Efficient Real-time Policies for Revenue Management Problems

by

Yanzhe Lei

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2018

Doctoral Committee:

Assistant Professor Stefanus Jasin, Co-Chair
Associate Professor Amitabh Sinha, Co-Chair
Professor Xiuli Chao
Assistant Professor Eric. M. Schwartz

Yanzhe Lei

leiyz@umich.edu

ORCID iD: 0000-0003-2292-7336

Acknowledgments

First and foremost, I would like to express my sincerest gratitude to my advisors, Professor Stefanus Jasin and Amitabh Sinha, for their immense support and unreserved mentorship during my Ph.D. study both in academic and in personal terms. Their work ethics profoundly influenced my character. Their enthusiasm has been and will always be constant inspirations in my life. I could not have imagined having better advisors and mentors. I would also like to thank my dissertation committee members, Professor Xiuli Chao and Professor Eric Schwartz, for their comments that significantly improved this dissertation and advice that sparks exciting future research directions.

My special thanks also go to the faculty members and Ph.D. colleagues at the University of Michigan. In particular, I would like to thank Professor Roman Kapuscinski, for his endless guidance throughout my Ph.D. study from the first day I join the program, thinking about selecting advisors, toward the end of my Ph.D. study, worrying about surviving as an assistant professor; to Professor Joline Uichanco, for her insightful suggestions and invaluable contribution to our joint research project; to Professor Cong Shi, for his guidance on finding research problems and cultivating research tastes; to Professor Jun Li, for her tremendous help on my teaching; to Professor Stephen Leider, for his years of support as a PhD coordinator; to Dr. Brian Jones and Ms. Sheila Green, for their support from the PhD Program Office; and to my friends, especially Ruthy Beer, Qi Chen, Hakjin Chung, Yao Cui, Liang Ding, Evgeny Kagan, Anyan Qi, Yixin Wang, for making my journey as a Ph.D. student much more enjoyable.

Last but not least, I am deeply indebted to my parents. Their unconditionally love and support made my academic career possible. A very special acknowledgment goes to my girlfriend, Yuezhu He, for her love and sacrifice during a five-year long-distance relationship - I am incredibly excited about starting a new chapter of my life with you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

This dissertation studies the development of provably near-optimal real-time prescriptive analytics solutions that are easily implementable in a dynamic business environment. We consider several stochastic control problems that are motivated by different applications of the practice of pricing and revenue management. Due to high dimensionality and the need for real-time decision making, it is computationally prohibitive to characterize the optimal controls for these problems. Therefore, we develop heuristic controls with simple decision rules that can be deployed in real-time at large scale, and then show theirs good theoretical and empirical performances. In particular, the first chapter studies the joint dynamic pricing and order fulfillment problem in the context of online retail, where a retailer sells multiple products to customers from different locations and fulfills orders through multiple fulfillment centers. The objective is to maximize the total expected profits, defined as the revenue minus the shipping cost. We propose heuristics where the real-time computations of pricing and fulfillment decisions are partially decoupled, and show their good performances compared to reasonable benchmarks. The second chapter studies a dynamic pricing problem where a firm faces price-sensitive customers arriving stochastically over time. Each customer consumes one unit of resource for a deterministic amount of time, after which the resource can be immediately used to serve new customers. We develop two heuristic controls and show that both are asymptotically optimal in the regime with large demand and supply. We further generalize both of the heuristic controls to the settings with multiple service types requiring different service times and with advance reservation. Lastly, the third chapter considers a general class of single-product dynamic

pricing problems with inventory constraints, where the price-dependent demand function is unknown to the firm. We develop nonparametric dynamic pricing algorithms that do not assume any functional form of the demand model and show that, for one of the algorithm, its revenue loss compared to a clairvoyant matches the theoretic lower bound in asymptotic regime. In particular, the proposed algorithms generalize the classic bisection search method to a constrained setting with noisy observations.

CHAPTER 1

Introduction

In the past few decades, information and computation technology has fundamentally changed every operational perspective of the modern business world. These advancements enable firms to control and optimize various instruments that have direct impacts on firms' economic outcome at a much granular level in real-time. In particular, the concept of Revenue Management (RM), which often refers to the applications of analytics to understand consumer behavior at a micro-economic level and optimize market performance, has been widely adopted by a broad spectrum of industries, including but not limited to the airline, car rental, hotel, retail, and on-demand platforms. When applying RM in practice, the capability of deploying analytics at scale can successfully help firms gain competitive advantages, especially when firms operate in a highly dynamic and delicately engineered market. However, given the complexity of the problems, it can be quite challenging for firms to maintain scalability while sustaining good performances. Motivated by these challenges, this dissertation develops and analyzes several real-time prescriptive analytics solutions that are easily implementable and have good performances both in theory and in practice. More specifically, the first two chapters are motivated by real-world problems in online retail and on-demand service platforms respectively; the third chapter investigates a more fundamental challenge faced by a price-setting firm facing demand model uncertainty.

The first chapter taps into a fundamental decision-making process of any e-commerce retailer (e-tailer): at the arrival of an incoming customer from specific demand location, the e-tailer offers a variety of products in stock quoted at competitive prices and, upon purchase request, ships the requested product from a specific warehouse in its fulfillment network. Throughout this process, the e-tailer's profit is directly affected by pricing and fulfillment decisions. In theory, joint dynamic optimization of these decisions is the most lucrative option since they are mutually affected by each other through the balance of demand and supply. Unfortunately, joint dynamic optimization is usually computationally challenging in practice, as the scale of either decision can be enormous, and the frequency at which either decision is made can be high. We propose a class of computationally efficient policies that are guaranteed to achieve the most of the benefit of joint optimization. The proposed class of policies only requires solving a relaxed joint optimization problem once before the selling season, and then decomposes the real-time pricing and fulfillment decisions thereafter. Using asymptotic analysis, we show that the performance of the proposed policy is very close to the optimal dynamic policy for problems of practical scales. From a managerial perspective, the proposed policy has at least two favorable features: first, being able to separate the real-time computation of pricing and fulfillment decisions allows these two decisions to be handled by different functions, which is a common reality; second, the proposed policy is effective even if the prices across different demand locations are quoted uniformly, which is a desirable pricing practice driven by consumers' fairness perception. From the implementation perspective, the proposed policy can be deployed in real-time at large scale since heavy optimization is required neither by the pricing decision nor by the fulfillment decision.

The second chapter focuses on the design and analysis of pricing policies for service systems with reusable resources, where firms manage a fixed amount of resources to serve customers arriving stochastically over time at a non-stationary price-dependent rate. The arriving customer requires a service that consumes a certain amount of resources for a deterministic amount of time. The resource is reusable in the sense that it can be immediately used to serve a new customer upon the completion of the previous service. The firm's objective is to maximize the total expected revenue by charging price dynamically. This problem captures the fundamental operational trade-off faced by firms providing different types of services. Examples are ubiquitous, ranging from traditional hotels and car rental companies to the emerging cloud computing and on-demand service providers. More importantly, the dynamics of a system of reusable resource differ significantly from those of the canonical dynamic pricing problem since resources can be "sold" repeatedly throughout the selling season as long as the service cycles do not overlap. It is not clear from the existing literature what kind of policy has guaranteed favorable performance. In particular, we provide the first provably near-optimal policy for the dynamic pricing problem of a service system with reusable resource. The policy uses the solution to a relaxed optimization problem as a baseline control and adapts to the realized randomness in real-time according to a novel adjustment scheme that only requires simple linear operations. The adjustment scheme judiciously controls the magnitude of deviation between the realized demand and the expected demand to achieve the right balance between service level and profitability.

Using large deviation analysis, we show that the gap between the revenue generated by our policy and that of the optimal policy is small for problems of practical scale. All the results can be generalized to more complex settings with multiple types of resources, heterogeneous service time requirements, and advance reservation.

In the third chapter, we investigate a fundamental question of dynamic pricing under model uncertainty. In this problem, a firm needs to sequentially choose prices from a continuous range when the underlying demand function is unknown, and the market response to any given prices can only be observed with statistical noise. The firm's objective is to maximize the total expected profit over a finite selling season under inventory constraints. To learn the demand function, conducting price experiments at different price levels is necessary. The critical question is how to do so effectively. More precisely, at which prices should the firm test and how frequently? A significant challenge in answering these questions is that experimenting opportunity is limited both by the finite selling horizon and the capacitated inventory level. Therefore, the firm needs to carefully balance the tradeoff between learning demand information through exploration at various price levels, and earning the maximum revenue by exploiting the market information gathered thus far. Existing methods suggest that, in this scenario, the firm earns higher revenue when it at least knows the functional class of the demand, yet the consequence of assuming the incorrect class of demand model can be disastrous. This leaves the firm in a quandary: either it risks model misspecification or suffers a worse performance guarantee under weaker demand assumptions. We address this problem by proposing rate-optimal policies that do not rely on any information regarding functional form. We formulate the single product dynamic pricing and learning problem as a continuous-armed bandit model, which is a classic machine learning model that explicitly characterizes the exploration-exploitation tradeoff. We then propose a family of policies that generalize the classic bisection search method to the setting with stochastic noises and constraints. The policies adaptively learn whether the resource is going to be depleted or not by the end of the selling horizon, and generate a sequence of pricing intervals that converges to the optimal static price with high probability. Under mild assumptions on the demand curve, we show that the performance of one of our policies is optimal in the sense that its gap from the optimal pricing policy is minimum when comparing all policies.

CHAPTER 2

Joint Dynamic Pricing and Order Fulfillment for E-commerce Retailers

2.1 Abstract

We consider an e-commerce retailer (e-tailer) who sells a catalog of products to customers from different regions during a finite selling season and fulfills orders through multiple fulfillment centers. The e-tailer faces a *Joint Pricing and Fulfillment* (JPF) optimization problem: At the beginning of each period, she needs to jointly decide the price for each product and also how to fulfill an incoming order (i.e., from which warehouse to ship the order). The objective of the e-tailer is to maximize her total expected profits defined as total expected revenues minus total expected shipping costs (all other costs are fixed in this problem). The exact optimal policy for JPF is difficult to solve; so, we propose two heuristic controls that have provably good performance compared to reasonable benchmarks. Our first heuristic control directly uses the solution of a deterministic approximation of JPF as its control parameters. Our second heuristic control improves the first one by adaptively adjusting the original control parameters according to the realized demand. An important feature of the second heuristic control is that it decouples the real-time pricing and fulfillment decisions, making it easy to implement. We show theoretically and numerically that the second heuristic control significantly outperforms the first heuristic control, and is very close to a benchmark that jointly re-optimizes the full deterministic problem at the beginning of every period.

2.2 Introduction

Driven by the growing population of internet users, the retailing industry has witnessed a boom in the e-commerce channel during the past decades. According to U.S. Census

Bureau (2016), for the year of 2015, the sales of e-commerce retail in the United States grew at an impressive rate of 14.63%, which accounted for 68% of the growth of the whole retail sector. While the growth statistic is impressive, online retailing has never been an easy business to run. As pointed out in Rigby (2014), Amazon.com, whose figure is similar to other e-tailers, has averaged only 1.3% in operating margin over the past three years; in contrast, the operating margins for department/discount stores typically run at about 6% to 10%. Despite the razor-thin margin, e-tailers have to spend heavily in expenditure to meet consumers' evolving expectations about their shopping experience. For example, from a logistics perspective, online shopping induces significantly higher fulfillment cost compared to in-store shopping since there are many more additional activities (e.g., packing, out-bound shipping, return handling, etc.) involved with every order made whose costs are not likely to be fully picked up by consumers (Howland, 2016). All these factors put together highlight the importance for e-tailers to operate in a way that maximizes their revenues while at the same time also minimizing their expenditures.

While running an e-commerce business introduces new operational challenges that do not previously exist compared to its brick-and-mortar counterpart, an e-tailer has extra flexibilities in responding to the market by being able to change prices frequently in real-time (Chen, 2014) and reducing outbound shipping cost through tactical order fulfillment (Agatz et al., 2008). Indeed, powered by a vast amount of data and efficient IT infrastructure, e-tailers nowadays actively adjust their prices according to the imbalance between supply and demand and other external factors in the market. This practice, also known as *dynamic pricing*, has been widely adopted in many industries including airlines, car rental, hotel, and cruise. The retailing industry is among the latest entrants, pioneered by Amazon.com, which is reported to adjust its price lists every ten minutes on average (Shpanya, 2014). As reported in the same article, at least 22% of retailers, including Sears, Bestbuy, and Walmart, have also chosen to implement automatic pricing solutions in their online channel and improved their gross margin by 10%.

Aside from the ability to adjust prices in real-time, an e-tailer also has the flexibility to optimize her fulfillment decisions. However, unlike the pricing decisions that are executed online and have an immediate impact on the revenue stream, an e-tailer's fulfillment decisions affect the physical distribution of inventories and have an immediate impact on its operating cost. Among the different parts of an e-tailer's fulfillment plan, outbound shipping is often cited as the primary source of cost (Dinlersoz and Li 2006). For example, Amazon.com spent \$11.54 billion in the fiscal year of 2015 on outbound shipping alone (including sortation and delivery center costs); this roughly represents 10% of its net revenue ($107.01 billion) and a 30% increment over its total costs in 2014 ($8.71 billion) (Amazon.com, 2015). Moreover, driven by consumers' expectation of cheap delivery (Sides and Hogan, 2015), many retailers now offer appealing shipping options for online shoppers such as unconditional free shipping (Nordstrom, Zappos), contingent free shipping (Amazon.com, Jet.com), and free in-store pickup (Macy's, Walmart). It should be noted that even when the shipping fee is applied, many e-tailers simply opt to offer a fixed shipping fee structure regardless of the actual shipping cost due to different shipment weights, speeds and distances (e.g., Overstock.com charges $4.95 to most locations in the United States), which means that the remaining costs are potentially absorbed by the e-tailers themselves. As a consequence, e-tailers are strongly incentivized to find the cheapest fulfillment plan on every single order, since every dollar saved goes directly to the bottom line.

Conceptually, the e-tailer's pricing and fulfillment decisions are closely tied together, since they both immediately affect the balance between supply and demand. On the one hand, an e-tailer's fulfillment strategy affects her pricing decision as the price that maximizes total revenues does not necessarily maximize total expected profits (i.e., revenue minus cost); on the other hand, the effectiveness of a fulfilment strategy heavily depends on the current inventory distribution and forecasted future demands, which in turn are determined by the pricing decision. This interdependency calls for a systematic study of joint pricing and fulfillment optimization.

To illustrate the potential benefit of managing pricing and fulfillment decisions jointly in an e-commerce environment, we describe a simple example. Consider an e-tailer selling a cast-iron grill pan weighing 7.1 lbs to Midwest and West Coast regions. Customers from both regions see the same price posted online. For the purpose of illustration, we assume that the demand is divisible and deterministically determined by $\lambda(p) = 116 - 2p$ for both regions. The price is restricted to within the range of $14.22 and $30.34 (see Camelcamelcamel.com 2016 for a price history of a similar product at Amazon.com). The e-tailer has a distribution network consisting of two fulfillment centers (FCs) located at California (CA) and Illinois (IL), which hold C_{CA} and C_{IL} units of inventory, respectively. Each customer purchases exactly one grill pan, which is to be shipped immediately from either FC using UPS' 3-day select service. Figure 2.1 describes the profit maximization problem faced by the e-tailer, where we use MI (Michigan) and OR (Oregon) as representatives of the Midwest and West Coast regions, respectively. Shipping cost data is gathered from UPS (2016).

Suppose that $C_{IL} = 60$ and $C_{CA} = 56$, i.e., the inventory level in IL is slightly higher than the inventory level in CA. If the e-tailer manages the pricing decision separately from fulfillment assignment, she would first solve a revenue maximization problem: $\max_{p\in[\$14.22,\$30.34]} \{p \cdot (116-2p) + p\cdot(116-2p) : 2\cdot(116-2p) \leq 60+56\}$. The optimal solution to this optimization is given by $p = \$29.00$, which results in 58 units of demand from each MI and OR, and yields a total revenue of $\$29 \times 58 \times 2 = \$3,364.00$. Next, she needs to decide how to fulfill these orders by solving the following cost minimization problem: $\min_{x_{ij}\geq 0}\{\sum_{i\in\{CA,IL\}}\sum_{j\in\{MI,OR\}} c_{ij}x_{ij} \ : \sum_{i\in\{CA,IL\}} x_{ij} = 58, \forall j, \sum_{j\in\{MI,OR\}} x_{ij} \leq C_i, \forall i\}$. The optimal solution is given by $x_{IL,MI} = 58, x_{IL,OR} = 2, x_{CA,MI} = 0, x_{CA,OR} = 56$, which yields a total shipping costs of $\$2,246.10$ and leaves a net total profit of $\$3,364.00 - \$2,246.10 = \$1,117.90$. Suppose now that the e-tailer manages the pricing and fulfillment decisions jointly by solving the following profit maximization problem:

$$\begin{aligned}
\max_{p\in[\$14.22,\$30.34],\ x_{ij}\geq 0} \quad & p(116-2p) + p(116-2p) - \sum_{i\in\{CA,IL\}}\sum_{j\in\{MI,OR\}} c_{ij}x_{ij} \\
\text{s.t.} \quad & \sum_{i\in\{CA,IL\}} x_{ij} = 116-2p, \ \forall j, \quad \sum_{j\in\{MI,OR\}} x_{ij} \leq C_i, \ \forall i.
\end{aligned}$$

The optimal solution to the joint optimization is $p = \$30.34, x_{IL,MI} = x_{CA,OR} = 55.32$, $x_{IL,OR} = x_{CA,MI} = 0$ and the corresponding total net profits is $\$1,249.13$. This stands for a 11.74% improvement in total net profits compared to optimizing price and fulfillment separately. At a closer look, we find that although the increment in price lowers the total revenues, it also reduces total demands so that we no longer ship on the IL-OR and CA-MI routes, which have negative profit margins.

The above example shows the effectiveness of joint pricing and fulfillment optimization, even when the future demands are known exactly. This benefit is further amplified

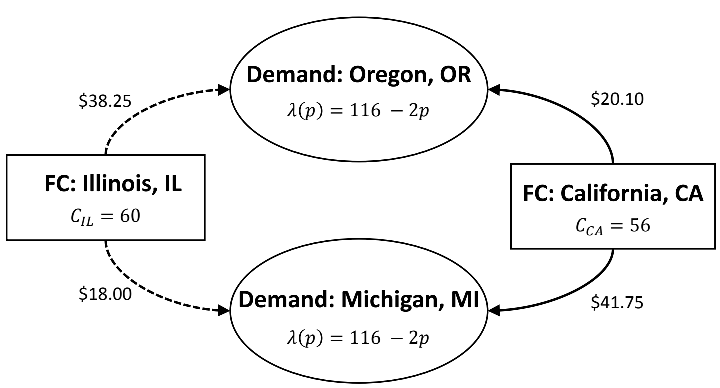


Figure 2.1: A 2-FC 2-Demand-Location Example

by the inventory imbalance across the FC network: if we set $C_{IL} = 18$ and $C_{CA} = 98$ in the above example, the net profits under joint optimization can be about twice as much as that under separate optimization. In reality, even if the initial inventory levels are carefully chosen, inventory imbalance may still happen within the replenishment cycle due to demand uncertainties and various operational difficulties (see Acimovic and Graves 2017 for a detailed identification of the potential causes using real data). On the other hand, the realized demand for an item may depend not only on its own price, but also on the price of other products that may either be complements or substitutes. Since typical e-tailers manage a large number of products, whose inventories are distributed across a large number of FCs, the task of dynamically optimizing the pricing and fulfillment decisions jointly becomes highly challenging and it is a priori not clear whether there is a computationally efficient way to do this. In this paper, we address this issue. We ask: *How should an e-tailer manage the pricing and fulfillment decision for multiple products jointly by utilizing the information regarding the current inventory distribution and future demand projection in a way that maximizes total expected profits?*

Our results and contributions. We consider a multi-period *Joint Pricing and Fulfillment* (JPF) problem where an e-tailer sells multiple products to customers coming from multiple demand locations and demands are fulfilled in real-time through multiple FCs. The decision variables are the price and fulfillment assignment; the objective is to maximize total expected profits. To the best of our knowledge, we are the first in the literature to consider the dynamic version of the JPF problem. This is surprising given the importance of pricing and fulfillment decisions as tactical levers to maximize total expected profits in e-tail setting (see Chapter 2.3 for extensive literature review). Our results and contributions can be summarized as follows:

- We propose a tractable deterministic approximation of JPF. In practice, it is not always feasible for the e-tailer to price-differentiate customers from different locations by charging different prices for the same product during the same period. This constraint introduces complexities that do not previously appear in the relevant literature (see discussions in Chapter 2.4 and 2.5). To address this problem, we propose a novel deterministic relaxation of the original stochastic control problem where all the random variables are approximated by their expected values and the pricing decision is approximated by a *randomization* over a fixed set of discrete prices. We show that there exists a set of discrete prices such that the optimal value of the resulting *Approximate Linear Program* (ALP) well approximates that of JPF (in some sense).

- We develop two easy-to-implement heuristic controls using the solution of ALP. The first heuristic control, *Randomized Pricing and Fulfillment* Control (RPF), simply uses the ALP solution to randomly sample the pricing and fulfillment decisions at each time period. The second heuristic control, which we call *Re-adjust and Re-optimize Pricing and Fulfillment* Control (R^2PF), refines RPF by dynamically updating the pricing and fulfillment decisions. To be precise, using RPF as the base control, at the beginning of every period, R^2PF first adjusts the set of discrete prices over which a new price will be sampled through a real-time perturbation scheme, and then solves several simple tranportation LPs for the fulfillment decisions. We show theoretically and numerically that R^2PF significantly improves RPF. Moreover, R^2PF achieves a performance that is very close to a benchmark control that re-optimizes ALP at the beginning of every period while having a much faster computation time (see Table A.2). To the best of our knowledge, our work is the first in the literature to study a combination of a real-time adjustment of some decision variables (i.e., price) and a re-optimized update of other decision variables (i.e., fulfillment). This idea is potentially useful for other applications where the number of decision variables is large and the problem has some structures that can be exploited.

Aside from the methodological contributions discussed above, our work also highlights the potential managerial benefit of an effective top-down policy for managing both demand (via pricing) and supply (via fulfillment). To put it differently, the purpose of the first step in R^2PF (i.e., price adjustment) is to maintain balance between total available inventories at all FCs and total forecasted future demands from all locations for every product. Moreover, it is done without taking into account total shipping costs. The second step of R^2PF deals with what is left of the first stage: It takes into account the actual inventory distribution across different FCs and computes a fulfillment assignment that minimizes total shipping costs. Our results suggest that these two steps are indispensable in general: Without the aggregate re-balancing in the first step, the fulfillment optimization in the second step will only be minimizing shipping cost without maximizing revenue; and, without the fulfillment optimization in the second step, the aggregate re-balancing in the first step may result in a high shipping cost, which leads to a lower net profit.

Organization of the paper. The related literature is reviewed in Chapter 2.3. In Chapter 2.4, we formally formulate the JPF problem and state our modeling assumptions. We propose an approximation scheme and our performance measure in Chapter 2.5. Chapter 2.6 and 2.7 are devoted to the analysis of our heuristic controls. Numerical simulations are presented in Chapter 2.8. Finally, in Chapter 2.9, we conclude the paper. The proof of all results and the remaining details of numerical experiments can be found in Chapter A.

2.3 Literature Review

In terms of topic, the problem studied in this paper is related to three streams of literature: dynamic pricing, e-commerce fulfillment, and the interaction between pricing and fulfillment-related decisions. In terms of methodology, our work is related to the study of asymptotic performance of re-optimization-based heuristic control and real-time control. We discuss them in turn.

Dynamic Pricing. In the revenue management (RM) literature, research on dynamic pricing studies how a firm should dynamically change their price to balance supply and demand during a finite selling season; see Talluri and van Ryzin (2006) and Özer and Phillips (2012) for comprehensive reviews. Although the idea was popularized by its application in airline ticket pricing, as argued by Boyd and Bilegan (2003), the classic dynamic pricing model can also cover the revenue maximization problem in e-commerce. Several works discuss how to design an optimal pricing policy for specific types of e-tailer's problems. For example, Netessine et al. (2006) and Aydin and Ziya (2008) explore the optimal policy for dynamic pricing and packaging when an e-tailer offers an additional product other than the product requested by consumers as a bundle; Ferreira et al. (2015a) and Fisher et al. (2015) devise pricing decision support systems for large e-tailers and illustrate their effectiveness by conducting field experiments. Compared to the existing models in the RM literature and the papers cited above, our model shares similarity in the price-induced nature of demand generation and some related assumptions (see Chapter 2.4). Unlike the existing literature, though, we jointly consider both the pricing and fulfillment decisions.

E-commerce Fulfillment. The advent of e-commerce has led to substantial research in various aspects of optimizing e-commerce supply chains; see Simchi-Levi et al. (2004) and Agatz et al. (2008) for comprehensive reviews. The fulfillment part of our model focuses exclusively on designing an outbound shipping assignment strategy that helps the e-tailer minimize total shipping costs. This problem was first studied by Xu et al. (2009); they construct a heuristic control that periodically re-evaluates the real-time assignment decisions based on the currently available information, and illustrate its effectiveness using numerical experiments. Their objective is to minimize the number of split shipments. Acimovic and Graves (2014) study a similar problem and develop a heuristic control that minimizes total shipping costs instead of the number of split shipments. Using industry data, they show that their approach captures 36% of the savings on costs induced by the optimal hindsight control. Jasin and Sinha (2015) consider a multi-item fulfillment cost minimization problem. They first propose a heuristic control based on the solution of a deterministic relaxation LP and then show how to improve its performance by carefully constructing a correlated rounding scheme. Since our focus in this work is on the benefit of joint optimization of pricing and fulfillment decisions, for the fulfillment part, we simplify the model in Jasin and Sinha (2015) by requiring that each order consists of exactly one item. However, the additional layer of the pricing decision, as well as the re-adjusting/re-optimization feature of our main heuristic control, precludes a direct generalization of the methodology used in Jasin and Sinha (2015).

Joint pricing and fulfillment-related decisions. There are a few works that study the interplay between an e-tailer's pricing decisions and shipping policy (i.e., the format and the nominal fee charged on deliveries); see, for example, Leng and Becerril-Arreola (2010), Becerril-Arreola et al. (2013) and Gümüş et al. (2013). In our work, we do not explicitly consider the design of shipping policy (the format and the extra charge for deliveries); instead, we simply assume a certain cost structure and analyze how to dynamically adjust both the price and fulfillment decisions given the structure. Closest to ours is Harsha et al. (2016), where a joint pricing and fulfillment planning problem is considered in the setting of omni-channel retail. Specifically, for an omni-channel retailer managing both online and physical channels, inventory held at the brick-and-mortar stores can also be used to fulfill e-commerce demand. There are, however, two critical features that differentiate their work from ours: First, in the omni-channel setting, the retailer can charge different prices at different brick-and-mortar stores and e-commerce channel at the same point in time whereas, in the pure e-commerce setting, the retailer is restricted to applying only a single price to customers from all locations at the same point in time. Second, Harsha et al. (2016) essentially assumes deterministic demand functions, which reduces the problem to a static optimization problem that can be solved before the selling season. In comparison, we assume stochastic demands and focus on the design of dynamic control. Thus, our work complements their work in different dimensions.

Re-optimization-based controls. In the broader dynamic optimization literature where a multi-period stochastic control problem is often intractable, re-optimization is typically used as a heuristic approach due to its simplicity. A re-optimization-based heuristic control first approximates the original stochastic control problem with a simple optimization problem (e.g., an LP) and, as time evolves and uncertainties are realized, the heuristic re-optimizes the approximate optimization problem by updating its parameters to the status quo. In the Operations Management (OM) literature, this idea has been applied to price-based RM (Maglaras and Meissner 2006, Jasin 2014), quantity-based RM (Reiman and Wang 2008, Ciocan and Farias 2012, Jasin and Kumar 2012, 2013), inventory control (Plambeck and Ward 2006, Secomandi 2008, Doğru et al. 2010, Ahn et al. 2015), and vehicle routing (Secomandi and Margot, 2009). In our setting, the proposed approximate optimization can be very large in size for a high-quality approximation. Therefore, full-scale frequent re-optimizations may not be practically feasible. To address this problem, we introduce a new methodological novelty by decoupling the pricing and fulfillment decisions. For our main heuristic control, only the fulfillment assignment decisions involve periodic re-optimization of an LP. The size of this LP is much smaller than the original approximate optimization problem and is decomposable over the products. This makes the re-optimization part of our heuristic control very time-efficient.

Real-time controls. Generally speaking, a real-time control consists of a simple decision rule that can be easily computed as a function (e.g., affine) of a baseline control and realized historical outcomes. Similar to re-optimization-based controls, a real-time control is often used to deal with an intractable multi-period stochastic control problem, and has been applied to robust optimization (Ben-Tal et al. 2004, Bertsimas et al. 2010), portfolio management (Calafiore 2009, Moallemi and Saglam 2012), and dynamic pricing (Atar and Reiman 2012, Jasin 2014, Chen et al. 2015). It is designed to adapt quickly to the observed uncertainties, especially in the setting where speed and time-efficiency are of utmost importance. Hence, it is sometimes preferable to re-optimization-based controls. In our main heuristic control, the pricing decisions are adjusted according to a simple updating rule akin to the one used in Jasin (2014) and Chen et al. (2015) (see Chapter 2.7). However, there is an important difference between our approach and their approach: In both Jasin (2014) and Chen et al. (2015), the adjustment is made directly to the price of each product whereas, in ours, the adjustment is made to the set of discrete prices from which the actual price will be sampled. Thus, our work generalizes the one-point adjustment scheme in the existing literature to a distribution adjustment scheme.

2.4 Problem Formulation

Consider a monopolistic e-tailer selling a catalog of K products to customers in J locations with sales fulfilled from I FCs. Throughout the paper, we will use $[N]$ to denote the set $\{1, \dots, N\}$ for any $N \in \mathbb{N}_+$. The selling season is finite and divided into $T \geq 1$ periods. (Although we assume a discrete-time setting in the analysis, our results can also be applied to a continuous-time setting with Poisson demand. Indeed, our numerical experiment in Chapter 2.8 is conducted in the continuous-time setting.) At the beginning of period t, the e-tailer posts the price vector $\boldsymbol{p}^t = (p_k^t)$ for K products. (We use a boldface letter to denote a vector and its light face with subscript i to denote its i^{th} entry.) For each location $j \in [J]$, the price vector induces a demand vector $\boldsymbol{D}_j^t(\boldsymbol{p}^t) = (D_{jk}^t(\boldsymbol{p}^t))$ with rate vector $\boldsymbol{\lambda}_j(\boldsymbol{p}^t) = (\lambda_{jk}(\boldsymbol{p}^t))$, where $\boldsymbol{\lambda}_j(\boldsymbol{p}_t) = \mathbb{E}\left[\boldsymbol{D}_j^t(\boldsymbol{p}^t)\right]$. (For convenience, we assume stationary rate functions. Our results can also be generalized to the case of non-stationary rates.) Demands across different periods are assumed to be independent, but can be correlated among different products within the same period. (In our model, cross-elasticity is the only thing that connects different products, not the inventory or the fulfillment.) Moreover, as is common in the literature, we allow at most one customer's arrival in each period across all demand locations, i.e., $\sum_{j=1}^{J} \sum_{k=1}^{n} D_{jk}^t(\boldsymbol{p}^t) \leq 1$. This is without loss of generality since we can always slice the selling season fine enough so that at most one customer arrives in each period across all locations. The quantity $\lambda_{jk}(\boldsymbol{p}^t)$ can thus be interpreted as the purchase probability of product k from demand location j in period t. We will also use $\boldsymbol{\lambda}^{tot}(\boldsymbol{p}) = (\sum_{j=1}^{J} \lambda_{jk}(\boldsymbol{p}))_{k=1}^{K}$ to denote the total purchase probability, or aggregate demand rate over all locations. Our model implicitly assumes that a customer only purchases at most one product at a time. (The case where customers purchase multiple products at the same time is challenging to analyze, even from the perspective of pure fulfillment decisions; see Jasin and Sinha 2015. We leave this for future research pursuit.)

A common feature of e-commerce retail is that, at a given time t, customers from all demand locations observe the same price vector $\boldsymbol{p}^t$ from the same website. Compared to brick-and-mortar retailers where prices could be different across different physical stores, this distinct feature limits the e-tailer's degree of freedom in controlling demand intensity from multiple locations. (Technically, the e-tailer can set different prices to different customers by exploiting their profiles. However, such practice may cause severe adverse effect since (1) it will lead to customer's unfair perception, psychological resistance, negative word-of-mouth, and brand switching (Zhan and Lloyd, 2014), and

(2) it is commonly considered as unethical if not unlawful (Reid, 2014).) Indeed, this is also the very feature that makes the analysis of JPF in e-commerce setting more challenging than in the classic RM setting. (See Chapter 2.5 for more discussions.) For each location $j \in [J]$, let $R_j^t(\boldsymbol{p}^t) := (\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)$ denote the realized revenue in period t, where $(\boldsymbol{p}^t)^\top$ indicates the transpose of $\boldsymbol{p}^t$. We call $r_j(\boldsymbol{p}^t) = \mathbb{E}\left[R_j^t(\boldsymbol{p}^t)\right]$ the revenue rate for location j in period t. We use $\mathcal{G}_f$ to denote the $K \times K$ Jacobian matrix for any $\boldsymbol{f} = (f_1, \ldots, f_K) : \mathbb{R}^K \to \mathbb{R}^K$, i.e., $\mathcal{G}_f(x) = [(\nabla f_1(x))^\top; \ldots; (\nabla f_K(x))^\top]$ where $\nabla f_k(x)$ is the gradient of f_k at x. Let $\Omega_p := [p_\ell, p_u]^K \subset \mathbb{R}^K$ and $\Omega_\lambda \subset \mathbb{R}^K$ denote the convex and compact sets of feasible prices and demand rates, respectively. (Without loss of generality, we assume that the domain of prices for all products and demand rates at all locations are the same.) To facilitate our analysis, we make the following assumptions on the underlying demand and revenue rate functions for all $j \in [J]$:

A1. The demand rates $\boldsymbol{\lambda}_j(\boldsymbol{p}) : \Omega_p \to \Omega_\lambda$ and $\boldsymbol{\lambda}^{tot}(\boldsymbol{p}) : \Omega_p \to [0,1]^K$ are invertible, twice-differentiable and monotonically decreasing in its individual argument.

A2. The revenue rates $r_j(\boldsymbol{p})$ are continuous and strictly unimodal with interior maximizers.

A3. For all $\boldsymbol{p} \in \Omega_p$, the absolute eigenvalues of $\mathcal{G}_{\boldsymbol{\lambda}^{tot}}(\boldsymbol{p})$ are bounded from below, whereas the absolute eigenvalues of $\nabla^2 r_j(\boldsymbol{p})$ are bounded from above.

Assumptions A1 and A2 are standard regularity conditions widely assumed in the RM literature (see similar assumptions in Gallego and van Ryzin 1997). The first part of A3 is a natural consequence of the invertibility of the demand function; the second part of A3 is easily satisfied, especially for a compact pricing decision region. Both of them have been assumed in the dynamic pricing literature (e.g., Wang et al. 2014, Chen et al. 2015). It can be easily shown that Assumptions A1 - A3 are satisfied by a broad class of demand functions such as linear, exponential, power and logit demand models. Note that we do *not* assume that the revenue rate is concave when viewed as a function of demand rate instead of price, which is a critical assumption in most existing studies on dynamic pricing. Instead, we simply require the revenue function to be strictly unimodal. As will be discussed in Chapter 2.5, we are able to sidestep the concavity assumption by a novel deterministic formulation of the original stochastic problem.

After a customer in location j makes a purchase of product k, the e-tailer chooses an FC i from which the order should be fulfilled immediately. In this paper, we do not allow any deliberate delay in shipment for further savings in cost, since it is in itself a complex research problem and beyond the scope of this work (see Xu et al. 2009 for further discussions on the same assumption). The shipping cost of product k from FC i to location j equals $c_{ijk} \geq 0$. Let $X_{ijk}^t \in \{0, 1\}$ denote the e-tailer's decision to fulfill an incoming order for product k from location j in period t using the inventory available at FC i. We assume that FC i carries $\boldsymbol{C}_i = (C_{ik}) \succeq \mathbf{1}$ units of initial inventory before the selling season starts and no replenishment occurs during the selling season. (We use $\mathbf{1}$ to denote a column vector with proper dimension whose entries are all ones, and $a \succeq b$ to denote $a_i \geq b_i$ for all i for any vectors a, b with the same dimension). The assumption on no replenishment opportunity is commonly made in the previous works on dynamic fulfillment optimization (e.g., Xu et al. 2009, Acimovic and Graves 2014, and Jasin and Sinha 2015). The justifications are as follows: (1) we can interpret our selling season as the time window between two replenishments and we focus on the tactical instead of strategic decisions; and, (2) the impact of stockout can be accounted for as explained shortly.

We define a *fictitious* FC 0 that has an infinite amount of initial inventory (i.e., $\boldsymbol{C}_0 = +\infty \cdot \mathbf{1}$), and shipping costs set by us at $c_{0jk} := \max\{2 \max_{i \in [I]} c_{ijk}, p_u\}$ for all j, k. The formulation of FC 0 serves the purpose of a backup facility when certain product is depleted at all real FCs, and technically guarantees that there is always a feasible solution to our problem. In practice, the e-tailer may also decide to simply announce that the product is unavailable when it is depleted at all real FCs; in this case, the cost of shipping from FC 0 can be interpreted as the cost of lost sales. It should be noted that our analysis does *not* depend on the specific cost of shipping from FC 0. For the purpose of this work, we simply set the cost to be no smaller than both the maximum revenue of a single product and all the other fulfillment options to emphasize the undesirability of fulfilling from FC 0.

In addition to having to make the pricing and fulfillment decisions, the e-tailer also needs to satisfy several constraints. First, any arriving order in period t must be fulfilled in the same period (i.e., no backorder or strategically delayed shipment) by a unit of inventory at a certain FC. Second, the number of orders for any given product that each FC can fulfill throughout the selling season cannot exceed the initial inventory level for that product at that FC. The e-tailer's objective is to maximize her total expected profits, which is defined as total expected revenues minus total expected fulfillment costs. We can write the optimal control formulation of JPF problem as follows:

$$
\begin{aligned}
\mathcal{J}^* \quad := \max_{\{\boldsymbol{p}^{t,\pi},\boldsymbol{X}^{t,\pi}\}\in\Pi} \quad & \mathbb{E}^{\pi}\left[\sum_{t=1}^{T}\sum_{j=1}^{J}\left(\boldsymbol{p}^{t,\pi}\right)^{\top}\boldsymbol{D}_j^t(\boldsymbol{p}^{t,\pi}) - \sum_{t=1}^{T}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}X_{ijk}^{t,\pi}\right] \\
\text{s.t.} \quad & \sum_{i=0}^{I} X_{ijk}^{t,\pi} = D_{jk}^t\left(\boldsymbol{p}^{t,\pi}\right), \quad \forall j,k,t && (2.1) \\
& \sum_{t=1}^{T}\sum_{j=1}^{J} X_{ijk}^{t,\pi} \le C_{ik}, \quad \forall i,k && (2.2) \\
& \boldsymbol{p}^{t,\pi} \in \Omega_p,\ X_{ijk}^{t,\pi} \in \{0,1\}, \quad \forall i,j,k,t && (2.3)
\end{aligned}
$$

where Π is the set of all non-anticipating controls and the constraints must hold almost surely. We denote by π^* the optimal control for **JPF**.

Remark 2.4.1 *In practice, e-tailers may offer different options for delivery speed. Our modeling framework is sufficiently general to cover this extra layer of complexity. Specifically, the requests of different shipping options can be modeled as demand nodes at the same demand location with adjusted cost capturing different nominal fees (e.g., {Fast, Oregon, Grill pan} and {Slow, Oregon, Grill pan}). Similarly, different supply nodes should be added at the same FCs with different shipping costs (e.g., {Fast, California, Grill pan} and {Slow, California, Grill pan}), and constraints limiting total consumption from supply nodes representing the same product at the same location should also be added (e.g., total grill pan fulfilled from California under both Fast and Slow shipping options cannot exceed the number of grill pans stored in FC California). All of our results can be easily generalized to the case of multiple shipping options.*

2.5 A Deterministic Approximation of JPF

In practice, the magnitude of demand intensity faced by an e-tailer is often high, especially during holiday and promotion seasons. (According to CNN 2015, Amazon.com sold 398 items per second during its global shopping event exclusively for Amazon Prime members on July 15, 2015.) This translates into the need for e-tailers to make fast real-time pricing and fulfillment decisions. This requirement, together with the well-known curse of dimensionality of dynamic programming, makes solving JPF optimally practically infeasible.

In the RM literature where a similar problem is encountered, many researchers turn their attention to developing heuristic controls that are both easy to implement and have a provably good performance under well-defined metrics. A popular framework is to first propose an approximate formulation of the original stochastic control problem, and then use its optimal solution as a heuristic control. A good approximate formulation usually has three characteristics: (1) its optimal solution is much easier to solve than that of the original stochastic control; (2) its optimal solution is easily implemented as an intuitive heuristic control that can be viewed as a simple approximation of the optimal control; and (3) its optimal objective value is not too much smaller than the optimal value of the original stochastic control problem (since the performance of the derived heuristic control tends to mimic the objective value of the approximate formulation). In what follows, we first discuss why an approximation scheme commonly used in the operations literature may not be appropriate for JPF. This motivates us to propose a novel approximation scheme based on the idea of price randomization.

Classic Certainty Equivalent Approximation. In the broad dynamic optimization literature, Certainty Equivalent (CE) approximation refers to the idea where random variables in the original stochastic problem are replaced by their expected values. Under the classic RM models, CE approximation has all the aforementioned characteristics and has been used to develop several high-performing heuristic controls; see e.g. Gallego and van Ryzin (1994, 1997), Ciocan and Farias (2012), Jasin (2014). In the JPF problem, applying the CE principle leads to the following deterministic formulation, which we call *Deterministic JPF* (DJPF):

$$
\begin{aligned}
\mathcal{J}^D \quad := \max_{\{\boldsymbol{p}^t,\boldsymbol{x}^t\}} \quad & \sum_{t=1}^{T}\sum_{j=1}^{J} r_j(\boldsymbol{p}^t) - \sum_{t=1}^{T}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk} x_{ijk}^t \\
\text{s.t.} \quad & \sum_{i=0}^{I} x_{ijk}^t = \lambda_{jk}(\boldsymbol{p}^t), \quad \forall j,k,t && (2.4)\\
& \sum_{t=1}^{T}\sum_{j=1}^{J} x_{ijk}^t \le C_{ik}, \quad \forall i,k && (2.5)\\
& \boldsymbol{p}^t \in \Omega_p,\ x_{ij}^t \in [0,1] && (2.6)
\end{aligned}
$$

Observe that the optimal solution of DJPF can be easily implemented: $\boldsymbol{p}^t$ can be used as the posted price vector in period t and $x_{ijk}^t/\lambda_{jk}(\boldsymbol{p}^t)$ can be used as the probability of fulfilling an order of product k from location j in period t using inventory in FC i. However, there are two serious drawbacks of the DJPF formulation. First, despite being a deterministic optimization problem, DJPF still has non-linear constraints and a potentially non-concave objective function, which means that it may not be easy (or time-efficient) to solve (see Table A.2. for a numerical example). Second, if the demand function is non-linear in price, it is possible that $\mathcal{J}^* > \mathcal{J}^D$ and, depending on the problem parameters, the gap can be quite large. This implies that the performance of a heuristic control derived directly from the solution of DJPF, as it is intended to mimic $\mathcal{J}^D$, may perform a lot worse than $\mathcal{J}^*$ (see Chapter 2.8 for numerical examples that confirms this conjecture). This is in sharp contrast with the standard RM models, where CE approximation serves as an upper bound of the optimal value of the original stochastic problem under a general class of non-linear demand functions (see Remark 2.5.1 for a discussion on the intuition).

Motivated by the preceding discussions, in this paper, we will use an alternative deterministic formulation based on the idea of price discretization. We will show that it is possible to construct a deterministic optimization problem whose optimal value is at most $\epsilon > 0$ smaller than $\mathcal{J}^*$ for any value of ϵ. We will use this alternative deterministic formulation to construct our heuristic controls. (Note that our approach in this paper can also be used in combination with DJPF if the e-tailer uses the DJPF formulation.)

An Approximate Linear Program. Our new formulation shares similarities with DJPF and other CE approximations in the literature, in that it also replaces all the random variables by their expected values. However, in the new formulation, the pricing decision is approximated by a randomization over a set of discrete prices instead of by a singleton. Formally, let $\mathcal{Q} := (\boldsymbol{q}_m)_{m=1}^M$ denote a set of M price vectors ($\boldsymbol{q}_m \in \Omega_p$) and $\boldsymbol{\alpha}^t = (\alpha_1^t, \ldots, \alpha_M^t)$ denote a weight vector whose entries are all non-negative and sum up to one. For a fixed discretization set $\mathcal{Q}$, we can define the following

Approximate Linear Program (ALP):

$$\begin{aligned}
\mathcal{J}^{ALP}(\mathcal{Q}) := \quad \max_{\{\boldsymbol{\alpha}^t, \boldsymbol{x}^t\}} \quad & \sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^t r_j(\boldsymbol{q}_m) - \sum_{t=1}^{T}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk} x_{ijk}^t \\
\text{s.t.} \quad & \sum_{i=0}^{I} x_{ijk}^t = \sum_{m=1}^{M} \alpha_m^t \lambda_{jk}(\boldsymbol{q}_m), \quad \forall j,k,t && (2.7)\\
& \sum_{t=1}^{T}\sum_{j=1}^{J} x_{ijk}^t \le C_{ik}, \quad \forall i,k && (2.8)\\
& 0 \le x_{ijk}^t \le 1, \quad \forall i,j,k,t && (2.9)\\
& \sum_{m=1}^{M} \alpha_m^t = 1, \quad \alpha_m^t \ge 0, \forall m,t && (2.10)
\end{aligned}$$

There are several nice features about the above ALP formulation. First, since FC 0 has infinite inventory, ALP always has a solution. Similar to DJPF, the solution of ALP can be easily implemented as an intuitive heuristic control, which is formally studied in Chapter 2.6. Second, if we include the optimal prices from the solution of DJPF in $\mathcal{Q}$, then $\mathcal{J}^D \le \mathcal{J}^{ALP}$ since the optimal solution to DJPF is also feasible for ALP. Thus, one can view ALP as a generalization of DJPF that allows the price vector to be sampled from a multi-point distribution instead of a singleton. The randomization over different price points brings additional benefit in increasing the expected profit. Third, since ALP is an LP and demand rates are stationary, there exists a stationary optimal solution satisfying $x_{ijk}^t = x_{ijk}^1$ and $\alpha_m^t = \alpha_m^1$ for all t. (Let $\{x_{ijk}^t, \alpha_m^t\}_{t=1}^T$ denote a pair of optimal solution of ALP. Define: $\bar{x}_{ijk}^t = \sum_{s=1}^T x_{ijk}^s/T$ and $\bar{\alpha}_m^t = \sum_{s=1}^T \alpha_m^s/T$. It is not difficult to check that $\{\bar{x}_{ijk}^t, \bar{\alpha}_m^t\}_{t=1}^T$ is also optimal for ALP.) Without loss of generality, throughout this paper we will be working with a stationary optimal solution of ALP, which is simply denoted as $\boldsymbol{x}^* := (x_{ijk}^*)$ and $\boldsymbol{\alpha}^* := (\alpha_m^*)$. We will also assume that $\alpha_m^* > 0$ for all $m \in [M]$, since if $\alpha_m = 0$ for some m, we can simply delete those $\boldsymbol{q}_m$ from the set $\mathcal{Q}$ without affecting any other α_m^* and x_{ijk}^*.

Note that the value of $\mathcal{J}^{ALP}$, $\boldsymbol{\alpha}^*$, and x^* depend on price discretization $\mathcal{Q}$. We neglect this dependency for notational simplicity. The following lemma tells us that, there exists a set of discrete price vectors $\mathcal{Q}$ such that ALP approximates JPF well (in some sense).

Lemma 2.5.1 *Assume that assumptions A1 and A2 hold. For any $\epsilon > 0$, there exists a discretization $\mathcal{Q}$ such that $\mathcal{J}^* - \mathcal{J}^{ALP} \le \epsilon$.*

In proving Lemma 2.5.1, we use a specific set of dicrete price vectors that forms a uniform grid on Ω_p and show that it satisfies the above approximation guarantee. Formally, we first divide the feasible set $[p_\ell, p_u]$ into $\lfloor m \rfloor$ sub-intervals of equal length and let $\bar{\mathcal{Q}}^u$ be the set of mid-points of the resulting sub-intervals. We then define our *uniform grid* as $\mathcal{Q}^u = \{(p_1, \ldots, p_K) \in \Omega_p : p_k \in \bar{\mathcal{Q}}^u \ \forall k \in [K]\}$. In the proof of Lemma 2.5.1, we show that to reach an ϵ-approximation stated in Lemma 2.5.1, the number of uniform grid points is at most $M = m^K = O\left(\left(\frac{IJKT}{\epsilon}\right)^K\right)$. (If demands are independent, this number reduces to $O\left(\frac{IJK^2T}{\epsilon}\right)$ since we only need to approximate K univariate functions.) volume of the smallest hyper-cubes sliced by the uniform grid.) Although this number can be large for problems of practical size, our numerical experiment suggests that it is not necessary to use too many price points to guarantee a good approximation (see Chapter 2.8 and our discussions below). Moreover, in practice, e-tailers often already work with a pre-determined price set (see Chapter 5.2.1.3 in Talluri and van Ryzin 2006 and Cohen et al. 2017a). In this context, Lemma 2.5.1 can be seen as providing a theoretical justification that this type of approximation (i.e., using price discretization) provides a good approximation of JPF, at least for a sufficiently fine discretization. (This is in contrast to DJPF, which can be a very inaccurate approximation of JPF.) Although our proposed heuristic controls can be applied with any price discretization $\mathcal{Q}$, in the remaining of this paper we will always use the set of uniform grids discussed above for consistency.

We want to underscore that, if the number of products is very large, it may not be possible to solve the corresponding ALP. In practice, this challenge can be resolved by first segmenting the products into clusters within which demands are strongly correlated and then applying our approach to each segment separately. The question of how to properly disaggregate products into clusters in a way that balances the trade-off between computational complexity and approximation quality is an important one; however, it is beyond the scope of this paper and we leave it for future research pursuit.

Remark 2.5.1 *In the typical RM literature, under the standard assumptions that demand rate is invertible in price and revenue is concave in demand rate, the CE-type formulation can be transformed into a concave optimization problem by using demand rate instead of price as the decision variable. In JPF problem, for any period, the price vectors observed by customers in all locations are the same, which results in new non-linear constraints that cannot be easily transformed into deterministic constraints by standard techniques. If demand rates are linear in prices, then DJPF is indeed a proper deterministic relaxation of JPF since it can be shown $\mathcal{J}^* \leq \mathcal{J}^D$. In this case, we can use DJPF as our deterministic relaxation and the ALP is not needed.*

Performance Measure and Asymptotic Regime. In this paper, we use the optimal value of ALP as the benchmark to evaluate the theoretical performance of our heuristic controls. Motivated by the typical large volume of sales faced by e-tailers, and for the purpose of theoretical performance analysis, we will consider a sequence of JPFs and ALPs where both the length of selling season and the amount of initial inventories are scaled proportionally by a factor of θ while keeping all the other parameters unchanged. More specifically, in the θ^{th} problem, the length of selling season is given by $T(\theta) = \theta T$ and the amount of initial inventories in FC i is given by $\boldsymbol{C}_i(\theta) = \theta \boldsymbol{C}_i$. Since we only allow at most one new arrival in each period, increasing the selling season by θ is equivalent to multiplying the number of potential demands by θ. So, in the prescribed *asymptotic* setting, we essentially scale both the potential demands and initial inventories proportionally. Naturally, we shall interpret the scaling parameter θ as the *size* of the problem.

For a problem with size θ, let $\mathcal{J}^{\pi}(\theta)$ denote the total expected profits collected under a specific heuristic control $\pi \in \Pi$. Similarly, let $\mathcal{J}^{ALP}(\theta)$ denote the optimal value of ALP with size θ. We use the *loss* of heuristic control π, defined as $\mathcal{L}^{\pi}(\theta) := \mathcal{J}^{ALP}(\theta) - \mathcal{J}^{\pi}(\theta)$, as our performance measure. (Again, for notational simplicity, we neglect the notational dependency of $\mathcal{L}^{\pi}(\theta)$ on $\mathcal{Q}$.) By definition, the loss of any control captures the difference in profit between the optimal control and that control. A control whose loss scales sublinearly in θ is *asymptotically optimal*. It is noteworthy that although there is no theoretical guarantee that an asymptotically optimal heuristic control will also perform well in non-asymptotic settings, existing works in the literature have found that they tend to also perform sufficiently well, if not extremely well, in non-asymptotic settings (see e.g. Ciocan and Farias 2012, Jasin 2014). In our case, we also observe sufficiently good performance for both of our heuristic controls in the non-asymptotic setting (see Chapter 2.8).

2.6 First Heuristic Control: Randomizing Pricing and Fulfillment Decisions

In this Chapter, we describe a simple non-adaptive heuristic control and discuss its asymptotic performance. Let $\sigma_k^t : [J] \to [I] \cup \{0\}$ denote the fulfillment assignment for period t, i.e., $\sigma_k^t(j) = i$ indicates that we fulfill an order of product k from location j in period t from FC i. Our first heuristic control directly uses the solution of ALP to construct a randomized heuristic. Note that, for a fixed set of discrete price vectors

$\mathcal{Q}$, $\boldsymbol{\alpha}^*$ and $\boldsymbol{x}^*$ are the optimal sampling vector and fulfillment vector given by ALP. The idea behind our first heuristic control is to sample a price vector $\boldsymbol{p}^t$ from $\mathcal{Q}^u$ according to $\boldsymbol{\alpha}^*$, and sample the fulfillment assignment σ^t according to $\boldsymbol{x}^*$. Let $\boldsymbol{C}_i^t$ denote the inventory level in FC i at the beginning of period t. We formally define our first heuristic control below.

Randomized Pricing and Fulfillment Heuristic (RPF)

1. **Initialization**: Fix a discretization $\mathcal{Q}$ and solve ALP to get $\boldsymbol{\alpha}^*$, $\boldsymbol{x}^*$.
2. During period $t \geq 1$, do:
   a. Sample $\boldsymbol{p}^t = \boldsymbol{q}_m$ with probability $\mathbb{P}\{\boldsymbol{p}^t = \boldsymbol{q}_m\} = \alpha_m^*$ and apply $\boldsymbol{p}^t$.
   b. Sample $\sigma_k^t(j)$ with probability $\mathbb{P}\{\sigma_k^t(j) = i\} = y_{ijk}^* := x_{ijk}^* / \sum_{i=0}^I x_{ijk}^*$.
   c. If there exists a $(j,k) \in [J] \times [K]$ such that $D_{jk}^t = 1$, do:
      i. If $C^t_{\sigma_k^t(j),k} > 0$, fulfill the order from FC $\sigma_k^t(j)$ and update $C^{t+1}_{\sigma_k^t(j),k} = C^t_{\sigma_k^t(j),k} - 1$;
      ii. Otherwise, fulfill the order from FC 0.

The following theorem characterizes the performance of the RPF.

Theorem 2.6.1 *Let $\mathcal{Q}^u$ be the uniform price grids discussed in Chapter 2.5. There exists a constant $\Psi_1 > 0$ independent of $\theta \geq 1$ such that $\mathcal{L}^{RPF}(\theta) \leq \Psi_1\sqrt{\theta}$.*

Two comments are in order. First, it is not difficult to show that $\mathcal{J}^{ALP}(\theta)$ is an upper bound for total expected profits under *any* feasible joint pricing and fulfillment control that restricts $\boldsymbol{p}^t \in \mathcal{Q}$ for all t and some $\mathcal{Q}$, and that the above bound is tight, i.e., for some problem instances, there exists a constant $\Psi_1' > 0$ independent of $\theta \geq 1$ such that $\mathcal{L}^{RPF}(\theta) \geq \Psi_1'\sqrt{\theta}$ (see Remark 2 in Jasin 2014 for an argument for a simple example where $I = J = K = 1$). This means that Theorem 2.6.1 completely characterizes the asymptotic performance of RPF. (The constant in Theorem 2.6.1 scales linearly in I, J and K. However, this is not surprising as $\mathcal{J}^{ALP}$ itself also scales linearly in I, J and K.) Second, although RPF is asymptotically optimal, a heuristic control that has a stronger performance guarantee than $\sqrt{\theta}$ is still highly desirable. Since RPF does not adjust its decisions dynamically depending on the realized observations, it may lose significant opportunities to boost total profits. The important question is how to construct a heuristic control that both significantly improves the performance guarantee of RPF while maintaining its tractability. One simple idea is to re-optimize ALP at the beginning of every period by updating its inventory parameters. Unfortunately, this approach may not be feasible in practice since ALP can be very large (for example, a 5-point discretization for each product for a catalog of ten products results in $5^{10} \approx 10^7$ price points). Therefore, we will not focus on the heuristic control that fully re-optimizes ALP. Instead, in the next Chapter, we will develop a novel re-adjust-and-re-optimize heuristic control based on the idea of combining real-time price adjustment with re-optimization of only the fulfillment part of ALP.

Remark 2.6.1 *Since RPF samples fulfillment assignment decisions randomly over a static distribution, it is possible that, at some point of the selling season, the assigned FC has zero inventory whereas other FCs have positive inventory. In other words, RPF may randomly deny demand although there is still inventory for the requested product in some of the FCs. In practice, a simple way to fix this is problem is to re-optimize the fulfillment part of the ALP (see (2.12) for a formal definition) whenever such event happens.*

2.7 Second Heuristic Control: Re-adjust and Re-optimize Pricing and Fulfillment Decisions

Our second heuristic control adaptively adjusts the discretization set $\mathcal{Q}$ and re-optimizes the fulfillment vector $\boldsymbol{x}$ in every period. An important feature of this modification is although both prices and fulfillment probabilities are still decided jointly at the beginning of the selling season via solving ALP, their updates during the selling season are done almost separately through a two-stage process. We show in this chapter that, under some conditions, our proposed modification guarantees a significant improvement over RPF.

We start by introducing a few more notations. For every period t, we let $\mathcal{Q}^t = (\boldsymbol{q}_m^t)$ be the set from which $\boldsymbol{p}^t$ is sampled, and $\boldsymbol{x}^t = (x_{ijk}^t)$ be the fulfillment vector. Given $\boldsymbol{x}^t$, define $y_{ijk}^t := x_{ijk}^t / \sum_{i=0}^{I} x_{ijk}^t$ to be the conditional probability of using FC i to fulfill an order of product k from location j conditioning on the arrival of such order. Let $\boldsymbol{C}^t := (\boldsymbol{C}_i^t)$ denote the vector of remaining inventory level at the beginning of period t. Recall from Chapter 2.4 that X_{ijk}^t is the actual fulfillment decision in period t. Let $\Delta C_{ik}^t := \sum_{j=1}^{J} [X_{ijk}^t - y_{ijk}^t (\sum_{m=1}^{M} \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^t))]$ denote the difference between the actual consumption of inventory of product k at FC i during period t and the expected consumption prescribed by the current control parameters (i.e., y^t_{ijk} and $\boldsymbol{q}^t_m$). (We suppress notational dependencies of ΔC^t_{ik} on $\boldsymbol{p}^t$ and $\mathcal{Q}^t$ for the sake of brevity.) Let $\Delta \boldsymbol{C}^t_i = [\Delta C^t_{i1}; \ldots; \Delta C^t_{iK}]^\top$. Define $\mathrm{proj}_A(x) := \arg\min_{y \in A} ||y - x||_2$ to be the Euclidean projection function. We are now ready to present our second heuristic control.

Re-adjust and Re-optimize Pricing and Fulfillment Heuristic (R^{2}PF)

1. **Initialization**: Fix discretization $\mathcal{Q}$ and solve ALP to get $\boldsymbol{\alpha}^*$, $\boldsymbol{x}^*$.
   Define $\mathcal{Q}^1 = \mathcal{Q}$ and $\hat{\boldsymbol{x}}^1 = \boldsymbol{x}^*$.

2. During period $t \geq 1$, do:

   a. **Adjust Price:** For each m, calculate $\boldsymbol{q}^t_m$ satisfying

$$\boldsymbol{\lambda}^{tot}\left(\boldsymbol{q}^t_m\right) = \operatorname*{proj}_{[0,1]^K}\left[\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m) - \frac{1}{M\alpha_m}\left(\sum_{i=0}^{I}\sum_{s=1}^{t-1}\frac{\Delta \boldsymbol{C}^s_i}{T-s}\right)\right]. \qquad (2.11)$$

   b. **Update Fulfillment:** Set $\hat{\mathbf{x}}^{t+1}$ equal to the optimal solution of the following *Fulfillment LP* (FLP):

$$\mathbf{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t) :=$$
$$\left\{\min_{x_{ijk}\geq 0} \mathbf{c}^\top \mathbf{x} : \sum_{i=0}^{I} x_{ijk} = \sum_{m=1}^{M} \alpha^*_m \lambda_{jk}\left(\boldsymbol{q}^t_m\right),\ \sum_{j=1}^{J} x_{ijk} \leq \frac{C^t_{ik}}{T-t+1}\right\}. \qquad (2.12)$$

   c. Sample $\boldsymbol{p}^t$ with probability $\mathbb{P}\{\boldsymbol{p}^t = \boldsymbol{q}^t_m\} = \alpha^*_m$ and apply $\boldsymbol{p}^t$.

   d. Sample $\sigma^t_k(j)$ with probability $\mathbb{P}\{\sigma^t_k(j) = i\} = y^t_{ijk} := \hat{x}^t_{ijk} / \sum_{i=0}^{I} \hat{x}^t_{ijk}$.

   e. If there exists a $(j,k) \in [J] \times [K]$ such that $D^t_{jk} = 1$, do:

      i. If $C^t_{\sigma^t_k(j),k} > 0$, fulfill the order from FC $\sigma^t_k(j)$ and update $C^{t+1}_{\sigma^t_k(j),k} = C^t_{\sigma^t_k(j),k} - 1$;

      ii. Otherwise, fulfill the order from FC 0.

Recall that ΔC^s_{ik} is the error from the expected consumption of product k in FC i at period s. In designing R^{2}PF, we wish to eliminate as much of these errors as possible such that, by the end of the selling season, the performance of R^{2}PF is very close to the deterministic benchmark $\mathcal{J}^{ALP}$. This is accomplished in two steps: (1) We first adjust the discretization set $\mathcal{Q}$ such that the new aggregate expected demands equal the original aggregate expected demands given by ALP minus a linear combination of inventory consumption errors caused by randomness up to period t; (2) we then update the fulfillment probabilities by re-optimizing the fulfillment part of ALP, which has a much smaller number of variables compared to the full ALP. In the price adjustment step, under the uniform pricing constraint, we can only precisely control the aggregate expected demands (over all locations) for each product. Therefore, at any period s, we aggregate the incurred consumption error (over all FCs) at the product level and correct them uniformly throughout the remaining periods—this is the intuition behind the term $\sum_{i=1}^{I} \Delta C_{ik}^{s}/(T-s)$. (The uniform error distribution may not be the optimal correction scheme; however, Jasin (2014) has shown in the context of dynamic pricing that it is sufficient to guarantee a very strong performance bound.) Thus, the total errors for product k that needs to be corrected up to period t is given by $\sum_{i=0}^{I}\sum_{s=1}^{t-1} \Delta C_{ik}^{s}/(T-s)$. We then perturb $\mathcal{Q}^u$ to $\mathcal{Q}^t$ such that the new aggregated expected demands for product k equals the original one under $(\boldsymbol{\alpha}^*, \boldsymbol{x}^*)$ minus the perturbation term. Mathematically, we want the following system of equations to hold:

$$\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* \lambda_{jk}\left(\boldsymbol{q}_m^t\right) \;=\; \sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* \lambda_{jk}\left(\boldsymbol{q}_m\right) - \sum_{i=0}^{I}\sum_{s=1}^{t-1} \frac{\Delta C_{ik}^{s}}{T-s}, \quad \forall k. \tag{2.13}$$

One can show that any interior solution to (2.11) is also a solution to (2.13). Moreover, by the invertibility of $\boldsymbol{\lambda}^{tot}(\cdot)$ (Assumption A1), the system in Step 2a always has a unique solution of $\mathcal{Q}^t$. Although we need to perturb potentially all price vectors in $\mathcal{Q}$, the computation in Step 2a can be done for each price vector in parallel very efficiently (e.g., using standard gradient-based methods). This decomposability is crucial for the time-efficiency of R^{2}PF.

We want to emphasize: Although the price adjustment helps balance future demands with remaining inventories, it only does so at the *aggregate* level across all FCs. To address the potential inventory imbalance across different FCs caused by the randomness in demand and fulfillment assignment, another layer of optimization is needed. To do so, given $\mathcal{Q}^t$ in the price adjustment step, we update the fulfillment vector by re-optimizing FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$. (For notational brevity, we will often write it as FLP^t whenever the values of $\mathcal{Q}^t$ and $\boldsymbol{C}^t$ are clear from the context.) FLP^t essentially solves the optimal static fulfillment decisions for the remaining $T-t+1$ periods, assuming that we always sample price from $\mathcal{Q}^t$. (The deleted constraints $x_{ijk} \le 1$ is redundant, since $x_{ijk} \le \sum_{m=1}^{M} \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^t) \le \sum_{m=1}^{M} \alpha_m^* \cdot 1 \le 1$.) This extra step is crucial for making sure that we are also minimizing total shipping costs while maximizing total revenues.

Before we evaluate the asymptotic performance of R^{2}PF, we need to first introduce a concept that will be useful for the analysis. Consider the initial transportation problem faced by the e-tailer, i.e., FLP^1. Since we assume that each customer only requests at most one product, FLP^1 can be decomposed into K transportation LPs defined as

$$\begin{aligned}&\mathrm{FLP}^1_k(\mathcal{Q}, \boldsymbol{C}_k)\\ &:= \left\{ \min_{x_{ijk}\ge 0} \sum_{i=0}^{I}\sum_{j=1}^{J} c_{ijk}x_{ijk} \;:\; \sum_{i=0}^{I} x_{ijk} = \sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^1_m),\ \sum_{j=1}^{J} x_{ijk} \le C_{ik}/T \right\}.\end{aligned}$$

We assume without loss of generality that $\sum_{j=1}^{J} x^*_{ijk} = C_{ik}$ (this is for the simplicity of the proof; otherwise, we can always define $\tilde{C}_{ik} := \sum_{j=1}^{J} x^*_{ijk}$ and replace the original initial inventory C_{ik} with $\tilde{C}_{ik}$ without changing anything else). In other words, the inventory constraints in FLP^1_k are all binding. From the study of transportation LP (e.g., Dantzig and Thapa 2006), we know that there is exactly one redundant constraint in every FLP^1_k. Moreover, if we delete an arbitrary constraint, the remaining constraints are always linearly independent. Let $\overline{\mathrm{FLP}}^1_k$ be the LP where we delete the inventory constraint regarding FC 0; since the deleted constraint is redundant, $\overline{\mathrm{FLP}}^1_k$ is equivalent to FLP^1_k. We call a basic solution to FLP^1 as *DR-degenerate* ("DR" is short for de-redundancy) if and only if the corresponding basic solution to $\overline{\mathrm{FLP}}^1_k$ is degenerate for some $k \in [K]$.

We state a theorem on the performance of R^{2}PF.

Theorem 2.7.1 *Let $\mathcal{Q}^u$ be the uniform price grids discussed in Chapter 2.4. Suppose that FLP^1($\mathcal{Q}^u, \boldsymbol{C}$) has a unique non-DR-degenerate optimal solution. There exists a constant $\Psi_2 > 0$ independent of $\theta \ge 1$ such that $\mathcal{L}^{R^2PF}(\theta) \;\le\; \Psi_2(1+\log\theta)$.*

Some comments are in order. First, since R^{2}PF may use different discretization sets in different periods, $\mathcal{J}^{ALP}(\theta)$ is not necessarily an upper bound for $\mathcal{J}^{R^2PF}(\theta)$; in other words, $\mathcal{L}^{R^2PF}$ can actually be negative. However, given that the expected loss of RPF relative to $\mathcal{J}^{ALP}(\theta)$ is of order $\sqrt{\theta}$, the bound in Theorem 2.7.1 is useful because it shows that R^{2}PF guarantees a significant improvement over RPF, at least asymptotically. (The constant in Theorem 2.7.1 scales linearly in J and K, but quadratically in I. In practice, due to its high installment cost, the number of FCs e-tailers own is usually much smaller than the stock level of products. Our numerical results in Chapter 2.8 show that the impact of I on revenue loss is dominated by the revenue improvement due to using R^2PF over RPF.) Second, the non-DR-degeneracy assumption only applies to the initial FLP^1 and is not required for the subsequent FLP^t for all $t \geq 2$. Similar conditions have been used in other works that study the performance of re-optimization-based controls with deterministic relaxation being an LP, e.g., Jasin and Kumar (2012, 2013), and Ferreira et al. (2015b). Although this assumption is critical for the tractability of the proof, our numerical results in Chapter 2.8 show that R^2PF still performs well even when all $\overline{\mathrm{FLP}}_k^1$'s are degenerate. Finally, the fact that R^2PF significantly improves RPF is not a trivial result. Although it is known in the literature that frequent re-optimization has the potential to significantly improve performance (see Chapter 2.3), it matters what is being re-optimized. In the case of R^2PF, the FLP^t takes as its input the perturbed $\mathcal{Q}^t$ that is chosen almost independently of the current inventory distribution and how it would affect total shipping costs (except the perturbations terms $\{\Delta C_{ik}^t\}$). It is, thus, not immediately clear that frequent re-optimizations of the fulfillment LP updated in this manner still yields the level of improvement that we want. In order to analyze R^2PF, we introduce a key concept of *balanced* FLP^t (see Step 1, Chapter A.3), meaning that the aggregate demand under the current price equals to the aggregate inventory across all the FCs. We show that joint optimization guarantees that FLP^t stays balanced during most of the selling season. This observation allows us to express the evolution of demand and inventory consumption levels at all FCs in closed form, which is instrumental to the following proof.

Remark 2.7.1 *Both RPF and R^2PF involve frequent randomizations in pricing decision. As discussed in Chapter 2.5, this is critical to help us overcome the difficulty in JPF caused by the requirement of uniform pricing. In practice, in order to avoid an adverse impact of randomization, it is recommended that the range of price points within the discretization set is limited to a reasonable interval. This is crucial for making sure that randomization does not result in a scenario where some customers are charged extreme prices.*

2.8 Numerical Experiment

Experiment Setup. We now conduct numerical simulation to illustrate the performance of the proposed heuristic controls in comparison to some natural benchmarks. We choose $I = 6$, $J = 15$ and $K = 5$ (i.e., the e-tailer sells five different products to fifteen different demand locations through six FCs) and select our fifteen demand locations to be the fifteen largest metropolitan statistical areas (MSAs) in the United States estimated by U.S. Census Bureau (2014a). The logistic network consists of six FCs selected from the list of the most efficient warehouses (in terms of transit lead-times) in the United States (Chicago Consulting, 2013) and spans the contiguous United States.

The demand process is determined by a two-step procedure: we first generate arrivals from all fifteen locations according to independent Poisson processes whose rates are proportional to total populations of the corresponding MSAs. We then set the purchase probability of an arriving customer according to an exponential function in price. The parameters of the purchase probability functions are set such that customers from locations with higher income are more likely to make a purchase compared to customers from locations with lower income. We set the feasible price range to be \$100 and \$250. The outbound shipping costs are set to be proportional to the distance between the demand location and the FC. The average shipping cost over all FC-MSA pair is \$9.55. (Since the annual outbound transportation costs as a percentage of net sales typically varies between 4% to 10%, our choice at least guarantees that the relative magnitude between revenue and cost is practical; see Tompkins Supply Chain Consortium 2012.) The costs of the fictitious FC, per Chapter 2.4, are calculated as $c_{0jk} := \max\{2\max_{i\in[I]} c_{ijk}, p_u\} = \$250$ for all j, k. We set the initial inventory levels to be balanced across FCs, taking into account for the market sizes of MSAs and the distances between all FC-MSA pairs. The details of parameters configuration can be found in Chapter A.4.

Implemented Heuristics. We now list all heuristic controls that are tested in the experiment. For any heuristic control $\textsc{Alg}$ that is motivated by ALP formulation, we denote by $\textsc{Alg}$-m the one that uses uniform price grid $\mathcal{Q}^u$ with size of m^K. (We choose $K = 5$ and $m \in \{2, 5, 8\}$, therefore $|\mathcal{Q}^u| \in \{32,\ 3125,\ 32768\}$.)

- RPF-m: RPF heuristic proposed in Chapter 2.6.
- R^2PF-m: R^2PF heuristic proposed in Chapter 2.7.
- ALP-$\textsc{Reopt}$-m: At the beginning of period t, re-optimize ALP by replacing the inventory parameters C_{ik} from the original ALP with C_{ik}^t.
- R^2PF-$\textsc{Ful}$-m: R^2PF-m without price re-adjustment, i.e., $\mathcal{Q}^t \equiv \mathcal{Q}^u$ for all t.
- R^2PF-$\textsc{Pr}$-m: R^2PF-m without fulfillment re-optimization, i.e., $\boldsymbol{x}^t \equiv \boldsymbol{x}^*$ for all t.

- DJPF-Reopt-k: At the beginning of periods $t \in \{1, \lfloor \frac{T}{k} \rfloor, 2 \cdot \lfloor \frac{T}{k} \rfloor, \ldots, (k-1) \cdot \lfloor \frac{T}{k} \rfloor\}$, re-optimize DJPF by updating the inventory parameters from the original DJPF from C_{ik} to C_{ik}^t and apply its solution.

- Sep-Reopt: At the beginning of period t, compute price $\boldsymbol{p}^t$ according to

$$\boldsymbol{p}^t := \{\max_{p \in \Omega} \sum_{j=1}^{J} r_j(\boldsymbol{p}) : \lambda_{jk}(\boldsymbol{p}) \leq \sum_{i=1}^{I} \frac{C_{ik}^t}{T-t+1}\};$$

then compute the fulfillment vector $\boldsymbol{x}^t$ according to

$$\boldsymbol{x}^t := \left\{ \min_{0 \leq x_{ijk} \leq 1} \sum_{i=0}^{I} \sum_{j=1}^{J} c_{ijk} x_{ijk} : \sum_{i=0}^{I} x_{ijk} = \lambda_{jk}(\boldsymbol{p}^t), \sum_{j=1}^{J} x_{ijk} \leq \frac{C_{ik}^t}{T-t+1} \right\}.$$

Several comments are in order. First, despite its long computation time, ALP-reopt is a good benchmark for heuristics based on the ALP formulation. (In fact, ALP-reopt-8 is not implemented since its computation time is too long, see Table A.2.) Second, implementing Sep-reopt allows us to illustrate the benefit of joint joint pricing and fulfillment optimization. To do this, we solve for the optimal price $\boldsymbol{p}^t$ first by aggregating the inventory for each product across all FCs, and then solve the FLP under $\boldsymbol{p}^t$ for the fulfillment assignment distribution $\boldsymbol{x}^t$. Since we re-optimize both decisions at every period, Sep-reopt is in fact a near-optimal heuristic if we are restricted to separate pricing and fulfillment optimization. Third, we test the performance of R^2PF-Ful-m and R^2PF-Pr-m to tease out the benefit of price optimization and fulfillment optimization, respectively. Lastly, the performance DJPF-Reopt-k can help us understand empirically whether the ALP formulation is indeed more beneficial than DJPF in providing a better approximate formulation of the original stochastic control problem.

All heuristic controls are tested under varying problem scales. For simplicity, we normalize T to 1. This means that the scaling factor θ is the same as the length of selling season and can be immediately interpreted as the size of potential market. The value of θ ranges from 200 to 2,000, which means that the average initial inventory level for each product in each FC ranges from 3 units to 30 units. Note that this scale allows us to highlight the performance of our heuristic controls in a non-asymptotic setting. For each θ, we simulate all heuristic controls for 500 runs to approximate their total expected profits. To understand the performance of the heuristics beyond the scenarios prescribed theoretically in previous chapters, we intentionally choose m to

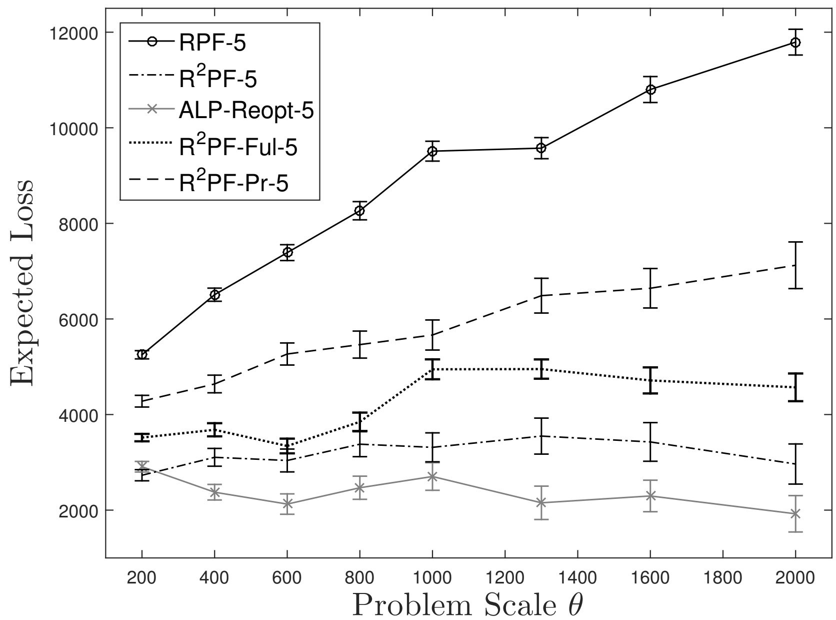


Figure 2.2: Performance of Heuristics Motivated by ALP θ

be small and the initial FLP to be DR-degenerate.

Results and Observations. We now present representative results of our experiments in Figures 2.2 to 2.5 and Table 2.1. The detailed numerical results can be found in Table A.1.

Figure 2.2 shows the expected losses of all heuristic controls that uses the solution to ALP and has a parameter of $m = 5$. (We group the results according to m since the benchmark $\mathcal{J}^{ALP}(\theta)$ depend on the granularity of price discretization.) The upper and lower bars around each instance form a 95% confidence interval. In Figure 2.2, the trends of the curves suggest that the expected losses of RPF and R^2PF grow sublinearly in θ, with R^2PF growing significantly slower; this empirically validates our theoretical results in Theorems 2.6.1 and 2.7.1. Also, the loss of R^2PF is the second smallest overall and is comparable to that of ALP-Reopt, which is the smallest (not surprisingly). This is achieved with a significant reduction in computation time; see Table A.2 for details. The performances of R^2PF-Ful and R^2PF-Pr suggest that, under our choice of parameters, the dynamic fulfillment optimization is more beneficial than the real-time price adjustment. We show later that this may be caused by the randomized nature of R^2PF-Pr and may not be true in general.

Figure 2.3 compares the expected losses of benchmark heuristics motivated by ALP and DJPF respectively. It confirms the conjecture that ALP is indeed a better approx-

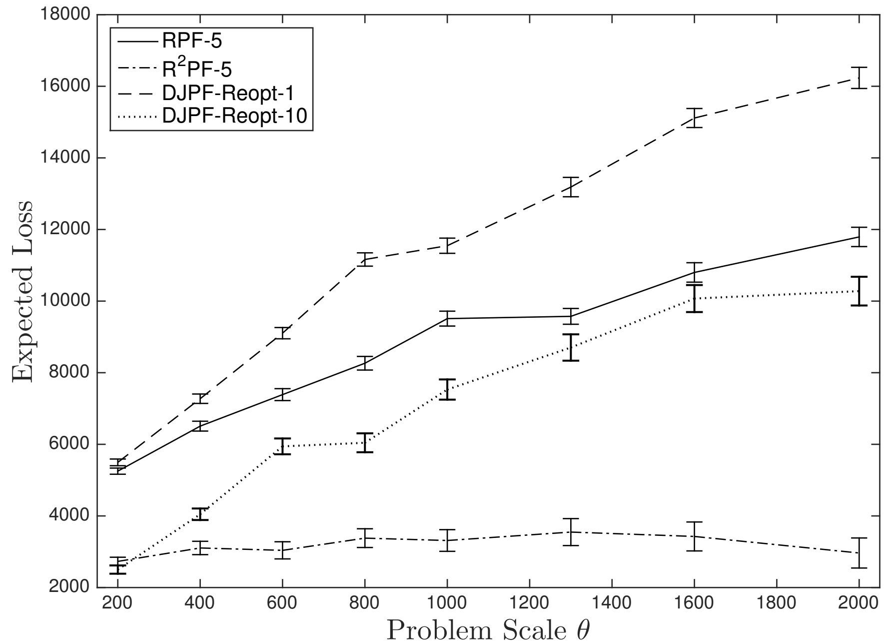


Figure 2.3: Performance of Heuristics Motivated by ALP and DJPF with Varying θ

imate formulation than DJPF since it leads to heuristics with better performances. In particular, we report the performance of the static control under DJPF (i.e. DJPF-Reopt-1, since DJPF is never re-optimized during the selling season) and another control that re-optimizes DJPF ten times throughout the selling season. We do not further increase the re-optimization frequency since frequently re-optimizing DJPF is very time consuming (see Table A.2) and the improvement in performance is only marginal. (Numerical results suggest that the loss of DJPF-Reopt-10 is smaller than that of DJPF-Reopt-1 by around 40%, but further increasing the re-optimization frequency to DJPF-Reopt-20 only brings additional 3% reduction in loss. This is not surprising. In the classic RM setting, it has also been shown that the marginal benefit of re-solving decreases as the frequency increases; see Jasin 2014.) In our simulation, we have $\mathcal{J}^D = 67.96$, and $\mathcal{J}^{ALP} = 69.52$ for $m = 5$. Since DJPF-Reopt and R^{2}PF tries to mimic the DJPF and ALP formulations respectively, it is not surprising that R^{2}PF perform significantly better than DJPF-Reopt.

For all heuristic controls, we also test a variant where, whenever the assigned FC for an incoming demand has no remaining inventory for the requested product, the seller simply denies that demand without incurring any penalty. All previous observations still hold in this new setting, which suggests that they are robust with respect to the change in the value of penalty parameter. To have a better understanding on the

	Profit ($)	Revenue ($)	Fulfillment Cost ($)	Denied Demands #
RPF-5	132561.1	141140.8	8579.7	114.22
R^2PF-5	135431.1	145944.6	10513.5	19.04
ALP-Reopt-5	137918.7	146892.6	8973.9	17.82
R^2PF-Ful-5	134664.4	146420.7	11756.3	22.84
R^2PF-Pr-5	133448.1	142010.8	8562.7	56.06
R^2PF-Pr*-5	134395.2	143932.7	9537.5	29.3

Table 2.1: Detailed Analysis of Performance of Different Heuristics ($\theta = 2000$)

effects of pricing and fulfillment optimization, we calculate the total revenues, total fulfillment costs, and total denied demands due to stockout (i.e., total lost sales). Table 2.1 reports a set of results for a specific problem instance. We see that all the heuristic controls with real-time adjustment significantly decreases the chances of stock-out. As a result, they are able to satisfy demand from more customers, which induces higher revenue and higher fulfillment cost. Moreover, compared to the two heuristic controls that only optimize one set of decision, R^2PF guarantees significantly more earning without sacrificing too much on the fulfillment cost. We also observe in our simulation that more than 50% of the lost sales under R^2PF-Pr happen when there is still some inventories left in some of the unassigned FCs (see Remark 2.6.1 for a discussion on the same issue for RPF). In contrast, R^2PF-Ful does not have this issue, since re-optimizing FLP guarantees that it only samples fulfillment assignment over FCs holding positive inventory. To reduce the number of lost sales, we implement a variant of R^2PF-Pr, denoted by R^2PF-Pr* in Table 2.1, as follows: whenever the assigned FC has zero inventory for the requested product and there are still some inventories left at some other FCs, we simply re-sample the fulfillment assignment decision among the FCs having positive inventory uniformly. Interestingly, this simple modification significantly reduces the number of denied demands and brings the profit of R^2PF-Pr very close to R^2PF-Ful. This example shows that the seller can perhaps couple R^2PF-Pr (or R^2PF-Ful) with a simple adjustment in fulfillment (or pricing) to achieve a better performance. (Although R^2PF-Ful appears to perform better than R^2PF-Pr in Table 2.1, it is not clear that R^2PF-Ful is necessarily superior than R^2PF-Pr since a simple modification to either of them may bring their total profits very close to each other.)

Figures 2.4 and 2.5 shows the absolute percentage improvement in total profits for both RPF and R^2PF relative to the profit of Sep-Reopt for $m \in \{2, 5, 8\}$. From the plots, it is easy to see that even RPF dominates Sep-Reopt. This illustrates the

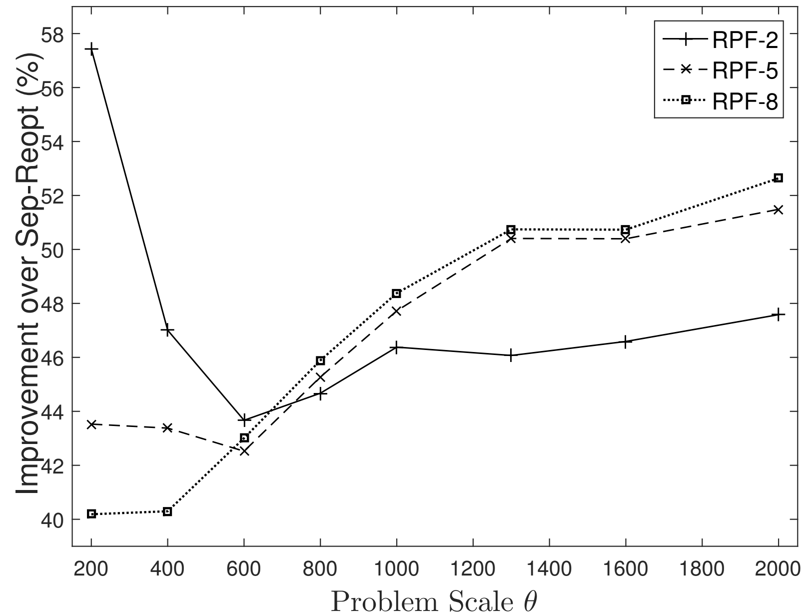


Figure 2.4: Impact of Finer Discretization on the Performance of RPF

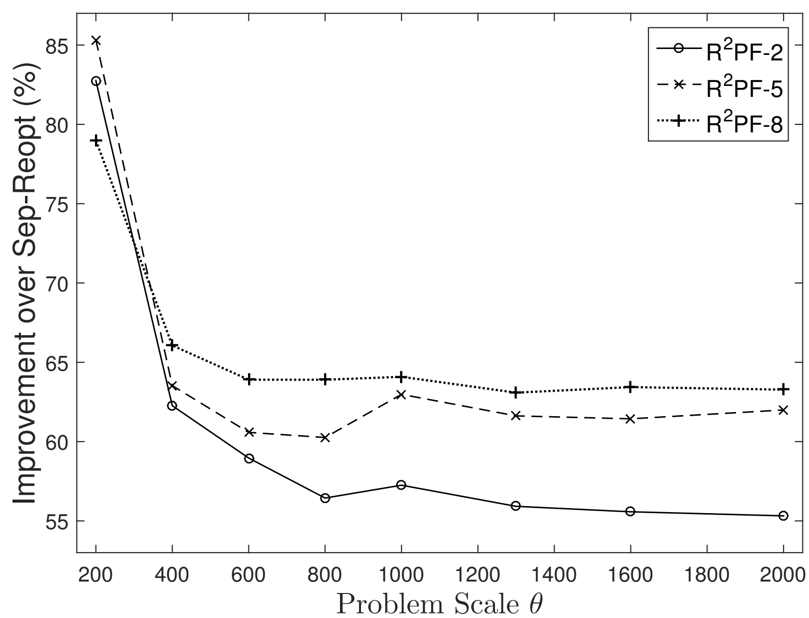


Figure 2.5: Impact of Finer Discretization on the Performance of R^2PF

benefit of joint pricing and fulfillment optimization, even if the e-tailer only does it once before the selling season, with a relatively sparse price discretization. In general, for both heuristic controls, finer discretization (i.e., larger m) leads to a higher profit when θ is large enough. However, the marginal benefit of finer discretization decreases as m increases. In our case, the improvement is large when we increase m from 2 to 5, and much smaller when we further increase m to 8. This is consistent with the value of $\mathcal{J}^{ALP}$ under varying m; see Figure 2.6. (We can see that $\mathcal{J}^{ALP}$ is easily a better approximation than $\mathcal{J}^{D}$ even for sparse discretization.) All of these suggest that the e-tailer may not need to use too many price discretizations.

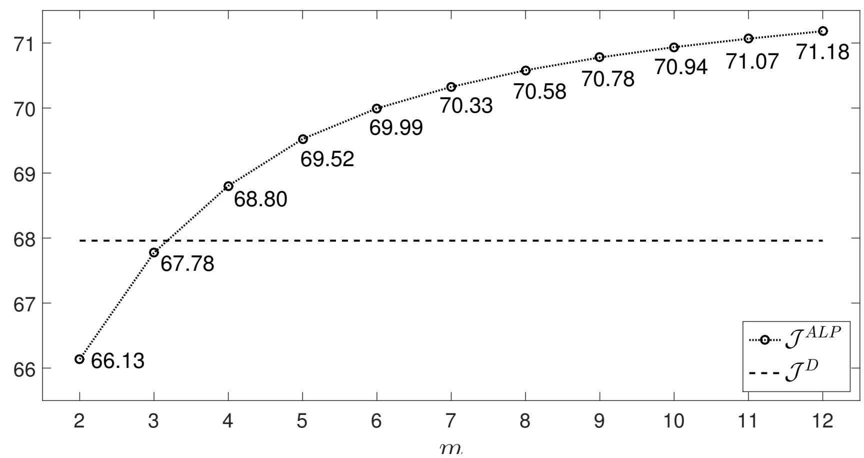


Figure 2.6: Optimal Value of Different Deterministic Formulations

2.9 Closing Remarks

This paper studies the dynamic joint pricing and order fulfillment problem for an e-commerce retailers. An LP-based approximation scheme is proposed to address the difficulty caused by the inability to charge different prices to customers from different regions, and two heuristic controls are analyzed. There are several possible extensions of our current work. For example, in our model, each order is restricted to contain exactly one item. In reality, numerous online orders contain multiple items and it is very common that e-tailers strategically split order fulfillment from different FCs (Jasin and Sinha, 2015). It would be interesting to see how our method can be generalized to incorporate this scenario. Another potential direction is to study dynamic pricing and fulfillment problem in the omnichannel environment, where retailers can either use online FCs or nearby brick-and-mortar stores to satisfy demand. This may potentially complicate the optimization problem. As discussed in Chapter 2.5, it would also be practically relevant and impactful to develop a way to apply our framework to the setting with a large number of products. From the technical point of view, we believe that our analytical framework can certainly be used to address other stochastic optimization problems in the broader OM context where many inter-related decisions have to be jointly made in real-time.

CHAPTER 3

Real-time Dynamic Pricing for Revenue Management with Reusable Resources and Deterministic Service Time Requirements

3.1 Abstract

We consider the setting of a firm that sells a finite amount of resources to price-sensitive customers who arrive randomly over time according to a specified non-stationary rate. Each customer requires a service that consumes one unit of resource for a *deterministic* amount of time, and the resource is *reusable* in the sense that it can be immediately used to serve a new customer upon the completion of the previous service. The firm's objective is to set the price dynamically to maximize its expected total revenues. This is a fundamental problem faced by many firms in many industries. We formulate this as an optimal stochastic control problem and develop two heuristic controls based on the solution of the deterministic relaxation of the original stochastic problem. The first heuristic control is static since the corresponding price sequence is determined before the selling horizon starts; the second heuristic control is dynamic, it uses the first heuristic control as its baseline control and adaptively adjusts the price based on previous demand realizations. We show that both heuristic controls are asymptotically optimal in the regime with large demand and large number of resources. Finally, we consider two important generalizations of the basic model to the setting with multiple service types requiring different service times and the setting with advance service bookings.

3.2 Introduction

Consider a firm managing a fixed amount of resources to satisfy time-varying price-dependent demand over a finite (selling) horizon. The resources are homogeneous, which means that customers do not have preference over a specific unit of resource, and each arriving customer requests a single unit of resource for a consecutive and *deterministic* amount of time (i.e., deterministic *service time*). If a resource is available at the time of a new arrival, the new customer is immediately admitted into the system at the current list price and the service is immediately started without delay. (Later in this paper we will also consider the case with advance service booking where the service can be started at a fixed future time.) After the service is completed, the corresponding resource is released and can be directly used to satisfy a new demand (i.e., resource is *reusable*). The firm's objective is to maximize her expected total revenues throughout the horizon by setting prices dynamically. This is a fundamental problem faced by many firms in many different industries and the nature of this problem is not exactly identical to the canonical revenue management problem with stochastic demand and limited inventory (e.g., the classic model proposed in Gallego and van Ryzin 1997). (To be precise, although it is mathematically possible to model revenue management with reusable resources and deterministic service time requirement using the same modeling approach as in the classic revenue management literature, the scale of the problem primitives for the applications considered in this paper is different from that considered in the standard revenue management literature. Hence, a different approach is needed to properly analyze this model; see Chapter 3.4 for more discussions.) Our main contribution in this paper is in developing a real-time dynamic pricing control that is easy to implement and has a provably good performance. We first show how to do this for a basic setting with one service type and immediate service requirement; we then show how our idea can be applied to more complicated settings with multiple service types with heterogenous service time requirements and advance service booking. We believe that the idea behind our proposed control can be potentially used to develop more sophisticated dynamic pricing controls for other complicated real-world problems.

Our formulation captures the critical operational trade-offs faced by firms in different industries. On the one hand, capacity needs to be sufficiently utilized throughout the selling horizon since, at any point of time, any unused or idle capacity constitutes immediate monetary loss; on the other hand, firms may also want to ration the capacity to anticipate potential peak periods in the future where the system is fused with incoming demands. The main challenge here is how to properly balance the capacity utilization during different *service cycles*. (The meaning of service cycle will be explained in Chapter 3.4. Note that, due to the difference in the scale of problem primitives as noted above, the classic revenue management problem effectively only has one service cycle as demands are typically modeled to be fulfilled only at the end of the selling horizon instead of on a rolling horizon basis. This is in contrast to the setting considered in this paper, which may have a large number of service cycles.) Although different cycles may appear to be independent of each other, they are connected through the realization of capacity utilization since capacity is finite and the utilization in one cycle affects the utilization in the subsequent cycle. This calls for a carefully designed dynamic pricing control to properly manage capacity utilization across different cycles.

In the queueing literature, a finite capacitated system similar to the one considered in our work is often termed as a *loss system*, since an arriving customer is rejected when the capacity is full (on the contrary, in a *delay system* model, incoming customers are allowed to wait in a queue, see Hampshire et al. 2009). Many firms providing virtual services such as telecommunication, smart grid, and Internet-based service (Voice-over-IP, wireless data transfer) can be appropriately modeled as loss systems. In all these examples, pricing decision is important not only because it serves as a marketing instrument that determines the total revenues collected by the firm, but also as a control instrument by which the firm continuously manages the utilization level of her finite resources. There are at least two salient features of the firms' operation problem that often complicate the pricing decision: on the demand size, demand rates tend to change dynamically and is better described as a time-inhomogeneous process (see Brown et al. 2005 for a statistical study in the setting of call center); on the supply side, capacity expansion is sometimes a long-term investment decision and the current capacity is not easily scalable within a short period of time, thus, they must be managed properly. The effectiveness of dynamic pricing in matching time-varying demand with limited capacity has been widely recognized and implemented by many firms, e.g., mobile service providers in Africa charges rates of voice-call dynamically to alleviate the burden on their bandwidth during peak periods and stimulate demand during low period (Economists, 2009); smart grids in United States and Europe experiment with programs that bill customers' consumption of electricity on a time-dependent rate (Hu et al., 2015).

In addition to the two salient features mentioned above, pricing decision is also often complicated by the fact that, once used, the same resource may continue to be used during a fixed period of time, and different customers may request to use the resource for different length of time (i.e., different service time). One of the emerging business that fits this feature is cloud computing, where firms deliver on-demand internet-based computing service to customers. Cloud computing service providers usually have a fixed amount of computation resources and lease their available resources to customers who arrive (either on spot or under subscription) randomly with specific request on usage time and capacity requirements. In the provision of cloud service, researchers and practitioners have advocated the economic benefit of dynamic pricing strategy for many cloud service settings. By and large, dynamic pricing has been implemented under the form of utilization-based pricing (CloudSigma, Jelastic, PiCloud), real-time bidding (Amazon Elastic Computing Cloud (EC2) Spot), and many others (Al-Roomi et al., 2013). As arguably the largest cloud computing service provider, Amazon launched its EC2 Spot service in 2009, whose per-hourly price is determined in real-time by a Vickrey-style auction. More specifically, after customers submit sealed bids, Amazon will computes a market clearing price (a.k.a "spot price"). All customers with bid above spot price win, and pay the lowest winning bid for the service with the requested features such as duration, memory size, etc. Not too surprisingly, the resulting price trajectory is often highly non-stationary (Xu and Li, 2013) and, in spite of its flexibility, the implementation of bidding mechanism has its own flaws. As an example, Cheng et al. (2016) shows empirically that, for the same type of computing service on Amazon EC2 Spot platform, network latency causes significant and consistent price difference between its East and West data center, which clearly opens an arbitrage opportunity. These problems would not have existed if the firm has a full control over the price trajectory. The key technical question is how to implement a dynamic pricing in a way that matches time-varying demand with fixed but reusable resources. This has motivated many researchers to investigate a proper dynamic pricing control under various settings where service providers fully control the price (e.g., Xu and Li 2013, Alzhouri and Agarwal 2015, Arshad et al. 2015).

Other than the examples discussed above, many firms managing physical resources are also well described by our model, including some classic examples that are well-known for their adoptions of dynamic pricing such as car rental and hotel reservation. (Due to the reusable nature of their resources, both car rental and hotel reservation are more properly modeled using the framework of revenue management with reusable resources and deterministic service time requirements instead of using the classic revenue management framework motivated by airline application.) Moreover, our model can also be used to address the demand-supply matching problem faced by many emerging so-called on-demand service firms. These firms usually control a finite number of resources and offer them to be consumed by customers who book services (either in advance or on the spot) through internet or smartphone. Industries that have seen the booming of on-demand service providers include vehicle rental (Zipcar, Citi-Bike), logistic (Project44), food delivery (Instacar, Sprig), car parking (Luxe), and beauty service (StyleSeat) (Bensinger, 2015). One prevalent feature in many on-demand service firms is that demand is characterized by both a specified service type and an intended usage time, including the starting and ending times of the service. Moreover, since customers mostly interact with the firm using digital platforms, existing digital user interfaces often enable the firm to effortlessly manage demand by dynamically changing prices. Indeed, some firms have already used dynamic pricing on a daily basis. For example, Project44 provides dynamic pricing solutions to third-party logistics company owning their own trucks and facilities (Project44, 2015); Tock provides ticketing systems to high-end restaurants where reservation of seats are dynamically priced (Businessweek, 2015); Sprig uses its own employees to deliver fresh made meals to customers at a delivery fee that changes dynamically (Chamlee, 2016). Other firms that have not yet deployed dynamic pricing have also acknowledged its advantage: According to Robin Chase, the founder of Zipcar, utilizing data to correctly and dynamically set the price on car-sharing platform can largely increase the efficiency and sustainability of the deployment of city services (GreenBiz, 2014). Thus, although not every firms under the banner of on-demand economy is currently using dynamic pricing, given its simplicity and good performance, we believe that our proposed real-time dynamic pricing control can provide a useful guidance on how to do real-time dynamic pricing when the firm finally needs it.

We want to re-emphasize that eventually different business models may have different complexities that require separate customized dynamic pricing solutions. In this paper, we simply focus on a simple model that captures the most fundamental aspects of dynamic pricing with reusable resources and deterministic service time requirements. We hope that our result can be used to design more sophisticated algorithms to be used in all of the aforementioned examples.

Our results and contributions. In this paper, we consider a multi-period dynamic pricing problem faced by a revenue-maximizing firm with finite reusable resources and deterministic service time requirements. Our analysis and results are summarized below:

1. We first consider a basic model where all customers have the same deterministic service time requirements, there is no delay in service fulfillment, and demand

rate as a function of time and price is non-stationary. We propose a deterministic relaxation of the optimal control formulation, and show that its objective value serves as an upper bound for the optimal expected total revenues under the original stochastic control problem. This allows us to evaluate the performance of any feasible pricing control by its *average regret*, defined as the average (over T periods) difference between the optimal value of the deterministic formulation and the expected total revenues collected under the prescribed control.

2. Our first heuristic control, which we call *Deterministic Price Control (DPC)*, applies price p_t in period t in such a way that that the expected demand in period t equals the computed deterministic demand rate under the deterministic formulation minus a constant. The constant serves as a buffer on random error, for the purpose of hedging against uncertainty. The size of this buffer needs to be carefully chosen: It needs to be large enough such that the resource is not depleted too often; yet, it cannot be too large otherwise the total revenues collected by the firm will deviate too far from the optimal one. We obtain a general bound on the average regret of DPC under arbitrary problem parameters and show that, under an optimal choice of buffer size, the average regret of DPC converges to zero at a rate of $\tilde{O}(n^{-\frac{1}{2}})$, where n is the size of the problem (i.e., the size of potential demand during a service cycle, which is to be defined later, and capacity are both of order n).

3. One drawback of DPC is that the price p_t to be applied during period t is already determined at the beginning of the selling horizon and it does not take into account to the realized demand observations during periods 1 to $t-1$. This suggests a room of improvement and motivates our second heuristic control, which we call *Deterministic Price Control with Batch Adjustment (DPC-Batch)*. DPC-Batch divides the selling horizon into batches of the same size. At each period, in addition to making sure that we have the buffer as in the case of DPC, we also set the price in such a way that the cumulatively demand errors (i.e., from expected demands) during the previous batch is uniformly corrected by the new demands in the current batch. We obtain a general bound for the average regret of DPC-Batch under arbitrary problem parameters and show that, under an optimal choice of buffer size and batch size, the average regret of DPC-Batch is of order $\tilde{O}(n^{-\frac{2}{3}})$, which significantly improves the performance of DPC. We conduct several numerical experiments that validate our theoretical findings.

4. Finally, we consider two extensions of the basic model to include two important

features often found in practice, namely heterogeneous service time requirements (where different service type may require different service time) and advance service booking (where different service type may be started at different time in the future). We focus our analysis on the generalization of DPC-Batch. For the sake of clarity how the analysis of our basic model can be extended to a more general model, we treat these two extensions as separate instances instead of one. Under properly chosen problem parameters, we show that the average regret of the generalized DPC-Batch for each of these extensions is still of the order $\tilde{O}(n^{-\frac{2}{3}})$.

Organization of the paper. The related literature is reviewed in Chapter 3.3. In Chapter 3.4, we formulate the basic model of dynamic pricing with reusable resource, and discuss our performance measure. We propose and analyze a static heuristic control (DPC) and its dynamic improvement (DPC-Batch) in Chapter 3.5 and 3.6, respectively. The performance of both DPC and DPC-Batch are tested in simple numerical experiments in Chapter 3.7. Chapter 3.8 and 3.9 discuss two extensions of the basic model that allow heterogeneous service time requirements and advance booking. Finally, in Chapter 10, we conclude the paper. The proof of some of the results and the details of the numerical experiments can be found in Appendix B.

3.3 Literature Review

Broadly speaking, our work is related to the extensive literature on dynamic pricing and revenue management, queueing and service operations, and on-demand service platforms. In terms of methodology, our work is related to the study of asymptotic performance of heuristic controls with real-time adjustment. We discuss them in turn.

Dynamic pricing and revenue management. Given the space limit, we will not attempt to discuss all the related literature but only highlight the most relevant works (interested readers are referred to the extensive surveys by Bitran and Caldentey 2003, Talluri and van Ryzin 2006 and Özer and Phillips 2012.) Instead, we discuss in details two papers that are most closely related to our work, both are motivated by the revenue management problem in cloud computing setting. Xu and Li (2013) study the dynampic pricing problem of a cloud service provider that leases resources to customers with exponential service time and price-dependent Poisson arrival. They obtain some structural properties for the capacitated system under stationary demand and also for the uncapacitated system under non-stationary demand. However, no dynamic pricing heuristic control is proposed and the optimal price is still time-consuming to compute, especially when demand is non-stationary. Our work complements their work: We explicitly address the capacitated system with non-stationary demand and deterministic service time requirement, and focus on developing an easy-to-implement heuristic control instead of studying the properties of the optimal solution. Borgs et al. (2014) study a similar problem under non-stationary demand with limited time-varying capacity and customers' strategic waiting. In their model, demands are assumed to be deterministic and the price trajectory for the whole season is announced at the beginning of the horizon. They show that the resulting optimization problem is non-convex and propose a dynamic programming-based algorithm that can be run in polynomial time. The key difference between our model and theirs is on the stochasticity of demand and customer's strategic waiting: In our model, demand is random and, thus, an adaptive heuristic control is needed to guarantee a near-optimal revenue. (In many service settings, especially for the on-demand platform, uncertainty in demand pervasively exists and introduces a significant difficulty in control design.) Unlike their model, we do not explicitly consider customer waiting behavior in our current work. Although customers' waiting is an important issue and needs to be properly taken into account when designing a dynamic pricing control, proposing a provably good heuristic control under a combination of stochastic demand, limited inventory, and customers' waiting is a notoriously difficult problem even in the traditional revenue management setting (see e.g., Liu and Cooper 2015, Chen et al. 2017b and Chen and Farias 2018 for recent progress) and in the reusable resource setting (Chen and Shi, 2018). Thus, we leave this for future research pursuit.

Queueing and service operations. As explained in the previous chapter, our model is similar to the loss system in the queueing literature. Pricing decision in such model has been studied extensively under various setting (e.g., Lanning et al. 1999, Courcoubetis et al. 2001 and Maglaras and Zeevi 2005). Most of these papers propose heuristic controls based on a fluid approximation of the original stochastic control problem under the assumption of stationary arrival and exponential service time. An exception to this is Hampshire et al. (2009), where demand follows a non-homogeneous Poisson process and the firm has to satisfy a Quality-of-Service constraints which requires the blocking probability to be bounded. They develop a dynamic pricing control using deterministic optimal control theory and show numerically that this control performs better than static or myopic pricing control; however, no theoretical performance guarantee is provided of their proposed control. Another major stream of literature studies the property of the optimal admission control of loss system, including Miller (1969), Kelly (1991), Altman et al. (2001), Örmeci et al. (2001), Savin et al. (2005), Gans and Savin (2007), Papier and Thonemann (2010) and Jain et al. (2015). Yet, none of them consider the design of practical and provably-good heuristic controls. There are two exceptions: Levi and Radovanovic (2010) propose a heuristic control based on a knapsack-type linear program and show the asymptotic optimality their proposed control under a general service time distribution, and Chen et al. (2017c) generalize this heuristic control to the setting with advance booking and provide an asymptotic upper bound on the blocking probability. However, both Levi and Radovanovic (2010) and Chen et al. (2017c) assume stationary demand and do not consider dynamic pricing.

Aside from the literature on loss model, dynamic pricing has also been studied in the literature on delay model. From the modeling perspective, researchers that study optimal dynamic pricing control either assume that customers are sensitive to price only but not delay (e.g., Low 1974, Paschalidis and Tsitsiklis 2000, Yoon and Lewis 2004, Maglaras 2006) or customers are sensitive to both price and delay (e.g., Chen and Frank 2001, and Ata and Shneorson 2006, Afèche and Ata 2013). Several papers study asymptotically optimal dynamic pricing controls: Çelik and Maglaras (2008) and Ata and Olsen (2009, 2013) study a revenue maximizing control when the firm dynamically quotes lead-times; Besbes and Maglaras (2009) study dynamic pricing where the market size varies stochastically over time; assuming observable queue length and stochastic customer valuation, Kim and Randhawa (2018) propose a heuristic control that continuously refines the baseline control given by a fluid approximation, and show (somewhat surprisingly) that the average regret is on the order of $\tilde{O}(n^{-\frac{2}{3}})$. (To the best of our knowledge, Kim and Randhawa (2018) is the only work in the queueing literature that shows dynamic pricing can achieve an average regret with order smaller than the more typical $\tilde{O}(n^{-1/2})$.) Aside from not permitting customers to wait, our model is different from the above cited works adopting asymptotic analysis in two aspects: (1) We assume that the service time is deterministic (earlier works assume that it is exponentially distributed) and the demand function can vary over time (earlier works assume a stationary willingness-to-pay distribution), and (2) we also consider an extension with advance service booking. The appropriateness of using either a deterministic service time or exponentially distributed service time is dictated by the application context. In this paper, we choose to work with deterministic service time because, in most of the applications that we are considering, service process is not memoryless (as would have been implied by an exponentially distributed service time). Thus, our work complements existing works in the queueing literature by developing a near-optimal dynamic pricing control that can be applied in the setting of non-stationary demand, deterministic service time requirements, and advance service booking. Moreover, we also complement the result of Kim and Randhawa (2018) by showing that the $\tilde{O}(n^{-2/3})$ bound is also achievable in our setting.

On-demand service platform. Our paper is also connected to the growing literature on the operational problems faced by firms providing various types of on-demand services. Most of the existing works focus on a specific industry and, henceforth, deal with more complicated models than ours. (Per our discussions in Chapter 3.2, our objective in this paper is to focus on the most fundamental aspects of revenue management with reusable resources and deterministic service time requirements instead of addressing a particular problem instance with all its complexities.) One line of research in this literature studies the logistic optimization problems for vehicle/bike sharing platforms (e.g., Raviv and Kolka 2013, Shu et al. 2013, Schuijbroek et al. 2017, O'Mahony 2015 and Kaspi et al. 2016). Existing works that study pricing decisions are Pfrommer et al. (2014) and Waserhole (2014). They both consider a network of shared mobility system and view price as an incentive to direct customers to allocate resources in a way that inventory balancing is properly maintained throughout the network. Different heuristic controls are proposed based on certainty equivalent principle and are tested using numerical experiments. In contrast to our work, Pfrommer et al. (2014) and Waserhole (2014) use platform's expected cost of repositioning vehicle as the objective. Another stream of literature studies the optimization of dynamic delivery fee for the attended home delivery firms, e.g., Campbell and Savelsbergh (2006), Asdemir et al. (2009), Klein et al. (2015). The key trade-off addressed in these works is how to use price to incentivize customers to allocate their demands to different delivery time slots such that the profit (delivery fee minus the cost associated with service type and time slots) is maximized. Moreover, their systems are capacitated in the sense that the delivery capacity within each time slots is fixed and known. In comparison to our model, this modeling framework embraces less uncertainty since, in our model, the available capacity at any time depends dynamically on the past demand realizations.

Real-time control. In the broader dynamic optimization literature where a multi-period stochastic control problem is often difficult (if not impossible) to solve optimally, researchers often resort to simple heuristic controls. A specific type of heuristic control, called real-time control, calculates the decision at the current period as a simple (e.g., affine) function of a baseline control and the historical information. Driven by its practicality (as the name suggests, a real-time control adaptively adjusts the control on the fly and does not require heavy re-optimizations) and good performance, real-time control has been investigated in various fields, including robust optimization (Ben-Tal et al. 2004, Bertsimas et al. 2010), portfolio management (Calafiore 2009, Moallemi and Saglam 2012), and revenue management (Atar and Reiman 2012, Chen and Farias 2013 Golrezaei et al. 2014, Chen et al. 2015, Lei et al. 2017). Closest to our paper are Jasin (2014) and Chen et al. (2015). They both consider the discrete-time version of the canonical dynamic pricing problem studied in Gallego and van Ryzin (1997), and propose real-time price controls with provable performance guarantees. As discussed in Chapter 3.4.2, in theory, our problem can also be formulated using the same framework as in Gallego and van Ryzin (1997); however, the reusability of resource in our setting introduces a non-trivial subtlety that prohibits a simple adoption of the heuristic controls proposed in Jasin (2014) and Chen et al. (2015) into our setting. (In fact, we show numerically in Chapter 3.7 that a simple adoption of this heuristic control performs very poorly.) Thus, our work complements existing works in the literature of real-time control by proposing a different real-time price control that is appropriate for the setting of revenue management with reusable resources and deterministic service time requirements.

3.4 Basic Model

In this chapter, we first discuss the setting and primitive of our basic model. Next, we discuss the stochastic and deterministic formulations of our dynamic pricing problem. Finally, we discuss our performance measure.

3.4.1 The Setting

We consider a discrete-time model with T periods and C units of resource. (Although we assume a discrete-time model, our results also hold for a continuous-time model with Poisson arrivals.) For our basic model, we assume that the firm only sells one service (or product) type where each request requires one unit of resource and n units of service time (or n periods). For example, if $n = 1$, then the service started in period 1 is completed at the end of period 1 and the resource used to fulfill this service is immediately available to fulfill a new request in period 2. Demand rate, as a function of price, in period t is given by $\lambda_t(p_t)$, and the corresponding revenue rate is given by $r_t(p_t) = p_t \cdot \lambda_t(p_t)$. Let $D_t(p_t)$ denote the realized demand in period t under price p_t. By definition, we have $\mathbf{E}[D_t(p_t)] = \lambda_t(p_t)$ and $\mathbf{E}[p_t \cdot D_t(p_t)] = r_t(p_t)$. It is typically assumed in the literature that demand rate is invertible in price (see Assumption A1 below). Thus, by abuse of notation, we will also write $D_t(\lambda_t) = D_t(p_t(\lambda_t))$ and $r_t(p_t) = p_t \cdot \lambda_t(p_t) = \lambda_t \cdot p_t(\lambda_t) = r_t(\lambda_t)$ to denote the direct dependency of realized demand and revenue rate on demand rate instead of on price (we use $p_t(\cdot)$ to denote the inverse of $\lambda_t(\cdot)$). We assume that demands across different periods are independent, but demand rate as a function of time may be non-stationary. As is typical in the revenue management literature (see e.g. Jasin 2014), we further assume that at most one request arrives during each period. (Thus, $\lambda_t(p_t)$ can be interpreted as the arrival probability of a new request in period t under price p_t.) This is without loss of generality since our analysis can also be applied to the setting where multiple requests arrive in each period. Let Ω_p and Ω_λ denote the convex feasible set of price and demand rate, respectively. (For simplicity, we assume the same feasible sets in all periods.) Below, we state some standard regularity conditions on $\lambda_t(\cdot)$ and $r_t(\cdot)$:

A1. $\lambda_t(p_t) : \Omega_p \to \Omega_\lambda$ is bounded, twice differentiable, and invertible.

A2. There exists a "turn-off" price $\bar{p}$ such that $p_t^k \to \bar{p}$ implies $\lambda_t(p_t^k) \to 0$ for all t.

A3. For all t, $\lambda_t^k \to 0$ implies $\lambda_t^k \cdot p_t(\lambda_t^k) \to 0$ for all feasible sequences $\{\lambda_t^k\}_{k=1}^{\infty}$.

A4. $r_t(\lambda_t)$ is bounded, strictly concave, and has a finite maximizer $\lambda_t^* \in \Omega_\lambda$.

The above assumptions are sufficiently general and are immediately satisfied by most commonly demand functions including linear, exponential, power, and logit. The existence of a turn-off price $\bar{p}$ allows the firm to effectively turn off demand whenever needed (e.g., when no resource is currently available). It should be noted that although the theoretical turn-off price can be infinite (e.g., for exponential demand function with $\lambda_t(p_t) = a \cdot e^{-p_t}$), since real-world price is never infinite, we can assume without loss of generality that $\bar{p} < \infty$. (To be precise, we can pick a sufficiently large $\bar{p}$ such that both $\lambda_t(\bar{p})$ and $r_t(\bar{p})$ are very small. The exact value of $\bar{p}$ does not affect our analysis.)

3.4.2 The Stochastic and Deterministic Formulations of Dynamic Pricing Problem

The dynamics of our pricing problem are as follows. First, a new request arrives at the beginning of period t with probability $\lambda_t(p_t)$. If a unit of resource is available, the service is immediately started (i.e., no waiting is allowed) and, once a service is started in period t, it will be completed at the end of period $t+n-1$. The corresponding resource is then immediately available for a new service in period $t+n$. No intervention or cancellation is allowed, i.e., neither the firm nor the customer can stop the service before it is completed. Since we assume at most one request arrives in each period, at most one service is completed at the end of any period.

Let Π denote the set of all non-anticipating controls (i.e, the control that decides the price at the beginning of period t using only the accumulated information up to, and including, the end of period $t-1$), and let p_t^π denote the price to be applied during period t under policy $\pi \in \Pi$. The optimal stochastic control formulation of our dynamic pricing problem is given below:

$$\textbf{OPT}: \; J^* \; = \; \left\{ \max_{\pi\in\Pi} \; \mathbf{E}\left[\sum_{t=1}^{T} r_t(p_t^\pi)\right] : \sum_{s=\max\{1,\, t-n+1\}}^{t} D_s(p_s^\pi) \le C \;\; \text{for all } t \le T \right\}$$

where the constraints must hold almost surely, or with probability one. To understand the intuition behind the above constraints, note that the number of units of resource available at the beginning of period t is given by $C - \sum_{s=\max\{1,\,t-n+1\}}^{t-1} D_s(p_s^\pi)$. Here, we only need to consider total demands in the previous $n-1$ periods because any resource being used in period $s < \max\{1,\, t-n+1\}$ must already complete its assigned service and is either at an idle state at the beginning of period t or currently being used to satisfy a new request arriving in period $s \in [\max\{1, t-n+1\},\, t-1]$, where by abuse of notation we use $[t_1,\, t_2]$ to denote $\{t_1, t_1+1, \ldots, t_2\}$. For a new service to be started in period t, we must satisfy capacity constraint $D_t(p_t^\pi) \le C - \sum_{s=\max\{1,\,t-n+1\}}^{t-1} D_s(p_s^\pi)$, or equivalently $\sum_{s=\max\{1,\,t-n+1\}}^{t} D_s(p_s^\pi) \le C$. This explains our constraints in **OPT**.

The exact stochastic formulation **OPT** is in general difficult to solve due to the famous "curse of dimensionality" of Dynamic Programming (DP). Our focus in this paper is on the construction of near-optimal heuristic controls using the solution of a deterministic analogue of **OPT**. We define a deterministic optimization **DET** as follows:

$$\textbf{DET}: \quad J^D \; = \; \left\{ \max_{p_t\in\Omega_p} \; \sum_{t=1}^{T} r_t(p_t) : \sum_{s=\max\{1,\, t-n+1\}}^{t} \lambda_s(p_s) \le C \;\; \text{for all } t \le T \right\}.$$

The above formulation is sometimes called a *fluid* model in the literature (e.g., Atar and Reiman 2012). Since demand is invertible in price (by Assumption A1), we can also re-write **DET** using demand rates as the immediate decision variables instead of prices as follows:

$$\textbf{DET}: \quad J^D = \left\{ \max_{\lambda_t \in \Omega_\lambda} \sum_{t=1}^{T} r_t(\lambda_t) : \sum_{s=\max\{1,\,t-n+1\}}^{t} \lambda_s \le C \;\text{ for all } t \le T \right\}.$$

One of the benefit of the above re-formulation is that the constraints are now linear in the decision variables and the objective is strongly concave by Assumption A4; so, **DET** can be efficiently solved using an off-the-shelf convex optimization solver. Note that the constraints in **DET** can be more compactly written as $A\lambda \le \mathbf{e}\cdot C$, where λ is a column vector of demand rates, $\mathbf{e}$ is a column vector of ones with an appropriate length, and A is an appropriate constant matrix. Although this compact formulation is similar to the canonical deterministic formulation in the standard revenue management literature (e.g., Gallego and van Ryzin 1997), it is important to note that the size of matrix A in our setting scales with T whereas the size of matrix A in the standard literature is independent of T. This seemingly minor difference has an important, non-trivial, consequence in heuristic design. This is the reason why a different approach is needed to properly analyze the general revenue management with reusable resources and deterministic service time requirements.

Let $p^D := (p_t^D)_{t=1}^T$ denote the optimal solution of **DET**, and let $\lambda^D := (\lambda_t^D)_{t=1}^T$ denote the corresponding optimal demand rates (i.e., $\lambda_t^D = \lambda_t(p_t^D)$ for all t). Unlike in the standard revenue management setting where the optimal deterministic price is static (i.e., $p_t^D = p_1^D$ for all t) when demand rates are stationary (see e.g. Gallego and van Ryzin 1997), the optimal solution of **DET** is not necessarily static even when demand rates are stationary (except for a special case T is a constant multiplicand of n). Below, we state additional assumptions on λ^D and the derivatives of revenue rate and price as functions of demand rate. There exist positive constants φ_L , φ_U, and Ψ such that:

A5. $[\lambda_t^D - \varphi_L,\, \lambda_t^D + \varphi_U] \subseteq \Omega_\lambda$ for all t.

A6. $|r_t'(\lambda)|$, $|r_t''(\lambda)|$, and $|p_t'(\lambda)|$ are bounded by Ψ on $[\lambda_t^D - \varphi_L,\, \lambda_t^D + \varphi_U]$ for all t.

The above assumptions are sufficiently general. Assumption A5 corresponds to the case where, at least in a deterministic world, the prices in all periods are neither too low that they collectively induce too many demands nor too high that they collectively induce too few demands. (This reflects what we find in most real-world settings as typical prices are neither extremely low nor outrageously high.) On another note, this assumption is also easily satisfied when λ_t^* lies in an interior of Ω_λ for all t, which is not at all uncommon given the strong concavity of $r_t(\cdot)$ as a function of λ_t. The boundedness of the derivatives of the revenue and price functions in an interior of Ω_λ as stated in Assumption A6 are also quite natural and easily satisfied by many demand functions including linear, exponential, power, and logit. Note that we only require that these derivatives are bounded in a certain compact subset of Ω_λ instead of the whole Ω_λ. The later is too restrictive and is not possible even for the case of power demand function $\lambda_t(p_t) = a \cdot p_t^{-b}$ since $r_t'(\lambda_t) \to \infty$ as $\lambda_t \to 0$.

The following lemma tells us that J^D is an upper bound of J^*. This result is analogous to a standard result in the revenue management literature (e.g., Gallego and van Ryzin 1997), and its proof utilizes a simple argument using Jensen's inequality. We state it here for the sake of completeness.

Lemma 3.4.1 *$J^* \leq J^D$.*

One of the benefit of Lemma 3.4.1 is that it allows us to use J^D as a proxy for J^*. This is particularly useful for the purpose of evaluating the performance of different heuristic controls since J^* is not practically computable. We discuss this next.

3.4.3 Performance Measure and Asymptotic Regime

Let R^π denote the total revenues collected under policy π throughout T periods. We are interested in measuring the average expected total losses, or average regret, of a given control with respect to the optimal control. However, since the optimal control is not computable as mentioned above, we will use the deterministic upper bound as a proxy. We thus defined the average regret of a non-anticipating control $\pi \in \Pi$ as follows:

$$\text{AvReg}(\pi) \;=\; \frac{J^D - \mathbf{E}[R^\pi]}{T}.$$

Intuitively, since the expected total revenues throughout T periods under the optimal control scales linearly with T, the above definition of average regret captures the order of relative regret with respect to the optimal control. In this paper, we are particularly interested in the case where n is large and $C = \Theta(n)$. This can be interpreted as the setting where total potential demands during a service cycle is large and we have just enough resources to satisfy the demands in one cycle. (For completeness, in Remarks 1 and 3 in Chapter 3.5 and 3.6, we also discuss what happens when

$C = o(n)$; this can be interpreted as the setting where either resources are very scarce or the length of service time is very long. The remaining case where we have a lot more resources than what we need to satisfy demands in one service cycle, i.e., $n = O(C)$, is less interesting as it reduces our dynamic pricing problem into an unconstrained problem and we can simply apply $p_t = p_t(\lambda_t^*)$ for all t.) This is not uncommon and is motivated by many practical applications discussed in Chapter 3.2. As the size of n can be very large (i.e., at least hundreds or thousands), we focus in constructing heuristic controls that are near-optimal in the so-called *asymptotic regime*. We would like to note that the setting where n is large and $C = \Theta(n)$ is also similar to the standard asymptotic setting in the queueing literature (e.g., Kim and Randhawa 2018) where both the demand and service rates are scaled by the same large constant.

We say that a control $\pi \in \Pi$ is *asymptotically optimal* if $\frac{J^D - \mathbf{E}[R^\pi]}{T} \to 0$ as $n \to \infty$ for a suitable value of T, which may also scale with n. In this paper, we prove that both DPC and DPC-Batch are asymptotically optimal. However, as n increases, the average regret of DPC-Batch converges to 0 faster than the average regret of DPC. (For our basic model, the convergence rate of DPC-Batch is approximately $n^{-2/3}$ whereas the convergence rate of DPC is approximately $n^{-1/2}$.) For ease of exposition, throughout the remaining of the paper, we will always assume that $\frac{T}{n} \in \mathcal{Z}^+$.

3.5 Deterministic Price Control

In this chapter, we first introduce a simple heuristic control called *Deterministic Price Control* (DPC) and then we analyze its performance.

3.5.1 Control Description and Statement of Result

Let C_t denote the number of units of resource available at the beginning of period t before the firm sets a new price p_t. The formal definition of DPC is given below.

Deterministic Price Control with Parameter ϵ (DPC(ϵ))

Step 1. Solve **DET** and get λ^D.
Step 2. At the beginning of each t, do:
a. If $C_t \geq 1$, set $p_t = \hat{p}_t^D$ where

$$\lambda_t(\hat{p}_t^D) \;=\; \lambda_t^D - \frac{\epsilon}{n};$$

b. Otherwise, set $p_t = \bar{p}$.

Note that DPC is parameterized by $\epsilon > 0$, and ϵ needs to be chosen such that $\lambda_t^D - \frac{\epsilon}{n} \in \Omega_\lambda$ (otherwise, the second step in DPC(ϵ) is not well-defined). Since the targeted demand rate in period t under DPC(ϵ) is $\lambda_t = \lambda_t^D - \frac{\epsilon}{n}$, the total targeted average demands in n consecutive periods (i.e., one service cycle) is at most $C - \epsilon$, which means that we are essentially holding back ϵ units of resource. We do this for the purpose of hedging against uncertainty: If total realized demands in the previous n periods turn out to be higher than expected, then we still have an extra ϵ units of resource that can be immediately used to satisfy demand. (From a theoretical perspective, having a positive ϵ is useful in making the analysis of DPC more tractable, though it may not be necessary for the actual implementation. In Chapter 3.7, we numerically test what happens when we set $\epsilon = 0$.) The following theorem tells us the performance of DPC.

Theorem 3.5.1 *There exists a constant $M_1 > 0$ such that for all T, C, n, and $\epsilon \in [1, n\varphi_L]$,*

$$\textsc{AvReg}(DPC) \;\le\; M_1 \cdot \left[\frac{\epsilon}{n} + \frac{T}{n} \cdot \exp\left\{-\frac{(\epsilon-1)^2}{36\min\{C-\epsilon,\, n\}}\right\}\right]. \tag{3.1}$$

In particular, if $C = a \cdot n$ for some $a > 0$, then using $\epsilon = 1 + 6\sqrt{b \cdot n \cdot \log n}$ for some $b > 0$ yields

$$\textsc{AvReg}(DPC) \;=\; O\left(\sqrt{\frac{b \cdot \log n}{n}} + \frac{T}{n^{1+\frac{b}{\max\{1,a\}}}}\right). \tag{3.2}$$

The first bound in Theorem 3.5.1 is very general; it highlights the impact of T, n, C, and $\epsilon \in [1, n\varphi_L]$ on performance. As for the second bound, as long as T grows at a polynomial rate in n (i.e., T can be very large, especially when n is large), we can always pick a proper b to make sure that the term $\frac{T}{n^{1+\frac{b}{\max\{1,a\}}}}$ in the second bound in Theorem 3.5.1 is of order $\frac{1}{n}$. Thus, for all practical purposes, the average regret of DPC when $C = \Theta(n)$ is of order $\sqrt{\frac{\log n}{n}}$. Note that we only need to have a buffer of order $\sqrt{n \cdot \log n}$. Since the magnitude of cumulative demand randomness in n consecutive periods is of order $\sqrt{n}$, this means that we only need to buffer a little bit more (i.e., by a factor of $\sqrt{\log n}$) to guarantee an asymptotically optimal performance under DPC.

Remark 1 (The Case of Scarce Resource). Although we have focused our discussions in Theorem 3.5.1 on the case $C = \Theta(n)$, the first bound in Theorem 3.5.1 also holds when $C = o(n)$. Suppose that demand rates are stationary and $\frac{T}{n} \in \mathcal{Z}^+$. It is not difficult to show in this case that the optimal deterministic solution is static, i.e., $\lambda_t^D = \frac{C}{n}$ for all t. Suppose that $C = n^\gamma$ for some $\gamma \in (0, 1)$ and let $\varphi_L = \varphi_U = \frac{1}{2n^{1-\gamma}}$. Then, using $\epsilon = 1 + 6\sqrt{b \cdot n^\gamma \cdot \log n}$ for some $b > 0$ yields an average regret of order $O\left(\sqrt{\frac{b \cdot \log n}{n^{2-\gamma}}} + \frac{T}{n^{1+b}}\right)$. If γ is close to 0 (but not exactly 0), then the average regret of DPC is practically of order $\frac{\sqrt{\log n}}{n}$. This means that DPC has a better performance in the setting of scarce resource. However, there is a caveat: If $C = \Theta(1)$ (e.g., $C = 1$), then the argument breaks down and the average regret of DPC is of order $\min\left\{1, \frac{T}{n}\right\}$ (i.e., the performance of DPC can be very poor). This is the setting of an *extremely* scarce resource and a different type of heuristic control seems to be needed to address this case. Since our focus in the paper is on the case $C = \Theta(n)$, we leave this for future research pursuit. (See also Remark 3 at the end of Chapter 3.6.)

3.5.2 Proof of Theorem 3.5.1

The proof of Theorem 3.5.1 can be separated into two steps. In the first step, we construct a high-probability event $\mathcal{G}$, and show that, on the set $\mathcal{G}$, we always have $C_t \geq 1$ and $p_t = \hat{p}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC(ϵ).

Step 1

We start with the first step. Let $\Delta_t(\hat{p}_t^D) = D_t(\hat{p}_t^D) - \lambda_t(\hat{p}_t^D)$ (i.e., $\Delta_t(\hat{p}_t^D)$ is the error from the expected demand in period t under price $\hat{p}_t^D$). For notational brevity, we will simply write $\lambda_t = \lambda_t(\hat{p}_t^D)$ and $\Delta_t = \Delta_t(\hat{p}_t^D)$. For some positive $\delta = o(n)$, whose exact value is to be determined later, define a sequence of events $\{\mathcal{A}_k(\epsilon, \delta)\}$ as follows:

$$\mathcal{A}_k(\epsilon, \delta) \;=\; \left\{ \max_{t \leq kn} \left| \sum_{s=(k-1)n+1}^{t} \Delta_s \right| \;<\; \delta \right\} \qquad \text{for all } k = 1, \ldots, \tfrac{T}{n}. \tag{3.3}$$

We now analyze $\mathbf{P}(\mathcal{A}_k(\epsilon, \delta))$. Note that, for all $r > 0$, we can bound:

$$
\begin{aligned}
&\mathbf{P}\left(\max_{t\le kn}\left|\sum_{s=(k-1)n+1}^{t}\Delta_s\right| \;\ge\; \delta\right)\\
&\le\; \frac{\mathbf{E}\left[\exp\left\{r\left|\sum_{s=(k-1)n+1}^{kn}\Delta_s\right|\right\}\right]}{\exp\{r\delta\}}\\
&\le\; \frac{\mathbf{E}\left[\exp\left\{r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right]+\mathbf{E}\left[\exp\left\{-r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right]}{\exp\{r\delta\}},
\end{aligned}
$$

where the first inequality follows from a sub-Martingale inequality (see e.g. Williams 1991) and the last inequality holds because $e^{|x|} \le e^x + e^{-x}$ for all x. Since $D_t(\lambda_t)$ is a Bernoulli random variable with success probability λ_t, by the Moment Generating Function of Bernoulli random variable,

$$
\begin{aligned}
\mathbf{E}\left[\exp\left\{r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right] &= \prod_{s=(k-1)n+1}^{kn}\mathbf{E}\left[\exp\{r\Delta_s\}\right]\\
&= \prod_{s=(k-1)n+1}^{kn}\left[e^r\cdot\lambda_t+1-\lambda_t\right]\cdot e^{-r\lambda_t} \;\le\; \prod_{s=(k-1)n+1}^{kn} e^{(e^r-1)\lambda_t}\cdot e^{-r\lambda_t}.
\end{aligned}
$$

Now, for all $|x|\le 1$, it holds that $e^x - 1 - x \le x^2$. Moreover, $\sum_{t=(k-1)n+1}^{kn}\lambda_t = \left(\sum_{t=(k-1)n+1}^{kn}\lambda_t^D\right) - \epsilon \le \min\{C-\epsilon, n\}$ (because at most one new request arrives in each period). So, we can bound:

$$
\mathbf{E}\left[\exp\left\{r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right] \;\le\; \exp\{r^2\min\{C-\epsilon,n\}\} \quad \text{for all } r\in[0,1].
$$

Note that similar arguments can also be applied to $\mathbf{E}\left[\exp\left\{-r\sum_{s=(k-1)n+1}^{kn}\Delta_s\right\}\right]$. Putting all things together, for $r\in[0,1]$, we have:

$$
\mathbf{P}(\bar{\mathcal{A}}_k(\epsilon,\delta)) \;\le\; 2\cdot\exp\{r^2\min\{C-\epsilon,n\}-r\delta\} \quad \text{for all } k=1,\ldots,\tfrac{T}{n}. \tag{3.4}
$$

Define $\mathcal{G}(\epsilon,\delta) := \cap_{k=1}^{T/n}\mathcal{A}_k(\epsilon,\delta)$. (Per our discussions above, $\mathcal{G}(\epsilon,\delta)$ is our high-probability event.) By the sub-additive property of probability,

$$
\mathbf{P}(\mathcal{G}(\epsilon,\delta)) \ge 1 - \frac{2T}{n}\exp\{r^2\min\{C-\epsilon,n\}-r\delta\}. \tag{3.5}
$$

We make an important observation— on the set $\mathcal{G}(\epsilon, \delta)$, we always have:

$$\sum_{s=t}^{t+n-1} D_s(\hat{p}_s^D) \;=\; \sum_{s=t}^{t+n-1} \left(\lambda_s^D - \frac{\epsilon}{n} + \Delta_s\right) \;\leq\; C - \epsilon + 3\delta \quad \text{for all } t+n-1 \leq T. \quad (3.6)$$

To see why, note that for any pair (t_1, t_2) with $t_1 \in [(k-1)n+1, kn]$ and $t_2 \in [kn+1, (k+1)n]$ for some $k \in \{1, \ldots, \frac{T}{n}\}$, we have: $\left|\sum_{s=t_1}^{t_2} \Delta_s\right| \leq \left|\sum_{s=t_1}^{kn} \Delta_s\right| + \left|\sum_{s=kn+1}^{t_2} \Delta_s\right| \leq 2\delta + \delta = 3\delta$, where the last inequality follows from the definition of δ in (3.3). This observation has an important implication: If we set $\delta = \frac{\epsilon - 1}{3}$, then we always have $C_t \geq 1$ and $p_t = \hat{p}_t^D$ for all t on the set $\mathcal{G}(\epsilon, \delta)$. For the remaining of the proof, we will therefore assume that $\delta = \frac{\epsilon-1}{3}$.

Step 2

We are now ready to bound the expected regret of DPC(ϵ). Let $\{p_t\}$ be the price sequence under DPC(ϵ) and let $r^u = \max_t \max_{\lambda_t \in \Omega_\lambda} r_t(\lambda_t)$. Note that

$$\begin{aligned}
J^D - \mathbf{E}[R^{DPC(\epsilon)}] &= J^D - \mathbf{E}\left[\sum_{t=1}^{T} r_t(p_t)\right] \\
&\leq J^D - \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\hat{p}_t^D)\right) \cdot \mathbf{1}\{\mathcal{G}(\epsilon, \delta)\}\right] \\
&= J^D - \mathbf{E}\left[\sum_{t=1}^{T} r_t(\hat{p}_t^D)\right] + \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\tilde{p}_t^D)\right) \cdot \mathbf{1}\{\bar{\mathcal{G}}(\epsilon, \delta)\}\right] \\
&\leq \sum_{t=1}^{T} \left[r_t(\lambda_t^D) - r_t\left(\lambda_t^D - \frac{\epsilon}{n}\right)\right] \;+\; r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon, \delta)) \\
&\leq T\Psi \cdot \frac{\epsilon}{n} \;+\; r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon, \delta)).
\end{aligned}$$

where the last inequality follows by the fact that $\epsilon \in [1, n\varphi_L]$ (which implies $\lambda_t^D - \frac{\epsilon}{n} \in [\lambda_t^D - \varphi_L, \lambda_t^D + \varphi_U]$) and by Assumption A6. Together with the bound in (3.5) and Assumption A6, we have for all $r \in [0, 1]$:

$$\begin{aligned}
\frac{J^D - \mathbf{E}[R^{DPC(\epsilon)}]}{T} &\leq \frac{1}{T} \cdot \left[\frac{T\Psi\epsilon}{n} \;+\; r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon, \delta))\right] \\
&\leq \frac{\Psi\epsilon}{n} + \frac{2r^u T}{n} \cdot \exp\{r^2 \min\{C - \epsilon, n\} - r\delta\}
\end{aligned}$$

Taking $r = \frac{\delta}{2\min\{C-\epsilon, n\}}$ and substituting $\delta = \frac{\epsilon - 1}{3}$ yields:

$$\frac{J^D - \mathbf{E}[R^{DPC(\epsilon)}]}{T} \;\; \leq \;\; M_1 \cdot \left[\frac{\epsilon}{n} + \frac{T}{n} \cdot \exp\left\{ - \frac{(\epsilon - 1)^2}{36 \min\{C - \epsilon,\, n\}} \right\} \right] \qquad (3.7)$$

for some $M_1 > 0$ independent of T, C, n, and $\epsilon \in [1,\, n\varphi_L]$. This completes the proof. ■

3.6 Deterministic Price Control with Periodic Batch Adjustments

We now discuss an improvement of DPC with periodic batch adjustments. We first provide a description of our heuristic control and then we analyze its performance.

3.6.1 Control Description and Statement of Result

Let m be a positive integer such that $\frac{n}{m} \in \mathcal{Z}^+$. (This is only exposition clarity and does not affect the key result of our analysis; we discuss this in more detail in Remark 2 at the end of this subsection.) The idea behind our periodic adjustments is to slice the interval $[1,\, T]$ into $\frac{T}{m}$ batches, each of length m periods, and then to adjust the prices in each batch in such a way that the cumulative errors in the previous batch is corrected in the current batch. To be precise, let $\{\mathcal{T}_i\}_{i=1}^{T/m}$ denote a partition of $[1,\, T]$, where $\mathcal{T}_i = [(i-1)m + 1,\, im]$ for all $i \geq 1$. For convenience, we assume that $\mathcal{T}_0 = \emptyset$. Define $\Delta_t(p_t) = D_t(p_t) - \lambda_t(p_t)$ (i.e., $\Delta_t(p_t)$ is the error from expected demand during period t under price p_t), where for notational brevity we will simply write $\Delta_t = \Delta_t(p_t)$. The complete definition of DPC with periodic batch adjustment (DPC-Batch) is given below.

DPC-Batch with Parameters m and ϵ (DPC-Batch(m, ϵ))

Step 1. Solve **DET** and get λ^D.

Step 2. At the beginning of each t, if $t \in \mathcal{T}_i$, do:

a. Compute $\hat{p}_t^D$ according to

$$\lambda_t(\hat{p}_t^D) \;\; = \;\; \lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m} \sum_{s \in \mathcal{T}_{i-1}} \Delta_s;$$

b. If $C_t \geq 1$ and $\lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m} \sum_{s \in \mathcal{T}_{i-1}} \Delta_s \in \Omega_\lambda$, set $p_t = \hat{p}_t^D$;

Otherwise, set $p_t = \bar{p}$.

Unlike the original DPC in Chapter 3.5, DPC-Batch is parameterized by two parameters m and ϵ. The value of these parameters must be carefully chosen. If m is too small, the price adjustment scheme under DPC-Batch may not have sufficient corrective power for re-balancing total demands in the current batch (e.g., cumulative errors in the previous batch may have the same order of magnitude as total potential demands in the current batch); if, on the other hand, m is too large, we already incur a lot of loss in the previous batch that is not recoverable by the adjustment in the current batch. The following theorem tells us the performance of DPC-Batch.

Theorem 3.6.1 *Suppose that* $\epsilon \in \left[1, \min\left\{n, m, n \cdot \frac{1+4m\cdot\min\{\varphi_L,\varphi_U\}}{4m+n}\right\}\right]$. *There exists a constant* $M_2 > 0$ *such that for all* T, C, n, m, *and* ϵ *we have*

$$\textsc{AvReg}(\textit{DPC-Batch}) \;\le\; M_2 \cdot \left[\frac{\epsilon}{n} + \frac{1}{m} + \frac{T}{m} \cdot \exp\left\{-\frac{(\epsilon-1)^2}{64\min\{C-\epsilon,\, m\}}\right\}\right]. \quad (3.8)$$

In particular, if $C = a \cdot n$ *for some* $a > 0$, *then using* $\epsilon = 1 + 8\sqrt{b \cdot n^c \cdot \log n}$ *and* $m = \lceil n^c \rceil$ *for some* $b > 0$ *and* $c \in \left(\frac{\log\log n}{\log n}, 1\right)$ *yields*

$$\textsc{AvReg}(\textit{DPC-Batch}) \;=\; O\left(\frac{\sqrt{b\cdot\log n}}{n^{1-\frac{c}{2}}} + \frac{1}{n^c} + \frac{T}{n^{c+\frac{b}{\max\{1,a\}}}}\right). \quad (3.9)$$

Similar to bound (3.2) in Theorem 3.5.1, as long as T grows polynomially in n, we can always pick a proper b such that the term $\frac{T}{n^{c+\frac{b}{\max\{1,a\}}}}$ is of order $\frac{1}{n}$. Thus, the performance of DPC-Batch when $C = \Theta(n)$ is largely affected by the choice of c. If c is too large (i.e., close to 1), then the bound is again of order $\sqrt{\frac{\log n}{n}}$ as in Theorem 3.5.1 (i.e., we do not get any benefit from batch adjustments); if, on the other hand, c is too small (i.e., close to 0), then the bound is of order 1. This means that, under our proposed periodic batch adjustment scheme, the length of each batch m should neither be too small nor too large for the most effective adjustment. Ignoring the logarithmic term in (3.9), the optimal bound is achieved when $c = 2/3$, which yields an average regret of order $\frac{\sqrt{\log n}}{n^{2/3}}$. This is a significant improvement over the bound in Theorem 3.5.1.

Remark 2 (The Case $\frac{n}{m} \notin \mathcal{Z}^+$). In the proof of Theorem 3.6.1, we assume that n is divisible by m for some $m > 1$. If, however, such m does not exist (i.e., n is a prime number), we only need to make a minor change in the definition of a batch. Formally, let $\mathcal{T}_i = [(i-1)m+1,\, im]$ for all $i = 1, \ldots, \lfloor \frac{T}{m} \rfloor - 1$, and $\mathcal{T}_{\lfloor \frac{T}{m} \rfloor} = [(\lfloor \frac{T}{m} \rfloor - 1)m + 1,\, T]$. Note that each of the first $\lfloor \frac{T}{m} \rfloor - 1$ batches still has the same length m, but the length of the last batch is between m and $2m$. With these new batches, the definition of $\hat{p}_t^D$ in Step 2 part a is re-defined as:

$$\lambda_t(\hat{p}_t^D) \;=\; \lambda_t^D - \frac{\epsilon}{n} - \frac{1}{|\mathcal{T}_i|} \sum_{s \in \mathcal{T}_{i-1}} \Delta_s.$$

Following the same arguments as in the proof of Theorem 3.6.1 (in Chapter 3.6.2), it is not difficult to check that the statement in Theorem 3.6.1 still holds under this minor alteration.

Remark 3 (The Case of Scarce Resource). Continuing our discussions in Remark 1 at the end of Chapter 3.5, if $C = n^{\gamma}$ for some $\gamma \in (0,1)$, then using $\epsilon = 1 + 8\sqrt{b \cdot n^{\min\{\gamma,c\}} \cdot \log n}$ yields an average regret of order $O\left(\frac{\sqrt{\log n}}{n^{1-\frac{\min\{\gamma,c\}}{2}}} + \frac{1}{n^c} + \frac{T}{n^{b+\min\{\gamma,c\}}}\right)$. Note that if γ is close to 0 (but not 0), we can choose c close to 1 and the average regret of DPC-Batch is practically of order $\frac{\sqrt{\log n}}{n}$, which is about the same order as the average regret of DPC. This means that, when resource is very scarce, periodic adjustment may not have a significant impact in improving performance.

3.6.2 Proof of Theorem 3.6.1

The proof of Theorem 3.6.1 follows similar arguments as the proof of Theorem 3.5.1. We still proceed in two steps: In the first step, we construct a high-probability event $\mathcal{G}$ and show that, on the set $\mathcal{G}$, we always have $C_t \geq 1$ and $p_t = \hat{p}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC-Batch(m, ϵ).

Step 1

We start with the first step. For some positive $\delta = o(m)$, whose exact value is to be determined later, define a sequence of events $\{\mathcal{A}_i(\epsilon, \delta)\}$ as follows:

$$\mathcal{A}_i(\epsilon,\delta) \;=\; \left\{ \max_{t \leq im} \left| \sum_{s=(i-1)m+1}^{t} \Delta_s \right| \;<\; \delta \right\} \qquad \text{for all } i \leq \tfrac{T}{m}. \tag{3.10}$$

Analogous to (3.4) in Chapter 3.5.2, it can be shown that for all $i \leq \frac{T}{m}$ and $r \in [0,1]$,

$$\mathbf{P}(\bar{\mathcal{A}}_i(\epsilon,\delta)) \quad \leq \quad 2 \cdot \exp\{r^2 \min\{C-\epsilon, m\} - r\delta\}. \tag{3.11}$$

Now, define $\mathcal{G}(\epsilon,\delta) = \cap_{i=1}^{T/m} \mathcal{A}_i(\epsilon,\delta)$. By the sub-additivity property of probability,

$$\mathbf{P}(\mathcal{G}(\epsilon,\delta)) \geq 1 - \frac{2T}{m} \exp\{r^2 \min\{C-\epsilon, m\} - r\delta\}. \tag{3.12}$$

We make some important observations. First, on the set $\mathcal{G}(\epsilon,\delta)$, we always have: $\left|\frac{\epsilon}{n} + \frac{1}{m}\sum_{s\in\mathcal{T}_i} \Delta_s\right| \leq \frac{\epsilon}{n} + \frac{\delta}{m}$ for all i. This means that, as long as the parameters ϵ, δ, and m are chosen such that $\frac{\epsilon}{n} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$, the condition $\lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}} \Delta_s \in \Omega_\lambda$ in Step 2 part a in the definition of DPC-Batch is always satisfied. For the remaining of the proof, we will therefore assume that $\frac{\epsilon}{n} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$. Now, suppose that $t \in \mathcal{T}_{j_1}$ and $t+n-1 \in \mathcal{T}_{j_2}$, where $j_1 < j_2$ and $t+n-1 \leq T$. We can write the total demands during $[t,\, t+n-1]$ as follows:

$$\begin{aligned}
&\sum_{s=t}^{t+n-1} D_s(\hat{p}_s^D) \\
&= \sum_{s\geq t, s\in\mathcal{T}_{j_1}} D_s(\hat{p}_s^D) \; + \sum_{j=j_1+1}^{j_2-1} \sum_{s\in\mathcal{T}_j} D_s(\hat{p}_s^D) \; + \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} D_s(\hat{p}_s^D) \\
&= \sum_{s\geq t, s\in\mathcal{T}_{j_1}} \left(\lambda_s^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{l\in\mathcal{T}_{j_1-1}} \Delta_l + \Delta_s\right) \\
&\qquad + \sum_{j=j_1+1}^{j_2-1} \sum_{s\in\mathcal{T}_j} \left(\lambda_s^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{l\in\mathcal{T}_{j-1}} \Delta_l + \Delta_s\right) \\
&\qquad + \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} \left(\lambda_s^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{l\in\mathcal{T}_{j_2-1}} \Delta_l + \Delta_s\right) \\
&= \sum_{s=t}^{t+n-1} \lambda_s^D \; - \; \epsilon \; - \; \frac{1}{m} \sum_{s\geq t, s\in\mathcal{T}_{j_1}} \sum_{l\in\mathcal{T}_{j_1-1}} \Delta_l \; - \sum_{s<t, s\in\mathcal{T}_{j_1}} \Delta_s \\
&\qquad + \left(\sum_{s\in\mathcal{T}_{j_2-1}} \Delta_s - \frac{1}{m} \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} \sum_{s\in\mathcal{T}_{j_2-1}} \Delta_s\right) + \sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}} \Delta_s.
\end{aligned}$$

Since $\mathcal{T}_j$ contains m periods for all j, on the set $\mathcal{G}(\epsilon,\delta)$, we can bound:

$$\left|\frac{1}{m}\sum_{s\geq t, s\in\mathcal{T}_{j_1}}\sum_{l\in\mathcal{T}_{j_1-1}}\Delta_l\right| \leq \delta, \quad \left|\sum_{s<t, s\in\mathcal{T}_{j_1-1}}\Delta_s\right| \leq \delta, \quad \left|\sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}}\Delta_s\right| \leq \delta$$
$$\text{and } \left|\sum_{s\in\mathcal{T}_{j_2-1}}\Delta_s - \frac{1}{m}\sum_{s\leq t+n-1,\, s\in\mathcal{T}_{j_2}}\sum_{s\in\mathcal{T}_{j_2-1}}\Delta_s\right| \leq \delta.$$

Putting the above four bounds together, on the set $\mathcal{G}(\epsilon,\delta)$, we have:

$$\sum_{s=t}^{t+n-1} D_s(\hat{p}_s^D) \quad \leq \quad C-\epsilon+4\delta \quad \text{for all } t+n-1\leq T. \tag{3.13}$$

It is worth noting that although (3.13) is similar to (3.6) in the proof of Theorem 3.5.1, the term δ in (3.6) represents a bound on cumulative errors during n periods whereas the term δ in (3.13) represents a bound on cumulative errors during $m<n$ periods (i.e., the δ in (3.13) is potentially much smaller than the δ in (3.6), which highlights the potential improvement due to batch adjustments).

Let $\delta = \frac{\epsilon-1}{4}$. Given this and the assumption that $\frac{\epsilon}{n}+\frac{\delta}{m} \leq \min\{\varphi_L,\varphi_U\}$, it is not difficult to see that the following always hold on $\mathcal{G}(\epsilon,\delta)$: $C_t \geq 1$ and $\lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s \in \Omega_\lambda$ for all i and $t\in\mathcal{T}_i$. As a consequence, we also have $p_t = \hat{p}_t^D$ for all t.

Step 2

We are now ready to bound the expected regret of DPC-Batch(m,ϵ). Let $\{p_t\}$ be the price sequence under DPC-Batch(m,ϵ). Note that

$$\begin{aligned}\mathbf{E}[R^{DPC-Batch(m,\epsilon)}] &= \mathbf{E}\left[\sum_{t=1}^T r_t(p_t)\right] \geq \mathbf{E}\left[\left(\sum_{t=1}^T r_t(\hat{p}_t^D)\right)\cdot \mathbf{1}\{\mathcal{G}(\epsilon,\delta)\}\right] \\ &= \mathbf{E}\left[\sum_{t=1}^T r_t(\hat{p}_t^D)\right] - \mathbf{E}\left[\left(\sum_{t=1}^T r_t(\hat{p}_t^D)\right)\cdot \mathbf{1}\{\bar{\mathcal{G}}(\epsilon,\delta)\}\right].\end{aligned}$$

The second expectation after the last equality above can be bounded by $r^u T\cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon,\delta))$ where $r^u = \max_t \max_{\lambda_t\in\Omega_\lambda} r_t(\lambda_t)$. As for the first expectation, suppose that $t\in\mathcal{T}_i$ for some $i\geq 2$. By Taylor's expansion and Assumption A6, we can bound $r_t(\hat{p}_t^D) = r_t\left(\lambda_t^D - \frac{\epsilon}{n} - \frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\right) \geq r_t(\lambda_t^D) - r_t'(\lambda_t^D)\cdot\left(\frac{\epsilon}{n}+\frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\right) - \Psi\cdot\left(\frac{\epsilon}{n}+\frac{1}{m}\sum_{s\in\mathcal{T}_{i-1}}\Delta_s\right)^2$. Taking expectation and applying Assumption A6 one more time yield $\mathbf{E}\left[r_t(\hat{p}_t^D)\right] \geq r_t(\lambda_t^D) - \frac{\Psi\epsilon}{n} - \Psi \cdot \left(\frac{2\epsilon^2}{n^2} + \frac{2}{m}\right)$, where the inequality follows because $(x+y)^2 \leq 2x^2+2y^2$ for all (x,y) and $\mathbf{E}\left[\left(\sum_{s\in\mathcal{T}_{i-1}} \Delta_s\right)^2\right] \leq m$ (by definition, $\{\Delta_s\}_{s\in\mathcal{T}_{i-1}}$ are independent zero-mean random variables and $|\Delta_s| \leq 1$).

Putting the bounds together, for all $r \in [0,1]$, we have:

$$\begin{aligned}
&\frac{J^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T} \\
\leq\ & \frac{1}{T} \cdot \left[\frac{T\Psi\epsilon}{n} + T\Psi \cdot \left(\frac{2\epsilon^2}{n^2} + \frac{2}{m}\right) + r^u T \cdot P(\bar{\mathcal{G}}(\epsilon,\delta))\right] \\
\leq\ & \frac{\Psi\epsilon}{n} + \frac{2\Psi\epsilon^2}{n^2} + \frac{2\Psi}{m} + \frac{2r^u T}{m} \cdot \exp\{r^2 \min\{C-\epsilon, m\} - r\delta\}.
\end{aligned}$$

Taking $r = \frac{\delta}{2\min\{C-\epsilon,m\}}$ and substituting $\delta = \frac{\epsilon-1}{4}$ yields:

$$\textsc{AvReg}(DPC - Batch) \leq M_2 \cdot \left[\frac{\epsilon}{n} + \frac{1}{m} + \frac{T}{m} \cdot \exp\left\{-\frac{(\epsilon-1)^2}{64\min\{C-\epsilon,\ m\}}\right\}\right]$$

for some $M_2 > 0$ independent of T, C, n, m, and

$$\epsilon \in \left[1,\ \min\left\{n, m, n \cdot \frac{1 + 4m \cdot \min\{\varphi_L, \varphi_U\}}{4m+n}\right\}\right]$$

(Note that $\delta = \frac{\epsilon-1}{4}$ and $\epsilon \leq \frac{n+4mn\cdot\min\{\varphi_L,\varphi_U\}}{4m+n}$ implies $\frac{\epsilon}{n} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$; $1 < \epsilon < m$ ensures that $r = \frac{\delta}{2\min\{C-\epsilon,m\}} = \frac{\epsilon-1}{8\min\{C-\epsilon,m\}} \in (0,1)$.) To get bound (3.9), we further require $c > \frac{\log\log n}{\log n}$ to ensure $r \in (0,1)$. This completes the proof. ■

3.7 Numerical Experiments

We now conduct simple numerical experiments to illustrate the performance of the proposed heuristic controls under different problem parameters. For simplicity, we assume that the demand function (i.e., purchasing probability) is stationary over time, and is exponentially decreasing in price, i.e., $\lambda(p) = \exp\left(\lambda_0 - \frac{p}{p_0}\right)$. We use $\lambda_0 = 0.8$ and $p_0 = 100$. The length of selling horizon T and resource capacity C are both set to be linear in n, and we vary n from $1,000$ to $8,000$. Specifically, we choose $C = 0.7 \cdot n$ and $T = 5 \cdot n$ (i.e., $n = 1,000$ corresponds to the problem instance with 700 units of resources and length of selling horizon equals 5,000 periods). The resulting deterministic problem has a stationary optimal solution $\lambda_t^D \equiv \lambda^D = 0.7$, with optimal objective value $J^D = 80.97$.

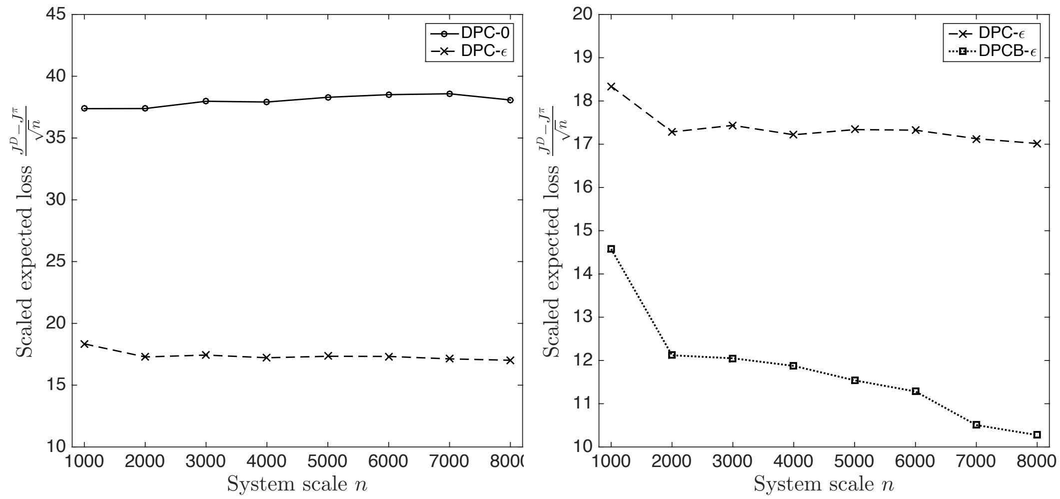
Table 3.1 summarizes the heuristic controls tested in all experiments. Several comments are in order regarding the implementation details. First, we implement DPC-0 to identify the impact of injecting a buffer for deterministic control. Second, in DPCB-ϵ, if $T/m \notin \mathbb{Z}_+$, we simply set the last period to be of size T mod m (given the discussions in Remark 2 in Chapter 3.6, this should not affect the performance of DPC-Batch by much). Third, LRC-k refers to a simple adoption of the self-adjusting control proposed in Jasin (2014), where we simply re-start the control at every k periods by setting the cumulative error to be zero (see Chapter B.4 for a detailed description). Fourth, for any combination of parameters, we simulate all the heuristic controls with 300 Monte Carlo runs to approximate their expected total revenues. Lastly, for DPC-ϵ and DPCB-ϵ, we simply use a grid-search method to find the optimal ϵ (between 0 and 1, with an increment of 0.01).

Table 3.1: Summary Description of All Heuristic Controls

Label	Description
DPC-0	DPC(ϵ) with $\epsilon = 0$ (defined in Chapter 3.5.1)
DPC-ϵ	DPC(ϵ) (defined in Chapter 3.5.1)
DPCB-ϵ	DPC-Batch(m, ϵ) with $m = \lceil n^{2/3} \rceil$ (defined in Chapter 3.6.1)
LRC-k	Linear rate control with re-starting at every k periods (see Section 4 in Jasin 2014)

Simulation results. Figures 3.1 and 3.1 show the expected total regrets of the first three heuristic controls, where the y-axis is the *scaled* total expected revenue loss $\frac{J^D - J^\pi}{\sqrt{n}}$. We do not plot the regret of LRC-k since it performs much worse than any of the other three heuristic controls under any k, but we report the complete numerical results in Appendix B.4. As expected, DPC-Batch dominates DPC, which in turn dominates DPC-0. Moreover, a closer look on the average regret confirms the asymptotic optimality of all three heuristic controls. The relatively poor performance (in fact, may not even be asymptotically optimal) of LRC suggests that a heuristic control that performs well in the setting of canonical revenue management cannot be directly adopted to the setting of revenue management with reusable resources and deterministic service time requirements. This reinforces our point in Chapter 3.2 that the setting considered in our work, though may appear identical, is not exactly the same as the setting in the standard revenue management literature.

Figure 3.1: Expected Average Regret with Varying n



3.8 Extension to Multiple Service Types with Heterogeneous Service Time Requirements

In this chapter, we discuss a generalization of the basic model in Chapter 3.4 that allows different service types with heterogeneous service time requirements. We first discuss the setting of the problem and then provide a generalization of DPC-Batch.

3.8.1 The Setting

The firm sells $K \geq 1$ service (or product) types where a request of service type k requires one unit of resource and n_k units of service time (or n_k periods). For ease of exposition, we will assume that $\frac{T}{n_k} \in \mathcal{Z}^+$ for all $k = 1, \ldots, K$. Moreover, without loss of generality, we also assume that the service types are labeled in such a way that $1 \leq n_1 \leq n_2 \leq \cdots \leq n_K$. The dynamics of the problems are as follows: At the beginning of period t, the firm sets the prices for all service types, denoted by a vector $\mathbf{p_t} := (p_{t,1}, \ldots, p_{t,K}) \in \Omega_p$. (Unless otherwise noted, all vectors are to be understood as column vectors.) For period t, a price vector $\mathbf{p}_t$ induces a demand vector $\mathbf{D}_t(\mathbf{p}_t) = (D_{t,1}(\mathbf{p}_t), \ldots, D_{t,K}(\mathbf{p}_t))$ with rate vector $\boldsymbol{\lambda}_t(\mathbf{p}_t) := (\lambda_{t,1}(\mathbf{p}_t), \ldots, \lambda_{t,K}(\mathbf{p}_t))$, where $\boldsymbol{\lambda}_t(\mathbf{p}_t) = \mathbf{E}\left[\mathbf{D}_t(\mathbf{p}_t)\right]$. The corresponding revenue rate is given by $r_t(\mathbf{p}_t) = \mathbf{E}[\mathbf{p}_t^\top \mathbf{D}_t(\mathbf{p}_t)] = \mathbf{p}_t^\top \boldsymbol{\lambda}_t(\mathbf{p}_t)$. By the invertibility assumption (see below), we will also use $\mathbf{D}_t(\boldsymbol{\lambda}_t) = \mathbf{D}_t(\mathbf{p}_t(\boldsymbol{\lambda}_t))$ and $r_t(\boldsymbol{\lambda_t}) = \boldsymbol{\lambda}_t^\top \mathbf{p}_t(\boldsymbol{\lambda}_t)$ to denote the realized demand vector and revenue rate as a function of demand rates, respectively. As in the basic model, we assume that demands across different periods are independent but demands over different service types within the same period may be correlated and demand rates as functions of time may be non-stationary. We assume at most one request arrives in each period, i.e., $\sum_{k=1}^{K} D_{t,k}(\mathbf{p}_t) \leq 1$ (this is without loss of generality). Let $\Omega_p = \otimes_{k=1}^{K}\Omega_{p,k}$ and $\Omega_\lambda = \otimes_{k=1}^{K}\Omega_{\lambda,k}$ denote the convex feasible set for price vector and demand rate vector, respectively. The following regularity conditions are the generalization of Assumptions A1-A4 in Chapter 3.4 to the multiple service types setting:

MA1. $\boldsymbol{\lambda}_t(\mathbf{p}_t) : \Omega_p \to \Omega_\lambda$ is bounded, twice differentiable and invertible.

MA2. For each k, there exists a "turn-off" price $\bar{p}_k$ such that $p_{k,t}^v \to \bar{p}_k$ implies $\lambda_{t,k}(\mathbf{p}_t^v) \to 0$.

MA3. $\boldsymbol{\lambda}_t^v \to \mathbf{0}$ implies $r_t(\boldsymbol{\lambda}_t^k) \to 0$ for any feasible sequence $\{\boldsymbol{\lambda}_t^v\}_{v=1}^{\infty}$.

MA4. $r_t(\boldsymbol{\lambda}_t)$ is bounded, strictly jointly concave, and has a finite maximizer $\boldsymbol{\lambda}_t^* \in \Omega_\lambda$.

The optimal stochastic control formulation of our dynamic pricing problem is given by:

$$\textbf{OPT-M}: \quad J_M^* = \left\{ \max_{\pi \in \Pi} \mathbf{E}\left[\sum_{t=1}^{T} r_t(\mathbf{p}_t^\pi) \right] : \sum_{k=1}^{K} \sum_{s=\max\{1,t-n_k+1\}}^{t} D_{s,k}(\mathbf{p}_s^\pi) \leq C \; \forall t \right\}$$

where the constraints must hold almost surely (or with probability one) and Π is the set of all non-anticipating controls. Using demand rate vector as the decision variable, the deterministic relaxation of **OPT-M** is given by:

$$\textbf{DET-M}: \quad J_M^D = \left\{ \max_{\boldsymbol{\lambda}_t \in \Omega_\lambda} \sum_{t=1}^{T} r_t(\boldsymbol{\lambda}_t) : \sum_{k=1}^{K} \sum_{s=\max\{1,t-n_k+1\}}^{t} \lambda_{s,k} \leq C \; \forall t \right\}$$

As in Lemma 1, it is not difficult to show that J_M^D is an upper bound of J_M^*. Therefore, the average regret defined in Chapter 3.4 can still be used as a proper performance measure. Let $\boldsymbol{\lambda}^D := (\boldsymbol{\lambda}_t^D)_{t=1}^T$ denote the optimal solution of **DET-M**, and let $\mathbf{p}^D := (\mathbf{p}_t^D)_{t=1}^T$ denote the corresponding optimal price vectors (i.e., $\mathbf{p}_t^D = \mathbf{p}_t(\boldsymbol{\lambda}_t^D)$). Let $\mathbf{e}$ be a vector of ones, with a proper dimension. Similar to Assumptions A5-A6, we assume that there exist positive constants φ_L , φ_U, and Ψ such that the following two conditions hold for all t:

MA5. $[\boldsymbol{\lambda}_t^D - \varphi_L \mathbf{e},\, \boldsymbol{\lambda}_t^D + \varphi_U \mathbf{e}] \subseteq \Omega_\lambda$.

MA6. $||\nabla r_t(\boldsymbol{\lambda})||_\infty$ and $||\nabla^2 r_t(\boldsymbol{\lambda})||_2$ are bounded from above by Ψ on $[\lambda_t^D - \varphi_L,\, \lambda_t^D + \varphi_U]$.

We are now ready to present the generalization of DPC-Batch in the setting with multiple service types and heterogeneous service time requirements.

3.8.2 A Generalized DPC-Batch and Its Performance

Let $\mathbf{m} = (m_1, \ldots, m_K)$ be a sequence of positive integers such that $\frac{n_k}{m_k} \in \mathcal{Z}^+$ for all k. (As in Chapter 3.6.1, the existence of such sequence is assumed for ease of exposition and does not affect our result. If a proper m_k satisfying $\frac{n_k}{m_k} \in \mathcal{Z}^+$ does not exist, then we can slightly modify our batch definition as in Remark 2 at the end of Chapter 3.6.1.) For each service type k, we slice the selling horizon into $\frac{T}{m_k}$ batches, each of length m_k periods. Let $\mathcal{T}_{k,i} = [(i-1)m_k + 1,\, im_k]$ denote the i^{th} batch for service type k. The key idea behind our generalized DPC-Batch is to manage the demand rate for each service type somewhat independently of the other service types. To be precise, the demand rates in each batch are adjusted in such a way that the cumulative errors for a given service type in the previous batch are corrected by the demands of the same service type in the current batch. (This does not mean that the controls are completely decoupled since demands over different service types are still connected through their prices, which means that the corresponding prices adjustments need to be computed jointly.) Let $\boldsymbol{\Delta}_t := (\Delta_{t,k})_{k=1}^K = (D_{t,k}(\mathbf{p}_t) - \lambda_{t,k}(\mathbf{p}_t))_{k=1}^K$ denote the vector of errors from expected demands in period t under price vector $\mathbf{p}_t$ (we suppress the dependency of $\boldsymbol{\Delta}_t$ on $\mathbf{p}_t$). Also, let $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_K)$ be a sequence of real-valued constants denoting the size of buffer for each service type, and define $i_k(t)$ such that $t \in \mathcal{T}_{k,i_k(t)}$ for all t and k. The complete definition of our generalized DPC-Batch with multiple service types and heterogeneous service time requirements is given below.

DPC-Batch with Parameters m and ϵ (DPC-Batch(m, ϵ))

Step 1. Solve **DET-M** and get $\boldsymbol{\lambda}^D$.

Step 2. At the beginning of each t, do:

a. Compute $\hat{\mathbf{p}}_t^D$ according to

$$\lambda_{t,k}(\hat{\mathbf{p}}_t^D) \;=\; \lambda_{t,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k} \sum_{s \in \mathcal{T}_{k,i_k(t)-1}} \Delta_{s,k} \quad \text{for all } k;$$

b. If $C_t \geq 1$ and $\lambda_{t,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)-1}} \Delta_{s,k} \in \Omega_{k,\lambda}$, set $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$; Otherwise, set $\mathbf{p}_t = \bar{\mathbf{p}}$.

Note that the price vector $\hat{\mathbf{p}}_t^D$ in Step 2 part a is well-defined by the invertibility assumption in MA1. Let $C_k^D := \max_{1\leq t\leq T}\sum_{s=\max\{1,\,t-n_k+1\}}^{t} \lambda_{s,k}^D$ denote the maximum amount of resource used by service type k in the deterministic model. The following theorem tells us the performance of DPC-Batch with heterogeneous service time requests; we defer its proof to the Appendix B.3.

Theorem 3.8.1 *Suppose that $0 < n_1 \leq \cdots \leq n_K \leq 1$. There exists a constant $M_3 > 0$ such that for all T, C, m_k, $\epsilon_k \in n_k \cdot \left[\frac{1}{Kn_1}, \min\left\{1, \frac{1}{K}\cdot\frac{1+4Km_k\cdot\min\{\varphi_L,\varphi_U\}}{4m_k+n_1}\right\}\right]$, and $n_1 \geq \frac{1}{K\min\{\varphi_L,\varphi_U\}}$ we have*

$$\begin{aligned}&\textsc{AvReg}(\textit{DPC-Batch})\\ \leq\;& M_3\cdot\sum_{k=1}^{K}\left[\frac{\epsilon_k}{n_k} + \frac{1}{m_k} + \frac{T}{m_k}\cdot\exp\left\{-\frac{(Kn_1\epsilon_k - n_k)^2}{64K^2n_k^2\min\{C_k^D-\epsilon_k,\, m_k\}}\right\}\right]. \quad (3.14)\end{aligned}$$

In particular, if $n_k = \alpha_k\cdot n$ and $C_k^D = \beta_k\cdot n$ for some $0<\alpha_1\leq\cdots\leq\alpha_K$ and $\beta_k > 0$ for all k, then using $\epsilon_k = \frac{\alpha_k}{K\alpha_1}(1+8\sqrt{b\cdot n^c\cdot\log n})$ and $m_k = \lceil n^c\rceil$ for all k, for some $b>0$ and $c\in[0,1)$, yields

$$\textsc{AvReg}(\textit{DPC-Batch}) = O\left(\frac{\sqrt{b\cdot n^c\cdot\log n}}{n} + \frac{1}{n^c} + \frac{T}{n^{c+\frac{b}{\max\{1,\max_k\{\beta_k\}\}}}}\right). \quad (3.15)$$

Two comments are in order. First, under a proper choice of b, setting $c = 2/3$ in (3.15) yields an average regret of order $\frac{\sqrt{\log n}}{n^{2/3}}$. This is the same order as the optimal bound as in Theorem 3.6.1 (with $c = 2/3$). Second, although the second bound in Theorem 3.8.1 only focuses on the case where $C_k^D = \Theta(n_k) = \Theta(n)$ for all k, the first bound in Theorem 3.8.1 holds in great generality. For example, if $n_k = \Theta(n^{\alpha_k})$ and $C_k^D = \beta_k\cdot n_k$ for some $\alpha_k,\beta_k>0$, we can use $\epsilon_k = \frac{n_k}{Kn_1}\left(1+8K\sqrt{b\cdot n_k^{c_k}\cdot\log n_k}\right)$ and $m_k = \lceil n_k^{c_k}\rceil$ for some $c_k>0$ for all k, for some $b>0$, and the bound in Theorem 3.8.1 becomes

$$\textsc{AvReg}(\textit{DPC-Batch}) \;=\; \sum_{k=1}^{K} O\left(\frac{\sqrt{b\cdot n_k^{c_k}\cdot\log n_k}}{n_1} + \frac{1}{n_k^{c_k}} + \frac{T}{n_k^{c_k+\frac{b}{\max\{1,\beta_k\}}}}\right).$$

Ignoring the logarithmic term in the bound above, an optimal c_k can be calculated by setting $n_k^{\frac{3}{2}c_k} = n_1$, or equivalently $c_k = \frac{2}{3} \cdot \frac{\log n_1}{\log n_k} := c_k^*$. Note that the number of batches in one service cycle for service type k under $c_k = c_k^*$ is approximately $n_k^{1-c_k^*} = n_1^{\frac{2}{3}(\frac{1}{c_k^*}-1)}$. Since the power term on n_1 is decreasing in c_k^* for all $c_k^* \in (0,1)$ and a larger n_k implies a smaller c_k^*, the service type with a longer service time requires a larger batch size and a more frequent price adjustments during one service cycle than the service type with a shorter service time. Overall, the average regret of DPC-Batch for the above scenario under c_k^* is of order $\frac{\sqrt{\log n_1}}{n_1^{2/3}}$.

3.9 Extension to Advance Service Bookings with Homogeneous Service Time Requirements

In this chapter, we consider a generalization of the basic model in Chapter 3.4 to the setting with advance service booking or scheduling. We first discuss the setting of the problem and then we provide a generalization of DPC-Batch.

3.9.1 The Setting

Similar to the basic model, the firm sells only a single service type where each request requires a single unit of resource and n units of service time. However, unlike in the basic model where a customer arriving in period t immediately starts her service in period t, she can now choose to start her service at time $t + \ell$, where $\ell \in [0, L]$. (For simplicity, we will call a request whose service starts ℓ periods later as type-ℓ request; this should not be confused with the meaning of "type" in the previous chapter.) The firm controls the arrival rates of all types of requests by setting a price vector $\mathbf{p}_t = (p_{t,0}, \dots, p_{t,L})$, where $p_{t,\ell}$ is the price of a type-ℓ request booked in period t (note that $p_{t,0}$ is the price of service that starts immediately in period t). Demand rates in period t is denoted by $\boldsymbol{\lambda}_t(\mathbf{p}_t) := (\lambda_{t,0}(\mathbf{p}_t), \dots, \lambda_{t,L}(\mathbf{p}_t))$. Let $\mathbf{D}_t(\mathbf{p}_t) = (D_{t,1}(\mathbf{p}_t), \dots, D_{t,L}(\mathbf{p}_t))$ denote the realized requests in period t (by definition, $\mathbf{E}[\mathbf{D}_t(\mathbf{p}_t)] = \boldsymbol{\lambda}_t(\mathbf{p}_t)$). By the invertibility assumption (see below), we can write the corresponding revenue rate as $r_t(\mathbf{p}_t) := \mathbf{p}_t^\top \boldsymbol{\lambda}_t(\mathbf{p}_t) = \boldsymbol{\lambda}_t^\top \mathbf{p}_t(\boldsymbol{\lambda}_t) = r_t(\boldsymbol{\lambda}_t)$. As in the basic model, we assume that demands across different periods are independent, though demands over different request types within the same period may be correlated, and at most one request arrives in each period, i.e., $\sum_{\ell=0}^{L} D_{t,\ell}(\mathbf{p}_t) \leq 1$. Let $\Omega_p = \otimes_{\ell=0}^{L} \Omega_{p,\ell}$ and $\Omega_\lambda = \otimes_{\ell=0}^{L} \Omega_{\lambda,\ell}$ denote the convex feasible set for price vector and demand rate vector, respectively. As in Chapter 3.4, we assume that MA1-MA4 hold. (Although the definition of service, or request, types in Chapter 3.8 and 3.9 are different, from the point of view of abstraction, the demand and revenue functions in Chapter 3.8 and 3.9 are essentially a multi-product variant of the functions in Chapter 3.4.)

The optimal stochastic control formulation of our dynamic pricing problem is given by:

$$\textbf{OPT-A}: \;\; J_A^* \;=\; \left\{ \max_{\pi\in\Pi} \; \mathbf{E}\left[\sum_{t=1}^{T} r_t(\mathbf{p}_t^\pi)\right] \; : \; \sum_{\ell=0}^{L} \; \sum_{s=\max\{1,\,t-n-\ell+1\}}^{t-\ell} D_{s,\ell}(\mathbf{p}_s^\pi) \leq C \;\forall t \right\}$$

where the constraints must hold almost surely (or with probability one) and Π is the set of all non-anticipating controls. Using demand rate vector as the decision variable, the deterministic relaxation of **OPT-A** is given by:

$$\textbf{DET-A}: \;\; J_A^D \;=\; \left\{ \max_{\boldsymbol{\lambda}_t\in\Omega_\lambda} \; \sum_{t=1}^{T} r_t(\boldsymbol{\lambda}_t) \; : \; \sum_{\ell=0}^{L} \; \sum_{s=\max\{1,\,t-n-\ell+1\}}^{t-\ell} \lambda_{s,\ell} \leq C \;\forall t \right\}$$

Let $\boldsymbol{\lambda}^D := (\boldsymbol{\lambda}_t^D)_{t=1}^T$ denote the optimal solution of **DET-A**, and let $\mathbf{p}^D := (\mathbf{p}_t(\boldsymbol{\lambda}_t^D))_{t=1}^T$ denote the corresponding price vectors. As in Chapter 3.8, we assume that MA 5 and MA 6 also hold for all t.

Lastly, we define our performance measure in the setting with advance booking as follows:

$$\textsc{AvReg}(\pi) \;=\; \frac{J_A^D - \mathbf{E}[R^\pi]}{T\cdot(L+1)}\,.$$

In the same spirit with Lemma 1, it is not difficult to show that $J_A^* \leq J_A^D$. However, unlike in the basic model in Chapter 3.4 where the expected total revenues under the optimal policy throughout T periods only scales linearly with T, the expected total revenues in the advance booking setting may scale linearly with $T\cdot(L+1)$, especially when T is large and the demand rate function $\lambda_{t,\ell}(\cdot)$ has the same order of magnitude for all t and ℓ (i.e., at any period t, we have the same intensity among customers who are requesting to start their service at period $t+\ell$ where $\ell = 0,1,\ldots,L$), because we are essentially collecting revenues from about $n\cdot(L+1)$ customers instead of n during each service cycle. This explains why we divide the expected total regrets with $T\cdot(L+1)$ instead of T in the above. We can alternatively interpret this as the average expected revenue loss *per* customer.

3.9.2 A Generalized DPC-Batch and Its Performance

Let $\{\mathcal{T}_i\}_{i=1}^{T/m}$ denote a partition of $[1, T]$, where $\mathcal{T}_i = [(i-1)m+1, im]$ for all $i \geq 1$. The key idea behind our generalized DPC-Batch with advance service booking is to correct the cumulative errors of type-ℓ request in the previous batch with the demands of type-ℓ request in the current batch. Let $\boldsymbol{\Delta}_t := (\Delta_{t,\ell})_{\ell=0}^{L} = (D_{t,\ell}(\mathbf{p}_t) - \lambda_{t,\ell}(\mathbf{p}_t))_{\ell=0}^{L}$ denote the vector of errors from expected demands in period t under price vector $\mathbf{p}_t$, where we suppress the dependency of $\boldsymbol{\Delta}_t$ on $\mathbf{p}_t$. For each t, let $i_\ell(t)$ be such that $\max\{t-\ell, 1\} \in \mathcal{T}_{i_\ell(t)}$. The complete definition of DPC-Batch with advance service booking is given below.

DPC-Batch with Parameters m and ϵ (DPC-Batch(m, ϵ))

Step 1. Solve **DET-A** and get $\boldsymbol{\lambda}^D$.

Step 2. At the beginning of each $t \geq 1$, do:

a. Compute $\hat{\mathbf{p}}_t^D$ according to

$$\lambda_{t,\ell}(\hat{\mathbf{p}}_t^D) \;=\; \lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{s \in \mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell} \quad \text{for all } \ell;$$

b. If $C_t \geq 1$ and $\lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m}\sum_{s \in \mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell} \in \Omega_{\lambda,\ell}$, set $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$; Otherwise, set $\mathbf{p}_t = \bar{\mathbf{p}}$.

The following theorem tells us the performance of DPC-Batch with advance service booking; we defer its proof to Appendix B.2.

Theorem 3.9.1 *The following two bounds hold for all C, L, n and m:*

1. *If $L \leq n$ and $\epsilon \in \left[1, \min\left\{n(L+1), m(L+1), n \cdot \frac{1+m(L+1)\min\{\varphi_L, \varphi_U\}}{8m+n}\right\}\right]$, there exists a constant $M_4 > 0$ such that*

$$\begin{aligned}&\textsc{AvReg}(\mathit{DPC\text{-}Batch}) \\ &\leq\; M_4 \cdot \left[\frac{\epsilon}{n(L+1)} + \frac{1}{m} + \frac{T}{m} \cdot \exp\left\{-\frac{(\epsilon-1)^2}{256(L+1)^2 \min\{C-\epsilon,\, m\}}\right\}\right]. \quad (3.16)\end{aligned}$$

2. *If $L > n$ and $\epsilon \in (L+1) \cdot \left[2, \min\left\{n, m, \frac{4mn\min\{\varphi_L, \varphi_U\}+2}{4m+n}\right\}\right]$, there exists a constant*

$M_4' > 0$ *such that*

$$\begin{aligned} &\textsc{AvReg}(\textit{DPC-Batch}) \\ &\leq\ M_4' \cdot \left[\frac{\epsilon}{n(L+1)} + \frac{1}{m} + \frac{T}{m} \cdot \exp\left\{-\frac{(\epsilon - 2(L+1)/n)^2}{64(L+1)^2 \min\{C-\epsilon,\, m\}}\right\}\right]. \end{aligned} \tag{3.17}$$

In particular, if $C = a \cdot nL$, $L = n^d$, $m = \lceil n^c \rceil$ *for some* $a > 0$, $d \geq 0$ *and* $c \in \left(\frac{\log\log n}{\log n}, 1\right)$, *we can bound the average regret of DPC-Batch as follows:*

1. *For* $d \leq 1$, *using* $\epsilon = 1 + 16\sqrt{b \cdot n^{2d+c} \cdot \log n}$ *for some* $b > 0$ *in bound (3.16) yields*

$$\textsc{AvReg}(\textit{DPC-Batch}) \;=\; O\left(\frac{\sqrt{b \log n}}{n^{1-\frac{c}{2}}} + \frac{1}{n^c} + \frac{T}{n^{c + \frac{b}{\max\{1,a\}}}}\right). \tag{3.18}$$

2. *For* $d > 1$, *using* $\epsilon = 2 \cdot \frac{L+1}{n} + 8\sqrt{b \cdot n^{2d+c} \cdot \log n}$ *for some* $b > 0$ *in bound (3.17) yields*

$$\textsc{AvReg}(\textit{DPC-Batch}) \;=\; O\left(\frac{\sqrt{b \cdot \log n}}{n^{1-\frac{c}{2}}} + \frac{1}{n^c} + \frac{T}{n^{c + \frac{b}{\max\{1,a\}}}}\right). \tag{3.19}$$

The two general bounds in Theorem 3.9.1 (i.e., (3.16) and (3.17)) are proved in a very similar manner under different requirements on ϵ and the relative magnitude of n and L. Together, they are the analogue of (3.8) in Theore 3.6.1 and holds for general problem parameters C, m, and n. Under slightly different choice of ϵ, the optimal order of (3.18) and (3.19) are both achieved when $c = 2/3$, which yields an average regret of order $\frac{\sqrt{\log n}}{n^{2/3}}$. Hence, Theorem 3.9.1 tells us DPC-Batch can be generalized to the setting with advance service bookings without worsening the performance.

3.10 Closing Remarks

In this paper, we address the dynamic pricing problem with reusable resources and deterministic service time requirements. Given the complexity of solving the stochastic control optimally, we focus on designing provably-good heuristic controls and evaluate their performances in the asymptotic regime. We also extend our result to the setting with heterogeneous service time requirements and advance booking length. Given its simplicity and generality, we believe that our heuristic controls can be tailored to address practical dynamic pricing problems faced by firms from various industries. Methodologically, our asymptotic analysis also shed lights on the difference between revenue management with reusable resources and deterministic service time requirements and the canonical revenue management problems. Many possible extensions are not addressed in this paper. For example, it is interesting to see how our analytical framework can be generalized to the setting with stochastic service time. Another potential future direction is to analyze the "network" version of our model, where resources can move dynamically between nodes, which is the setting of many on-demand ride sharing models such as UBER and Lyft.

CHAPTER 4

Near-Optimal Bisection Search for Nonparametric Dynamic Pricing with Inventory Constraint

4.1 Abstract

We consider a single-product revenue management problem with an inventory constraint and unknown, noisy, demand function. The objective of the firm is to dynamically adjust the prices to maximize total expected revenue. We restrict our scope to the *nonparametric* approach where we only assume some common regularity conditions on the demand function instead of a specific functional form. We propose a family of novel pricing heuristics that successfully balance the tradeoff between *exploration* and *exploitation*. The idea is to generalize the classic bisection search method to a problem that is affected both by stochastic noise and an inventory constraint. Our algorithm extends the bisection method to produce a sequence of pricing intervals that converge to the optimal static price with high probability. Using *regret* (the relative revenue loss compared to the optimal dynamic pricing solution for a clairvoyant) as the performance metric, we show that one of our heuristics exactly matches the theoretical asymptotic lower bound that has been previously shown to hold for any feasible pricing heuristic. Although the results are presented in the context of revenue management problems, our analysis of the bisection technique for stochastic optimization with learning can be potentially applied to other application areas.

4.2 Introduction

Dynamic pricing has became a common practice in many firms nowadays. It plays a central role in the revenue optimization of many industries including airlines, hotels, car rentals, and retails (Talluri and van Ryzin 2006, Özer and Phillips 2012). In the typical dynamic pricing problem, firms adaptively adjust their prices in response to market demand and try to maximize their expected revenue. The success of this approach relies heavily on the firms' knowledge about the relationship between market demand and the posted price, which is characterized by a demand function. Although in reality firms may not know the exact demand function, firms can still dynamically price their products through a combination of active learning (e.g., price experimentation) and dynamic optimization. The challenge, however, is obvious: Given the limited time window of opportunity and the limited on-hand inventory, firms have to balance the effort spent on probing the true demand function (exploration) and generating near-optimal revenue (exploitation).

The literature on dynamic pricing with demand learning can be broadly divided into two categories: *parametric* and *nonparametric* models. (See den Boer 2015 for a recent overview of the field.) In the parametric model, it is assumed that the firms know the functional form of the underlying demand function (e.g., linear, exponential, logit, etc.). The key challenges in such setting are to estimate the unknown demand parameters and to develop a price optimization scheme utilizing this estimate. Some popular estimation procedures that have been studied in the literature include Bayesian method (Araman and Caldentey 2009; Farias and van Roy 2010; Harrison et al. 2012), Maximum Likelihood estimation (Broder and Rusmevichientong 2012; den Boer 2014; den Boer and Zwart 2013; den Boer and Zwart 2015), and Least Squares approach (Keskin and Zeevi 2014). In contrast to parametric model, nonparametric model does not assume that the firms know the functional form of the demand function; instead, it only assumes a certain set of mild regularity conditions such as the decreasing property of demand as a function of price, the boundedness of the first and second derivatives of the demand function, and the unimodality of the revenue function. In such setting, the firms' tasks are further complicated by the fact that there is no explicit function to optimize.

Current literature suggests that parametric approaches outperform nonparametric approaches for general class of demand function, at least asymptotically. Given that parametric approach assumes a precise knowledge of the functional form of the underlying demand function, this observation is hardly surprising. Let $\theta > 0$ denote the relative size of the problem (i.e., the amount of initial inventory). A common way to evaluate the performance of a heuristic is to quantify the relationship between θ and the *regret*, which is the revenue loss compared to the optimal dynamic pricing policy for a clairvoyant (we will define it formally in Chapter 4.3). It is know that the information-theoretic lower bound on the regret is $\Omega(\sqrt{\theta})$ (see e.g. Wang et al. (2014)). Under the parametric model, this lower bound has been repeatedly shown to be tight under different scenarios using different heuristics; see e.g. Keskin and Zeevi (2014), den Boer and Zwart (2013), and Broder and Rusmevichientong (2012) in the setting without inventory constraints and Chen et al. (2017a) in the setting with inventory constraints. Under the nonparametric demand model, Wang et al. (2014) proposed a heuristic whose regret is on the order of $O(\sqrt{\theta}\log^{4.5}\theta)$ for a fairly general class of demand function. Under tighter regularity conditions (e.g. smoothness of demand function), Chen et al. (2018) proposed a heuristic whose regret matches the lower bound. Therefore, under mild regularity conditions on the demand functions, there is a performance gap between the parametric approaches and the nonparametric approaches, at least asymptotically.

The question is whether a parametric approach is always applicable in practice. To illustrate, suppose that the underlying demand function is actually a logit function. What will happen if we mistakenly assume a linear function instead of a logit function when estimating the demand parameters? As shown in Besbes and Zeevi (2015), although model mis-specification is not always detrimental, it can lead to suboptimal prices, which yield a large loss in revenue. It remains an open research problem whether there is a way to make parametric approach more robust with respect to model mis-specification for a general class of demand function. This leaves the firms in a quandary of having to choose between a parametric approach, with the risk of model mis-specification, or a nonparametric approach, with a weaker performance guarantee. The purpose of this paper is to address this issue. In particular, we will consider a nonparametric approach and study a scheme that will be shown to match the theoretical performance guarantee of the best known parametric approaches.

The proposed heuristics and their performances. Under uncertainty in demand information, a good pricing policy must balance the tradeoff between demand learning (exploration) and revenue maximization (exploitation) while also successfully dealing with the dynamics caused by stochastic demands and inventory constraints. Our heuristics achieve these objectives by generating a sequence of shrinking intervals that converge to the optimal static price calculated via a deterministic relaxation of the original dynamic pricing problem. More specifically, we generalize the standard bisection search algorithm to stochastic and constrained setting. (Our heuristics actually generalize the trisection search. However, for consistency with the existing optimization literature, we will simply call it a bisection instead of a trisection.) We use empirical mean of the observed demands as an estimate of the true demand rate to shrink the intervals accordingly. The sampling frequencies are chosen carefully: If they are too low, the resulting estimates are not very accurate; if, on the other hand, they are too high, we spend too much time on the sub-optimal prices, which incurs a large revenue loss.

For the single-product dynamic pricing problem, the implementation of our heuristics can be essentially divided into two phases: the exploration phase and the exploitation phase. Since it is known in this setting that the optimal static price can be written as the maximum of the unconstrained maximizer and the clearance price (see Gallego and van Ryzin 1994), the purpose of the exploration phase is to determine the identity of the optimal static price via bisection search. We show that it is possible to distinguish this identity quickly with a very high probability. During the exploitation phase, we apply another bisection search to more efficiently shrink the intervals according to the identity of the optimal price. We show that, if the heuristic uses bisection search methods in both phases, then the asymptotic regret is $O(\sqrt{\theta}\log\theta)$. This is already very close to the $\Omega(\sqrt{\theta})$ lower bound, and dominates the performance of the best known nonparametric scheme for single-product problem in Wang et al. (2014) under mild demand assumptions. It turns out that it is possible to remove the logarithm dependency in the upper bound completely: If we use Stochastic Approximation algorithms (i.e., Kiefer-Wolfowitz and Robbins-Monro, see Broadie et al. 2011) during the exploitation phase instead of another bisection search, then the resulting revenue loss is exactly $\Theta(\sqrt{\theta})$. Therefore, we have provided an asymptotically optimal nonparametric pricing heuristic for the setting of a single-product problem with inventory constraint.

Related literature. Apart from the standard parametric and nonparametric approaches, there are also works in the literature that consider robust optimization approach. Lim and Shanthikumar (2007) study a robust formulation of the classic single-product pricing problem where nature adversarially chooses the distribution governing the demand realization. They use a conservative max-min formulation that does not involve real-time demand learning and bears no closed form solution in general. Eren and Maglaras (2010) also study the robust setting and use a competitive ratio formulation. However, they only deal with the setting without inventory constraint and assume deterministic demand. Perakis and Roels (2010) adopt both the maximin and minimax formulation. Their focus is on deriving structural insights instead of proving a performance bound. As has been noted in Cohen et al. (2017b), the robust optimization literature mainly focuses on static problems and the previously realized uncertainty is not utilized to adjust the pricing decision; this may result in a rather conservative pricing decision. Cohen et al. (2017b) try to bridge the gap between robust approach and data-driven optimization by proposing algorithms that utilize the realized demands and converges to the optimal robust solution. However, there is no theoretical guarantee on the convergence rate of their algorithm. Rusmevichientong et al. (2006) also adopt a data-driven approach. They provide a bound on the number of samples required to guarantee a near-optimal revenue if one uses the empirical optimal price under general consumer choice model. Their approach is restricted to static setting, i.e., the pricing decision does not depend on the previously realized demand uncertainties. Therefore, there is no trade-off between revenue earning and demand learning.

On the technical side, our work is also related to three other streams of literature. The first one is the continuum-armed bandit literature (e.g., Agrawal 1995; Auer et al. 2007; Cope 2009; and Kleinberg 2004, Badanidiyuru et al. 2013). While there are some high-level connections between our approach and the bandit approach, our problems is fundamentally different from theirs because (i) there exists an inventory constraint, and (ii) the feasible pricing region is continuous. Another stream of related literature is the study of bisection search. Despite its long history and broad prevalence, there is little work that studies its generalization into stochastic setting. To the best of our knowledge, Waeber et al. 2013 is the only work that attempts to generalize the deterministic bisection search into a stochastic setting. However, the scope of their application is restricted to a root-finding problem. Thus, compared to the existing studies on bisection search method, our work is the only one that combines the challenge of stochastic setting and constrained optimization. These distinctions do not allow any direct comparison to the existing literature. Finally, our work is also related to the Online Convex Optimization (OCO) literature (see Cesa-Bianchi and Lugosi 2006 for a review). OCO considers a setting where at each time period, after a decision has been made, *nature* choose a cost function adversarially. The performance of a given policy is then compared to the policy that uses the best *static* action in hindsight. Although there are some similarities in the problem formulation, the vast majority of the OCO literature restricts its scope to convex cost functions and unconstrained setting; this clearly differentiates our work from OCO.

Remainder of this paper. The remainder of the paper is organized as follows. In Chapter 4.3, we introduce the problem formulation. In Chapter 4.4 and 4.5, we discuss our heuristics and prove their asymptotic bounds. Chapter 4.6. summarizes the paper and potential future research directions. Unless otherwise noted, the details of the proofs can be found in the Appendix.

4.3 Problem Formulation

In this chapter, we first describe the problem setting and discuss general modeling assumptions. We then introduce the deterministic analog of the original stochastic pricing problem and discuss our performance metric.

4.3.1 Model setting

We consider a monopolist selling a single product with C units of initial inventory. The selling horizon is discrete and divided into T periods. Without loss of generality, we assume that at most one customer arrives during each period. At the beginning of period t, the firm first posts the price p_t and in turn induces a stochastic demand $D_t(p_t)$ with a stationary rate $\lambda(p_t) = \mathbb{E}[D_t(p_t)]$. Note that, since at most one customer arrives during each period, the term $\lambda(p_t)$ can be interpreted as the probability of a purchase request during period t given p_t. Demands across different periods are assumed to be independent. Let $r(p) = p\lambda(p)$ denote the revenue rate and p^u its unique maximizer. Also, let Ω_p and Ω_λ denote the convex set of feasible prices and demand rates, respectively. We make the following assumptions on the underlying demand and revenue rate functions:

A1. The function $\lambda(\cdot) : \Omega_p \to \Omega_\lambda$ is invertible and twice-differentiable. Moreover, $\lambda(p)$ is strictly decreasing in p, i.e., there exists a constant $L > 0$ such that $|\lambda'(p)| \geq L$. We will use $p(\cdot) : \Omega_\lambda \to \Omega_p$ to denote the inverse of $\lambda(\cdot)$.

A2. The function $r(p)$ is strictly unimodal. In addition, $r(\lambda) := p(\lambda)\lambda$ is strictly concave in λ. (By abuse of notation, we will often write $r(\lambda)$ instead of $r(p)$ to denote the direct dependency of revenue on demand rate instead of price.)

A3. $\lambda(p)$ and $p(\lambda)$ are Lipschitz continuous with a factor $K > 0$, i.e., $\forall p, p' \in \Omega_p, |\lambda(p) - \lambda(p')| \leq K|p - p'|$, and $\forall \lambda, \lambda' \in \Omega_\lambda, |p(\lambda) - p(\lambda')| \leq K|\lambda - \lambda'|$.

A4. There exists a "shut-off" price p^∞ such that if $\{p^k\}$ is any price sequence satisfying $p^k \to p^\infty$, then we have $\lambda(p^k) \to 0$.

A5. There exist positive constants $M_L < M_U$ such that $0 > -M_L \geq r''(\lambda) \geq -M_U$ and $M_L|p - p^u| \leq |r'(p)| \leq M_U|p - p^u|$.

Assumptions A1-A4, together with the first part of A5, are quite natural and have been repeatedly used in the literature (cf. Besbes and Zeevi 2009, Wang et al. 2014). In particular, the existence of shut-off price allows the firm to effectively shut down the demand whenever desired. The second part of A5 is needed only for the analysis of Stochastic Approximation algorithms in Chapter 4.4.3. (They are standard assumptions in the Stochastic Approximation literature, e.g., Broadie et al. 2011.)

4.3.2 The stochastic and deterministic pricing problems

We say that a pricing policy $\pi := (p_t^\pi : 0 \le t \le T)$ is *non-anticipating* if the decision p_t^π at the beginning of period t only depends on past prices $\{p_s^\pi : 0 \le s < t\}$ and past demand observations $\{D_s(p_s^\pi) : 0 \le s < t\}$. Furthermore, we also say that a pricing policy π is *admissible* if $p_t^\pi \in \Omega_p$ for all t and π is non-anticipating. Let Π denote the set of all admissible pricing policies. The stochastic formulation of the dynamic pricing problem is given by

$$J^* = \max_{\pi \in \Pi} \mathbb{E}\left[\sum_{t=1}^{T} p_t^\pi \cdot D_t(p_t^\pi)\right] \text{ such that } \sum_{t=1}^{T} D_t(p_t^\pi) \le C \qquad a.s. \tag{4.1}$$

The deterministic analog of the above pricing problem is

$$J^D = \max_{p_t \in \Omega_p} \sum_{t=1}^{T} r(p_t) \text{ such that } \sum_{t=1}^{T} \lambda(p_t) \le C. \tag{4.2}$$

By assumption A1, the above deterministic problem can also be written as

$$J^D = \max_{\lambda_t \in \Omega_\lambda} \sum_{t=1}^{T} r(\lambda_t) \text{ such that } \sum_{t=1}^{T} \lambda_t \le C. \tag{4.3}$$

Let $\{p_t^D\}$ denote the unique optimal solution of (4.2); correspondingly, we also define $\lambda_t^D := \lambda(p_t^D)$. ($\lambda_t^D$ and p_t^D are uniquely determined since (4.3) is a concave optimization problem with linear constraint.) Since the demand function is time-homogeneous, it can be shown that $p_t^D = p^D$ for all t (see Gallego and van Ryzin 1994 for proof). Thus, the optimal deterministic price is static. For analytical tractability, we will assume that both p^D and p^u lie in a proper interior of Ω_p. We state this assumption formally below.

A6. There exists $0 < \underline{p} < \bar{p}$ such that such that p^D, $p^u \in [\underline{p}, \bar{p}] \subset \Omega_p$.

4.3.3 Performance metric and asymptotic setting

Let J^π denote the expected revenue earned under pricing policy π. It is known that J^D is an upper bound for the expected revenue under any admissible policy, i.e., $J^D \geq J^\pi$ for all $\pi \in \Pi$ (see Gallego and van Ryzin 1994 for proof, we omit the details). Thus, following the convention in the literature, as our performance metric, we will define the revenue loss of an admissible policy π as $\mathcal{R}^\pi = J^D - J^\pi$. Since it is typical for revenue management firms to sell a large inventory during a selling season, following the standard setting in the literature, in this paper we will consider a sequence of increasing problems where we scale both the size of the initial inventory level and the number of selling periods by a factor of $\theta > 0$. To be precise, the θ^{th} problem is parameterized by $(C_\theta, T_\theta) = (\theta C, \theta T)$. Let J_θ^D denote the optimal value of the deterministic problem (4.2) with scaling factor θ (it is not difficult to see that $J_\theta^D = \theta J^D$) and let J_θ^π denote the expected revenue under policy π for a problem with scaling factor θ. (Throughout this paper, the subscript θ will be consistently used as a reference to the problem with scaling factor θ.) Our objective is to study the asymptotic behavior of $\mathcal{R}_\theta^\pi = J_\theta^D - J_\theta^\pi$ as θ grows large. The scaling parameter θ can be interpreted as the size of the potential market, which is often large in the application of dynamic pricing. Ideally, we would expect that a good policy will have an expected revenue loss which grows relatively slowly with respect to θ. Notationwise, we will use $f(\theta) = O(g(\theta))$ to mean that $f(\theta) \leq M_1 g(\theta)$ for some constant $M_1 > 0$ and for all large n. Likewise, $f(\theta) = \Theta(g(\theta))$ means that there exist constants $0 < M_2 < M_3$ such that $M_2 g(\theta) \leq f(\theta) \leq M_3 g(\theta)$ for large enough n and $f(\theta) = \Omega(g(\theta))$ means that there exists a constant $M_4 > 0$ such that $f(\theta) \geq M_4 g(\theta)$ for all large n. For notational simplicity, whenever there is no confusion, we will often suppress the dependency on θ.

4.4 Main Results

In this chapter, we first introduce a generalization of the standard bisection search heuristic to a stochastic and constrained problem. We then discuss two improvements of the basic bisection heuristic to further reduce the asymptotic revenue loss bound. (The proofs of these results can be found in Chapter 4.5.)

4.4.1 Preliminary ideas

The departure point for the construction of our heuristics is a structural property of the optimal solution of the deterministic problem (4.2). It is known (e.g., Gallego and van Ryzin 1994) that the optimal deterministic policy is a static price control where the firms apply the same price $p^D = \max\{p^u, p^c\}$ until stock-out, where $p^c = \text{argmin}_{p \in \Omega_p} |\lambda(p) - C/T|$. For analytical tractability, we will assume that $\lambda(\bar{p}) < C/T$, which implies $p^c = p\,(C/T)$. (This is the original static price control in Gallego and van Ryzin 1994 and can be easily satisfied, for example, if the feasible set Ω_p is sufficiently large.) Intuitively, the static control prescribes that the firms apply the unconstrained optimal price if inventory is abundant, and the clearance price if inventory is scarce. If the firm knows p^D and applies it to the stochastic pricing problem until the inventory is depleted, then it incurs a revenue loss of order $O(\sqrt{\theta})$ (Gallego and van Ryzin 1994). Jasin (2014) show that this bound cannot be improved in general, i.e., the revenue loss of static price policy is $\Theta(\sqrt{\theta})$. Motivated by the good performance of static price policy in the case where p^D is known, one fruitful idea that has been exploited in the literature (e.g., Besbes and Zeevi 2009; Wang et al. 2014) is to design an algorithm whose resulting price sequence converges to p^D in the long run. In this paper, we will follow the same strategy and try to efficiently estimate p^D.

4.4.2 First Heuristic: Generalized Bisection Search

The key idea behind our first heuristic is to generalize the classical bisection search into a stochastic setting with constraint. Before presenting the complete algorithm for our heuristic, we first define a price experimentation subroutine that will be repeatedly used throughout the paper. We parametrize the subroutine with $I \subset [\underline{p}, \bar{p}]$ and $N \in \mathbb{R}$, where I denotes the sampling price range and N denotes the sampling frequency.

Bisection Sampling Subroutine. BiSamp(I, N)

a. Divide I into 3 intervals of equal length.

Let $S := \{p_l, l = 1, 2, 3, 4\}$ be the resulting endpoints of each interval.

b. For each l, apply p_l for N consecutive periods.

c. Compute the empirical mean rates

$$\begin{aligned}
\hat{r}(p_l) &= \frac{\text{total revenue received by applying } p_l}{N} \quad \text{and} \\
\hat{\lambda}(p_l) &= \frac{\text{total demand received by applying} p_l}{N}, \quad \text{for } l = 1, 2, 3, 4.
\end{aligned}$$

Note that $\hat{r}(\cdot)$ denotes the empirical revenue rate and $\hat{\lambda}(\cdot)$ denotes the empirical demand rate. The complete algorithm for our first heuristic is given below.

Bisection Dynamic Pricing Algorithm (BDPA).

Step 1: Initialization

Define $\underline{p}_1 = \underline{p}$, $\bar{p}_1 = \bar{p}$ and $I_1 = [\underline{p}_1, \bar{p}_1]$ to be the starting interval.

Step 2: Shrinking the Interval

For $k = 1, ..., \tau_\theta$, do:

a. Execute $\textsc{BiSamp}(I_k, N_{k,\theta})$ as long as the inventory level is still positive.

If the inventory is depleted, then apply p_∞ until time T_θ.

b. If $\hat{r}(p_{k,2}) < \hat{r}(p_{k,3})$, then define $\underline{p}_{k+1} = p_{k,2}$, $\bar{p}_{k+1} = p_{k,4}$;

If $\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})$ and $\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta}$, then define $\underline{p}_{k+1} = p_{k,1}$, $\bar{p}_{k+1} = p_{k,3}$;

If $\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})$ and $\hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}$, then define $\underline{p}_{k+1} = p_{k,2}$, $\bar{p}_{k+1} = p_{k,4}$;

If $\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})$ and $|\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}$, then define $\underline{p}_{k+1} = p_{k,2}$, $\bar{p}_{k+1} = p_{k,4}$;

c. Define the price range for the next iteration $I_{k+1} = [\underline{p}_{k+1}, \bar{p}_{k+1}]$.

Step 3: Applying Near-Optimal Static Price

Apply $\hat{p}^D_\theta = \frac{1}{2}(\underline{p}_{\tau_\theta+1} + \bar{p}_{\tau_\theta+1})$ until the end of selling horizon.

Apply p_∞ if inventory is depleted.

The above algorithm is defined by three groups of parameters: τ_θ, which is the total number of rounds of bisection search performed; $\Delta_{k,\theta}$, which serves as the tolerance level for stochastic error and will be elaborated in Chapter 4.5; and $N_{k,\theta}$, which denotes the sampling frequency. The value of these parameters must be carefully chosen. For example, if $N_{k,\theta}$ is too large, we would be spending too much time on sampling sub-optimal prices instead of converging to the optimal static price. If, on the other hand, $N_{k,\theta}$ is too small, we may not be able to accurately estimate the revenues and demand rates at different prices, which may lead to mis-identification of the optimal static price. If $\Delta_{k,\theta}$ is too large, we will not be able to know with a high enough probability whether certain price violates the capacity constraint; if $\Delta_{k,\theta}$ is too small, we will need to increase the sampling frequencies accordingly. Below, we provide an explicit choice of parameters that will be used in our analysis:

$$\begin{aligned} N_{k,\theta} &= \left\lceil \left(\frac{3}{2}\right)^{4k} \log^2 T_\theta \right\rceil, \;\; \Delta_{k,\theta} = \left(\frac{2}{3}\right)^{2k} \log^{-1/4} T_\theta, \;\; \textit{and} \\ \tau_\theta &= \sup \left\{ n \in \mathbb{N} \;:\; 4 \cdot \sum_{k=1}^{n} N_{k,\theta} \leq T_\theta \right\}, \end{aligned}$$

where $\lceil x \rceil = \inf\{y \geq x : \; y \in \mathbb{N}\}$. We make two observations: First, we define τ_θ to be the maximum number of full-rounds bisection search until the end of the selling season. Since the intervals generated by BDPA keep shrinking to the optimal static price with a high probability, such choice potentially has the smallest revenue loss. Second, the sampling frequencies $N_{k,\theta}$ are increasing in k, whereas the error tolerances $\Delta_{k,\theta}$ are decreasing in k. The reasoning behind these choices are intuitive: As the price interval shrinks, the revenue difference at two different prices within the interval decreases and yet the magnitude of stochastic noise does not change. Thus, more samples are needed to guarantee a more accurate estimate of the revenue rate, and smaller error tolerances are required. We state our first result below.

Theorem 4.4.1 *Under the aforementioned choice of parameters, we have:*

$$\mathcal{R}_\theta^{BDPA} = O\left(\theta^{3/4} \log^{1/2} \theta\right).$$

It is noteworthy that the performance guarantee in Theorem 4.4.1 is of the same order as the performance of nonparametric algorithm in Besbes and Zeevi (2009). This result, however, is not very satisfactory as there is still a big gap between the upper bound on the revenue loss and the theoretical lower bound of $\Omega(\sqrt{\theta})$. The reason behind this relatively poor performance is that BDPA tries to estimate p^u and p^c simultaneously and utilize the fact that p^D is the maximum of the two prices to estimate p^D. However, if we know the true identity of p^D, the original pricing problem can be simplified into either a unconstrained optimization problem (when $p^D = p^u$) or a root-finding problem (when $p^D = p^c$). Both problems can be solved in more loss-efficient manners than the original pricing problem. This enlightens us to first explore the identity of p^D, then exploit this identity using a more loss-efficient algorithm. The following two subsections are devoted to expanding this idea and achieving a better performance.

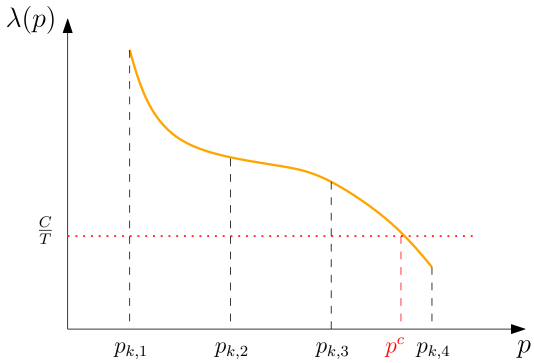


Figure 4.1: Deterministic Demand

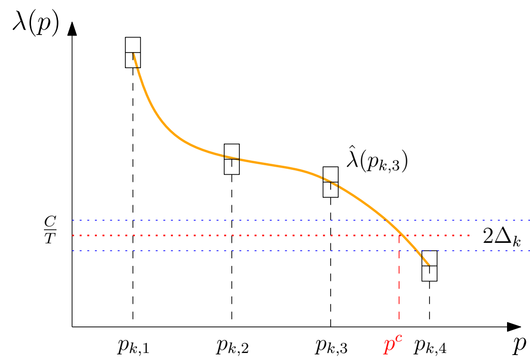


Figure 4.2: Stochastic Demand

Remark 4.4.1 *The iterative procedure in Step 2 helps us to shrink the size of price range while at the same time making sure that the new interval still contains the optimal static price. The key idea is to distinguish which of the three intervals does not contain p^u (or p^c) through revenue (or demand) rates comparison. To understand the reasoning behind the four scenarios in Step 2b, suppose that demand is deterministic and $p^D \in I_k$ for some $k \geq 1$. (In this case, the Bisection Sampling Routine gives us the true demand and revenue rate, i.e., $\hat{r}(p) = r(p)$, $\hat{\lambda}(p) = \lambda(p)$.) Now, if $r(p_{k,2}) < r(p_{k,3})$, by unimodality of $r(\cdot)$ we know that $p^u \geq p_{k,2}$. Then we know that $p^D = \max\{p^u, p^c\} \geq p_{k,2}$ and can safely delete $[p_{k,1}, p_{k,2})$ for the next round. This explains the intuition behind the first scenario. As for the second scenario, if $r(p_{k,2}) \geq r(p_{k,3})$, then $p^u \leq p_{k,3}$. Moreover, if $\lambda(p_{k,3}) < C/T - \Delta_{k,\theta}$, then $p^c \leq p_{k,3}$ (because $\lambda(\cdot)$ is decreasing). This implies that $p^D \leq p_{k,3}$ and, thus, we can safely delete $[p_{k,3}, p_{k,4})$ for the next round. If, on the other hand, $\lambda(p_{k,3}) \geq C/T - \Delta_{k,\theta}$, then for a sufficiently small $\Delta_{k,\theta}$, p^c belongs to a small region near $p_{k,3}$ such that $p^c \geq p_{k,2}$. Then we know $p^D = \max\{p^u, p^c\} \geq p_{k,2}$ and can safely delete $[p_{k,1}, p_{k,2})$ for the next round. This explains the intuition behind the third and fourth scenarios. If the demand observations are stochastic, as long as the empirical mean rates ($\hat{r}(\cdot)$ and $\hat{\lambda}(\cdot)$) are close enough to the true rates ($r(\cdot)$ and $\lambda(\cdot)$), we can infer the true order relationships with high probability. As an example, Figure 4.1 to 4.4 illustrate the intuition behind scenario 3. The black boxes in Figure 4.2 and 4.4 denote the ranges where $\hat{\lambda}(\cdot)$ and $\hat{r}(\cdot)$ fall with high probability, while Figure 4.1 and 4.3 show their respective deterministic counterparts. If N_k and Δ_k are well-chosen, the upper blue dotted line in Figure 4.2 will not cross the third box, and we can thus make correct prediction of the position of p^c. Also, in Figure 4.4, the prediction of the order relationship between $r(p_{k,2})$ and $r(p_{k,3})$ is correct as long as the middle two boxes do not overlap along the vertical axis. As a consequence, the shrinking strategy in stochastic setting (Figure 4.4) is the same with those in deterministic setting (Figure 4.3).*

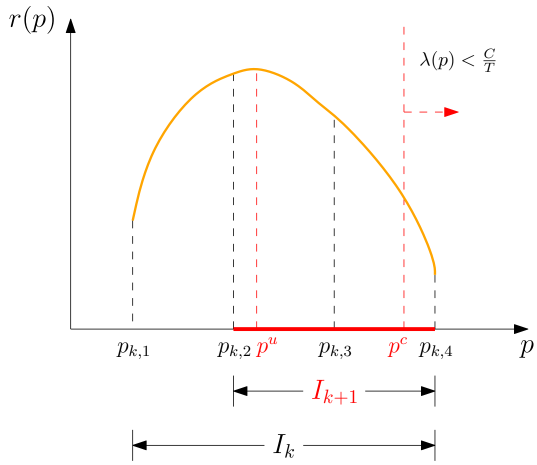


Figure 4.3: Deterministic Revenue

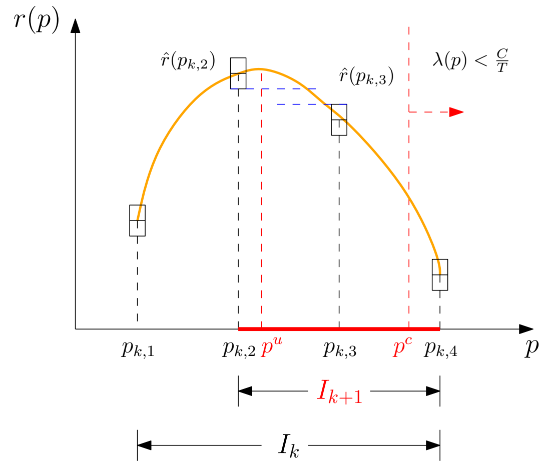


Figure 4.4: Stochastic Revenue

4.4.3 Second Heuristic: Double Bisection Search

It is important to note that, if $p^u \neq p^c$, then the functional behavior of $r(p)$ around p^u and p^c are different. To be precise, $r(p)$ is approximately quadratic around p^u and is approximately linear around p^c. This suggests that an efficient algorithm must take into account the distinction between p^u and p^c. Broadly speaking, our heuristics can be divided into two phases: (1) an *exploration* phase, during which we try to identify whether the optimal static price is p^u or p^c, and (2) an *exploitation* phase, during which we implement a more efficient search algorithm exploiting the identity of the optimal static price. For the exploration phase, we will use the generalized bisection search in BDPA. For the exploitation phase, we will use more efficient bisection search method depending on the identity of p^D distinguished by the exploration phase. The algorithm will accordingly generate a sequence of shrinking intervals that contain the optimal static price with a very high probability.

Double-Bisection Dynamic Pricing Algorithm (D-BDPA).

Step 1-2: Same as BDPA

Step 3: Identifying the Optimal Price

If $\hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}$, go to Step 4a; else, go to Step 4b.

Step 4a: Converge to p^u when $p^D = p^u > p^c$.

Define $I_1^u = [\underline{p}_1^u, \bar{p}_1^u] = I_{\tau_\theta+1}$. For $k = 1, ..., \tau_\theta^u$, do:

a. Execute $\textsc{BiSamp}(I_k^u, N_{k,\theta}^u)$.

b. If $\hat{r}(p^u_{k,2}) < \hat{r}(p^u_{k,3})$, then define $\underline{p}^u_{k+1} = p^u_{k,2}$, $\bar{p}^u_{k+1} = p^u_{k,4}$; else define $\underline{p}^u_{k+1} = p^u_{k,1}$, $\bar{p}^u_{k+1} = p^u_{k,3}$.

c. Define the price range for next iteration as $I^u_{k+1} = [\underline{p}^u_{k+1}, \bar{p}^u_{k+1}]$.

Apply $\hat{p}^D_\theta = \frac{1}{2}\left(\underline{p}^u_{\tau^u_\theta+1} + \bar{p}^u_{\tau^u_\theta+1}\right)$. If inventory is depleted, then apply p_∞.

Step 4b: Converge to p^c when $p^D = p^c \geq p^u$.

Define $I^c_1 = [\underline{p}^c_1, \bar{p}^c_1] = I_{\tau_\theta+1}$. For $k = 1, ..., \tau^c_\theta$, do:

a. Execute BiSamp$(I^c_k, N^c_{k,\theta})$.

b. If $\hat{\lambda}(p^c_{k,2}) > C/T + \Delta^c_{k,\theta}$, define $\underline{p}^c_{k+1} = p^c_{k,2}$, $\bar{p}^c_{k+1} = p^c_{k,4}$;
else, define $\underline{p}^c_{k+1} = p^c_{k,1}$, $\bar{p}^c_{k+1} = p^c_{k,3}$.

c. Define price range of next iteration $I^c_{k+1} = [\underline{p}^c_{k+1}, \bar{p}^c_{k+1}]$.

Apply $\hat{p}^D_\theta = \frac{1}{2}(\underline{p}^c_{\tau^c_\theta+1} + \bar{p}^c_{\tau^c_\theta+1})$. If inventory is depleted, then apply p_∞.

We introduce some more parameters: τ^u_θ, and τ^c_θ, which are the numbers of rounds of bisection search performed during exploitation phase (Step 4), respectively; $\Delta^c_{k,\theta}$, which serve as the tolerance level for stochastic error; and $N^c_{k,\theta}$ and $N^u_{k,\theta}$, which denote the sampling frequencies. As for the old parameters, we use the same $N_{k,\theta}$ and $\Delta_{k,\theta}$, but different τ_θ, since now the exploration phase only lasts for a few periods. Below, we provide an explicit choice of parameters which will be used in our analysis:

$$
\begin{aligned}
N^c_{k,\theta} &= \left\lceil \left(\frac{3}{2}\right)^{2k} \log^2 T_\theta \right\rceil, \quad N^u_{k,\theta} = \left\lceil \left(\frac{3}{2}\right)^{4k} \log^3 T_\theta \right\rceil, \quad \Delta^c_{k,\theta} = \left(\frac{2}{3}\right)^k \log^{-3/8} T_\theta, \\
\tau_\theta &= \sup\left\{ n \in \mathbb{N} \ : \ 4 \cdot \sum_{k=1}^n N_{k,\theta} \leq \log^3 T_\theta \right\}, \\
\tau^u_\theta &= \sup\left\{ n \in \mathbb{N} \ : \ 4 \cdot \sum_{k=1}^n N^u_{k,\theta} \leq T_\theta - 4 \cdot \sum_{k=1}^n N_{k,\theta} \right\}, \\
\tau^c_\theta &= \sup\left\{ n \in \mathbb{N} \ : \ 4 \cdot \sum_{k=1}^n N^c_{k,\theta} \leq T_\theta - 4 \cdot \sum_{k=1}^n N_{k,\theta} \right\},
\end{aligned}
$$

We make several observations here. First, we set τ_θ such that the length of the exploration phase does not exceed $\log^3 T_\theta$, which is relatively short for large θ. This means that only a small number of price experimentations are needed to correctly identify (with a very high probability) whether $p^D = p^u$ or $p^D = p^c$. Secondly, the definitions of τ_θ^u and τ_θ^c follow from the fact that, during the exploration phase, we try to perform as many full-rounds of bisection search as possible until the end of the selling season. Thirdly, the sampling frequencies $(N_{k,\theta}^u, N_{k,\theta}^c)$ and tolerance of error $(\Delta_{k,\theta}^c)$ are different in exploitation phase comparing with those parameters in exploration phase $N_{k,\theta}$. These along with different shrinking strategy provide better performance. We state our result regarding the performance of D-BDPA below.

Theorem 4.4.2 *Under the aforementioned choice of parameters, we have:*

$$\mathcal{R}_\theta^{D-BDPA} = O\left(\sqrt{\theta}\log\theta\right).$$

Theorem 4.4.2 tells us that D-BDPA is asymptotically optimal. Moreover, its performance guarantee dominates the performance guarantee of any existing nonparametric algorithm in the literature, including the $O(\sqrt{\theta}\log^{4.5}\theta)$ of Wang et al. (2014), and is very close to the known theoretical lower bound of $\Omega(\sqrt{\theta})$. In the next subsection we will show that if we replace the bisection search during the exploitation phase with Stochastic Approximation algorithm, then we can exactly match the lower bound.

4.4.4 Third Heuristic: Combining Bisection Search with Stochastic Approximation

Stochastic Approximation refers to a class of iterative stochastic optimization algorithms. We refer to Kushner and Yin (2003) for a comprehensive review. Broadly speaking, stochastic approximation algorithms can be divided into two different types: those that who try to solve a root-finding problem and those who try to stochastically estimate the maximum of a unimodal function. In this work, we consider the first and prototypical algorithms of this kind, i.e. Robbins-Monro (Robbins and Monro 1951) and Kiefer-Wolfowitz algorithms (Kiefer and Wolfowitz 1952). Let $R_t(p_t) = p_t \cdot D_t(p_t)$ denotes the realized revenue during period t under p_t, and define $P_X(x) = \arg\min_{y\in X} \; ||y-x||$ to be the geometric projection function. The complete description of the combined bisection search and Stochastic Approximation algorithm is given below.

SA-Bisection Dynamic Pricing Algorithm (SA-BDPA).

Steps 1 - 3: Same as D-BDPA

Step 4a: Converge to p^u when $p^u > p^c$. (Kiefer-Wolfowitz Scheme)
Let $p_1^u = \underline{p}_{\tau_\theta+1}$. For $k = 1, ..., \tau_\theta^u$, do:

a. Sample the revenue rate at price $p_k^u + c_k^u$ at period $4\sum_{k=1}^{\tau_\theta} N_k + 2k - 1$, and $p_k^u - c_k^u$
at period $4\sum_{k=1}^{\tau_\theta} N_k + 2k$ respectively; if inventory is depleted, apply p_∞.

b. Update the price according to

$$p_{k+1}^u = P_{I_{\tau_\theta+1}}\left[p_k^u + a_k^u \frac{R_k(p_k^u + c_k^u) - R_k(p_k^u - c_k^u)}{c_k^u}\right].$$

Step 4b: Converge to p^c when $p^c \geq p^u$. (Robbins-Monro Scheme)
Let $p_1^c = \underline{p}_{\tau_\theta+1}$. For $k = 1, ..., \tau_\theta^c$, do:

a. Sample the revenue rate at price p_k^c for one period; if inventory is depleted, apply p_∞.

b. Update the price according to

$$p_{k+1}^c = P_{I_{\tau_\theta+1}}\left[p_k^c + a_k^c \left(\frac{C}{T} - D_k(p_k^c)\right)\right].$$

Note that SA-BDPA is parameterized by τ_θ, $\Delta_{k,\theta}$, $N_{k,\theta}$, a_k^u, a_k^c, and c_k^u. (The a_k^u, a_k^c, and c_k^u are standard parameters in Stochastic Approximation algorithm, see Broadie et al. (2011).) We state a theorem.

Theorem 4.4.3 *Under the same choice of τ_θ, $\Delta_{k,\theta}$, and $N_{k,\theta}$ as in Theorem 4.4.1 and a proper choice of a_k^u, a_k^c, and c_k^u, we have:*

$$\mathcal{R}_\theta^{SA-BDPA} = O\left(\sqrt{\theta}\right). \tag{4.4}$$

It is noteworthy that Besbes and Zeevi (2009) also discuss a potential application of SA algorithms in their work. Specifically, they propose to apply the two types of SA schemes consecutively during the exploration phase to estimate p^u and p^c. At the end of the exploration phase, they propose that we choose the maximum of the two estimates and apply it during the remaining selling season until stock-out. The difference between their proposal and ours is obvious: They intend to use SA as an exploration algorithm while we use it as an exploitation algorithm. They conjecture that the revenue loss of their proposed SA-based dynamic pricing heuristic would be $O(\theta^{2/3})$, which is worse than ours.

4.5 Proof of Results

This chapter contains the proof of Theorem 4.4.2 and 4.4.3. We start by providing an outline of the proof in Chapter 4.5.1. The remaining details of the proof can be found in Chapter 4.5.2 - Chapter 4.5.7 and in the Appendix at the end of this paper. As for the proof of Theorem 4.4.1, since it is very similar with proof of Theorem 4.4.2, we only discuss the outline briefly in Chapter 4.5.1.

4.5.1 Outline of the Proofs and Key Lemmas

We first discuss the outline of the proofs. For analytical convenience, we will consider a slightly modified pricing policy called *Modified D-BDPA* (MD-BDPA) and *Modified SA-BDPA* (MSA-BDPA), respectively, which operate exactly as D-BDPA and SA-BDPA with the exception that it does not apply p_∞ when the seller runs out of inventory. Under MD-BDPA and MSA-BDPA, any excess demand beyond the available inventory can be outsourced at a unit price of $2\bar{p}$. Since $p_t < 2\bar{p}$ for all $p_t \in [\underline{p}, \bar{p}]$, obviously, we have $J^{MD-BDPA} \leq J^{D-BDPA}$ and $J^{MSA-BDPA} \leq J^{SA-BDPA}$. Thus, in order to bound $J^* - J^{D-BDPA}$ and $J^* - J^{SA-BDPA}$, it suffices that we compute a bound for each $J^* - J^{MD-BDPA}$ and $J^* - J^{MSA-BDPA}$. The outline of the proof of Theorems 4.4.2 and 4.4.3 is as follows:

1. Bounding the Probability of Converging to p^D in Step 2

We compute a lower bound for the probability that the optimal deterministic price p^D lies in I_k for all k in Step 2. This is critical to ensure that the final interval in the exploration phase contains p^D with a high probability. Define $E_1 := \cap_{k=1}^{\tau_\theta+1}\{p^D \in I_k\}$. We state a lemma.

Lemma 4.5.1 *Under the choice of parameters given in Chapter 4.5.2, there exists a constant* $C_1 > 0$ *independent of* $\theta \geq 1$ *such that* $P(E_1) \geq 1 - C_1 \frac{(\log\log\theta)^2}{\theta}$.

The proof of Lemma 4.5.1 can be found in Chapter 4.5.2. It is not difficult to show that, after τ_θ rounds of bisection search in Step 2, the length of the remaining feasible price interval is of order $\log^{-1/4}\theta$ (see Chapter 4.4.2). So, Lemma 4.5.1 tells us that, by the end of the exploration phase, we are already sufficiently "close" to the optimal price (not close enough for us to ignore the exploitation phase and simply apply fixed price throughout the remaining selling horizon as in Besbes and Zeevi (2009), but close enough for us to distinguish the identity of the optimal price).

2. Bounding the Probability of Distinguishing the Identity of p^D in Step 3

Once we guarantee that $p^D \in I_{\tau_\theta+1}$ with a high probability, we also need to guarantee that the action in Step 3 correctly distinguishes the identity of the optimal deterministic price with a high probability. If $p^D = p^u > p^c$, then we expect that the empirical demand rate at a point close to p^D will be much smaller than C/T. Similarly, if $p^D = p^c \geq p^u$, the empirical demand rate at a point close to p^D will be very close to C/T. Define $E_2 := \{\hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}\}$ if $p^u > p^c$ and $E_2 := \{\hat{\lambda}(\underline{p}_{\tau_\theta+1}) \geq C/T - \Delta_{\tau_\theta,\theta}\}$ otherwise. We state our second lemma.

Lemma 4.5.2 *Under the choice of parameters given in Chapter 4.4.2, there exists a constant* $C_2 > 0$ *independent of* $\theta \geq 1$ *such that* $P(E_1 \cap E_2) \geq 1 - C_2 \frac{(\log\log\theta)^2}{\theta}$.

The proof of Lemma 4.5.2 can be found in Chapter 4.5.3.

3. Bounding the Revenue Loss in Step 4

After we know the identity of p^D, we can then properly bound the revenue loss incurred during the exploitation phase. Note that, by definition of τ_θ, the total revenue loss incurred during the exploration phase is only $O(\log^3\theta)$. So, all that matters is the revenue loss incurred during the exploitation phase. In particular, by definition of $\pi \in \{$MD-BDPA, MSA-BDPA$\}$, we can write:

$$J_\theta^\pi \;=\; \mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] - 2\bar{p}\,\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right].$$

(Above, we suppress the notational dependency on π.) The bulk of the arguments in the rest of the analysis are in showing that

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \;=\; r(p^D)\,T_\theta - O(\sqrt{\theta}\log\theta) \quad \text{(for Theorem 4.4.2)}$$

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \;=\; r(p^D)\,T_\theta - O(\sqrt{\theta}) \quad \text{(for Theorem 4.4.3)}$$

$$\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \;=\; O(\sqrt{\theta}) \quad \text{(for Theorems 4.4.2 and 4.4.3),}$$

which completes the proof. We now briefly explain how D-BDPA achieves this order of performance. (See Chapter 4.5.4 and Chapter 4.5.5 for the parts regarding Theorem 4.4.2. We defer the proof of Theorem 4.4.3 in appendix since there are some similarities.) Assuming that the sequence of price intervals produced by D-BDPA converges to p^D, which happens with high probability. Since the exploration phase is relatively short, we can simply lower bound the collected revenue by zero. Now for the exploitation phase, notice that if p^u is the optimal static price, the revenue function is relatively "flat" near p^u in the sense that it is approximately quadratic (see Lemma 4.5.3(i)). Hence, to correctly distinguish the order relationship of the demand rates at two different prices, we need to sample more, i.e. $N^u_{k,\theta} = \Theta((\frac{3}{2})^{4k}\log^2\theta)$. On the other hand, the convergence of revenue rate around p^u can be shown to be quadratic (see Lemma 4.5.3(iii)). Now, assume without loss of generality that the selling season ends at the last period of the $(\tau^u_\theta)^{th}$ round of bisection search. Notice that $|I^u_k| = \Theta\left(\left(\frac{2}{3}\right)^{2k}\log^{-1/4}\theta\right)$ (see Chapter 4.5.4) and contains p^D with high probability, the revenue loss during Step 4a is on the order of

$$\begin{aligned} O\left(\sum_{k=1}^{\tau^u_\theta}\left(\frac{2}{3}\right)^{2k} N^u_{k,\theta}\right) &= O\left(\sum_{k=1}^{\tau^u_\theta}\left(\frac{3}{2}\right)^{2k}\log^{3/2}\theta\right) \\ = \; O\left(\left(\frac{3}{2}\right)^{2\tau^u_\theta}\log^{3/2}\theta\right) &= O(\sqrt{\theta}\log\theta), \end{aligned}$$

where the last inequality follows from Lemma 4.5.4. Now, if p^c is the optimal static price, the demand function is relatively "steep" near p^c in the sense that it is approximately linear (see Lemma 4.5.3(ii)). And accordingly we sample less frequently i.e.

$N^c_{k,\theta} = \Theta((\frac{3}{2})^{2k}\log^2\theta)$. However, the convergence of revenue rate around p^c can be shown to be linear (see Lemma 4.5.3(iii)), which is slower comparing with the case that $p^D = p^u$. Again, notice that $|I^c_k| = \Theta\left((\frac{2}{3})^{2k}\log^{-1/4}\theta\right)$ (see Chapter 4.5.4) and contains p^D with high probability, the revenue loss during Step 4b is on the order of

$$O\left(\sum_{k=1}^{\tau^c_\theta}\left(\frac{2}{3}\right)^k N^c_{k,\theta}\right) = O\left(\sum_{k=1}^{\tau^c_\theta}\left(\frac{3}{2}\right)^k \log^{7/4}\theta\right)$$
$$= O\left(\left(\frac{3}{2}\right)^{\tau^c_\theta}\log^{7/4}\theta\right) = O\left(\sqrt{\theta}\log\theta\right),$$

where the last inequality follows from Lemma 4.5.4.

Building upon the intuition, we briefly explain the intuition behind the order of the performance guarantee of BDPA. Notice that BDPA executes bisection search without distinguishing the identity of p^D. As a consequence, it has to sample with higher frequency ($N_{k,\theta} = N^u_{k,\theta} > N_{k,\theta}$, since $r(p)$ is flat around p^u) without knowing if the revenue convergence rate is quadratic ($p^D = p^u$) or linear ($p^D = p^c$). Then, if the optimal price is p^c, BDPA will clearly suffer from oversampling. Quantitatively speaking, the revenue loss of BDPA is of the order of

$$O\left(\sum_{k=1}^{\tau^{BDPA}_\theta} N_{k,\theta}\left(\frac{2}{3}\right)^k\right) = O\left(\sum_{k=1}^{\tau^{BDPA}_\theta}\left(\frac{3}{2}\right)^{3k}\log^2\theta\right)$$
$$O\left(\left(\frac{3}{2}\right)^{3\tau^{BDPA}_\theta}\log^2\theta\right) = O\left(\theta^{3/4}\log^{1/2}\theta\right),$$

where $\tau^{BDPA}_\theta = \sup\{n \in \mathbb{N} : 4\sum_{k=1}^n N_{k,\theta} \le T_\theta\}$ is the rounds of bisection search performed in BDPA and satisfies $(3/2)^{\tau^{BDPA}_\theta} = \Theta(\theta^{1/4}\log^{-1/2}\theta)$.

Below, we state two lemmas that will be repeatedly used in the proof.

Lemma 4.5.3 *(i) There exists a constant $K_u > 0$ such that for all $p_a, p_b \in [\underline{p}, \bar{p}]$, if $p^u > p_a > p_b$ (or $p_b > p_a > p^u$), then $r(p_a) - r(p_b) \ge K_u(p_a - p_b)^2$.*

(ii) For any $p_a, p_b \in [\underline{p}, \bar{p}]$, we have $|\lambda(p_a) - \lambda(p_b)| \ge L|p_a - p_b|$ for some positive constant L.

(iii) For any $p \in [\underline{p}, \bar{p}]$, we have $r(p^u) - r(p) \le \frac{M_u K^2}{2}(p^u - p)^2$ and $r(p^c) - r(p) \le (1 + 2K\bar{p})|p^c - p|$.

Lemma 4.5.4 *The following identities hold:* $\tau_\theta = \Theta(\log\log\theta)$, $\tau_\theta^u = \Theta(\log\theta)$, *and* $\tau_\theta^c = \Theta(\log\theta)$. *Moreover,*

$$\left(\frac{3}{2}\right)^{\tau_\theta} = \Theta\left(\log^{1/4}\theta\right), \ \left(\frac{3}{2}\right)^{4\tau_\theta^u} = \Theta\left(\frac{\theta}{\log^3\theta}\right), \quad and \quad \left(\frac{3}{2}\right)^{2\tau_\theta^c} = \Theta\left(\frac{\theta}{\log^2\theta}\right).$$

The first two parts of the first lemma tells us the "distinctiveness" of the revenue and demand function. They will provide useful guidelines for the choice of sampling frequencies. The third part of the first lemma provides upper bounds on the revenue loss depending on the identity of p^D. The second lemma quantifies the exact order of τ_θ, τ_θ^u, and τ_θ^c.

4.5.2 Proof of Lemma 4.5.1

By De Morgan's law and sub-additivity of probability measure, we have

$$P\left(\bar{E}_1\right) = P(\cup_{k=1}^{\tau_\theta+1}\{p^D \notin I_k\}) \leq \sum_{k=1}^{\tau_\theta+1} P(p^D \notin I_k),$$

where $\bar{E}$ is the complement of E. For $k > 1$, we can bound:

$$\begin{aligned}
& P(p^D \notin I_k) \\
= \ & P(p^D \notin I_k | p^D \in I_{k-1})P(p^D \in I_{k-1}) + P(p^D \notin I_k | p^D \notin I_{k-1})P(p^D \notin I_{k-1}) \\
\leq \ & P(p^D \notin I_k, \, p^D \in I_{k-1}) + P(p^D \notin I_{k-1}) \\
\leq \ & \cdots \\
\leq \ & \sum_{j=1}^{k-1} P(p^D \notin I_{j+1}, \, p^D \in I_j)
\end{aligned}$$

where the last inequality follows from $P(p^D \notin I_1) = 0$. Substituting them back into the bound for $P(\bar{E}_1)$ and using the fact that $P(p^D \notin I_1) = 0$, we get:

$$P\left(\bar{A}_1\right) \leq \sum_{k=2}^{\tau_\theta+1}\sum_{j=1}^{k-1} P(p^D \notin I_{j+1}, \, p^D \in I_j) = \sum_{k=1}^{\tau_\theta}(\tau_\theta - k + 1)P(p^D \notin I_{k+1}, \, p^D \in I_k).$$

We will now proceed to bound the term $P(p^D \notin I_{k+1}, \ p^D \in I_k)$ for $k = 1, ..., \tau_\theta$.

Define five groups of events $B_k^1, ..., B_k^5$ as follows:

$$
\begin{aligned}
B_k^1 &= \{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}),\, p^u < p_{k,2}\}, \\
B_k^2 &= \{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}),\, p^u > p_{k,3}\}, \\
B_k^3 &= \{\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta},\, p^c > p_{k,3}\}, \\
B_k^4 &= \{\hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta},\, p^c < p_{k,3}\}, \\
B_k^5 &= \{|\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta},\, p^c < p_{k,2}\}.
\end{aligned}
$$

We claim that:

$$
P(p^D \notin I_{k+1},\, p^D \in I_k) \leq \sum_{l=1}^{5} P(B_k^l),\ \forall k \tag{4.5}
$$

To prove this, first, note that, per the description of our algorithm, there are four different cases in Step 2(b) that we can enter in round k. So, we can bound:

$$
\begin{aligned}
& P\left(p^D \notin I_{k+1},\, p^D \in I_k\right) \\
\leq\ & P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}),\, p^D \notin I_{k+1},\, p^D \in I_k\right\} \\
& +P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3})\ ,\ \hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta},\, p^D \notin I_{k+1},\, p^D \in I_k\right\} \\
& +P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}),\, \hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta},\, p^D \notin I_{k+1},\, p^D \in I_k\right\} \\
& +P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}),\, |\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta},\, p^D \notin I_{k+1},\, p^D \in I_k\right\}.
\end{aligned}
$$

Now, if $p^D = p^u > p^c$, we have:

$$
\begin{aligned}
& P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}),\, p^u \notin I_{k+1},\, p^u \in I_k\right\} \\
&\qquad = P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}),\, p^u \in [p_{k,1}, p_{k,2}),\, p^u \in I_k\right\} \\
&\qquad \leq P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}),\, p^u < p_{k,2},\, p^u \in I_k\right\} \\
&\qquad \leq P\left(B_k^1\right);
\end{aligned}
$$

$$
\begin{aligned}
& P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}),\, \hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta},\, p^u \notin I_{k+1},\, p^u \in I_k\right\} \\
&\qquad \leq P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), p^u \in (p_{k,3}, p_{k,4}],\, p^u \in I_k\right\} \\
&\qquad \leq P\left(B_k^2\right);
\end{aligned}
$$

$$
\begin{aligned}
&P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), \hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^u \notin I_{k+1},\, p^u \in I_k\right\} \\
&\qquad = P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), \hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^u \in [p_{k,1}, p_{k,2}),\, p^u \in I_k\right\} \\
&\qquad \leq P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), \hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^u < p_{k,2}\right\} \\
&\qquad \leq P\left\{\hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^c < p_{k,2}\right\} \quad \text{(because } p^D = p^u > p^c\text{)} \\
&\qquad \leq P\left(B_k^4\right);
\end{aligned}
$$

$$
\begin{aligned}
&P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), |\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}, p^u \notin I_{k+1},\, p^u \in I_k\right\} \\
&\qquad = P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), |\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}, p^u \in [p_{k,1}, p_{k,2}),\, p^u \in I_k\right\} \\
&\qquad \leq P\left\{|\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}, p^c < p_{k,2}\right\} \quad \text{(because } p^D = p^u > p^c\text{)} \\
&\qquad \leq P\left(B_k^5\right).
\end{aligned}
$$

If, on the other hand, $p^D = p^c \geq p^u$, we have:

$$
\begin{aligned}
&P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}), p^c \notin I_{k+1},\, p^c \in I_k\right\} \\
&\qquad = P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}), p^c \in [p_{k,1}, p_{k,2}),\, p^c \in I_k\right\} \\
&\qquad \leq P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}), p^c < p_{k,2}\right\} \\
&\qquad \leq P\left\{\hat{r}(p_{k,2}) < \hat{r}(p_{k,3}), p^u < p_{k,2}\right\} \quad \text{(because } p^D = p^c \geq p^u\text{)} \\
&\qquad = P\left(B_k^1\right).
\end{aligned}
$$

$$
\begin{aligned}
&P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), \hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta}, p^c \notin I_{k+1},\, p^c \in I_k\right\} \\
&\qquad \leq P\left\{\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta}, p^c \in (p_{k,3}, p_{k,4}],\, p^c \in I_k\right\} \\
&\qquad \leq P\left(B_k^3\right).
\end{aligned}
$$

$$
\begin{aligned}
&P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), \hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^c \notin I_{k+1},\, p^c \in I_k\right\} \\
&\qquad = P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), \hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^c \in [p_{k,1}, p_{k,2}),\, p^c \in I_k\right\} \\
&\qquad \leq P\left\{\hat{\lambda}(p_{k,3}) > C/T + \Delta_{k,\theta}, p^c < p_{k,2}\right\} \\
&\qquad = P\left(B_k^4\right).
\end{aligned}
$$

$$
\begin{aligned}
&P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), |\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}, p^c \notin I_{k+1},\ p^c \in I_k\right\} \\
&\qquad = P\left\{\hat{r}(p_{k,2}) \geq \hat{r}(p_{k,3}), |\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}, p^c \in [p_{k,1}, p_{k,2}),\ p^c \in I_k\right\} \\
&\qquad \leq P\left\{|\hat{\lambda}(p_{k,3}) - C/T| \leq \Delta_{k,\theta}, p^c < p_{k,2}\right\} \\
&\qquad = P\left(B_k^5\right).
\end{aligned}
$$

Thus, in either case (i.e., $p^D = p^u > p^c$ or $p^D = p^c \geq p^u$), the bound in (4.5) holds. Put this together with our earlier bound for $P\left(\bar{A}_1\right)$, we get:

$$
P\left(\bar{A}_1\right) \leq \sum_{k=1}^{\tau_\theta} (\tau_\theta - k + 1)\left[\sum_{l=1}^{5} P(B_k^l)\right].
$$

To complete the proof of Lemma 4.5.1, it suffices that we compute a bound for $P(B_k^l)$ for $k = 1, ..., \tau_\theta$, $l = 1, ..., 5$, which is our remaining focus.

Upper bound for $\mathbf{P}(B_k^1)$ and $\mathbf{P}(B_k^2)$

The probabilities $P(B_k^1)$ and $P(B_k^2)$ can be bounded in a similar manner. So, we will only show how to bound $P(B_k^1)$. Fix $k \in \{1, ..., \tau_\theta\}$. Note that $p^u < p_{k,2} < p_{k,3}$ on B_1^k. Then by Lemma 4.5.3 part (ii), on B_k^1, $r(p_{k,2}) - r(p_{k,3}) \geq K_u(p_{k,3} - p_{k,2})^2 = K_u(\frac{|I|}{3})^2(\frac{2}{3})^{2(k-1)}$. Since $|\hat{r}(p_{k,l}) - r(p_{k,l})| < \frac{1}{4}K_u(\frac{|I|}{3})^2(\frac{2}{3})^{2(k-1)}$ for $l \in \{2, 3\}$ implies

$$
\begin{aligned}
\hat{r}(p_{k,2}) - \hat{r}(p_{k,3}) &= (r(p_{k,2}) - r(p_{k,3})) + (\hat{r}(p_{k,2}) - r(p_{k,2})) - (\hat{r}(p_{k,3}) - r(p_{k,3})) \\
&\geq (r(p_{k,2}) - r(p_{k,3})) - |\hat{r}(p_{k,2}) - r(p_{k,2})| - |\hat{r}(p_{k,3}) - r(p_{k,3})| \\
&> K_u\left(\frac{|I|}{3}\right)^2\left(\frac{2}{3}\right)^{2(k-1)} - \frac{2}{4}K_u\left(\frac{|I|}{3}\right)^2\left(\frac{2}{3}\right)^{2(k-1)} > 0,
\end{aligned}
$$

by Hoeffding's inequality (Hoeffding 1963), we can bound

$$
\begin{aligned}
P(B_k^1) &\leq P\left(|\hat{r}(p_{k,l}) - r(p_{k,l})| \geq \frac{1}{4}K_u\left(\frac{|I|}{3}\right)^2\left(\frac{2}{3}\right)^{2(k-1)} \quad \text{for some } l \in \{2, 3\}\right) \\
&\leq \sum_{l=2}^{3} P\left(|\hat{r}(p_{k,l}) - r(p_{k,l})| \geq \frac{1}{4}K_u\left(\frac{|I|}{3}\right)^2\left(\frac{2}{3}\right)^{2(k-1)}\right) \\
&\leq 4\exp\left(-2\,\frac{N_{k,\theta}\ \frac{1}{4^2}\ K_u^2\ \left(\frac{|I|}{3}\right)^4 \left(\frac{2}{3}\right)^{4(k-1)}}{\bar{p}^2}\right).
\end{aligned}
$$

By definition, $N_{k,\theta} = \Theta(\left(\frac{3}{2}\right)^{4k}\log^2\theta)$. So, for all sufficiently large θ, $P(B_k^1) \leq \frac{4}{\theta}$.

The same bound also holds for $P(B_k^2)$.

Upper bound for $\mathbf{P}(B_k^3)$ and $\mathbf{P}(B_k^4)$

The probabilities $P(B_k^3)$ and $P(B_k^4)$ can be bounded in a similar manner. So, we will only show how to bound $P(B_k^3)$. Note that $p^c > p_{k,3}$ implies $\lambda(p_{k,3}) > C/T$. So,

$$\begin{aligned} P(B_k^3) &\leq P\left(\hat{\lambda}(p_{k,3}) < C/T - \Delta_{k,\theta} \,, \lambda(p_{k,3}) > C/T\right) \\ &\leq P\left(\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3}) < -\Delta_{k,\theta}\right) \\ &\leq P\left(|\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3})| > \Delta_{k,\theta}\right). \end{aligned}$$

Again, by Hoeffding's inequality, since $\Delta_{k,\theta} = \Theta((\frac{2}{3})^{2k} \log^{-1/4}\theta)$ and $N_{k,\theta} = \Theta(\left(\frac{3}{2}\right)^{4k} \log^2\theta)$, for all large θ, we have $P(|\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3})| \geq \Delta_{k,\theta}) \leq 2\exp(-2N_{k,\theta}\Delta_{k,\theta}^2) \leq \frac{2}{\theta}$. The same bound also holds for $P(B_k^4)$.

Upper bound for $\mathbf{P}(B_k^5)$

By the decreasing property of demand function, $p^c < p_{k,2}$ implies $\lambda(p_{k,2}) \leq C/T$. By Lemma 4.5.3 part (i), $\lambda(p_{k,2}) - \lambda(p_{k,3}) \geq L|p_{k,2} - p_{k,3}| = L\frac{|I|}{3}(\frac{2}{3})^{k-1}$. So, on B_k^5,

$$\begin{aligned} \lambda(p_{k,3}) - \hat{\lambda}(p_{k,3}) &\leq \lambda(p_{k,2}) - L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} - \Delta_{k,\theta}\right) \\ &\leq \frac{C}{T} - L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} - \Delta_{k,\theta}\right) \\ &\leq -\frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}, \end{aligned}$$

where the last inequality follows because, by definition, $\Delta_{k,\theta} \leq \frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}$ for all sufficiently large θ. Now, by similar arguments as above,

$$\begin{aligned} P(B_k^5) &\leq P\left(\lambda(p_{k,3}) - \hat{\lambda}(p_{k,3}) < -\frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}\right) \\ &\leq P\left(|\hat{\lambda}(p_{k,3}) - \lambda(p_{k,3})| > \frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}\right) \\ &\leq 2\exp\left(-2N_k\left[\frac{1}{2}L\frac{|I|}{3}\left(\frac{2}{3}\right)^{k-1}\right]^2\right) \;\leq\; \frac{2}{\theta} \quad \text{(for all sufficiently large } \theta\text{)}. \end{aligned}$$

Put all the bounds together, we have

$$P\left(\bar{E}_1\right) \;\leq\; \sum_{k=1}^{\tau_\theta}(\tau_\theta - k + 1)\left[\sum_{l=1}^{5} P(B_k^l)\right] \;\leq\; \frac{\tau_\theta(\tau_\theta+1)}{2}\cdot\frac{4}{\theta}\cdot 5 \;=\; \frac{10\tau_\theta(\tau_\theta+1)}{\theta}.$$

Since $\tau_\theta = \Theta(\log\log\theta)$ (see Lemma 4.5.4), we conclude that there exists a constant C_1 such that

$$P(E_1) = 1 - P\left(\bar{E}_1\right) \geq 1 - C_1\frac{(\log\log\theta)^2}{\theta}. \qquad \square$$

4.5.3 Proof of Lemma 4.5.2

The proof is similar to that of Lemma 4.5.1. We will analyze the two cases (i.e., $p^D = p^u > p^c$ and $p^D = p^c \geq p^u$) separately.

Case 1: $p^D = p^c \geq p^u$

If $p^u \leq p^c$, then the optimal deterministic price p^D equals p^c. On E_1, we know that $p^c = p^D \in [\underline{p}_{\tau_\theta+1}, \bar{p}_{\tau_\theta+1}]$. This implies $\lambda(\underline{p}_{\tau_\theta+1}) \geq \lambda(p^c) = C/T$. So, we can bound:

$$\begin{aligned}
1 - P(E_1 \cap E_2) &= 1 - P(E_1) + P\left(E_1\right) - P\left(E_1 \cap E_2\right) \\
&= P\left(\bar{E}_1\right) + P\left(E_1 \cap \bar{E}_2\right) \\
&\leq P\left(\bar{E}_1\right) + P\left(p^c \in [\underline{p}_{\tau_\theta+1}, \bar{p}_{\tau_\theta+1}],\ \hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}\right) \\
&\leq P\left(\bar{E}_1\right) + P\left(\lambda(\underline{p}_{\tau_\theta+1}) \geq C/T,\ \hat{\lambda}(\underline{p}_{\tau_\theta+1}) < C/T - \Delta_{\tau_\theta,\theta}\right) \\
&\leq P\left(\bar{E}_1\right) + P\left(\hat{\lambda}(\underline{p}_{\tau_\theta+1}) - \lambda(\underline{p}_{\tau_\theta+1}) < -\Delta_{\tau_\theta,\theta}\right) \\
&\leq P\left(\bar{E}_1\right) + P\left(|\hat{\lambda}(\underline{p}_{\tau_\theta+1}) - \lambda(\underline{p}_{\tau_\theta+1})| > \Delta_{\tau_\theta,\theta}\right) \\
&\leq P\left(\bar{E}_1\right) + 2\exp\left(-2N_{\tau_\theta,\theta}\ \Delta^2_{\tau_\theta,\theta}\right) \quad \text{(by Hoeffding's inequality)} \\
&\leq C_1\frac{(\log\log\theta)^2}{\theta} + \frac{2}{\theta} \quad \text{(by Lemma 4.5.1)},
\end{aligned}$$

where the last inequality holds for all sufficiently large θ.

Case 2: $p^D = p^u > p^c$

If $p^u > p^c$, then $p^D = p^u$ and $\lambda(p^u) < \lambda(p^c) = C/T$. By definition of τ_θ, $|I_{\tau_\theta+1}|$ and $\Delta_{\tau_\theta,\theta}$ decrease to zero as $\theta \to \infty$. Since we always have $p^u = p^D \in [\underline{p}_{\tau_\theta+1}, \bar{p}_{\tau_\theta+1}]$ on E_1, it must also hold for all sufficiently large θ on E_1 that $p^c < \underline{p}_{\tau_\theta+1} < p^u$, $\lambda(\underline{p}_{\tau_\theta+1}) - \lambda(p^u) \leq (\lambda(p^c) - \lambda(p^u))/4$, and $\Delta_{\tau_\theta,\theta} \leq (\lambda(p^c) - \lambda(p^u))/4$. Arguing as in case 1, for all large θ, we can bound:

$$
\begin{aligned}
& 1-P(E_1\cap E_2) \\
= \; & P\left(\bar{E}_1\right)+P\left(E_1\cap \bar{E}_2\right) \\
\le \; & P\left(\bar{E}_1\right)+P\Bigg(\max\left\{\Delta_{\tau_\theta,\theta}, \lambda(\underline{p}_{\tau_\theta+1})-\lambda(p^u)\right\} \le \frac{\lambda(p^c)-\lambda(p^u)}{4}, \\
& \qquad\qquad\qquad\qquad\qquad\qquad \hat{\lambda}(\underline{p}_{\tau_\theta+1}) \ge \frac{C}{T}-\Delta_{\tau_\theta,\theta}\Bigg) \\
\le \; & P\left(\bar{E}_1\right)+P\Bigg(\max\left\{\Delta_{\tau_\theta,\theta}, \lambda(\underline{p}_{\tau_\theta+1})-\lambda(p^u)\right\} \le \frac{\lambda(p^c)-\lambda(p^u)}{4}, \\
& \qquad\qquad\qquad\qquad\qquad\qquad \hat{\lambda}(\underline{p}_{\tau_\theta+1}) \ge \lambda(p^c)-\Delta_{\tau_\theta,\theta}\Bigg) \\
\le \; & P\left(\bar{E}_1\right)+P\left(\hat{\lambda}(\underline{p}_{\tau_\theta+1})-\lambda(\underline{p}_{\tau_\theta+1}) \ge \frac{\lambda(p^c)-\lambda(p^u)}{2}\right) \\
\le \; & C_1\frac{(\log\log\theta)^2}{\theta}+\frac{1}{\theta},
\end{aligned}
$$

where the last inequality follows by Lemma 4.5.1 and Hoeffding's inequality (for sufficiently large θ).

Put the bounds from case 1 and case 2 together, we conclude that there exists a constant $C_2>0$ independent of $\theta\ge 1$ such that $P(E_1\cap E_2)\ge 1-C_2\frac{(\log\log\theta)^2}{\theta}$. $\square$

4.5.4 Bounding the Revenue Loss of D-BDPA Upon Entering Step 4a

Since $p^D=p^u>p^c$, for all sufficiently large θ, the following two conditions must hold: (i) $p^c\notin I_1^u$ and (ii) $r(p)$ is strictly concave in $I_1^u=I_{\tau_\theta+1}$. The first condition holds because p^u is strictly larger than p^c and the interval $I_{\tau_\theta+1}$ can be arbitrarily small for large θ. The second condition follows from the fact that $r(p)$ is locally strictly concave in the neighborhood of p^u (see Lemma 4.5.3 part (i)).

Let $E_u := \cap_{k=1}^{\tau_\theta^u}\{p^u\in I_k^u\}$. The following lemma is analogous to Lemma 4.5.1.

Lemma 4.5.5 *There exists a constant $C_3>0$ such that $P(E_1\cap E_2\cap E_u)\ge 1-C_3\frac{(\log\theta)^2}{\theta}$.*

We defer the proof of Lemma 4.5.5 to the appendix. Per our discussions in Chapter 4.5.1, the net revenue of MD-BDPA is the direct revenue minus the penalty, i.e.,

$$
J_\theta^{MD-BDPA} \;=\; \mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right]-2\bar{p}\;\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+\right].
$$

We will now proceed to bound the two expectations separately.

Step 1: Lower Bound for Direct Revenue Collected by MD-BDPA

We claim that there exists a constant $\tilde{C}_1 > 0$ such that

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \geq\ r(p^u)T_\theta - \tilde{C}_1\sqrt{\theta}\log\theta.$$

We focus our analysis on the sample path in $E_1 \cap E_2 \cap E_u$. Define $\tilde{T}^u_{\theta,1} = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta}$ and $\tilde{T}^u_{\theta,2} = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta} + \sum_{k=1}^{\tau^u_\theta} 4N^u_{k,\theta}$. The collected revenue can be lower bounded by two components as follows:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \\
\geq\ &\mathbb{E}\left[\sum_{t=1+\tilde{T}^u_{\theta,1}}^{\tilde{T}^u_{\theta,2}} p_t D_t(p_t)\mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
\geq\ &\mathbb{E}\left[\sum_{k=1}^{\tau^u_\theta}\sum_{l=1}^{4} N^u_{k,\theta}\ \hat{r}(p^u_{k,l})\mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&+\mathbb{E}\left[\left(T_\theta - \tilde{T}^u_{\theta,2}\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right].
\end{aligned} \tag{4.6}$$

For the first term, note that

$$\mathbb{E}\left[\sum_{k=1}^{\tau^u_\theta}\sum_{l=1}^{4} N^u_{k,\theta}\ \hat{r}(p^u_{k,l})\ \middle|\ E_1 \cap E_2 \cap E_u\right]\ =\ \sum_{k=1}^{\tau^u_\theta}\sum_{l=1}^{4} N^u_{k,\theta}\ \mathbb{E}\left[r(p^u_{k,l})|\ E_1 \cap E_2 \cap E_u\right].$$

Since on event $E_1 \cap E_2 \cap E_u$, $|p^u_{k,l} - p^u| \leq |I^u_1|(\frac{2}{3})^{k-1}$, then by Lemma 4.5.3(iii) we know that $r(p^u_{k,l}) \geq r(p^u) - \frac{9M_U K^2}{8}|I^u_1|^2\left(\frac{2}{3}\right)^{2k}$. Put this together with Lemma 4.5.5 and the fact that $\sum_{k=1}^{\tau_\theta^u} 4N_{k,\theta}^u \geq \tilde{T}_{\theta,2}^u - \log^3 T_\theta$, we have

$$
\begin{aligned}
&\sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \mathbb{E}\left[r(p_{k,l}^u)|E_1\cap E_2\cap E_u\right] P(E_1\cap E_2\cap E_u)\\
&\geq \left[\sum_{k=1}^{\tau_\theta^u} 4N_{k,\theta}^u\left(r(p^u)-\frac{9M_UK^2}{8}|I_1^u|^2\left(\frac{2}{3}\right)^{2k}\right)\right]\left(1-C_3\frac{(\log\theta)^2}{\theta}\right)\\
&\geq r(p^u)\left(\tilde{T}_{\theta,2}^u-\log^3 T_\theta\right)- C_3\ \bar{p}\ \frac{\log^2\theta}{\theta}\left(\sum_{k=1}^{\tau_\theta^u}4N_{k,\theta}^u\right)\\
&\qquad -\frac{9M_UK^2}{8}|I_1^u|^2\left[\sum_{k=1}^{\tau_\theta^u}4N_{k,\theta}^u\left(\frac{2}{3}\right)^{2k}\right]\\
&\geq r(p^u)\tilde{T}_{\theta,2}^u-\bar{p}\log^3 T_\theta - C_3\bar{p}T\log^2\theta-\frac{81}{10}M_UK^2|I_1^u|^2\log^3 T_\theta\left(\frac{3}{2}\right)^{2(\tau_\theta^u+1)}\\
&\geq r(p^u)\tilde{T}_{\theta,2}^u-\Theta(\sqrt{\theta}\log\theta),
\end{aligned}
$$

where the last inequality follows because $|I_1^u| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{4\tau_\theta^u} = \Theta\left(\frac{\theta}{\log^3\theta}\right)$ (see Lemma 4.5.4).

As for the second term in the RHS of (4.6), by the same arguments as above,

$$
\begin{aligned}
&\mathbb{E}\left[\left(T_\theta-\tilde{T}_{\theta,2}^u\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1\cap E_2\cap E_u\}\right]\\
&\geq \left(T_\theta-\tilde{T}_{\theta,2}^u\right)\left(r(p^u)-\frac{9M_UK^2}{8}|I_1^u|^2\left(\frac{2}{3}\right)^{2\tau_\theta^u}\right)\left(1-C_3\frac{(\log\theta)^2}{\theta}\right)\\
&\geq r(p^u)\left(T_\theta-\tilde{T}_{\theta,2}^u\right)- C_3\ \bar{p}\ T\log^2\theta-\frac{9M_UK^2}{8}|I_1^u|^2\ T_\theta\left(\frac{2}{3}\right)^{2\tau_\theta^u}\\
&\geq r(p^u)\left(T_\theta-\tilde{T}_{\theta,2}^u\right)-\Theta(\sqrt{\theta}\log\theta),
\end{aligned}
$$

where the last inequality follows because $|I_1^u| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{4\tau_\theta^u} = \Theta\left(\frac{\theta}{\log^3\theta}\right)$. Put the bounds for the two terms together proves our initial claim.

Step 2: Upper Bound for Total Penalty Incurred by Capacity Violation

We claim that there exists a constant $\tilde{C}_2 > 0$ such that

$$2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t)-C_\theta\right)^+\right] \leq \tilde{C}_2\sqrt{\theta}.$$

We first analyze the sample path on $E_1 \cap E_2 \cap E_u$. We know that

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\quad \leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\qquad + \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} \lambda(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\quad \leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \mathbb{E}\left[\left(\sum_{t=1}^{\tilde{T}_{\theta,1}^u} \lambda(p_t) - \tilde{T}_{\theta,1}^u \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\qquad + \mathbb{E}\left[\left(\sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \lambda(p_{k,l}^u) - \left(\tilde{T}_{\theta,2}^u - \tilde{T}_{\theta,1}^u\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\qquad + \mathbb{E}\left[\left(\sum_{t=\tilde{T}_{\theta,2}^u+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - \tilde{T}_{\theta,2}^u\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\quad \leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \tilde{T}_{\theta,1}^u \\
&\qquad + \mathbb{E}\left[\left(\sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \left(\lambda(p_{k,l}^u) - \frac{C}{T}\right)\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\qquad + \mathbb{E}\left[\left(\sum_{t=\tilde{T}_{\theta,2}^u+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - \tilde{T}_{\theta,2}^u\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\quad \leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] + \log^3 T_\theta \\
&\qquad + \sum_{k=1}^{\tau_\theta^u}\sum_{l=1}^{4} N_{k,\theta}^u \,\mathbb{E}\left[\left(\lambda(p_{k,l}^u) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\qquad + \mathbb{E}\left[\left(T_\theta - \tilde{T}_{\theta,2}^u\right)\left(\lambda\left(\hat{p}^D\right) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right],
\end{aligned}$$

where the first and second inequalities follow from Jensen's Inequality; the third inequality follows from the boundedness of demand observation and the definition of $\tilde{T}^u_{\theta,1}$, $\tilde{T}^u_{\theta,2}$; the last inequality follows from Jensen's Inequality and the definition of τ_θ. Basically, we break the capacity violation into four parts: stochastic randomness, and the capacity violation during Step 2, during bisection search in Step 4 and applying $\hat{p}^D$ in Step 4.

By Cauchy-Schwarz's inequality and the boundedness of demand observation, the first term can be easily bounded as follows:

$$
\begin{aligned}
\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D(p_t) - \lambda(p_t)\right)^+\right] &\leq \left\{\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D(p_t) - \lambda(p_t)\right)^2\right]\right\}^{1/2} \\
= \left\{\sum_{t=1}^{T_\theta} \mathbb{E}\left[(D(p_t) - \lambda(p_t))^2\right]\right\}^{1/2} &\leq \sqrt{T_\theta}.
\end{aligned}
$$

As for the third term, since $p^u > p^c$, which implies $\lambda(p^u) < \lambda(p^c) = C/T$, and $p^c \notin I^u_k$ for all k (for all large θ), we always have $\lambda(p^u_{k,l}) < \lambda(p^c) = C/T$. So, $(\lambda(p^u_{k,l}) - C/T)^+ = 0$ for all k and l. Similarly, since $\hat{p}^D \in I^u_{\tau^u_\theta+1}$, we have $\lambda(\hat{p}^D) < C/T$ for all large θ. So, the last term also equals to 0. Put the bounds together we have:

$$
\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] = O(\sqrt{\theta}).
$$

Thus, the total penalty for capacity violation satisfies

$$
\begin{aligned}
&2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \\
&\quad = 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_u\}\right] \\
&\qquad + 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2 \cap E_u}\}\right] \\
&\quad \leq 2\bar{p}\ O\left(\sqrt{\theta}\right) + 2\bar{p}\ T_\theta P\left(\overline{E_1 \cap E_2 \cap E_u}\right) = O(\sqrt{\theta}),
\end{aligned}
$$

where the last inequality follows the boundedness of demand observation.

Finally, combining our results from **Steps 1** and **2** above we conclude that

$$J_\theta^{MD-BDPA} \;\geq\; r(p^u)T_\theta - \tilde{C}_1\sqrt{\theta}\log\theta - \tilde{C}_2\sqrt{\theta} \;=\; r(p^u)T_\theta - O(\sqrt{\theta}\log\theta). \quad \square$$

4.5.5 Bounding the Revenue Loss of D-BDPA Upon Entering Step 4b

The proof is similar to those in Chapter 4.5.2. Let $E_c = \cap_{k=1}^{\tau_\theta^c}\{p_c \in I_k^c\}$. The following lemma is the analog of Lemma 4.5.5.

Lemma 4.5.6 *There exists a constant $C_4 > 0$ such that $P\left(\,\overline{E_1 \cap E_2 \cap E_c}\,\right) \geq 1 - C_4\frac{(\log\theta)^2}{\theta}$.*

We defer the proof of Lemma 4.5.6 to the appendix. We again consider MD-BDPA. The net revenue generated by MD-BDPA is given by:

$$J_\theta^{MD-BDPA} \;\;\geq\;\; \mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] - 2\bar{p}\;\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right].$$

Step 1: Lower Bound for Direct Revenue Collected by MD-BDPA

We claim that there exists a constant $\tilde{C}_3 > 0$ such that

$$\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \geq r(p^c)T_\theta - \tilde{C}_3\sqrt{\theta}\log\theta.$$

The proof is similar to Step 1 in Chapter 4.5.1. We break up the revenue on the sample path of $E_1 \cap E_2 \cap E_c$ into two parts:

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{T_\theta} p_t D_t(p_t)\right] \;\;\geq\;\; & \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c\;\hat{r}(p_{k,l}^c)\mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\ & + \mathbb{E}\left[\left(T_\theta - \tilde{T}_{\theta,2}^c\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right], \qquad (4.7)\end{aligned}$$

where $\tilde{T}_{\theta,1}^c = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta}$ and $\tilde{T}_{\theta,2}^c = \sum_{k=1}^{\tau_\theta} 4N_{k,\theta} + \sum_{k=1}^{\tau_\theta^c} 4N_{k,\theta}^c$. For the first term, note that

$$\mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c \hat{r}(p_{k,l}^c)\middle| E_1\cap E_2\cap E_c\right] \;=\; \sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c\, \mathbb{E}\left[r(p_{k,l}^c)\middle| E_1\cap E_2\cap E_c\right].$$

Since on $E_1\cap E_2\cap E_c$, $|p_{k,l}^c - p^c| \le \frac{3|I_1^c|}{2}(\frac{2}{3})^k$, by Lemma 4.5.3(iii) we know that $r(p^c)-r(p_{k,l}^c) \le \frac{3}{2}(1+2K\bar{p})|I_1^c|\left(\frac{2}{3}\right)^k$. Put this together with Lemma 4.5.6 and the fact that $\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c \ge \tilde{T}_{\theta,2}^c - \log^3 T_\theta$ we have

$$\begin{aligned}
&\sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c\, \mathbb{E}\left[r(p_{k,l}^c)\middle| E_1\cap E_2\cap E_c\right] P(E_1\cap E_2\cap E_c)\\
&\quad\ge\; \sum_{k=1}^{\tau_\theta^c} 4N_{k,\theta}^c\left[r(p^c)-\frac{3}{2}(1+2K\bar{p})|I_1^c|\left(\frac{2}{3}\right)^k\right]\left[1-C_4\frac{\log^2\theta}{\theta}\right]\\
&\quad\ge\; r(p^c)(\tilde{T}_{\theta,2}^c-\log^3 T_\theta) - \bar{p}\; C_4\frac{\log^2\theta}{\theta}\left(\sum_{k=1}^{\tau_\theta^c} 4N_{k,\theta}^c\right)\\
&\qquad -\;\frac{3}{2}(1+2K\bar{p})|I_1^c|\left[\sum_{k=1}^{\tau_\theta^c} 4N_{k,\theta}^c\left(\frac{2}{3}\right)^k\right]\\
&\quad\ge\; r(p^c)\tilde{T}_{\theta,2}^c - \bar{p}\log^3 T_\theta - \bar{p}TC_4\log^2\theta - 18(K\bar{p}+1)|I_1^c|\log^2 T_\theta\left(\frac{3}{2}\right)^{\tau_\theta^c}\\
&\quad= r(p^c)\tilde{T}_{\theta,2}^c - O(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality follows since $|I_1^c| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{2\tau_\theta^c} = \Theta\left(\frac{\theta}{\log^2\theta}\right)$, or equivalently $\left(\frac{3}{2}\right)^{\tau_\theta^c} = \Theta\left(\frac{\sqrt{\theta}}{\log\theta}\right)$. (See Lemma 4.5.4)

As for the second term in the RHS of (4.7), by the same argument as above,

$$\begin{aligned}
&\mathbb{E}\left[\left(T_\theta-\tilde{T}_{\theta,2}^c\right)\hat{r}(\hat{p}^D)\mathbf{1}\{E_1\cap E_2\cap E_c\}\right]\\
&\quad\ge\; \left(T_\theta-\tilde{T}_{\theta,2}^c\right)\left[r(p^c)-\frac{3}{2}(1+2K\bar{p})|I_1^c|\left(\frac{2}{3}\right)^{\tau_\theta^c}\right]\left(1-C_4\frac{\log^2\theta}{\theta}\right)\\
&\quad\ge\; r(p^c)\left(T_\theta-\tilde{T}_{\theta,2}^c\right) -\; C_4\;\bar{p}\;T\log^2\theta - \frac{3}{2}(1+2K\bar{p})|I_1^c|\;T_\theta\left(\frac{2}{3}\right)^{\tau_\theta^c}\\
&\quad\ge r(p^c)\left(T_\theta-\tilde{T}_{\theta,2}^u\right) - O(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality follows since $|I_1^c| = \Theta(\log^{-1/4}\theta)$ and $\left(\frac{3}{2}\right)^{2\tau^c} = \Theta\left(\frac{\theta}{\log^2\theta}\right)$. Put the bounds for the two terms in together proves the initial claim.

Step 2: Upper Bound for Total Penalty Incurred by Capacity Violation

We claim that there exists a constant $\tilde{C}_4 > 0$ such that

$$2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \leq \tilde{C}_4\sqrt{\theta}\log\theta. \tag{4.8}$$

We first analyze the sample path on $E_1 \cap E_2 \cap E_c$. We break the amount of capacity violation into several different parts. Following the same arguments as in Step 2 in Chapter 4.5.2,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\
&\leq \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] \\
&\quad + \log^3 T_\theta + \sum_{k=1}^{\tau_\theta^c}\sum_{l=1}^{4} N_{k,\theta}^c\ \mathbb{E}\left[\left(\lambda(p_{k,l}^c) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\
&\quad + \mathbb{E}\left[\left(T_\theta - \tilde{T}_{\theta,2}^c\right)\left(\lambda\left(\hat{p}^D\right) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right].
\end{aligned}$$

By Cauchy-Schwarz's inequality again, the first term can be upper bounded by $\sqrt{T_\theta}$. Then since for the sample paths on event $E_1 \cap E_2 \cap E_c$, $|\lambda(p_{k,l}^c) - \lambda(p^c)| \leq \frac{3}{2}K|I_1^c|(\frac{2}{3})^k$ for all k and l and $\lambda(p^c) = C/T$, we can bound

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\
&\leq \sqrt{T_\theta} + \log^3 T_\theta + \sum_{k=1}^{\tau_\theta^c} 4 \cdot N_{k,\theta}^c \cdot \frac{3}{2}K|I_1^c|\left(\frac{2}{3}\right)^k + \left(T_\theta - \tilde{T}_{\theta,2}^c\right)\cdot\frac{3}{2}K|I_1^c|\left(\frac{2}{3}\right)^{\tau_\theta^c} \\
&\leq \sqrt{T_\theta} + \log^3 T_\theta + 18K|I_1^c|\left(\frac{3}{2}\right)^{\tau_\theta^c}\log^2 T_\theta + \frac{3}{2}KT_\theta|I_1^c|\left(\frac{2}{3}\right)^{\tau_\theta^c} \\
&= O(\sqrt{\theta}\log\theta),
\end{aligned}$$

where the last inequality the same argument as in **Step 1** above.

Then, the total penalty for capacity violation satisfies

$$\begin{aligned}
&2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \\
&= 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2 \cap E_c\}\right] \\
&\quad + 2\bar{p}\ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2 \cap E_c}\}\right] \\
&\le O\left(\sqrt{\theta}\log\theta\right) + 2\bar{p}\ T_\theta P(\overline{E_1 \cap E_2 \cap E_c}) = O(\sqrt{\theta}\log\theta).
\end{aligned}$$

Finally, combining our results from **Step 1** and **2** above we have

$$J_\theta^{MD-BDPA} \ \ge\ r(p^u)T_\theta - \tilde{C}_3\sqrt{\theta}\log\theta - \tilde{C}_4\sqrt{\theta}\log\theta \ =\ r(p^c)T_\theta - O(\sqrt{\theta}\log\theta). \quad \square$$

4.6 Closing Remarks

This paper presents a scheme of nonparametric dynamic pricing with demand learning. Our scheme generalizes the classical bisection search algorithm into a stochastic setting with a constraint. We show that the performance of one of our heuristics exactly matches the theoretical lower bound for any feasible pricing policy. Thus, we have closed the gap (in asymptotic sense) between the performance of parametric approach and nonparametric approach for the single product problem.

There are several possible extensions of this work. One important direction is a generalization to the multiproduct setting. Although we have focused our analysis in the paper only on the single product setting, it is an open question whether our bisection search heuristic can also be applied to multiproduct problem. There are at least two challenges for such an extension: First, it is not immediately clear how to do bisection in high dimensional spaces. To the best of our knowledge, there is no existing literature on applying bisection search to multidimensional constrained optimization problem, even in the deterministic setting. Second, in multiproduct setting, nonparametric approach might suffer from *curse of dimensionality*, since it has to estimate a multidimensional function. In fact, the order of the revenue loss of the best known nonparametric scheme for multiproduct setting depends on the number of products in a non-trivial way (cf. Besbes and Zeevi 2012). It is curious to see whether applying bisection search algorithm to multiproduct setting can reduce the curse of dimensionality on revenue loss.

Additionally, throughout the paper, we have assumed that the demand function is stationary, i.e., it does not vary with time. In reality however, this assumption might not hold, which suggests that a good pricing heuristic should ideally take into account this possibility in its learning algorithm. The challenge, however, is obvious. For dynamic pricing with non-stationary demand, it is no longer true that the optimal solution to the deterministic problem is static pricing. This limits the ability to exploit the structure of the optimal solution, as we did in this paper. Actually, all of the works in non-stationary setting (Besbes et al. 2015, Keskin and Zeevi 2016a) consider only the problem without inventory constraint. Moreover, it is not clear how one can generalize the bisection search heuristic to non-stationary setting. Obviously this is an important research topic; we leave this as future research project.

CHAPTER 5

Conclusion

This dissertation studies the design and analysis of real-time heuristic controls in three different settings. The problems we studied and the solution approaches we adopted share some similarity. From the modeling perspective, the investigated problems can all be formulated as dynamic controls for which characterizing optimal policies are computationally infeasible. From the technical perspective, the parameters of the proposed heuristics can all be viewed as the sum of some baseline parameters (which are given by some approximated optimization problems) and some adjustment parameters (which are computed adaptively according to the realized randomness). One contribution of this thesis to the broader dynamic optimization literature is that it illustrates the usefulness of the types of heuristic controls as mentioned above, in terms of its simplicity and effectiveness. This observation further motivates us to consider other dynamic optimization problems arise in related settings. We elaborate the potential future research directions here.

Firstly, there are many potential research questions in the area of online retail that have similar flavor with the first chapter, where decisions that are closely related to each other is better made jointly. One direct question is how to incorporate the assortment decision, or more broadly, the decision of how the products are displayed on the online retailer's website (see Aouad and Segev 2015 and Gallego et al. 2016 for the optimization of product display decisions). Another operational decision that directly affects the balance between demand and supply is the inventory decision. In the setting of online retail setting, the inventory level at each FC are usually made according to a two-step procedure: the retailer first decides on how many units of products to source from her supplier, and then decide on how to allocate the products into her network of FCs. It is important to characterize a near-optimal policy that decides on these decisions together with the pricing and fulfillment decision. Given the complexity of the corresponding dynamic optimization problems, developing real-time heuristic controls will be particularly useful. Moreover, recent advancement on the design of asymptotically optimal inventory policies shed lights on the potential usefulness on this type of controls; see e.g. Reiman and Wang (2015), Goldberg et al. (2016),Xin and Goldberg (2016) and Wei et al. (2018).

Secondly, although the revenue management problem of reusable resources that we considered in chapter 2 already captures a wide range of applications, there are numerous other applications where resources are also reusable in nature but more complex models need to be proposed to fully characterize the problem dynamics. One potential model is where customer specifies a time window during which she wants to enjoy the service, which has applications for the on-demand delivery firms. The key modeling difference here is that, given the flexibility in the timing of demand fulfillment, the firm also need to make scheduling decisions which dictate the orders by which different requests are served. Another related direction is the two-sided market, where the supply is consist of self-scheduling agents; see e.g. the growing literature studying the pricing and matching problems in the two-side market setting (e.g. Banerjee et al. 2015, Bimpikis et al. 2016, Zhou 2017, Afèche et al. 2018, Ma et al. 2018). This direction introduces uncertainty in terms of the capacity level, and therefore calls for additional effort in designing appropriate controls.

Lastly, there are many open research questions in the field of dynamic pricing with demand learning. In particular, there are two extensions that have draw a lot of attentions: incorporating the non-stationarity of the arrival process (e.g. Besbes et al. 2015 and Keskin and Zeevi 2016b), and how to efficiently utilize customer's feature information (e.g. Cohen et al. 2016, Javanmard and Nazerzadeh 2016, and Ban and Keskin 2017). A major un-answered question in both of the extensions is what happens if there is inventory constraints present. Moreover, in the setting of online retail and advertisement industry, practitioners and researcher has been studying the design of efficient learning mechanism for a long time, which leads to many celebrated models; e.g. recommender system, clickthrough attribution model, etc. How to incorporate pricing decision while respecting firm's operational constraint is certainly an interesting question to answer.

APPENDIX A

Appendix to Chapter 2

A.1 Proof of Lemma 2.5.1

In what follows, we will only show the existence of a proper set $\mathcal{Q}$ under the single-product setting; the argument can be easily extended to the multiple-product setting. Let $F^t : \Omega_p \to [0,1]$ denote the CDF for pricing decision during period t under the optimal control π^*. Also, let $\bar{r}_j^t$ and $\bar{\lambda}_j^t$ denote the expected revenue and demand rate from location j during period t under π^* (since we only consider the single-product setting, there is no need to use subscript k), i.e.,

$$\bar{r}_j^t := \mathbb{E}^{\pi^*}[R_j^t(p^t)] = \int_{\Omega_p} r_j(p)\ dF^t(p) \quad \text{and} \quad \bar{\lambda}_j^t := \mathbb{E}^{\pi^*}[D_j^t(p^t)] = \int_{\Omega_p} \lambda_j(p)\ dF^t(p).$$

To prove Lemma 2.5.1, we first show that there exist weight vectors $\{\boldsymbol{\alpha}^t\}$ such that, for the uniform grid $\mathcal{Q}^u$ defined in Chapter 2.5 and some sufficiently small $\epsilon_r, \epsilon_\lambda > 0$, the following hold:

$$\begin{aligned}
\left|\bar{r}_j^t - \sum_{m=1}^M \alpha_m^t r_j(q_m^u)\right| &= \left|\int_{\Omega_p} p\lambda_j(p)\ dF^t(p) - \sum_{m=1}^M \alpha_m r_j(q_m^u)\right| \le \epsilon_r \quad \forall j,t, && \text{(A.1)}\\
\left|\bar{\lambda}_j^t - \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u)\right| &= \left|\int_{\Omega_p} \lambda_j(p)\ dF^t(p) - \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u)\right| \le \epsilon_\lambda \quad \forall j,t, && \text{(A.2)}\\
\sum_{m=1}^M \alpha_m^t = 1, &\qquad \alpha_m^t \ge 0, \quad \forall m,t. && \text{(A.3)}
\end{aligned}$$

Define a uniform partition of the interval Ω_p as

$$\Omega_p = \cup_{m=1}^M \mathcal{P}_m := \left[\cup_{m=1}^{M-1} \left[p_\ell + (m-1)\,\Delta_q, p_\ell + m\Delta_q\right)\right] \cup \left[p_u - \Delta_q, p_u\right]$$

where $\Delta_q := (p_u - p_\ell)/M$ is the length of the sub-intervals. Then the uniform price grid can be defined as $\mathcal{Q}^u := (p_\ell + (m-1/2)\Delta_q)_{m=1}^M$. Consider a choice of weight vector $\alpha_m^t = \int_{\mathcal{P}_m} dF^t(p)$. Note that (A.3) is satisfied immediately by definition. We now show that the combination of $\mathcal{Q}^u$ and $\boldsymbol{\alpha}^t$ defined above satisfy (A.1) and (A.2). By definition, for all $j \in [J]$, we have

$$
\begin{aligned}
& \left|\bar{\lambda}_j^t - \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u)\right| \\
= \; & \left|\int_{\Omega_p} \lambda_j(p)\ dF^t(p) - \sum_{m=1}^M \boldsymbol{\alpha}_m^t \lambda_j(q_m^u)\right| = \left|\sum_{m=1}^M \int_{\mathcal{P}_m} \left(\lambda_j(p) - \lambda_j(q_m^u)\right)\ dF^t(p)\right| \\
\leq \; & \sum_{m=1}^M \int_{\mathcal{P}_m} \left|\lambda_j(p) - \lambda_j(q_m^u)\right|\ dF^t(p) \leq \lambda_u \Delta_q.
\end{aligned}
$$

where the first inequality follows from triangular inequality and the last inequality follows from Assumption A1 together with $\lambda_u := \max_{j\in[J], p\in\Omega_p} |\lambda_j'(p)|$. By similar argument, since $|r_j'(p)| \leq |\lambda_j(p) + p\lambda_j'(p)| \leq 1 + p_u\lambda_u$ for all $p \in \Omega_p$, it is not difficult to show that (A.1) is satisfied for $\epsilon_r = (1+p_u\lambda_u)\Delta_q$.

We now show that the choices of $\mathcal{Q}^u$ and $\boldsymbol{\alpha}^t$ above guarantees a good approximation. The fulfillment LP under the uniform discretization we construct is as follows:

$$
\mathbf{FC}^A := \min_{\{0\leq x_{ij}^t \leq 1\}} \left\{ \sum_{t=1}^T \sum_{i=0}^I \sum_{j=1}^J c_{ij} x_{ij}^t : \sum_{i=0}^I x_{ij}^t = \sum_{m=1}^M \alpha_m^t \lambda_j(q_m^u), \sum_{t=1}^T \sum_{j=1}^J x_{ij}^t \leq C_i \right\}.
$$

On the other hand, the CDF of the fulfillment assignment under π^* can be solve by the following LP:

$$
\mathbf{FC}^O := \min_{\{0\leq x_{ij}^t \leq 1\}} \left\{ \sum_{t=1}^T \sum_{i=0}^I \sum_{j=1}^J c_{ij} x_{ij}^t : \sum_{i=0}^I x_{ij}^t = \bar{\lambda}_j^t,\ \sum_{t=1}^T \sum_{j=1}^J x_{ij}^t \leq C_i \right\}.
$$

The only difference between $\mathbf{FC}^A$ and $\mathbf{FC}^O$ is on the RHS of fulfillment constraint. Note that both $\mathbf{FC}^A$ and $\mathbf{FC}^O$ have stationary optimal solution. Then given (A.2) and the perturbation theory of the optimal objective value of LP (see e.g. Theorem 10.5 in Schrijver 1998), $\mathbf{FC}^A - \mathbf{FC}^O \leq IJT\lambda_u\Delta_q$. So the approximation error is bounded as follows:

$$
\begin{aligned}
& \mathcal{J}^* - \mathcal{J}^{ALP} \\
\leq\ & \left[\sum_{t=1}^{T}\sum_{j=1}^{J} \bar{r}_j^t - \mathbf{FC}^O\right] - \left[\sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^t r_j(q_m^u) - \mathbf{FC}^A\right] \\
\leq\ & \sum_{t=1}^{T}\sum_{j=1}^{J} \left| \bar{r}_j^t - \sum_{m=1}^{M} \alpha_m^t r_j(q_m^u) \right| + \left(\mathbf{FC}^A - \mathbf{FC}^O\right) \\
\leq\ & \left[JT(1+p_u\lambda_u) + IJT\lambda_u\right]\Delta_q \leq \frac{(p_u - p_\ell)JT\left[1+\lambda_u(p_u+I)\right]}{M}.
\end{aligned}
$$

The proof is concluded by letting $M = \lceil (p_u - p_\ell)JT\left[1+\lambda_u(p_u+I)\right]/\epsilon \rceil$. For general K, the number of discrete prices required to reach an error of ϵ is at most $\lceil \left[(p_u - p_\ell)JT(1+K\Phi_1(p_u+IK))\right]^K/\epsilon^K \rceil$, $\Phi_1 = \max\limits_{p\in\Omega_p,\ j\in[J],\ k,\ell\in[K]} |\partial\lambda_{jk}(\boldsymbol{p})/\partial p_\ell| > 0$ (it is finite by Assumption A1) $\square$

A.2 Proof of Theorem 2.6.1

Let $\mathcal{Q}^u$ be uniform grid defined in Chapter 2.5. Without loss of generality, we assume that $T=1$. We consider a variant of RPF (V-RPF) defined as follow: during period t, fulfill the order from location j according to $\sigma_k^t(j)$ regardless of the availability of the corresponding FC; if the FC runs out of inventory, the retailer incurs a penalty cost of $\bar{c} := 2\cdot \max\limits_{j\in[J],k[K]} c_{0jk}$. In other words, V-RPF incurs the same revenue as RPF, yet no smaller fulfillment cost. Consequently, the loss can be bounded as follows:

$$
\begin{aligned}
& \mathcal{J}^{ALP}(\theta) - \mathcal{J}^{RPF}(\theta) \\
\leq\ & \mathcal{J}^{ALP}(\theta) - \mathcal{J}^{V-RPF}(\theta) \\
=\ & \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^*) - \sum_{t=1}^{\theta}\sum_{j=1}^{J} (\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)\right] \\
& \quad +\ \bar{c}\,\mathbb{E}\left[\sum_{i=1}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J} X_{ijk}^t - C_{ik}(\theta)\right)^+\right] \\
& \quad +\ \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}X_{ijk}^t - \sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}x_{ijk}^*\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\Delta R_j^t\right] + \bar{c}\,\mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K} + \left(\sum_{t=1}^{\theta}\sum_{j=1}^{J} X_{ijk}^t - C_{ik}(\theta)\right)^+\right] \\
&\qquad + \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}\Delta X_{ijk}^t\right],
\end{aligned}
$$

where $\Delta R_j^t := \sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^*) - (\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)$, and $\Delta X_{ijk}^t := X_{ijk}^t - x_{ijk}^*$. By definition of RPF, $\mathbb{E}\,[\Delta R_j^t] = \mathbb{E}\,[\Delta X_{ijk}^t] = 0$. As for the last term, by triangular inequality,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J} X_{ijk}^t - C_{ik}(\theta)\right)^+\right] \\
\le\;& \mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\sum_{t=1}^{\theta}\sum_{j=1}^{J} X_{ijk}^t - \theta\sum_{j=1}^{J} x_{ijk}^*\right)^+\right] + \mathbb{E}\left[\sum_{i=0}^{I}\sum_{k=1}^{K}\left(\theta\sum_{j=1}^{J} x_{ijk}^* - C_i(\theta)\right)^+\right] \\
\le\;& \sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\mathbb{E}\left[\left(\sum_{t=1}^{\theta} X_{ijk}^t - x_{ijk}^*\right)^+\right] + 0 \\
\le\;& \sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left[\mathbf{Var}\left(\sum_{t=1}^{\theta}\Delta X_{ijk}^t\right)\right]^{1/2} \\
=\;& O\left(\sqrt{\theta}\right),
\end{aligned}
$$

where the second inequality follows from the inventory constraint in ALP, the last inequality follows because ΔX_{ijk}^t's are independent and bounded from above by $D_{jk}^t \le 1$. This completes the proof. $\square$

A.3 Proof of Theorem 2.7.1

Let $T = 1$. Per our discussion in Chapter 2.7, we can assume $\sum_{j=1}^{J} x_{ijk}^* = C_{ik}$ without loss of generality. Let $\boldsymbol{C}_i^t(\theta)$ be the on-hand inventory level in FC i at the beginning of period t for a problem with size θ. By definition, we have $\boldsymbol{C}_i^1(\theta) = \theta\boldsymbol{C}_i$. Fix $\theta > 0$. We divide our proof into several steps.

Step 1

In this step, we state and prove two key observations that are useful in helping us to express the evolution of pricing and fulfillment decisions over time. We call an $\text{FLP}^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ to be "balanced" if it satisfies (i) $\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^*\lambda_{jk}(\boldsymbol{q}_m^t) =$

$\sum_{i=0}^{I} C_{ik}^t/(T-t+1)$ for all k, and (ii) $C_{ik}^t > 0$ for all i,k. We make our first observation regarding the solution of a balanced FLP^t.

Observation A.1. *The optimal solution $\boldsymbol{x}^t$ to a non-DR-degenerate balanced FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ has the following property: For every $k \in [K]$, there are exactly $I+J$ strictly positive components in $(x_{ijk}^t)_{i\in\{0\}\cup[I], j\in[J]}$, with the other components equal to zero. Moreover, the inventory constraints are all binding.*

Proof. Note that FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is separable over k, so solving FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is equivalent to solving K sub-problems defined below:

$$\text{FLP}_k^t(\mathcal{Q}^t, C_k^t) :=$$
$$\left\{ \min_{x_{ijk}\geq 0} \sum_{i=0}^{I}\sum_{j=1}^{J} c_{ijk}x_{ijk} \;:\; \sum_{i=0}^{I} x_{ijk} = \sum_{m=1}^{M} \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^t),\; \sum_{j=1}^{J} x_{ijk} \leq \frac{C_{ik}^t}{T-t+1} \right\}.$$

Since FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is balanced, all the inventory constraints in FLP$_k^t(\mathcal{Q}^t, \boldsymbol{C}_k^t)$ must be binding. Since FLP$^t(\mathcal{Q}^t, \boldsymbol{C}^t)$ is non-DR-degenerate and separable over k, FLP$_k^t(\mathcal{Q}^t, \boldsymbol{C}_k^t)$ is also non-degenerate for each k. Thus, Observation A.1 follows directly from the standard result on transportation LP (see Corollary 7.2 in Dantzig and Thapa 2006). $\square$

Let $\boldsymbol{x}_k = (x_{ijk})_{i\in\{0\}\cup[I], j\in[J]}$ and $\boldsymbol{c}_k = (c_{ijk})_{i\in\{0\}\cup[I], j\in[J]}$. Given our assumptions in the statement of Theorem 2.7.1 and at the beginning of this chapter, FLP$^1(\mathcal{Q}^u, \boldsymbol{C})$ is non-DR-degenerate and balanced. Thus, for all k, FLP$_k^1(\mathcal{Q}^u, \boldsymbol{C}_k)$ are non-degenerate and has $I+J$ non-zero components in $\boldsymbol{x}_k^*$ (since there are $I+J+1$ constraints with exactly one redundant). Let A_k and $\boldsymbol{V}_k$ denote the coefficient matrix and the RHS of inventory constraints in FLP$_k^1$. Let $\bar{A}_k$ be the matrix where we delete the $(J+1)^{th}$ row from A_k, i.e., the row corresponding to the inventory constraint on FC 0, and $\bar{\boldsymbol{V}}_k$ be the vector where we delete C_{0k}/θ from $\boldsymbol{V}_k$. This constraint is redundant, since any $\boldsymbol{x}_k$ satisfying the system of equations $\bar{A}_k\boldsymbol{x}_k = \bar{\boldsymbol{V}}_k$ automatically satisfies $\sum_{j=1}^{J} x_{0jk}^t = C_{0k}/\theta$ (the deleted constraint). Since the deleted constraint is redundant, FLP$_k^1$ is equivalent to $\left\{\min\ \boldsymbol{c}_k^\top \boldsymbol{x}_k :\ \bar{A}_k\boldsymbol{x}_k = \bar{\boldsymbol{V}}_k,\ \boldsymbol{x} \succeq \boldsymbol{0}\right\}$; moreover, by Lemma 7.1 in Dantzig and Thapa (2006), $\bar{A}_k$ has linearly independent rows. Let $\mathcal{B}_k = \{(i,j) :\ 0 < x_{ijk}^* < 1\}$ and $\mathcal{N}_k = \{(i,j) :\ x_{ijk}^* = 0\}$ be the indices of the optimal basic and non-basic variables respectively. Without loss of generality, we assume that $\bar{A}_k$ is written as $[B_k, N_k]$ where B_k and N_k are the sub-matrices of $\bar{A}_k$ corresponding to the basic and non-basic indices in $\mathcal{B}_k$ and $\mathcal{N}_k$ respectively. Following the same decomposition, the optimal solution can be represented as $\boldsymbol{x}_k^* = [\boldsymbol{x}_{k,B}^*, \boldsymbol{x}_{k,N}^*]$, where $\boldsymbol{x}_{k,B}^* = B_k^{-1}\bar{\boldsymbol{V}}_k$ and $\boldsymbol{x}_{k,N}^* = \boldsymbol{0}$ (the invertibility of B_k is proved in Theorem 7.6 in Dantzig and Thapa 2006). Thus, the unique optimal solution to FLP^1 can be accordingly written as $\boldsymbol{x}^* = [\boldsymbol{x}^*_B; \boldsymbol{x}^*_N]$, where $\boldsymbol{x}^*_B = (\boldsymbol{x}^*_{k,B})_{k=1}^K$, $\boldsymbol{x}^*_N = (\boldsymbol{x}^*_{k,N})_{k=1}^K$. Note that if we define $B = \operatorname{diag}(B_1, \ldots, B_K)$ as a block diagonal matrix with $(B_k)_{k=1}^K$ as its main diagonal blocks and zero matrices as off-diagonal blocks, and $\bar{\boldsymbol{V}} = [\bar{\boldsymbol{V}}_1; \ldots; \bar{\boldsymbol{V}}_K]$, we can write $\boldsymbol{x}^*_B = B^{-1}\bar{\boldsymbol{V}}$. Let $\boldsymbol{V}^t_k$ be the RHS of FLP^t_k and $\bar{\boldsymbol{V}}^t_k$ be the vector where we delete $C^t_{0k}/(\theta - s)$ from $\boldsymbol{V}^k_t$. Define

$$\delta \boldsymbol{V}^t_k := \left(\left(\sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^t_m) - \sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^u_m) \right)_{j=1}^{J}, \ \left(-\sum_{s=1}^{t-1} \Delta C^s_{ik}/(\theta - s) \right)_{i=0}^{I} \right)$$

and let $\delta \bar{\boldsymbol{V}}^t_k$ be the vector where we delete $-\sum_{s=1}^{t-1} \Delta C^s_{0k}/(\theta - s)$ from $\delta \boldsymbol{V}^t_k$. Let $\delta \bar{\boldsymbol{V}}^t = (\delta \bar{\boldsymbol{V}}^t_k)_{k=1}^K$. Following the same decomposition, we will also write $\boldsymbol{c} = [\boldsymbol{c}_B; \boldsymbol{c}_N]$. Per our definition in Chapter 2.4, $\boldsymbol{\lambda}^{tot}(\boldsymbol{p})$ is the aggregated purchase probability given a price vector $p \in \Omega_p$. We make our second observation below:

Observation A.2. *At period t, as long as the following conditions hold:*

$$\sum_{j=1}^{J} \boldsymbol{\lambda}_j(\boldsymbol{q}^t_m) = \hat{\boldsymbol{\lambda}}^t_m := \boldsymbol{\lambda}^{tot}(\boldsymbol{q}^u_m) - \frac{1}{M\alpha^*_m} \left(\sum_{i=0}^{I} \sum_{s=1}^{t-1} \frac{\Delta \boldsymbol{C}^s_i}{T - s} \right) \in [0,1]^K, \tag{A.4}$$

$$C^t_{ik}(\theta) = \hat{C}^t_{ik}(\theta) := (\theta - t + 1) \left[C_{ik} - \sum_{s=1}^{t-1} \frac{\Delta C^s_{ik}}{\theta - s} \right] \geq 0, \tag{A.5}$$

$$\boldsymbol{x}^*_{k,B} + B_k^{-1} \left(\delta \bar{\boldsymbol{V}}^t_k \right) \succeq \boldsymbol{0}, \tag{A.6}$$

then the unique optimal solution to FLP^t is given by $\boldsymbol{x}^t_{k,B} = \boldsymbol{x}^*_{k,B} + B_k^{-1} \left(\delta \bar{\boldsymbol{V}}^t_k \right)$ and $\boldsymbol{x}^t_{k,N} = \boldsymbol{0}$ for all k.

Proof. Under condition (A.4), FLP^t is balanced. This is so because, for all k,

$$\begin{aligned} \sum_{j=1}^{J} \sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^t_m) &= \sum_{j=1}^{J} \sum_{m=1}^{M} \alpha^*_m \lambda_{jk}(\boldsymbol{q}^u_m) - \sum_{i=0}^{I} \sum_{s=1}^{t} \frac{\Delta C^s_{ik}}{T - s} \\ &= \sum_{i=0}^{I} C_{ik} - \sum_{i=0}^{I} \sum_{s=1}^{t-1} \frac{\Delta C^s_{ik}}{\theta - s} = \sum_{i=0}^{I} \frac{C^t_{ik}(\theta)}{\theta - t + 1}, \end{aligned}$$

where the second equality follows from our assumption in the beginning of this chapter, and the last equality follows from the definition of ΔC^t_{ik}. As a result, for all k, the inventory constraints in FLP^t_k are all binding. Notice that condition (A.4) and (A.5) implies that $\boldsymbol{V}^t_k = \boldsymbol{V}_k + \delta \boldsymbol{V}^t_k \succeq \boldsymbol{0}$, and thus FLP$^t_k$ is equivalent to

$\left\{\min_{\mathbf{x}_k^t}\ \mathbf{c}_k^\top \mathbf{x}_k^t\ :\ \bar{A}_k \boldsymbol{x}_k^t = \bar{\boldsymbol{V}}_k + \delta\bar{\boldsymbol{V}}_k^t,\ \mathbf{x}_k^t \succeq \mathbf{0}\right\}$. The feasibility of the proposed optimal solution can be directly verified under condition (A.6); its optimality follows from Karush-Kuhn-Tucker (KKT) conditions; and its uniqueness follows from the invertibility of B_k. $\square$

Step 2

Define $\hat{\mathbf{x}}^t := (\hat{\boldsymbol{x}}_B^t, \boldsymbol{x}_N) = (\boldsymbol{x}_B^* + B^{-1}\left(\delta\bar{\boldsymbol{V}}^t\right), \mathbf{0})$. Let $\phi_x = \min_{k\in[K]}\min_{(i,j)\in\mathcal{B}_k} x_{ijk}^* > 0$ (by non-degeneracy assumption); $\Phi_1 = \max\limits_{p\in\Omega_p,\ j\in[J],\ k,\ell\in[K]} |\partial\lambda_{jk}(\boldsymbol{p})/\partial p_\ell| > 0$ (it is finite by Assumption A1); $\Phi_2 = \max_{k\in[K]} ||B_k^{-1}||_\infty > 0$ (it is also finite by the invertibility of B_k); $\phi_\lambda := \max\{x > 0 : \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^u) + x\cdot\mathbf{1} \in [0,1]^K,\ \forall m\} > 0$ (by Assumption A1 and the fact that $\boldsymbol{q}_m^u$ lies in the interior of Ω_p); and $v > 0$ denote the smallest absolute eigenvalue of $\mathcal{G}_{\boldsymbol{\lambda}^{tot}}$ (by Assumption A3). Without loss of generality, $\boldsymbol{\alpha}^* \succ \mathbf{0}$ since we can delete any α_m^* with zero value without changing anything else. We state a lemma.

Lemma A.3.1 *Suppose that $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^s) = \hat{\boldsymbol{\lambda}}_m^s \in [0,1]^K$, $\mathbf{x}_s = \hat{\mathbf{x}}_s \succeq \mathbf{0}$ and $\boldsymbol{C}_i^s(\theta) = \hat{\boldsymbol{C}}_i^s(\theta) \succeq \mathbf{0}$ for all $s < t$. Then $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t) = \hat{\lambda}_m^t$, $\mathbf{x}^t = \hat{\mathbf{x}}^t$ and $\boldsymbol{C}_i^t(\theta) = \hat{\boldsymbol{C}}_i^t(\theta)$ hold if the following two conditions hold at time t*

$$(\dagger):\qquad \left|\sum_{i=1}^{I}\sum_{s=1}^{t-1}\frac{\Delta C_{ik}^s}{\theta - s}\right| \le \min\left\{\frac{\phi_x}{\Phi_2}\left(1+\frac{K\Phi_1}{v}\right)^{-1},\ \phi_\lambda M\cdot\min_{m\in[M]}\alpha_m^*\right\},\ \forall k,$$

$$(\dagger\dagger):\qquad \left|\sum_{s=1}^{t-1}\frac{\Delta C_{ik}^s}{\theta - s}\right| \le C_{ik},\ \forall i,k,$$

Proof. We proceed by induction. The base case ($t = 1$) is verified directly by definition. Now, consider $t > 1$. Assume the identity holds for $s \le t-1$. Given condition ($\dagger$) and the definition of ϕ_λ, it is not difficult to show that $\hat{\boldsymbol{\lambda}}_m^t \in [0,1]^K$. Since $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t)$ is simply the projection of $\hat{\boldsymbol{\lambda}}_m^t$ onto $[0,1]^K$ (see Step 2a), $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t) = \hat{\boldsymbol{\lambda}}_m^t$.

We now show that $C_{ik}^t(\theta) = \hat{C}_{ik}^t(\theta)$. Suppose that, in Step 2c of R^{2}PF, we sample m^t for some $m^t \in [M]$. Remember that, in period $t-1$, the probability of using FC i to fulfill the request of product k from location j conditioned on $D_{jk} = 1$ is $y_{ijk}^{t-1} = x_{ijk}^{t-1}/\sum_{i=0}^{I} x_{ijk}^{t-1}$. Moreover, since conditions (A.4) - (A.6) are implied for all $s \le t$ by the inductive assumption, by Observation A.2, the inventory constraints in

FLP^{t-1} are binding. So, the remaining inventory at the beginning of period t satisfies:

$$\begin{aligned} C_{ik}^t(\theta) &= C_{ik}^{t-1}(\theta) - \sum_{j=1}^J X_{ijk}^{t-1} = C_{ik}^{t-1}(\theta) - \sum_{j=1}^J y_{ijk}^{t-1}\left(\sum_{m=1}^M \alpha_m^* \lambda_{jk}(\boldsymbol{q}_m^{t-1})\right) - \Delta C_{ik}^{t-1} \\ &= C_{ik}^{t-1}(\theta) - \sum_{j=1}^J x_{ijk}^{t-1} - \Delta C_{ik}^{t-1} = C_{ik}^{t-1}(\theta) - \frac{C_{ik}^{t-1}(\theta)}{\theta - t + 2} - \Delta C_{ik}^{t-1} \\ &= (\theta - t + 2 - 1)\left[C_{ik}(\theta) - \sum_{s=1}^{t-2} \frac{\Delta C_{ik}^s}{\theta - s}\right] - \Delta C_{ik}^{t-1} = \hat{C}_{ik}^t(\theta), \end{aligned}$$

where the second equality follows from the definition of ΔC_{ik}^t; the third equality follows from the fulfillment constraint in FLP^t; the fourth constraint follows since the inventory constraints in FLP^{t-1} are binding; and, the fifth constraints follows from the inductive assumption.

At last, to show that $\boldsymbol{x}^t = \hat{\boldsymbol{x}}^t$, by Observation A.2, it suffices to show conditions (A.4) - (A.6) are satisfied for period t. Condition (A.4) is implied by $\boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t) = \hat{\boldsymbol{\lambda}}_m^t$. Since condition (††) implies $\hat{C}_{ik}^t(\theta) \geq 0$, and we have shown that $C_{ik}^t(\theta) = \hat{C}_{ik}^t(\theta)$, condition (A.5) is satisfied. To check condition (A.6), define $\delta\boldsymbol{q}_m^t = \boldsymbol{q}_m^t - \boldsymbol{q}_m^u$. By Assumption A1 and Mean Value Theorem, $\delta\boldsymbol{q}_m^t = [\mathcal{G}_{\boldsymbol{\lambda}^{tot}}(\boldsymbol{\xi}_m^t)]^{-1}\left(\sum_{i=0}^I \sum_{s=1}^{t-1} \Delta\boldsymbol{C}_i^s/(\theta - s)\right)/(M\alpha_m^*)$ for some $\boldsymbol{\xi}_m^t \in \Omega_p$. By Mean Value Theorem again, there exist $\boldsymbol{\zeta}_{mk}^t \in \Omega_p$ such that

$$\begin{aligned} &\left|\sum_{m=1}^M \alpha_m^* \left[\lambda_{jk}(\boldsymbol{q}_m^t) - \lambda_{jk}(\boldsymbol{q}_m^u)\right]\right| \\ =\ &\left|\sum_{m=1}^M \frac{(\nabla\lambda_{jk}(\boldsymbol{\zeta}_{mk}^t))^\top [\mathcal{G}_{\boldsymbol{\lambda}^{tot}}(\boldsymbol{\xi}_m^t)]^{-1}}{M}\left(\sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta\boldsymbol{C}_i^s}{\theta - s}\right)\right| \\ \leq\ &\frac{K\Phi_1}{v} \max_{k\in[K]} \left|\sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s}\right| \end{aligned}$$

where the inequality holds by Assumption A3 and the definition of Φ_1. So,

$$\left|\left|B_k^{-1}\left(\delta\bar{\boldsymbol{V}}_k^t\right)\right|\right|_\infty \leq \left|\left|B_k^{-1}\right|\right| \cdot \left|\left|\delta\bar{\boldsymbol{V}}_k^t\right|\right| \leq \Phi_2 \cdot \left(1 + \frac{K\Phi_1}{v}\right) \max_{k\in[K]} \left|\sum_{i=0}^I \sum_{s=1}^{t-1} \frac{\Delta C_{ik}^s}{\theta - s}\right| \leq \phi_x,$$

where the last inequality follows from condition (†). This implies condition (A.6). □

Step 3

In this step, we show that the conditions in Lemma A.3.1 hold for the majority of the selling season. Define a stopping time $\tau(\theta)$ to be the first t such that either

(†) or (††) is violated. According to Lemma A.3.1, for any period before $\tau(\theta)$, we can explicitly characterize the evolution of price, fulfillment assignment, and inventory consumption. The following lemma provides a lower bound on the length of $\tau(\theta)$.

Lemma A.3.2 *There exists a constant* $\Psi_3 > 0$ *independent of* θ *such that* $\mathbb{E}\,[\theta - \tau(\theta)] \leq \Psi_3(1 + \log \theta)$.

Proof. Define $\tau_1(\theta)$ and $\tau_2(\theta)$ to be the first period t such that conditions (†) and (††) are violated, respectively. By definition $\tau(\theta) = \min_{i\in\{1,2\}} \tau_i(\theta)$. We only bound $\tau_1(\theta)$, since $\tau_2(\theta)$ can be bounded using a similar argument. Let Γ_k denote the RHS of the inequality in condition (†) in Lemma A.3.1. The sequence

$$\left\{ S_k^t = \sum_{i=0}^{I} \frac{\Delta C_{ik}^{t-1}}{\theta - (t-1)} + \sum_{i=0}^{I} \frac{\Delta C_{ik}^{t-2}}{\theta - (t-2)} + \cdots + \sum_{i=0}^{I} \frac{\Delta C_{ik}^{1}}{\theta - 1} \right\}_{t \leq \theta}$$

is a Martingle with respect to the natural filtration $\{\mathcal{H}^t\}$, where $\mathcal{H}^t$ is the history of all information up to the beginning of period t. This implies that the sequence $\{|S_k^t|\}_{t\leq\theta}$ is a sub-Martingle. By Doob's submartingle inequality (see for example Williams 1991) and union bound,

$$\begin{aligned}
\mathbb{P}(\tau_1(\theta) \leq t) &\leq \mathbb{P}\left(|S_k^s| \geq \Gamma_k \text{ for some } s \leq t, k \in [K]\right) \\
&\leq \sum_{k=1}^{K} \mathbb{P}\left(\max_{s\leq t} |S_k^s| \geq \Gamma_k\right) \\
&\leq \sum_{k=1}^{K} \frac{\mathbb{E}\left[(S_k^t)^2\right]}{\Gamma_k^2}.
\end{aligned}$$

Note that ΔC_{ik}^s and ΔC_{jk}^t are independent for all $s \neq t$ and $i, j \in \{0\} \cup I$. So,

$$\begin{aligned}
\mathbb{E}[\left(S_k^t\right)^2] = \mathbb{E}\left[\left(\sum_{s=1}^{t-1}\sum_{i=0}^{I} \frac{\Delta C_{ik}^s}{\theta - s}\right)^2\right] &= \sum_{s=1}^{t-1} \frac{\mathbb{E}\left[\left(\sum_{i=0}^{I} \Delta C_{ik}^s\right)^2\right]}{(\theta - s)^2} \\
= \sum_{s=1}^{t-1} \frac{\sum_{i,j\in\{0\}\cup[I]} \mathbb{E}\left[\Delta C_{ik}^s \Delta C_{jk}^s\right]}{(\theta - s)^2} &= O\left(\frac{1}{\theta - t}\right),
\end{aligned}$$

where the last inequality follows from the boundedness of $\mathbb{E}\left[\Delta C_{ik}^s \Delta C_{jk}^s\right]$. The proof is complete by noting that $\mathbb{E}\,[\theta - \tau_1(\theta)] = \sum_{t=2}^{\theta} \mathbb{P}(\tau_1(\theta) \leq t) = 1 + \sum_{t=2}^{\theta-1} O\left(\frac{1}{\theta - t}\right) = O(\log \theta)$. $\square$

Step 4

We now bound the loss of R²PF. First, note that we can decouple the loss into two terms as follows:

$$
\begin{aligned}
&\mathcal{J}^{ALP}(\theta) - \mathcal{J}^{R^2PF}(\theta) \\
&= \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^*) - \sum_{t=1}^{\theta}\sum_{j=1}^{J} (\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)\right] \\
&\quad + \mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk} X_{ijk}^t - \sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk} x_{ijk}^*\right].
\end{aligned}
$$

The two terms on the RHS of the equation above are the loss in revenue and the loss in fulfillment cost of R²PF, respectively. We start with providing an upper bound for the loss in revenue:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^*) - \sum_{t=1}^{\theta}\sum_{j=1}^{J} (\boldsymbol{p}^t)^\top \boldsymbol{D}_j^t(\boldsymbol{p}^t)\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J} R_j^t(\boldsymbol{p}^t)\right] \\
&\quad + \mathbb{E}\left[\sum_{t=\tau(\theta)}^{\theta}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^u)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J} R_j^t(\boldsymbol{p}^t)\right] \\
&\quad + \mathbb{E}\left[(\theta - \tau(\theta) + 1)\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^u)\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J} R_j^t(\boldsymbol{p}^t)\right] \\
&\quad + K p_u (1 + \Psi_3 + \Psi_3 \log\theta), \qquad \text{(A.7)}
\end{aligned}
$$

where the last inequality follows from Lemma A.3.2, the boundedness of price, and the assumption of at most one arrival per period. Let $\hat{\Delta}_j^t = \sum_{m=1}^{M} \alpha_m^* r_j(\boldsymbol{q}_m^t) - (\boldsymbol{p}^t)^\top D_j(\boldsymbol{p}^t)$. Define $r^{tot}(\boldsymbol{p}) = \sum_{j=1}^{J} r_j(\boldsymbol{p}) = \boldsymbol{p}^\top \boldsymbol{\lambda}^{tot}(\boldsymbol{p})$. By Assumption A1, there exists an inverse of $\boldsymbol{\lambda}^{tot}(\boldsymbol{p})$, which we will denote as $\boldsymbol{p}(\boldsymbol{\lambda}^{tot}) : [0,1]^K \to \Omega_p$. With slight abuse of notation, we will use $r^{tot}(\boldsymbol{\lambda}^{tot}) = (\boldsymbol{p}(\boldsymbol{\lambda}^{tot}))^\top \boldsymbol{\lambda}^{tot}$ to denote total revenue rate as a function of aggregate demand. Let $\boldsymbol{\lambda}_m^* = \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^u)$, $\boldsymbol{\lambda}_m^t = \boldsymbol{\lambda}^{tot}(\boldsymbol{q}_m^t)$, and $\boldsymbol{\epsilon}^t = \sum_{i=0}^{I}\sum_{s=1}^{t-1}\Delta \boldsymbol{C}_i^s/(\theta - s)$. For $t \le \tau(\theta)$, we know that $\boldsymbol{\lambda}_m^t = \boldsymbol{\lambda}_m^* - \boldsymbol{\epsilon}^t/(M\alpha_m^*)$. By Taylor's expansion at λ_m^*, we have

$$
\begin{aligned}
r^{tot}(\boldsymbol{q}_m^t) &= r^{tot}(\boldsymbol{\lambda}_m^t) \\
&= r^{tot}(\boldsymbol{\lambda}_m^*) - \frac{(\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top \boldsymbol{\epsilon}_m^t}{M\alpha_m^*} + \frac{(\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}(\boldsymbol{\eta}^t)\boldsymbol{\epsilon}^t}{2M^2(\alpha_m^*)^2} \\
&= r^{tot}(\boldsymbol{q}_m^u) - \frac{(\nabla r^{tot}(\boldsymbol{\lambda}_{\boldsymbol{m}}^*))^\top \boldsymbol{\epsilon}_m^t}{M\alpha_m^*} + \frac{(\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}(\boldsymbol{\eta}^t)\,\boldsymbol{\epsilon}^t}{2M^2(\alpha_m^*)^2}
\end{aligned}
$$

for some $\boldsymbol{\eta}_m^t \in [0,1]^K \in \Omega_p$. So, the first term in (A.7) can be bounded as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}R_j^t(\boldsymbol{p}^t)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{m=1}^{M}\alpha_m^* r^{tot}(\boldsymbol{q}_m^u) - \sum_{t=1}^{\tau(\theta)-1}\sum_{m=1}^{M}\alpha_m^* r^{tot}(\boldsymbol{q}_m^t)\right] \\
&\quad + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{m=1}^{M}\alpha_m^* r_j(\boldsymbol{q}_m^t) - \sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}(\boldsymbol{p}^t)^\top D_j(p^t)\right] \\
&\le \mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}\frac{(\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top \boldsymbol{\epsilon}^t}{M}\right] - \mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}\frac{(\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}(\boldsymbol{\eta}_m^t)\,\boldsymbol{\epsilon}^t}{2M^2 \min_{m\in[M]}\alpha_m^*}\right] \\
&\quad + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)}\sum_{j=1}^{J}\hat{\Delta}_j^t\right] + Kp_u,
\end{aligned} \tag{A.8}
$$

where the last inequality holds because $\mathbb{E}[\hat{\Delta}_j^{\tau(\theta)}] \le Kp_u$. Note that $\{\sum_{s=1}^{t}\sum_{j=1}^{J}\hat{\Delta}_j^s\}_{t\le\theta}$ is a Martingale with respect to $\{\mathcal{H}^t\}_{t\le\theta}$ and $\tau(\theta)$ is bounded. So, by stopping time theorem (Williams, 1991), $\mathbb{E}[\sum_{t=1}^{\tau(\theta)}\sum_{j=1}^{J}\hat{\Delta}_j^t] = 0$. We are left to bound the first two terms in (A.8). Note that $\mathbb{E}[\sum_{t=2}^{\tau(\theta)-1}\boldsymbol{\epsilon}^t] = \mathbb{E}[\sum_{t=2}^{\tau(\theta)}\boldsymbol{\epsilon}^t] - \mathbb{E}[\sum_{t=\tau(\theta)}^{\theta}\boldsymbol{\epsilon}^t] = -\mathbb{E}[\sum_{t=\tau(\theta)}^{\theta}\boldsymbol{\epsilon}^t]$. By stopping time theorem again, $\mathbb{E}[\boldsymbol{\epsilon}^{\tau(\theta)}] = \mathbf{0}$, and $\mathbb{E}[\boldsymbol{\epsilon}^t] = \mathbf{0}$ for all $t > \tau(\theta)$. Consequently, $\mathbb{E}[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}(\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top\boldsymbol{\epsilon}^t] = (\sum_{m=1}^{M}\nabla r^{tot}(\boldsymbol{\lambda}_m^*))^\top\mathbb{E}[\sum_{t=2}^{\tau(\theta)-1}\boldsymbol{\epsilon}^t] = 0$. As for the second term in (A.8), let $\Phi_3 > 0$ be the largest absolute eigenvalue of $\nabla^2 r^{tot}$.

By Assumption A3, Φ_3 is finite. We thus have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{m=1}^{M}(\boldsymbol{\epsilon}^t)^\top \nabla^2 r^{tot}\left(\boldsymbol{\eta}_m^t\right)\boldsymbol{\epsilon}^t\right] \le \Phi_3\ \mathbb{E}\left[\sum_{t=2}^{\tau(\theta)-1}\sum_{k=1}^{K}\left(\sum_{i=1}^{I}\sum_{s=1}^{t-1}\frac{\Delta C_{ik}^s}{\theta - s}\right)^2\right] \\
\le\ & \Phi_3 \sum_{t=2}^{\theta}\sum_{k=1}^{K}\sum_{1\le s,v\le t-1}\frac{\mathbb{E}\left[\left(\sum_{i=1}^I \Delta C_{ik}^s\right)^2\left(\sum_{i=1}^I \Delta C_{ik}^v\right)^2\right]}{(\theta-s)(\theta-v)} \\
=\ & \Phi_3 \sum_{t=2}^{\theta}\sum_{k=1}^{K}\sum_{s=1}^{t-1}\frac{\mathbb{E}\left[\left(\sum_{i=1}^I \Delta C_{ik}^s\right)^2\right]}{(\theta-s)^2} \\
=\ & O(\log\theta).
\end{aligned}$$

At last we bound the loss of fulfillment cost. By Lemma A.3.1, for $t < \tau(\theta)$, $\mathbf{x}^t = [x_B^* + B^{-1}\delta\bar{\boldsymbol{V}}^t; \mathbf{0}]$. By definition, $\bar{c}$ is larger than all unit shipping costs. So,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] \\
&\quad + \mathbb{E}\left[\sum_{t=\tau(\theta)}^{\theta}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] \\
&\quad + \bar{c}IJK\mathbb{E}[\theta - \tau(\theta) + 1] \\
&\le \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}X_{ijk}^t - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}c_{ijk}x_{ijk}^*\right] \\
&\quad + \bar{c}IJK(1 + \Psi_3 + \Psi_3\log\theta). \qquad \text{(A.9)}
\end{aligned}$$

We are left to bound the first term in (A.9). Let $\Delta x_{ijk}^t = X_{ijk}^t - x_{ijk}^t$. Since $\mathbf{x}^t = \hat{\mathbf{x}}^t$

for all $t < \tau(\theta)$, we have:

$$\begin{aligned}
& \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}X_{ijk}^{t} - \sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}x_{ijk}^{*}\right] \\
= \; & \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}\left(x_{ijk}^{t} - x_{ijk}^{*}\right)\right] + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ij}\Delta x_{ijk}^{t}\right] \\
= \; & \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1} \boldsymbol{c}_B^{\top}B^{-1}\left(\delta\bar{\boldsymbol{V}}^{t}\right)\right] + \mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} c_{ijk}\Delta x_{ijk}^{t}\right] \\
\leq \; & \bar{c}(I+J)K||B^{-1}||_1\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{m=1}^{M} \alpha_m^{*}\left(\lambda_{jk}(\boldsymbol{q}_m^{t}) - \lambda_{jk}(\boldsymbol{q}_m^{u})\right)\right] \\
& -\bar{c}(I+J)K||B^{-1}||_1\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{k=1}^{K}\sum_{s=1}^{t} \frac{\Delta C_{ik}^{s}}{\theta - s}\right] \\
& +\bar{c}\,\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \Delta x_{ijk}^{t}\right] \\
= \; & -2\bar{c}(I+J)K||B^{-1}||_1\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=0}^{I}\sum_{k=1}^{K}\sum_{s=1}^{t} \frac{\Delta C_{ik}^{s}}{\theta - s}\right] + \bar{c}\,\mathbb{E}\left[\sum_{t=1}^{\tau(\theta)-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \Delta x_{ijk}^{t}\right]
\end{aligned}$$

where the second inequality follows from the definition of $\delta\bar{\boldsymbol{V}}^{t}$, the second equality follows from the definition of $\tau(\theta)$ and Lemma A.3.1. Note that $\{\sum_{s=1}^{t}\Delta x_{ijk}^{s}\}_{t\leq\theta}$ is Martingale with respect to the filtration $\{\mathcal{H}^{t}\}_{t\leq\theta}$. Following a similar argument as in bounding the revenue loss, it is not difficult to see that the terms after the above equation can be bounded by a constant independent of θ. □

A.4 Remaining Details of Numerical Experiment

The Poisson process that models the arrival from location j has rate $\gamma_j = \textit{pois-rate} \times \textit{mkt-share}_j$, where *pois-rate*$\in (0,1]$ is the total arrival rate and $\textit{mkt-share}_j$ is the conditional probability that this arrival comes from region j. We set *pois-rate* to be 0.9 and $\textit{mkt-share}_j$ to be the ratio between the total population in the j^{th} largest MSA and the total population of all fifteen MSA. A customer arriving from location j makes a purchase with probability $\exp(A_j + B_j p)$. The parameters of purchasing probabilities are chosen as follows: We first set "baseline" demand parameters A_1 and B_1. For all

$j \geq 2$, we then set $A_j = \frac{income_1}{income_j} \times A_1$ and $B_j = \frac{income_1}{income_j} \times B_1$, where $income_j$ represents the medium household income of the j^{th} largest MSA, as reported in U.S. Census Bureau (2014b). Since we want $\exp(A_j + B_j p) \leq \mathbf{1}$ for all $p \in \Omega_p$, we set A_j's to be vectors with negative components, and B_j's to be diagonally dominated matrices with negative diagonal components. The baseline parameters shown below are generated to satisfy these constraints. The absolute magnitude of their entries depends on the price range, which, in our setting, depends on the shipping cost.

$$A_1 = \begin{bmatrix} -1.0071 \\ -1.2603 \\ -1.3228 \\ -1.5005 \\ -1.4810 \end{bmatrix}, \quad B_1 = \begin{bmatrix} -9.5 & 1.1 & 1.1 & 1.1 & 1.7 \\ 1.9 & -10.5 & 2.0 & 1.4 & 1.0 \\ 1.1 & 1.6 & -11.5 & 2.0 & 1.9 \\ 2.0 & 2.0 & 1.5 & -12.6 & 2.0 \\ 1.6 & 2.0 & 1.8 & 2.0 & -12.1 \end{bmatrix} \times 10^{-3}$$

The transportation cost is calculated using the cost equation estimated in Section EC.3 in Jasin and Sinha (2015) assuming that each product weighs exactly one pound. To be precise, $c_{ijk} = \bar{c}_k \cdot (9.182 + 0.000541 \cdot d_{ij})$, where d_{ij} is the distance in miles from FC i to demand location j, and $\bar{c}_k$ is uniformly distributed in $[0.9, 1.1]$. We set the inventory levels of the FCs to minimize the likelihood that we use FCs that are far away from the demand location even under a myopic fulfillment policy; this is to prevent the separate optimization heuristic from performing too bad. To do so, we first match between FCs and MSAs such that (1) each FC serves five MSA, (2) each MSA is served by 2 FCs, and (3) the total mileage between all the assigned FC-MSA pairs is minimized. We then approximate an average purchase quantity from MSA j by $\hat{\lambda}_j = \textit{pois-rate} \times \textit{mkt-share}_j \times 0.9$. (The factor 0.9 means that the initial inventory levels are set to be slightly below the expected total arrivals; this reflects the common reality where firm stocks neither too low such that the induced demand has to be really scarce, nor too high as if there is no inventory constraint at all.) Each of the two FCs serving MSA_j fulfill a portion of the $\hat{\lambda}_j$, where the portion is decided by a random number drawn uniformly from $[0.4, 0.6]$. (Our results are robust with respect to perturbation in the numbers 0.9, 0.4, and 0.6.) The total initial inventory at each of the FC is then calculated as the sum of all the demand portions from the five MSAs it serves. At last, we distribute the initial inventory at each of the FC uniformly across all of the products. As a result, the initial inventory level is $C_{1k} = 0.0337$, $C_{2k} = 0.0218$, $C_{3k} = 0.0217$, $C_{4k} = 0.0276$ $C_{5k} = 0.0196$ and $C_{6k} = 0.0196$ for all $k = 1, \ldots, 5$. The fictitious FC is set to hold abundant initial inventories so that they will never be depleted. For a specific θ, we always round down θC_{ik}.

Table A.1 reports the expected profits of all the heuristics implemented in Chapter 2.8. The coefficient of variations are consistently small (less than 0.5% for all instances); due to the space constraint, we will not report them in the paper.

Table A.2 reports the running time of a single simulation for several different heuristics when $\theta = 2000$. The computation time for the last two heuristics is very long, therefore it is not feasible to implement them in practice. All simulations were implemented on a desktop computer with 3.40GHz Intel Core i7-3770 CPU and 8 GB of RAM.

θ	RPFC-2	R^2PF-2	R^2PF-Ful-2	R^2PF-Pr-2	ALP-Reopt-2
200	9430.3	11188.2	10685.9	9573.5	11082.3
400	21637.1	24584.6	23772.8	22762.9	24509.0
600	34556.3	38487.0	37361.6	35861.0	38445.3
800	47128.5	51096.7	50420.3	49233.1	51269.1
1000	59421.7	63915.5	63231.5	61764.5	64106.5
1300	78355.8	83891.0	83296.8	81244.5	84184.8
1600	97707.0	104095.7	103189.5	101556.2	104209.7
2000	123868.1	130583.3	129244.4	127404.6	131023.8

θ	RPF-5	R^2PF-5	R^2PF-Ful-5	R^2PF-Pr-5	ALP-Reopt-5
200	8654.1	11174.6	10243.5	9625.1	10875.5
400	21301.9	24704.3	23913.0	23170.0	25219.7
600	34326.4	38675.3	38190.4	36447.4	39409.5
800	47354.4	52239.4	51661.2	50155.6	53119.2
1000	60013.1	66210.0	64542.2	63859.5	66763.0
1300	80807.3	86832.0	85417.4	83893.7	88029.6
1600	100438.0	107811.5	106475.0	104595.7	108747.1
2000	127255.7	136083.0	134312.8	131924.8	136999.7

θ	RPF-8	R^2PF-8	R^2PF-Ful-8	R^2PF-Pr-8
200	8452.7	10944.4	10293.8	9095.4
400	20844.2	25130.5	24415.4	21773.3
600	34442.5	39717.1	38684.6	35845.4
800	47555.6	53663.8	52378.8	49914.0
1000	60280.1	66685.4	65582.1	63264.4
1300	80988.6	87803.4	86609.7	84359.5
1600	100663.0	109225.9	108053.1	105571.4
2000	128223.2	137391.0	135841.8	132286.8

θ	Sep-reopt	DJPF-Reopt-1	DJPF-Reopt-10
200	6029.6	8409.3	11402.7
400	14857.1	20535.8	23762.8
600	24084.7	32610.1	35771.6
800	32596.9	44457.0	49576.2
1000	40625.2	57977.1	61993.9
1300	53726.0	77196.4	81677.3
1600	66782.8	96125.3	101167.5
2000	84005.0	122814.6	128769.5

Table A.1: Expected Profit of Different Heuristics with Varying θ

R^2PFC-2	R^2PF-5	R^2PF-8	DJPF-Reopt-θ
14.98	23.12	23.69	58376.24
ALP-Reopt-2	ALP-Reopt-5	ALP-Reopt-8	
26.98	992.87	10814.08	

Table A.2: Typical Running Time (in seconds) for a Single Simulation for Selected Heuristics

APPENDIX B

Appendix to Chapter 3

B.1 Proof of Lemma 3.4.1

Consider any admissible control $\pi \in \Pi$. Per our notations above, π essentially corresponds to the demand rate sequence $\{\lambda_t^\pi\}_{t=1}^T$. By definition, the sequence $\{\lambda_t^\pi\}_{t=1}^T$ satisfies the capacity constraints in DET. Moreover, we know from Assumption A4 and Jensen's inequality that

$$\begin{aligned}\mathbf{E}\left[\sum_{t=1}^{T} r(\lambda_t^\pi)\right] &= \sum_{t=1}^{T}\mathbf{E}\left[p_t(\lambda_t^\pi)\cdot D_t(\lambda_t^\pi)\right] = \sum_{t=1}^{T}\mathbf{E}\left[\mathbf{E}\left[p_t(\lambda_t^\pi)\cdot D_t(\lambda_t^\pi)|\mathcal{H}_t\right]\right]\\ &= \sum_{t=1}^{T}\mathbf{E}\left[r(\lambda_t^\pi)\right] \le \sum_{t=1}^{T} r\left(\mathbf{E}\left[\lambda_t^\pi\right]\right).\end{aligned}$$

Therefore, we conclude that $J^* = J^{\pi^*} \le J^D$. ■

B.2 Proof of Theorem 3.8.1

The proof of Theorem 3.8.1 follows similar arguments as in the proofs of Theorems 3.5.1 and 3.6.1. We still proceed in two steps: In the first step, we construct a high-probability event $\mathcal{G}$, and show that, on the set $\mathcal{G}$, we always have $C_t \ge 1$ and $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC-Batch($\mathbf{m}, \boldsymbol{\epsilon}$).

Step 1

We start with the first step. For some $\delta_k = o(m_k)$, whose value is to be determined later, define a sequence of events $\{\mathcal{A}_{k,i}(\epsilon_k, \delta_k)\}$ as follows:

$$\mathcal{A}_{k,i}(\epsilon_k,\delta_k) \;=\; \left\{ \max_{t\le i m_k} \left| \sum_{s=(i-1)m_k+1}^{t} \Delta_{s,k} \right| \;<\; \delta_k \right\} \quad \forall i,k \tag{B.1}$$

Analogous to (3.4), it can be shown that

$$\mathbf{P}\left(\bar{\mathcal{A}}_{k,i}(\epsilon_k,\delta_k)\right) \;\le\; 2\cdot\exp\{r_k^2\ \min\{C_k^D-\epsilon_k, m_k\} - r_k\delta_k\} \quad \forall r_k\in[0,1]. \tag{B.2}$$

Define $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta}) = \cap_{k=1}^{K} \cap_{i=1}^{T/m_k} \mathcal{A}_{i,k}(\epsilon_k,\delta_k)$, where $\boldsymbol{\delta} = (\delta_1,\ldots,\delta_K)$. By the sub-additivity property of probability, we have:

$$\mathbf{P}(\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})) \ge 1-2T\ \sum_{k=1}^{K} \frac{\exp\{r_k^2\ \min\{C_k^D-\epsilon_k,m_k\}-r_k\delta_k\}}{m_k}. \tag{B.3}$$

Now, we make some observations. First, on the set $\mathcal{G}(\boldsymbol{\epsilon},\boldsymbol{\delta})$, we always have: $\left|\frac{\epsilon_k}{n_k}+\frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i}}\Delta_{s,k}\right| \le \frac{\epsilon_k}{n_k}+\frac{\delta_k}{m}$ for all i and k. This means that, as long as the parameters ϵ_k, δ_k, and m_k are chosen such that $\frac{\epsilon_k}{n_k}+\frac{\delta_k}{m_k} \le \min\{\varphi_L,\varphi_U\}$, the condition $\lambda_{t,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k}\sum_{s\in\mathcal{T}_{k,i_k(t)}}\Delta_{s,k} \in \Omega_{\lambda,k}$ in Step 2 part a of DPC-Batch is always satisfied for all t. For the remaining of the proof, we will therefore assume that $\frac{\epsilon_k}{n_k}+\frac{\delta_k}{m_k} \le \min\{\varphi_L,\varphi_U\}$. Now, suppose that $t\in\mathcal{T}_{k,i_k}$ and $\max\{1,t-n_k+1\}\in\mathcal{T}_{k,j_k}$, where $t\in[n_1,T]$. We can write the total resource consumption by the end of period t as follows:

$$\begin{aligned}
&\sum_{k=1}^{K}\ \sum_{s=\max\{1,t-n_k+1\}}^{t} D_{s,k}(\hat{\mathbf{p}}_s^D)\\
&=\sum_{k=1}^{K}\left[\ \sum_{\substack{s\ge\max\{1,t-n_k+1\}\\ s\in\mathcal{T}_{k,j_k}}} D_{s,k}(\hat{\mathbf{p}}_s^D) \;+\; \sum_{j=j_k+1}^{i_k-1}\ \sum_{s\in\mathcal{T}_{k,j}} D_{s,k}(\hat{\mathbf{p}}_s^D) \;+\; \sum_{s\le t,\, s\in\mathcal{T}_{k,i_k}} D_{s,k}(\hat{\mathbf{p}}_s^D)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{K} \sum_{\substack{s \geq \max\{1, t-n_k+1\} \\ s \in \mathcal{T}_{k,j_k}}} \left(\lambda_{s,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k} \sum_{l \in \mathcal{T}_{k,j_k-1}} \Delta_{l,k} + \Delta_{s,k} \right) \\
&\quad + \sum_{k=1}^{K} \sum_{j=j_k+1}^{i_k-1} \sum_{s \in \mathcal{T}_{k,j}} \left(\lambda_{s,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k} \sum_{l \in \mathcal{T}_{k,j-1}} \Delta_{l,k} + \Delta_{s,k} \right) \\
&\quad + \sum_{k=1}^{K} \sum_{s \leq t,\, s \in \mathcal{T}_{k,i_k}} \left(\lambda_{s,k}^D - \frac{\epsilon_k}{n_k} - \frac{1}{m_k} \sum_{l \in \mathcal{T}_{k,i_k-1}} \Delta_{l,k} + \Delta_{s,k} \right) \\
&\leq \sum_{k=1}^{K} \sum_{s=\max\{1,\, t-n_k+1\}}^{t} \lambda_{s,k}^D - \sum_{k=1}^{K} \frac{n_1}{n_k} \epsilon_k \\
&\quad - \sum_{k=1}^{K} \left(\frac{j_k \cdot m_k - \max\{1, t-n_k+1\}}{m_k} \right) \cdot \left(\sum_{s \in \mathcal{T}_{k,j_k-1}} \Delta_{s,k} \right) \\
&\quad + \sum_{k=1}^{K} \left(1 - \frac{t-(i_k-1)m_k}{m_k} \right) \cdot \left(\sum_{s \in \mathcal{T}_{k,i_k-1}} \Delta_{s,k} \right) \\
&\quad + \sum_{k=1}^{K} \sum_{s \leq t,\, s \in \mathcal{T}_{k,i_k}} \Delta_{s,k} - \sum_{k=1}^{K} \sum_{\substack{s < \max\{1,\, t-n_k+1\} \\ s \in \mathcal{T}_{k,j_k}}} \Delta_{s,k} \\
&\leq C - \sum_{k=1}^{K} \frac{n_1}{n_k} \epsilon_k + \sum_{k=1}^{K} \left| \sum_{s \in \mathcal{T}_{k,j_k-1}} \Delta_{s,k} \right| + \sum_{k=1}^{K} \left| \sum_{\substack{s < \max\{1,\, t-n_k+1\} \\ s \in \mathcal{T}_{k,j_k}}} \Delta_{s,k} \right| \\
&\quad + \sum_{k=1}^{K} \left| \sum_{s \in \mathcal{T}_{k,i_k-1}} \Delta_{s,k} \right| + \sum_{k=1}^{K} \left| \sum_{s \leq t, s \in \mathcal{T}_{k,i_k}} \Delta_{s,k} \right|
\end{aligned}
$$

where the last inequality follows by the definition of i_k and j_k. On the set $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$, for each k, each of the terms with $|\cdot|$ above is at most δ_k. So, we can bound:

$$
\sum_{k=1}^{K} \sum_{s=\max\{1,\, t-n_k+1\}}^{t} D_{s,k}(\hat{\mathbf{p}}_s^D) \;\; \leq \;\; C - \sum_{k=1}^{K} \left(\frac{n_1}{n_k} \epsilon_k - 4\delta_k \right) \;\; \forall t \in [n_1, T]. \qquad \text{(B.4)}
$$

(For $t < n_1$, we can bound the total resource consumption by the end of period t with the total resource consumption by the end of period n_1. So, the above bound also holds.) Note that (B.4) is the analogue of (3.13) in the proof of Theorem 3.6.1. An immediate choice of $\boldsymbol{\delta}$ that guarantees our resource will never run out on the set $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$ is therefore $\delta_k = \frac{n_1 \epsilon_k}{4n_k} - \frac{1}{4K}$. Given this and the assumption $\frac{\epsilon_k}{n_k} + \frac{\delta_k}{m_k} \leq \min\{\varphi_L, \varphi_U\}$, we conclude that the following always hold on $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$: (i) $C_t \geq 1$ and (ii) $\lambda^D_{t,k} - \frac{\epsilon_k}{n_k} - \frac{1}{m_k} \sum_{s \in \mathcal{T}_{k, i_k(t)-1}} \Delta_{s,k} \in \Omega_{k,\lambda}$ for all t. Consequently, $\mathbf{p}_t = \hat{\mathbf{p}}^D_t$ for all t.

Step 2

We now ready bound the average regret of DPC-Batch$(\mathbf{m}, \boldsymbol{\epsilon})$. Let $\{\mathbf{p}_t\}$ be the sequence of price vector under DPC-Batch$(\mathbf{m}, \boldsymbol{\epsilon})$. As in Step 2 in the proof of Theorem 3.6.1, we have:

$$\mathbf{E}[R^{DPC-Batch(\mathbf{m},\boldsymbol{\epsilon})}] \;\; \geq \;\; \mathbf{E}\left[\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}^D_t)\right] - \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}^D_t)\right) \cdot \mathbf{1}\{\bar{\mathcal{G}}(\boldsymbol{\epsilon}, \boldsymbol{\delta})\}\right].$$

The second expectation after the last equality above can be bounded by $r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\boldsymbol{\epsilon}, \boldsymbol{\delta}))$ where $r^u = \max_t \max_{\boldsymbol{\lambda_t} \in \Omega_\lambda} r_t(\boldsymbol{\lambda}_t)$. As for the first expectation, by Taylor's expansion and Assumption MA6, we can bound:

$$\begin{aligned}
& \mathbf{E}[r_t(\hat{\mathbf{p}}^D_t)] \\
= \;\; & \mathbf{E}\left[r_t\left(\lambda^D_{t,1} - \frac{\epsilon_1}{n_1} - \frac{1}{m_1}\sum_{s \in \mathcal{T}_{1,i_1(t)-1}} \Delta_{s,1}, \;\cdots\;, \lambda^D_{t,K} - \frac{\epsilon_K}{n_K} - \frac{1}{m_K}\sum_{s \in \mathcal{T}_{K,i_K(t)-1}} \Delta_{s,K}\right)\right] \\
\geq \;\; & r_t\left(\boldsymbol{\lambda}^D_t\right) - \Psi \sum_{k=1}^{K} \frac{\epsilon_k}{n_k} - \Psi \cdot \sum_{k=1}^{K} \mathbf{E}\left[\left(\frac{\epsilon_k}{n_k} + \frac{1}{m_k}\sum_{s \in \mathcal{T}_{k,i_k(t)-1}} \Delta_{s,k}\right)^2\right] \\
\geq \;\; & r_t\left(\boldsymbol{\lambda}^D_t\right) - \Psi \sum_{k=1}^{K}\left(\frac{\epsilon_k}{n_k} + \frac{2\epsilon_k^2}{n_k^2} + \frac{2}{m_k}\right)
\end{aligned}$$

where the first inequality follows from Assumption MA6; the last inequality follows because $(x+y)^2 \leq 2x^2 + 2y^2$ for all (x,y) and $\mathbf{E}\left[\left(\sum_{s \in \mathcal{T}_{k,i_k(t)-1}} \Delta_{s,k}\right)^2\right] \leq m_k$. Putting the bounds together, for all $r_k \in [0,1]$, we have:

$$\begin{aligned}
& \frac{J^D_M - \mathbf{E}[R^{DPC-Batch(\mathbf{m},\boldsymbol{\epsilon})}]}{T} \\
\leq \;\; & \frac{1}{T} \cdot \left[T\Psi \sum_{k=1}^{K}\left(\frac{\epsilon_k}{n_k} + \frac{2\epsilon_k^2}{n_k^2} + \frac{2}{m_k}\right) + r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon, \delta))\right] \\
\leq \;\; & \sum_{k=1}^{K}\left(\frac{\Psi\epsilon_k}{n_k} + \frac{2\Psi\epsilon_k^2}{n_k^2} + \frac{2\Psi}{m_k} + \frac{2r^u T}{m_k} \cdot \exp\{r_k^2 \min\{C^D_k - \epsilon_k, m_k\} - r_k \delta_k\}\right).
\end{aligned}$$

Taking $r_k = \frac{\delta_k}{2\min\{C_k^D - \epsilon_k, m_k\}}$ and substituting $\delta_k = \frac{n_1\epsilon_k}{4n_k} - \frac{1}{4K}$ yields:

$$\begin{aligned}
&\frac{J_M^D - \mathbf{E}[R^{DPC-Batch(\mathbf{m},\boldsymbol{\epsilon})}]}{T} \\
\leq \quad & M_3 \cdot \sum_{k=1}^{K} \left[\frac{\epsilon_k}{n_k} + \frac{1}{m_k} + \frac{T}{m_k} \cdot \exp\left\{ - \frac{(Kn_1\epsilon_k - n_k)^2}{64K^2 n_k^2 \min\{C_k^D - \epsilon, m_k\}} \right\} \right]
\end{aligned}$$

for some $M_3 > 0$ independent of T, C, m_k, $n_1 \geq \frac{1}{K\min\{\varphi_L, \varphi_U\}}$, and

$$\epsilon_k \in n_k \cdot \left[\frac{1}{Kn_1}, \min\left\{ 1, \frac{1}{K} \cdot \frac{1 + 4Km_k \cdot \min\{\varphi_L, \varphi_U\}}{4m_k + n_1} \right\} \right]$$

(Note: $\epsilon_k \geq \frac{n_k}{Kn_1}$ is needed to guarantee that $\delta_k = \frac{n_1\epsilon_k}{4n_k} - \frac{1}{4K} \geq 0$ and $n_1 \geq \frac{1}{K\min\{\varphi_L,\varphi_U\}}$ is needed to guarantee that $\frac{1}{Kn_1} \leq \frac{1}{K} \cdot \frac{1+4Km_k \cdot \min\{\varphi_L,\varphi_U\}}{4m_k+n_1}$.) ■

B.3 Proof of Theorem 3.9.1

In this section, we prove the two bounds presented in (3.16) and (3.17). The proofs of these bounds are similar, and follow similar arguments as in the proof of Theorems 3.6.1. In what follows, we first show (3.16) in two steps: In the first step, we construct a high-probability event $\mathcal{G}$, and show that, on the set $\mathcal{G}$, we always have $C_t \geq 1$ and $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ for all t. In the second step, we bound the total revenue losses under DPC-Batch(m, ϵ) followed by a brief discussion on a crucial observation for deriving the bound in (3.18). Finally, we will comment on which parts of the proof of (3.16) need to be modified to show (3.17).

Proof of (3.16): Step 1

For some $\delta = o(m)$ whose exact value is to be determined later, define $\{\mathcal{A}_{i,\ell}(\epsilon,\delta)\}$ as follows:

$$\mathcal{A}_{i,\ell}(\epsilon, \delta) \;=\; \left\{ \max_{t \leq im} \left| \sum_{s=(i-1)m+1}^{t} \Delta_{s,\ell} \right| \;<\; \delta \right\} \quad \forall i, \ell. \tag{B.5}$$

Analogous to (3.4), it can be shown that

$$\mathbf{P}\left(\bar{\mathcal{A}}_{i,\ell}(\epsilon, \delta)\right) \;\leq\; 2 \cdot \exp\{r^2 \min\{C - \epsilon, m\} - r\delta\} \quad \forall r \in [0, 1]. \tag{B.6}$$

Define $\mathcal{G}(\epsilon,\delta) = \cap_{\ell=0}^{L} \cap_{i=1}^{T/m} \mathcal{A}_{i,\ell}(\epsilon,\delta)$. By the sub-additivity property of probability,

$$\mathbf{P}(\mathcal{G}(\epsilon,\delta)) \geq 1 - \frac{2T(L+1)}{m} \exp\{r^2 \min\{C-\epsilon, m\} - r\delta\}. \tag{B.7}$$

Note that, on the set $\mathcal{G}(\epsilon,\delta)$, we always have: $\left|\frac{\epsilon}{n(L+1)} + \frac{1}{m}\sum_{s\in\mathcal{T}_i} \Delta_{s,\ell}\right| \leq \frac{\epsilon}{n(L+1)} + \frac{\delta}{m}$ for all i and ℓ. This means that, as long as the parameters ϵ, δ, and m are chosen such that $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$, the condition $\lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m}\sum_{s\in\mathcal{T}_i} \Delta_{s,\ell} \in \Omega_{\lambda,\ell}$ in Step 2 part a of DPC-Batch is always satisfied for all i. For the remaining of the proof, we will therefore assume that $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$. Now, suppose that $t-\ell \in \mathcal{T}_{i_\ell}$ (if $t-\ell \leq 0$, then we set $i_\ell = 0$) and $\max\{1, t-\ell-n+1\} \in \mathcal{T}_{j_\ell}$, where $n \leq t \leq T$. (For $t < n$, we can bound total resource consumption by the end of period t with the total resource consumption by the end of period n.) We can bound total consumption of resource by the end of period t as follows:

$$\begin{aligned}
&\sum_{\ell=0}^{L} \sum_{s=\max\{1,\, t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \\
&= \sum_{\ell=0}^{L} \left[\sum_{\substack{s\geq\max\{1,t-\ell-n+1\} \\ s\in\mathcal{T}_{j_\ell}}} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \; + \sum_{j=j_\ell+1}^{i_\ell-1} \sum_{s\in\mathcal{T}_j} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \; + \sum_{s\leq t-\ell,\, s\in\mathcal{T}_{i_\ell}} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \right] \\
&= \sum_{\ell=0}^{L} \sum_{\substack{s\geq\max\{1,t-\ell-n+1\} \\ s\in\mathcal{T}_{j_\ell}}} \left(\lambda_{s,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{v\in\mathcal{T}_{j_\ell-1}} \Delta_{v,\ell} + \Delta_{s,\ell} \right) \\
&\quad + \sum_{\ell=0}^{L} \sum_{j=j_\ell+1}^{i_\ell-1} \sum_{s\in\mathcal{T}_j} \left(\lambda_{s,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{v\in\mathcal{T}_{j-1}} \Delta_{v,\ell} + \Delta_{s,\ell} \right) \\
&\quad + \sum_{\ell=0}^{L} \sum_{s\leq t,\, s\in\mathcal{T}_{i_\ell}} \left(\lambda_{s,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m} \sum_{v\in\mathcal{T}_{i_\ell-1}} \Delta_{v,\ell} + \Delta_{s,\ell} \right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\ell=0}^{L} \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} \lambda_{s,\ell}^{D} - \sum_{\ell=0}^{L} \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} \frac{\epsilon}{n(L+1)} \\
&\quad - \sum_{\ell=0}^{L} \left(\frac{j_\ell \cdot m - \max\{1, t-\ell-n+1\}}{m} \right) \cdot \left(\sum_{s \in \mathcal{T}_{j_\ell - 1}} \Delta_{s,\ell} \right) \\
&\quad + \sum_{\ell=0}^{L} \left(1 - \frac{t-\ell-(i_\ell - 1)m}{m} \right) \cdot \left(\sum_{s \in \mathcal{T}_{i_\ell - 1}} \Delta_{s,\ell} \right) \\
&\quad - \sum_{\ell=0}^{L} \sum_{\substack{s < \max\{1, t-\ell-n+1\}, \\ s \in \mathcal{T}_{j_\ell}}} \Delta_{s,\ell} + \sum_{\ell=0}^{L} \sum_{s \le t-\ell, s \in \mathcal{T}_{i_\ell}} \Delta_{s,\ell} \\
&\le C - \frac{\sum_{s=\max\{1,n-L\}}^{n} s}{n(L+1)} \cdot \epsilon + \sum_{\ell=0}^{L} \left| \sum_{s \in \mathcal{T}_{j_\ell}} \Delta_{s,\ell} \right| + \sum_{\ell=0}^{L} \left| \sum_{\substack{s < \max\{1, t-\ell-n+1\} \\ s \in \mathcal{T}_{j_\ell}}} \Delta_{s,\ell} \right| \\
&\quad + \sum_{\ell=0}^{L} \left| \sum_{s \in \mathcal{T}_{i_\ell - 1}} \Delta_{s,\ell} \right| + \sum_{\ell=0}^{L} \left| \sum_{s \le t, s \in \mathcal{T}_{i_\ell}} \Delta_{s,\ell} \right| . \qquad \text{(B.8)}
\end{aligned}
$$

where the inequality follows from the definition of i_ℓ, j_ℓ, and the fact that

$$
\sum_{\ell=0}^{L} \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} 1 \ge \sum_{s=\max\{1,n-L\}}^{n} s \text{ for all } t \ge n.
$$

Note that $L \le n$ implies

$$
\begin{aligned}
\sum_{s=\max\{1,n-L\}}^{n} s &= \sum_{s=n-L}^{n} s = \frac{n(n+1)}{2} - \frac{(n-L)(n-L-1)}{2} \\
&= \frac{(2n-L)(L+1)}{2} \ge \frac{n(L+1)}{2}.
\end{aligned}
$$

Moreover, on the set $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$, the terms with $|\cdot|$ in (B.8) are all bounded by δ. Thus, we have

$$
\sum_{\ell=0}^{L} \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \;\le\; C - \frac{1}{2} \cdot \epsilon + 4(L+1)\delta \quad \text{for all } t \ge n \; . \qquad \text{(B.9)}
$$

(B.9) is the analogue of (3.13) in the proof of Theorem 3.6.1. An immediate choice of δ that guarantees our resource will never run out on the set $\mathcal{G}(\boldsymbol{\epsilon}, \boldsymbol{\delta})$ is therefore $\delta = \frac{\epsilon-1}{8(L+1)}$. Given this and the assumption $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$, we conclude that the following always hold on $\mathcal{G}(\epsilon, \delta)$: (i) $C_t \geq 1$ and (ii) $\lambda_{t,\ell}^D - \frac{\epsilon}{n(L+1)} - \frac{1}{m}\sum_{s \in \mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell} \in \Omega_{\lambda,\ell}$. Consequently, we have $\mathbf{p}_t = \hat{\mathbf{p}}_t^D$ for all t.

Proof of (3.16): Step 2

We now ready to bound the average regret of DPC-Batch(m, ϵ). Let $\{\mathbf{p}_t\}$ be the price sequence under DPC-Batch(m, ϵ). By the same argument as in Step 2 of the proof of Theorem 3.6.1, we have

$$\mathbf{E}[R^{DPC-Batch(m,\epsilon)}] \;\geq\; \mathbf{E}\left[\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}_t^D)\right] - \mathbf{E}\left[\left(\sum_{t=1}^{T} r_t(\hat{\mathbf{p}}_t^D)\right) \cdot \mathbf{1}\{\bar{\mathcal{G}}(\epsilon, \delta)\}\right].$$

The second expectation after the last equality above can be bounded by $r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\boldsymbol{\epsilon}, \boldsymbol{\delta}))$ where $r^u = \max_t \max_{\boldsymbol{\lambda_t} \in \Omega_\lambda} r_t(\boldsymbol{\lambda}_t)$. As for the first expectation, by Taylor's expansion and Assumption MA6, we can bound

$$\begin{aligned}
&\mathbf{E}[r_t(\hat{\mathbf{p}}_t^D)] \\
\geq\;& r_t(\boldsymbol{\lambda}_t^D) - \Psi\frac{\epsilon}{n} - \Psi \cdot \sum_{\ell=0}^{L} \mathbf{E}\left[\left(\frac{\epsilon}{n(L+1)} + \frac{1}{m}\sum_{s \in \mathcal{T}_{i_\ell(t)-1}} \Delta_{s,\ell}\right)^2\right] \\
\geq\;& r_t(\boldsymbol{\lambda}_t^D) - \Psi\left[\frac{\epsilon}{n} + \frac{2\epsilon^2}{n^2(L+1)} + \frac{2(L+1)}{m}\right]
\end{aligned}$$

where the first inequality follows from Assumption MA6; the last inequality follows because $(x+y)^2 \leq 2x^2 + 2y^2$ for all (x, y) and $\mathbf{E}\left[\left(\sum_{s \in \mathcal{T}_{i_\ell - 1}} \Delta_{s,\ell}\right)^2\right] \leq m$ for all ℓ.

Putting the bounds together, for all $r \in [0, 1]$, we have:

$$\begin{aligned}
&\frac{J_A^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T(L+1)} \\
\leq\;& \frac{1}{T(L+1)} \cdot \left[T\Psi\left(\frac{\epsilon}{n} + \frac{2\epsilon^2}{n^2(L+1)} + \frac{2(L+1)}{m}\right) + r^u T \cdot \mathbf{P}(\bar{\mathcal{G}}(\epsilon, \delta))\right] \\
\leq\;& \frac{\Psi\epsilon}{n(L+1)} + \frac{2\Psi\epsilon^2}{n^2(L+1)^2} + \frac{2\Psi}{m} + \frac{2r^u T}{m}\; \exp\{r^2\; \min\{C - \epsilon, m\} - r\delta\}.
\end{aligned}$$

Taking $r = \frac{\delta}{2\min\{C-\epsilon, m\}}$ and substituting $\delta = \frac{\epsilon-1}{8(L+1)}$ yield:

$$\begin{aligned}
&\frac{J_A^D - \mathbf{E}[R^{DPC-Batch(m,\epsilon)}]}{T(L+1)} \\
\leq\ & M_4 \left[\frac{\epsilon}{n(L+1)} + \frac{1}{m} + \frac{T}{m} \cdot \exp\left\{ -\frac{(\epsilon-1)^2}{256(L+1)^2 \min\{C-\epsilon,\, m\}} \right\} \right]
\end{aligned}$$

for some $M_4 > 0$ for all $\epsilon \in \left[1, \min\left\{ n(L+1), m(L+1), n \cdot \frac{8m(L+1)\min\{\varphi_L,\varphi_U\}+1}{8m+n} \right\}\right]$, T, C, m, and $L < n$. (Note that $\epsilon \leq \min\left\{ m(L+1), n \cdot \frac{8m(L+1)\min\{\varphi_L,\varphi_U\}+1}{8m+n} \right\}$ and $\delta = \frac{\epsilon-1}{8(L+1)}$ imply $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$ and $r \in (0,1)$.)

Proof of (3.17)

We now prove the bound for the case $L > n$. The major difference between the proof of (3.16) and (3.17) lies in the way we bound the total resource consumption in (B.8). We first discuss why this is important in dealing with large L. On the RHS of (B.8), the negative term after C is an upper bound for negative total buffers in DPC-Batch (i.e., the term $-\frac{\epsilon}{n(L+1)}$ in the definition of $\lambda_{t,\ell}(\hat{\mathbf{p}}_t^D)$) and the remaining four positive terms is an upper bound for total random errors. If $L > n$, the term $\frac{\sum_{s=\max\{1,n-L\}}^{n} s}{n(L+1)}$ in (B.8) equals $\frac{\sum_{s=1}^{n} s}{n(L+1)} = \frac{n+1}{2(L+1)}$ and the bound in (B.9) becomes

$$\sum_{\ell=0}^{L} \sum_{s=\max\{1,\, t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \ \leq\ C - \frac{n+1}{2(L+1)} \cdot \epsilon + 4(L+1)\delta.$$

Since $\epsilon \leq n(L+1)$ (otherwise $\hat{\mathbf{p}}_t^D$ is not well-defined), the size of ϵ is at most on the order of n^2. Per our argument in Step 1 in the proof of (3.16), δ represents an upper bound of the total errors of m Bernoulli random variables (for some m), which means that $\delta = \Omega(1)$. But then, $4(L+1)\delta$ is $\Omega(L)$ and we cannot always guarantee $C - \frac{n+1}{2(L+1)} \cdot \epsilon + 4(L+1)\delta \leq C$ for all large $L > n^2$ (i.e., we may not be able to find a feasible $\epsilon \leq n(L+1)$ such that $-\frac{n+1}{2(L+1)} \cdot \epsilon + 4(L+1)\delta \leq 0$). This calls for a more careful analysis on the bound of total resource consumption.

Note that, assuming we never apply $\bar{p}$ up to and including period $t - \ell \geq 0$, total resource consumption of type-ℓ request by the end of period t is $\sum_{s=\max\{1,\, t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D)$. We divide our analysis into three cases: $n \leq t \leq L+1$, $L+1 < t \leq n+L$, and $t > n+L$. (For $t < n$, we can bound total resource consumption by the end of period t with the total resource consumption by the end of period n.) When $n \leq t \leq L+1$, all type-ℓ requests with $\ell \geq t$ have not consumed any resource yet. For $0 \leq \ell < t$, following similar arguments as in (B.8), we have

$$\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \begin{cases} \sum_{s=t-\ell-n+1}^{t-\ell} \lambda_{s,\ell}^D - \frac{n}{n(L+1)} \cdot \epsilon + 4\delta & \text{if } 0 \leq \ell < t-n \\ \sum_{s=1}^{t-\ell} \lambda_{s,\ell}^D - \frac{t-\ell}{n(L+1)} \cdot \epsilon + 2\delta & \text{if } t-n \leq \ell < t \end{cases}$$

When $L+1 < t \leq n+L$, all type-ℓ requests (for all $\ell \in \{0,1,\ldots,L\}$) have already consumed some of the resources. Following similar arguments as in (B.8), we have

$$\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \begin{cases} \sum_{s=t-\ell-n+1}^{t-\ell} \lambda_{s,\ell}^D - \frac{n}{n(L+1)} \cdot \epsilon + 4\delta & \text{if } 0 \leq \ell < t-n \\ \sum_{s=1}^{t-\ell} \lambda_{s,\ell}^D - \frac{t-\ell}{n(L+1)} \cdot \epsilon + 2\delta & \text{if } t-n \leq \ell \leq L \end{cases}$$

At last, when $t > n+L$, following similar arguments as in (B.8), we have

$$\sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \sum_{s=t-\ell-n+1}^{t-\ell} \lambda_{s,\ell}^D - \frac{n}{n(L+1)} \cdot \epsilon + 4\delta.$$

Given all the above bounds, the total resource consumption by the end of period $t \geq n$ can be bounded as follows:

$$\sum_{\ell=0}^{L} \sum_{s=\max\{1,t-\ell-n+1\}}^{t-\ell} D_{s,\ell}(\hat{\mathbf{p}}_s^D) \leq \begin{cases} C \;-\; \frac{2t-n+1}{2(L+1)} \cdot \epsilon \;+\; 2 \cdot (2t-n+1) \cdot \delta & \text{if } n \leq t \leq L+1 \\ C \;-\; \frac{n(L+1) - \frac{(L-t+n)(L-t+n+1)}{2}}{n(L+1)} \cdot \epsilon \;+\; 2 \cdot (t+L-n+2) \cdot \delta & \\ & \text{if } L+1 < t \leq n+L \\ C \;-\; \epsilon \;+\; 4(L+1)\delta & \text{if } n+L < t \leq T \end{cases}$$

We claim that, if we set $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ and $\epsilon \geq 2(L+1)$, total resource consumption by the end of period $t \geq n$ is at most $C-1$. To see this, when $n \leq t \leq L+1$, substituting $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ yields

$$\frac{2t-n+1}{2(L+1)} \cdot \epsilon \;-\; 2 \cdot (2t-n+1) \cdot \delta = 2 \cdot (2t-n+1) \left(\frac{\epsilon}{4(L+1)} - \delta \right)$$
$$\frac{2t-n+1}{n} \geq 1.$$

When $L+1 < t \leq n+L$, substituting $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ yields

$$\begin{aligned}
&\frac{n(L+1) - \frac{(L-t+n)(L-t+n+1)}{2}}{n(L+1)} \cdot \epsilon \;-\; 2 \cdot (t+L-n+2) \cdot \delta \\
&\quad = \epsilon \;-\; 4(L+1)\delta \;+\; \frac{(L-t+n)}{2n(L+1)} \left[4n(L+1)\delta \;-\; (L-t+n+1) \cdot \epsilon\right] \\
&\quad = \frac{2(L+1)}{n} \;+\; \frac{(L-t+n)}{2n(L+1)} \left[(t-L-1)\epsilon \;-\; 2(L+1)\right] \\
&\quad \geq \frac{2(L+1)}{n} + \frac{(L-t+n)(t-L-2)}{n} \;\geq\; 1
\end{aligned}$$

where the last inequality holds since $t > L+1, \epsilon > 2(L+1)$. Finally, when $t > n+L$, substituting $\delta = \frac{\epsilon}{4(L+1)} - \frac{1}{2n}$ yields $\epsilon \;-\; 4(L+1)\delta = 2 \cdot \frac{L+1}{n} > 1$.

Now, plug the choice of δ into (B.7) and substituting $r = \frac{\delta}{2\min\{C-\epsilon,\, m\}}$, we can bound

$$\mathbf{P}(\mathcal{G}(\epsilon, \delta)) \geq 1 - \frac{2T(L+1)}{m} \;\exp\left\{-\frac{(\epsilon - 2(L+1)/n)^2}{64(L+1)^2 \min\{C-\epsilon,\, m\}}\right\}.$$

The remaining arguments are the same as in Step 2 of the proof of (3.17). Note that $\epsilon \in (L+1) \cdot \left[2, \min\left\{n, m, \frac{4mn\min\{\varphi_L, \varphi_U\}+2}{4m+n}\right\}\right]$ ensures $\frac{\epsilon}{n(L+1)} + \frac{\delta}{m} \leq \min\{\varphi_L, \varphi_U\}$ and $r \in (0,1)$. ■

B.4 Remaining Details of Numerical Experiment

We first give a detailed definition of LRC-k. Similar with DPC-Batch, we slice the selling horizon into batches, each of which is of size k (except for the last one), i.e. $\mathcal{T}_i = [(i-1)\cdot k+1, \min\{i\cdot k, T\}]$, for all $i = 1,\ldots,\lceil T/k\rceil$. LRC-$k$ is defined as follows

Linear Rate Control with Batch Restarting (LRC-k)

Step 1. Solve **DET** and get λ^D.

Step 2. At the beginning of each t, if $t \in \mathcal{T}_i$, do:

a. If $C_t \geq 1$ and $\lambda_t^D - \frac{1}{m}\sum_{s\in\mathcal{T}_i, s<t} \Delta_s/[\mathcal{T}_i - (s-(i-1)\cdot k)] \in \Omega_\lambda$, set $p_t = \hat{p}_t^D$ where

$$\lambda_t(\hat{p}_t^D) \;=\; \lambda_t^D - \frac{1}{m}\sum_{s\in\mathcal{T}_i\;,\;s<t}\frac{\Delta_s}{\mathcal{T}_i - (s-(i-1)\cdot k)};$$

b. Otherwise, set $p_t = \bar{p}$.

At last, we provide the numerical results of experiment 1 in Table B.1. We only show the results LRC-n, since, compared to the other heuristics, LRC-k is similarly worse as LRC-n for any k.

Table B.1: Expected Regret of Different Heuristics with Varying n

n	DPC-0				DPC-ϵ				
	Regret	Std	AvgReg(%)	Runtime (ms)	Regret	Std	AvgReg(%)	Runtime (ms)	Opt. ϵ
500	824	0.68	2.03	0.5	420	1.27	1.04	0.5	0.29
1000	1182	0.67	1.46	1.1	580	1.26	0.72	0.9	0.32
2000	1672	0.65	1.03	2.1	773	1.18	0.48	2.0	0.26
3000	2080	0.66	0.86	3.2	955	1.20	0.39	3.0	0.26
4000	2398	0.65	0.74	4.2	1089	1.14	0.34	4.0	0.27
5000	2708	0.67	0.67	5.1	1226	1.12	0.30	4.8	0.29
6000	2983	0.66	0.61	6.4	1342	1.17	0.28	5.7	0.31
7000	3228	0.73	0.57	7.3	1433	1.27	0.25	6.4	0.31
8000	3406	0.70	0.53	8.4	1522	1.27	0.23	7.5	0.28
n	LRC-n				DPCB-ϵ				
	Regret	Std	AvgReg(%)	Runtime (ms)	Regret	Std	AvgReg(%)	Runtime (ms)	Opt. ϵ
500	13662	8.8	33.75	19.34	390	1.27	0.96	4.4	0.17
1000	28971	15.6	35.78	39.04	461	1.26	0.57	9.5	0.17
2000	60868	25.3	37.59	75.51	542	1.18	0.33	17.5	0.16
3000	91693	31.6	37.75	119.82	660	1.20	0.27	25.7	0.17
4000	123529	36.5	38.14	157.39	751	1.14	0.23	35.4	0.15
5000	156829	44.3	38.74	197.40	816	1.12	0.20	42.8	0.13
6000	183936	42.3	37.86	236.34	874	1.17	0.18	54.2	0.15
7000	222871	53.9	39.32	277.82	879	1.27	0.16	62.3	0.14
8000	248279	49.8	38.33	328.38	919	1.27	0.14	70.1	0.13

APPENDIX C

Appendix to Chapter 4

C.1 Proof of Theorem 4.4.3

C.1.1 Bounding the Revenue Loss in SA-BDPA Upon Entering Step 4a

Following the same arguments as in the proof of Theorem 4.4.2, we know that

$$
\begin{aligned}
J_\theta^{MSA-BDPA} \quad \geq \quad & \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^u} \left[R_k(p_k^u + c_k^u) + R_k(p_k^u - c_k^u)\right] \mathbf{1}\left\{E_1 \cap E_2\right\}\right] \\
& -2\bar{p}\,\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \qquad \text{(C.1)}
\end{aligned}
$$

where $\tau_\theta^u := \left\lfloor \frac{T_\theta - 4\sum_{k=1}^{\tau} N_k}{2} \right\rfloor$ and E_1 and E_2 are as defined in Lemma 4.5.1 and Lemma 4.5.2, respectively.

We start with bounding the first term, which is the direct revenue incurred by MSA-DPA. Note that, for all p, we have $r(p^u) - r(p) \leq \frac{M_U K^2}{2}(p^u - p)^2$. So,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{k=1}^{\tau_\theta^u}\left[R_k(p_k^u+c_k^u)+R_k(p_k^u-c_k^u)\right]\mathbf{1}\left\{E_1\cap E_2\right\}\right] \\
&\geq \sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p_k^u+c_k^u)\mathbf{1}\left\{E_1\cap E_2\right\}\right]+\sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p_k^u-c_k^u)\mathbf{1}\left\{E_1\cap E_2\right\}\right] \\
&\geq \sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p^u)-\frac{M_U K^2}{2}(p^u-p_k^u-c_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right] \\
&\quad+\sum_{k=1}^{\tau_\theta^u}\mathbb{E}\left[r(p^u)-\frac{M_U K^2}{2}(p^u-p_k^u+c_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right] \\
&\geq 2\ \tau_\theta^u\ r(p^u)-2M_U K^2\left[\sum_{k=1}^{\tau_\theta^u}E\left[(p^u-p_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right]+(c_k^u)^2\right] \\
&\geq r(p^u)T_\theta-\bar{p}\ (2+\log^3 T_\theta)-2M_U K^2\left[\sum_{k=1}^{\tau_\theta^u}E\left[(p^u-p_k^u)^2\mathbf{1}\left\{E_1\cap E_2\right\}\right]+(c_k^u)^2\right],
\end{aligned}
$$

where the last inequality follows because, by definition of τ_θ and τ_θ^u, we have $2\tau_\theta^u \geq T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta} - 2$. As for the second term in (C.1), which is the total penalty incurred by capacity violation, similar to the arguments in Step 2 in section 4.2, for sample paths on $E_1 \cap E_2$, we can bound

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta}D_t(p_t)-C_\theta\right)^+\mathbf{1}\{E_1\cap E_2\}\right] \\
\leq\ &\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta}D_t(p_t)-\lambda(p_t)\right)^+\right] \\
&+\mathbb{E}\left[\left(\sum_{t=1}^{4\sum_{k=1}^{\tau_\theta}N_{k,\theta}}\lambda(p_t)-4\sum_{k=1}^{\tau_\theta}N_{k,\theta}\frac{C}{T}\right)^+\mathbf{1}\{E_1\cap E_2\}\right] \\
&+\mathbb{E}\left[\left(\sum_{t=4\sum_{k=1}^{\tau_\theta}N_{k,\theta}+1}^{T_\theta}\lambda(p_t)-\left(T_\theta-4\sum_{k=1}^{\tau_\theta}N_{k,\theta}\right)\frac{C}{T}\right)^+\mathbf{1}\{E_1\cap E_2\}\right] \\
\leq\ &\sqrt{T_\theta}+\log^3 T_\theta+0=O(\sqrt{\theta}),
\end{aligned}
$$

where the third inequality follows from Cauchy-Schwarz inequality, the definition of τ_θ, and the fact that p^c is to the left of I_k^u for all k.

Then, the total penalty for capacity violation satisfies

$$
\begin{aligned}
& \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \\
= \;& \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] + \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2}\}\right] \\
\leq \;& O(\sqrt{\theta}) + \; T_\theta P(\overline{E_1 \cap E_2}) = O(\sqrt{\theta}).
\end{aligned}
$$

Finally, combining the bounds for the two terms in (C.1), we get

$$
\begin{aligned}
& J_\theta^{MSA-BDPA} \\
\geq \;& r(p^u)T_\theta - 2M_U K^2 \left[\sum_{k=1}^{\tau_\theta^u} \mathbb{E}\left[(p^u - p_k^u)^2 \mathbf{1}\left\{E_1 \cap E_2\right\}\right] + (c_k^u)^2\right] - O(\sqrt{\theta}).
\end{aligned}
$$

Applying the standard result in Stochastic Approximation (e.g. Proposition 1 in Broadie et al. 2011), there exists positive constants C_a^u and C_c^u such that if $a_k^u = C_a/k$ and $c_k^u = C_c/k^{1/4}$ we have $\mathbb{E}[(p^u - p_k^u)^2 \mathbf{1}\left\{E_1 \cap E_2\right\}] \leq C_u/\sqrt{k}$, for all $k \geq 1$, where $C_u > 0$ is also a constant. Substitute this into the above bound, we get

$$
\begin{aligned}
J_\theta^{MSA-BDPA} \;&\geq\; r(p^u)T_\theta - 2M_U K^2 \sum_{k=1}^{\tau_\theta^u} \left(\frac{C_u}{\sqrt{k}} + \frac{C_c^2}{\sqrt{k}}\right) - O(\sqrt{\theta}) \\
&\geq\; r(p^u)T_\theta - 2M_U K^2 (C_u + C_c^2)\sqrt{\tau_\theta^u} - O(\sqrt{\theta}) \\
&\geq\; r(p^u)T_\theta - O(\sqrt{\theta}). \quad \square
\end{aligned}
$$

C.1.2 Bounding the Revenue Loss in SA-BDPA Upon Entering for Step 4b

Following the same arguments as in Step 1 in section 4.2, we know that

$$
J_\theta^{MSA-BDPA} \;\geq\; \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c} R_k(p_k^c)\mathbf{1}\left\{E_1 \cap E_2\right\}\right] - 2\bar{p}\mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \quad \text{(C.2)}
$$

where $\tau_\theta^c := T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}$. For the first term in (C.2), note that $r(p^c) - r(p_k^c) \leq$

$(1+K\bar{p})|p^c - p_k^c|$. So, we can bound

$$
\begin{aligned}
& \mathbb{E}\left[\sum_{k=1}^{\tau_\theta^c} R_k(p_k^c)\mathbf{1}\left\{E_1 \cap E_2\right\}\right] \\
= & \sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[r(p_k^c)\mathbf{1}\left\{E_1 \cap E_2\right\}\right] \\
\geq & \sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[\left\{r(p^c) - (1+K\bar{p})|p^c - p_k^c|\right\}\ \mathbf{1}\left\{E_1 \cap E_2\right\}\right] \\
\geq & \ r(p^c)T_\theta - \bar{p}\log^3 T_\theta - (1+K\bar{p})\sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[|p^c - p_k^c|\, \mathbf{1}\left\{E_1 \cap E_2\right\}\right] \\
\geq & \ r(p^c)T_\theta - \bar{p}\log^3 T_\theta - (1+K\bar{p})\sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2\, \mathbf{1}\{E_1 \cap E_2\}\right]},
\end{aligned}
$$

where the second inequality follows by definition of τ_θ^c and the last inequality follows from Jensen's inequality. As for the second term in (C.2), following the same arguments as in Step 2 in section 4.2, we know that for the sample paths on $E_1 \cap E_2$,

$$
\begin{aligned}
& \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
\leq & \ \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - \lambda(p_t)\right)^+\right] \\
& + \mathbb{E}\left[\left(\sum_{t=1}^{4\sum_{k=1}^{\tau_\theta} N_{k,\theta}} \lambda(p_t) - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
& + \mathbb{E}\left[\left(\sum_{t=4\sum_{k=1}^{\tau_\theta} N_{k,\theta}+1}^{T_\theta} \lambda(p_t) - \left(T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}\right)\frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
\leq & \qquad \leq \sqrt{T_\theta} + \log^3 T_\theta + \sum_{k=1}^{\tau_\theta^c} \mathbb{E}\left[\left(\lambda(p_k^c) - \frac{C}{T}\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] \\
\leq & \ O(\sqrt{\theta}) + \sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(\lambda_k(p_k^c) - \lambda(p^c))^2\, \mathbf{1}\{E_1 \cap E_2\}\right]} \\
\leq & \ O(\sqrt{\theta}) + K\sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2\, \mathbf{1}\{E_1 \cap E_2\}\right]}.
\end{aligned}
$$

Thus, the total penalty for capacity violation satisfies

$$\begin{aligned}
& \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+\right] \\
= & \quad \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{E_1 \cap E_2\}\right] + \mathbb{E}\left[\left(\sum_{t=1}^{T_\theta} D_t(p_t) - C_\theta\right)^+ \mathbf{1}\{\overline{E_1 \cap E_2}\}\right] \\
\leq & \quad O(\sqrt{\theta}) + \ K \sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2 \, \mathbf{1}\{E_1 \cap E_2\}\right]} + 2\bar{p}\ T_\theta P(\overline{E_1 \cap E_2}) \\
= & \quad O(\sqrt{\theta}) + K \sum_{k=1}^{\tau_\theta^c} \sqrt{\mathbb{E}\left[(p_k^c - p^c)^2 \, \mathbf{1}\{E_1 \cap E_2\}\right]}.
\end{aligned}$$

Combining the results above we get

$$J_\theta^{MSA-BDPA} \ \geq \ r(p^c)T_\theta - O(\sqrt{\theta}) - (1 + 3K\bar{p}) \sum_{k=1}^{\tau_\theta^c} \sqrt{E\left[(p_k^c - p^c)^2 \, \mathbf{1}\{E_1 \cap E_2\}\right]}.$$

Applying the established convergence result for Robbins-Monro type of Stochastic Approximation, by Theorem 1 in the electronic companion in Broadie et al. 2011, we know that there exists positive constant C_a^c such that when $a_k^c = C_a^c/k$, we have $E[(p_k^c - p^c)^2 \, \mathbf{1}\{E_1 \cap E_2\}] \leq C_c/k, \quad$ for all $k \geq 1$, where $C_c > 0$ is also a constant. Substitute this back into the previous bound, we have

$$\begin{aligned}
J_\theta^{MSA-BDPA} \ & \geq \ r(p^c)T_\theta - \Theta(\sqrt{\theta}) - (1 + 3K\bar{p}) \sum_{k=1}^{\tau_\theta^c} \sqrt{C_c/k} \\
& \geq \ r(p^c)T_\theta - \Theta(\sqrt{\theta}) - (1 + 3K\bar{p})\sqrt{C_c \tau_\theta^c} = r(p^c)T_\theta - \Theta(\sqrt{\theta}). \quad \square
\end{aligned}$$

C.2 Proof of Key Lemmas in Chapter 4.5

C.2.1 Proof of Lemma 4.5.3

(i) We assume without loss of generality that $p_b > p_a > p^u$. Let $\lambda_a = \lambda(p_a)$, $\lambda_b = \lambda(p_b)$, and we have $\lambda_b < \lambda_a < \lambda^u$ since demand is decreasing in price. Now, by Assumption

A5, we know that (see Boyd and Vandenberghe 2004)

$$\begin{aligned} r(p_a) - r(p_b) &= r(\lambda_a) - r(\lambda_b) \geq \frac{M_l}{2}(\lambda_b - \lambda_a)^2 - r'(\lambda_a)(\lambda_b - \lambda_a) \\ &\geq \frac{M_l L^2}{2}(p_b - p_a)^2 - r'(\lambda_a)(\lambda_b - \lambda_a) \\ &\geq \frac{M_l L^2}{2}(p_b - p_a)^2, \end{aligned}$$

where the first inequality follows from Assumption A5x, second inequality follows from Lemma 4.5.3 part (ii), and the third inequality follows from Assumption A2. Setting $K_u = \frac{M_l L^2}{2}$ completes the proof of part (i).

(ii) Follows directly from Assumption A1.

(iii) Let us denote $\lambda = \lambda(p)$. Notice that $r(\lambda)$ is strictly concave in λ, by Taylor's expansion, there exists $\xi \in [\lambda, \lambda^u]$ (or possibly $[\lambda^u, \lambda)$) such that

$$\begin{aligned} r(p) = r(\lambda) &= r(\lambda^u) + r'(\lambda^u)(\lambda - \lambda^u) + \frac{r''(\xi)}{2}(\lambda - \lambda^u)^2 \\ &\geq r(p^u) - \frac{M_U}{2}(\lambda - \lambda^u)^2 \geq r(\lambda^u) - \frac{M_U K^2}{2}(p - p^u)^2, \end{aligned}$$

where the first and the second inequalities follow by Assumptions A2 and A4, respectively.

As for the second part, we know that

$$\begin{aligned} r(p^c) - r(p) &= r(p^c) - (p^c + p - p^c)[\lambda(p^c) + \lambda(p) - \lambda(p^c)] \\ &= \lambda(p^c)(p^c - p) + p\ (\lambda(p^c) - \lambda(p)) - (p - p^c)(\lambda(p) - \lambda(p^c)) \\ &\leq |p^c - p| + K\bar{p}\ |p^c - p| + K|p^c - p|^2 \\ &\leq (1 + 2K\bar{p})|p^c - p|, \end{aligned}$$

where the first inequality follows from the boundedness of demand and price and Assumption A3. □

C.2.2 Proof of Lemma 4.5.4

We start with τ_θ. Define:

$$t_1 = \left\lceil \frac{1}{4} \log_{3/2}\left(\frac{1}{6}\log T_\theta + 1\right)\right\rceil - 3 \quad \text{and} \quad t_2 = \left\lceil \frac{1}{4} \log_{3/2}\left(\frac{65}{324}\log T_\theta + 1\right)\right\rceil + 1.$$

Note that $t_1 < t_2$ when θ is large and they are both $\Theta(\log\log\theta)$. Moreover, we also have

$$\begin{aligned}
4\cdot\sum_{k=1}^{t_2} N_{k,\theta} &\geq 4\sum_{k=1}^{t_2}\left(\frac{3}{2}\right)^{4k}\log^2 T_\theta = \frac{324}{65}\left[\left(\frac{3}{2}\right)^{4t_2}-1\right]\log^2 T_\theta > \log^3 T_\theta \quad \text{and} \\
4\cdot\sum_{k=1}^{t_1+1} N_{k,\theta} &< \left[4\sum_{k=1}^{t_1+1}\left(\frac{3}{2}\right)^{4k}\log^2 T_\theta\right] + 4t_1 < \left(\frac{3}{2}\right)^{4(t_1+2)}\log^2 T_\theta + 4t_1 \\
&\leq \frac{1}{6}\log^3 T_\theta + \log^2 T_\theta + \Theta(\log\log T_\theta) < \log^3 T_\theta \quad \text{(for all large } \theta).
\end{aligned}$$

Since $\sum_{k=1}^{t} N_{k,\theta}$ is increasing in t, we must have $t_1 < \tau_\theta < t_2$. We conclude that $\tau_\theta = \Theta(\log\log\theta)$ and $(2/3)^{\tau_\theta} = \Theta(\log^{-1/4}\theta)$. We now calculate the order of τ_θ^u. Define:

$$t_1^u = \left\lceil \frac{1}{4}\log_{3/2}\left(\frac{65T_\theta}{648\log^3 T_\theta}+1\right)\right\rceil - 1 \quad \text{and} \quad t_2^u = \left\lceil \frac{1}{4}\log_{3/2}\left(\frac{65T_\theta}{162\log^3 T_\theta}+1\right)\right\rceil.$$

By definition of τ_θ^u and $N_{k,\theta}^u$, for all large enough θ, we have

$$\begin{aligned}
4\sum_{k=1}^{t_1^u} N_{k,\theta}^u &\leq 4\sum_{k=1}^{t_1^u}\left[\left(\frac{3}{2}\right)^{4k}\log^3 T_\theta + 1\right] \leq 4t_1^u + \frac{324}{65}\left[\left(\frac{3}{2}\right)^{4t_1^u}-1\right]\log^3 T_\theta \\
&\leq \frac{1}{2}T_\theta + \Theta\left(\log\left(\frac{T_\theta}{\log^3 T_\theta}\right)\right) \leq T_\theta - \log^3 T_\theta \leq T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta} \quad \text{and} \\
4\sum_{k=1}^{t_2^u} N_{k,\theta}^u &\geq 4\sum_{k=1}^{t_2^u}\left[\left(\frac{3}{2}\right)^{4k}\log^3 T_\theta - 1\right] \geq \frac{324}{65}\log^3 T_\theta\left[\left(\frac{3}{2}\right)^{4t_2^u}-1\right] - 4t_2^u \\
&\geq 2T_\theta - \Theta\left(\log\left(\frac{T_\theta}{\log^3 T_\theta}\right)\right) \geq T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta},
\end{aligned}$$

which implies that $t_1^u \leq \tau_\theta^u \leq t_2^u$. Since t_1^u and t_2^u are both $\Theta(\log\theta)$, we conclude that $\tau_\theta^u = \Theta(\log\theta)$. Moreover, $(2/3)^{4\tau_\theta^u} = \Theta(\theta^{-1}\log^3\theta)$. Finally, we calculate τ_θ^c. Define:

$$t_1^c = \left\lceil \frac{1}{2}\log_{3/2}\left(\frac{5T_\theta}{72\log^2 T_\theta}+1\right)\right\rceil - 1 \quad \text{and} \quad t_2^c = \left\lceil \frac{1}{2}\log_{3/2}\left(\frac{5T_\theta}{18\log^2 T_\theta}+1\right)\right\rceil.$$

By definition of τ_θ^c and $N_{k,\theta}^c$, for all large enough θ, we have

$$\begin{aligned}
4\sum_{k=1}^{t_1^c} N_{k,\theta}^c &\le 4\sum_{k=1}^{t_1^c}\left[\left(\frac{3}{2}\right)^{2k}\log^2 T_\theta + 1\right] \le 4t_1^c + \frac{36}{5}\left[\left(\frac{3}{2}\right)^{2t_1^c} - 1\right]\log^2 T_\theta \\
&\le \frac{1}{2}T_\theta + \Theta\left(\log\left(\frac{T_\theta}{\log^2 T_\theta}\right)\right) \le T_\theta - \log^3 T_\theta \le T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta} \quad \text{and} \\
4\sum_{k=1}^{t_2^c} N_{k,\theta}^c &\ge 4\sum_{k=1}^{t_2^c}\left[\left(\frac{3}{2}\right)^{2k}\log^2 T_\theta - 1\right] \ge \frac{36}{5}\left[\left(\frac{3}{2}\right)^{2t_2^c} - 1\right]\log^2 T_\theta - 4\tau_\theta^c \\
&\ge 2T_\theta - \Theta\left(\log\left(\frac{T_\theta}{\log^2 T_\theta}\right)\right) \ge T_\theta - 4\sum_{k=1}^{\tau_\theta} N_{k,\theta}
\end{aligned}$$

which implies $t_1^c \le \tau_\theta^c \le t_2^c$. Since t_1^c and t_2^c are both $\Theta(\log\theta)$, we conclude that $\tau^c = \Theta(\log\theta)$. Moreover, $(2/3)^{2\tau_\theta^c} = \Theta(\theta^{-1}\log^2\theta)$. This completes the proof. □

C.2.3 Proof of Lemma 4.5.5

By the same arguments as in the proof of Lemma 4.5.1, $P\left(\bar{E}_u | E_1 \cap E_2\right) \le \sum_{k=1}^{\tau_\theta^u}(\tau_\theta^u - k + 1)P(p^u \notin I_{k+1}^u,\ p^u \in I_k^u)$. So, we can bound

$$\begin{aligned}
P\left(\ \overline{E_1 \cap E_2 \cap E_u}\ \right) &\le P\left(\ \overline{E_1 \cap E_2}\ \right) + P(E_1 \cap E_2 \cap \bar{E}_u) \\
&\le P\left(\ \overline{E_1 \cap E_2}\ \right) + \sum_{k=1}^{\tau_\theta^u}(\tau_\theta^u - k + 1)P(p^u \notin I_{k+1}^u,\, p^u \in I_k^u).
\end{aligned}$$

The remaining task then is to bound the term $P(p^u \notin I_{k+1}^u,\ p^u \in I_k^u)$ for $k = 1, ..., \tau_\theta^u$. Define:

$$B_{k,1}^u = \{\hat{r}(p_{k,2}^u) < \hat{r}(p_{k,3}^u),\, p^u < p_{k,2}^u\} \quad \text{and} \quad B_{k,2}^u = \{\hat{r}(p_{k,2}^u) \ge \hat{r}(p_{k,3}^u),\, p^u > p_{k,3}^u\}.$$

Observe that, for all k, we have

$$\begin{aligned}
P(p^u \notin I^u_{k+1},\, p^u \in I^u_k) &\leq P\left(\hat{r}(p^u_{k,2}) < \hat{r}(p^u_{k,3}),\, p^u \notin I^u_{k+1},\, p^u \in I^u_k\right) \\
&\quad + P\left(\hat{r}(p^u_{k,2}) \geq \hat{r}(p^u_{k,3}),\, p^u \notin I^u_{k+1},\, p^u \in I^u_k\right) \\
&= P(\hat{r}(p^u_{k,2}) < \hat{r}(p^u_{k,3}),\, p^u \in [p^u_{k,1}, p^u_{k,2}),\, p^u \in I^u_k) \\
&\quad + P(\hat{r}(p^u_{k,2}) \geq \hat{r}(p^u_{k,3}),\, p^u \in (p^u_{k,3}, p^u_{k,4}],\, p^u \in I^u_k) \\
&\leq P(\hat{r}(p^u_{k,2}) < \hat{r}(p^u_{k,3}),\, p^u < p^u_{k,2},\, p^u \in I^u_k) \\
&\quad + P(\hat{r}(p^u_{k,2}) \geq \hat{r}(p^u_{k,3}),\, p^u > p^u_{k,3},\, p^u \in I^u_k) \\
&\leq P\left(B^u_{k,1}\right) + P\left(B^u_{k,2}\right).
\end{aligned}$$

By Lemma 4.5.3 part (i), we have

$$r(p^u_{k,2}) - r(p^u_{k,3}) \geq K_u (p^u_{k,2} - p^u_{k,3})^2 = K_u \frac{|I^u_1|^2}{9} \left(\frac{2}{3}\right)^{2(k-1)} = \frac{1}{4} K_u |I^u_1|^2 \left(\frac{2}{3}\right)^{2k}.$$

Arguing as in the proof of Lemma 4.5.1, if $|\hat{r}(p_{k,l}) - r(p_{k,l})| < \frac{1}{8} K_u |I^u_1|^2 (\frac{2}{3})^{2k}$ for all k and $l \in \{2, 3\}$, then we can correctly predict whether $r(p^u_{k,2}) \geq r(p^u_{k,3})$ or $r(p^u_{k,2}) < r(p^u_{k,3})$. (This guarantees that the deleted segment does not contain p^u.) So, applying Hoeffding's inequality together with the facts that $\hat{r}(p_{k,l}) < \bar{p}$ and $|I^u_1| = |I| \left(\frac{2}{3}\right)^{\tau_\theta} = \Theta(\log^{-1/4} \theta)$ (see Lemma 4.5.4), we can bound $P(B^u_{k,l})$ as follows:

$$\begin{aligned}
P(B^u_{k,l}) &\leq P\left(|\hat{r}(p_{k,j}) - r(p_{k,j})| \geq \frac{1}{8} K_u |I^u_1|^2 \left(\frac{2}{3}\right)^{2k} \text{ for some } j \in \{2, 3\}\right) \\
&\leq \sum_{j=2}^{3} P\left(|\hat{r}(p_{k,j}) - r(p_{k,j})| \geq \frac{1}{8} K_u |I^u_1|^2 \left(\frac{2}{3}\right)^{2k}\right) \\
&\leq 4 \cdot \exp\left(-2 \frac{N^u_{k,\theta}\ [\frac{1}{8} K_u\, |I^u_1|^2 \left(\frac{2}{3}\right)^{2k}]^2}{\bar{p}^2}\right) \\
&\leq 4 \cdot \exp(-\log \theta) = \frac{4}{\theta}, \quad \text{for } l = 1, 2 \text{ and sufficiently large } \theta.
\end{aligned}$$

Since it can be shown that $\tau^u_\theta = \Theta(\log \theta)$, put the above bounds together with our earlier bound for $P\left(\,\overline{E_1 \cap E_2 \cap E_u}\,\right)$ and $P\left(\,\overline{E_1 \cap E_2}\,\right)$ (from Lemma 4.5.2), we conclude that

$$\begin{aligned} P\left(\overline{E_1 \cap E_2 \cap E_u} \right) &\le P\left(\overline{E_1 \cap E_2} \right) + \sum_{k=1}^{\tau_\theta^u} (\tau_\theta^u - k + 1) \sum_{l=1}^{2} P(B_{k,l}^u) \\ &= \Theta\left(\frac{\log^2 \theta}{\theta} \right). \quad \square \end{aligned}$$

C.2.4 Proof of Lemma 4.5.6

Define two events:

$$\begin{aligned} B_{k,1}^c &= \{\hat{\lambda}(p_{k,2}^c) > C/T + \Delta_{k,\theta}^c,\ p^c < p_{k,2}^c\} \text{ and} \\ B_{k,2}^c &= \{\hat{\lambda}(p_{k,2}^c) \le C/T + \Delta_{k,\theta}^c,\ p^c > p_{k,3}^c\}. \end{aligned}$$

By similar arguments as in the proof of Lemma 4.5.5, we know that $P(\bar{E}_c \mid E_1 \cap E_2) \le \sum_{k=1}^{\tau_\theta^c} (\tau_\theta^c - k + 1) \left[\sum_{l=1}^{2} P(B_{k,l}^c)\right]$. For event $B_{k,1}^c$, note that $p^c < p_{k,2}^c$ implies $\lambda(p_{k,2}^c) < C/T$. So,

$$\begin{aligned} P(B_{k,1}^c) &\le P\left(\hat{\lambda}(p_{k,2}^c) > C/T + \Delta_{k,\theta}^c\ ,\ \lambda(p_{k,2}^c) < C/T \right) \\ &\le P\left(\hat{\lambda}(p_{k,3}^c) - \lambda(p_{k,3}^c) > \Delta_{k,\theta}^c \right). \end{aligned}$$

Since $N_{k,\theta}^c = \Theta(\left(\frac{3}{2}\right)^{2k} \log^2 \theta)$ and $\Delta_{k,\theta}^c = \Theta(\left(\frac{2}{3}\right)^k \log^{-3/8} \theta)$, by Hoeffding's inequality,

$$\begin{aligned} P(B_{k,1}^c) &\le P\left(\hat{\lambda}(p_{k,2}^c) - \lambda(p_{k,2}^c) > \Delta_{k,\theta}^c \right) \\ &\le \exp\left(-2N_{k,\theta}^c (\Delta_{k,\theta}^c)^2\right) \le \exp(-\log\theta) = \frac{1}{\theta}. \end{aligned}$$

As for event $B_{k,2}^c$, note that $p^c > p_{k,3}^c$ implies $\lambda(p_{k,3}^c) > C/T$. By Lemma 4.5.3 part (ii), $\lambda(p_{k,2}^c) - \lambda(p_{k,3}^c) \ge L \cdot |p_{k,2}^c - p_{k,3}^c| = L\frac{|I_1^c|}{3}(\frac{2}{3})^{k-1}$. So, for the sample path in $B_{k,2}^c$, we have:

$$\begin{aligned} \lambda(p_{k,2}^c) - \hat{\lambda}(p_{k,2}^c) &\ge \lambda(p_{k,3}^c) + L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} + \Delta_{k,\theta}^c\right) \\ &> \frac{C}{T} + L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1} - \left(\frac{C}{T} + \Delta_{k,\theta}^c\right) \\ &= L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1} - \Delta_{k,\theta}^c > \frac{1}{2}L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1} \end{aligned}$$

where the last inequality follows because $|I_c^1| = |I| \left(\frac{2}{3}\right)^{\tau_\theta} = \Theta(\log^{-1/4}\theta)$ and so $\Delta_{k,\theta}^c <$

$\frac{1}{2}L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1}$ for all large θ. By similar argument as above,

$$\begin{aligned} P(B_{k,2}^c) &\le P\left(\lambda(p_{k,2}^c) - \hat{\lambda}(p_{k,2}^c) > \frac{1}{2}L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1}\right) \\ &\le \exp\left(-2\frac{N_{k,\theta}^c\ [\frac{1}{2}L\frac{|I_1^c|}{3}\left(\frac{2}{3}\right)^{k-1}]^2}{\bar{p}^2}\right) \le \exp(-\log\theta) = \frac{1}{\theta}. \end{aligned}$$

Since it can be shown that $\tau_\theta^c = \Theta(\log\theta)$, put the above bounds together with our earlier bound for $P\left(\ \overline{E_1 \cap E_2 \cap E_c}\ \right)$ and $P\left(\ \overline{E_1 \cap E_2}\ \right)$ (from Lemma 4.5.2), we conclude that

$$\begin{aligned} P\left(\ \overline{E_1 \cap E_2 \cap E_c}\ \right) &\le P\left(\ \overline{E_1 \cap E_2}\ \right) + \sum_{k=1}^{\tau_\theta^c}(\tau_\theta^c - k + 1)\sum_{l=1}^{2} P(B_{k,l}^c) \\ &= \Theta\left(\frac{\log^2\theta}{\theta}\right). \quad \square \end{aligned}$$

BIBLIOGRAPHY

J Acimovic and SC Graves. Making better fulfillment decisions on the fly in an online retail environment. *M&SOM*, 17(1):34–51, 2014.

J Acimovic and SC Graves. Mitigating spillover in online retailing via replenishment. *M & SOM*, 19(3):419–436, 2017.

P Afèche and B Ata. Bayesian dynamic pricing in queueing systems with unknown delay cost characteristics. *M&SOM*, 15(2):292–304, 2013.

Philipp Afèche, Zhe Liu, and Costis Maglaras. Ride-hailing networks with strategic drivers: The impact of platform control capabilities on performance. Working paper, 2018.

NAH Agatz, M Fleischmann, and JA van Nunen. E-fulfillment and multi-channel distribution–a review. *Eur. J. Oper. Res.*, 187(2):339–356, 2008.

R Agrawal. The continuum-armed bandit problem. *SIAM J. Control Optim.*, 33(6): 1926–1951, 1995.

HS Ahn, S Jasin, P Kaminsky, and Y Wang. Certainty equivalent planning for multi-product batch differentiation: Analysis and bounds. *Available at SSRN 2704595*, 2015.

M Al-Roomi, S Al-Ebrahim, S Buqrais, and I Ahmad. Cloud computing pricing models: a survey. *Int. J. Grid Dist. Comput.*, 6(5):93–106, 2013.

R Altman, T Jiménez, and G Koole. On optimal call admission control in resource-sharing system. *IEEE T. Commun.*, 49(9):1659–1668, 2001.

F Alzhouri and A Agarwal. Dynamic pricing scheme: Towards cloud revenue maximization. In *2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom)*, pages 168–173. IEEE, 2015.

Amazon.com. Amazon annual report. `http://phx.corporate-ir.net/phoenix.zhtml?c=97664&p=irol-newsArticle&ID=2133281`, 2015.

Ali Aouad and Danny Segev. Display optimization for vertically differentiated locations under multinomial logit choice preferences. Working paper, 2015.

VF Araman and R Caldentey. Dynamic pricing for nonperishable products with demand learning. *Oper. Res.*, 57(5):1169–1188, 2009.

S Arshad, S Ullah, SA Khan, MD Awan, and MSH Khayal. A survey of cloud computing variable pricing models. In *2015 International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE)*, pages 27–32. IEEE, 2015.

K Asdemir, VS Jacob, and E Krishnan. Dynamic pricing of multiple home delivery options. *Eur. J. Oper. Res.*, 196(1):246–257, 2009.

B Ata and TL Olsen. Near-optimal dynamic lead-time quotation and scheduling under convex-concave customer delay costs. *Oper. Res.*, 57(3):753–768, 2009.

B Ata and TL Olsen. Congestion-based leadtime quotation and pricing for revenue maximization with heterogeneous customers. *Queueing Syst.*, 73(1):35–78, 2013.

B Ata and S Shneorson. Dynamic control of an m/m/1 service system with adjustable arrival and service rates. *Manage. Sci.*, 52(11):1778–1791, 2006.

R Atar and MI Reiman. Asymptotically optimal dynamic pricing for network revenue management. *Stochastic Systems*, 2(2):232–276, 2012.

P. Auer, R. Ortner, and C. Szepesvári. Improved rates for the stochastic continuum-armed bandit problem. In *Learning Theory*, pages 454–468. Springer, 2007.

G Aydin and S Ziya. Pricing promotional products under upselling. *M&SOM*, 10(3):360–376, 2008.

A Badanidiyuru, R Kleinberg, and A Slivkins. Bandits with knapsacks. In *FOCS*, 2013.

GY Ban and NB Keskin. Personalized dynamic pricing with machine learning. Working paper, 2017.

S Banerjee, C Riquelme, and R Johari. Pricing in ride-share platforms: A queueing-theoretic approach. Working paper, 2015.

R Becerril-Arreola, M Leng, and M Parlar. Online retailers promotional pricing, free-shipping threshold, and inventory decisions: A simulation-based analysis. *Eur. J. Oper. Res.*, 230(2):272–283, 2013.

A Ben-Tal, A Goryashko, E Guslitzer, and A Nemirovski. Adjustable robust solutions of uncertain linear programs. *Math. Program.*, 99(2):351–376, 2004.

G Bensinger. Startups scramble to define employee. `https://www.wsj.com/articles/startups-scramble-to-define-employee-1438228860`, 2015.

D Bertsimas, DA Iancu, and PA Parrilo. Optimality of affine policies in multistage robust optimization. *Math. of Oper. Res.*, 35(2):363–394, 2010.

O Besbes and C Maglaras. Revenue optimization for a make-to-order queue in an uncertain market environment. *Oper. Res.*, 57(6):1438–1450, 2009.

O Besbes and A Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Oper. Res.*, 57(6):1407–1420, 2009.

O Besbes and A Zeevi. Blind network revenue management. *Oper. Res.*, 60(6):1537–1550, 2012.

O Besbes and A Zeevi. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Manage. Sci.*, 61(4):723–739, 2015.

O Besbes, Y Gur, and A Zeevi. Non-stationary stochastic optimization. *Oper. Res.*, 63(5):1227–1244, 2015.

K Bimpikis, O Candogan, and S Daniela. Spatial pricing in ride-sharing networks. Working paper, 2016.

G Bitran and R Caldentey. An overview of pricing models for revenue management. *M&SOM*, 5(3):203–229, 2003.

C Borgs, O Candogan, J Chayes, I Lobel, and H Nazerzadeh. Optimal multiperiod pricing with service guarantees. *Manage. Sci.*, 60(7):1792–1811, 2014.

EA Boyd and IC Bilegan. Revenue management and e-commerce. *Manage. Sci.*, 49 (10):1363–1386, 2003.

S Boyd and L Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

M Broadie, D Cicek, and A Zeevi. General bounds and finite-time improvement for the kiefer-wolfowitz stochastic approximation algorithm. *Oper. Res.*, 59(5):1211–1224, 2011.

J Broder and P Rusmevichientong. Dynamic pricing under a general parametric choice model. *Oper. Res.*, 60(4):965–980, 2012.

L Brown, N Gans, A Mandelbaum, A Sakov, H Shen, S Zeltyn, and L Zhao. Statistical analysis of a telephone call center: A queueing-science perspective. *J. Ame. Stat. Assoc.*, 100(469):36–50, 2005.

Bloomberg Businessweek. Nick kokonas is selling tickets to dinner. `http://goo.gl/0St8EU`, 2015.

GC Calafiore. An affine control method for optimal dynamic asset allocation with transaction costs. *SIAM J. Control Optim.*, 48(4):2254–2274, 2009.

Camelcamelcamel.com. Amazon price history for lodge l8sgp3 pre-seasoned cast-iron square grill pan. `https://camelcamelcamel.com/Lodge-L8SGP3-Square-Pre-Seasoned-10-5-inch/product/B0000CF66W?context=browse`, 2016. Accessed: 2016-03-16.

AM Campbell and M Savelsbergh. Incentive schemes for attended home delivery services. *Transport. Sci.*, 40(3):327–341, 2006.

S Çelik and C Maglaras. Dynamic pricing and lead-time quotation for a multiclass make-to-order queue. *Manage. Sci.*, 54(6):1132–1146, 2008.

N Cesa-Bianchi and G Lugosi. *Prediction, Learning, and Games*, volume 1. Cambridge University Press Cambridge, 2006.

V Chamlee. Why surge pricing could hit food delivery apps next. `http://goo.gl/qbv80S`, 2016.

H Chen and MZ Frank. State dependent pricing with a queue. *IIE Trans.*, 33(10):847–860, 2001.

J Chen. Pricing strategies in the electronic marketplace. In *Handbook of Strategic e-Business Management*, pages 489–522. Springer, 2014.

Q Chen, S Jasin, and I Duenyas. Real-time dynamic pricing with minimal and flexible price adjustment. *Manage. Sci.*, 62(8):2437–2455, 2015.

Q Chen, S Jasin, and I Duenyas. Analysis of self-adjusting controls with unknown demand parameters.. Working paper, 2017a.

Q Chen, S Jasin, and I Duenyas. A nonparametric self-adjusting control for joint learning and optimization of multi-product pricing with finite resource capacity. Technical report, 2018. Accepted at Math. Oper. Res..

Y Chen and VF Farias. Simple policies for dynamic pricing with imperfect forecasts. *Oper. Res.*, 61(3):612–624, 2013.

Y Chen and VF Farias. Robust dynamic pricing with strategic customers. Accepted at Math. Oper. Res.., 2018.

Y Chen and C Shi. Near-optimal pricing policy for service systems with reusable resources and forward-looking customers. Working paper, 2018.

Y Chen, VF Farias, and NK Trichakis. On the efficacy of static prices for revenue management in the face of strategic customers. Working paper, 2017b.

Y Chen, R Levi, and C Shi. Revenue management of reusable resources with advanced reservations. *Prod. Oper. Manag.*, 26(5):836–859, 2017c.

HK Cheng, Z Li, and A Naranjo. Research notecloud computing spot pricing dynamics: Latency and limits to arbitrage. *Inform. Syst. Res.*, 27(1):145–165, 2016.

Chicago Consulting. The 10 best warehouse network 2013. `http://www.chicago-consulting.com/ten-best-warehouse-networks-consultant-usa/`, 2013.

DF Ciocan and V Farias. Model predictive control for dynamic resource allocation. *Math. of Oper. Res.*, 37(3):501–525, 2012.

CNN. Amazon 'prime day' shattered global sales records. http://money.cnn.com/2015/07/15/news/amazon-walmart-sales/, 2015.

M Cohen, I Lobel, and R Paes Leme. Feature-based dynamic pricing. Working paper, 2016.

MC Cohen, NHZ Leung, K Panchamgam, G Perakis, and A Smith. The impact of linear optimization on promotion planning. *Oper. Res.*, page Forthcoming, 2017a.

MC Cohen, R Lobel, and G Perakis. Dynamic pricing though data sampling. Working paper, 2017b.

EW Cope. Regret and convergence bounds for a class of continuum-armed bandit problems. *IEEE T. Automat. Contr.*, 54(6):1243–1253, 2009.

CA Courcoubetis, A Dimakis, and MI Reiman. Providing bandwidth guarantees over a best-effort network: call-admission and pricing. In *Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, volume 1, pages 459–467. IEEE, 2001.

GB Dantzig and MN Thapa. *Linear programming 1: Introduction*. Springer, 2006.

AV den Boer. Dynamic pricing with multiple products and partially specified demand distribution. *Math. Oper. Res.*, 39(3):863 – 888, 2014.

AV den Boer. Dynamic pricing and learning: historical origins, current research, and new directions. *Surv. Oper. Res. Manage. Sci.*, 20(1):1–18, 2015.

AV den Boer and B. Zwart. Simultaneously learning and optimizing using controlled variance pricing. *Manage. Sci.*, 60(3):770–783, 2013.

AV den Boer and B Zwart. Dynamic pricing and learning with finite inventories. *Oper. Res.*, 63(4):965–978, 2015.

EM Dinlersoz and H Li. The shipping strategies of internet retailers: Evidence from internet book retailing. *Quant. Marketing Econom.*, 4(4):407–438, 2006.

MK Doğru, MI Reiman, and Q Wang. A stochastic programming based inventory policy for assemble-to-order systems with application to the W model. *Oper. Res.*, 58(4-part-1):849–864, 2010.

The Economists. The mother of invention: Network operators in the poor world are cutting costs and increasing access in innovative ways. http://goo.gl/K7b8xX, 2009.

SS Eren and C Maglaras. Monopoly pricing with limited demand information. *J. Revenue & Pricing Manage.*, 9(1):23–48, 2010.

VF Farias and B van Roy. Dynamic pricing with a prior on market response. *Oper. Res.*, 58(1):16–29, 2010.

KJ Ferreira, BHA Lee, and D Simchi-Levi. Analytics for an online retailer: Demand forecasting and price optimization. *M&SOM*, 18(1):69–88, 2015a.

KJ Ferreira, D Simchi-Levi, and H Wang. Online network revenue management using thompson sampling. *Available at SSRN 2588730*, 2015b.

M Fisher, S Gallino, and J Li. Competition-based dynamic pricing in online retailing: A methodology validated with field experiments. *Available at SSRN 2547793*, 2015.

G Gallego and GJ van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Manage. Sci.*, 40(8):999–1020, 1994.

G Gallego and GJ van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Oper. Res.*, 45(1):24–41, 1997.

G Gallego, A Li, VA Truong, and X Wang. Approximation algorithms for product framing and pricing. Working paper, 2016.

N Gans and S Savin. Pricing and capacity rationing for rentals with uncertain durations. *Manage. Sci.*, 53(3):390–407, 2007.

DA Goldberg, DA Katz-Rogozhnikov, Y Lu, M Sharma, and MS Squillante. Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times. *Math. Oper. Res.*, 41(3):898–913, 2016.

N Golrezaei, H Nazerzadeh, and P Rusmevichientong. Real-time optimization of personalized assortments. *Manage. Sci.*, 60(6):1532–1551, 2014.

GreenBiz. Zipcar founder robin chase: Whats next for urban mobility? `https://goo.gl/HwFx1Q`, 2014.

M Gümüş, S Li, W Oh, and S Ray. Shipping fees or shipping free? A tale of two price partitioning strategies in online retailing. *Prod. Oper. Manag.*, 22(4):758–776, 2013.

RC Hampshire, WA Massey, and Q Wang. Dynamic pricing to control loss systems with quality of service targets. *Probab. Eng. Inform. Sc.*, 23(02):357–383, 2009.

JM Harrison, NB Keskin, and A Zeevi. Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Manage. Sci.*, 58(3):570–586, 2012.

P Harsha, S Subramanian, and J Uichanco. Omni-channel revenue management through integrated pricing and fulfillment planning. Available at http://bit.ly/2avJTca, 2016.

W Hoeffding. Probability inequalities for sums of bounded random variables. *J. Am. Stat Assoc.*, 58(301):13–30, 1963.

D Howland. Online operating costs crushing brick-and-mortar retailers. `http://bit.ly/23pFWoF`, 2016.

Z Hu, JH Kim, J Wang, and J Byrne. Review of dynamic pricing programs in the us and europe: Status quo and policy recommendations. *Renew. Sust. Energ. Rev.*, 42: 743–751, 2015.

A Jain, K Moinzadeh, and A Dumrongsiri. Priority allocation in a rental model with decreasing demand. *M&SOM*, 17(2):236–248, 2015.

S Jasin. Reoptimization and self-adjusting price control for network revenue management. *Oper. Res.*, 62(5):1168–1178, 2014.

S Jasin and S Kumar. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Math. of Oper. Res.*, 37(2):313–345, 2012.

S Jasin and S Kumar. Analysis of deterministic lp-based booking limit and bid price controls for revenue management. *Oper. Res.*, 61(6):1312–1320, 2013.

S Jasin and A Sinha. An LP-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Oper. Res.*, 63(6):1336–1351, 2015.

A Javanmard and H Nazerzadeh. Dynamic pricing in high-dimensions. Working paper, 2016.

M Kaspi, T Raviv, M Tzur, and H Galili. Regulating vehicle sharing systems through parking reservation policies: Analysis and performance bounds. *Eur. J. Oper. Res.*, 251(3):969–987, 2016.

FP Kelly. Effective bandwidths at multi-class queues. *Queueing Syst.*, 9(1-2):5–15, 1991.

N B Keskin and A Zeevi. Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Oper. Res.*, 62(5):1142–1167, 2014.

N B Keskin and A Zeevi. Chasing demand: Learning and earning in a changing environment. *Math. Oper. Res.*, 42(2):277–307, 2016a.

NB Keskin and A Zeevi. Chasing demand: Learning and earning in a changing environment. *Math. Oper. Res.*, 42(2):277–307, 2016b.

J Kiefer and J Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, 23(3):462–466, 1952.

J Kim and RS Randhawa. The value of dynamic pricing in large queueing systems. *Oper. Res.*, 66(2):409–425, 2018.

R Klein, M Neugebauer, D Ratkovitch, and C Steinhardt. Differentiated time slot pricing under routing considerations in attended home delivery. Draft available at SSRN 2674061, 2015.

RD Kleinberg. Nearly tight bounds for the continuum-armed bandit problem. In *NIPS*, 2004.

HJ Kushner and G Yin. *Stochastic Approximation and Recursive Algorithms and Applications*, volume 35. Springer, 2003.

S Lanning, WA Massey, B Rider, and W Qiong. Optimal pricing in queueing systems with quality of service constraints. *Teletraf. Sci. Eng.*, pages 747–756, 1999.

Y Lei, S Jasin, and A Sinha. Dynamic joint pricing and order fulfillment for e-commerce retailers. Accepted at M&SOM, 2017.

M Leng and R Becerril-Arreola. Joint pricing and contingent free-shipping decisions in B2C transactions. *Prod. Oper. Manag.*, 19(4):390–405, 2010.

R Levi and A Radovanovic. Provably near-optimal lp-based policies for revenue management in systems with reusable resources. *Oper. Res.*, 58(2):503–507, 2010.

AEB Lim and JG Shanthikumar. Relative entropy, exponential utility, and robust dynamic pricing. *Oper. Res.*, 55(2):198–214, 2007.

Y Liu and WL Cooper. Optimal dynamic pricing with patient customers. *Oper. Res.*, 63(6):1307–1319, 2015.

DW Low. Optimal dynamic pricing policies for an M/M/s queue. *Oper. Res.*, 22(3): 545–561, 1974.

Hongyao Ma, Fei Fang, and David C Parkes. Spatio-temporal pricing for ridesharing platforms. Working paper, 2018.

C Maglaras. Revenue management for a multiclass single-server queue via a fluid model analysis. *Oper. Res.*, 54(5):914–932, 2006.

C Maglaras and J Meissner. Dynamic pricing strategies for multiproduct revenue management problems. *M&SOM*, 8(2):136–148, 2006.

C Maglaras and A Zeevi. Pricing and design of differentiated services: Approximate analysis and structural insights. *Oper. Res.*, 53(2):242–262, 2005.

BL Miller. A queueing reward system with several customer classes. *Manage. Sci.*, 16 (3):234–245, 1969.

CC Moallemi and M Saglam. Dynamic portfolio choice with linear rebalancing rules. *Available at SSRN 2011605*, 2012.

S Netessine, S Savin, and W Xiao. Revenue management through dynamic cross selling in e-commerce retailing. *Oper. Res.*, 54(5):893–913, 2006.

ED O'Mahony. *Smarter Tools For (Citi) Bike Sharing*. PhD thesis, Cornell University, 2015.

EL Örmeci, A Burnetas, and J van der Wal. Admission policies for a two class loss system. *Stoch. Models*, 17(4):513–540, 2001.

Ö Özer and R Phillips. *The Oxford handbook of pricing management*. Oxford University Press, 2012.

F Papier and UW Thonemann. Capacity rationing in stochastic rental systems with advance demand information. *Oper. Res.*, 58(2):274–288, 2010.

IC Paschalidis and JN Tsitsiklis. Congestion-dependent pricing of network services. *IEEE ACM Trans. Network*, 8(2):171–184, 2000.

G Perakis and G Roels. Robust controls for network revenue management. *M&SOM*, 12(1):56–76, 2010.

J Pfrommer, J Warrington, G Schildbach, and M Morari. Dynamic vehicle redistribution and online price incentives in shared mobility systems. *IEEE Trans. Intelli. Transp.*, 15(4):1567–1578, 2014.

EL Plambeck and AR Ward. Optimal control of a high-volume assemble-to-order system. *Math. of Oper. Res.*, 31(3):453–477, 2006.

Project44. Project44 introduces dynamic pricing technology to freight industry. `http://goo.gl/Ayk2mG`, 2015.

T Raviv and O Kolka. Optimal inventory management of a bike-sharing station. *IIE Trans.*, 45(10):1077–1093, 2013.

B Reid. Avoiding everyday antitrust law violations. `http://huff.to/2a9Ingj`, 2014.

MI Reiman and Q Wang. An asymptotically optimal policy for a quantity-based network revenue management problem. *Math. of Oper. Res.*, 33(2):257–282, 2008.

MI Reiman and Q Wang. Asymptotically optimal inventory control for assemble-to-order systems with identical lead times. *Oper. Res.*, 63(3):716–732, 2015.

D Rigby. Online shopping isnt as profitable as you think. `https://hbr.org/2014/08/online-shopping-isnt-as-profitable-as-you-think`, 2014.

H Robbins and S Monro. A stochastic approximation method. *Ann. Math. Statist.*, pages 400–407, 1951.

P Rusmevichientong, B van Roy, and PW Glynn. A nonparametric approach to multiproduct pricing. *Oper. Res.*, 54(1):82–98, 2006.

SV Savin, MA Cohen, N Gans, and Z Katalan. Capacity management in rental businesses with two customer bases. *Oper. Res.*, 53(4):617–631, 2005.

A Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, 1998.

J Schuijbroek, RC Hampshire, and WJ Van Hoeve. Inventory rebalancing and vehicle routing in bike sharing systems. *Eur. J. Oper. Res.*, 257(3):992–1004, 2017.

N Secomandi. An analysis of the control-algorithm re-solving issue in inventory and revenue management. *M&SOM*, 10(3):468–483, 2008.

N Secomandi and F Margot. Reoptimization approaches for the vehicle-routing problem with stochastic demands. *Oper. Res.*, 57(1):214–230, 2009.

A Shpanya. Why dynamic pricing is a must for ecommerce retailers. `https://econsultancy.com/blog/65327-why-dynamic-pricing-is-a-must-for-ecommerce-retailers/`, 2014.

J Shu, MC Chou, Q Liu, CP Teo, and IL Wang. Models for effective deployment and redistribution of bicycles within public bicycle-sharing systems. *Oper. Res.*, 61(6): 1346–1359, 2013.

R Sides and HK Hogan. Deloitte's 2015 holiday survey: Embracing retail disruption. `https://www2.deloitte.com/tr/en/pages/consumer-business/articles/holiday-retail-sales-consumer-survey.html`, 2015.

D Simchi-Levi, SD Wu, and ZJM Shen. *Handbook of quantitative supply chain analysis: Modeling in the e-business era*, volume 74. Springer, 2004.

KT Talluri and GJ van Ryzin. *The theory and practice of revenue management*, volume 68. Springer, 2006.

Tompkins Supply Chain Consortium. Supply chain metrics: Data for you to compare against. `https://www.supplychainconsortium.com/Report/GetReport/Supply-Chain-Metrics`, 2012.

UPS. 2016 UPS rate and service guide. `https://www.ups.com/media/en/daily_rates.pdf`, 2016.

U.S. Census Bureau. Annual estimates of the resident population: April 1, 2010 to july 1, 2014 - United States - metropolitan and micropolitan statistical area. `https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk`, 2014a.

U.S. Census Bureau. Income and poverty in the United States: 2014. `https://www.census.gov/content/dam/Census/library/publications/2015/demo/p60-252.pdf`, 2014b.

U.S. Census Bureau. Latest quarterly e-commerce report. `https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf`, 2016.

R Waeber, PI Frazier, and SG Henderson. Bisection search with noisy responses. *SIAM J. Control Optim.*, 51(3):2261–2279, 2013.

Z Wang, S Deng, and Y Ye. Close the gaps: A learning-while-doing algorithm for single-product revenue management problems. *Oper. Res.*, 62(2):318–331, 2014.

A Waserhole. *Vehicle sharing systems pricing optimization*. PhD thesis, Université de Grenoble, 2014.

L Wei, S Jasin, and L Xin. Asymptotic optimality of order-up-to control for stochastic inventory systems with sequential probabilistic service level constraints. Working paper, 2018.

D Williams. *Probability with martingales*. Cambridge university press, 1991.

L Xin and DA Goldberg. Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. *Oper. Res.*, 64(6):1556–1565, 2016.

H Xu and B Li. Dynamic cloud pricing for revenue maximization. *IEEE T. Cloud Comp.*, 1(2):158–171, 2013.

PJ Xu, R Allgor, and SC Graves. Benefits of reevaluating real-time order fulfillment decisions. *M&SOM*, 11(2):340–355, 2009.

S Yoon and ME Lewis. Optimal pricing and admission control in a queueing system with periodically varying parameters. *Queueing Syst.*, 47(3):177–199, 2004.

L Zhan and AE Lloyd. Customers asymmetrical responses to variable pricing. *J. Revenue Pricing Management*, 13(3):183–198, 2014.

Y Zhou. *Pricing and Matching in the Sharing Economy*. PhD thesis, 2017.