathbf{p}_t \\ \mathbf{r}_t \end{bmatrix} + F_t,$$

where parameter matrices K_t, G_t and F_t are calculated via backward recursion, which we explain later in this section.

The boundary conditions at stage T are:

$$
\begin{aligned}
K_T &= \begin{bmatrix} K_T^{11} & K_T^{12} \\ K_T^{21} & K_T^{22} \end{bmatrix} \\
&= \begin{bmatrix} -\Phi^{-1}[\nu\eta(\mathbf{1}\mathbf{1}') + \pi\eta I_b] & -\Phi^{-1}\eta\pi \\ \frac{\tau+\eta}{\eta}K_T^{11} + I_b & \frac{\tau+\eta}{\eta}K_T^{12} \end{bmatrix}, \\
G_T &= \begin{bmatrix} G_T^{11} & G_T^{12} \\ G_T^{21} & G_T^{22} \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\mu+\beta}{2}\Phi^{-1} & \frac{\beta}{2}\Phi^{-1} \\ \frac{1}{2\eta}(\mu+\beta)I_b - \frac{\tau+\eta}{\eta}\frac{\mu+\beta}{2}\Phi^{-1} & \frac{\beta}{2}\frac{\tau+\eta}{\eta}\Phi^{-1} - \frac{\beta}{2\eta}I_b \end{bmatrix}, \\
F_T &= \begin{bmatrix} F_T^1 \\ F_T^2 \end{bmatrix} \\
&= \begin{bmatrix} \Phi^{-1}\{\frac{\eta}{2}\alpha + \eta(\nu\bar{c} + \frac{\lambda}{2})\mathbf{1}\} \\ \frac{\tau+\eta}{\eta}\Phi^{-1}\{\frac{\eta}{2}\alpha + \eta(\nu\bar{c} + \frac{\lambda}{2})\mathbf{1}\} \end{bmatrix},
\end{aligned}
$$

where $\Phi = (\eta\tau + \eta\pi + \tau\pi)I_b + \nu(\tau+\eta)(\mathbf{1})(\mathbf{1})'$.

In addition, we have the following boundary conditions for Q_T, Π_T and Λ_T:

$$
\begin{aligned}
Q_T &= \begin{bmatrix} Q_T^{11} & Q_T^{12} \\ Q_T^{21} & Q_T^{22} \end{bmatrix} \\
&= \tau \begin{pmatrix} (K_T^{11})' \\ (K_T^{12})' \end{pmatrix} \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} + (\nu) \begin{pmatrix} (K_T^{21})' \\ (K_T^{22})' \end{pmatrix} (\mathbf{1})(\mathbf{1})' \begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} \\
&\quad + h \begin{pmatrix} I_{b\times b} & 0 \\ 0 & 0 \end{pmatrix} + \eta \begin{pmatrix} (I_b + K_T^{11} - K_T^{21})' \\ (K_T^{12} - K_T^{22})' \end{pmatrix} \begin{pmatrix} I_b + K_T^{11} - K_T^{21} & K_T^{12} - K_T^{22} \end{pmatrix} \\
&\quad + \pi \begin{pmatrix} (K_T^{21})' \\ (K_T^{22})' \end{pmatrix} \begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} + \pi \begin{pmatrix} 0 & 0 \\ 2K_T^{21} & I_{b\times b} + 2K_T^{22} \end{pmatrix}, \\
\Pi_T &= \{-2\tau(G_T^{11}, G_T^{12}) + \beta[-I, I]\}' \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} \\
&\quad - \{2\nu[(G_T^{21}, G_T^{22})]'(\mathbf{1}\mathbf{1}')\} \begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} \\
&\quad - 2\eta\{[G_T^{11} - G_T^{21}, G_T^{12} - G_T^{22}]\}'[I_b + K_T^{11} - K_T^{21}, K_T^{12} - K_T^{22}] \\
&\quad - \mu[I, 0]' \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} - 2\pi\{[G_T^{21}, G_T^{22}]\}'[0, I_b], \\
\Lambda_T &= -2\{\tau F_T^1\}' \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} \\
&\quad - \{2\nu\{F_T^2\}'(\mathbf{1}\mathbf{1}') - \alpha' - 2\nu\bar{c}\mathbf{1}'\} \begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} + 2[h\xi', 0] \\
&\quad - 2\eta[F_T^1 - F_T^2]'[I_b + K_T^{11} - K_T^{21}, K_T^{12} - K_T^{22}] \\
&\quad + \lambda\mathbf{1}'[K_T^{21}, I_b + K_T^{21}] - 2\pi\{F_T^2\}'[0, I_b].
\end{aligned}
$$

We now present the general recursion at time period t, $\forall t \in \{T-1, T-2, \ldots, 2, 1\}$. For any time period t, the parameter matrices in the optimal policy are given as:

$$K_t = \begin{pmatrix} K_t^{11} & K_t^{12} \\ K_t^{21} & K_t^{22} \end{pmatrix} = A_t^{-1} B_t,$$

where

$$A_t = \begin{pmatrix} 2\tau I_b + 2Q_{t+1}^{11} & -2Q_{t+1}^{11} + 2Q_{t+1}^{12} \\ -2Q_{t+1}^{11} + Q_{t+1}^{21} & 2\nu \mathbf{1}\mathbf{1}' + 2Q_{t+1}^{11} - Q_{t+1}^{21} - Q_{t+1}^{12} + 2Q_{t+1}^{22} \end{pmatrix},$$

$$B_t = \begin{pmatrix} -2Q_{t+1}^{11} & -2Q_{t+1}^{12} \\ 2Q_{t+1}^{11} - Q_{t+1}^{21} & Q_{t+1}^{12} - 2Q_{t+1}^{22} \end{pmatrix},$$

$$\begin{aligned} G_t &= \begin{bmatrix} G_t^1 \\ G_t^2 \end{bmatrix} \\ &= A_t^{-1} \begin{bmatrix} (-(\Pi_{t+1}^1)' | -(\beta+\mu) I_b, \beta I_b) \\ ((\Pi_{t+1}^1)' - (\Pi_{t+1}^2)' | 0_b, 0_b) \end{bmatrix}, \\ F_t &= \begin{bmatrix} F_t^1 \\ F_t^2 \end{bmatrix} \\ &= A_t^{-1} \begin{bmatrix} -(\Lambda_{t+1}^1)' \\ (\Lambda_{t+1}^1)' - (\Lambda_{t+1}^2)' + (\alpha + 2\nu \bar{c} \mathbf{1}) \end{bmatrix}. \end{aligned}$$

Furthermore, we have the following recursions for Q_t, Π_t and Λ_t:

$$\begin{aligned} Q_t = {} & \tau \begin{pmatrix} (K_t^{11})' \\ (K_t^{12})' \end{pmatrix} \begin{pmatrix} K_t^{11} & K_t^{12} \end{pmatrix} + \nu \begin{pmatrix} (K_t^{21})' \\ (K_t^{22})' \end{pmatrix} \mathbf{1} \cdot \mathbf{1}' \begin{pmatrix} K_t^{21} & K_t^{22} \end{pmatrix} + h \begin{bmatrix} I_b & 0_b \\ 0_b & 0_b \end{bmatrix} \\ & + \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}, \end{aligned}$$

$$
\begin{aligned}
\Pi_t &= ([0_{b\times 2(T-t)}| - (\mu+\beta)I_b, \beta I_b] + 2\tau G_t^1)'(K_t^{11}, K_t^{12}) - 2\nu(G_t^2)'\mathbf{1}\cdot\mathbf{1}'(K_t^{21}, K_t^{22}) \\
&\quad +2\begin{pmatrix} G_t^1 - G_t^2 \\ G_t^2 \end{pmatrix}' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
&\quad + \begin{bmatrix} \Pi_{t+1}^1 & \Pi_{t+1}^2 \\ 0_{b\times 1} & 0_{b\times 1} \\ 0_{b\times 1} & 0_{b\times 1} \end{bmatrix} \begin{pmatrix} I + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}, \\
\Lambda_t &= 2\tau(F_t^1)'(K_t^{11}, K_t^{12}) \\
&\quad +[(\alpha + \bar{c}\mathbf{1})' - 2\nu(F_t^2)'\mathbf{1}\cdot\mathbf{1}'](K_t^{21}, K_t^{22}) + 2[h\xi', 0] \\
&\quad +2[\begin{pmatrix} F_t^1 - F_t^2 \\ F_t^2 \end{pmatrix}]' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
&\quad +(\Lambda_{t+1}^1, \Lambda_{t+1}^2) \begin{pmatrix} I + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}.
\end{aligned}
$$

For any given set of price trajectories, after the parameter matrices are calculated off-line in advance, we can compute the customer's optimal purchase and consumption trajectories in a forward fashion. For given initial reference prices $\boldsymbol{r_1}$ and price trajectories $\{\boldsymbol{p_t}\}_{t=1}^T$, we calculate the corresponding reference price trajectories, $\{\boldsymbol{r_t}\}_{t=1}^T$. Starting from $t = 1$, for given $\boldsymbol{y_t}$, $\boldsymbol{\mathcal{C}}_t$, $\boldsymbol{p_t}$ and $\boldsymbol{r_t}$, we calculate $\boldsymbol{x_t}$, which along with inventory level $\boldsymbol{y}_t$, and aggregate consumption level $\boldsymbol{\mathcal{C}}_t$, determine $\mathbf{y}_{t+1}$ and $\boldsymbol{\mathcal{C}}_{t+1}$. We move to the next period and repeat the process until we reach the end of the horizon.

3.4.3 Retailer's Profit Maximization Problem

Using the derived demand and consumption response functions, $\{\boldsymbol{x}_t\}_{t=1}^T$ and $\{\boldsymbol{c}_t\}_{t=1}^T$, we can now reformulate the retailer's profit maximization problem, considering price trajectories over the entire policy space. We incorporate the customer's optimal purchase response in the form of constraints in the retailer's price optimization problem that relate $x_{it}^*(\mathbf{P})$ to state variables, reference prices, current prices, and future prices.

The reformulation of the retailer's profit maximization problem is:

$$(P2B) \qquad \max_{p_{it}, \forall i,t} n \sum_{t=1}^{T} \sum_{i=1}^{N} (p_{it} - v_i) x_{it}$$

$$\text{s.t.} \begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{c}_t \end{bmatrix} = K_t \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix} + G_t \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \dots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix} + F_t, \qquad t = 1, ..., T,$$

$$\boldsymbol{r}_{t+1} = \Gamma \boldsymbol{r}_t + \Psi \boldsymbol{p_t}, \ t = 1, ..., T, \ R_1 \text{ given},$$
$$\boldsymbol{\mathcal{C}}_{t+1} = \boldsymbol{\mathcal{C}}_t + \boldsymbol{c}_t, \ t = 1, ..., T, \ \boldsymbol{\mathcal{C}}_1 = 0,$$
$$\boldsymbol{y}_{t+1} = \boldsymbol{y}_t + \boldsymbol{x}_t - \boldsymbol{c}_t, \ t = 1, ..., T, \ \boldsymbol{y}_1 = 0.$$

As the customer's purchase quantities in each period t are linear functions of the price trajectory for the products, the objective function is quadratic and the constraints are linear functions of the elements in the price trajectories. Problem (P2B) is therefore a quadratic program. Using modern commercial optimization solvers, both concave and nonconcave quadratic programs can be solved to optimality tractably. A similar approach can be used to evaluate the retailer's profit for any given set of price trajectories.

3.4.4 Extension: Heterogeneous Customers

In this subsection, we extend our methodology to allow K distinct customer segments, each of which is homogeneous. We use n_k to denote the number of customers in segment k. For each segment k, the representative customer's utility maximization problem has the same form as (P1) and we can simply index the pertinent parameters and variables by k. By Theorem 2, the optimal purchasing-consumption policy of a representative customer in the k-th segment, after model reformulation, can be expressed as a linear function of the elements in the retailer's price trajectories.

Accounting for K distinct customer segments, the retailer's profit maximization problem can be expressed as follows:

$$(\text{P2C}) \qquad \max_{p_{it}, \forall i,t} \sum_{k=1}^{K} \sum_{t=1}^{T} \sum_{i=1}^{N} n_k (p_{it} - v_i) x_{it}^*(\mathbf{P}),$$

where **P** is the matrix of prices. The retailer can solve (P2C) using the solution methodology described in Section 3.4.3. Note that segments of customers who are loyal to a specific product or to a set of products can be handled by appropriately specifying their utilities for the various products. Finally, we note that our model can also be extended to consider different customer segments who arrive in heterogeneous sets of time periods and updating of each segment's reference price only when the segment visits the store and observes the price (as in Wang 2016).

3.5 Numerical Study

In this section, we present numerical examples that illustrate the effects of separating the purchasing and consumption decisions to shed light on both the strengths and limitations of our approach. In our first example, the consumption target is constant over the horizon. We compare the solution of the full model (P1B) and that of a variant of it, which we call (P1C), in which the customer does not hold any inventory and must consume everything in the period of purchase. A formulation of (P1C) appears below.

$$
\begin{aligned}
\text{(P1C)} \quad \max_{x_{it}\forall i,t} & \Big[\sum_{t=1}^{T}\sum_{i=1}^{N}(\beta(R_{it}-p_{it})+\alpha_i)x_{it}-\tau\sum_{t=1}^{T}\sum_{i=1}^{N}x_{it}^2\Big]-\mu m_{T+1} \\
\text{s.t.} \quad & r_{i,t+1}=\sum_j \gamma_{ji}r_{jt}+\sum_j \psi_{ji}p_{jt}, \quad \forall i,t \\
& m_{t+1}=m_t+\sum_i p_{it}x_{it} \quad \forall t \\
& m_{T+1}\leq \bar{M} \\
& m_1=0
\end{aligned}
$$

It is possible to include consumption targets in the foregoing formulation, but we have excluded them here.

In the numerical example, the retailer sells two substitutable products for which the reference prices evolve independently. The retailer sells each product at its regular price for most of the fifteen-period horizon, but offers a promotion in the fourth period and has a final markdown phase from the eleventh period onward. The price trajectories are:

$$\{p_{1t}\}_{t=1}^{15}=15,15,15,10,15,15,15,15,15,15,12,7.5,7.5,7.5,7.5$$

$$\{p_{2t}\}_{t=1}^{15}=16,16,16,13,16,16,16,16,16,16,13,8.5,8.5,8.5,8.5$$

The other problem parameters are: $\bar{M}=2000$, $\tau=1.2$, $\alpha_1=11$, $\alpha_2=13$, $\beta=0.9$, $\gamma_{ii}=0.75\ \forall i$, $\psi_{ii}=0.25\ \forall i$, $\bar{c}=15$, $\nu=40$, $h=2.0$, $\eta=25$, $\xi=10$, $\lambda=3.0$ and $\pi=0.15$. In order to provide a sensible comparison of the customer's optimal responses in the two scenarios, we assume that she has the same budget in both cases. We wish to model a setting in which the customer desires smooth consumption and has an incentive to consume all purchased units by the end of the selling season, so we set both the coefficient for disutility associated with non-smooth consumption (ν) and the coefficient for discounts on inventory consumed beyond the selling horizon (η) to high values in order to induce these desired effects. For simplicity, we have also assumed that the inventory holding costs are identical for both products, i.e. $h_1 = h_2 = h$.

The customer's optimal responses from the purchase-only model and the full model are shown in Figures 3.1 and 3.2, respectively. In Figure 3.1 we present, from top to bottom, the price and reference price trajectory, the customer's optimal purchase, cumulative aggregate consumption, and cumulative aggregate expenditures. Figure 3.2 is organized in a similar fashion with the following two modifications: (i) the second panel also shows the the customer's aggregate consumption of the two products in each period and (ii) the third panel shows the customer's inventory level rather than cumulative aggregate consumption.

We can see that under the purchase-only model, the customer purchases both products at an almost constant rate during regular-price periods, and increases her purchases very slightly during the in-season promotion. Thus, reference prices alone do not induce large demand "spikes' that are commonly observed in practice. Consumption (which is equal to purchases in the purchase-only model) turns out not to be very smooth (which is difficult to discern from the graph due to scaling issues) because the customer purchases more than 40% of her aggregate consumption during the four final markdown periods that constitute less than 27% of the time horizon, so she purchases and consumes at a higher rate during the markdown periods. Under the full model, the customer, aiming to simultaneously satisfy her budget constraint and meet her high per-period consumption target, significantly increases her purchases during promotion periods, stockpiles much of the discounted purchases, and consumes at a constant rate. In addition, (small) pre- and post-promotion dips that are observed in practice and have been reported in empirical studies (Neslin and van Heerde 2008) do appear in the results from the full model.

We have experimented with many combinations of parameters and have found, not surprisingly, that introducing stable consumption targets (which are pertinent for some products) into the purchase-only model dampens fluctuations in purchases. In general, we have found that the only realistic way to induce significant spikes in purchases when consumption is required to be equal to purchases in each period is to set high consumption targets corresponding to special periods (e.g., holidays). A less realistic way is to introduce very strong gain-seeking effects without decreasing marginal returns. However, simply by separating the purchase and consumption decisions, as in the full model, demand spikes, whose amplitudes obviously depend upon the specific parameters, always arise when discounts are offered, with much of the purchases being stockpiled.

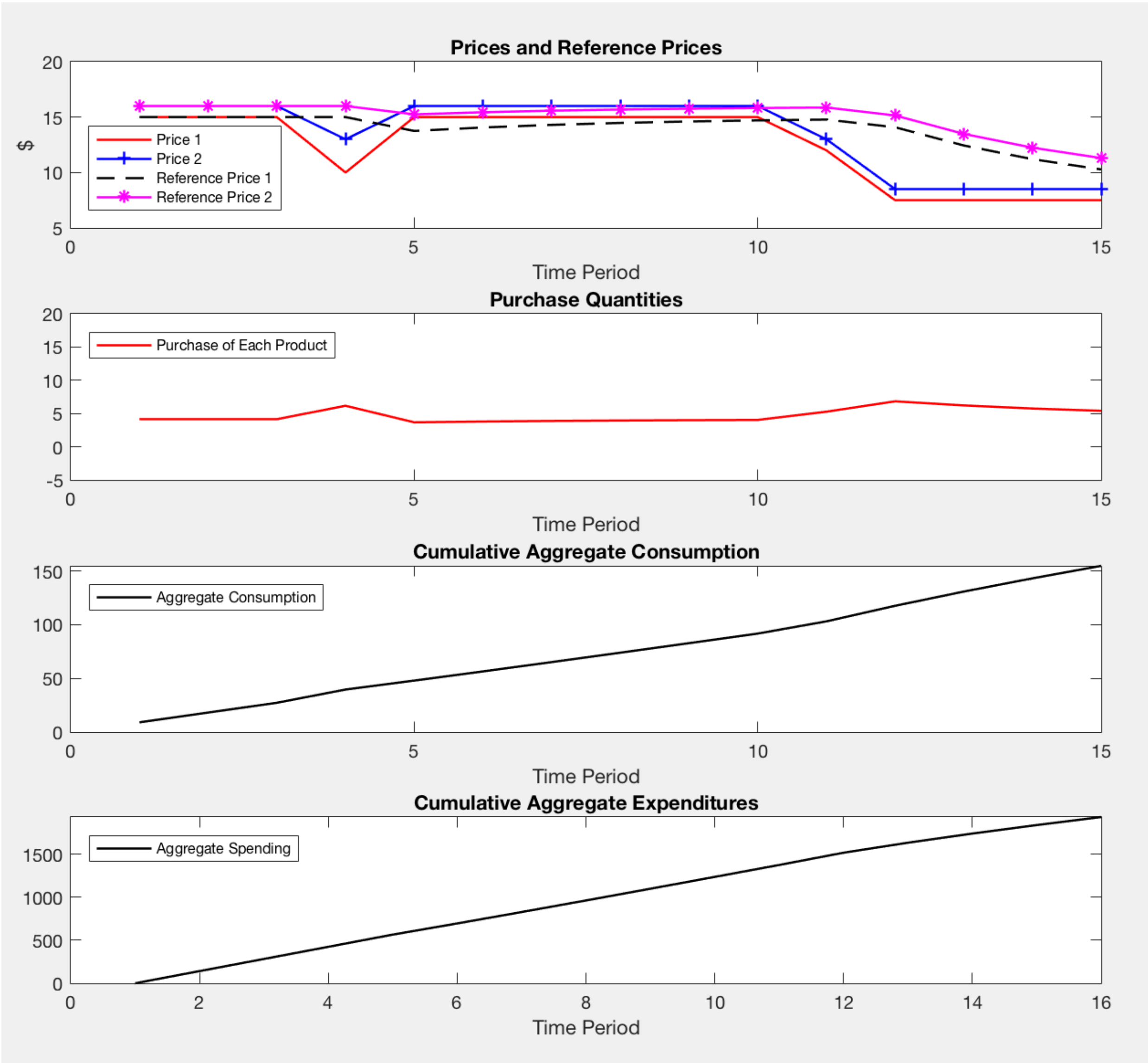


Figure 3.1 Customer's Optimal Response (Purchase-Only Model)

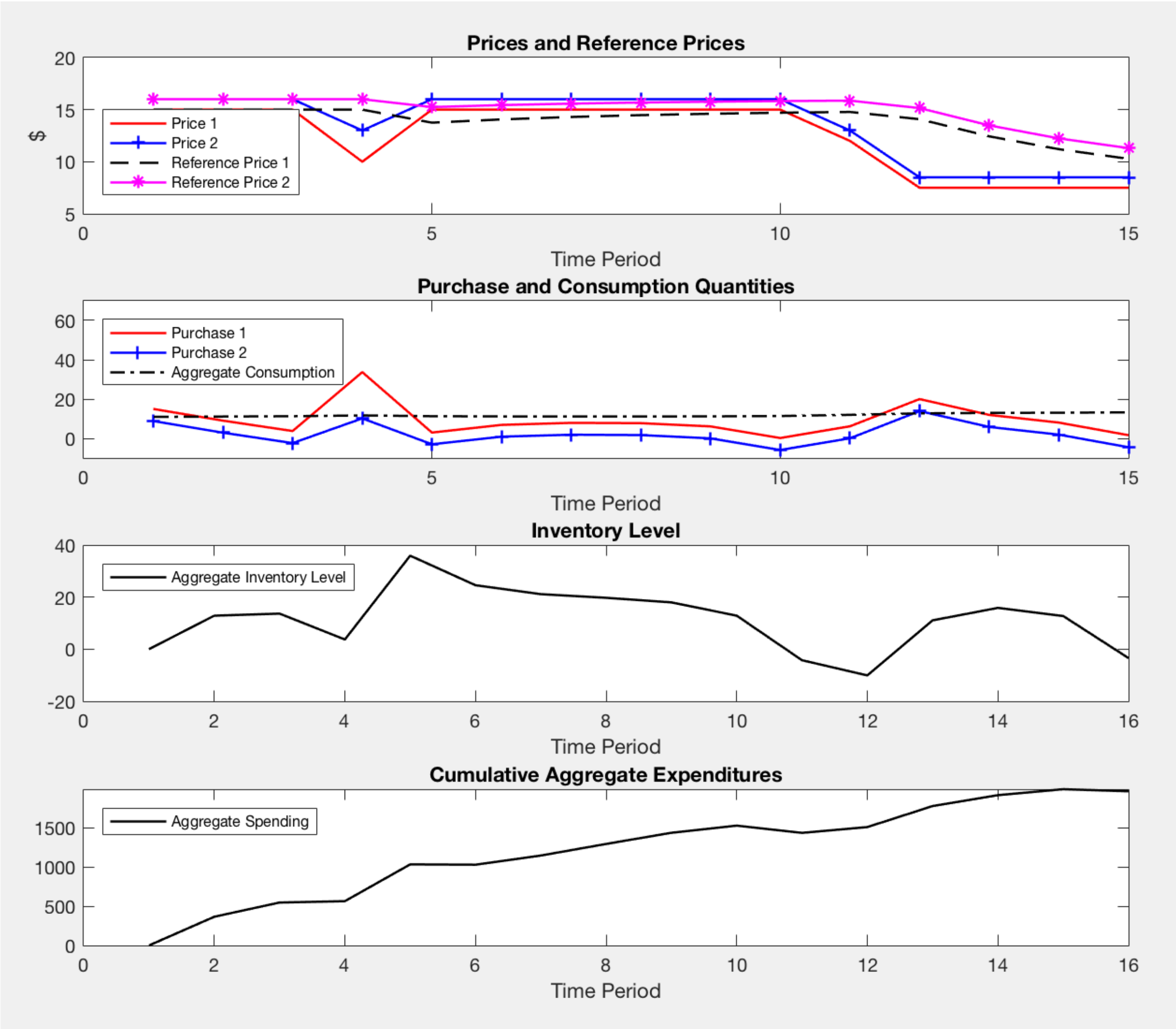


Figure 3.2 Customer's Optimal Response (Full Model)

Although the full model can generate realistic customer purchase trajectories, the quadratic function that we use to penalize both shortages and inventory in the customer's objective function, even if it is centered at a non-zero value, is extremely difficult, if not impossible, to calibrate for this purpose. For example, the inventory holding cost parameter, h, and the center of the quadratic inventory cost function, ξ, must be sufficiently high to ensure that the non-negativity constraint on inventory level is not severely violated. As an example, in the solution shown in Figure 3.2 for a case where h is quite high and $\xi = 10$, the customer's inventory level is close to ξ when the true optimal values should be zero, but later in the horizon, the inventory level becomes negative prior to the end-of-season markdown period. Here, h is too high to accurately represent true inventory holding costs, thus leading to distortions in the tradeoffs in the customer's objective function, and the positive value of ξ distorts the purchasing pattern of the customer early in the horizon as she has an incentive to bring her inventory level up to ξ.

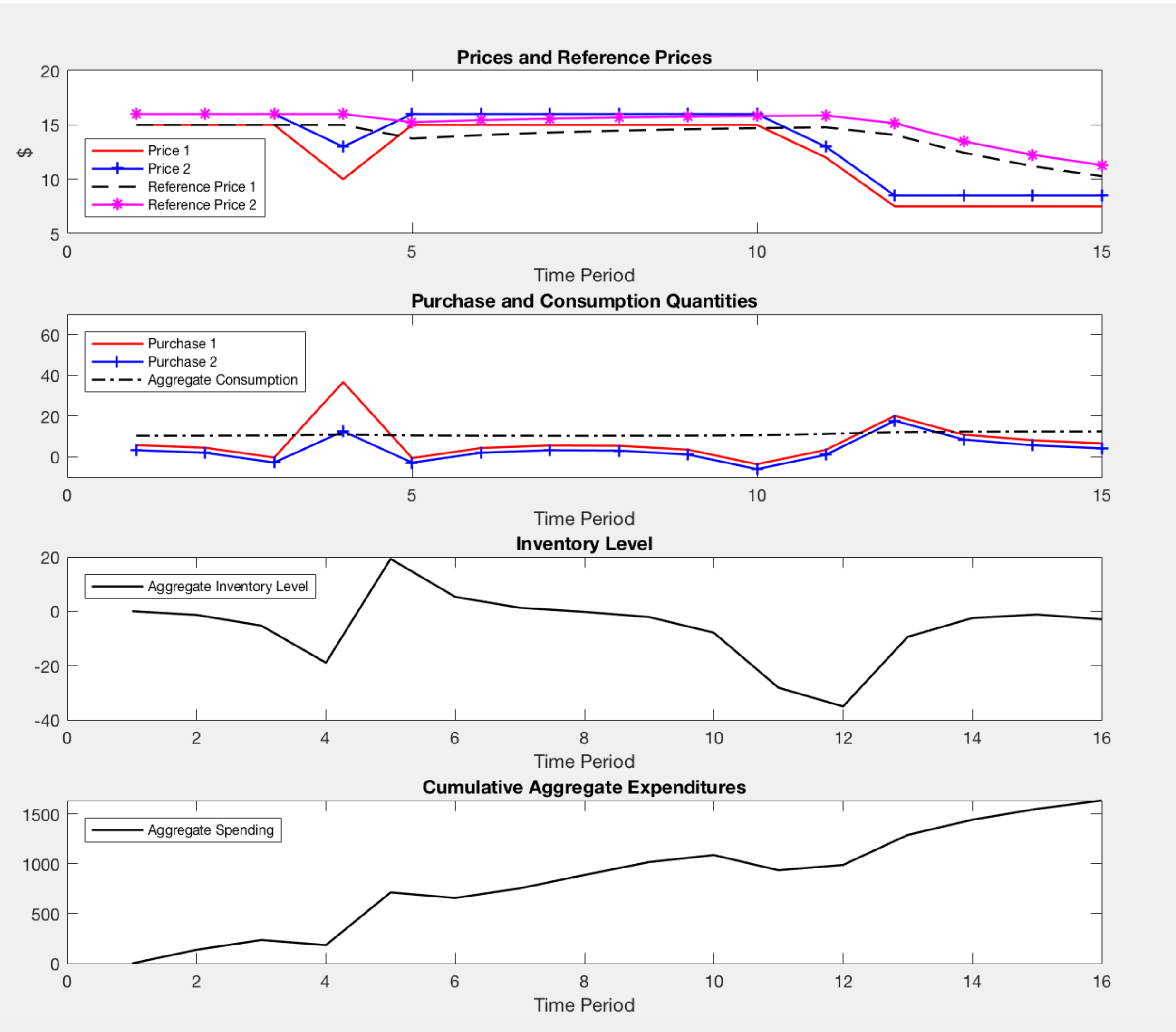


Figure 3.3 Customer's Optimal Response (Full Model with $\xi = 0$)

To further illustrate these effects, in Figure 3.3, we present the customer's optimal responses for the full model under the same parameters, except that $\xi = 0$. Observe that the inventory trajectory shifts downward, when compared to the results shown in Figure 3.2, and the inventory level becomes quite negative, especially prior to the final markdown periods. These results elucidate the limitations of our approach. It would be extremely challenging to specify the coefficients in the customer's objective function even when each has an economic or physical meaning; any further calibration seems to be beyond the scope of what even a sophisticated retailer can do. We discuss ideas for a path forward in the concluding section.

3.6 Conclusions and Future Research Directions

In this chapter, we consider a problem retailers face when setting prices, including promotional prices, over a multi-period planning horizon for multiple substitutable products. We consider the problem in a setting in which customers anticipate the retailer's pricing strategy and the retailer anticipates the customers' purchasing decisions. We formulate the problem as a two-stage game in which the retailer chooses prices and the customers respond by selecting whether, when, and how much to purchase and consume. We develop a unified leader-follower framework in which the retailer chooses price trajectories for the products over some time horizon and customers make purchasing and consumption decisions. The separation of the customer's problem from the retailer's problem allows us to capture a much richer set of considerations in the customer's optimization problem than in past research. We are also able to incorporate fairly general reference price formation processes that include cross-product effects. We explicitly separate customer purchasing and consumption decisions, thus also capturing implicit stockpiling decisions.

We develop a methodology for solving the customer's problem for arbitrary price trajectories based on a linear quadratic control formulation of an approximation of the customer's utility maximization problem. We derive analytical representations for the customer's optimal decisions as simple linear functions of prices, reference prices, inventory levels (as state variables), and the cumulative aggregate consumption level (as a state variable). The separation of the customer's two types of decisions turns out to be important in the ability to induce realistic purchasing patterns, including demand spikes like those observed in practice and the associated stockpiling, as well as pre- and post-promotion dips, which are also observed in practice. We then embed the consumer's optimal policy (in analytic form) in the retailer's profit maximization problem and show that the retailer's problem can be reformulated as a quadratic program which is easy to solve numerically. Despite the additional generality of our problem context, our solution methodology is computationally tractable. Our approach can be extended to allow multiple customer segments.

We hope that our research, which provides a new framework for retail price optimization in which the customer's complicated multi-product, multi-period decisions are modeled via an optimization problem, serves as a catalyst for more research in this direction. Our approach has several limitations. Accordingly, we conclude this chapter by discussing several future research directions. As noted in the previous section, more research is needed to develop mechanisms to handle the non-negativity constraints on inventory without meaningfully distorting the customer's objective function. One possibility is to replace the quadratic inventory cost function with a different second-order expression that explicitly accounts for the asymmetric consequences of shortages and inventory. As the resulting optimization problem is no longer a linear quadratic control problem, it will be necessary to develop a new representation of the customer's approximately optimal solution in closed form, which poses a challenging technical problem. We also note that determining the values of coefficients in the customer's optimization problem is a difficult task. One would need sales observations for a variety of price trajectories for each customer segment, and the fitted values of the coefficients may not work well for a wide range of price trajectories. It would be desirable for each coefficient to represent a value that has an economic or substantive interpretation so that a retailer can specify it directly. Finally, it would also be interesting explore whether our approach can be generalized to consider multiple related product categories that may be linked via reference prices or budget constraints, and multiple product generations in settings where valuations and reference prices carry over to some extent from one generation to the next.

Appendix

Appendix C: Proof of Theorem 2

In this section, we prove Theorem 2 via mathematical induction. As problem (P1B) is stage-wise separable, it can be solved recursively by dynamic programming. For ease of exposition, we assume that the holding cost is the same for all products within the product category and provide our proof in matrix form for brevity.

First, we consider the problem at Stage T:

$$\begin{aligned} J_T(\boldsymbol{r}_T, m_T, \boldsymbol{\mathcal{C}}_T, \boldsymbol{y}_T)) \quad &= \max_{\boldsymbol{x_T},\boldsymbol{c_T}} -\tau\|\boldsymbol{x}_T\|^2 + \beta(\boldsymbol{r}_T - \boldsymbol{p}_T)'\boldsymbol{x}_T - \nu\boldsymbol{c}_T'((\mathbf{1})(\mathbf{1})')\boldsymbol{c}_T + [\alpha + 2\nu\bar{c}\mathbf{1}]'\boldsymbol{c}_T \\ &- h\|\boldsymbol{y}_T\|_2^2 + 2h\xi'\boldsymbol{y}_T - \eta\|\mathbf{y}_T + \boldsymbol{x}_T - \boldsymbol{c}_T\|_2^2 - \mu[m_T + (\boldsymbol{p}_T)'\boldsymbol{x}_T] \\ &+ \lambda(\mathbf{1})'[\boldsymbol{\mathcal{C}}_T + \boldsymbol{c_T}] - \pi\|\boldsymbol{\mathcal{C}}_T + \boldsymbol{c}_T\|_2^2. \end{aligned}$$

Setting the derivatives of J_T with respect to $\boldsymbol{x}_T$ and $\boldsymbol{c}_T$ equal to zero yields the first order necessary conditions:

$$2\tau\boldsymbol{x}_T - \beta\boldsymbol{r}_T + \beta\boldsymbol{p}_T + 2\eta(\boldsymbol{y}_T + \boldsymbol{x}_T - \boldsymbol{c_T}) + \mu\boldsymbol{p}_T = 0, \tag{3.1}$$

$$2\nu((\mathbf{1})(\mathbf{1})')\boldsymbol{c}_T - \alpha - 2\nu\bar{c}\mathbf{1} - 2\eta(\boldsymbol{y}_T + \boldsymbol{x}_T - \boldsymbol{c}_T) - \lambda\mathbf{1} + 2\pi(\boldsymbol{\mathcal{C}}_T + \boldsymbol{c}_T) = 0. \tag{3.2}$$

Solving both (3.1) and (3.2) for c_T, we have

$$\boldsymbol{c}_T = \frac{\tau+\eta}{\eta}\boldsymbol{x}_T + \boldsymbol{y}_T + \frac{1}{2\eta}[(\mu+\beta)\boldsymbol{p}_T - \beta\boldsymbol{r}_T], \tag{3.3}$$

$$\boldsymbol{c}_T = \Sigma^{-1}\{\eta\boldsymbol{y}_T + \eta\boldsymbol{x}_T + (\nu\bar{c} + \frac{\lambda}{2})\mathbf{1} + \frac{\alpha}{2} - \pi\boldsymbol{\mathcal{C}}_T\}, \tag{3.4}$$

where $\Sigma = \nu(\mathbf{1})(\mathbf{1})' + (\pi+\eta)I_b$ is positive definite.

Equalizing the right hand sides of (3.3) and (3.4) yields

$$\frac{\tau+\eta}{\eta}\boldsymbol{x}_T + \boldsymbol{y}_T + \frac{1}{2\eta}[(\mu+\beta)\boldsymbol{p}_T - \beta\boldsymbol{r}_T] = \Sigma^{-1}\{\eta\boldsymbol{y}_T + \eta\boldsymbol{x}_T + (\nu\bar{c} + \frac{\lambda}{2})\mathbf{1} + \frac{\alpha}{2} - \pi\boldsymbol{\mathcal{C}}_T\},$$

from which we obtain:

$$\boldsymbol{x}_T^* = \Phi^{-1}\{\frac{1}{2}\{\Sigma[\beta\boldsymbol{r}_T - (\mu+\beta)\boldsymbol{p}_T] + \alpha\eta\} - (\nu\eta(\mathbf{1})(\mathbf{1})' + \pi\eta)\boldsymbol{y}_T + \eta(\nu\bar{c} + \frac{\lambda}{2})\mathbf{1} - \eta\pi\boldsymbol{\mathcal{C}}_T\}, \tag{3.5}$$

where $\Phi = (\eta\tau + \eta\pi + \tau\pi)I_b + \nu(\tau+\eta)(\mathbf{1})(\mathbf{1})'$ is positive definite.

Let us write the expression of x_T^* in a more compact form:

$$\boldsymbol{x}_T^* = [K_T^{11}, K_T^{12}]\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} + D_T^1,$$

where

$$
\begin{aligned}
K_T^{11} &= -\Phi^{-1}[\nu\eta(\mathbf{1}\mathbf{1}') + \pi\eta I_b],\ K_T^{12} = -\Phi^{-1}\eta\pi, \\
D_T^1 &= \Phi^{-1}\{\frac{1}{2}[\beta \boldsymbol{r_T} - (\mu+\beta)P_T] + \frac{\eta}{2}\alpha + \eta(\nu\bar{c} + \frac{\lambda}{2})\mathbf{1}\} \\
&= [G_T^{11}, G_T^{12}]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + F_T^1,
\end{aligned}
$$

with

$$
\begin{aligned}
G_T^{11} &= -\frac{\mu+\beta}{2}\Phi^{-1}, \\
G_T^{12} &= \frac{\beta}{2}\Phi^{-1}, \\
F_T^1 &= \Phi^{-1}\{\frac{\eta}{2}\alpha + \eta(\nu\bar{c} + \frac{\lambda}{2})\mathbf{1}\}.
\end{aligned}
$$

Furthermore,

$$
\boldsymbol{c}_T^* = [K_T^{21}, K_T^{22}]\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} + D_T^2, \tag{3.6}
$$

where

$$
\begin{aligned}
K_T^{21} &= \frac{\tau+\eta}{\eta}K_T^{11} + I_b,\ K_T^{22} = \frac{\tau+\eta}{\eta}K_T^{12}, \\
D_T^2 &= \frac{\tau+\eta}{\eta}D_T^1 + \frac{1}{2\eta}[(\mu+\beta)\boldsymbol{p_T} - \beta\boldsymbol{r_T}] \\
&= [G_T^{21}, G_T^{22}]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + F_T^2,
\end{aligned}
$$

with

$$
\begin{aligned}
G_T^{21} &= \frac{1}{2\eta}(\mu+\beta)I_b - \frac{\tau+\eta}{\eta}\frac{\mu+\beta}{2}\Phi^{-1}, \\
G_T^{22} &= \frac{\beta}{2}\frac{\tau+\eta}{\eta}\Phi^{-1} - \frac{\beta}{2\eta}I_b, \\
F_T^2 &= \frac{\tau+\eta}{\eta}\Phi^{-1}\{\frac{\eta}{2}\alpha + \eta(\nu\bar{c} + \frac{\lambda}{2})\mathbf{1}\}.
\end{aligned}
$$

Thus, we can express the optimal decision at stage T as:

$$
\begin{aligned}
\begin{bmatrix} \boldsymbol{x}_T^* \\ \boldsymbol{c}_T^* \end{bmatrix} &= K_T\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} + D_T \\
&= \begin{bmatrix} K_T^{11} & K_T^{12} \\ K_T^{21} & K_T^{22} \end{bmatrix}\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} + \begin{bmatrix} G_T^{11} & G_T^{12} \\ G_T^{21} & G_T^{22} \end{bmatrix}\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + \begin{bmatrix} F_T^1 \\ F_T^2 \end{bmatrix},
\end{aligned} \tag{3.7}
$$

where the matrix K_T is independent of $\boldsymbol{p}_T$ and $\boldsymbol{r}_T$ and vector D_T is linear in $\boldsymbol{p}_T$ and $\boldsymbol{r}_T$.

Note that

$$\begin{aligned}
\boldsymbol{y}_T + \boldsymbol{x}_T^* - \boldsymbol{c}_T^* &= \boldsymbol{y}_T + K_T^{11}\boldsymbol{y}_T + K_T^{12}\boldsymbol{\mathcal{C}}_T - K_T^{21}\boldsymbol{y}_T - K_T^{22}\boldsymbol{\mathcal{C}}_T + D_T^1 - D_T^2 \\
&= [I_b + K_T^{11} - K_T^{21}, K_T^{12} - K_T^{22}]\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} + [G_T^{11} - G_T^{21}, G_T^{12} - G_T^{22}]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} \\
&\quad + [F_T^1 - F_T^2].
\end{aligned}$$

Substituting the expressions for $\boldsymbol{x}_T^*$ and $\boldsymbol{c}_T^*$ from (3.7) into $J_T(\boldsymbol{r}_T, m_T, \boldsymbol{\mathcal{C}}_T, \boldsymbol{y}_T)$ yields:

$$J_T(\boldsymbol{r}_T, m_T, \boldsymbol{\mathcal{C}}_T, \boldsymbol{y}_T) = -[\boldsymbol{y}_T', \boldsymbol{\mathcal{C}}_T]Q_T\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} + L_T\begin{bmatrix} \boldsymbol{y}_T \\ \boldsymbol{\mathcal{C}}_T \end{bmatrix} - \mu m_T + E_T, \tag{3.8}$$

where Q_T is independent of $\boldsymbol{p}_T$ and $\boldsymbol{r}_T$, and is given as follows:

$$\begin{aligned}
Q_T &= \begin{bmatrix} Q_T^{11} & Q_T^{12} \\ Q_T^{21} & Q_T^{22} \end{bmatrix} \\
&= \tau\begin{pmatrix} (K_T^{11})' \\ (K_T^{12})' \end{pmatrix}\begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} + (\nu)\begin{pmatrix} (K_T^{21})' \\ (K_T^{22})' \end{pmatrix}(\mathbf{1})(\mathbf{1})'\begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} \\
&\quad + h\begin{pmatrix} I_{b\times b} & 0 \\ 0 & 0 \end{pmatrix} + \eta\begin{pmatrix} (I_b + K_T^{11} - K_T^{21})' \\ (K_T^{12} - K_T^{22})' \end{pmatrix}\begin{pmatrix} I_b + K_T^{11} - K_T^{21} & K_T^{12} - K_T^{22} \end{pmatrix} \\
&\quad + \pi\begin{pmatrix} (K_T^{21})' \\ (K_T^{22})' \end{pmatrix}\begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} + \pi\begin{pmatrix} 0 & 0 \\ 2K_T^{21} & I_{b\times b} + 2K_T^{22} \end{pmatrix}.
\end{aligned}$$

L_T is linear in $\boldsymbol{p}_T$ and $\boldsymbol{r}_T$, and is given as follows:

$$\begin{aligned}
L_T &= [L_T^1, L_T^2] \\
&= \{-2\tau[(G_T^{11}, G_T^{12})\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + F_T^1] + \beta[-I, I]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix}\}'\begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} \\
&\quad -\{2\nu[(G_T^{21}, G_T^{22})\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + F_T^2]'(\mathbf{1}\mathbf{1}') - \alpha' - 2\nu\bar{c}\mathbf{1}'\}\begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} + 2[h\xi', 0] \\
&\quad -2\eta\{[G_T^{11} - G_T^{21}, G_T^{12} - G_T^{22}]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + [F_T^1 - F_T^2]\}'[I_b + K_T^{11} - K_T^{21}, K_T^{12} - K_T^{22}] \\
&\quad -\{\mu[I, 0]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix}\}'\begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} + \lambda\mathbf{1}'[K_T^{21}, I_b + K_T^{21}] \\
&\quad -2\pi\{[G_T^{21}, G_T^{22}]\begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \end{bmatrix} + F_T^2\}'[0, I_b] \\
&= [\boldsymbol{p}_T', \boldsymbol{r}_T']\Pi_T + \Lambda_T,
\end{aligned}$$

with

$$
\begin{aligned}
\Pi_T &= \{-2\tau(G_T^{11}, G_T^{12}) + \beta[-I, I]\}' \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} \\
&\quad -\{2\nu[(G_T^{21}, G_T^{22})]'(\mathbf{1}\mathbf{1}')\} \begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} \\
&\quad -2\eta\{[G_T^{11} - G_T^{21}, G_T^{12} - G_T^{22}]\}'[I_b + K_T^{11} - K_T^{21}, K_T^{12} - K_T^{22}] \\
&\quad -\mu[I, 0]' \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} - 2\pi\{[G_T^{21}, G_T^{22}]\}'[0, I_b], \\
\Lambda_T &= -2\{\tau F_T^1\}' \begin{pmatrix} K_T^{11} & K_T^{12} \end{pmatrix} \\
&\quad -\{2\nu\{F_T^2\}'(\mathbf{1}\mathbf{1}') - \alpha' - 2\nu\bar{c}\mathbf{1}'\} \begin{pmatrix} K_T^{21} & K_T^{22} \end{pmatrix} + 2[h\xi', 0] \\
&\quad -2\eta[F_T^1 - F_T^2]'[I_b + K_T^{11} - K_T^{21}, K_T^{12} - K_T^{22}] \\
&\quad +\lambda\mathbf{1}'[K_T^{21}, I_b + K_T^{21}] - 2\pi\{F_T^2\}'[0, I_b],
\end{aligned}
$$

and E_T is quadratic in $\boldsymbol{p}_T$ and $\boldsymbol{r}_T$ and independent of state variables $\boldsymbol{y}_T$ and $\boldsymbol{\mathcal{C}}_T$. Because E_T does not play a role in the backward recursion, we omit its detailed expression.

General DP recursion:
We now derive the complete solution of the problem using backward induction. Let us assume that the optimal utility-to-go in time period $t+1$ takes the following form:

$$
\begin{aligned}
& J_{t+1}(\boldsymbol{r}_{t+1}, m_{t+1}, \boldsymbol{\mathcal{C}}_{t+1}, \boldsymbol{y}_{t+1}) && (3.9)\\
= & -[\boldsymbol{y}_{t+1}', \boldsymbol{\mathcal{C}}_{t+1}]Q_{t+1} \begin{bmatrix} \boldsymbol{y}_{t+1} \\ \boldsymbol{\mathcal{C}}_{t+1} \end{bmatrix} + L_{t+1} \begin{bmatrix} \boldsymbol{y}_{t+1} \\ \boldsymbol{\mathcal{C}}_{t+1} \end{bmatrix} - \mu m_{t+1} + E_{t+1}, && (3.10)
\end{aligned}
$$

where

$$
\begin{aligned}
L_{t+1} &= [\boldsymbol{p}_T', \boldsymbol{r}_T', \boldsymbol{p}_{T-1}', \boldsymbol{r}_{T-1}', \ldots, \boldsymbol{p}_{t+2}', \boldsymbol{r}_{t+2}', \boldsymbol{p}_{t+1}', \boldsymbol{r}_{t+1}']\Pi_{t+1} + \Lambda_{t+1} \\
&= [\boldsymbol{p}_T', \boldsymbol{r}_T', \boldsymbol{p}_{T-1}', \boldsymbol{r}_{T-1}', \ldots, \boldsymbol{p}_{t+2}', \boldsymbol{r}_{t+2}', \boldsymbol{p}_{t+1}', \boldsymbol{r}_{t+1}'][\Pi_{t+1}^1, \Pi_{t+1}^2] + [\Lambda_{t+1}^1, \Lambda_{t+1}^2].
\end{aligned}
$$

Note Π_{t+1} is of dimension $2(T-t) \times 2b$, both Π_{t+1}^2 and Π_{t+1}^2 are of dimension $2(T-t) \times b$, Λ_{t+1} is of dimension $1 \times 2b$, both Λ_{t+1}^1 and Λ_{t+1}^2 are of dimension $1 \times b$, and E_{t+1} is independent of both $\boldsymbol{y}_{t+1}$ and $\boldsymbol{\mathcal{C}}_{t+1}$. When $t = T-1$, the form of the optimal utility-to-go in time period T satisfies this form (see (3.8)) and serves as the base case in the mathematical induction.

The optimal utility-to-go in period t is then given by:

$$
\begin{aligned}
J_t(\boldsymbol{r}_t, m_t, \boldsymbol{\mathcal{C}}_t, \boldsymbol{y}_t) &= \max_{\boldsymbol{x}_t, \boldsymbol{c}_t}\{-\tau\|\boldsymbol{x}_t\|_2^2 + \beta(\boldsymbol{r}_t - \boldsymbol{p}_t)'\boldsymbol{x}_t - \nu\boldsymbol{c}_t'((\mathbf{1})(\mathbf{1})')\boldsymbol{c}_t + [\alpha + 2\nu\bar{c}\mathbf{1}]'(\boldsymbol{c}_t) \\
&\quad -h\|\boldsymbol{y}_t\|_2^2 + 2h\xi'\boldsymbol{y}_t + J_{t+1}(\Gamma\boldsymbol{r}_t + \Psi\boldsymbol{p}_t, m_t + \boldsymbol{p}_t'\boldsymbol{x}_t, \boldsymbol{\mathcal{C}}_t + \boldsymbol{c}_t, \boldsymbol{y}_t + \boldsymbol{x}_t - \boldsymbol{c}_t)\} \\
&= \max_{\boldsymbol{x}_t, \boldsymbol{c}_t}\{-\tau\|\boldsymbol{x}_t\|_2^2 + \beta(\boldsymbol{r}_t - \boldsymbol{p}_t)'\boldsymbol{x}_t - \nu\boldsymbol{c}_t'((\mathbf{1})(\mathbf{1})')\boldsymbol{c}_t + [\alpha + 2\nu\bar{c}\mathbf{1}]'(\boldsymbol{c}_t) \\
&\quad +2h\xi'\boldsymbol{y}_t - [\boldsymbol{y}_t + \boldsymbol{x}_t - \boldsymbol{c}_t, \boldsymbol{\mathcal{C}}_t + \boldsymbol{c}_t]' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_t + \boldsymbol{x}_t - \boldsymbol{c}_t \\ \boldsymbol{\mathcal{C}}_t + \boldsymbol{c}_t \end{bmatrix} \\
&\quad +[L_{t+1}^1, L_{t+1}^2] \begin{bmatrix} \boldsymbol{y}_t + \boldsymbol{x}_t - \boldsymbol{c}_t \\ \boldsymbol{\mathcal{C}}_t + \boldsymbol{c}_t \end{bmatrix} - h\|\boldsymbol{y}_t\|_2^2 - \mu(m_t + \boldsymbol{p}_t'\boldsymbol{x}_t) + E_{t+1},
\end{aligned}
$$

Taking partial derivatives with respect to $\boldsymbol{x}_t$ and $\boldsymbol{c}_t$ and setting them equal to zero yields the following first-order necessary conditions:

$$\begin{aligned}
&2\tau\boldsymbol{x}_t-\beta(\boldsymbol{r}_t-\boldsymbol{p}_t)]+2Q^{11}_{t+1}(\boldsymbol{y}_t+\boldsymbol{x}_t-\boldsymbol{c}_t)+2Q^{12}_{t+1}(\boldsymbol{\mathcal{C}}_t+\boldsymbol{c}_t)-(L^1_{t+1})'+\mu\boldsymbol{p}_t=0,\\
&2\nu(\mathbf{1}\mathbf{1}')\boldsymbol{c}_t-[\alpha+2\nu\bar{c}\mathbf{1}]-2Q^{11}_{t+1}(\boldsymbol{y}_t+\boldsymbol{x}_t-\boldsymbol{c}_t)+Q^{21}_{t+1}(\boldsymbol{y}_t+\boldsymbol{x}_t-\boldsymbol{c}_t)-Q^{12}_{t+1}(\boldsymbol{\mathcal{C}}_t+\boldsymbol{c}_t)\\
&\quad+2Q^{22}_{t+1}(\boldsymbol{\mathcal{C}}_t+\boldsymbol{c}_t)-(L^1_{t+1})'+(L^2_{t+1})'=0.
\end{aligned}$$

We can express the first-order necessary conditions above in a more compact form:

$$\begin{bmatrix} A^{11}_t & A^{12}_t \\ A^{21}_t & A^{22}_t \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{c}_t \end{bmatrix}=\begin{bmatrix} B^{11}_t & B^{12}_t \\ B^{21}_t & B^{22}_t \end{bmatrix}\begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix}+\begin{bmatrix} H^1_t \\ H^2_t \end{bmatrix},$$

where

$$\begin{aligned}
A^{11}_t &= 2\tau I_b+2Q^{11}_{t+1},\quad A^{12}_t=-2Q^{11}_{t+1}+2Q^{12}_{t+1},\\
A^{21}_t &= -2Q^{11}_{t+1}+Q^{21}_{t+1},\quad A^{22}_t=2\nu\mathbf{1}\mathbf{1}'+2Q^{11}_{t+1}-Q^{21}_{t+1}-Q^{12}_{t+1}+2Q^{22}_{t+1},\\
B^{11}_t &= -2Q^{11}_{t+1},\quad B^{12}_t=-2Q^{12}_{t+1},\\
B^{21}_t &= 2Q^{11}_{t+1}-Q^{21}_{t+1},\quad B^{22}=Q^{12}_{t+1}-2Q^{22}_{t+1},
\end{aligned}$$

$$\begin{aligned}
H^1_t &= \beta(\boldsymbol{r}_t-\boldsymbol{p}_t)-(L^1_{t+1})'-\mu\boldsymbol{p}_t\\
&= [-(\Pi^1_{t+1})'|-(\beta+\mu)I_b,\beta I_b]\begin{bmatrix}\boldsymbol{p}_T\\ \boldsymbol{r}_T\\ \boldsymbol{p}_{T-1}\\ \boldsymbol{r}_{T-1}\\ \dots\\ \boldsymbol{p}_{t+1}\\ \boldsymbol{r}_{t+1}\\ \boldsymbol{p}_t\\ \boldsymbol{r}_t\end{bmatrix}-(\Lambda^1_{t+1})',\\
H^2_t &= (L^1_{t+1})'-(L^2_{t+1})'+(\alpha+2\nu\bar{c}\mathbf{1})\\
&= [(\Pi^1_{t+1})'-(\Pi^2_{t+1})'|0_b,0_b]\begin{bmatrix}\boldsymbol{p}_T\\ \boldsymbol{r}_T\\ \boldsymbol{p}_{T-1}\\ \boldsymbol{r}_{T-1}\\ \dots\\ \boldsymbol{p}_{t+1}\\ \boldsymbol{r}_{t+1}\\ \boldsymbol{p}_t\\ \boldsymbol{r}_t\end{bmatrix}+(\Lambda^1_{t+1})'-(\Lambda^2_{t+1})'+(\alpha+2\nu\bar{c}\mathbf{1}).
\end{aligned}$$

To summarize, the customer's optimal decision and utility-to-go function at any stage t can be written as linear and quadratic functions of the state variables, respectively, i.e.,

$$\begin{bmatrix} \boldsymbol{x}_t^* \\ \boldsymbol{c}_t^* \end{bmatrix} = K_t \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix} + D_t,$$

where

$$K_t = A_t^{-1} B_t = \begin{pmatrix} K_t^{11} & K_t^{12} \\ K_t^{21} & K_t^{22} \end{pmatrix},$$

and

$$\begin{aligned} D_t &= \begin{pmatrix} D_t^1 \\ D_t^2 \end{pmatrix} \\ &= A_t^{-1} H_t = A_t^{-1} \begin{bmatrix} H_t^1 \\ H_t^2 \end{bmatrix} \\ &= \begin{bmatrix} G_t^1 \\ G_t^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \cdots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix} + \begin{bmatrix} F_t^1 \\ F_t^2 \end{bmatrix}, \end{aligned}$$

with

$$\begin{aligned} G_t &= \begin{bmatrix} G_t^1 \\ G_t^2 \end{bmatrix} \\ &= A_t^{-1} \begin{bmatrix} (-(\Pi_{t+1}^1)' | -(\beta+\mu) I_b, \beta I_b) \\ ((\Pi_{t+1}^1)' - (\Pi_{t+1}^2)' | 0_b, 0_b) \end{bmatrix}, \\ F_t &= \begin{bmatrix} F_t^1 \\ F_t^2 \end{bmatrix} \\ &= A_t^{-1} \begin{bmatrix} -(\Lambda_{t+1}^1)' \\ (\Lambda_{t+1}^1)' - (\Lambda_{t+1}^2)' + (\alpha + 2\nu\bar{c}\mathbf{1}) \end{bmatrix}. \end{aligned}$$

Note that

$$\begin{aligned}
\begin{pmatrix} \boldsymbol{y}_t + \boldsymbol{x}_t - \boldsymbol{c}_t \\ \boldsymbol{\mathcal{C}}_t + \boldsymbol{c}_t \end{pmatrix} &= \begin{pmatrix} \boldsymbol{y}_t + K_t^{11}\boldsymbol{y}_t + K_t^{12}\boldsymbol{\mathcal{C}}_t - K_t^{21}\boldsymbol{y}_t - K_t^{22}\boldsymbol{\mathcal{C}}_t \\ \boldsymbol{\mathcal{C}}_t + K_t^{21}\boldsymbol{y}_t + K_t^{22}\boldsymbol{\mathcal{C}}_t \end{pmatrix} + \begin{pmatrix} D_t^1 - D_t^2 \\ D_t^2 \end{pmatrix} \\
&= \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{pmatrix} \\
&\quad + \begin{pmatrix} G_t^1 - G_t^2 \\ G_t^2 \end{pmatrix} \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \cdots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix} + \begin{pmatrix} F_t^1 - F_t^2 \\ F_t^2 \end{pmatrix}.
\end{aligned}$$

We can now express the optimal utility-to-go in period t as

$$J_t(\boldsymbol{r}_t, m_t, \boldsymbol{\mathcal{C}}_t, \boldsymbol{y}_t) = -[\boldsymbol{y}_t', \boldsymbol{\mathcal{C}}_t]Q_t \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix} + L_t \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{\mathcal{C}}_t \end{bmatrix} - \mu m_t + E_t,$$

where

$$\begin{aligned}
Q_t &= \tau \begin{pmatrix} (K_t^{11})' \\ (K_t^{12})' \end{pmatrix} \begin{pmatrix} K_t^{11} & K_t^{12} \end{pmatrix} + \nu \begin{pmatrix} (K_t^{21})' \\ (K_t^{22})' \end{pmatrix} \mathbf{1} \cdot \mathbf{1}' \begin{pmatrix} K_t^{21} & K_t^{22} \end{pmatrix} + h \begin{bmatrix} I_b & 0_b \\ 0_b & 0_b + \end{bmatrix} \\
&\quad \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix},
\end{aligned}$$

$$
\begin{aligned}
L_t &= [\beta(\boldsymbol{r}_t - \boldsymbol{p}_t) - \mu \boldsymbol{p}_t - 2\tau(D_t^1)]'(K_t^{11}, K_t^{12}) + [(\alpha + \bar{c}\mathbf{1})' - 2\nu(D_t^2)'\mathbf{1}\cdot\mathbf{1}'](K_t^{21}, K_t^{22}) \\
&\quad +2[h\xi', 0] + 2\begin{pmatrix} D_t^1 - D_t^2 \\ (D_t^2)' \end{pmatrix}' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
&\quad +(L_{t+1}^1, L_{t+1}^2)\begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
&= [([0_{b\times 2(T-t)}| - (\mu+\beta)I_b, \beta I_b] + 2\tau G_t^1) \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \cdots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix} + 2\tau F_t^1]'(K_t^{11}, K_t^{12}) \\
&\quad +[(\alpha + \bar{c}\mathbf{1})' - 2\nu(G_t^2 \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \cdots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix} + (F_t^2))'\mathbf{1}\cdot\mathbf{1}'](K_t^{21}, K_t^{22}) + 2[h\xi', 0]
\end{aligned}
$$

$$
\begin{aligned}
& +2[\begin{pmatrix} G_t^1 - G_t^2 \\ G_t^2 \end{pmatrix} \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \cdots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix} \\
& + \begin{pmatrix} F_t^1 - F_t^2 \\ F_t^2 \end{pmatrix}]' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
& + \begin{bmatrix} \boldsymbol{p}_T \\ \boldsymbol{r}_T \\ \boldsymbol{p}_{T-1} \\ \boldsymbol{r}_{T-1} \\ \cdots \\ \boldsymbol{p}_{t+1} \\ \boldsymbol{r}_{t+1} \\ \boldsymbol{p}_t \\ \boldsymbol{r}_t \end{bmatrix}' \begin{bmatrix} \Pi_{t+1}^1 & \Pi_{t+1}^2 \\ 0_{b\times 1} & 0_{b\times 1} \\ 0_{b\times 1} & 0_{b\times 1} \end{bmatrix} + (\Lambda_{t+1}^1, \Lambda_{t+1}^2)] \begin{pmatrix} I + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
= \; & [\boldsymbol{p}_T', \boldsymbol{r}_T', \boldsymbol{p}_{T-1}', \boldsymbol{r}_{T-1}', \ldots, \boldsymbol{p}_{t+1}', \boldsymbol{r}_{t+1}', \boldsymbol{p}_t', \boldsymbol{r}_t']\Pi_t + \Lambda_t,
\end{aligned}
$$

with

$$
\begin{aligned}
\Pi_t \;=\; & ([0_{b\times 2(T-t)}| - (\mu+\beta)I_b, \beta I_b] + 2\tau G_t^1)'(K_t^{11}, K_t^{12}) - 2\nu(G_t^2)'\mathbf{1}\cdot\mathbf{1}'(K_t^{21}, K_t^{22}) \\
& +2\begin{pmatrix} G_t^1 - G_t^2 \\ G_t^2 \end{pmatrix}' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
& + \begin{bmatrix} \Pi_{t+1}^1 & \Pi_{t+1}^2 \\ 0_{b\times 1} & 0_{b\times 1} \\ 0_{b\times 1} & 0_{b\times 1} \end{bmatrix} \begin{pmatrix} I + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}, \\
\Lambda_t \;=\; & 2\tau(F_t^1)'(K_t^{11}, K_t^{12}) \\
& +[(\alpha + \bar{c}\mathbf{1})' - 2\nu(F_t^2)'\mathbf{1}\cdot\mathbf{1}'](K_t^{21}, K_t^{22}) + 2[h\xi', 0] \\
& +2[\begin{pmatrix} F_t^1 - F_t^2 \\ F_t^2 \end{pmatrix}]' \begin{bmatrix} Q_{t+1}^{11} & Q_{t+1}^{12} \\ Q_{t+1}^{21} & Q_{t+1}^{22} \end{bmatrix} \begin{pmatrix} I_b + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix} \\
& +(\Lambda_{t+1}^1, \Lambda_{t+1}^2) \begin{pmatrix} I + K_t^{11} - K_t^{21} & K_t^{12} - K_t^{22} \\ K_t^{21} & I_b + K_t^{22} \end{pmatrix}.
\end{aligned}
$$

Thus, the form of the optimal solution remains the same for period t as for period $t+1$, and by induction, for all periods. This completes the proof of Theorem 2.

Bibliography

Agrawal, N., S. A. Smith. 1996. Estimating Negative Binomial Demand for Retail Inventory Management with Unobservable Lost Sales. *Naval Research Logistics* **43**(6) 839–861.

Agrawal, N., S. A. Smith. 2009. Multi-Location Inventory Models for Retail Supply Chain Management. N. Agrawal, S. A. Smith, eds., *Retail Supply Chain Management: Quantitative Models and Empirical Studies*, Chap. 12. Springer, Boston, MA, 319–347.

Agrawal, N., S. A. Smith. 2013. Optimal Inventory Management for a Retail Chain with Diverse Store Demands. *European Journal of Operational Research* **225**(3) 393–403.

Ailawadi, K., B. Harlam, J. Cesar, D. Trounce. 2007. Quantifying and Improving Promotion Profitability at CVS. *Marketing Science* **26**(4) 566–575.

Anderson, E., N. Vilcassim. 2001. Structural Demand Models for Retailer Category Pricing. Working paper, University of Chicago.

Arslan, H., S. Kachani. 2010. Dynamic Pricing Under Consumer Reference-Price Effects. J. J. Cochran, L. A. Cox, P. Kesinocak, J. P. Kharoufeh, J. C. Smith, eds., *The Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, Hoboken, NJ, 1–8.

Baltas, G. 2005. Modeling Category Demand in Retail Chains. *Journal of Operational Research Society* **56**(11) 1258–1264.

Bell, D. R., J. Chiang, V. Padmanabhan. 1999. The Decomposition of Promotional Response: An Empirical Generalization. *Marketing Science* **18**(4) 504–526.

Bell, D. R., J. M. Lattin. 2000. Looking for Loss Aversion in Scanner Panel Data: The Confounding Effect of Price Response Heterogeneity. *Marketing Science* **19**(2) 185–200.

Belobaba, P. P. 1992. Optimal vs Heuristic Methods for Nested Seat Allocation. Presentation at ORSA/TIMS Joint National Meeting.

Belobaba, P. P., L. R. Weatherford. 1996. Comparing Decision Rules that Incorporate Customer Diversion in Perishable Asset Revenue Management Situations. *Decision Sciences* **27**(2) 343–363.

Bertsekas, D. 1995. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA.

Bertsimas, D., S. De Boer. 2005. Dynamic Pricing and Inventory Control for Multiple Products. *Journal of Revenue and Pricing Management* **3**(4) 303–319.

Besbes, O., I. Lobel. 2015. Intertemporal Price Discrimination: Structure and Computation of Optimal Policies. *Management Science* **60**(1) 92–110.

Blattberg, R. C., G. Eppen, J. Lieberman. 1981. A Theoretical and Empirical Evaluation of Price Deals for Consumer Nondurables. *Journal of Marketing* **45**(1) 116–129.

Briesch, R. A., L. Krishnamurthi, T. Mazumdar, S. P. Raj. 1997. A Comparative Analysis of Reference Price Models. *Journal of Consumer Research* **24**(2) 202–212.

Brumelle, S. L., J. I. McGill. 1993. Airline Seat Allocation with Multiple Nested Fare Classes. *Operations Research* **41**(1) 127–137.

Cachon, G. P., A. G. Kök. 2007. Implementation of Newsvendor Model with Clearance Pricing: How to (and How Not to) Estimate a Salvage Value. *Manufacturing & Service Operations Management* **9**(3) 276–290.

Caro, F., V. Martínez-de-Albéniz. 2013. Operations Management in Apparel Retailing: Processes, Frameworks and Optimization. *BEIO, Boletín de Estadística e Investigación Operativa* **29**(2) 103–116.

Chen, N., J. Nasiry. 2018. Does Loss Aeversion Preclude Price Variation. *Manufacturing & Service Operations Management (Forthcoming)* .

Chen, X., P. Hu, Z. Hu. 2017. Efficient Algorithms for Dynamic Pricing with Reference Price Effect. *Management Science* **63**(12) 4389–4408.

Chew, E. P., C. Lee, R. Liu. 2009. Joint Inventory Allocation and Pricing Decisions for Perishable Products. *International Journal of Production Economics* **120**(1) 139–150.

Chiang, J. 1991. A Simultaneous Approach to the Whether, What and How Much to Buy Questions. *Marketing Science* **10**(4) 297–315.

Chintagunta, P. K. 1993. Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households. *Marketing Science* **12**(2) 184–208.

Chintagunta, P. K. 2002. Investigating Category Pricing Behavior at a Retail Chain. *Journal of Marketing Research* **39**(2) 141–154.

Choi, T. 2007. Pre-season Stocking and Pricing Decisions for Fashion Retailers with Multiple Information Updating. *International Journal of Production Economics* **106**(1) 146–170.

Chung, C., J. Flynn. 2011. Simply Structured Policies for a Dynamic Pricing Problem with Constant Price Elasticity Demand. *2011 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, Singapore, December 6-9, 2011*. 211–215.

Chung, C., J. Flynn, J. Zhu. 2009. The Newsvendor Problem with an In-Season Price Adjustment. *European Journal of Operational Research* **198**(1) 148–156.

Cohen, M. C., J. Kalas, G. Perakis. 2017a. Optimizing Promotions for Multiple Items in Supermarkets. Working paper, Leonard N. Stern School of Business, New York University.

Cohen, M. C., N. Z. Leung, K. Panchamgam, G. Perakis, A. Smith. 2017b. The Impact of Linear Optimization on Promotion Planning. *Operations Research* **65**(2) 446–468.

Cohen, M. C., G. Perakis. 2018. Promotion Optimization in Retail. Working paper, Leonard N. Stern School of Business, New York University.

Cooper, L. G., P. Baron, W. Levy, M. Swisher, P. Gogos. 1999. Promocast: A New Forecasting Method for Promotion Planning. *Marketing Science* **18**(3) 301–316.

Curry, R. E. 1990. Optimal Airline Seat Allocation with Fare Classes Nested by Origins and Destinations. *Transportation Science* **24**(3) 193–204.

Eppen, G., Y. Liebermann. 1984. Why Do Retailers Deal? An Inventory Explanation. *Journal of Business* **57**(4) 519–530.

Erdem, T., S. Imai, M. P. Keane. 2003. Brand and Quantity Choice Dynamics under Price Uncertainty. *Quantitative Marketing and Economics* **1**(1) 5–64.

Fibich, G., A. Gavious, O. Lowengart. 2003. Explicit Solutions of Optimization Models and Differential Games with Nonsmooth (Asymmetric) Reference-Price Effects. *Operations Research* **51**(5) 721–734.

Foekens, E. W., S. P. Leeflang, D. R. Wittink. 1998. Varying Parameter Models to Accommodate Dynamic Promotion Effects. *Journal of Econometrics* **89**(1-2) 249–268.

Fredrickson, B. L., D. Kahneman. 1993. Duration Neglect in Retrospective Evaluations of Affective Episodes. *Journal of Personality and Social Psychology* **65**(1) 45–55.

Gallego, G., G. van Ryzin. 1997. A Multiproduct Dynamic Pricing Problem and its Applications to Network Yield Management. *Operations Research* **45**(1) 24–41.

Garbarino, E., R. Slonim. 2003. Interrelationships and Distinct Effects of Internal Reference Prices on Perceived Expensiveness and Demand. *Psychology and Marketing* **20**(3) 227–249.

Gönül, F., K. Srinivasan. 1996. Estimating the Impact of Consumer Expectations of Coupons on Purchase Behavior: A Dynamic Structural Model. *Marketing Science* **15**(3) 262–279.

Graves, S. C. 1996. A Multiechelon Inventory Model with Fixed Replenishment Intervals. *Management Science* **42**(1) 1–18.

Greenleaf, E. A. 1995. The Impact of Reference Price Effects on the Profitability of Price Promotions. *Marketing Science* **14**(1) 82–104.

Gupta, S. 1988. Impact of Sales Promotions on When, What, and How Much to Buy. *Journal of Marketing Research* **25**(4) 342–355.

Hardie, B. G. S., E. J. Johnson, P.S. Fader. 1993. Modeling Loss Aversion and Reference Dependent Effects on Brand Choice. *Marketing Science* **12**(4) 378–394.

Hendel, I., A. Nevo. 2006a. Measuring the Implications of Sales and Consumer Stockpiling Behavior. *Econometrica* **74**(6) 1637–1673.

Hendel, I., A. Nevo. 2006b. Sales and Consumer Inventory. *RAND Journal of Economics* **37**(3) 543–561.

Hinterhuber, A., M. Bertini. 2011. Profiting when Consumers Choose Value over Price. *Business Strategy Review* **22**(1) 46–49.

Howell, D. C. 2007. *Statistical Methods for Psychology*. Thomson, Belmont, CA.

Jackson, P. L. 1988. Stock Allocation in a Two-Echelon Distribution System or "What to Do Until Your Ship Comes In". *Management Science* **34**(7) 880–895.

Jackson, P. L., J. A. Muckstadt. 1989. Risk Pooling in a Two Period, Two Echelon Inventory Stocking and Allocation Problem. *Naval Research Logistics* **36**(1) 1–26.

Jeuland, A. P., C. Narasimhan. 1985. Dealing-Temporary Price Cuts-by Seller as a Buyer Discrimination Mechanism. *The Journal of Business* **58**(3) 295–308.

Kahneman, D., A. Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **47**(2) 263–291.

Kim, B., R. C. Blattberg, P. E. Rossi. 1995. Modeling the Distribution of Price Sensitivity and Implications for Optimal Retail Pricing. *Journal of Business & Economic Statistics* **13**(3) 291–303.

Kopalle, P. K., A. G. Rao, J. L. Assunção. 1996. Asymmetric Reference Price Effects and Dynamic Pricing Policies. *Marketing Science* **15**(1) 60–85.

Krishnamurthi, L., T. Mazumdar, S. P. Raj. 1992. Asymmetric Response to Price in Consumer Brand Choice and Purchase Quantity Decisions. *The Journal of Consumer Research* **19**(3) 387–400.

Kumar, V., K. Karande, W. J. Reinartz. 1998. The Impact of Internal and External Reference Prices on Brand Choice: the Moderating Role of Contextual Variables. *Journal of Retailing* **74**(3) 401–426.

Lautenbacher, C. J., S. J. Stidham. 1999. The Underlying Markov Decision Process in the Single-Leg Airline Yield Management Problem. *Transportation Science* **33**(2) 136–146.

Lee, T. C., M. Hersh. 1993. A Model for Dynamic Airline Sear Inventory Control with Multiple Seat Bookings. *Transportation Science* **27**(3) 252–265.

Littlewood, K. 1972. Forecasting and Control of Passenger Bookings. *In Proceedings of the Twelfth Annual AGIFORS Symposium*, Vol. 12. Nathanya, Israel, 95–117.

Liu, Q., T. Otter, G. M. Allenby. 2009. Measurement of Self- and Cross-Price Effect. V. R. Rao, ed., *Handbook of Pricing Research in Marketing*, Chap. 3. Edward Elgar Publishers, Cheltenham, UK, 61–75.

Maglaras, C., J. Meissner. 2006. Dynamic Pricing Strategies for Multiproduct Revenue Management Problems. *Manufacturing & Service Operations Management* **8**(2) 136–148.

Marklund, J., K. Rosling. 2012. Lower Bounds and Heuristics for Supply Chain Stock Allocation. *Operations Research* **60**(1) 92–105.

Mazumdar, T., S. P. Raj, I. Sinha. 2005. Reference Price Research: Review and Propositions. *Journal of Marketing* **69**(4) 84–102.

McGavin, E. J., L. B. Schwarz, J. E. Ward. 1993. Two-interval Inventory-allocation Policies in a One-Warehouse N-identical-retailer Distribution System. *Management Science* **39**(9) 1092–1107.

McGill, J., G. van Ryzin. 1999. Revenue Management: Research Overview and Prospects. *Transportation Science* **33**(2) 233–256.

Monahan, G. E., N. C. Petruzzi, W. Zhao. 2004. The Dynamic Pricing Problem from a Newsvendor's Perspective. *Manufacturing & Service Operations Management* **6**(1) 73–91.

Monroe, K. B. 1973. Buyers' Subjective Perception of Price. *Journal of Marketing Research* **10**(1) 70–80.

Nasiry, J., I. Popescu. 2011. Dynamic Pricing with Loss-Averse Consumers and Peak-End Anchoring. *Operations Research* **59**(6) 1361–1368.

Neslin, S., C. Henderson, J. Quelch. 1985. Consumer Promotions and the Acceleration of Product Purchases. *Marketing Science* **4**(2) 147 – 165.

Neslin, S., H. J. van Heerde. 2008. Promotion Dynamics. *Foundations and Trends in Marketing* **3**(4) 177 – 268.

Popescu, I., Y. Wu. 2007. Dynamic Pricing Strategies with Reference Effects. *Operations Research* **55**(3) 413–429.

Powell, W. B. 2014. Clearing the Jungle of Stochastic Optimization. *In INFORMS Tutorials in Operations Research* 109–137.

Psendorfer, M. 2002. Retail Sales: A Study of Pricing Behavior in Supermarkets. *Journal of Business* **75**(1) 33–66.

Putler, D. S. 1992. Incorporating Reference Price Effects into a Theory of Consumer Choice. *Marketing Science* **11**(3) 287–309.

Robinson, L. W. 1995. Optimal and Approximate Control Policies for Airline Booking with Sequential Nonmonotonice Fare Classes. *Operations Research* **43**(2) 252–263.

Shoemaker, R. W. 1979. An Analysis of Consumer Reactions to Product Promotions. N. Beckwith, M. Houston, R. Mittelstaedt, K. Monroe, S. Ward, eds., *AMA Conference Proceedings*. American Marketing Association, 244–248.

Su, X. 2010. Intertemporal Pricing and Consumer Stockpiling. *Operations Research* **58**(4) 1133–1147.

Thaler, R. H. 2017. Behavioral Economics: Past, Present, and Future. *American Economic Review* **106**(7) 1577–1600.

Tian, Z. 2010. Dynamic Pricing of Seasonal Product without Replenishment: A Discrete Time Analysis. T.C. Cheng, T. Choi, eds., *Innovative Quick Response Programs in Logistics and Supply Chain Management*, Vol. 15. Springer, Boston, MA, 159–180.

van Heerde, H., S. Gupta, D. R. Wittink. 2004. Is 75% of the Sales Promotion Bump Due to Brand Switching? No, only 33% Is. *Journal of Marketing Research* **40**(4) 481–491.

von Neumann, J., O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Wang, Q. 2011. Control Policies for Multi-echelon Inventory Systems with Stochastic Demand. T. Choi, T.C. Cheng, eds., *Supply Chain Coordination under Uncertainty*. Springer, Boston, MA, 83–108.

Wang, Z. 2016. Technical Note - Intertemporal Price Discrimination via Reference Price Effects. *Operations Research* **64**(2) 290–296.

Ward, R. W., J. E. Davis. 1978. Coupon Redemption. *Journal of Advertising Research* **18**(4) 51 – 58.

Winer, R. S. 1986. A Reference Price Model of Brand Choice for Frequently Purchased Products. *Journal of Consumer Research* **13**(2) 250–256.

Wollmer, R. D. 1992. An Airline Seat Management Model for a Single Leg Route When Lower Fare Classes Book First. *Operations Research* **40**(1) 26–37.