MULTI-DIMENSIONAL PROBLEMS IN SINGLE-RESOURCE REVENUE MANAGEMENT

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Chao Ding

August 2012

MULTI-DIMENSIONAL PROBLEMS IN SINGLE-RESOURCE REVENUE MANAGEMENT

Chao Ding, Ph.D.
Cornell University 2012

This dissertation addresses three multi-dimensional problems in single-resource revenue management. Many problems in revenue management involve managing a single resource but their dynamic programming formulation still requires using a multi-dimensional vector as the state variable. For instance, single-flight-leg revenue management is an extensively studied problem where airlines use the same seat resource to accommodate different classes of customers. In these problems, due to the multi-dimensional state space, the exact formulation is either intractable, or the resulting optimal policy is inconvenient to implement operationally. We develop approximations for the exact formulations, which generate tractable and operationally attractive policies.

The first problem we study is a strategic decision problem about whether or not to discontinue a product sold under warranty, whose failure probabilities are unknown initially and are learnt as sales take place and failure information is accumulated. Since there are multiple types of failures that the product can fail from, we formulate the problem as a multi-dimensional optimal stopping problem with Bayesian learning. Two approximations based on dynamic programming decomposition and deterministic approximation are developed, and insights about the value of learning are extracted from asymptotic analysis.

Next we study a dual-channel pricing and capacity allocation problem for hotel revenue management. While one channel is the spot market in which we

can adopt dynamic pricing, the other is a conference market with a fixed price offered for conference participants. Remaining rooms in the conference market will be released to the spot market if not booked by a deadline. Tactical decisions on number of rooms to reserve and fixed price to offer for the conference market need to be made at the beginning of the selling horizon. For the operational pricing problem in the spot market, because the two markets will join together in a future time, we need a two-dimensional dynamic program which tracks the remaining capacities in both markets to make optimal pricing decisions in the spot market. We develop a single-dimensional approximation to the exact two-dimensional formulation, which generates a robust and operationally attractive policy. For the tactical problem of finding the optimal capacity allocation between spot and conference markets and choosing the fixed price to charge in the conference market, we construct an asymptotically optimal policy through a deterministic formulation.

Finally, we consider a revenue management problem where we sell a product to multiple markets with heterogeneous price sensitivities. We can allocate the capacity to different markets and charge different prices in different markets (separate pricing), in which case we gain pricing flexibility, or we can merge all markets together and serve them with a common price (joint pricing), in which case we obtain capacity flexibility. We study the tradeoff between pricing and capacity flexibilities and establish conditions under which one is more important than the other. For a hybrid model where separate pricing is adopted early in the selling horizon and joint pricing is used towards the end, we develop a single-dimensional approximation which gives rise to a policy with remarkably good performance especially for problem instances with tight capacity, large number of markets, or drastically different price sensitivities.

BIOGRAPHICAL SKETCH

Chao Ding was born on October 22, 1984 in Luoyang, Henan Province, China. He did his undergraduate studies in the Department of Industrial Engineering at Tsinghua University. He earned a Bachelor of Engineering degree in June of 2007.

In August of 2007, he joined the School of Operations Research and Information Engineering at Cornell University. He received a M.S. degree in 2010. His research interests are revenue management and stochastic dynamic optimization. He has been working with Professor Paat Rusmevichientong and Professor Huseyin Topaloglu on revenue management problems since 2009.

Upon completion of his Ph.D., he will join Google as a Quantitative Analyst.

To my parents Mingzheng Ding, Minai Sun,

and my brother Jifeng Ding

ACKNOWLEDGEMENTS

First of all, I want to heartily thank my advisors Professor Paat Rusmevichientong and Professor Huseyin Topaloglu for their invaluable guidance, encouragement and patience during my doctoral studies. Without their support, it would be impossible to complete this work. Their enthusiasm on research sets an example for me. Meeting with them has always been delightful and inspiring. Their encouragements and positive attitudes help me through depressed times. Paat is very kind to keep our online research meetings after he moved to California. Huseyin is always accessible to me for academic or personal issues. I am very thankful to have two great advisors during my time at Cornell.

I would also like to thank Professor David Shmoys and Professor David Williamson for serving on my committee and for their invaluable inputs to my research. I have greatly benefited from my interactions with all the faculty members. In particular, I would like to thank Professor Peter Jackson for being my first year advisor. His kindness and passion for teaching set a role model. My sincere gratitude also goes to Professor John A. Muckstadt for his fun classes and insightful career advice.

In addition, I want to thank Kathryn King, Selene Cammer and Eric Johnson for their kindness and numerous helps. Special thanks go to all my Ph.D. fellows and friends, who made my stay in Ithaca much more enjoyable. To name a few, Fan Zhu, Jie Chen, Tuohua Wu, Maurice Cheung, Yuemeng Sun, Yi Shen, Yinan Huang, Jiawei Qian, Sophia Liu, Collin Chan, Chao Fang, Jialie Chen.

Last but not least, nothing would have been possible without the unconditional love and support from my parents and my brother. This thesis is dedicated to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Revenue management deals with strategic and tactical demand-management decisions in order to increase revenue when we sell scarce resources to customers in an uncertain environment. Starting from 1950's, it has found wide applications in various areas such as passenger airlines, hotel chains, car rental companies, sports and entertainment industry and retail industry (see Talluri and Van Ryzin, 2005).

If we are managing multiple interdependent resources at the same time, then we have a multi-resource revenue management problem. An extensively studied such problem is the network revenue management problem in airline industry, where the resources are seats on different flight legs in an airline network. Such problems are usually modeled as a multi-dimensional dynamic program, whose state variable keeps track of remaining capacities of different resources and thus suffers from the so-called "curse of dimensionality". When we are managing a single resource, we have a single-recourse revenue management problem. A typical such problem is the dynamic pricing of a perishable product studied by Gallego and Van Ryzin (1994). While the single-resource problems are in general easier than their multi-resource counterparts, there are single-resource revenue management problems which have to be modeled by multi-dimensional dynamic programs, hence share similar difficulties with multi-resource problems. We study three such problems in this dissertation.

In many cases, the exact dynamic programming formulations for multi-dimensional single-resource problems are intractable to solve; in other cases, even if the dynamic programming formulation can be solved exactly, the optimal policy from the exact formulation could be inconvenient to implement. The common theme of the dissertation is to develop tractable approximations to the exact formulations, which give rise to operationally attractive policies.

This dissertation is structured into four chapters beyond this introduction. Chapter 2 studies a strategic decision problem to decide whether or not to discontinue a product that is sold under warranty and is subject to failure due to multiple causes. Chapter 3 studies a dual-channel pricing and capacity allocation problem for a hotel when a block of rooms have to be reserved for a special event such as a conference. Chapter 4 studies a problem where we can allocate the total capacity into different markets and charge separate prices at different markets, or serve all the customers in a single market with a common price. In the former case, we achieve pricing flexibility at the cost of capacity flexibility; whereas in the latter case, we achieve capacity flexibility at the cost of pricing flexibility. Chapter 5 presents concluding remarks and suggestions for future research. A more detailed overview follows.

In Chapter 2, we consider the problem faced by a company selling a product with warranty, under partial information about the reliability of the product. The product can fail from multiple failure types, each of which is associated with an inherently different repair cost. If the product fails within the warranty duration, then the company is responsible to pay the repair costs. The company does not know the probabilities associated with different failure types, but it learns the failure probabilities as sales occur and failure information is accumulated. If the failure probabilities turn out to be too high and it becomes costly to fulfill the warranty coverage, then the company may decide to stop selling the product, or replace it with a more reliable alternative. The objective is to decide if and when to stop. The multiple failure types make this problem a high dimensional decision problem.

By formulating the problem as an optimal stopping problem with Bayesian learning, we establish structural properties of the optimal policy. The state variable of the dynamic program is a vector representing our estimates for the probabilities of different failure types, and the high-dimensional state space makes it intractable to compute the optimal policy. We propose two approximation methods. The first method is based on decomposing the problem by failure types, and it provides upper bounds on the value functions. The second method provides lower bounds on the value functions, and it is based on a deterministic approximation. By using asymptotic analysis, we extract insights on when it is important to actively learn the failure probabilities. Computational experiments are carried out to evaluate the performance of the proposed policies.

In Chapter 3, we study the pricing and capacity allocation problem in hotel revenue management when a block of rooms have to be reserved for a special event such as a conference. In this setting, the conference organizer requests to reserve a block of rooms for conference participants at a fixed price, and the reservation expires if the reserved rooms are not fully booked by a deadline before the date of conference. This adds a conference market channel parallel to the spot market. The price at the conference market is agreed upon and fixed, whereas dynamic pricing can be employed in the spot market. In addition to finding the optimal pricing policy in the spot market, at a tactical level, the hotel manager needs to decide the number of rooms to reserve and the fixed price to offer for conference participants.

For the operational level pricing problem in the spot market, while the exact formulation is a reasonably tractable two-dimensional dynamic program, the optimal policy is not attractive because we need to keep track of sales in both the spot market and conference market to make pricing decisions in the spot market. We construct an asymptotically tight single dimensional approximation for the exact formulation. In the single dimensional approximation, we only keep track of the sales in the spot market and model the remaining capacity in the conference market by a static random variable. The policy from the single dimensional approximation is not only appealing from an operational perspective, its performance is also robust to load factor and tactical level decision inputs. For the tactical level decision problem, based on a deterministic formulation of a relaxed problem, we develop a heuristic to make decisions on conference market pricing and capacity allocation between the spot and conference markets simultaneously. For the overall problem, the policy by combing the tactical level heuristic and the operational level approximation is asymptotically optimal, and it also provides remarkably good performance in numerical experiments.

In Chapter 4, we extend the problem in Chapter 3 to multiple markets and more general setting. We sell a product to different markets with heterogeneous price sensitivities. A strategic decision is whether we should allocate the capacity to different markets and charge different prices in different markets (separate pricing), or we should merge all the markets together and serve them with a common price (joint pricing). If the former is chosen, we gain pricing flexibility while lose capacity flexibility by committing certain number of inventories to each market. If the latter is chosen, we lose pricing flexibility since a common price needs to be offered, however, we have capacity flexibility since there is no market separation and we can use any remaining capacity to satisfy the demand from any markets. We characterize the region where the benefit of pricing flexibility outweighs the benefits of capacity flexibility. The characterization also motivates a hybrid model where we adopt separate pricing at the early stage and switch to joint pricing towards the end of the selling horizon. Such hybrid settings occur when a number of hotel rooms are sold simultaneously in a spot market and a conference market, and the spot market and conference market are kept separate until a predetermined deadline. In the hybrid model, we need to keep track of the remaining inventories in every markets and the state variable in the dynamic programming formulation ends up being a multi-dimensional vector, which destroys the tractability of the model. We develop heuristics based on a single-dimensional approximation for each market.

Computational experiments compare the performance of the policy from the single-dimensional approximation with policies that charge fixed prices obtained from a deterministic approximation. Our single-dimensional approximation provides significant improvements especially when the capacity is tight, the number of markets is large, or the price sensitivities of customers in different markets differ from each other substantially.

CHAPTER 2

BALANCING REVENUES AND REPAIR COSTS UNDER PARTIAL INFORMATION ABOUT PRODUCT RELIABILITY

2.1 Introduction

We consider the problem faced by a company selling a product with warranty to customers, while the reliability information of the product covered by the warranty is only partially available to the company. The product can be a physical product or a warranty-type service agreement. Whenever a sale occurs, the company receives a one-time or monthly payment from its customers. In return, it is responsible to cover the repair costs of the product during a certain warranty duration. Typically, a product can fail due to multiple failure types, each of which is associated with an inherently different repair cost and an unknown failure probability. Initially, the company only has rough estimates of these failure probabilities, usually from experience with similar products or a short test marketing stage that involves a small number of units. If the true failure probabilities turn out to be too high and it becomes costly to fulfill the warranty coverage, then the company may want to stop selling such product, or replace it with a more reliable alternative. As shown in the following examples, problems of this flavor naturally arise in a variety of industries.

Example 1: Printer Technical Support Agreement. To minimize downtime of a large printing equipment, customers, such as schools and business organizations, often buy a technical support agreement from either the manufacturer or a certified service provider for their printers. Companies providing such services include Xerox, MIDCOM Service Group and CPLUS Critical Hardware Support. Customers pay either a one-time or monthly fee to the service provider, and the service provider is responsible for on-site or depot repairs of the printer during the agreed time period ranging from one to five years (see Xerox Corporation for a sample service agreement). The service provider usually maintains a team of technicians and carries out the repair service itself. The printer can break down due to fuser, motor, network failures and so on. The labor and part costs associated with different failures can vary widely. If the failure probabilities prove to be very high, then the company may choose to stop selling this service or increase its fee for future contracts. Similar technical support services for other types of equipments widely exist.

Example 2: Extended Vehicle Warranty. Most new vehicles come with a manufacturer warranty, which typically covers the repair of components for a limited duration, for example, three years or 36,000 miles, whichever comes first. When the initial manufacturer warranty expires, the customer has the option to buy an extended vehicle warranty, mostly from a third party service provider such as Warranty Direct and CARCHEX (see, for example, Top Extended Vehicle Warranty Providers). The warranty contract works essentially like an insurance. The service provider charges a monthly fee to the customer for a certain warranty duration ranging from several months to ten years, and if anything covered by the warranty breaks within this time period, then the service provider pays the repair cost, including material and labor costs. Depending on the type of warranty contract and the failed components, the repair costs can vary from a few hundred dollars to several thousands. If the failure probabilities turn out to be too high, then the service provider may choose to increase the payment rate of such warranty service for future demands.

Example 3: Cell Phone Warranty. Cell phone service providers routinely introduce new cell phone models into the market. These companies make a certain revenue with the sale of each cell phone, while each sold cell phone comes with a warranty that covers failures for a certain duration. Generally, a cell phone can fail due to failures in five to fifteen different categories, while the repair costs for failures in different categories can vary substantially. For instance, according to a nationwide repair service provider Mission Repair, the screen repair for iPhone 4S costs $99, the home button repair costs $79, while the battery repair costs $39. If the failure probabilities turn out to be too high and the warranty coverage cost offsets a large portion of the sales revenue, then the company may decide to stop selling a particular cell phone model, and turn its attention to other possible alternatives.

The above motivating applications share common features. First, the product covered by the warranty is fairly complex and it can fail in multiple failure types, while the repair costs associated with these failure types can vary substantially. Second, the company selling the product has limited information about the failure probabilities. Third, the company has the option to stop selling current product, possibly with some follow up actions such as increasing the payment rate, terminating the product, or switching to another more reliable alternative. In case a stopping decision is made, the sold warranty contracts usually need to be honored by the company. The fundamental tradeoff is that if the company waits too long to get a better understanding of the reliability of the product, then it may introduce a lot of units into the market and incur excessive refurbishment costs due to them. On the other hand, if the company does not spend enough time collecting data, then it may base its decisions on unreliable estimates of the failure probabilities.

In this chapter, motivated by the applications above, we analyze an optimal stopping model that balances revenues and repair costs under partial information about product reliability. We formulate the problem as a dynamic program with a high-dimensional state variable that keeps track of our beliefs about the probabilities of different failure types. As new failure information is accumulated at each time period, we adjust our beliefs about the probabilities of different failure types according to a Bayesian updating scheme. At each time period, we decide whether to continue or stop selling the product. The objective is to maximize the total expected profit, which is given by the difference between the revenue from the sales and the total expected repair cost.

We give a characterization of the optimal policy by showing that the value functions in our dynamic programming formulation are decreasing and convex in our estimates of the failure probabilities (Proposition 2.4.2). By using this result, we establish that the optimal policy is of boundary type (Propositions 2.4.3) and compare the optimal stopping boundaries at different time periods (Proposition 2.4.4). To deal with the high-dimensional state variable, we give two tractable approximation methods. The first method decomposes the problem into a collection of one-dimensional dynamic programs, one for each failure type. By combining the value functions obtained from each one-dimensional dynamic program, we obtain upper bounds on the original value functions (Proposition 2.5.1). The second approximation method is based on a deterministic formulation that ignores the benefits from future learning. Complementing the first method, we show that the second approximation method provides lower bounds (Proposition 2.6.1). By using the second approximation method, we demonstrate that the value of learning the failure probabilities is larger when the demand quantities in the different time periods are smaller

(Proposition 2.6.2). Finally, we establish that the lower bounds from the second approximation method are asymptotically tight as the demand in each time period scales linearly with the same rate (Proposition 2.6.3). Our numerical experiments compare the performances of the policies from the two approximation methods with a heuristic policy that is based on aggregating all failure types into a single failure. The policy from upper bound approximation provides impressive improvement with reasonable computational effort, especially when the number of possible failure types is large.

At a higher level, we formulate a joint adaptive learning and decision making problem that can appear in a variety of industries. The focus of the chapter is then to develop solutions for the model and analyze the learning dynamics. With respect to methodological contributions, first, it turns out that our model is a generalized one-armed bandit problem, where the expected reward of the unknown arm depends on n unknown parameters. This is a useful generalization of the one-armed bandit problem, which does not appear in the literature as far as we know. Second, we characterize various properties of the learning dynamics by using stochastic ordering and asymptotic analysis. Finally, we develop computationally tractable approximation methods that can provide good policies for use in practice.

The rest of the chapter is organized as follows. An overview of the related literature is provided in Section 2.2. We derive a dynamic programming formulation in Section 2.3, followed by the structural properties of the optimal policy in Section 2.4. In Section 2.5, we develop an upper bound approximation based on a dynamic programming decomposition idea. Section 2.6 provides a lower bound approximation based on a deterministic formulation, along with detailed analysis for the benefits of learning. Computational experiments comparing the policies obtained by the two approximation methods and a heuristic based on single-failure approximation appear in Section 2.7. Section 2.8 gives possible extensions for our model. We conclude with directions for future research in Section 2.9.

2.2 Literature Review

Our work in this chapter is related to several streams of literature. The first stream of work involves operations management models that have embedded optimal stopping or Bayesian learning features. Feng and Gallego (1995) consider optimal timing of a single price change from an initial price to a fixed lower or higher price. They show that the optimal policy is characterized by sequences of time thresholds that depend on the number of unsold units. Aviv and Pazgal (2005) use partially observed Markov decision processes to analyze a dynamic pricing problem while learning an unknown demand parameter. They develop upper bounds on the expected revenue and propose heuristics based on modifications of the available information structure. Bertsimas and Mersereau (2007) give a model for adaptively learning the effectiveness of ads in interactive marketing. The authors formulate the problem as a dynamic program with a Bayesian learning mechanism and propose a decomposition based approximation approach. Our decomposition idea resembles theirs. Caro and Gallien (2007) study a stylized dynamic assortment problem by using the multi-armed bandit framework. They develop a closed form dynamic index policy and extend the policy to more realistic environments by incorporating lead times, switching costs and substitution effects. Araman and Caldentey

(2009) study a dynamic pricing problem under incomplete information for the market size. They formulate the problem as a Poisson intensity control problem, derive structural properties of the optimal policy and give tractable approximation methods. Farias and Van Roy (2010) study a dynamic pricing problem, where the willingness to pay distribution of the customers is known, but the customer arrival rate is unknown. The authors develop a heuristic to learn the customer arrival rate dynamically and give a performance guarantee for their heuristic. Harrison, Keskina, and Zeevi (2011) consider a pricing problem where one of the two demand models is known to apply, but they do not know which one is the one. They give asymptotic analysis for a family learning policies. Arlotto, Chick, and Gans (2011) study the optimal hiring and retention problem via an infinite armed bandit model. They characterize the optimal policy by a Gittins index policy and develop tractable approximations to the Gittins index.

The second stream of related work involves decomposition methods for multi-dimensional dynamic programs. Hawkins (2003) develops a Lagrangian relaxation method for the so-called weakly coupled dynamic programs, where the original problem can be decomposed into a collection of independent single-dimensional subproblems except for a set of linking constraints on the action space. Adelman and Mersereau (2008) compare the approximate linear programming approach and the Lagrangian relaxation method for weakly coupled dynamic programs. Topaloglu (2009) explores Lagrangian relaxation method in the network revenue management setting to come up with tractable policies for controlling airline ticket sales. A similar approach is adopted in Topaloglu and Kunnumkal (2010) to compute time-dependent bid prices in network revenue management problems, whereas Erdelyi and Topaloglu (2010a) extend this approach to handle joint capacity allocation and overbooking problems over an airline network. Brown, Smith, and Sun (2010) propose a dual approach to compute bounds on value functions by relaxing the non-anticipativity constraint and provide an analysis based on duality theory.

Finally, since our problem involves choosing between continuing and stopping under incomplete information, it is related to the bandit literature. The one-armed bandit problem considers sampling from two options, where the reward from one option is known, but the expected reward from the other one depends on an unknown parameter. The goal is to maximize the total expected reward from sampling. Our model can be viewed as a generalized one-armed bandit model where the expected reward from the unknown arm depends on n different unknown parameters. Bradt, Johnson, and Karlin (1956) consider a setting that allows finite number of sampling opportunities, and show that the optimal policy is to stick with the known option until the end once we switch to this option. The optimal policy for the infinite horizon one-armed bandit problem is characterized as an index policy by Gittins (1979). Burnetas and Katehakis (1998) generalize the results in Bradt et al. (1956) to single-parameter exponential family, whereas Burnetas and Katehakis (2003) provide asymptotic approximations when the number of time periods goes to infinity. Goldenshluger and Zeevi (2011, 2009) consider a minimax formulation of the one-armed bandit problem, establish lower bound on the minimum regret under an arbitrary policy, and propose a policy that achieves a matching upper bound. When there are multiple options to choose from, the problem is referred to as a multi-armed bandit problem. Berry and Fristedt (1985) and Gittins (1989) give early analysis of multi-armed bandit problems. Such problems tend to be significantly more difficult than their one-armed counterparts, since one needs to keep track of the beliefs about the rewards of multiple options.

2.3 Problem Formulation

We sell a product over a finite selling horizon. The product can fail from multiple failure types, and we sell the product with a warranty that covers the repairs of failed units within a certain warranty duration. We incur a repair cost when a unit fails within its warranty. The probabilities of different failure types are unknown at the beginning of the selling horizon, but we obtain more accurate estimates of the failure probabilities as failure information is accumulated over time. If the failure probabilities turn out to be high and the repairs become too costly, then we may decide to stop selling the product and avoid introducing new units into circulation, while honoring the warranty coverage of existing units (Instead of terminating a product, we can easily extend our model to allow switching to a standard more reliable alternative. See Section 2.8 for more details). On the other hand, if the failure probabilities turn out to be low, then we may continue selling the product until the end of selling horizon. The objective is to decide if and when to stop selling the product so as to maximize the total expected profit, which is the difference between the total revenue from sales and the total expected repair cost.

There are n failure types indexed by $1, \ldots, n$, each of which is associated with an unknown failure probability. We assume that the true failure probabilities do not change overtime, but they are unknown to us. The selling horizon consists of the time periods $\{0, 1, \ldots, \tau\}$. At time period zero, we sell the product to obtain an initial belief for the failure probabilities, and we do not make a decision at the beginning of this time period. At the beginning of the other time periods, we need to decide whether to stop or continue selling the product. Therefore, we can view time period zero as a test marketing stage during which we form an initial belief about the failure probabilities. If we already have a prior belief for the failure probabilities, then we can modify our model to skip time period zero altogether and start directly at time period one with a prior belief. During each time period, each unit may suffer from multiple failure types, but we assume that the failures of different units and the failures of each unit from different failure types are independent of each other. Section 2.8 discusses an alternative model where each unit can fail only from a single failure type, inducing dependence across failure types.

We generate a revenue of r from each unit sold. A sold unit is covered by warranty for K consecutive time periods. If a unit under warranty fails from failure type i, then we incur a repair cost of c_i. A repaired unit remains in warranty only for the duration left in the original warranty contract. In other words, the warranty for a particular unit does not start from scratch after each repair. For simplicity, we ignore the possible lead time when a unit is being repaired. This is a reasonable assumption in the application settings we consider because companies usually have a number of spare products that they can use to immediately replace a failed unit as the failed unit is being repaired, or a technician is sent immediately when an on-site repair is requested. We assume that the repaired units have the same unknown failure probabilities as the brand new units. This assumption is reasonable when the failure types are mostly electrical or the repair is carried out by replacing the broken parts with new ones, which is again the case for the application settings that motivate our work.

We use D_t to denote the demand for the product at time period t, which is assumed to be deterministic. The deterministic demand assumption allows us to focus on the learning dynamics for the failure probabilities and to identify structural insights from our model. Extending our model to stochastic demand is an important avenue for further investigation. Under this assumption the number of units under warranty coverage in time period t, denoted by W_t, is simply the total demand in the last K time periods, that is, $W_t = \sum_{s=0}^{t} \mathbf{1}(t - s < K)\, D_s$, where $\mathbf{1}(\cdot)$ is the indicator function. In this case, W_t corresponds to the number of units that we can potentially receive as failed units for repairs at time period t. We naturally expect to receive only a fraction of these units as failed units because not all of them fail at the same time. Finally, if we let $M_t = W_0 + W_1 + \ldots + W_{t-1}$, then M_t is the total number of units that we could have potentially received as failed units for repairs up until time period t. Since the demand is a deterministic quantity, W_t and M_t are deterministic quantities as well.

2.3.1 Learning Dynamics

The learning process is based on a Bayesian update of the failure probabilities. At each time period, our prior belief about the probability of a particular failure type has a beta distribution. After observing the number of failed units from a particular failure type, we apply the Bayes rule to obtain an updated posterior belief about the failure probability. Since each unit fails independently and the beta distribution is a conjugate prior of the binomial distribution, our posterior belief continues to have a beta distribution.

Let P_{it} denote our (random) prior belief at the beginning of time period t for the probability of failure type i. We recall that M_t corresponds to the total number of units that we could have potentially received as failed units up until time period t. Using θ_{it} to denote the proportion of the M_t units that have actually failed from failure type i, we assume that the random variable P_{it} has a beta distribution with parameters $(\theta_{it}\, M_t, (1-\theta_{it})\, M_t)$. The parameters $\theta_{it}\, M_t$ and $(1-\theta_{it})\, M_t$ are the number of units that have failed and have not failed, respectively, from failure type i up until time period t. The expected value of P_{it} is $\theta_{it}\, M_t/[\theta_{it}\, M_t + (1-\theta_{it})\, M_t] = \theta_{it}$, which agrees with the intuition that the expected value of our prior belief for the probability of failure type i is equal to the proportion of units that have failed from failure type i up until time period t.

Recall that W_t is the number of units that is still under warranty at time period t. If we let the random variable Y_{it} denote the number of units that we actually receive as failed units from failure type i at time period t, then our prior belief implies that Y_{it} has the binomial distribution with parameters (W_t, P_{it}), where the second parameter P_{it} is itself a beta random variable with parameters $(\theta_{it}\, M_t, (1-\theta_{it})\, M_t)$. Binomial random variables whose second parameter has a beta distribution are commonly referred to as beta-binomial random variables. In this case, since the beta distribution is a conjugate prior for the binomial distribution, it is well known that our posterior belief at time period t for the probability of failure type i has a beta distribution with parameters $(\theta_{it}\, M_t + Y_{it}, (1-\theta_{it})\, M_t + W_t - Y_{it})$. The two parameters of this distribution correspond to the number of units that have failed and have not failed, respectively, from failure type i up until time period $t+1$. Throughout the rest of the chapter, we prefer writing the random variables P_{it} and Y_{it} as $P_{it}(\theta_{it})$ and $Y_{it}(\theta_{it})$, respectively, to explicitly emphasize the fact that the distributions of these random variables depend on θ_{it}. By conditioning on $P_{it}(\theta_{it})$, we compute the expected value of $Y_{it}(\theta_{it})$ as $\mathbb{E}\{Y_{it}(\theta_{it})\} = \mathbb{E}\{\mathbb{E}\{Y_{it}(\theta_{it}) \mid P_{it}(\theta_{it})\}\} = \mathbb{E}\{W_t\, P_{it}(\theta_{it})\} = W_t\, \theta_{it}$, which corresponds to the expected number of units that we receive as failed units from failure type i at time period t. This computation shortly becomes useful when constructing the cost function.

2.3.2 Dynamic Programming Formulation

Our prior belief at time period t for the probability of failure type i has a beta distribution with parameters $(\theta_{it} \, M_t, (1-\theta_{it}) \, M_t)$. Noting that M_t is a deterministic quantity, we only need to know θ_{it} in order to keep track of our prior belief for the probability of failure type i. Therefore, we can use $\boldsymbol{\theta}_t = (\theta_{1t}, \ldots, \theta_{nt})$ as the state variable in our dynamic programming formulation. Since θ_{it} corresponds to the proportion of the M_t units that we have actually received as failed units from failure type i up until time period t, the dynamics of θ_{it} is given by

$$\theta_{i,t+1} = \frac{\theta_{it} \, M_t + Y_{it}(\theta_{it})}{M_{t+1}} = \frac{M_t}{M_{t+1}} \, \theta_{it} + \left[1 - \frac{M_t}{M_{t+1}}\right] \frac{Y_{it}(\theta_{it})}{W_t}, \tag{2.1}$$

where we use the fact that $M_{t+1} - M_t = W_t$. Using the vector $\boldsymbol{Y}_t(\boldsymbol{\theta}_t) = (Y_{1t}(\theta_{1t}), \ldots, Y_{nt}(\theta_{nt}))$ and defining the deterministic quantity $\lambda_t = M_t / M_{t+1}$, we write the dynamics of $\boldsymbol{\theta}_t$ in vector notation as $\boldsymbol{\theta}_{t+1} = \lambda_t \, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t)$. The quantity $\lambda_t \in [0,1]$ is called the shrinkage factor.

To develop the cost structure of our dynamic program, we begin by considering the case where we continue selling the product at time period t. In this case, the number of units that fail from failure type i at time period t is given by the random variable $Y_{it}(\theta_{it})$. Since $\mathbb{E}\{Y_{it}(\theta_{it})\} = W_t \, \theta_{it}$, we incur an expected repair cost of $C_t(\boldsymbol{\theta}_t) = \sum_{i=1}^{n} c_i \, W_t \, \theta_{it}$, if we continue selling the product. If there are additional costs, such as goodwill or reputation costs, associated with failure probabilities of the product, then these costs can be incorporated into $C_t(\boldsymbol{\theta}_t)$ as well. On the other hand, if we decide to stop selling the product at time period t, then all future demands are lost, while the warranties of the already sold units need to be honored. In this case, the number of units that remain under warranty coverage at a future time period $\ell \geq t$ is given by $\sum_{s=0}^{t-1} \mathbf{1}(\ell - s < K)\, D_s$. Therefore, the number of units that we receive as failed units from failure type i at the future time period ℓ is given by a beta-binomial random variable with parameters $\big(\sum_{s=0}^{t-1} \mathbf{1}(\ell - s < K)\, D_s, P_{it}(\theta_{it})\big)$, whose expectation is given by $\sum_{s=0}^{t-1} \mathbf{1}(\ell - s < K)\, D_s\, \theta_{it}$, where we use the same conditioning argument that we use to compute the expectation of $Y_{it}(\theta_{it})$. Adding over all of the future time periods and all of the failure types, this implies that we incur an expected repair cost of $S_t(\boldsymbol{\theta}_t) = \sum_{i=1}^{n} \sum_{\ell=t}^{\infty} \sum_{s=0}^{t-1} c_i\, \mathbf{1}(\ell - s < K)\, D_s\, \theta_{it}$ from time period t onwards, given that we stop selling the product at time period t. The implicit assumption in the last cost expression is that if a unit is covered by the warranty beyond the selling horizon, then we are responsible from fulfilling the repairs for this unit until its warranty coverage expires.

We can formulate the problem as a dynamic program by using $\boldsymbol{\theta}_t$ as the state variable at time period t. If we continue selling the product at time period t, then we generate a revenue of rD_t and incur an expected repair cost of $C_t(\boldsymbol{\theta}_t)$, whereas if we stop selling the product at time period t, then we incur an expected repair cost of $S_t(\boldsymbol{\theta}_t)$. Using the learning dynamics given in (2.1), the value functions satisfy the optimality equation

$$\vartheta_t(\boldsymbol{\theta}_t) = \max \Big\{ rD_t - C_t(\boldsymbol{\theta}_t) + \mathbb{E}\{\vartheta_{t+1}(\lambda_t\, \boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}, -S_t(\boldsymbol{\theta}_t) \Big\},$$

with the boundary condition $\vartheta_{\tau+1}(\cdot) = -S_{\tau+1}(\cdot)$. If the first term in the $\max$ operator is larger than the second term, then it is optimal to continue; otherwise, it is optimal to stop. The expectation operator involves the random variable $\boldsymbol{Y}_t(\boldsymbol{\theta}_t)$. It turns out that we can simplify the optimality equation by using a relationship between $C_t(\cdot)$ and $S_t(\cdot)$. Adding $S_t(\boldsymbol{\theta}_t)$ to both sides of the optimality equation and letting $V_t(\boldsymbol{\theta}_t) = \vartheta_t(\boldsymbol{\theta}_t) + S_t(\boldsymbol{\theta}_t)$, we obtain

$$V_t(\boldsymbol{\theta}_t) = \max \Big\{ r D_t - C_t(\boldsymbol{\theta}_t) + S_t(\boldsymbol{\theta}_t) + \mathbb{E}\{V_{t+1}(\lambda_t \, \boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} - \mathbb{E}\{S_{t+1}(\lambda_t \, \boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}, 0 \Big\},$$

with the boundary condition $V_{\tau+1}(\cdot) = 0$. From the definition of $S_{t+1}(\cdot)$, we see that it is a linear function, hence we have $\mathbb{E}\{S_{t+1}(\lambda_t \, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} = S_{t+1}(\mathbb{E}\{\lambda_t \, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t)\}) = S_{t+1}(\boldsymbol{\theta}_t)$, where we use the fact that $\mathbb{E}\{Y_{it}(\theta_{it})\} = W_t \, \theta_{it}$. Therefore, in the above optimality equation we can write the expression $C_t(\boldsymbol{\theta}_t) - S_t(\boldsymbol{\theta}_t) + \mathbb{E}\{S_{t+1}(\lambda_t \, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t} \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}$ as $C_t(\boldsymbol{\theta}_t) - S_t(\boldsymbol{\theta}_t) + S_{t+1}(\boldsymbol{\theta}_t)$. In addition, using the definitions of $C_t(\cdot)$ and $S_t(\cdot)$, a simple algebraic manipulation given in Appendix A.1 shows that $C_t(\boldsymbol{\theta}_t) - S_t(\boldsymbol{\theta}_t) + S_{t+1}(\boldsymbol{\theta}_t) = K \sum_{i=1}^n c_i \, \theta_{it} \, D_t = K \boldsymbol{c}^\mathsf{T} \boldsymbol{\theta}_t \, D_t$, where we use $\boldsymbol{c} = (c_1, \ldots, c_n)$. Thus, we can write the last optimality equation as

$$V_t(\boldsymbol{\theta}_t) = \max \Big\{ (r - K \boldsymbol{c}^\mathsf{T} \boldsymbol{\theta}_t) \, D_t + \mathbb{E}\{V_{t+1}(\lambda_t \, \boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}, 0 \Big\}. \tag{2.2}$$

The optimality equation above has an intuitive explanation. According to our belief at time period t, we expect a unit to fail from failure type i with probability θ_{it} and the expression $K \boldsymbol{c}^\mathsf{T} \boldsymbol{\theta}_t$ can be interpreted as the expected repair cost of a unit over its whole warranty coverage. Therefore, the expression $r - K \boldsymbol{c}^\mathsf{T} \boldsymbol{\theta}_t$ in the optimality equation above corresponds to the expected net profit contribution of a sold unit. The optimality equation in (2.2) indicates that for each unit sold at time period t, the total expected repair cost over the whole warranty duration can be charged immediately at time period t according to our belief at this time period. Shifting the timing of costs appears in the literature frequently, but it is surprising that we can shift the timing even when we learn certain problem parameters.

The structure of the optimality equation in (2.2) is the same as that for the classical one-armed bandit problem, in the sense that the first term in the maximum can be thought as the expected reward of pulling the unknown arm. The difference is that in classical one-armed bandit problem, the expected reward of the unknown arm depends on a single unknown parameter, while in the above optimality equation, the expected reward of the unknown arm depends on n unknown parameters. From this prospective, our model can be viewed as a generalization of the classical one-armed bandit problem. This observation motivates the single failure approximation strategy we use as a benchmark policy in Section 2.7.

In the optimality equation in (2.2), it is optimal to continue selling the product at time period t whenever the state $\boldsymbol{\theta}_t$ satisfies $(r - K\boldsymbol{c}^\top\boldsymbol{\theta}_t)\, D_t + \mathbb{E}\{V_{t+1}(\lambda_t\, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} > 0$. Furthermore, the state at time period t satisfies the last inequality if and only if $V_t(\boldsymbol{\theta}_t) > 0$. Therefore, the set of states at which it is optimal to continue selling the product at time period t is given by $\mathcal{C}_t = \{\boldsymbol{\theta}_t \in [0,1]^n : V_t(\boldsymbol{\theta}_t) > 0\}$. We obtain an optimal policy by continuing selling the product at time period t if and only if the state $\boldsymbol{\theta}_t$ at this time period satisfies $\boldsymbol{\theta}_t \in \mathcal{C}_t$.

From the optimality equation in (2.2), it is clear that if there is only one failure type, then the formulation becomes a single-dimensional dynamic program, which is easy to solve. Furthermore, if two failure types have the same repair cost, they can be combined as a single failure type. In particular, if the repair costs of different failure types are all the same, then the formulation can be reduced to a single-dimensional dynamic program that uses $\sum_i \theta_{it}$ as the state variable. The presence of multiple failure types and the heterogeneity of repair costs are the primary reasons that the problem we study is difficult to solve.

2.4 Structural Properties

In this section, we first show that the value functions are componentwise decreasing and convex in the state variable. By using this structural property, we establish the monotonicity of the optimal decision in the state variable and time periods. Shaked and Shanthikumar (2007) give a detailed overview for the concept of stochastic monotonicity and convexity. Following their terminology, we say that a family of random variables $\{X(\gamma) : \gamma \in \Re\}$ is stochastically increasing if $\mathbb{E}\{\phi(X(\gamma))\}$ is increasing in γ for any increasing function $\phi(\cdot)$ on $\Re$. Similarly, a family of random variables $\{X(\gamma) : \gamma \in \Re\}$ is stochastically convex if $\mathbb{E}\{\phi(X(\gamma))\}$ is convex in γ for any convex function $\phi(\cdot)$ on $\Re$. Following an argument similar to those in the proofs of Theorems 8.A.15 and 8.A.17 in Shaked and Shanthikumar (2007), we can show the following closure properties for stochastically increasing and convex families.

Lemma 2.4.1 (Closure Properties)**.** *We have the following properties.*

(1) *Assume that the families of random variables $\{X(\gamma) : \gamma \in \Re\}$ and $\{Y(\gamma) : \gamma \in \Re\}$ are stochastically increasing and stochastically convex. Furthermore, assume that $X(\gamma)$ and $Y(\gamma)$ are independent of each other for all $\gamma \in \Re$. Then, for any $a, b \in \Re_+$, the family of random variables $\{a\, X(\gamma) + b\, Y(\gamma) : \gamma \in \Re\}$ is stochastically increasing and stochastically convex.*

(2) *Let $\{X(\gamma) : \gamma \in \Re\}$ be a family of real-valued random variables. Assume that the families of random variables $\{X(\gamma) : \gamma \in \Re\}$ and $\{Y(\theta) : \theta \in \Re\}$ are stochastically increasing and stochastically convex. Furthermore, assume that $X(\gamma)$ and $Y(\theta)$ are independent of each other for all $\gamma \in \Re$ and $\theta \in \Re$. Then, the family of random variables $\{Y(X(\gamma)) : \gamma \in \Re\}$ is stochastically increasing and stochastically convex.*

If we use $\mathsf{Binomial}(n, q)$ to denote a binomial random variable with parameters (n, p), then Example 8.B.3 in Shaked and Shanthikumar (2007) establishes that the family of random variables $\{\mathsf{Binomial}(n, p) : p \in [0, 1]\}$ is stochastically increasing and linear in the sample path sense, and hence stochastically increasing and stochastically convex. Similarly, if we use $\mathsf{Beta}(\theta\, m, (1-\theta)\, m)$ to denote a beta random variable with parameters $(\theta\, m, (1-\theta)\, m)$, and define $\mathsf{Beta}(0, m) \equiv 0$ and $\mathsf{Beta}(m, 0) \equiv 1$, then Adell, Badia, and de la Cal (1993) show that the family of beta random variables $\{\mathsf{Beta}(\theta\, m, (1-\theta)\, m) : \theta \in [0, 1]\}$ is stochastically increasing and stochastically convex. By using these two results together with Lemma 2.4.1, we show the following monotonicity and convexity result for the value functions. We defer the details of the proof to Appendix A.2

Proposition 2.4.2 (Monotonicity and Convexity of the Value Functions)**.** *For $t = 1, \ldots, \tau$ and $i = 1, \ldots, n$, the value function $V_t(\boldsymbol{\theta}_t)$ is componentwise decreasing and convex in θ_{it}.*

Based on the monotonicity property, if $V_t(\boldsymbol{\theta}_t) > 0$, then for any $\boldsymbol{\theta}'_t$ where $\theta'_{it} \leq \theta_{it}, \forall i = 1, \ldots, n$, we have $V_t(\boldsymbol{\theta}'_t) \geq V_t(\boldsymbol{\theta}_t) > 0$. In this case, we immediately obtain the following proposition, which gives a comparison between the decisions made by the optimal policy for different values of the state variable. In the statement of the next proposition, we recall that $\mathcal{C}_t = \{\boldsymbol{\theta}_t \in [0, 1]^n : V(\boldsymbol{\theta}_t) > 0\}$ denotes the set of states for which it is optimal to continue selling the product.

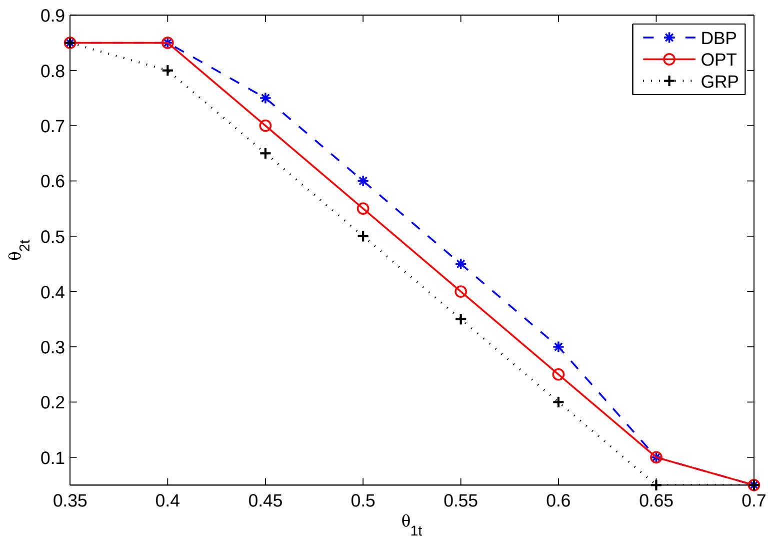
Proposition 2.4.3 (Shape of the Continuation Region)**.** *For all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t, \boldsymbol{\theta}'_t \in [0, 1]^n$ that satisfy $\theta'_{it} \leq \theta_{it}, \forall i = 1, \ldots, n$, if $\boldsymbol{\theta}_t \in \mathcal{C}_t$, then $\boldsymbol{\theta}'_t \in \mathcal{C}_t$.*

The proposition makes intuitive sense. If the proportions of the units that failed in the past from different failure types are lower, then our estimates of the failure probabilities are also lower and we are more likely to continue selling the product, all else being equal. The implication of this proposition is that the optimal policy is a boundary type policy. In other words, there is an optimal stopping boundary at time period t and if the proportions of the units that failed in the past from different failure types are above the optimal stopping boundary, then it is optimal to stop selling the product. Otherwise it is optimal to continue. The optimal stopping boundary at time period t is an $(n-1)$-dimensional hypersurface determined by the values of $\boldsymbol{\theta}_t$ that satisfy $(r - K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\theta}_t)\, D_t + \mathbb{E}\{V_{t+1}(\lambda_t\, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} = 0$. For a problem instance with two failure types, the solid line in Figure 2.1 shows the shape of the optimal stopping boundary at a particular time period t. The horizontal and vertical axes in this figure give our expected beliefs about the probabilities of the two failure types, which we indicate by $(\theta_{1t}, \theta_{2t})$. The solid line shows the values of the expected beliefs about the probabilities of the two failure types such that we would be indifferent between stopping and continuing to sell the product. To the lower left side of the stopping boundary, the optimal decision is to continue selling the product. The dashed and dotted lines show approximations to the optimal stopping boundary that we obtain by using the methods developed in Sections 2.5 and 2.6. We dwell on these stopping boundaries later in the chapter.

By making use of the monotonicity of the value functions, the following proposition gives a comparison between the decisions made by the optimal policy at different time periods. The proof of this proposition appears in Appendix A.2.

Proposition 2.4.4 (Absence of Failures Leads to Continuation)**.** *Assume that the state at time period t is $\boldsymbol{\theta}_t$ and $\boldsymbol{\theta}_t \in \mathcal{C}_t$. If no units fail at time period t, then it is still optimal to continue selling the product at time period $t+1$.*

Figure 2.1: Optimal stopping boundary and approximations to the optimal stopping boundary at $\boldsymbol{t} = \mathbf{4}$ for a problem instance with the following parameters: $N = 2, \tau = 5, K = 1, r = 1, c = (1.5, 0.5), D_l = 5, l = 0, \ldots, \tau$.



Proposition 2.4.4 makes intuitive sense because the absence of failures at time period t strengthens our belief that failures are rare. In this case, if it is optimal to continue selling the product at time period t and we do not observe any failures at this time period, then it is sensible to continue selling the product at time period $t + 1$.

The expectation in the optimality equation in (2.2) involves the random variable $Y_{it}(\theta_{it})$, which has a beta-binomial distribution with parameters $(W_t, P_{it}(\theta_{it}))$, where the random $P_{it}(\theta_{it})$ itself is a beta random variable. Computing expectations that involve such a beta-binomial random variable requires calculating beta functions, which can be problematic in practice when the parameters of the beta functions are large. Teerapabolarn (2008) demonstrates that a beta-binomial distribution can be approximated well by a binomial distribution, especially when the expectation of the beta random variable is small. This result brings up the possibility of replacing the beta-binomial random variable $Y_{it}(\theta_{it})$ in the optimality equation in (2.2) with a binomial random variable $Z_{it}(\theta_{it})$ with parameters (W_t, θ_{it}), which is a binomial random variable with the same expectation as $Y_{it}(\theta_{it})$. With this substitution, we can simplify the computation of the expectation in (2.2). Furthermore, it is possible to show that this approximation provides lower bounds on the original value functions. Details are given in Appendix A.3. We use this approximation for our numerical experiments in Section 2.7.

As we discuss at the end of Section 2.3, the difficulty of the problem is due to the high-dimensional state space in the optimality equation in (2.2), making it difficult to compute the value functions exactly when the number of failure types exceeds two or three. In the next two sections, we develop computationally tractable methods to construct approximations to the value functions. These methods scale gracefully with the number of failure types.

2.5 Upper Bounds and Decomposition-Based Policy

In this section, we develop a tractable method for approximating the value functions by decomposing the problem into a collection of one-dimensional dynamic programs, each involving a single failure type. We observe that the expected repair cost $K\boldsymbol{c}^\top\boldsymbol{\theta}_t\, D_t = K\, \sum_{i=1}^n c_i\, \theta_{it}\, D_t$ in the optimality equation in (2.2) decomposes by the failure types. Furthermore, the dynamics of $\boldsymbol{\theta}_t$ given in (2.1) implies that our beliefs about the different failure probabilities evolve independently of each other. These observations motivate writing the revenue expression in the optimality equation in (2.2) as $r D_t = \sum_{i=1}^n \rho_i D_t$, where we assume that the vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)$ satisfies $\sum_{i=1}^n \rho_i = r$, but leave it unspecified otherwise for the time being. In this case, for each $i = 1, \ldots, n$, we can solve the optimality equation

$$V_{it}^U(\theta_{it} \,|\, \boldsymbol{\rho}) = \max\Big\{(\rho_i - K c_i \, \theta_{it}) \, D_t + \mathbb{E}\{V_{i,t+1}^U(\lambda_t \, \theta_{it} + \tfrac{1-\lambda_t}{W_t} \, Y_{it}(\theta_{it}) \,|\, \boldsymbol{\rho})\}, 0\Big\}, \tag{2.3}$$

with the boundary condition $V_{i,\tau+1}^U(\cdot \,|\, \boldsymbol{\rho}) = 0$. The optimality equation in (2.3) finds the optimal policy for the case where the failures are only of type i and the revenue that we generate from each sold unit is given by ρ_i. We use the subscript i in the value functions to emphasize that the optimality equation in (2.3) focuses only on a single failure type i. The argument $\boldsymbol{\rho}$ in the value functions emphasizes that the solution to the optimality equation depends on the choice of the vector $\boldsymbol{\rho}$. As shown in the following proposition, the optimality equation in (2.3) provides upper bounds on the original value functions. The superscript U in the value functions emphasizes this upper bound property. We note that the optimality equation in (2.3) can be solved in a tractable manner because its state variable is a scalar.

Proposition 2.5.1 (Upper Bounds). *For any vector $\boldsymbol{\rho}$ satisfying $\sum_{i=1}^n \rho_i = r$, we have $V_t(\boldsymbol{\theta}_t) \leq \sum_{i=1}^n V_{it}^U(\theta_{it} \,|\, \boldsymbol{\rho})$ for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0, 1]^n$.*

Proof. We show the result by using induction over the time periods. The result trivially holds at time period $\tau+1$. Assuming that the result holds at time period

$t+1$, we have

$$
\begin{aligned}
&\sum_{i=1}^{n} V_{it}^{U}(\theta_{it}\,|\,\boldsymbol{\rho}) \\
&= \sum_{i=1}^{n} \max\Big\{(\rho_i - Kc_i\,\theta_{it})\,D_t + \mathbb{E}\{V_{i,t+1}^{U}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}, 0\Big\} \\
&\geq \max\left\{\sum_{i=1}^{n}(\rho_i - Kc_i\,\theta_{it})\,D_t + \sum_{i=1}^{n}\mathbb{E}\{V_{i,t+1}^{U}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}, 0\right\} \\
&= \max\Big\{(r - K\boldsymbol{c}^{\mathsf{T}}\,\boldsymbol{\theta}_t)\,D_t + \mathbb{E}\{\textstyle\sum_{i=1}^{n} V_{i,t+1}^{U}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}, 0\Big\} \\
&\geq \max\Big\{(r - K\boldsymbol{c}^{\mathsf{T}}\,\boldsymbol{\theta}_t)\,D_t + \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}, 0\Big\} = V_t(\boldsymbol{\theta}_t),
\end{aligned}
$$

where the second equality uses the fact that $\sum_{i=1}^{n}\rho_i = r$, the second inequality follows from the induction assumption and the fact that components of $\boldsymbol{Y}_t(\boldsymbol{\theta}_t)$ are independent, and the last equality follows from the optimality equation in (2.2). □

From the proof of the proposition above, we observe that sufficient conditions for the upper bound property to hold are the following. First, the immediate profit function should be of the form a constant plus a separable function of the state. In our problem setting, this profit is of the form $r\,D_t - K\,D_t\sum_{i=1}^{n} c_i\,\theta_{it}$. Second and perhaps more importantly, our beliefs about the probabilities of the different failure types should evolve independently of each other. No matter what learning model we use, as long as these two conditions are satisfied, the decomposition method provides upper bounds on the original value functions. Furthermore, this same decomposition method can be used in other high dimensional stopping problems as long as the continuation profit is of the form a constant plus a separable function of the state, and the different components of the state variable evolve independently of each other.

It is natural to ask how we can use Proposition 2.5.1 to choose a value for

$\boldsymbol{\rho}$ and how we can use the optimality equation in (2.3) to decide whether we should continue or stop selling the product at a particular time period. By Proposition 2.5.1, the optimal objective value of the problem

$$V_t^U(\boldsymbol{\theta}_t) = \min_{\boldsymbol{\rho}} \left\{ \sum_{i=1}^n V_{it}^U(\theta_{it} \,|\, \boldsymbol{\rho}) : \sum_{i=1}^n \rho_i = r \right\} \tag{2.4}$$

provides the tightest possible upper bound on $V_t(\boldsymbol{\theta}_t)$. In this case, we can mimic the optimal policy by defining the set of states $\mathcal{C}_t^U = \{\boldsymbol{\theta}_t \in [0,1]^n : V_t^U(\boldsymbol{\theta}_t) > 0\}$ and continuing selling the product at time period t if and only if the state $\boldsymbol{\theta}_t$ at this time period satisfies $\boldsymbol{\theta}_t \in \mathcal{C}_t^U$. We refer to this policy as the decomposition based policy. Since we have $V_t^U(\boldsymbol{\theta}_t) \geq V_t(\boldsymbol{\theta}_t)$, we obtain $\mathcal{C}_t^U \supseteq \mathcal{C}_t$. Therefore, the decomposition based policy is more likely to continue selling the product when compared with the optimal policy. The dashed line in Figure 2.1 shows the approximation to the optimal stopping boundary that we obtain by using the decomposition based policy. We observe from this figure that we indeed have $\mathcal{C}_t^U \supseteq \mathcal{C}_t$.

By using induction over the time periods, in Appendix A.4, we show that $V_{it}^U(\theta_{it} \,|\, \boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$. Furthermore, we establish that the total expected demand we observe until we stop selling the product according to the optimality equation in (2.3) gives a subgradient of $V_{it}^U(\theta_{it} \,|\, \boldsymbol{\rho})$ with respect to ρ_i. This result enables us to solve the optimization problem in (2.4) by using a standard subgradient method.

2.6 Lower Bounds and Greedy Policy

The goal of this section is to complement the approach in the previous section by providing lower bounds on the value functions. We begin this section by motivating our lower bounds through Jensen's inequality. We then show that our lower bounds correspond to the expected profits obtained by a greedy policy that makes its decisions based only on the current estimation, ignoring future learning altogether. By using the lower bounds, we construct a policy to decide if and when to stop selling the product. Finally, we establish that our lower bounds become asymptotically tight as we scale the demand in each time period linearly with the same rate.

2.6.1 A Deterministic Approximation

Our approach is based on exchanging the order in which we compute the expectation and the value function on the right side of the optimality equation in (2.2). In particular, noting that we have $\mathbb{E}\{\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t)\} = \boldsymbol{\theta}_t$, by replacing $\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}$ in the optimality equation in (2.2) with $V_{t+1}(\mathbb{E}\{\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t)\}) = V_{t+1}(\boldsymbol{\theta}_t)$, we obtain the optimality equation

$$V_t^L(\boldsymbol{\theta}_t) = \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + V_{t+1}^L(\boldsymbol{\theta}_t), 0\Big\}, \tag{2.5}$$

with the boundary condition $V_{\tau+1}^L(\cdot) = 0$. The optimality equation above does not involve any uncertainty, and the state does not change as long as we continue selling the product. In this case, letting $[\cdot]^+ = \max\{\cdot, 0\}$ and using induction over the time periods, it is possible to show that the value functions computed through the optimality equation in (2.5) are explicitly given by

$$V_t^L(\boldsymbol{\theta}_t) = [r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \sum_{s=t}^{\tau} D_s, \tag{2.6}$$

for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$. We defer the details of this simple induction argument to Appendix A.5. As shown in the following proposition, the value function $V_t^L(\cdot)$ provides a lower bound on the original value function. The superscript L emphasizes this lower bound property.

Proposition 2.6.1 (Lower Bounds). *For all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$, we have $V_t(\boldsymbol{\theta}_t) \geq V_t^L(\boldsymbol{\theta}_t)$.*

Proof. We show the result by using induction over the time periods. The result trivially holds at time period $\tau+1$. Assuming that the result holds at time period $t+1$, we have

$$\begin{aligned}
\mathbb{E}\{V_{t+1}^L(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} &= \mathbb{E}\{[r - K\boldsymbol{c}^\mathsf{T}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))]^+\} \sum_{s=t+1}^{\tau} D_s \\
&\geq [r - K\boldsymbol{c}^\mathsf{T}\mathbb{E}\{\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t)\}]^+ \sum_{s=t+1}^{\tau} D_s \\
&= (r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)^+ \sum_{s=t+1}^{\tau} D_s = V_{t+1}^L(\boldsymbol{\theta}_t),
\end{aligned} \tag{2.7}$$

where the equalities follow from the closed form expression for $V_t^L(\cdot)$ in (2.6) and the inequality follows by noting that $[\cdot]^+$ is a convex function and using Jensen's inequality. In this case, we obtain

$$\begin{aligned}
V_t(\boldsymbol{\theta}_t) &= \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}, 0\Big\} \\
&\geq \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + \mathbb{E}\{V_{t+1}^L(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}, 0\Big\} \\
&\geq \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + V_{t+1}^L(\boldsymbol{\theta}_t), 0\Big\} = V_t^L(\boldsymbol{\theta}_t),
\end{aligned}$$

where the first inequality follows from the induction hypothesis and the second one is by (2.7). □

From the discussion at the end of Section 2.3, we can interpret $r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t$ as the expected net profit contribution of a sold unit. Therefore, the expression for $V_t^L(\boldsymbol{\theta}_t)$ given in (2.6) corresponds to the total expected profit obtained by a policy that continues selling the product at all future time periods whenever the expected net profit contribution of a sold unit is positive. This policy does not consider the benefits from learning the failure probabilities at future time periods.

We can use the optimality equation in (2.5) to come up with a policy to decide whether we should continue or stop selling the product at a particular time period. In particular, if we use $V_t^L(\cdot)$ as an approximation to $V_t(\cdot)$, then we can mimic the optimal policy by defining the set of states $\mathcal{C}_t^L = \{\boldsymbol{\theta}_t \in [0,1]^n : V_t^L(\boldsymbol{\theta}) > 0\}$ and continuing selling the product at time period t if and only if the state $\boldsymbol{\theta}_t$ at this time period satisfies $\boldsymbol{\theta}_t \in \mathcal{C}_t^L$. We refer to this policy as the greedy policy. Since we have $V_t^L(\boldsymbol{\theta}_t) \leq V_t(\boldsymbol{\theta}_t)$, we obtain $\mathcal{C}_t^L \subseteq \mathcal{C}_t$, which implies that the greedy policy is more likely to stop selling the product when compared with the optimal policy. Furthermore, noting (2.6), we can write $\mathcal{C}_t^L$ as $\mathcal{C}_t^L = \{\boldsymbol{\theta}_t \in [0,1]^n : r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t > 0\}$ and the stopping boundary from the greedy policy is an $(n-1)$-dimensional hyperplane determined by $r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t = 0$. Therefore, we do not even need to solve an optimality equation to find the decisions made by the greedy policy. The dotted line in Figure 2.1 shows the approximation to the optimal stopping boundary that we obtain by using the greedy policy and it demonstrates that we indeed have $\mathcal{C}_t^L \subseteq \mathcal{C}_t$.

2.6.2 Asymptotic Analysis

Although the greedy policy is simple to compute, it does not consider the benefits from learning the failure probabilities and a natural question is when we can expect the greedy policy to perform reasonably well. In this section, we consider an asymptotic regime where we scale the demand at each time period linearly with the same rate, and show that the performance of the greedy policy becomes optimal under this regime. The asymptotic regime we consider is interesting in the following sense. On the one hand, if the demand at each time period is scaled up, then we collect a large amount of information right after the first time period and our estimates of the failure probabilities immediately become accurate. Thus, it may not be a huge problem to make decisions without considering the benefits from learning the failure probabilities and the greedy policy is expected to perform well. On the other hand, since the demand quantities at the future time periods are also large, we have the potential to collect a large amount of information about failure probabilities in the future, which may change our current belief about failure probabilities dramatically. Furthermore, since the demand quantities are large, small errors in estimating the failure probabilities may have drastic consequences. Thus, ignoring the benefits from learning may cause serious problems and the greedy policy may perform poorly. From these two conflicting perspectives, it is not clear a priori whether the greedy policy is expected to perform well or not. The rest of this section resolves this question by showing that the greedy policy is optimal under our asymptotic scaling regime.

We consider a family of problems $\{\mathcal{P}^m : m = 1, 2, \ldots\}$ indexed by the parameter $m \in \mathbb{Z}_+$. In problem $\mathcal{P}^m$, the demand at time period t is mD_t and all other problem parameters are the same as those in Section 2.3. Accordingly, in problem $\mathcal{P}^m$, we have mM_t units that we could potentially have received as failed units up until time period t. We continue using θ_{it} to denote the proportion of the mM_t units that we have actually received as failed units from failure type i. In this case, if we let $P^m_{it}(\theta_{it})$ be our prior belief at time period t for the probability of failure type i, then $P_{it}^m(\theta_{it})$ has a beta distribution with parameters $(\theta_{it}\, mM_t, (1-\theta_{it})\, mM_t)$. Similarly, in problem $\mathcal{P}^m$, we have mW_t units that we can potentially receive as failed units for repairs at time period t. We use $Y_{it}^m(\theta_{it})$ to denote the number of units that we actually receive as failed units from failure type i at time period t. In this case, $Y_{it}^m(\theta_{it})$ has a beta-binomial distribution with parameters $(mW_t, P_{it}^m(\theta_{it}))$.

We use $\{V_t(\cdot \,|\, m) : t = 1, \ldots, \tau\}$ to denote the value functions that we obtain by solving the optimality equation in (2.2) for problem $\mathcal{P}^m$. In other words, these value functions are obtained by replacing D_t with mD_t and $\boldsymbol{Y}_t(\boldsymbol{\theta}_t)$ with $\boldsymbol{Y}_t^m(\boldsymbol{\theta}_t) = (Y_{1t}^m(\theta_{1t}), \ldots, Y_{nt}^m(\theta_{nt}))$ in the optimality equation in (2.2) and solving this optimality equation. We note that λ_t in problem $\mathcal{P}^m$ does not depend on m since this quantity is given by $\lambda_t = mM_t/mM_{t+1} = M_t/M_{t+1}$. Similarly, we use $\{V_t^L(\cdot \,|\, m) : t = 1, \ldots, \tau\}$ to denote the value functions that we obtain by solving the optimality equation in (2.5) for problem $\mathcal{P}^m$. Noting (2.6), we have $V_t^L(\boldsymbol{\theta}_t \,|\, m) = m\, V_t^L(\boldsymbol{\theta}_t \,|\, 1)$ for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$.

For problem $\mathcal{P}^m$, it is optimal to continue selling the product at time period t whenever the state $\boldsymbol{\theta}_t$ at this time period is in the set $\mathcal{C}_t(m) = \{\boldsymbol{\theta}_t \in [0,1]^n : V_t(\boldsymbol{\theta}_t \,|\, m) > 0\}$. Replacing $V_t(\boldsymbol{\theta}_t \,|\, m)$ with the lower bound $V_t^L(\boldsymbol{\theta}_t \,|\, m)$, we can obtain an approximate policy for problem $\mathcal{P}^m$ by continuing selling the product at time period t whenever the state $\boldsymbol{\theta}_t$ at this time period is in the set $\mathcal{C}_t^L(m) = \{\boldsymbol{\theta}_t \in [0,1]^n : V_t^L(\boldsymbol{\theta}_t \,|\, m) > 0\}$. Since we have $V_t^L(\boldsymbol{\theta}_t \,|\, m) \leq V_t(\boldsymbol{\theta}_t \,|\, m)$ for all $\boldsymbol{\theta}_t \in [0,1]^n$, we naturally obtain $\mathcal{C}_t^L(m) \subseteq \mathcal{C}_t(m)$. Furthermore, since $V_t^L(\boldsymbol{\theta}_t \,|\, m) = m\, V_t^L(\boldsymbol{\theta}_t \,|\, 1)$, we have $\mathcal{C}_t^L(m) = \mathcal{C}_t^L(1)$ by the definition of $\mathcal{C}_t^L(m)$. Therefore, we have $\mathcal{C}_t^L(1) \subseteq \mathcal{C}_t(m)$ for all $m \in \mathbb{Z}_+$. In the following proposition, we give an ordering for $\{\mathcal{C}_t(m) : m \in \mathbb{Z}_+\}$, showing that $\mathcal{C}_t(m)$ shrinks as m increases. The proof requires two intermediate results and it is deferred to Appendix A.6.

Proposition 2.6.2 (Learning is More Beneficial for Smaller Demand)**.**
For $t = 1, \ldots, \tau$*,* $\boldsymbol{\theta}_t \in [0,1]^n$ *and* $m \in \mathbb{Z}_+$*, we have* $\mathcal{C}_t^L(1) \subseteq \mathcal{C}_t(m+1) \subseteq \mathcal{C}_t(m)$*.*

Proposition 2.6.2 indicates that the optimal policy is more likely to continue selling the product when m is small. Noting that the demand quantities are smaller when m is smaller, this result builds the intuition that if the demand quantities are smaller, then we should be more willing to learn the failure probabilities by continuing to sell the product. In other words, we intuitively expect the value of learning to be large when the demand quantities are small. In addition, Proposition 2.6.2 also shows that our deterministic approximation is at the extreme end in the sense that no matter how large the demand quantities are, our deterministic approximation is more likely to stop selling the product when compared with the optimal policy.

The following proposition shows that $V_t(\boldsymbol{\theta}_t \,|\, m)$ deviates from $V_t^L(\boldsymbol{\theta}_t \,|\, m)$ by a term that grows in the order of $\sqrt{m}$. We defer the proof to Appendix A.7.

Proposition 2.6.3 (Benefits of Learning Vanish with Increasing Demand)**.**
For all $t = 1, \ldots, \tau$ *and* $\boldsymbol{\theta}_t \in [0,1]^n$*, we have*

$$V_t^L(\boldsymbol{\theta}_t \,|\, m) \leq V_t(\boldsymbol{\theta}_t \,|\, m) \leq V_t^L(\boldsymbol{\theta}_t \,|\, m) + \bar{G}_t \sqrt{m},$$

where $\bar{G}_t$ *is a constant that depends on* $(D_0, \ldots, D_\tau)$*,* $(c_1, \ldots, c_n)$ *and* K*.*

Since we have $V_t^L(\boldsymbol{\theta}_t \,|\, m) = m\, V_t^L(\boldsymbol{\theta}_t \,|\, 1)$ for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$, Proposition 2.6.3 implies that $m\, V_t^L(\boldsymbol{\theta}_t \,|\, 1) \leq V_t(\boldsymbol{\theta}_t \,|\, m) \leq m\, V_t^L(\boldsymbol{\theta}_t \,|\, 1) + \bar{G}_t \sqrt{m}$. Therefore, as long as $V_t^L(\boldsymbol{\theta}_t \,|\, 1)$ is strictly positive, both $V_t^L(\boldsymbol{\theta}_t \,|\, m)$ and $V_t(\boldsymbol{\theta}_t \,|\, m)$ grow linearly with m, but the difference between $V_t^L(\boldsymbol{\theta}_t \,|\, m)$ and $V_t(\boldsymbol{\theta}_t \,|\, m)$ grows only in the order of $\sqrt{m}$. In other words, we have $\lim_{m\to\infty} \frac{V_t(\boldsymbol{\theta}_t \,|\, m)}{V_t^L(\boldsymbol{\theta}_t \,|\, m)} = 1$ as long as $V_t^L(\boldsymbol{\theta}_t \,|\, 1)$ is strictly greater than zero. Intuitively speaking, when the demand quantities get sufficiently large, the value of learning vanishes and the problem become reasonably easy.

2.7 Computational Experiments

In this section, we provide computational experiments to test the performance of the policies developed in Sections 2.5 and 2.6. We begin by describing our benchmark policies and experimental setup. These are followed by our computational results.

2.7.1 Benchmark Policies

We compare the following four benchmark policies.

Ideal Benchmark Policy (IDE). This benchmark policy corresponds to an idealized decision rule that is computed under the assumption that the true failure probabilities are known. If we know that the true failure probabilities are given by the vector $\boldsymbol{p} = (p_1, \ldots, p_n)$, then we generate a revenue of r and incur an expected repair cost of $K\boldsymbol{c}^\top \boldsymbol{p}$ for each sold unit. Therefore, if we have $r > K\boldsymbol{c}^\top \boldsymbol{p}$, then it is optimal to continue selling the product until the end of the selling horizon, whereas if we have $r \leq K\boldsymbol{c}^\top \boldsymbol{p}$, then it is optimal to stop selling the product as early as possible. Since there is no decision to be made at the beginning of time period zero, the expected profit obtained by IDE is simply $D_0\,[r - K\boldsymbol{c}^\top \boldsymbol{p}] + \sum_{t=1}^{\tau} D_t\,[r - K\boldsymbol{c}^\top \boldsymbol{p}]^+$. This expected profit corresponds to the idealized scenario of knowing the true failure probabilities, and it provides an unattainable upper bound on the expected profit that can be obtained by any policy that tries to learn the failure probabilities. Nevertheless, by comparing the performance of a particular policy with this upper bound, we can get a feel for how difficult the problem is and how well the policy performs.

Decomposition Based Policy (DBP). This benchmark policy corresponds to the one described in Section 2.5. In particular, if our belief for the failure probabilities at time period t is captured by the vector $\boldsymbol{\theta}_t$, then DBP solves problem (2.4) to compute $V_t^U(\boldsymbol{\theta}_t)$, and continues selling the product if and only if $\boldsymbol{\theta}_t$ is in the set $\mathcal{C}_t^U = \{\boldsymbol{\theta}_t' \in [0,1]^n : V_t^U(\boldsymbol{\theta}_t') > 0\}$.

Greedy Policy (GRP). This benchmark policy corresponds to the one described in Section 2.6. If our belief for the failure probabilities at time period t is captured by the vector $\boldsymbol{\theta}_t$, then GRP continues selling the product if and only if $\boldsymbol{\theta}_t$ is in the set $\mathcal{C}_t^L = \{\boldsymbol{\theta}_t' \in [0,1]^n : r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t' > 0\}$. We note that implementing GRP does not require solving an optimality equation at all. Furthermore, the set $\mathcal{C}_t^L$ does not depend on the time period.

Single Failure Approximation Policy (SFA). This policy is motivated by the observation that our problem is a generalized one-armed bandit problem where the expected reward of the unknown arm depends on n unknown parameters. If we can aggregate different failure types as a single failure type, then the problem reduces to a classical one-armed bandit problem. The idea behind SFA is to approximate our belief for the failure probabilities by using a single weighted average of these probabilities. The weights that we put on the different failure probabilities are the repair costs so that the more costly failures get more weight. In particular, SFA works as follows. We solve the single-dimensional optimality equation

$$V_t^s(\theta_t^s) = \max\Big\{(r - K\, c^s\, \theta_t^s)\, D_t + \mathbb{E}\{V_{t+1}^s(\lambda_t\, \theta_t^s + \tfrac{1-\lambda_t}{W_t}\, Y_t^s(\theta_t^s))\}, 0\Big\},$$

where we let $c^s = \sum_{i=1}^n c_i$, $Y_t^s(\theta_t^s) = \mathsf{Binomial}(W_t, \mathsf{Beta}(\theta_t^s\, M_t, (1-\theta_t^s)\, M_t))$ and $V_{\tau+1}^s(\cdot) = 0$. We define $\mathcal{C}_t^s = \{\theta_t^s \in [0,1] : V_t^s(\theta_t^s) > 0\}$, $t = 1, \ldots, \tau$ to capture the set of states for which we are willing to continue selling the product. If our belief for the failure probabilities at time period t is captured by the vector $\boldsymbol{\theta}_t$, then SFA continues selling the product if and only if $\theta_t^s(\boldsymbol{\theta}_t) = \frac{\boldsymbol{c}^\top \boldsymbol{\theta}_t}{c^s}$ is in the set $\mathcal{C}_t^s$. Note that $\theta_t^s(\boldsymbol{\theta}_t)$ is a weighted average of the components of $\boldsymbol{\theta}_t$.

2.7.2 Experimental Setup

We use simulation to test the performance of the four benchmark policies. For each test problem, we simulate the performance for 500 sample paths. At the beginning of each sample path, we sample the true failure probabilities $(p_1, \ldots, p_n)$ such that p_i has a beta distribution with mean μ_i and standard deviation σ_i and the components of the vector $\boldsymbol{p} = (p_1, \ldots, p_n)$ are independent of each other. Once we fix the true failure probabilities, we use these probabilities to generate the numbers of failed units throughout the selling horizon. In particular, letting $\tilde{Y}_{it}$ be the number of units that fail from failure type i at time period t, we generate $\tilde{Y}_{it}$ from the binomial distribution with parameters (W_t, p_i). At each time period, we update the state of the system by using the dynamics $\boldsymbol{\theta}_{t+1} = \lambda_t\, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\, \tilde{\boldsymbol{Y}}_t$, where $\tilde{\boldsymbol{Y}}_t = (\tilde{Y}_{1t}, \ldots, \tilde{Y}_{nt})$. The initial value of the state variable is given by $\theta_{i1} = \frac{1}{W_0} \tilde{Y}_{i0}$ for all $i = 1, \ldots, n$, which corresponds to the fraction of units under warranty that fail from failure type i at time period zero. If we are testing DBP, then we continue selling the product at time period t if and only if the state $\boldsymbol{\theta}_t$ at this time period satisfies $\boldsymbol{\theta}_t \in \mathcal{C}_t^U$. Similarly, if we are testing GRP, then we continue selling the product at time period t if and only if we have $\boldsymbol{\theta}_t \in \mathcal{C}_t^L$. If we are testing SFA, then we continue selling the product at time period t if and only if $\theta_t^s(\boldsymbol{\theta}_t) \in \mathcal{C}_t^s$. Finally, if we are testing IDE, then we continue selling the product until the end of the selling horizon if and only if $r > K\boldsymbol{c}^\top \boldsymbol{p}$. Otherwise, we stop selling the product as early as possible, which is the beginning of time period 1. By simulating the decisions of the benchmark policies that we are testing, we accumulate the profit obtained over the selling horizon. Averaging the accumulated profits on 500 different sample paths, we estimate the expected profit obtained by a particular benchmark policy. We note that IDE has access to the true failure probabilities to make its decisions, whereas DBP, GRP and SFA use only the samples of failed units given by $\{\tilde{\boldsymbol{Y}}_t : t = 0, 1, \dots, \tau\}$. This way of testing the performance of the benchmark policies corresponds to a frequentist framework, where the true failure probabilities are fixed at the beginning of a sample path and all failures occur according to these true failure probabilities. However, we note that the dynamic programming formulation in (2.2) is under a Bayesian framework, where the true failure probabilities are assumed to evolve over the time periods according to the Bayes rule. As a result, although one can use the optimality equation in (2.2) to obtain a policy to make stopping and continuing decisions over time, if the performance of this policy is tested in a frequentist framework, then the total expected profit obtained by the optimal policy can be larger than or smaller than the one predicted by the value function $V_1(\cdot)$ evaluated at the initial belief. Similarly, although we obtain tractable upper bounds $\{V_t^U(\cdot) : t = 1, \dots, \tau\}$ on $\{V_t(\cdot) : t = 1, \dots, \tau\}$ by using the decomposition approach in Section 2.5, the total expected profits collected by our benchmark policies can be larger or smaller than the upper bound $V_1^U(\cdot)$ on the value function $V_1(\cdot)$ evaluated at the initial belief.

A few setup runs indicated that if r is substantially larger than $K\boldsymbol{c}^\top\boldsymbol{p}$, then it is clearly optimal to continue selling the product and DBP, GRP and SFA are all quick to realize this fact. In this case, the performance of the three benchmark policies is comparable. Similarly, if r is substantially smaller than $K\boldsymbol{c}^\top\boldsymbol{p}$, then it is clearly optimal to stop selling the product, in which case, DBP GRP and SFA end up performing comparably as well. Therefore, test problems tend to be more difficult when r is roughly equal to $K\boldsymbol{c}^\top\boldsymbol{p}$ so that it is not easy to detect immediately whether it is optimal to continue or stop selling the product. To generate test problems with this characteristic, we set the mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ of $\boldsymbol{p}$ to satisfy $r = K\boldsymbol{c}^\top\boldsymbol{\mu}$. In this case, if the standard deviation $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ of $\boldsymbol{p}$ is small, then the sampled value of $\boldsymbol{p}$ in many sample paths roughly satisfies $r = K\boldsymbol{c}^\top\boldsymbol{p}$ and we obtain test problems for which it is difficult to detect whether it is optimal to continue or stop selling the product. Therefore, we generally expect the test problems to be more difficult as $\boldsymbol{\sigma}$ gets smaller.

In all of our test problems, the number of time periods in the selling horizon is 12 and the warranty coverage is for six time periods so that we have $\tau = 12$ and $K = 6$. We work with test problems with two or five failure types, corresponding to $n = 2$ or $n = 5$. For the demand at each time period, one set of test problems assume that $D_0 = D_1 = \ldots = D_\tau = \bar{D}$ and we vary $\bar{D}$ over $50, 100, 200, 500$ and 1000, whereas another set of test problems assume that $D_0 = 50$ and $D_1 = \ldots = D_\tau = \bar{D}$, varying $\bar{D}$ over $50, 100, 200, 500$ and 1000. Therefore, the first set of test problems correspond to the case where the demand in the test marketing stage is proportional to the demand at subsequent time periods.

As $\bar{D}$ gets larger, the demand at each time period is scaled up linearly with the same rate. This way of scaling the demand corresponds to the asymptotic regime in Section 2.6 and we expect the performance of GRP to get better as $\bar{D}$ increases. On the other hand, the second set of test problems assume that the test marketing stage always involves a small number of units, irrespective of the magnitude of the demand at the subsequent time periods. The second set of test problems arguably correspond to a more realistic case where firms go through limited test marketing effort, whereas the first set of test problems are useful to verify how quickly our asymptotic optimality result in Section 2.6 comes into play. In all of our test problems, we normalize the unit revenue to $r = 1$. For the test problems with two failure types, we set the repair costs as $\boldsymbol{c} = (0.4, 0.8)$, whereas we use $\boldsymbol{c} = (0.4, 0.5, 0.6, 0.7, 0.8)$ for the test problems with five failure types. To choose the vector $\boldsymbol{\mu}$, we assume that $\mu_1 = \mu_2 = \ldots = \mu_n$ and pick the common value by solving $r = K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\mu}$ for $\boldsymbol{\mu}$. Finally, we choose the value of $\boldsymbol{\sigma}$ by letting $\sigma_i = CV\,\mu_i$ so that CV gives the coefficient of variation for each component of the vector $\boldsymbol{p}$. We vary CV over $0.6, 0.6^2, 0.6^3$ and 0.6^4. We label our test problems by using the tuple $(D_0, \bar{D}, n, CV)$. In our first set of test problems, we have $D_0 = \bar{D}$, whereas in the second set of test problems, we have $D_0 = 50$. We vary $\bar{D}$ over $\{50, 100, 200, 500, 1000\}$, n over $\{2, 5\}$ and CV over $\{0.6, 0.6^2, 0.6^3, 0.6^4\}$. This setup provides 80 test problems for our computational experiments.

2.7.3 Computational Results

We give our computational results in Tables 2.1 and 2.2. Table 2.1 focuses on the test problems where the demand at the test marketing stage is as large as the demand at the subsequent time periods, whereas Table 2.2 focuses on the test problems where the demand at the test marketing stage is always 50. The layouts of the two tables are the same. The left and right portions of the tables respectively give the results for the test problems with two and five failure types. In each portion, the first column shows the characteristics of the test problems by using the tuple $(D_0, \bar{D}, n, CV)$. The second column shows the percent gaps between the expected profits obtained by IDE and DBP, whereas the third column shows the percent gaps between the expected profits obtained by DBP and GRP. Finally, the fourth column shows the percent gaps between the expected profits obtained by DBP and SFA. Positive gaps in the last two columns favor DBP, whereas negative gaps favor GRP or SFA. Since IDE makes its decisions with full knowledge of the true failure probabilities, it always obtains larger expected profits than the other three policies. Our computational experiments indicate that DBP generally provides improvements over the remaining two benchmark policies, GRP and SFA. Therefore, we design the presentation of our computational results to demonstrate the benefits from using DBP.

The results in Table 2.1 indicate that DBP generally provides better performance than GRP. The average performance gap between DBP and GRP is 10.44% and there are test problems where the gap between the two benchmark policies can be quite drastic. For example, DBP outperforms GRP by more than 50% in expected profit for the test problem $(50, 50, 5, 0.6^4)$. The performances of DBP and SFA are generally comparable for the test problems with two failure types, while DBP provides better performance than SFA for the problems with five failure types. It is not too surprising that SFA does not perform very well when the number of failure types gets larger because SFA makes its decisions by aggregating all failure types together.

We observe three trends in Table 2.1. First, the performance gaps between IDE and DBP decrease as the demand at each time period increases. This trend is sensible since we collect more information on the failure probabilities when the demand is larger, in which case, DBP is able to assess the failure probabilities quickly and make a sound decision on whether to continue or stop selling the product. Similarly, the performance gaps between DBP and GRP decrease as the demand at each time period increases. This observation is in agreement with the results in Section 2.6, which show that the lower bounds used by GRP become asymptotically tight as the demand at each time period increases linearly with the same rate.

Second, the performance gap between DBP and GRP increases as CV decreases. As mentioned above, if CV gets smaller, then it becomes more likely that the true failure probabilities $\boldsymbol{p}$ roughly satisfy $r = K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{p}$. In this case, it is more difficult to detect whether it is profitable to continue selling the product. It is encouraging that DBP, which captures the learning process by using a dynamic programming formulation, performs better than GRP as the underlying problem gets more difficult. A similar trend holds when we compare the performances of DBP and SFA for the test problems with five failure types.

Third, comparing the left and right portions of the table, we see that the performance gaps between IDE and DBP get larger when we have a larger number of failure types. This indicates that the problem becomes more difficult when there are more failure types. Furthermore, we observe that the performance gaps between DBP and GRP, and also between DBP and SFA tend to increase as the number of failure types increases. Overall, DBP provides noticeable improvements over GRP for the more difficult test problems, corresponding to the cases where demand at each time period is smaller, or CV is smaller, or the number of failure types is larger. DBP also provides noticeable improvements over SFA for problems with larger number of failure types and smaller CV.

Table 2.2 gives our computational results for the test problems where the demand at the test marketing stage is always 50. In these test problems, DBP uniformly performs better than GRP. Comparing Table 2.2 with Table 2.1, we also observe that the performance gaps between DBP and GRP in Table 2.2 are larger than those in Table 2.1. For the test problems in Table 2.1, GRP can make use of the larger demand quantity at the test marketing stage to form a good estimate of the failure probabilities and it is less crucial to learn the failure probabilities at the subsequent time periods. For the test problems in Table 2.2, the demand of 50 at the test marketing stage does not seem to be enough to form a good estimate of the failure probabilities and it is important to consider the benefits from learning the failure probabilities at the subsequent time periods. The trends that we observe in Table 2.1 generally appear in Table 2.2. In particular, the performance gap between DBP and GRP increases as CV decreases or the number of failure types increases. The performances of DBP and SFA are comparable for the test problems with two failure types, while DBP provides noticeable improvements over SFA for the test problems with five failure types. Furthermore, for the test problems with five failure types, the performance gap between DBP and SFA demonstrates increasing trend as CV decreases.

To comment on the computational effort for the different benchmark policies, GRP is easy to implement and the computation time per decision is negligible. SFA needs to solve a single-dimensional dynamic program once at the beginning, after which the computation time per decision is also negligible. The

Table 2.1: Computational result for the test problems with $D_0 = \bar{D}$ and $D_\ell = \bar{D}$ for $\ell = 1, \ldots, \tau$. In the last two columns, cells shaded in gray are statistically significant at 5% level.

Test Problem				Percent Gaps		
D_0	$\bar{D}$	n	CV	IDE vs. DBP	DBP vs. GRP	DBP vs. SFA
50	50	2	0.6	3.67%	2.70%	1.15%
50	50	2	0.6^2	8.95%	8.63%	1.05%
50	50	2	0.6^3	23.29%	17.32%	-1.78%
50	50	2	0.6^4	27.44%	28.92%	-5.00%
100	100	2	0.6	2.49%	0.63%	0.40%
100	100	2	0.6^2	4.51%	3.98%	0.79%
100	100	2	0.6^3	11.97%	9.48%	1.02%
100	100	2	0.6^4	17.95%	28.82%	-1.20%
200	200	2	0.6	1.27%	0.84%	0.33%
200	200	2	0.6^2	3.66%	2.68%	0.03%
200	200	2	0.6^3	8.01%	5.86%	1.22%
200	200	2	0.6^4	14.65%	12.47%	0.24%
500	500	2	0.6	0.52%	0.36%	0.09%
500	500	2	0.6^2	1.75%	0.75%	0.13%
500	500	2	0.6^3	4.39%	2.04%	0.27%
500	500	2	0.6^4	7.33%	4.34%	0.50%
1000	1000	2	0.6	0.52%	0.34%	0.01%
1000	1000	2	0.6^2	0.75%	0.48%	0.18%
1000	1000	2	0.6^3	2.77%	1.86%	0.81%
1000	1000	2	0.6^4	4.51%	6.09%	0.80%
Average				7.52%	6.93%	0.05%

Test Problem				Percent Gaps		
D_0	$\bar{D}$	n	CV	IDE vs. DBP	DBP vs. GRP	DBP vs. SFA
50	50	5	0.6	10.05%	8.87%	1.89%
50	50	5	0.6^2	19.00%	25.03%	-0.71%
50	50	5	0.6^3	32.11%	43.08%	-0.46%
50	50	5	0.6^4	37.98%	54.99%	5.21%
100	100	5	0.6	6.38%	3.80%	3.74%
100	100	5	0.6^2	12.09%	19.70%	4.13%
100	100	5	0.6^3	24.12%	32.26%	6.19%
100	100	5	0.6^4	33.40%	24.87%	6.10%
200	200	5	0.6	4.89%	2.57%	1.55%
200	200	5	0.6^2	9.21%	5.35%	3.56%
200	200	5	0.6^3	17.76%	20.26%	4.72%
200	200	5	0.6^4	24.99%	25.62%	6.13%
500	500	5	0.6	3.44%	-0.45%	1.77%
500	500	5	0.6^2	6.00%	1.96%	3.00%
500	500	5	0.6^3	13.14%	6.20%	5.90%
500	500	5	0.6^4	20.32%	8.79%	5.50%
1000	1000	5	0.6	2.32%	-0.60%	0.97%
1000	1000	5	0.6^2	5.55%	-2.80%	2.11%
1000	1000	5	0.6^3	10.74%	-0.50%	4.48%
1000	1000	5	0.6^4	17.16%	0.16%	5.59%
Average				15.53%	13.96%	3.57%

Table 2.2: Computational result for the test problems with $D_0 = 50$ and $D_\ell = \bar{D}$ for $\ell = 1, \ldots, \tau$. In the last two columns, cells shaded in gray are statistically significant at 5% level.

Test Problem				Percent Gaps		
D_0	$\bar{D}$	n	CV	IDE vs. DBP	DBP vs. GRP	DBP vs. SFA
50	50	2	0.6	3.67%	2.70%	1.15%
50	50	2	0.6^2	8.95%	8.63%	1.05%
50	50	2	0.6^3	23.29%	17.32%	-1.78%
50	50	2	0.6^4	27.44%	28.92%	-5.00%
50	100	2	0.6	3.23%	3.08%	0.71%
50	100	2	0.6^2	8.18%	8.01%	0.85%
50	100	2	0.6^3	19.70%	19.30%	0.06%
50	100	2	0.6^4	22.05%	35.82%	-3.03%
50	200	2	0.6	2.57%	3.75%	0.91%
50	200	2	0.6^2	7.40%	7.58%	0.78%
50	200	2	0.6^3	17.39%	21.93%	-1.71%
50	200	2	0.6^4	20.24%	34.31%	-5.01%
50	500	2	0.6	2.29%	3.72%	1.08%
50	500	2	0.6^2	6.35%	7.68%	1.18%
50	500	2	0.6^3	15.77%	17.75%	-1.38%
50	500	2	0.6^4	17.79%	30.42%	-3.70%
50	1000	2	0.6	2.17%	3.81%	0.82%
50	1000	2	0.6^2	5.99%	7.95%	0.61%
50	1000	2	0.6^3	14.69%	16.10%	-2.32%
50	1000	2	0.6^4	16.05%	29.47%	-4.46%
Average				12.26%	15.41%	-0.96%

Test Problem				Percent Gaps		
D_0	$\bar{D}$	n	CV	IDE vs. DBP	DBP vs. GRP	DBP vs. SFA
50	50	5	0.6	10.05%	8.87%	1.89%
50	50	5	0.6^2	19.00%	25.03%	-0.71%
50	50	5	0.6^3	32.11%	43.08%	-0.46%
50	50	5	0.6^4	37.98%	54.99%	5.21%
50	100	5	0.6	9.85%	8.51%	0.88%
50	100	5	0.6^2	14.11%	24.28%	4.16%
50	100	5	0.6^3	25.23%	39.71%	2.86%
50	100	5	0.6^4	37.74%	51.75%	10.29%
50	200	5	0.6	7.88%	8.80%	3.17%
50	200	5	0.6^2	12.69%	26.36%	4.52%
50	200	5	0.6^3	22.21%	41.02%	5.19%
50	200	5	0.6^4	31.20%	34.29%	5.96%
50	500	5	0.6	6.97%	9.40%	2.23%
50	500	5	0.6^2	10.84%	26.20%	3.79%
50	500	5	0.6^3	19.95%	35.56%	4.43%
50	500	5	0.6^4	26.68%	42.79%	8.61%
50	1000	5	0.6	6.67%	8.65%	3.84%
50	1000	5	0.6^2	10.25%	24.66%	7.28%
50	1000	5	0.6^3	17.72%	37.49%	9.93%
50	1000	5	0.6^4	24.90%	30.97%	11.34%
Average				19.20%	29.12%	4.72%

computation time of DBP grows roughly linearly with the number of failure types. For the test problems with five failure types, DBP takes about one minute per decision on a Pentium IV Desktop PC with 2.4 GHz CPU and 4 GB RAM running Windows XP. This computation time is mostly spent on solving problem (2.4) with a subgradient search method.

To sum up, our computational experiments indicate that while GRP and SFA is easy to implement, DBP can provide improvements with reasonable computational effort. In addition, the performance improvement of DBP over GRP and SFA is especially impressive for the test problems that we expect to be more difficult. In particular, DBP performs quite well when the demand at the test marketing stage is small, or the number of failure types is large, or the true failure probabilities are such that it is not immediately clear whether it is optimal to continue or stop selling the product.

2.8 Extensions

In this section, we describe several extensions that are relatively straightforward to incorporate into our model. All of the results in the chapter go through with almost no modification under the first three extensions. The last extension is based on the Dirichlet-multinomial learning model, which allows correlations in our beliefs about the probabilities of different failure types. We briefly summarize which results in the chapter still hold under this learning model.

Standard Substitute Product. We can extend our model to deal with the case where there is a standard substitute product that brings a known net profit contribution of r_0 per sold unit and stopping selling the current product means switching to this standard product. In this case, all of our results go through once we replace r with $r - r_0$ in the optimality equation in (2.2).

Other Warranty Structures. The warranty coverage in the chapter is for K time periods, but it is possible to work with other forms of warranty. For example, units may be covered until the end of the selling horizon irrespective of when they were sold or the warranty duration may depend on when the unit was actually sold to a customer. The key observation is that since the demands at different time periods are deterministic, it is straightforward to compute the number of units that are under warranty coverage at any time period. All of our results go through as long as we can compute the number of units that are under warranty coverage at any time period.

Other Cost Structures. The cost structure in our model assumes that the cost of repairing a unit that fails from a particular failure type is constant, irrespective how long the unit has been with the owner. It is not difficult to use repair costs that depend on how long the unit has been with the owner. In particular, if we use c_{ik} to denote the repair cost of a unit that fails from failure type i after having been with the owner for k time periods, then a close inspection of the dynamic programming formulation in Section 2.3 shows that we can capture such age-dependent repair costs by replacing $K \boldsymbol{c}^\top \boldsymbol{\theta}_t$ in (2.2) with $\sum_{i=1}^{n} \sum_{k=0}^{K-1} c_{ik}\, \theta_{it}$. More generally, all of our results go through if we replace $K \boldsymbol{c}^\top \boldsymbol{\theta}_t$ in (2.2) with an additively separable concave increasing function. This allows us to model possible nonlinear costs associated with product failures.

Dirichlet-Multinomial Learning Model. By using the Dirichlet-multinomial model, we can capture the situation where each product can fail from one failure at a time. This induces correlations in our beliefs about the probabilities of different failure types. Under the Dirichlet-multinomial learning model, the learning dynamics is similar to that under the beta-binomial model we use in this chapter. In particular, at time period t, the learning dynamics can still be summarized by $\boldsymbol{\theta}_{t+1} = \lambda_t \, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t} \, \boldsymbol{Y}_t(\boldsymbol{\theta}_t)$, but the difference is that $\boldsymbol{Y}_t(\boldsymbol{\theta}_t)$ is now a Dirichlet-multinomial random variable instead of a random vector consisting of n independent beta-binomial random variables. With this new interpretation of the learning dynamics, we reach the same dynamic programming formulation as in (2.2). Under the Dirichlet-multinomial learning model, the monotonicity in Propositions 2.4.2 and Proposition 2.4.3 2.4.4 continue to hold. That is, the value functions are componentwise decreasing, an optimal stopping boundary exists and the optimality of continuing decision is preserved in the next time period if there are no failures in the current time period. The upper bound in Proposition 2.5.1 may not hold anymore, since the components of $\boldsymbol{Y}_t(\boldsymbol{\theta}_t)$ are not independent. It requires more advanced analysis in multi-dimensional stochastic orders to prove or disprove the upper bound, and we have to leave it as an open question for future research. The lower bound in Proposition 2.6.1 still holds, which implies that we still obtain a lower bound from the deterministic approximation. Proposition 2.6.2 uses properties of beta-binomial random variables, and it may not hold under the Dirichlet-multinomial learning model. Proposition 2.6.3 still holds so that the lower bound is asymptotically tight when we scale the demand in each time period.

2.9 Conclusions

In this chapter, we studied a problem faced by a firm selling a product under limited information about the probabilities of different failure types. The goal is to learn the failure probabilities as the sales take place and dynamically decide whether it is profitable to continue or stop selling the product. Our approach builds on a dynamic programming formulation with embedded optimal stopping and learning features, which is a generalized one-armed bandit problem, where the expected reward of the unknown arm depends on multiple unknown parameters. This dynamic programming formulation has a high-dimensional state variable, and we proposed two approximation methods to address the computational difficulties due to the high-dimensional state variable. The first approximation method focuses on each failure type individually and solves a sequence of dynamic programs with scalar state variables. The second approximation method is based on a deterministic formulation that ignores the benefits from learning the failure probabilities. The two approximation methods are complementary to each other in the sense that while the first approximation method provides upper bounds on the value functions, the second one provides lower bounds. Our computational experiments indicated that although the approximation method based on upper bounds is more computationally intensive, it can provide significant performance improvements.

Other than the extensions discussed in the previous section, there are other practically important directions that would like to pursue and we leave these directions as questions for future research. First, our model assumes that the failure probabilities are unknown constants, but it is conceivable that the failure probabilities may depend on age or usage of a unit. Second, our beta-binomial learning model assumes that different types of failures occur independently of each other, while Dirichlet-multinomial model described in the previous section ensures that each product can fail from one failure at a time, inducing negative correlations between numbers of failures of different types. In reality, there may be general correlations among the probabilities of different failure types and one may desire to adopt other learning models that allow general correlations. For example, instead of using Dirichlet distribution, we may use generalized Dirichlet distribution which allows general correlations and is still a conjugate prior for multinomial distribution. Third, it is worthwhile to investigate the possibility that it may be better for the company not to serve all of the demand in a time period. By rationing supply in the early time periods, the company may be able to learn about the reliability of the product while controlling the risk of facing too many returns. Lastly, our approach in this chapter assumes that the demand is deterministic. There are some settings with accurate forecasts or advance demand information that make this assumption reasonable, but it is certainly desirable to relax this assumption and incorporate stochastic demands into the model. While analytical analysis may be difficult, some computational work may be interesting and useful for practical applications.

CHAPTER 3

PRICING AND CAPACITY ALLOCATION IN DUAL-CHANNEL HOTEL REVENUE MANAGEMENT

3.1 Introduction

In this chapter we study a pricing and capacity allocation problem arising in hotel revenue management setting with two sales channels. One channel is the spot market where the hotel manager can adjust the price of hotel rooms dynamically, while in the other channel, the hotel manger needs to post a fixed price at the beginning of the selling horizon. The problem arises when a conference is to be held at location near the hotel. The capacity allocation between the two channels also needs to be decided at the beginning of the selling horizon, which can not be changed later on. The hotel operates in the two parallel channels until some deadline, after which the two channels join together and the price can be set dynamically.

For a mid-to-large size conference, the conference organizer usually negotiates with the nearby hotels to reserve a block of rooms for conference participants at a fixed price. Such hotel offer information is posted on the conference web site and accessible to participants. This type of arrangement creates a conference market with fixed price parallel to the ordinary spot market with dynamic pricing. In the conference market, constrained by the availability of the reserved block of hotel rooms, conference participants can book rooms at the posted fixed price before certain deadline, after which the fixed price offer in the conference market expires. In case the reserved block in the conference market does not get fully booked by the deadline, the remaining rooms in the conference and spot market join together, which can be priced dynamically and are sold until the targeted night of stay. We call the final sales period after the deadline the final market.

From the hotel's perspective, there are several decisions to make under such circumstances. At a tactical level, the hotel needs to decide the number of rooms to reserve for the conference and the fixed price to offer in the conference market. At an operational level, the hotel needs to decide its pricing policy in the spot and final markets. Note that the tactical decisions are inputs to the operational pricing problem. We first study the operational pricing problem given the tactical decisions. Building on the formulation for the operational pricing problem, we then study the optimization problem in the tactical level to choose the optimal capacity allocation between the spot and conference markets and the fixed price in the conference market.

For the operational pricing problem, the exact formulation uses a two-dimensional dynamic program whose state variable keeps track of remaining capacities in both spot market and conference market. While the exact formulation is reasonably tractable, the resulting policy is not convenient to implement since we need to keep track of sales in both markets to make pricing decisions in the spot market. We construct a single dimensional approximation to the exact two dimensional formulation, which is asymptotically tight when the capacity and number of time periods scale up linearly in the same rate. The policy from the approximation is very appealing from an operational perspective, since we only needs to keep track of the number of rooms left in the spot market to make pricing decision. In numerical experiments, the policy has a very robust performance with respect to both load factor and higher level decision. In particular, the policy based on our approximation captures more than 99% of the optimal revenue for any given tactical decisions on average. In contrast, the fixed price policy based on a deterministic formulation can lose more than 13% of the revenue if the tactical decision is not optimized.

The tactical decision problem turns out to be difficult to solve exactly due to lack of structural properties. Based on a deterministic formulation of a relaxed problem, we develop a heuristic to make decisions on capacity allocation and conference market pricing simultaneously. For the overall problem, we construct a mixed policy by combing the tactical level heuristic and the operational level approximation. The policy is shown to be asymptotically optimal and provides promising performance in numerical experiments.

The rest of the chapter is organized as follows. An overview of the related literature is provided in Section 3.2. In Section 3.3, we focus on the operational level pricing problem in the spot and final markets, given the tactical decisions on capacity allocation and fixed price for the conference market. We characterize structural properties of the exact formulation, and develop a single-dimensional approximation. The policy based on our single-dimensional approximation is shown to be asymptotic optimality, and its performance is evaluated through numerical experiments. In section 3.4, we focus on the tactical decision problem on capacity allocation and conference market pricing. We construct a heuristic policy to find near optimal capacity allocation and conference market price based on a deterministic formulation of a relax problem. We show that the mixed policy combing the heuristic in the tactical level and the policy based on our single-dimensional approximation in the operational level is asymptotically optimal. Numerical experiments are provided to evaluate the performance of the mixed policy for the overall problem. Section 3.5 concludes with future research directions.

3.2 Literature Review

The dynamic pricing literature in revenue management is closely related to our work. Gallego and Van Ryzin (1994) study the optimal dynamic pricing for a single product over a finite horizon with a continuous time Poisson demand model. They characterize the form of the optimal policy and show that a fixed price policy based on a deterministic version of the problem is asymptotically optimal. When the set of allowable prices is a discrete set, they propose a two-price policy with a single switch and show its asymptotical optimality. Feng and Gallego (1995) consider the optimal timing of a single price change from an initial price to a fixed lower or higher price. Gallego and Van Ryzin (1997) extend Gallego and Van Ryzin (1994) to a much more general setting that allows multi-product with a network structure and time-dependent demand models. Zhao and Zheng (2000) extend Gallego and Van Ryzin (1994) to allow nonhomogeneous demand processes and identify sufficient conditions under which the optimal price decreases over time for a given inventory level. Assuming a markup or markdown strategy, Feng and Xiao (2000b) study the dynamic pricing problem when the price has to be chosen from a predetermined discrete set. They characterize the structure properties of the value function and optimal price, and provide an exact solution for the continuous-time model. Feng and Xiao (2000a) extend the model to allow price reversal.

Maglaras and Meissner (2006) show that multi-product dynamic pricing and capacity allocation problems can be modeled in a common framework, and develop asymptotically optimal policies through fluid approximations. Zhang and Cooper (2009) consider the problem of pricing parallel flights that are substitutable with each other. They build upper and lower bounds on the value functions and use these bounds to construct heuristic policies. Erdelyi and Topaloglu (2010b) propose dynamic programming decomposition methods to solve pricing problem in network revenue management. Recent papers on multi-product dynamic pricing such as Dong, Kouvelis, and Tian (2009) and Akçay, Natarajan, and Xu (2010) incorporate customer choice into the pricing model where customers choose from a set of substitutable products according to certain utility maximization rule. Extensive overviews of pricing models can be found in papers by McGill and Van Ryzin (1999), Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003), and the book by Talluri and Van Ryzin (2005).

3.3 Dynamic Pricing in Spot Market

In this section we study the dynamic pricing problem in the spot market, given the tactical level decisions on capacity allocation and fixed price for the conference market. We assume a perfect market separation, in the sense that before the end time of the conference market, conference customers only book hotel rooms in the conference market at posted fixed price, while spot market customer only book hotel rooms in the spot market. While it is understandable that spot market customers may not be eligible for the posted fixed price in the conference market, we also assume that conference customers only book in the conference market. This assumption does not restrict the applicability of our model, as long as a certain (not necessarily fixed) portion of the conference customers book rooms in the spot market. We discuss the relaxation of this assumption in Section 3.5.

We first build a two dimensional dynamic programming model for the problem and study structural properties of the value functions and optimal policy. Then we construct a more computationally efficient approximation to the original value function and show that the approximation is asymptotically exact when the capacity and number of time periods in the selling season scale up linearly in the same rate. Besides being computationally efficient, the policy from this approximation is very attractive from an operational perspective, because we only need to keep track of the sales in the spot market to make pricing decisions. In addition, the policy based on approximation shows robust performance with respect to tactical decisions.

3.3.1 Formulation and Structural Properties

We manage a limited number of hotel rooms which are booked in a finite selling horizon from time period 1 to T. The targeted night of stay is on time period $T+1$, and the total number of rooms available is C. A block of b hotel rooms are reserve at fixed price p_c for the conference participants until time $\tau < T$. From time period 1 to τ, conference participants will book rooms in the conference market at the fixed price p_c, while other customers book rooms in the spot market. After time period τ, the remaining rooms in conference and spot market join together to be sold in the final market from time period $\tau+1$ until

T. The price in the spot market and final market can be adjusted by the hotel dynamically.

We adopt a Bernoulli model for the random demand, where there is exactly one customer arriving during one time period. In particular, for $t = 1, \ldots, \tau$, let λ_{ct} and λ_{st} be the customer arrival probabilities at time period t from the conference and spot market correspondingly, which satisfy $\lambda_{ct} + \lambda_{st} = 1$. Let $d_c(\cdot)$ and $d_s(\cdot)$ be the demand functions in conference and spot market respectively, which represent the probability that an arriving customer makes a booking given the offered price in the corresponding market. The inverse demand functions are $p_c(\cdot)$ and $p_s(\cdot)$, which map the demand rates back to the offered prices. We use demand rates as the decision variables instead of the offered prices. For $t = 1, \ldots, \tau$, in the conference market, given the fixed demand rate d_c in each time period, the revenue rate is $r_c(d_c) = d_c\, p_c(d_c)$ for any t. In the spot market, letting d_{st} denote the demand rate at time period t, the revenue rate is $r_s(d_{st}) = d_{st}\, p_s(d_{st})$.

For $t = \tau + 1, \ldots, T$ in the final market, exactly one demand arrives during time t. Let $d_f(\cdot)$ be the demand function and $p_f(\cdot)$ be the inverse demand function. Given a demand rate d_{ft}, the revenue rate is $r_f(d_{ft}) = d_{ft}\, p_f(d_{ft})$. In accordance with dynamic pricing literature, we assume that the revenue functions $r_c(\cdot)$, $r_s(\cdot)$ and $r_f(\cdot)$ are continuous, bounded and concave, while satisfying $\lim_{d\to 0} r_c(d) = 0, \lim_{d\to 0} r_s(d) = 0$, and $\lim_{d\to 0} r_f(d) = 0$.

Using the notation defined above, the following two dimensional dynamic programming recursion maximizes the total expected revenue from time period t onwards, if the number of rooms left in the spot market is x and the number of rooms left i the conference market is y. We use x to denote the remaining capacity in the spot market and y to denote the remaining capacity in the conference market. Then the sate space of the problem can be written as $\mathcal{D} = \{(x, y) \in \mathbb{Z}_+^2 \mid x + y \leq C\}$. When the remaining capacities at time period t in the two markets are given by $(x, y) \in \mathcal{D}$, we can obtain the optimal policy by solving the dynamic programming recursion

$$\begin{aligned} V_t^{d_c}(x, y) &= \max_{d_{st} \in [0,1]} \Big\{ \lambda_{st}[r_s(d_{st}) + d_{st}\, V_{t+1}^{d_c}(x-1, y)] \\ &\qquad + \lambda_{ct}[r_c(d_c) + d_c\, V_{t+1}^{d_c}(x, y-1)] + (1 - \lambda_{st}\, d_{st} - \lambda_{ct}\, d_c) V_{t+1}^{d_c}(x, y) \Big\} \\ &= \max_{d_{st} \in [0,1]} \Big\{ \lambda_{st}[r_s(d_{st}) - d_{st}\, \Delta_x V_{t+1}^{d_c}(x, y)] \Big\} \\ &\qquad + \lambda_{ct}[r_c(d_c) + d_c\, V_{t+1}^{d_c}(x, y-1)] + (1 - \lambda_{ct}\, d_c) V_{t+1}^{d_c}(x, y), \end{aligned} \tag{3.1}$$

where $\Delta_x V_t^{d_c}(x, y) = V_t^{d_c}(x, y) - V_t^{d_c}(x-1, y)$. $V_t^{d_c}(x, y)$ stands for the optimal expected revenue from time period t onwards if the remaining capacity at time period t is given by (x, y). Given the demand rate in the conference market d_c and the capacity allocated for the conference market b, the optimal expected revenue in the selling horizon is $V_1^{d_c}(C - b, b)$. We use superscript d_c to emphasize that the value function depends on the demand rate d_c determined by the fixed price p_c in the conference market. For notational simplicity, we will drop the superscript d_c throughout the rest of this section. In order to keep optimality equation (3.1) uniform when $x = 0$ or $y = 0$, we define $V_t(x, -1) = -p_c + V_t(x, 0)$ for $x \in [0, C]$, and $V_t(-1, y) = -\infty$ for $y \in [0, C]$. The boundary condition is $V_t(0, 0) = 0$, $V_{\tau+1}(x, y) = \Psi_{\tau+1}(x + y)$, where $\Psi_t(\cdot)$ is the expected revenue function in the final market defined by the following dynamic programming recursion:

$$\begin{aligned} \Psi_t(x) &= \max_{d_{ft} \in [0,1]} \Big\{ r_f(d_{ft}) + d_{ft}\, \Psi_{t+1}(x-1) + (1 - d_{ft}) \Psi_{t+1}(x) \Big\} \\ &= \max_{d_{ft} \in [0,1]} \Big\{ r_f(d_{ft}) - d_{ft}\, \Delta \Psi_{t+1}(x) \Big\} + \Psi_{t+1}(x), \end{aligned} \tag{3.2}$$

where $\Delta\Psi_{t+1}(x) = \Psi_{t+1}(x) - \Psi_{t+1}(x-1)$ and $\Psi_{T+1}(\cdot) \equiv 0$. Note that $\Psi_t(\cdot), t = \tau+1, \ldots, T$ are the value functions from a classical single dimensional dynamic pricing problem from time period $\tau+1$ to T. As shown in Gallego and Van Ryzin (1994), $\Psi_t(\cdot)$ is an increasing and concave function. The next proposition shows some structural properties for the value functions in optimality equation (3.1).

Proposition 3.3.1. *We have the following structural properties hold for* $\forall t = 1, \ldots, \tau$ *and any feasible pair of remaining capacities.*

$$(i)\ \Delta_x V_t(x+2, y) \leq \Delta_x V_t(x+1, y)$$

$$(ii)\ \Delta_x V_t(x+1, y+1) \leq \Delta_x V_t(x+1, y)$$

$$(iii)\ \Delta_x V_{t+1}(x+1, y) \leq \Delta_x V_t(x+1, y).$$

We defer the proof to Appendix B.1. Property (i) states that the marginal value of one unit capacity in the spot market is decreasing in x, or in other words, the value function is concave with respect to x. Define $\Delta_y V_t(x, y) = V_t(x, y) - V_t(x, y-1)$, and notice that $\Delta_y V_t(x+1, y) - \Delta_y V_t(x, y) = \Delta_x V_t(x+1, y+1) - \Delta_x V_t(x+1, y) \leq 0$ by (ii), hence the value function $V_t(x, y)$ is submodular. Property (iii) states that the marginal value of one unit capacity in the spot market is decreasing with respect to t.

Let $d^V_{st}(x, y)$ be the optimal demand rate at time period t with state (x, y), and $p^V_{st}(x, y)$ be the corresponding optimal price. From (3.1) we know that $d^V_{st}(x, y)$ is characterized by $r'_s(d^V_{st}(x, y)) = \Delta_x V_{t+1}(x, y)$, hence, corresponding to the three properties, we get that $d^V_t(x, y)$ is increasing in x, y and t, hence the optimal price $p^V_{st}(x, y)$ is decreasing in x, y and t. In other words, the optimal price in the spot market is smaller when the remaining capacity in either the spot market or the conference market is larger. Given the remaining capacities in both markets fixed, the optimal prices decrease over time.

3.3.2 A Single-Dimensional Approximation

Because the remaining rooms in the conference and spot market will join together in the final market at time period $\tau+1$, the sales taking place in the conference market has an impact on the pricing policy in the spot market. Intuitively, if the reserved rooms in the conference market are getting booked quickly, then we are confident that the reserved rooms will be fully booked by τ, hence we will want to keep the price in the spot market at a high level. On the other hand if the demand in the conference market is not very large and we anticipate that there will be many rooms left in the conference market, then we will want to price lower in the spot market to stimulate the demand in order to make use of the remaining capacity in the conference market after time τ to achieve higher revenue. Hence we use the two dimensional dynamic programming recursion (3.1) to keep track of the number of rooms left in both spot and conference market. While this formulation can make optimal pricing decisions in the spot market considering the booking dynamics in both markets, it is not operationally convenient for hotel managers, since we have to constantly watch both markets to make decisions in the spot market. Besides, while the two dimensional dynamic programming computation is manageable with one conference market, the computation of optimal policy quickly becomes unmanageable if, for example, there are multiple conferences to be held at the same time and the hotel reserves rooms for multiple conference markets. Essentially the formulation will be a high dimensional dynamic program and the "curse of dimensionality" prevents us to compute the optimal policy efficiently.

Starting from this observation, we want to construct an operationally simple policy. One idea is to model the dynamic pricing problem in spot market with a single dimensional dynamic program which only tracks the number of rooms left in the spot market, while the boundary condition at time period $\tau + 1$ is modified appropriately to account for the possible capacity joining from the conference market.

Given fixed demand rate d_c and the initial capacity of b in the conference market, according to the Bernoulli demand model we adopt, the total demand D_c in the conference market from time period 1 to τ has a Poisson binomial distribution with parameters $(\lambda_{c1} d_c, \ldots, \lambda_{c\tau} d_c)$. If the arrival probabilities are stationary, i.e., there exists λ_c such that $\lambda_{ct} = \lambda_c, \forall t = 1, \ldots, \tau$, then D_c has a Binomial distribution with parameters $(\tau, \lambda_c d_c)$. The remaining capacity at the end of time period τ in the conference market can be written as $b - \min\{b, D_c\} = (b - D_c)^+$ where we use $(\cdot)^+$ to denote $\max\{0, \cdot\}$. Starting from this idea, we construct the following dynamic program in the spot market from period 1 to period τ:

$$\begin{aligned}\Phi_t(x) &= \max_{d_{st} \in [0,1]} \Big\{ \lambda_{st} [r_s(d_{st}) + d_{st}\, \Phi_{t+1}(x-1)] + (1 - \lambda_{st}\, d_{st}) \Phi_{t+1}(x) \Big\} \\ &= \max_{d_{st} \in [0,1]} \Big\{ \lambda_{st} [r_s(d_{st}) - d_{st}\, \Delta\Phi_{t+1}(x)] \Big\} + \Phi_{t+1}(x), \qquad (3.3)\end{aligned}$$

with the boundary condition $\Phi_t(0) = \mathbb{E}\, \Psi_{\tau+1}((b - D_c)^+), \Phi_{\tau+1}(x) = \mathbb{E}\, \Psi_{\tau+1}(x + (b - D_c)^+)$. Note that $(b - D_c)^+$ represents the random remaining capacity in the conference market at time period $\tau + 1$. So we have modified the boundary condition to account for the possible capacity joining from the conference market at time period $\tau + 1$. Besides, $\Phi_t(x)$ represents the the total expected revenue in the spot and final market from time period t until the end of the selling horizon T, but it does not include the revenue collected in the conference market.

Let $d_{st}^{\Phi}(x)$ be the optimal demand rate derived from (3.3) at time period t with state x, and $p_{st}^{\Phi}(x)$ be the corresponding price. We call the policy that offers price $p_{st}^{\Phi}(x)$ at time period t in state x the Single Dimensional Approximation policy (SDA). By following this policy, the total expected revenue during the entire selling horizon is $\Phi_1(C-b) + p_c\mathbb{E}(\min\{b, D_c\})$, where the second term represents the expected revenue in the conference market. Observing that the optimal policy from (3.3) is feasible for (3.1), we have a lower bound on the optimal total expected revenue $V_1(C-b, b)$.

Proposition 3.3.2. *We have* $\Phi_1(C-b) + p_c\mathbb{E}(\min\{b, D_c\}) \leq V_1(C-b, b)$.

Since the single-dimensional dynamic program (3.3) is computationally easier to solve compared with (3.1), it is of interest to see the performance of the dynamic pricing policy from (3.3). Before that we need to develop several bounds and inequalities for the value functions in (3.1) and (3.3).

3.3.3 Bounds and Inequalities Based on Deterministic Problems

Consider a deterministic version of the pricing problem in the spot and final market. The hotel has $C-b$ rooms available in the spot market at time period 1, which is a *continuous* quantity. The hotel can control the demand rate in the spot and final market through pricing. For $t = 1, \ldots, \tau$, if controlled demand rate is $d_{st} \leq 1$, then the realized demand in time period t is $\lambda_{st}d_{st}$. At time period $\tau + 1$, a continuous capacity $\mathbb{E}[(b - D_c)^+]$ joins the available inventory of rooms, where D_c is the demand in conference market from time period 1 to τ. For $t = \tau + 1, \ldots, T$, if the controlled demand rate is d_{ft}, then the realized demand is d_{ft} exactly. The hotel wants to maximize its total revenue through controlling the demand rates from time period 1 to T. The following nonlinear programming solves the revenue maximization problem.

$$
\begin{aligned}
(\mathbf{NLP1})\ Z_1 = \max \quad & \sum_{t=1}^{\tau} \lambda_{st}\, r_s(d_{st}) + \sum_{t=\tau+1}^{T} r_f(d_{ft}) \\
\text{subject to} \quad & \sum_{t=1}^{\tau} \lambda_{st}\, d_{st} \qquad\qquad \le C - b \\
& \sum_{t=1}^{\tau} \lambda_{st}\, d_{st} + \sum_{t=\tau+1}^{T} d_{ft} \le C - b + \mathbb{E}(b - D_c)^+ \qquad (3.4) \\
& 0 \le d_{st}, d_{ft} \le 1, \forall t = 1, \ldots, T.
\end{aligned}
$$

The next proposition shows that the above nonlinear program provides an upper bound for revenues in the spot market through both value function (3.1) and (3.3). The proof is deferred to Appendix B.2.

Proposition 3.3.3. *For given b and d_c, we have $V_1(C-b,b) \le Z_1 + p_c \mathbb{E}\min\{b, D_c\}$, and $\Phi_1(C-b) \le Z_1$.*

Here we state three equivalent facts that will be used throughout this section. For a deterministic scalar z and a real-valued random variable Z with finite mean μ and finite variance σ^2, we have

(I) $\mathbb{E}\{[z-Z]^+\} \le [\sqrt{\sigma^2 + (z-\mu)^2} + (z-\mu)]/2 \le [z-\mu]^+ + \sigma/2$,

(II) $\mathbb{E}\{[Z-z]^+\} \le [\sqrt{\sigma^2 + (z-\mu)^2} - (z-\mu)]/2 \le [-z+\mu]^+ + \sigma/2$,

(III) $\mathbb{E}\min\{Z, z\} \ge \min\{\mu, z\} - \sigma/2$.

The first two are shown by Gallego (1992), while Fact (III) can be easily derived from Fact (II).

By moving the expectation inside the operator $(\cdot)^+$ in (3.4) and adding one additional term δ, we get a slightly different nonlinear program

$$
\begin{aligned}
(\mathbf{NLP2})\ Z_2(\delta) = \max \quad & \sum_{t=1}^{\tau} \lambda_{st} r_s(d_{st}) + \sum_{t=\tau+1}^{T} r_f(d_{ft}) \\
\text{subject to} \quad & \sum_{t=1}^{\tau} \lambda_{st} d_{st} \qquad\qquad \leq C - b \\
& \sum_{t=1}^{\tau} \lambda_{st} d_{st} + \sum_{t=\tau+1}^{T} d_{ft} \leq C - b + (b - \mathbb{E}(D_c))^+ + \delta \quad (3.5) \\
& 0 \leq d_{st}, d_{ft} \leq 1, \forall t = 1, \ldots, T
\end{aligned}
$$

Due to convexity of the function $(\cdot)^+$, Jensens' inequality tells that $(b - \mathbb{E}(D_c))^+ \leq \mathbb{E}(b - D_c)^+$. Thus when $\delta = 0$, the constraint (3.5) in (NLP2) is tighter than the constraint (3.4) in (NLP1), hence we have $Z_2(0) \leq Z_1$. On the other hand, letting $\sigma(D_c)$ be the standard deviation of D_c, and using Fact (I) we have $\mathbb{E}(b - D_c)^+ \leq (b - \mathbb{E}(D_c))^+ + \sigma(D_c)/2$. Thus constraint (3.5) with $\delta = \sigma(D_c)/2$ is looser than constraint (3.4), and we have $Z_1 \leq Z_2(\sigma_c/2)$. Thus we obtain

$$Z_2(0) \leq Z_1 \leq Z_2(\sigma(D_c)/2). \tag{3.6}$$

Noting that the objective function in (NLP2) is concave while the constraints are linear, we can show the following property.

Proposition 3.3.4. *There exists an optimal solution for (NLP2) which satisfies that* $d^*_{st}(\delta) = d^*_s(\delta), t = 1, \ldots, \tau, d^*_{ft}(\delta) = d^*_f(\delta), t = \tau + 1, \ldots, T$ *for some* $d^*_s(\delta)$ *and* $d^*_f(\delta)$. *Furthermore,* $Z_2(\delta)$ *is a concave function with respect to* δ.

Now we can rewrite (NLP2) as follows:

$$(\mathbf{NLP2'})\ Z_2(\delta) = \max \quad (\sum_{t=1}^{\tau} \lambda_{st}) r_s(d_s) + (T-\tau) r_f(d_f) \tag{3.7}$$

$$\text{subject to} \quad (\sum_{t=1}^{\tau} \lambda_{st}) d_s \leq C - b \tag{3.8}$$

$$(\sum_{t=1}^{\tau} \lambda_{st}) d_s + (T-\tau) d_f \leq C - b + (b - \mathbb{E}(D_c))^+ + \delta \tag{3.9}$$

$$0 \leq d_s, d_f \leq 1. \tag{3.10}$$

Let (d_s^*, d_f^*) be the optimal primal solution for (NLP2′) with $\delta = 0$, and (μ^*, η^*) be the corresponding optimal Lagrange multipliers associated with constraints (3.8) and (3.9). From the concavity of (NLP2′), we know that $Z_2(\delta)$ is a concave function whose subgradient at $\delta = 0$ is η^*. Hence we have

$$Z_2(\delta) \leq Z_2(0) + \delta\, \eta^*. \tag{3.11}$$

Combining (3.6) and (3.11), we have

$$Z_2(0) \leq Z_1 \leq Z_2(0) + \eta^*\, \sigma(D_c)/2. \tag{3.12}$$

From now on we will focus on (NLP2′) with $\delta = 0$. For notational simplicity, we let $Z_2 = Z_2(0) = (\sum_{t=1}^{\tau} \lambda_{st}) r_s(d_s^*) + (T-\tau) r_f(d_f^*)$. Let (p_s^*, p_f^*) be the prices corresponding to (d_s^*, d_f^*). This solution suggests a simple policy: We use a fixed price p_s^* in spot market from time period 1 to τ, and another fixed price p_f^* from time period $\tau + 1$ to T. We call this policy Fixed Price (FP) policy. If we let W be the total expected revenue collected in the spot and final market by following this policy for the original problem with stochastic demand, then we have

$$W = p_s^*\, \mathbb{E} \min\{C - b, D_s(d_s^*)\} + p_f^*\, \mathbb{E} \min\{(C - b - D_s(d_s^*))^+ + (b - D_c)^+, D_f(d_f^*)\}, \tag{3.13}$$

where $D_s(d_s^*)$ is the total demand in spot market from time period 1 to τ, which follows a Poisson binomial distribution with parameters $(\lambda_{s1}d_s^*, \lambda_{s2}d_s^*, \ldots, \lambda_{s\tau}d_s^*)$ (or a Binomial distribution with parameters $(\tau, \lambda_s d_s^*)$ with stationary arrival probabilities), and $D_f(d_f^*)$ is the total demand from time period $\tau+1$ to T, which follows a binomial distribution with parameters $(T-\tau,\ d_f^*)$. We have the following chain of inequalities:

$$1 \geq \frac{\Phi_1(C-b) + p_c\,\mathbb{E}\min\{b, D_c\}}{V_1(C-b,b)} \geq \frac{W + p_c\,\mathbb{E}\min\{b, D_c\}}{V_1(C-b,b)}$$
$$\geq \frac{W + p_c\mathbb{E}\min\{b, D_c\}}{Z_1 + p_c\,\mathbb{E}\min\{b, D_c\}} \geq \frac{W}{Z_1} \geq \frac{W}{Z_2 + \eta^*\,\sigma_c/2}. \tag{3.14}$$

The first inequality follows from Proposition 3.3.2, the second inequality follows from the observation that the FP policy is feasible for dynamic program recursion (3.3) hence $W \leq \Phi_1(C-b)$, the third inequality follows from Proposition 3.3.3, the fourth inequality follows the since $W \leq \Phi_1(C-b) \leq Z_1$, while the last inequality follows from (3.12).

3.3.4 Asymptotic Analysis

We consider a sequence of problems $\{\mathcal{P}^m : m = 1, 2, \ldots\}$ indexed by parameter $m \in \mathbb{Z}_+$. In problem $\mathcal{P}^m$, the total capacity is $m\,C$ and the number of rooms reserved for conference market is $m\,b$. The selling horizon starts from time period 1 and lasts until time period $m\,T$, while the conference market is open from time period 1 until time period $m\,\tau$, and the remaining capacity in the conference market is released to the spot market at time period $m\tau+1$. The conference market has a fixed price p_c and corresponding demand rate d_c. Let $\lceil\cdot\rceil$ be the round up function and $\lambda_{st}^m = \lambda_{s\lceil t/m\rceil}, \lambda_{ct}^m = \lambda_{c\lceil t/m\rceil}$. For time period $t \leq m\tau$ in problem $\mathcal{P}^m$, the arrival probability in the spot market is λ_{st}^m while the arrival probability in the conference market is λ_{ct}^m. We essentially repeat each time period in problem $\mathcal{P}^1$ m times to get problem $\mathcal{P}^m$, hence we use the ceiling function to establish the correspondence between the arrival probabilities in problem $\mathcal{P}^m$ and $\mathcal{P}^1$.

With this definitions, we note that the problem studied so far in this section is $\mathcal{P}^1$. For $t \leq \tau$, the arrival probabilities in both spot and conference markets at time periods $\{m(t-1)+1, \ldots, mt\}$ in problem $\mathcal{P}^m$ are the same as the arrival probabilities in corresponding markets at time period t in problem $\mathcal{P}^1$. This is a standard method in revenue management literature to scale the problem to show asymptotic optimality results. Our goal here is to show that the pricing policy derived from the single dimensional approximation in (3.3) is asymptotically optimal for problem $\mathcal{P}^m$ when m approaches infinity. For simplicity of notation, we assume stationary arrival probabilities in this subsection, i.e., $\lambda_{st}^m = \lambda_s, \lambda_{ct}^m = \lambda_c, \forall t = 1, \ldots, m\tau$, althrough the result in this subsection holds with non-stationary arrival probabilities.

For problem $\mathcal{P}^m$, the two dimensional dynamic program value function $V_1^m(m(C-b), mb)$ is the optimal total expected revenue. The single dimensional dynamic program value function $\Phi_1^m(m(C-b))$ is the expected revenue of the SDA policy in the spot and final market. D_c^m is the total demand in the conference market, which has a Binomial distribution with parameters $(m\tau, \lambda_s\, d_c)$. Z_1^m is the optimal objective value of (NLP1) for problem $\mathcal{P}^m$ and it provides an upper bound for the maximum revenue in the spot and final market, while Z_2^m is the optimal objective value of (NLP2′) for problem $\mathcal{P}^m$ with $\delta = 0$. W^m is the expected revenue of the Fixed Price policy derived from (NLP2′). The following proposition states that as m goes to infinity, the expected revenue achieved by the SDA policy together with the expected revenue in the conference market is asymptotically optimal.

Proposition 3.3.5. *We have* $\lim_{m\to\infty} \frac{\Phi_1^m(m(C-b))+p_c\,\mathbb{E}\min\{mb,D_c^m\}}{V_1^m(m(C-b),mb)} = 1.$

Proof. Note that in (NLP2′) with $\delta = 0$ for problem $\mathcal{P}^m$, all the coefficients scale linearly with m, hence the optimal primal and dual variables do not scale with m, i.e., the fixed price policy for $\mathcal{P}^m$ is the same as the fixed price policy for $\mathcal{P}^1$. In particular, the optimal solution for $\mathcal{P}^m$ is (d_s^*, d_f^*) with corresponding prices (p_s^*, p_f^*), the optimal dual variables are (μ^*, η^*), and we have $Z_2^m = mZ_2$. The chain of inequalities (3.14) for problem $\mathcal{P}^m$ can be written as

$$1 \geq \frac{\Phi_1^m(m(C-b)) + p_c\,\mathbb{E}\min\{mb, D_c^m\}}{V_1^m(m(C-b), mb)} \geq \frac{W^m}{Z_2^m + \eta^*\,\sigma(D_c^m)/2}, \tag{3.15}$$

where

$$\begin{aligned} W^m = {} & p_s^*\,\mathbb{E}\min\{m(C-b), D_s^m(d_s^*)\} \\ & + p_f^*\,\mathbb{E}\min\{[m(C-b) - D_s^m(d_s^*)]^+ + (m\,b - D_c^m)^+, D_f^m(d_f^*)\}, \end{aligned} \tag{3.16}$$

where $D_s^m(d_s^*)$ is the total demand in the spot market from time period 1 to $m\tau$, which follows a binomial distribution with parameters $(m\tau, \lambda_s\, d_s^*)$, and $D_f^m(d_f^*)$ is the total demand from time period $m\tau + 1$ to mT, which follows a binomial distribution with parameters $(m(T-\tau),\ d_f^*)$. For the first term in (3.13), we have

$$\begin{aligned} & p_s^*\,\mathbb{E}\min\{m(C-b), D_s^m(d_s^*)\} \\ & \geq p_s^*\ \{\min\{m(C-b), \mathbb{E}(D_s^m(d_s^*))\} - \sigma(D_s^m(d_s^*))/2\} \\ & = p_s^*\ \left\{\min\{m(C-b), m\tau\lambda_s d_s^*\} - \sqrt{m\tau\,\lambda_s d_s^*(1-\lambda_s d_s^*)}/2\right\} \\ & = p_s^*\ \left\{m\tau\lambda_s d_s^* - \sqrt{m\tau\,\lambda_s d_s^*(1-\lambda_s d_s^*)}/2\right\} \\ & = m\tau\lambda_s r_s(d_s^*) - G_1\sqrt{m}, \end{aligned}$$

where $G_1 = p_s^* \sqrt{\tau \lambda_s d_s^*(1-\lambda_s d_s^*)}/2$. The first inequality follows from Fact (III) and the second equality follows from constraint (3.8). Let $K(D_s^m(d_s^*), D_c^m) = [m(C-b) - D_s^m(d_s^*)]^+ + (mb - D_c^m)^+$, then the second term in (3.13) can be written as $p_f^* \mathbb{E} \min\{K(D_s^m(d_s^*), D_c^m), D_f^m(d_f^*)\}$. We have the following chain of inequalities:

$$
\begin{aligned}
& p_f^* \Big\{ \mathbb{E} \min\big\{K(D_s^m(d_s^*), D_c^m), D_f^m(d_f^*)\big\} \Big\} \\
& \geq p_f^* \Big\{ \mathbb{E} \min\big\{\mathbb{E}[K(D_s^m(d_s^*), D_c^m)], D_f^m(d_f^*)\big\} - \sigma(K(D_s^m(d_s^*), D_c^m))/2 \Big\} \\
& \geq p_f^* \ \big\{ \min\big\{\mathbb{E}[K(D_s^m(d_s^*), D_c^m)], \mathbb{E}(D_f^m(d_f^*))\big\} \\
& \qquad\qquad\qquad -\sigma(D_f^m(d_f^*))/2 - \sigma(K(D_s^m(d_s^*), D_c^m))/2 \big\} \\
& \geq p_f^* \ \big\{ \min\big\{[m(C-b) - \mathbb{E}(D_s^m(d_s^*))]^+ + [mb - \mathbb{E}(D_c^m)]^+, \mathbb{E}(D_f^m(d_f^*))\big\} \\
& \qquad\qquad\qquad -\sigma(D_f^m(d_f^*))/2 - \sigma(K(D_s^m(d_s^*), D_c^m))/2 \big\} \\
& = p_f^* \ \big\{ \min\big\{[m(C-b) - m\tau \lambda_s d_s^*]^+ + [mb - \mathbb{E}(D_c^m)]^+, m(T-\tau)\, d_f^*\big\} \\
& \qquad\qquad\qquad -\sigma(D_f^m(d_f^*))/2 - \sigma(K(D_s^m(d_s^*), D_c^m))/2 \big\} \\
& = p_f^* \ \big\{ \min\big\{m(C-b) - m\tau \lambda_s d_s^* + [mb - \mathbb{E}(D_c^m)]^+, m(T-\tau)\, d_f^*\big\} \\
& \qquad\qquad\qquad -\sigma(D_f^m(d_f^*))/2 - \sigma(K(D_s^m(d_s^*), D_c^m))/2 \big\} \\
& = p_f^* \Big\{ m(T-\tau)\, d_f^* - \sigma(D_f^m(d_f^*))/2 - \sigma(K(D_s^m(d_s^*), D_c^m))/2 \Big\} \\
& \geq p_f^* \Big\{ m(T-\tau)\, d_f^* - \sigma(D_f^m(d_f^*))/2 - \sigma(D_s^m(d_s^*))/2 - \sigma(D_c^m)/2 \Big\} \\
& = m(T-\tau)\, r_f(d_f^*) - G_2\sqrt{m},
\end{aligned}
$$

where $G_2 = p_f^* \Big\{ \sqrt{(T-\tau)\, d_f^*(1-d_f^*)} + \sqrt{\tau \lambda_s d_s^*(1-\lambda_s d_s^*)} + \sqrt{\tau \lambda_c d_c(1-\lambda_c d_c)} \Big\}/2$. The first and second inequalities follow from Fact (III), the third inequality follows from convexity of $(\cdot)^+$, the second equality follows from constraint (3.8), the third equality follows from constraint (3.9), and the last inequality follows from Lemma B.3.1 we give in Appendix B.3. Combing the above two terms, we have

$$W^m \geq m\tau\lambda_s r_s(d_s^*) + m(T-\tau)\, r_f(d_f^*) - G_1\sqrt{m} - G_2\sqrt{m} = mZ_2 - (G_1+G_2)\sqrt{m}.$$

Now we continue the chain of inequality (3.15) for problem $\mathcal{P}^m$ as

$$\begin{aligned} 1 &\geq \frac{\Phi_1^m(m(C-b)) + p_c\,\mathbb{E}\min\{mb, D_c^m\}}{V_1^m(m(C-b), mb)} \geq \frac{W^m}{Z_2^m + \eta^*\,\sigma(D_c^m)/2} \\ &\geq \frac{mZ_2 - (G_1+G_2)\sqrt{m}}{mZ_2 + \eta^*\sqrt{m\tau\,\lambda_{ct}d_c(1-\lambda_{ct}d_c)}} = \frac{Z_2 - (G_1+G_2)/\sqrt{m}}{Z_2 + \eta^*\sqrt{\tau\,\lambda_c d_c(1-\lambda_c d_c)}/\sqrt{m}} \xrightarrow{m\to\infty} 1. \end{aligned}$$

□

3.3.5 Numerical Experiments

In this section, we provide numerical experiments to evaluate the performance of the policies developed for the operational level problem on dynamic pricing in the spot and final markets. We begin by describing the benchmark policies and the experimental setup.

Benchmark Policies

We compare the following three benchmark policies.

Optimal Policy (OPT). This benchmark policy corresponds to solving the two dimensional dynamic programming recursion in (3.1) and recording the optimal prices at every possible state. In hotel operation, before the conference market offer expires, the hotel manager keeps track of the number of rooms remaining in both conference and spot market, and offers price $p_{st}^V(x,y)$ if there are x rooms left in the spot market and y rooms left in the conference market at time period t. After the conference market offer expires, the hotel manager sets prices according to the single dimensional dynamic programming recursion in (3.2) for the final market. Given conference market price p_c with corresponding demand rate d_c, and capacity allocation b, the total expected revenue achieved by OPT policy is $V_1^{d_c}(C-b,b)$. Since this is the optimal expected revenue, we use the gap between the revenues achieved by other policies and the revenue achieved by OPT policy to indicate the performance of other policies.

Single Dimensional Approximation Policy (SDA). This benchmark policy corresponds to solving the single dimensional dynamic programming recursion in (3.3). In hotel operation, before the conference market offer expires, the hotel manager only keeps track of the number of rooms remaining in the spot market and offers price $p_{st}^{\Phi}(x)$ if there are x rooms left in the spot market. After the conference market offer expires, the hotel manager sets prices according to the single dimensional dynamic programming recursion in (3.2) for the final market. Given conference market price p_c with corresponding demand rate d_c and capacity allocation b, the total expected revenue achieved by SDA policy is $\Phi_1^{d_c}(C-b)+p_c\mathbb{E}\min\{b,D_c(d_c)\}$.

Fixed Price Policy (FP). This benchmark policy corresponds to solving the nonlinear program (NLP2) with $\delta=0$, and offering fixed price p_s^* from time period 1 to τ in the spot market and fixed price p_f^* from time period $\tau+1$ to T in the final market. Given conference market price p_c with corresponding demand rate d_c and capacity allocation b, the total expected revenue achieved by FP policy is $W(b,d_c)+p_c\mathbb{E}\min\{b,D_c(d_c)\}$, where $W(b,d_c)$ is given in (3.13) as W.

Experimental Setup

The overall problem is defined by the following parameters: total number of rooms available C, total number of time periods T, deadline of conference market offer τ, arrival probabilities in the spot market $\lambda_{st}, t = 1, \ldots, \tau$, and the demand functions $d_c(\cdot)$, $d_s(\cdot)$ and $d_f(\cdot)$. In our experiments, we set $\tau = \frac{3}{4}T$ and use stationary arrival probabilities $\lambda_{st} = \lambda_s = 0.5, \lambda_{ct} = \lambda_c = 0.5, t = 1, \ldots, \tau$. We adopt exponential demand models $d_c(p) = e^{-\beta_c p}$, $d_s(p) = e^{-\beta_s p}$ and $d_f(p) = e^{-\beta_f p}$, where $\beta_c, \beta_s, \beta_f$ are price sensitivity parameters for the corresponding demand models. We set $(\beta_s, \beta_c, \beta_f) \in \{(1, 1, 0.5), (1, 0.75, 0.5)\}$ to represent two price sensitivity scenarios. The case with $(\beta_s, \beta_c, \beta_f) = (1, 1, 0.5)$ represents a situation where the price sensitivities of conference participants and spot market customers are the same, while the price sensitivity is lower in the final market periods. The case with $(\beta_s, \beta_c, \beta_f) = (1, 0.75, 0.5)$ represents a situation where conference participants are less sensitive compared to spot market customers, while the overall price sensitivity in the final market is lower.

Notice that with these demand models, the unconstraint maximizer of the three revenue functions are the same. In particular, $\arg\max_d r_c(d) = \arg\max_d r_s(d) = \arg\max_d r_f(d) = e^{-1} \triangleq d^0$. For an overall problem instance with capacity C and number of time periods T, we define the load factor as $l(C, T) = \frac{T\,d^0}{C}$, where the numerator is the total expected demand if we set the price to maximize the expected one period revenue in each of the markets at any given time period. The load factor measures the tightness of the capacity of the problem. We want to explore the performance of different policies under different load factors. We vary T among $\{80, 140, 200\}$. For $T = 80$, we vary C in $\{24, 20, 16, 12, 8\}$ so that the load factors are in the set

$L = \{1.23, 1.47, 1.84, 2.45, 3.68\}$. For $T = 140$, we vary C in $\{42, 35, 28, 21, 14\}$ and for $T = 200$ we vary C in $\{60, 50, 40, 30, 20\}$ so that the load factors are always in the same set L. Note that in this way we are scaling T and $C's$ linearly in the same rate, which corresponds to the asymptotic regime in Section 3.3.4, hence we can expect that both FP policy and SDA policy perform better as T increases from 80 to 200. In total we have $2 * 3 * 5 = 30$ problem instances.

Since the performance of different policies is contingent on the tactical decisions on conference market demand rate d_c and capacity allocation b, for each overall problem with a set of parameters, we compute the expected revenue of different policies at different (d_c, b) combinations. Since the unconstraint maximizer for $r_c(\cdot)$ is d^0, we know that the optimal demand rate in conference market must be within the interval $[0, d^0]$. We divide the interval into N equally spaced subintervals and vary d_c in $\{\frac{i}{N} d^0 : i = 0, 1, \ldots, N\}$, while b is varied in set $\mathcal{C} = \{0, 1, \ldots, C\}$. Thus we generate a set of $N(C + 1)$ operational level subproblems for each problem instance.

We report three performance statistics of FP and SDA policy for each set of subproblems under one overall problem: revenue gap compared to the OPT policy at the optimal (d_c, b) combination, which achieves the maximum $V_1^{d_c}(C - b, b)$ among the $N(C + 1)$ subproblems; maximum revenue gap compared to the OPT policy across all $N(C + 1)$ subproblems; and average revenue gap of the $N(C + 1)$ subproblems compared to the OPT policy. All the revenue gaps are reported as a percentage of the corresponding revenue achieved by OPT policy. The first statistic reflects the performance we can expect from different policies if we solve the tactical level decision problems to optimality and use the optimal conference pricing and capacity allocation; the second statistic reflects the performance of different policies in the worst case scenario if (d_c, b) are chosen randomly in their feasible region; while the third statistic reflects the average performance of different policies if (d_c, b) are chosen randomly in their feasible region. It turns out that the three key performance statistics do not change much when we increase N beyond 20, hence we set $N = 20$.

Numerical Results

The results of the numerical experiments are reported in Tables 3.1 and 3.2. Table 3.1 shows the results for problem instances with $(\beta_s, \beta_c, \beta_f) = (1, 1, 0.5)$, while Table 3.2 shows the results for problem instances with $(\beta_s, \beta_c, \beta_f) = (1, 0.75, 0.5)$. Each table consists of three portions: the top portion reports revenue gaps of FP and SDA policy compared to OPT policy at the optimal (d_c, b) combination according to $V_1^{d_c}(C - b, b)$; the middle portion reports maximum revenue gaps across the set of subproblems for each overall problem instance; the bottom portion reports the average revenue gaps across the subproblems for each overall problem instance. Within each portion, the structures of the table are the same. An overall problem instance is characterized by two parameters in each portion of the table: number of time periods T in the first column and load factor in the second row. The second column indicates which policy's revenue gaps we are reporting in the corresponding row, while the revenue gaps are reported from column three to seven and row three to eight. Finally the last row reports the mean improvements of SDA policy compared with FP policy across the three overall problem instances with the same load factor. All revenue gaps and mean improvements are reported as a percentage of the revenue obtained by OPT policy.

Table 3.1: Revenue gaps of FP and SDA policy compared to OPT policy with $(\beta_s, \beta_c, \beta_f) = (1, 1, 0.5)$.

gap at optimal		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FP	5.99%	5.52%	5.07%	4.89%	5.16%
	SDA	0.68%	0.67%	0.72%	0.70%	0.60%
140	FP	5.26%	4.89%	4.66%	4.65%	5.01%
	SDA	0.48%	0.58%	0.54%	0.56%	0.47%
200	FP	4.73%	4.46%	4.29%	4.39%	4.78%
	SDA	0.37%	0.47%	0.43%	0.43%	0.39%
mean impr.		4.82%	4.39%	4.11%	4.08%	4.50%
max gap		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FP	6.95%	8.10%	7.81%	10.33%	13.68%
	SDA	0.95%	1.17%	1.19%	1.02%	0.80%
140	FP	5.92%	6.69%	6.33%	9.63%	12.52%
	SDA	0.72%	0.95%	1.05%	0.91%	0.73%
200	FP	5.21%	5.81%	5.48%	9.08%	13.06%
	SDA	0.58%	0.80%	0.93%	0.82%	0.65%
mean impr.		5.28%	5.90%	5.49%	8.76%	12.36%
average gap		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FP	2.83%	4.43%	4.75%	5.02%	5.51%
	SDA	0.13%	0.20%	0.25%	0.25%	0.18%
140	FP	2.19%	3.66%	4.10%	4.49%	5.05%
	SDA	0.08%	0.13%	0.18%	0.18%	0.13%
200	FP	1.86%	3.20%	3.69%	4.12%	4.69%
	SDA	0.06%	0.10%	0.14%	0.14%	0.10%
mean impr.		2.20%	3.62%	3.99%	4.35%	4.95%

From Table 3.1 we can make several observations. First, when (d_c, b) are chosen to maximize $V_1^{d_c}(C-b, b)$, the performance of FP and SDA policy is not very sensitive to load factors. On average, SDA policy achieves 99.46% the optimal revenue, while it offers 4.38% improvement over FP policy. Second, when (d_c, b) are chosen randomly within their feasible regions, the maximum revenue gap of FP policy tends to increase with load factor, while it increases dramatically when the load factor is larger than 1.84. In contrast, the maximum revenue gap of SDA policy is not very sensitive to load factor, while it actually has a decreasing trend when the load factor is larger than 1.84. On average SDA policy achieves 99.12% the optimal revenue in the worst case, and provides 7.56% improvement over FP policy, with the largest improvement valued at 12.88%. Third, when (d_c, b) are chosen randomly within their feasible regions, the average revenue gap of FP policy increases with load factor, while the average revenue gap of SDA policy is not very sensitive to load factor. Over all the subproblems, SDA policy achieves 99.85% the optimal revenue, and gives 3.82% improvement over FP policy. Finally, for all the three statistics, we observe that given fixed load factor, the revenue gaps of both FP and SDA policy decrease as T increases. This is consistent with the asymptotic optimality result in Section 3.3.4. We have similar observations in Table 3.2 where the conference participants are less sensitive to price compared with spot market customers.

Compared to OPT policy, SDA policy captures more than 99% of the optimal revenue on average by solving a single dimensional dynamic program which only tracks the number of rooms in the spot market. This is very attractive from an operational point of view. Compared to FP policy, other than providing sizable improvements, we find that SDA policy is robust in the following two senses: it is sensitive neither to load factor, nor to higher level tactical decisions on conference market pricing and capacity allocation. This robustness could make SDA an attractive policy for hotel managers. In particular, as shown in Section 3.4 below, the tactical decision problem is a difficult optimization problem which lacks structural properties, and heuristics are usually used to compute a near optimal (d_c, b) combination. Even if we are able to find the true optimal conference market price and capacity allocation in the tactical level, this is done at the very beginning of the selling horizon, so the "optimal" tactical decisions are based on rough estimates of various parameters, which might be

Table 3.2: Revenue gaps of FP and SDA policy compared to OPT policy with $(\beta_s, \beta_c, \beta_f) = (1, 0.75, 0.5)$.

gap at optimal		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FP	5.44%	4.85%	4.52%	4.35%	4.70%
	SDA	0.62%	0.73%	0.78%	0.66%	0.61%
140	FP	4.78%	4.38%	4.22%	4.33%	4.59%
	SDA	0.44%	0.54%	0.61%	0.59%	0.49%
200	FP	4.30%	4.00%	3.97%	4.13%	4.40%
	SDA	0.34%	0.43%	0.48%	0.51%	0.39%
mean impr.		4.38%	3.84%	3.61%	3.68%	4.07%
max gap		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FP	6.45%	7.89%	7.81%	8.73%	11.55%
	SDA	0.85%	1.05%	1.04%	0.90%	0.70%
140	FP	5.42%	6.54%	6.33%	8.23%	10.28%
	SDA	0.65%	0.85%	0.93%	0.79%	0.64%
200	FP	4.77%	5.69%	5.48%	7.60%	10.61%
	SDA	0.53%	0.71%	0.82%	0.71%	0.57%
mean impr.		4.87%	5.84%	5.61%	7.38%	10.18%
average gap		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FP	2.62%	4.16%	4.46%	4.68%	5.11%
	SDA	0.12%	0.18%	0.23%	0.23%	0.16%
140	FP	2.02%	3.43%	3.84%	4.17%	4.66%
	SDA	0.07%	0.12%	0.16%	0.16%	0.11%
200	FP	1.71%	3.00%	3.45%	3.83%	4.32%
	SDA	0.05%	0.09%	0.12%	0.13%	0.09%
mean impr.		2.04%	3.40%	3.75%	4.05%	4.58%

changed later on in the selling horizon, and this will ruin the optimality of the original tactical decisions. However, due to practical constraints, the tactical decisions cannot be changed later on once made at the very beginning, hence it is important to have a policy that is robust to the tactical decisions.

3.4 Capacity Allocation and Conference Market Pricing

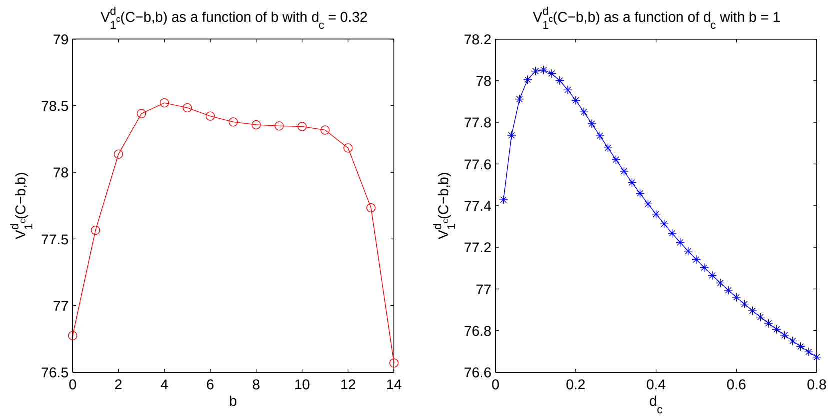
In this section, building on the formulation for the operational level dynamic pricing problem in the spot market, we study the overall problem with a focus on the tactical decisions about capacity allocation between spot and conference markets and finding optimal fixed price for conference market. While it turns out that the tactical decision problem lacks desirable structures, we construct a heuristic based on a deterministic formulation for a relaxed version of the overall problem, and show its asymptotically optimality when the capacity and number of time periods scale up linearly in the same rate. Finally, we construct a mixed policy by combining the heuristic for tactical decisions and SDA policy for operational level pricing decisions, and the mixed policy is also asymptotically optimal. The performance of different policies are evaluated through numerical experiments.

3.4.1 Relaxed Problem and Bounds

From the discussion following equation (3.2), we know that given the price in the conference market p_c (or the demand rate d_c) and the number of rooms reserved for conference market b, the total expected revenue in the selling horizon is $V_1^{d_c}(C-b,b)$. Hence the tactical decision problem can be written as

$$\begin{aligned} (\mathbf{HIGH})\ Z_h = \max \quad & V_1^{d_c}(C-b,b) \\ \text{subject to} \quad & 0 \le d_c \le 1 \\ & 0 \le b \le C, b \in \mathbb{Z}. \end{aligned} \tag{3.17}$$

Figure 3.1: Optimal revenue as a function of b and d_c for a problem instance with the following parameters: $T = 40, \tau = 20, C = 14, \lambda_c = \lambda_s = 0.5, d_s(p) = e^{-p}, d_c(p) = e^{-p}, d_f(p) = e^{-0.1p}$.



Unfortunately, $V_1^{d_c}(C-b, b)$ is not necessarily concave in either b or d_c. For a problem instance with exponential demand models, Figure 3.1 shows $V_1^{d_c}(C-b, b)$ as a function of b and d_c when the other parameter is fixed. It is clear from the figure that $V_1^{d_c}(C-b, b)$ is not concave in either component.

Due to the lack of structural properties, in order to solve problem (3.17) optimally, we essentially need to use discrete approximation to the feasible space of d_c, i.e., the interval $[0, 1]$. In particular, letting $\mathcal{S} = \{i/N : i = 0, 1, \ldots, N\}$, $\mathcal{C} = \{0, 1, \ldots, C\}$, for each $d_c \in \mathcal{S}$, we solve the two-dimensional dynamic program recursion (3.1) to obtain $V_1^{d_c}(C-b, b)$ for all values of b. There we can solve $\max_{b \in \mathcal{C}} V_1^{d_c}(C-b, b)$ to find the optimal capacity allocation for fixed value of d_c. Given a fixed N value, this approach takes $\mathcal{O}(N(\tau C^2 + (T-\tau)C))$ time. We can increase N to get more accurate approximation to the optimal value, with an increased computational complexity. This approach is computationally inefficient, and might be difficult for hotel managers to adopt. We next develop a relaxation of the original problem and its deterministic formulation, which provides a simple heuristic for problem (3.17).

In the original problem, the price in the conference market is fixed while dynamic pricing can be used in the spot market. We relax this constraint to allow dynamic pricing in both markets. For this relaxed problem, the following dynamic programming recursion maximizes the expected revenue from time period t onwards if the number of rooms left in spot and conference market are x and y respectively. For $t = 1, \dots, \tau$ and $(x, y) \in \mathcal{D}$, we solve

$$J_t(x,y) = \max_{d_{st}, d_{ct} \in [0,1]} \Big\{ \lambda_{st}[r_s(d_{st}) + d_{st} J_{t+1}(x-1, y)] + \lambda_{ct}[r_c(d_{ct}) + d_{ct} J_{t+1}(x, y-1)] \\ + (1 - \lambda_{st}\, d_{st} - \lambda_{ct}\, d_{ct}) J_{t+1}(x,y) \Big\}, \quad (3.18)$$

with boundary condition $J_t(0,0) = 0, J_{\tau+1}(x,y) = \Psi_{\tau+1}(x+y)$. For any given d_c, if we add the constraint $d_{ct'} = d_c, \forall t' = t, \dots, \tau$, then the recursion (3.18) will coincide with (3.1). Hence we have $V_t^{d_c}(x,y) \leq J_t(x,y), \forall d_c \in [0,1]$. Intuitively, this is saying that problem (3.18) is indeed a relaxation of the original problem. For this relaxed problem, the tactical level decision is to choose the capacity allocation parameter b in order to achieve maximum revenue, i.e., $\max_{b \in \mathcal{C}} J_1(C - b, b)$.

Consider a deterministic version of the relaxed problem: The hotel has C rooms available in total at time period 1, which is a *continuous* quantity. The hotel can control the capacity allocation b and the prices in both spot and conference market. For $t = 1, \dots, \tau$, if demand rates in spot and conference market are d_{st} and d_{ct}, then the realized demands during time period t in the two markets are $\lambda_{st} d_{st}$ and $\lambda_{ct} d_{ct}$ correspondingly. Given the number of rooms reserved for conference market b, the demand realized from time period 1 to τ can not exceed b in the conference market, while it can not exceed $C - b$ in the spot market. Starting from time period $\tau + 1$, the two markets join together as a single market and share the remaining capacity. For $t = \tau + 1, \ldots, T$, if the demand rate is d_{ft}, then the realized demand is d_{ft} exactly. The hotel wants to maximize its total revenue by controlling capacity allocation b, the demand rates in both markets from time period 1 to τ, and the demand rates from time period $\tau + 1$ to T. The following nonlinear program solves the revenue maximization problem:

$$
\begin{aligned}
(\textbf{NLP3})\ Z_3 = \max \quad & \sum_{t=1}^{\tau} \lambda_{ct}\, r_c(d_{ct}) + \sum_{t=1}^{\tau} \lambda_{st}\, r_s(d_{st}) + \sum_{t=\tau+1}^{T} r_f(d_{ft}) && (3.19)\\
\text{subject to} \quad & \sum_{t=1}^{\tau} \lambda_{ct}\, d_{ct} && \leq b \\
& \sum_{t=0}^{\tau} \lambda_{st}\, d_{st} && \leq C - b \\
& \sum_{t=1}^{\tau} \lambda_{ct}\, d_{ct} + \sum_{t=1}^{\tau} \lambda_{st}\, d_{st} + \sum_{t=\tau+1}^{T} d_{ft} && \leq C \\
& 0 \leq d_{ct}, d_{st} \leq 1, \quad \forall t = 1, \ldots, \tau \\
& 0 \leq d_{ft} \leq 1, \qquad \forall t = \tau + 1, \ldots, T \\
& 0 \leq b \leq C.
\end{aligned}
$$

Using very similar arguments as in the proof of Proposition 3.3.3, we can show that the above nonlinear program provides an upper bound for the total expected revenue of the relaxed problem, hence it is also an upper bound for the total expected revenue of the original problem.

Proposition 3.4.1. *For any $b \in \mathcal{C}$ and $d_c \in [0, 1]$, we have $V_1^{d_c}(C - b, b) \leq J_1(C - b, b) \leq Z_3$. In particular $Z_h \leq Z_3$.*

We skip the detail of the proof because it is very similar to the proof of Proposition 3.3.3. We make a note here that it is difficult to show $V_1^{d_c}(C - b, b) \leq Z_3$ directly without introducing the relaxed problem (3.18). Noting that the constraints in (NLP3) are linear while the objective function is concave, we can show the following property.

Proposition 3.4.2. *There exists an optimal solution* $(\bar{b}, \bar{d}_{st}, \bar{d}_{ct}, \bar{d}_{ft})$ *for (NLP3), which satisfies* $\bar{d}_{st} = \bar{d}_s, \bar{d}_{ct} = \bar{d}_c, t = 1, \ldots, \tau$, *and* $\bar{d}_{ft} = \bar{d}_f, t = \tau + 1, \ldots, T$.

Let $(\bar{p}_s, \bar{p}_c, \bar{p}_f)$ be the optimal prices corresponding to $(\bar{d}_s, \bar{d}_c, \bar{d}_f)$. We construct a simple Fixed Price with Capacity Allocation (FPCA) policy for the original problem based on the optimal solution of (NLP3) as follows: From time period 1 to τ, we use a fixed price $\bar{p}_c$ in the conference market and a fixed price $\bar{p}_s$ in the spot market; from time period $\tau+1$ to T, we use a fixed price $\bar{p}_f$. For the capacity allocation, we choose $\bar{b}' = \arg\max_b \left\{ \bar{W}(b) : b \in \{\lfloor \bar{b} \rfloor, \lceil \bar{b} \rceil\} \right\}$, where $\bar{W}(b)$ is the expected revenue achieved by the fixed prices $(\bar{p}_s, \bar{p}_c, \bar{p}_f)$ given capacity allocation b. In particular, we have

$$\begin{aligned} \bar{W}(b) = {} & \bar{p}_c \, \mathbb{E} \min\{b, D_c(\bar{d}_c)\} + \bar{p}_s \, \mathbb{E} \min\{C - b, D_s(\bar{d}_s)\} \\ & + \bar{p}_f \, \mathbb{E} \min\{(b - D_c(\bar{d}_c))^+ + (C - b - D_s(\bar{d}_s))^+, D_f(\bar{d}_f)\}, \end{aligned} \tag{3.20}$$

where $D_c(\bar{d}_c)$ is the total demand in conference market from time period 1 to τ, which follows a Poisson binomial distribution with parameters $(\lambda_{c1}\bar{d}_c, \lambda_{c2}\bar{d}_c, \ldots, \lambda_{c\tau}\bar{d}_c)$ (or a Binomial distribution with parameters $(\tau, \lambda_c\bar{d}_c)$ with stationary arrival probabilities); $D_s(\bar{d}_s)$ is the total demand in spot market from time period 1 to τ, which follows a Poisson binomial distribution with parameters $(\lambda_{s1}\bar{d}_s, \lambda_{s2}\bar{d}_s, \ldots, \lambda_{s\tau}\bar{d}_s)$ (or a Binomial distribution with parameters $(\tau, \lambda_s\bar{d}_s)$ with stationary arrival probabilities); and $D_f(\bar{d}_f)$ is the total demand from time period $\tau + 1$ to T, which follows a binomial distribution with parameters $(T - \tau, \bar{d}_f)$.

The following property holds for $\bar{W}(b)$. The proof is deferred to Appendix B.4.

Lemma 3.4.3. *$\bar{W}(b)$ is a piecewise linear function for $0 \leq b \leq C$, whose points of non-differentiability are integers in $\mathcal{C}$.*

An immediate result from this lemma is that $\bar{W}(\bar{b}') \geq \bar{W}(\bar{b})$. Hence we have the following chain of inequalities:

$$1 \geq \frac{\bar{W}(\bar{b}')}{Z_h} \geq \frac{\bar{W}(\bar{b})}{Z_h} \geq \frac{\bar{W}(\bar{b})}{Z_3}. \tag{3.21}$$

The first inequality holds since FPCA is a feasible policy for the original problem, while the third inequality holds due to Proposition 3.4.1.

3.4.2 Asymptotic analysis

Similar to the asymptotic analysis for the operational level spot market pricing problem in Section 3.3.4, we consider a series of problems $\mathbb{P}^m$ indexed by $m = 1, 2, \ldots$, where the problem we study above in this section can be written as $\mathbb{P}^1$. Notice that $\mathcal{P}^m$ is a operational level problem given certain exogenous conference market price p_c and capacity allocation $m\,b$, while $\mathbb{P}^m$ is an overall problem where the conference market price and capacity allocation are inherent decision variables. Problem $\mathbb{P}^m$ has the same parameters as problem $\mathcal{P}^m$ except that $\mathbb{P}^m$ does not have fixed conference market price or capacity allocation.

For problem $\mathbb{P}^m$, we let Z_h^m be the optimal expected revenue, Z_3^m be the deterministic upper bound from (NLP3), and $\bar{b}^m$ be the optimal capacity allocation from (NLP3). Since b in (NLP3) is treated as a continuous variable, the objective function and optimal capacity allocation for $\mathbb{P}^m$ scale linearly with m, while the optimal demand rates $(\bar{d}_c, \bar{d}_s, \bar{d}_f)$ do not scale with m. Let $\bar{W}^m(\cdot)$ be the total expected revenue achieved by the FPCA policy,

$\bar{b}'^m = \arg\max_b \left\{\bar{W}^m(b) : b \in \{\lfloor \bar{b}^m \rfloor, \lceil \bar{b}^m \rceil\}\right\}$, and $D_c^m(\bar{d}_c), D_s^m(\bar{d}_s), D_f^m(\bar{d}_f)$ be the demands in the conference, spot and final markets by the FPCA policy. Following an argument similar to that in the proof of Proposition 3.3.5, we have

$$
\begin{aligned}
\bar{W}^m(\bar{b}'^m) \geq & \sum_{t=1}^{m\tau} \lambda_{ct}^m \, r_c(\bar{d}_c) - \frac{1}{2}\sigma(D_c^m(\bar{d}_c)) + \sum_{t=1}^{m\tau} \lambda_{st}^m \, r_s(\bar{d}_s) - \frac{1}{2}\sigma(D_s^m(\bar{d}_s)) \\
& + \sum_{t=m\tau+1}^{mT} r_f(\bar{d}_f) - \frac{1}{2}\sigma(D_c^m(\bar{d}_c)) - \frac{1}{2}\sigma(D_s^m(\bar{d}_s)) - \frac{1}{2}\sigma(D_f^m(\bar{d}_f)) \\
= & Z_3^m - \sigma(D_c^m(\bar{d}_c)) - \sigma(D_s^m(\bar{d}_s)) - \frac{1}{2}\sigma(D_f^m(\bar{d}_f)). \qquad (3.22)
\end{aligned}
$$

We have $Z_3^m = m\, Z_3, \bar{b}^m = m\, \bar{b}$. Furthermore, the standard deviations of $D_c^m(\bar{d}_c), D_s^m(\bar{d}_s)$ and $D_f^m(\bar{d}_f)$ all scale linearly with $\sqrt{m}$, i.e., there exists a constant G where $\sigma(D_c^m(\bar{d}_c)) + \sigma(D_s^m(\bar{d}_s)) + \frac{1}{2}\sigma(D_f^m(\bar{d}_f)) = G\sqrt{m}$. Thus, we can continue the chain of inequalities (3.21) for problem $\mathbb{P}^m$ as follows:

$$
\begin{aligned}
1 \geq \frac{\bar{W}^m(\bar{b}'^m)}{Z_h^m} \geq \frac{\bar{W}^m(m\,\bar{b})}{Z_3^m} \geq \frac{Z_3^m - G\sqrt{m}}{Z_3^m} \\
= \frac{m\, Z_3 - G\sqrt{m}}{m\, Z_3} = 1 - \frac{G}{Z_3\sqrt{m}} \xrightarrow{m\to\infty} 1, \quad (3.23)
\end{aligned}
$$

where the third inequality follows from (3.22). Hence we have shown the following result.

Proposition 3.4.4. *FPCA policy is asymptotically optimal for the overall problem when the capacity and number of time periods scale up linearly in the same rate.*

We can construct a mixed policy by combing FPCA and the SDA policy for spot market pricing in the following way: We solve (NLP3) to get the fixed price in conference market $\bar{p}_c$ and capacity allocation $\bar{b}$, from which we derive $\bar{b}'$. We choose $(\bar{p}_c, \bar{b}')$ as the tactical level decisions. Then we adopt SDA policy to do dynamic pricing in the spot market for fixed values of $\bar{p}_c$ and $\bar{b}'$. We call this policy the Mixed policy (MXD). Since the fixed prices of FPCA in the spot and final markets are feasible for the dynamic programming recursion (3.3) used by SDA, we get that the MXD policy performs as least as good as FPCA. Thus we have the following corollary.

Corollary 3.4.5. *MXD policy performs at least as good as FPCA policy, hence it is also asymptotically optimal.*

3.4.3 Numerical Experiments

In this section, we evaluate the performance of FPCA and MXD policy for the overall problem through numerical experiments. We use the same experimental setup and overall problem instances as in Section 3.3.5. For each problem instance, we report the revenue gaps of the two policies compared to OPT policy. Note that FPCA and MXD use the same (d_c, b) combination from (NLP3), which may not be the same as the one used by OPT policy. The numerical results are provided in Table 3.3. The top portion of the table contains results for problem instances with $(\beta_s, \beta_c, \beta_f) = (1, 1, 0.5)$, while the bottom portion contains results for problem instances with $(\beta_s, \beta_c, \beta_f) = (1, 0.75, 0.5)$.

The observation from Table 3.3 is very similar to that for the top portion of Table 3.1 and 3.2 in Section 3.3.5, where we were comparing FP and SDA policy at the optimal (d_c, b) combination used by OPT policy. In particular, MXD achieves more than 99% of the optimal revenue, while offering 4.12% improvement over FPCA policy on average. This means that by solving (NLP3), we obtain a very good approximation to the optimal (d_c, b) combination, which is difficult to get through the exact formulation in (3.17). Also, we still observe that given fixed load factor, the revenue gaps of FPCA and MXD policy decrease as

Table 3.3: Revenue gaps of FPCA and MXD policy compared to OPT policy.

$(\beta_s, \beta_c, \beta_f) = (1, 1, 0.5)$						
gaps		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FPCA	6.32%	5.63%	5.21%	5.00%	5.13%
	MXD	1.06%	1.01%	0.96%	0.82%	0.60%
140	FPCA	5.46%	5.01%	4.77%	4.75%	4.92%
	MXD	0.70%	0.71%	0.67%	0.64%	0.41%
200	FPCA	4.84%	4.55%	4.37%	4.47%	4.70%
	MXD	0.51%	0.54%	0.49%	0.49%	0.33%
mean impr.		4.78%	4.31%	4.08%	4.09%	4.47%

$(\beta_s, \beta_c, \beta_f) = (1, 0.75, 0.5)$						
gaps		load factor				
T	Policy	1.23	1.47	1.84	2.45	3.68
80	FPCA	5.63%	5.02%	4.68%	4.74%	4.84%
	MXD	0.99%	0.98%	0.96%	1.03%	0.66%
140	FPCA	4.91%	4.55%	4.33%	4.42%	4.69%
	MXD	0.67%	0.71%	0.67%	0.70%	0.50%
200	FPCA	4.39%	4.13%	4.03%	4.15%	4.51%
	MXD	0.50%	0.57%	0.57%	0.59%	0.44%
mean impr.		4.26%	3.81%	3.61%	3.66%	4.15%

the number of time periods increases, which is consistent with the asymptotic optimality results in Section 3.4.2.

3.5 Conclusions and Extensions

We study the pricing and capacity allocation problem in a dual channel hotel revenue management setting. When a conference is to be held at a location near the hotel, the conference organizer often requests to reserve a block of rooms for conference participants at a fixed price, and the reservation expires if the reserved rooms are not fully booked by a deadline before the date of conference. This adds a conference market channel parallel to the spot market, where dynamic pricing practice is widely adopted. The hotel manager faces several decision in this setting. At a tactical level, the hotel manager need to decide the number of rooms to reserve and the fixed price to offer for conference participants. At an operational level, given the tactical decisions, the dynamic pricing policy in the spot market needs to be optimized considering the conference market channel.

For the operational level problem on spot market pricing, we construct a single dimensional approximation to the exact two dimensional dynamic programming formulation, by modifying the boundary condition of a classical single dimensional dynamic pricing recursion to incorporate the possible remaining capacity from the conference market. We show that the approximation is asymptotically tight when the capacity and number of time periods scale up linearly in the same rate. The novel part of the analysis is that the natural deterministic formulation of the operational level problem does not provide a scalable upper bound. We utilize an scalable approximation to the natural deterministic formulation, which provides a an upper bound with a small perturbation. The SDA policy from the approximation is not only appealing from an operational perspective, it also shows robust performance in numerical experiments with respect to load factor and tactical level decision inputs.

For the tactical problem in the higher level, based on a deterministic formulation of a relaxed problem, we develop a heuristic to make decisions on conference market pricing and capacity allocation simultaneously. For the overall problem, we construct a MXD policy by combining the tactical level heuristic and the operational level SDA policy, and show that the MXD policy is asymptotically optimal. Numerical experiments indicate that the MXD policy provides promising performance.

There are several possible extensions that are worth future investigation. First, the asymptotic analysis in Section 3.3.4 and 3.4.2 is based on the assumption that the demand model in each of the three markets is time invariant. Under this assumption, the deterministic formulations suggest fixed prices in each market, which we can analyze conveniently. An important extension is to study whether the results in this chapter hold under time dependent demand models. By taking advantage of properties of Poisson Process, Gallego and Van Ryzin (1997) show that a pricing policy derived from the deterministic formulation of the stochastic problem is asymptotically optimal under a continuous time dynamic pricing model. This makes us believe that our results in this chapter would hold with time dependent demand models under continuous time framework. However, it is not clear whether this is true under discrete time framework as we used in this chapter. We leave this as an open problem for future research.

Second, in this chapter we focus on the problem with single-night stay. It is a natural extension to consider the problem with multiple-night stay. The multiple-night stay problem is significantly more complicated and usually studied through heuristics. Since we can set different prices for different nights in the spot market, by solving each night's problem individually, our MXD policy provides a heuristic to solve the multiple-night stay problem, although it ignores the interdependence of demands for different nights. A possible modification is to add some penalty if the capacity of one night runs out, so that the capacities of different nights can be depleted in a more or less synchronized fashion.

Third, in this chapter, we assume that conference market and spot market are separated perfectly. In particular, the demand function in each market only depends on the price offered in its own market, and all future conference market demands are lost once the reserved rooms for the conference are fully booked. It is of interest to relax this assumption to incorporate the behavior of conference participants into the model. For instance, conference participants may compare the current price in the spot market with the fixed price in the conference market and choose the one with the smaller price. This complicates the problem in the following way. In our current model, given the conference market price and capacity allocation, the revenue in the conference market is not affected by the pricing policy in the spot market. However this will not be true if the customer behavior is incorporated in the model, since the demand rate in the conference market is affected by the price in the spot market. However the idea behind the SDA policy might be still very appealing, which is to approximate the original value function by a single dimensional dynamic program with modified boundary conditions.

CHAPTER 4

REVENUE MANAGEMENT IN MULTIPLE MARKETS: PRICING FLEXIBILITY VERSUS CAPACITY FLEXIBILITY

4.1 Introduction

A firm can sell a product in different markets, where customers in the same market share similar characteristics so that their demand can be modeled by a common function of the offered price. One important strategic decision for the firm is whether it should adopt price discrimination by charging different prices to customers in different markets. While it is believed that price discrimination benefits the seller in a monopoly setting in general, it is not necessarily true if the seller is required to allocate its capacity to different markets at the beginning of the selling horizon.

More specifically, the firm has two options to set up its operations. The first option is to allocate its capacity to different markets and charge different prices for the inventories allocated to different markets. In other words, the firm obtains pricing flexibility in this setting. However, the demand in each market can only be satisfied by the allocated inventory to its own market. If the firm runs out of inventory in a market, then all future demands in this market will be lost. In other words the firms loses capacity flexibility in this setting. We call this option with pricing flexibility but no capacity flexibility the "Separate Pricing" setting.

The second option is to sell the product in a single market to all the customers. In contrast to the Separate Pricing setting, the firm needs to offer a single common price, while capacity allocation is not needed and all of its capacity can be used to satisfy the demand from any customers. Hence, the firm loses pricing flexibility while obtains capacity flexibility in this setting. We call this option the "Joint Pricing" setting.

In this chapter, using consistent demand models, we study the trade off between pricing flexibility and capacity flexibility based on dynamic programming formulations. Under general conditions, we show that when the capacity and number of time periods in the selling horizon scale up linearly in the same rate, the benefit of pricing flexibility eventually outweighs the cost of capacity flexibility. This result motivates a hybrid model where we adopt Separate Pricing at the early stage and switch to Joint Pricing towards the end of the selling horizon. The exact dynamic programming formulation uses a multi-dimensional vector to keep track of the number of remaining inventories in every markets, hence is intractable to solve due to the high-dimensional state space. We develop heuristics using a single dimensional approximation for each market, whose boundary condition is based on a deterministic formulation. Computational experiments indicate that the heuristics perform extremely well compared to other policies when the capacity is tight, the number of markets is large, or the price sensitivities of customers in different markets are widely dispersed.

The rest of the chapter is organized as follows. An overview of the related literature is provided in Section 4.2. In Section 4.3, building on dynamic programming formulations for Separate Pricing and Joint Pricing problems, we show under general conditions that Separate Pricing is better when the capacity and the length of selling horizon scale up. We also construct examples under which Joint Pricing achieves larger expected revenue compared to Separate Pricing. Thus, neither Separate Pricing nor Joint Pricing is uniformly superior to the other. In Section 4.4, we study a hybrid model, where Separate Pricing is adopted from the beginning of the selling horizon to some deadline, after which Joint Pricing is used. While the hybrid model is intractable due to the high-dimensional state space, we construct asymptotically optimal policies using single dimensional approximations. In Section 4.5, we conduct numerical experiments to evaluate the performance of different policies, while conclusion and future research directions are presented in Section 4.6.

4.2 Literature Review

There are two streams of literature that are related to our work. The first stream of literature is the study of pricing discrimination versus uniform pricing in economics literature, where deterministic demand models are adopted widely. When a monopolist serves two independent markets with linear demands, Robinson (1933) shows that compared to uniform pricing, third-degree price discrimination leaves total output unchanged and therefore reduces social welfare if discriminatory prices are different. Schmalensee (1981) shows that for a monopolist with constant marginal cost facing independent demands, third degree price discrimination raises social welfare only if it increases total output. Malueg (1993) provides bounds on the ratio of social welfare achieved by pricing discrimination and uniform pricing given particular restrictions on market demands. Malueg and Snyder (2006) develop bounds of the ratio of a monopolist's profit achieved by pricing discrimination and uniform pricing. All of these models use deterministic demand models and do not have a capacity constraint.

Our problem can be thought as a stochastic version of the problem with fixed capacity constraint and dynamic demand revealed over time.

The second stream of work involves dynamic pricing in revenue management literature. Gallego and Van Ryzin (1994) study the optimal dynamic pricing policy for a single product over a finite horizon with a continuous time Poisson demand model. They characterize the form of the optimal policy and show that a fixed price policy based on a deterministic version of the problem is asymptotically optimal as the volume of expected sales tends to infinity. The deterministic bounds in this chapter resemble theirs. Gallego and Van Ryzin (1997) extend Gallego and Van Ryzin (1994) to a more general setting that allows time-dependent demand models and multiple products assembled from multiple resources according to a network structure. Zhao and Zheng (2000) extend Gallego and Van Ryzin (1994) to allow nonhomogeneous demand processes and identify sufficient conditions under which the optimal price decreases over time for a given inventory level. Assuming a markup or markdown strategy, Feng and Xiao (2000b) study the dynamic pricing problem when the price has to be chosen from a predetermined discrete set. They characterize the structural properties of the value function and optimal policy, and provide an exact solution for the continuous-time model. Feng and Xiao (2000a) extend the model to allow price changes in both increasing and decreasing directions. Maglaras and Meissner (2006) show that multi-product dynamic pricing and capacity allocation problems can be modeled in a common framework, and develop asymptotically optimal policies through fluid approximations. Zhang and Cooper (2009) consider the problem of pricing parallel flights that are substitutable with each other. They build upper and lower bounds on the value functions and use these bounds to construct heuristic policies. Erdelyi and Topaloglu (2010b) propose dynamic programming decomposition methods to solve pricing problem in network revenue management. Recent papers on multi-product dynamic pricing such as Dong et al. (2009) and Akçay et al. (2010) incorporate customer choice into the pricing model where customers choose from a set of substitutable products according to certain utility maximization rule. Extensive overviews of pricing models can be found in papers by McGill and Van Ryzin (1999), Bitran and Caldentey (2003) and Elmaghraby and Keskinocak (2003), and the book by Talluri and Van Ryzin (2005).

4.3 Separate Pricing vs. Joint Pricing

A firm sells a product to customers in n different markets within a finite selling horizon consisting of time periods $1, \dots, T$. At the beginning of the selling horizon, the total capacity available is C units. For $t = 1, \dots, T$, there is exactly one customer arriving during time period t. In particular, we let λ_i be the probability that an arriving customer is from market i in a time period, where $\sum_{i=1}^{n} \lambda_i = 1$. Time invariant arrival probabilities are assumed only for the sake of notational simplicity, while all the results in this chapter hold with time dependent arrival probabilities. Customers in different markets have different sensitivities to the offered price. More specifically, for market i with demand function $d_i(\cdot)$, $i = 1, \dots, n$, given the offered price p_i, the probability that a market i customer makes a purchase is $d_i(p_i)$. The inverse demand function is $p_i(\cdot)$, which maps the demand rate back to the offered price. We use demand rates as the decision variables instead of the offered prices. Given demand rate d_i, the revenue rate is given by $r_i(d_i) = d_i \, p_i(d_i)$. In accordance with dynamic pricing literature, we assume that the revenue functions $r_i(\cdot), i = 1, \dots, n$ are continuous, bounded and concave, and satisfy $\lim_{d_i \to 0} r_i(d_i) = 0$.

If the firms chooses the Separate Pricing setting, it needs to choose a capacity allocation characterized by $\boldsymbol{b} = (b_1, \ldots, b_n)$ with $\sum_{i=1}^{n} b_i = C$ at the beginning of the selling horizon, and future demands in market i will be lost if b_i units have already been sold in this market. For market i, the firm adopts a dynamic pricing strategy. At time period $t = 1, \ldots, T$ with remaining capacity $x_i \leq b_i$, the total expected revenue onwards in market i can be calculated by the following dynamic programming recursion:

$$\begin{aligned} H_{it}(x_i) &= \max_{d_{it} \in [0,1]} \Big\{ \lambda_i [r_i(d_{it}) + d_{it}\, H_{i,t+1}(x_i - 1)] + (1 - \lambda_i\, d_{it}) H_{i,t+1}(x_i) \Big\} \\ &= \max_{d_{it} \in [0,1]} \Big\{ \lambda_i [r_i(d_{it}) - d_{it}\, \Delta H_{i,t+1}(x_i)] \Big\} + H_{i,t+1}(x_i), \end{aligned} \tag{4.1}$$

where $\Delta H_{i,t+1}(x_i) = H_{i,t+1}(x_i) - H_{i,t+1}(x_i - 1)$. The boundary condition is $H_{it}(0) = 0, H_{i,T+1}(\cdot) = 0$. If a demand rate d_{it} is chosen at time t, it means that we offer price $p_i(d_i t)$ at this time period. The total expected revenue the firm can collect in this Separate Pricing setting with capacity allocation $\boldsymbol{b}$ is $\sum_{i=1}^{n} H_{i1}(b_i)$. Letting R_{sep} be the total expected revenue obtained from Separate Pricing, and $\mathcal{B} = \{\boldsymbol{b} \in \mathbb{Z}^n : 0 \leq b_i \leq C, \sum_{i=1}^{n} b_i = C\}$, we have

$$R_{sep} = \max_{\boldsymbol{b} \in \mathcal{B}} \sum_{i=1}^{n} H_{i1}(b_i). \tag{4.2}$$

If the firm chooses the Joint Pricing setting, a common price needs to be offered at any time period. Here we still use the demand rate for each market as the decision variable, hence the demand rates $d_{it}, i = 1, \ldots, n$ need to satisfy the condition that $p_1(d_{1t}) = p_2(d_{2t}) = \ldots = p_n(d_{nt})$. Letting $\mathcal{S} = \{\boldsymbol{d} \in [0,1]^n : p_1(d_1) = p_2(d_2) = \ldots = p_n(d_n)\}$, at time period period $t = 1, \ldots, T$ with remaining capacity $y \leq C$, the following dynamic programming recursion maximizes the total expected revenue onwards:

$$\Psi_t(y) = \max_{\boldsymbol{d}_t \in \mathcal{S}} \Big\{ \sum_{i=1}^{n} \lambda_i r_i(d_{it}) + \sum_{i=1}^{n} \lambda_i d_{it}\, \Psi_{t+1}(y-1) + (1 - \sum_{i=1}^{n} \lambda_i\, d_{it}) \Psi_{t+1}(y) \Big\}$$
$$= \max_{\boldsymbol{d}_t \in \mathcal{S}} \Big\{ \sum_{i=1}^{n} \lambda_i r_i(d_{it}) - \sum_{i=1}^{n} \lambda_i d_{it}\, \Delta\Psi_{t+1}(y) \Big\} + \Psi_{t+1}(y), \tag{4.3}$$

where $\Delta\Psi_{t+1}(y) = \Psi_{t+1}(y) - \Psi_{t+1}(y-1)$. The boundary condition is $\Psi_t(0) = 0, \Psi_{T+1}(\cdot) = 0$. Letting R_{joint} be the total expected revenue from the Joint Pricing setting, we have $R_{joint} = \Psi_1(C)$.

We want to know whether $R_{sep} \geq R_{joint}$ or $R_{sep} < R_{joint}$, and under which conditions each of the two inequalities holds. First, we show that the following two nonlinear programs provide up bounds for R_{sep} and R_{joint} respectively.

$$(\mathbf{NLP1})\ Z_{sep} = \max \quad \sum_{t=1}^{T} \sum_{i=1}^{n} \lambda_i\, r_i(d_{it}) \tag{4.4}$$
$$\text{subject to} \quad \sum_{t=1}^{T} \lambda_i\, d_{it} \leq b_i, \quad i = 1, \ldots, n \tag{4.5}$$
$$\sum_{i=1}^{n} b_i = C \tag{4.6}$$
$$0 \leq b_i \leq C, \qquad i = 1, \ldots, n \tag{4.7}$$
$$\boldsymbol{d}_t \in [0,1]^n, \qquad t = 1, \ldots, T, \tag{4.8}$$

$$(\mathbf{NLP2})\ Z_{joint} = \max \quad \sum_{t=1}^{T} \sum_{i=1}^{n} \lambda_i\, r_i(d_{it}) \tag{4.9}$$
$$\text{subject to} \quad \sum_{t=1}^{T} \sum_{i=1}^{n} \lambda_i\, d_{it} \leq C \tag{4.10}$$
$$\boldsymbol{d}_t \in \mathcal{S}, \qquad t = 1, \ldots, T. \tag{4.11}$$

Notice that (NLP1) and (NLP2) share the same objective function. The constraints (4.5)-(4.7) in (NLP1) indicate that a capacity allocation is needed, while the constraint (4.8) implies that we are allowed to set different prices in different markets. In contrast, constraint (4.10) in (NLP2) indicates that there is no need to do capacity allocation in this setting, while constraint (4.11) implies that a common price needs to be set across different markets. We have the following result regarding to (NLP1) and (NLP2).

Proposition 4.3.1. *$R_{sep} \le Z_{sep}$, $R_{joint} \le Z_{joint}$ and $Z_{joint} \le Z_{sep}$.*

Proof. We skip the detail of proof for $R_{sep} \le Z_{sep}$ and $R_{joint} \le Z_{joint}$, since the approach is very similar to the proof of Proposition 4.4.1 in Section 4.4.2. The idea is to construct a feasible solution for the nonlinear program based on the optimal pricing policy, whose objective value is the same as the expected revenue of the optimal policy. For $Z_{joint} \le Z_{sep}$, we note that a feasible solution for (NLP1) can be constructed using the optimal solution of (NLP2). In particular, if $\boldsymbol{d}_t^2, t = 1, \ldots, T$ is the optimal solution for (NLP2), then we let $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_n)$ where $\eta_i = \frac{\sum_{t=1}^T \lambda_i d_{it}^2}{\sum_{t=1}^T \sum_{j=1}^n \lambda_{jt} d_{jt}^2}, i = 1, \ldots, n$. If we set $\boldsymbol{b}^1 = \boldsymbol{\eta} C$ and $\boldsymbol{d}_t^1 = \boldsymbol{d}_t^2, t = 1, \ldots, T$, it is easy to verify that $(\boldsymbol{d}_t^1, \boldsymbol{b}^1)$ is a feasible solution for (NLP1) with the same objective value. Hence we have $Z_{joint} \le Z_{sep}$. □

Next, using the optimal solution of (NLP2), we construct a pair of nonlinear programs with the same objective function, but the constraint in one is tighter than the constraint in the other. This pair of nonlinear programs are used to connect R_{sep} and R_{joint}. We first state an useful fact here which can be derived from a similar result shown in Gallego (1992).

Fact (I): For a deterministic scalar z and a real-valued random variable Z with finite mean μ and finite variance σ^2, we have $\mathbb{E} \min\{Z, z\} \ge \min\{\mu, z\} - \sigma/2$.

As in the proof of Proposition 4.3.1, we let $\{\boldsymbol{d}_t^2 : t = 1, \ldots, T\}$ be the optimal solution for (NLP2), and $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_n)$ where $\eta_i = \frac{\sum_{t=1}^T \lambda_i d_{it}^2}{\sum_{t=1}^T \sum_{j=1}^n \lambda_{jt} d_{jt}^2}, i = 1, \ldots, n$. Now consider the following two nonlinear programs:

$$
\begin{aligned}
(\mathbf{NLP3})\; Z_3 = \max \quad & \sum_{t=1}^{T} \sum_{i=1}^{n} \lambda_i\, r_i(d_{it}) \\
\text{subject to} \quad & \sum_{t=1}^{T} \lambda_i\, d_{it} \leq \eta_i\, C,\; i = 1, \ldots, n && (4.12) \\
& \boldsymbol{d}_t \in [0,1]^n, \qquad t = 1, \ldots, T, && (4.13)
\end{aligned}
$$

$$
\begin{aligned}
(\mathbf{NLP4})\; Z_4 = \max \quad & \sum_{t=1}^{T} \sum_{i=1}^{n} \lambda_i\, r_i(d_{it}) \\
\text{subject to} \quad & \sum_{t=1}^{T} \lambda_i\, d_{it} \leq \eta_i\, C,\; i = 1, \ldots, n \\
& \boldsymbol{d}_t \in \mathcal{S}, \qquad t = 1, \ldots, T. && (4.14)
\end{aligned}
$$

According to the definition of $\boldsymbol{\eta}$, we see that (NLP4) is a rewrite of (NLP2), hence $Z_{joint} = Z_4$. Besides, (NLP3) and (NLP4) are the same except that (NLP4) uses the tighter constraint (4.14) while (NLP3) uses the looser constraint (4.13). Hence we have the following chain of inequalities:

$$R_{joint} \leq Z_{joint} = Z_4 \leq Z_3. \tag{4.15}$$

Due to the concavity assumption of $r_i(\cdot)$, we know that there exists an optimal solution for (NLP3) $\{\bar{\boldsymbol{d}}_t : t = 1, \ldots, T\}$ which satisfies $\bar{\boldsymbol{d}}_1 = \bar{\boldsymbol{d}}_2 = \ldots = \bar{\boldsymbol{d}}_T \doteq \bar{\boldsymbol{d}}$. Besides, it is easy to verify that $(\bar{\boldsymbol{d}}, \boldsymbol{\eta}\, C)$ is a feasible solution for (NLP1) with an objective value Z_3. Hence we have $Z_3 \leq Z_{sep}$. In addition, we can construct feasible policies for the Separate Pricing problem (4.2) based on $\bar{\boldsymbol{d}}$. Essentially we offer demand rate $\bar{d}_i$ and corresponding price $p_i(\bar{d}_i)$ in market i throughout the selling horizon. We call this policy as the Fixed Price policy. Given a capacity allocation vector $\boldsymbol{b}$, letting $W(\boldsymbol{b})$ be the expected revenue of the Fixed Price policy, we have

$$W(\boldsymbol{b}) = \sum_{i=1}^{n} \mathbb{E}\min\{b_i, D_i(\bar{d}_i)\}\, p_i(\bar{d}_i),$$

where $D_i(\bar{d}_i)$ is the total demand in market i under the Fixed Price policy, which follows a Binomial distribution with parameters $(T, \lambda_i \bar{d}_i)$. Letting $\mathcal{B}' = \{\boldsymbol{b} \in [0, C]^n : \sum_{i=1}^n b_i = C\}$, i.e., $\mathcal{B}'$ is the continuous relaxation of $\mathcal{B}$, it is not hard to verify the following fact.

$$\max_{\boldsymbol{b}\in\mathcal{B}} W(\boldsymbol{b}) = \max_{\boldsymbol{b}\in\mathcal{B}'} W(\boldsymbol{b}).$$

Let $\bar{\boldsymbol{b}}$ be the maximizer of $W(\boldsymbol{b})$ over $\mathcal{B}$ and $\mathcal{B}'$, then the Fixed Price policy with capacity allocation $\bar{\boldsymbol{b}}$ is feasible for Separate Pricing problem (4.2). Hence we have the following chain of inequalities:

$$\begin{aligned} R_{sep} \geq W(\bar{\boldsymbol{b}}) \geq W(\boldsymbol{\eta}\, C) &= \sum_{i=1}^{n} \mathbb{E}\min\{\eta_i C, D_i(\bar{d}_i)\}\, p_i(\bar{d}_i) \\ &\geq \sum_{i=1}^{n} \left[\min\{\eta_i C, T\,\lambda_i \bar{d}_i\} - \frac{1}{2}\sigma(D_i(\bar{d}_i))\right] p_i(\bar{d}_i) \\ &= \sum_{i=1}^{n} T\,\lambda_i \bar{d}_i\, p_i(\bar{d}_i) - \frac{1}{2}\sum_{i=1}^{n} \sigma(D_i(\bar{d}_i))\, p_i(\bar{d}_i) \\ &= Z_3 - \frac{1}{2}\sum_{i=1}^{n} \sigma(D_i(\bar{d}_i))\, p_i(\bar{d}_i). \end{aligned} \tag{4.16}$$

The third inequality holds due to Fact (I). The second equality follows from constraint (4.12), and the last equality follows from the definition of $\bar{\boldsymbol{d}}$. Combining inequality (4.15) and (4.16), we get

$$R_{sep} - R_{joint} \geq Z_3 - Z_4 - \frac{1}{2}\sum_{i=1}^{n} \sigma(D_i(\bar{d}_i))\, p_i(\bar{d}_i). \tag{4.17}$$

Next we show that under some general conditions, Separate Pricing collects more revenue eventually when the capacity and the number of time periods scale up. In order to do so, we consider a sequence of problems $\{\mathcal{P}^m : m =$

$1, 2, \ldots\}$ indexed by parameter $m \in \mathbb{Z}_+$. In problem $\mathcal{P}^m$, the total capacity is $m\,C$, while the selling horizon starts from time period 1 until time period $m\,T$. For time period $t = 1, \ldots, mT$ in problem $\mathcal{P}^m$, the probability that an arriving customer is of type i is $\lambda_i^m = \lambda_i$. With this definitions, we note that the problem studied above is $\mathcal{P}^1$. For Problem $\mathcal{P}^m$, letting R_{sep}^m and R_{joint}^m to denote the expected revenues from Separate Pricing and Joint Pricing respectively, we have the following theorem:

Theorem 4.3.2. *There exist* $\omega \geq 0, \nu \geq 0$ *such that*

$$\frac{R_{sep}^m - R_{join}^m}{R_{join}^m} \geq \omega - \frac{\nu}{\sqrt{m}}. \tag{4.18}$$

Proof. For Problem $\mathcal{P}^m$, we let $Z_{sep}^m, Z_{joint}^m, Z_3^m, Z_4^m$ be the optimal objective values of (NLP1) - (NLP4) respectively. It is not hard to verify that $\boldsymbol{\eta}$ based on the optimal solution of (NLP2) does not scale with m, neither does the optimal solution $\bar{\boldsymbol{d}}$ for (NLP3). Besides, $Z_{sep}^m, Z_{joint}^m, Z_3^m$ and Z_4^m all scale linearly with m. In particular, we have $Z_3^m = m\,Z_3, Z_4^m = m\,Z_4$. Similar to (4.15) and (4.17) for Problem $\mathcal{P}^1$, we have

$$R_{joint}^m \leq Z_{joint}^m = Z_4^m \leq Z_3^m \tag{4.19}$$

$$R_{sep}^m - R_{joint}^m \geq Z_3^m - Z_4^m - \frac{1}{2}\sum_{i=1}^{n} \sigma(D_i^m(\bar{d}_i))\, p_i(\bar{d}_i), \tag{4.20}$$

where $D_i^m(\bar{d}_i)$ is the total demand in market i throughout the selling horizon given the fixed demand rate $\bar{d}_i$. Hence $D_i^m(\bar{d}_i)$ follows a Binomial distribution with parameters $(m\,T, \lambda_i \bar{d}_i)$, thus $\sigma(D_i^m(\bar{d}_i)) = \sqrt{m\,T\,\lambda_i \bar{d}_i(1 - \lambda_i \bar{d}_i)}$. Now combining (4.19) and (4.20) we have

$$\begin{aligned}\frac{R^m_{sep} - R^m_{joint}}{R^m_{joint}} &\geq \frac{Z^m_3 - Z^m_4 - \frac{1}{2}\sum_{i=1}^n \sqrt{m\,T\,\lambda_i\bar{d}_i(1-\lambda_i\bar{d}_i)}}{Z^m_4} \\ &= \frac{m\,Z_3 - m\,Z_4 - \frac{1}{2}\sum_{i=1}^n \sqrt{m\,T\,\lambda_i\bar{d}_i(1-\lambda_i\bar{d}_i)}}{m\,Z_4} \\ &= \frac{Z_3 - Z_4}{Z_4} - \frac{\sum_{i=1}^n \sqrt{T\,\lambda_i\bar{d}_i(1-\lambda_i\bar{d}_i)}}{2\,Z_4\sqrt{m}}.\end{aligned}$$

Letting $\omega = \frac{Z_3 - Z_4}{Z_4} \geq 0$ and $\nu = \frac{\sum_{i=1}^n \sqrt{T\,\lambda_i\bar{d}_i(1-\lambda_i\bar{d}_i)}}{2\,Z_4} \geq 0$, the result follows. □

Recall that Z_3 and Z_4 are optimal objective values of (NLP3) and (NLP4), where the constraint in (NLP4) is tighter than that in (NLP3). When the demand functions in all the markets are the same, we have $Z_3 = Z_4$. However, in general we have $Z_3 \geq Z_4$, and it is not hard to construct problem instances with $Z_3 > Z_4$, in which case $\omega \neq 0$, and for any $m \geq (\frac{\nu}{\omega})^2$, we have $R^m_{sep} \geq R^m_{join}$. In other words, as long as $Z_3 > Z_4$, Separate Pricing achieves larger expected revenue than Joint Pricing when m is large enough.

Intuitively, the above results says that pricing flexibility is more important than capacity flexibility when the capacity and the number of time periods are large enough. On the other hand, when the capacity and number of time periods in the selling horizon is small, in particular, when the total capacity is 1, if we choose Separate Pricing, then we can only generate revenue from one of the markets. However, if we choose Joint Pricing, we can always capture the same revenue from this market while generating revenue from other markets with positive probability. Hence we have $R^m_{sep} < R^m_{join}$ in this case. Thus neither Separate Pricing nor Joint Pricing is always superior to the other. It is sensible that for a general problem, it may be beneficial to adopt a hybrid method where we use Separate Pricing initially with a large capacity, and switch to Joint Pricing towards the end of selling horizon when the remaining capacity is small. We study this hybrid model in the next section.

4.4 Hybrid model

Motivated by the observation in last section, we consider a hybrid model where the firm uses Separate Pricing in the early stage of the selling horizon and switches to Joint Pricing towards the end of the selling horizon. This model has some potential applications. For instance, in fashion goods retail industry, it is difficult to carry out replenishment within the selling season. The firm usually distributes the available inventory to retail stores in different regions at the beginning of the selling season. Once such distribution is finished, redistribution of inventories within the season is rare due to high transportation and overhead cost. When the end of the selling season approaches with a smaller total inventory, the firm may collect the remaining inventories in these retail stores and sell them in a few outlet stores, or through a cental online outlet store. In this section we study such a hybrid model and develop approximate solutions.

4.4.1 Formulation

In the hybrid model we have two additional two parameters: The end time for Joint Pricing τ and the capacity allocation vector $\boldsymbol{b} = (b_1, \ldots, b_n)$ with $\sum_i b_i = C$. For a $\tau \in \{1, \ldots, T\}$, Separate Pricing is adopted from time period 1 to τ while Joint Pricing is used from time period $\tau + 1$ to T. During time period 1 to τ, the firm operates in n separate markets. The available capacity in market i is

b_i, and excess demands will be lost if the total demand in market i exceeds b_i. Remaining capacities in all the markets will join together at time $\tau + 1$, and the Joint Pricing will be adopted there afterwards until time period T, at which time the remaining capacity becomes obsolete. We call the market in which the firm operates from time period $\tau + 1$ to T the joint market or the final market interchangeably. The firm's problem is to find a pricing policy that maximizes the total expected revenue generated over the entire selling horizon.

For $t = 1, \dots, \tau$, we use a n-dimensional vector $\boldsymbol{x} = (x_1, \dots, x_n)$ to represent the remaining capacities in all the markets, where x_i is the remaining capacity in market i. At time period t, if the remaining capacity is $\boldsymbol{x}$, then the following dynamic programming recursion maximizes the total expected revenue from time period t onwards:

$$\begin{aligned} V_t(\boldsymbol{x}) &= \max_{\boldsymbol{d}} \left\{ \sum_{i=1}^{n} \lambda_i (r_i(d_{it}) + d_{it}\, V_t(\boldsymbol{x} - \boldsymbol{e}_i)) + (1 - \sum_{i=1}^{n} \lambda_i d_{it})\, V_{t+1}(\boldsymbol{x}) \right\} \\ &= \sum_{i=1}^{n} \max_{d_{it}} \{\lambda_i (r_i(d_{it}) - d_{it}\, \Delta_i V_{t+1}(\boldsymbol{x}))\} + V_{t+1}(\boldsymbol{x}), \end{aligned} \tag{4.21}$$

where $\Delta_i\, V_{t+1}(\boldsymbol{x}) = V_{t+1}(\boldsymbol{x}) - V_{t+1}(\boldsymbol{x} - \boldsymbol{e}_i)$, and $\boldsymbol{e}_i$ is an n-dimensional unit vector with 1 at the i-th component and 0 at all other components. The boundary condition is $V_t(0, 0) = 0$, $V_{\tau+1}(\boldsymbol{x}) = \Psi_{\tau+1}(\sum_{i=1}^{n} x_i)$, where $\Psi_t(\cdot)$ is the expected revenue from Joint Pricing defined by optimality equation (4.3). Note that $\Psi_t(\cdot)$ is needed only for $t = \tau + 1, \dots, T$, i.e., in the final market periods.

In the final market time periods from $\tau + 1$ to T, the decision in each time period is essentially a single common price to offer. Hence instead of controlling an n-dimensional decision vector $\boldsymbol{d}_t \in \mathcal{S} = \{\boldsymbol{d} \in [0,1]^n : p_1(d_1) = p_2(d_2) = \dots = p_n(d_n)\}$ as in equation (4.3), we prefer to control a single decision variable $d_{ft} = d_{1t}$ with implied price $p_{ft} = p_1(d_{ft})$. For $i = 1, \dots, n$, the demand rate in market i at time period t is $d_i(p_{ft}) = d_i(p_1(d_{ft})) \doteq d_i(d_{ft})$, while the one period revenue function is $r_f(d_{ft}) = \sum_{i=1}^n \lambda_i r_i(d_i(d_{ft})) = \sum_{i=1}^n \lambda_i d_i(d_{ft})\, p_{ft}$. With these new notations, we can rewrite (4.3) as

$$\Psi_t(y) = \max_{d_{ft}\in[0,1]} \Big\{ r_f(d_{ft}) - \sum_{i=1}^{n} \lambda_i d_i(d_{ft})\, \Delta\Psi_{t+1}(y) \Big\} + \Psi_{t+1}(y). \tag{4.22}$$

In order to make sure (4.22) can be solved in an efficient manner, we assume $r_f(d_{ft})$ is continuous, bounded and concave, and $\sum_{i=1}^n \lambda_i d_i(d_{ft})$ is convex. Note that the implication of the assumption is that $r_f(d_{ft}) - \Delta \sum_{i=1}^n \lambda_i d_i(d_{ft})$ is concave for any $\Delta \geq 0$, hence we have a concave maximization problem in the dynamic programming recursion, which is easy to solve. For instance, when all the markets use linear demand moedel with the same price range, the condition is satisfied.

While the dynamic programming recursion (4.22) in the final market is tractable to solve, the value function (4.21) from time period 1 to τ has an n-dimensional state space, and it is impractical to solve exactly for any real size problems when the dimension exceeds 3 or 4. We will develop and analyze tractable approximations for (4.21) in the rest of this chapter.

Note that if $\Psi_{\tau+1}(\sum_{i=1}^n x_i)$ is additively separable, then the high dimensional dynamic program (4.21) can be decomposed into n independent single dimensional dynamic programs, which are tractable to solve. In other words, the intractability of the dynamic programming (4.21) is due to a non-separable boundary condition at time period $\tau + 1$. Intuitively, since the remaining capacities from different markets will join together at time period $\tau + 1$, and the expected revenue from the total remaining capacity at $\tau + 1$ is a non-separable function, the evolutions of sales in different markets have an impact on each other's pricing policy. For instance, when $n = 2$, if the capacity in market 1 is depleted quickly, then it is unlikely that we will have a large amount of remaining capacity from market 1 at the end of time period τ. Hence we may want to keep the price in market 2 at a high level, since we have $(T - \tau)$ sales opportunities in the final market and the capacity will be scarce. On the other hand if the demand in market 1 is not very large and we anticipate that there will be a large number of remaining capacity in market 1 at the end of time period τ, then we may want to price lower in market 2 to stimulate demands in order to make use of the plenty remaining capacity at time period $\tau + 1$ in the final market to achieve higher revenue.

From the discussion we see that it is the interaction among different markets that creates an intractable dynamic programming recursion. Besides, even if we were able to solve (4.21) and get the optimal policy, it will not be a practically attractive policy, since we need to monitor the remaining capacities in all the markets to make pricing decision in each market. If we can break up the interaction among different markets and still have a good approximation to the original value function (4.21), then we have a tractable approximation and an operationally attractive policy that only needs the local capacity information to make local pricing decisions.

Note that the interaction among different markets is realized at time period $\tau + 1$ when all the remaining capacities join together. For a specific market i, the impact of other markets on market i's pricing policy is through the total remaining capacity at time period $\tau + 1$. In other words, if we know exactly the total remaining capacity at time period $\tau + 1$ from markets other than i, then we can set the prices in market i from time period 1 to τ optimally. Following this idea, we first develop a nonlinear program based on the formulation of a deterministic counterpart of the original stochastic problem, from which we extract a simple Fixed Price policy. Then we use the resulting capacity from the Fixed Price policy as an approximation to the remaining capacities in different markets at time period $\tau+1$, based on which we construct better pricing policies for time period 1 to τ.

4.4.2 Tractable Policies Based on a Deterministic Formulation

Consider a deterministic version of the hybrid model: The firm has C units of products to sell at time period 1, which is a *continuous* quantity. There are n separate markets initially, and the firm has allocates b_i units to market i at the beginning of the selling horizon, where $\sum_{i=1}^{n} b_i = C$. For $t = 1, \ldots, \tau, i = 1, \ldots, n$, if the offered demand rate in market i is d_{it}, then the realized demand is exactly $\lambda_i d_{it}$ with a revenue of $r_i(d_{it})$. At time period $\tau + 1$, the remaining capacities in all the markets join together as a single final market. For $t = \tau + 1, \ldots, T$, if the controlled demand rate is d_{ft}, then the realized demand is $\sum_{i=1}^{n} \lambda_i d_i(d_{ft})$ with a revenue of $r_f(d_{ft})$. The firm wants to maximize its total revenue by controlling $d_{it}, t = 1, \ldots, \tau, i = 1, \ldots, n$ and $d_{ft}, t = \tau + 1, \ldots, T$. The following nonlinear program solves the revenue maximization problem:

$$(\mathbf{NLP5})\ Z_5 = \max \quad \sum_{t=1}^{\tau}\sum_{i=1}^{n} \lambda_i\, r_i(d_{it}) + \sum_{t=\tau+1}^{T} r_f(d_{ft}) \tag{4.23}$$

$$\text{subject to} \quad \sum_{t=1}^{\tau} \lambda_i\, d_{it} \le b_i, i = 1, \ldots, n \tag{4.24}$$

$$\sum_{t=1}^{\tau}\sum_{i=1}^{n} \lambda_i\, d_{it} + \sum_{t=\tau+1}^{T}\sum_{i=1}^{n} \lambda_i d_i(d_{ft}) \quad \le C \tag{4.25}$$

$$0 \le d_{it} \le 1, t = 1, \ldots, \tau, i = 1, \ldots, n \tag{4.26}$$

$$0 \le d_{ft} \le 1, t = \tau + 1, \ldots, T. \tag{4.27}$$

The next proposition shows that Z_5 from (NLP5) provides an upper bound for the optimal total expected revenue of the stochastic version of the hybrid model. The proof is deferred into Appendix C.1.

Proposition 4.4.1. *We have $V_1(\boldsymbol{b}) \le Z_5$ for any given $\boldsymbol{b}$ satisfying $\sum_{i=1}^{n} b_i = C$.*

Due to the assumptions about $r_i(\cdot), r_f(\cdot)$ and $\sum_{i=1}^{n} d_i(d_{ft})$, we know that (NLP5) is a concave maximization problem, and the proof for the following lemma is straightforward.

Lemma 4.4.2. *There exists an optimal solution $\{\boldsymbol{d}_t^* : t = 1, \ldots, \tau\}, \{d_{ft}^* : t = \tau + 1, \ldots, T\}$ for (NLP5) which satisfies that $\boldsymbol{d}_t^* = \boldsymbol{d}^*, t = 1, \ldots, \tau$ and $d_{ft}^* = d_f^*, t = \tau + 1, \ldots, T$ for some $\boldsymbol{d}^*$ and d_f^*.*

An immediate result based on Lemma 4.4.2 is that we can write $Z_5 = \sum_{i=1}^{n} \tau\, \lambda_i r_i(d_i^*) + (T - \tau) r_f(d_f^*)$. Besides, Lemma 4.4.2 also implies a Fixed Price policy where we use a fixed price $p_i^* = p_i(d_i^*)$ in market i from time period 1 to τ, $i = 1, \ldots, n$, and another fixed price $p_f^* = p_1(d_f^*)$ in the joint market from time period $\tau + 1$ to T. If this Fixed Price policy is adopted, the total number of demand in market i from time period 1 to τ is $D_i(d_i^*)$, which follows a Binomial distribution with parameters $(\tau, \lambda_i d_i^*)$. The revenue in market i from time period 1 to τ can be written as $p_i^* \min\{b_i, D_i(d_i^*)\}$. At time period $\tau + 1$, the remaining capacity from market i can be written as $(b_i - D_i(d_i^*))^+$, where we use $(\cdot)^+$ to denote the function $\max\{0, \cdot\}$. Then the revenue in the joint market from time period $\tau + 1$ to T can be written as $p_f^* \min\{\sum_{i=1}^n (b_i - D_i(d_i^*))^+, D_f(d_f^*)\}$, where $D_f(d_f^*)$ is the demand in the joint market, which follows a Binomial distribution with parameters $(T - \tau, \sum_{i=1}^n \lambda_i d_i(d_f^*))$. Hence the total expected revenue achieved by the Fixed Price policy can be written as

$$R_{FP} = \sum_{i=1}^{n} p_i^* \mathbb{E} \min\{b_i, D_i(d_i^*)\} + p_f^* \mathbb{E} \min \left\{ \sum_{i=1}^{n} (b_i - D_i(d_i^*))^+, D_f(d_f^*) \right\}. \quad (4.28)$$

While the Fixed Price policy is an attractive policy from operational point of view, we use it as a building block to construct other tractable policies with better performance. The idea is as follows. We pick one market j, and for any market $i \neq j$, we offer the fixed price p_i^* from time period 1 to τ as in the Fixed Price policy. For market j, when the remaining capacity at time period t is $x_j \leq b_j$, we use a dynamic pricing policy based on the following single dimensional dynamic programming recursion:

$$\begin{aligned}
\Phi_{jt}(x_j) &= \max_{d_{jt} \in [0,1]} \Big\{ \lambda_j [r_j(d_{jt}) + d_{jt}\, \Phi_{j,t+1}(x_j - 1)] + (1 - \lambda_j\, d_{jt}) \Phi_{j,t+1}(x_j) \Big\} \\
&= \max_{d_{jt} \in [0,1]} \Big\{ \lambda_t [r_j(d_{jt}) - d_{jt}\, \Delta\Phi_{j,t+1}(x_j)] \Big\} + \Phi_{j,t+1}(x_j), \qquad (4.29)
\end{aligned}$$

with the boundary condition $\Phi_{jt}(0) = \mathbb{E}\, \Psi_{\tau+1}(\sum_{i \neq j} (b_i - D_i(d_i^*))^+)$ and $\Phi_{j,\tau+1}(x_j) = \mathbb{E}\, \Psi_{\tau+1}(x_j + \sum_{i \neq j} (b_i - D_i(d_i^*))^+)$, where $\sum_{i \neq j} (b_i - D_i(d_i^*))^+$ is the total number of remaining capacities at time period $\tau + 1$ in all markets other than j, which use fixed prices from time period 1 to τ as in the Fixed Price policy. When any market $i \neq j$ uses the fixed price p_i^* from time period 1 to τ, given the remaining capacity x_j in market j at time period t, $\Phi_{jt}(x_j)$ represents the maximum expected revenue we can achieve in market j from time period t to τ and in the final market from time period $\tau + 1$ to T. Notice that d_j^* and d_f^* with their corresponding prices p_j^* and p_f^* are feasible for the dynamic programming recursion (4.29) and (4.22) respectively, which implies that

$$\Phi_{j1}(b_j) \geq p_j^* \mathbb{E} \min\{b_j, D_j(d_j^*)\} + p_f^* \mathbb{E} \min \left\{ \sum_{i=1}^{n} (b_i - D_i(d_i^*))^+, D_f(d_f^*) \right\}.$$

Now we have constructed a policy using fixed price p_i^* in market $i \neq j$ from time period 1 to τ, dynamic pricing policy implied by (4.29) in market j from time period 1 to τ, and the dynamic pricing policy implied by (4.22) in the joint market from time period $\tau + 1$ to T. The total expected revenue from this policy can be written as

$$\begin{aligned} R_j &= \Phi_{j1}(b_j) + \sum_{i \neq j} p_i^* \mathbb{E} \min\{b_i, D_i(d_i^*)\} \\ &\geq \sum_{i=1}^{n} p_i^* \mathbb{E} \min\{b_i, D_i(d_i^*)\} + p_f^* \mathbb{E} \min \left\{ \sum_{i=1}^{n} (b_i - D_i(d_i^*))^+, D_f(d_f^*) \right\} = R_{FP}. \end{aligned} \tag{4.30}$$

Note that (4.30) holds for any $j = 1, \ldots, n$, thus by varying our choice of j, we can find the best such policy which achieves the maximum total expected revenue

$$R_{FBD} = \max_{j \in \{1, \ldots, n\}} R_j. \tag{4.31}$$

We call this policy the Fixed Price Based Dynamic Pricing policy (FBD). Clearly we have $R_{FBD} \geq R_{FP}$. We analyze the performance of this policy in next subsection.

4.4.3 Asymptotic Analysis

In this section, we show that when the number of time periods and the initial capacities scale up linearly in the same rate, the FBD policy is asymptotically optimal.

We consider a sequence of problems $\{\mathbb{P}^m : m = 1, 2, \ldots\}$ indexed by parameter $m \in \mathbb{Z}_+$. In problem $\mathbb{P}^m$, the total capacity is $m\,C$, while the initial capacity allocation vector is $m\,\boldsymbol{b}$. The selling horizon starts from time period 1 and lasts until time period $m\,T$. From time period 1 to $m\tau$, the firm operates in n separate markets and is able to set different prices in different markets. The arrival probability in market i at any time period is given by λ_i, which satisfies $\sum_{i=1}^n \lambda_i = 1$. At time period $m\,\tau + 1$, the remaining capacities from all the markets join together, which is used to satisfy the demand in the joint market from time period $m\,\tau + 1$ to $m\,T$.

With this definitions, we note that the problem we study in the previous section is $\mathbb{P}^1$. Our goal here is to show that both the Fixed Price policy and the FBD policy are asymptotically optimal for problem $\mathbb{P}^m$ when m approaches infinity. For problem $\mathbb{P}^m$, letting $V_1^m(m\,\boldsymbol{b})$ denote the optimal expected revenue, while R_{FP}^m and R_{FBD}^m denote the expected revenue achieved by the Fixed Price policy and the FBD policy respectively, we have the following theorem:

Theorem 4.4.3. *For any given* $\boldsymbol{b}$ *satisfying* $\sum_{i=1}^n b_i = C$, *we have* $\lim_{m\to\infty} \frac{R_{FBD}^m}{V_1^m(m\,\boldsymbol{b})} = \lim_{m\to\infty} \frac{R_{FP}^m}{V_1^m(m\,\boldsymbol{b})} = 1$.

Proof. Let Z_5^m be the optimal objective function of (NLP5) for problem $\mathbb{P}^m$, which upper bounds $V_1^m(m\,\boldsymbol{b})$. Note that in (NLP5) for problem $\mathbb{P}^m$, all the coefficients scale linearly with m. Thus the optimal solution does not scale with m while the objective function scales linearly with m, i.e., the optimal solution is $(\boldsymbol{d}^*, d_f^*)$ with corresponding prices $(\boldsymbol{p}^*, p_f^*)$, and we have $Z_5^m = m\, Z_5 = m \sum_{i=1}^n \tau\, \lambda_i r_i(d_i^*) + m(T-\tau) r_f(d_f^*)$. Similar to (4.28) for problem $\mathbb{P}^1$, we have

$$R_{FP}^m = \sum_{i=1}^n p_i^* \mathbb{E} \min\{m\, b_i, D_i^m(d_i^*)\} + p_f^* \mathbb{E} \min \left\{ \sum_{i=1}^n [m\, b_i - D_i^m(d_i^*)]^+, D_f^m(d_f^*) \right\}, \tag{4.32}$$

where $D_i^m(d_i^*)$ is the total demand in market i from time period 1 to $m\tau$, which follows a binomial distribution with parameters $(m\tau, \lambda_i\, d_i^*)$; and $D_f^m(d_f^*)$ is the total demand from time period $m\tau + 1$ to mT, which follows a Binomial distribution with parameters $(m(T-\tau), \sum_{i=1}^n \lambda_i d_i(d_f^*))$. The first summation in (4.32) represents the total expected revenue in the separate markets from time period 1 to τ, while the second term is the total expected revenue in the joint market from time period $\tau + 1$ to T. For each $i = 1, \ldots, n$, we have

$$\begin{aligned} p_i^*\, \mathbb{E} \min\{m\, b_i, D_i^m(d_i^*)\} &\geq p_i^* \,\left\{\min\{m\, b_i, \mathbb{E}(D_i^m(d_i^*))\} - \sigma(D_i^m(d_i^*))/2\right\} \\ &= p_i^* \left\{ \min\{m\, b_i, m\, \tau\, \lambda_i d_i^*\} - \sqrt{m\, \tau\, \lambda_i d_i^* (1 - \lambda_i d_i^*)}/2 \right\} \\ &= p_i^* \left\{ m\, \tau\, \lambda_i d_i^* - \sqrt{m\, \tau\, \lambda_i d_i^* (1 - \lambda_i d_i^*)}/2 \right\} \\ &= m\, \tau\, \lambda_i r_i(d_i^*) - G_i \sqrt{m}, \end{aligned}$$

where $G_i = p_i^* \sqrt{\tau\, \lambda_i d_i^* (1 - \lambda_i d_i^*)}/2$. The first inequality follows from Fact (I) and the second equality follows from feasibility constraint (4.24).

Let $K = \sum_{i=1}^n [m\, b_i - D_i^m(d_i^*)]^+$, then $\mathbb{E}(K) \geq \sum_{i=1}^n [m\, b_i - \mathbb{E}(D_i^m(d_i^*))]^+$ due to Jensen's Inequality. Besides, the second term in (4.32) can be written as

$p_f^* \,\mathbb{E}\, \min\{K, D_f^m(d_f^*)\}$. We have the following chain of inequalities:

$$
\begin{aligned}
& p_f^* \, \left\{\mathbb{E}\,\min\left\{K, D_f^m(d_f^*)\right\}\right\} \geq p_f^* \, \left\{\mathbb{E}\,\min\left\{\mathbb{E}(K), D_f^m(d_f^*)\right\} - \frac{1}{2}\sigma(K)\right\} \\
\geq \; & p_f^* \left\{\min\left\{\mathbb{E}(K), \mathbb{E}(D_f^m(d_f^*))\right\} - \frac{1}{2}\sigma(D_f^m(d_f^*)) - \frac{1}{2}\sigma(K)\right\} \\
\geq \; & p_f^* \left\{\min\left\{\sum_{i=1}^{n}[m\, b_i - \mathbb{E}(D_i^m(d_i^*))]^+, \mathbb{E}(D_f^m(d_f^*))\right\} - \frac{1}{2}\sigma(D_f^m(d_f^*)) - \frac{1}{2}\sigma(K)\right\} \\
= \; & p_f^* \left\{\min\left\{\sum_{i=1}^{n}[m\, b_i - m\tau\, \lambda_i d_i^*]^+, m(T-\tau)\sum_{i=1}^{n}\lambda_i d_i(d_f^*)\right\}\right. \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad \left. -\frac{1}{2}\sigma(D_f^m(d_f^*)) - \frac{1}{2}\sigma(K)\right\} \\
= \; & p_f^* \left\{\min\left\{\sum_{i=1}^{n}[m\, b_i - m\tau\, \lambda_s d_s^*], m(T-\tau)\sum_{i=1}^{n}\lambda_i d_i(d_f^*)\right\}\right. \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad \left. -\frac{1}{2}\sigma(D_f^m(d_f^*)) - \frac{1}{2}\sigma(K)\right\} \\
= \; & p_f^* \left\{m(T-\tau)\sum_{i=1}^{n}\lambda_i d_i(d_f^*) - \frac{1}{2}\sigma(D_f^m(d_f^*)) - \frac{1}{2}\sigma(K)\right\} \\
\geq \; & p_f^* \left\{m(T-\tau)\sum_{i=1}^{n}\lambda_i d_i(d_f^*) - \frac{1}{2}\sigma(D_f^m(d_f^*)) - \frac{1}{2}\sum_{i=1}^{n}\sigma(D_i^m(d_i^*))\right\} \\
= \; & m(T-\tau)\, r_f(d_f^*) - G_f\sqrt{m},
\end{aligned}
$$

where

$$
G_f = \frac{p_f^*}{2}\left\{\sqrt{(T-\tau)\, \sum_{i=1}^{n}\lambda_i d_i(d_f^*)(1 - \sum_{i=1}^{n}\lambda_i d_i(d_f^*))} + \sum_{i=1}^{n}\sqrt{\tau\, \lambda_i d_i^*(1-\lambda_i d_i^*)}\right\}.
$$

The first and second inequalities follow from Fact (I), the third inequality holds due to the fact about $\mathbb{E}(K)$ shown above, the second equality follows from constraint (4.24), the third equality follows from constraint (4.25), and the last inequality follows from Lemma B.3.1 shown in Appendix B.3. Combing the above two terms, we have

$$
R_{FP}^m \geq m\sum_{i=1}^{n}\tau\lambda_i r_i(d_i^*) + m(T-\tau)\, r_f(d_f^*) - \sum_{i=1}^{n}G_i\sqrt{m} - G_f\sqrt{m} = m\, Z_5 - G\sqrt{m},
$$

where $G = \sum_{i=1}^{n} G_i + G_f$. Therefore we have the following chain of inequalities:

$$1 \geq \frac{R_{FBD}^m}{V_1^m(m\,\boldsymbol{b})} \geq \frac{R_{FP}^m}{V_1^m(m\,\boldsymbol{b})} \geq \frac{R_{FP}^m}{Z_5^m} \geq \frac{m\,Z_5 - G\sqrt{m}}{m\,Z_5} = 1 - \frac{G}{Z_5\sqrt{m}} \xrightarrow{m\to\infty} 1.$$

□

4.5 Numerical Experiments

In this section, we evaluate the performance of different policies through numerical experiments. We begin by describing the benchmark policies and the experimental setup.

4.5.1 Benchmark and Policies

Upper Bound (UPB). This corresponds to the optimal objective function Z_5 from (NLP5), which is an upper bound for $V_1(\boldsymbol{b})$. Since we can not get the exact value of the optimal expected revenue $V_1(\boldsymbol{b})$ for high dimensional problems, we use this upper bound to get a feel for optimality gaps.

Fixed Price Policy (FP). This benchmark policy corresponds to solving the nonlinear program (NLP5), and offering fixed price p_i^* in market i from time period 1 to τ and fixed price p_f^* from time period $\tau + 1$ to T in the joint market. The total expected revenue achieved by FP policy is given by R_{FP} in equation (4.28). Note that R_{FP} can be calculated exactly.

Fixed Price Based Dynamic Pricing Policy (FBD). This benchmark policy corresponds to solving the single dimensional dynamic programming recursion in (4.29) for each $j \in \{1, \ldots, n\}$, calculating the expected revenue R_j given in (4.30), and finding the maximum total expected revenue R_{FBD} given in (4.31). Suppose $k = \arg\max_{j \in \{1,\ldots,n\}} R_j$, then for any market $i \neq k$, FBD uses the fixed price p_i^* based on (NLP5) from time period 1 to τ. For market k, FBD uses a dynamic pricing strategy implied by $\Phi_{kt}(x_k)$ defined in (4.29). In other words, letting $d_{kt}^{\Phi}(x_k)$ be the optimal decision variable at time period t with state x_k in (4.29), FBD offers price $p_k(d_{kt}^{\Phi}(x_k))$ in market k when there are x_k units remaining at time period t. Note that FBD only needs to keep track of the the remaining capacity in market k from time period 1 to τ. From timer period $\tau + 1$ to T, FBD uses the dynamic pricing policy implied by dynamic programming recursion (4.22). In other words, letting $d_{ft}^{\Psi}(y)$ be the optimal decision variable of (4.22) at time period t with state y, FBD offers price $p_1(d_{ft}^{\Psi}(y))$ in the joint market at time period t when there are y units remaining in total. The total expected revenue of FBD is R_{FBD}, which can be calculated exactly too.

Mixed Policy (MXD). This benchmark policy is also based on solving the single dimensional dynamic programming recursion in (4.29) for each $j \in \{1, \ldots, n\}$. Instead of picking one market and adopting dynamic pricing in the chosen market while keeping fixed prices in all other markets as in FBD policy, MXD uses dynamic pricing in all the markets, pretending that every other market is using fixed price from the optimal solution of (NLP5). More specifically, for $t = 1, \ldots, \tau$ and $j = 1, \ldots, n$, letting $d_{jt}^{\Phi}(x_j)$ be the optimal decision variable at time period t with state x_j in (4.29), MXD offers price $p_j(d_{jt}^{\Phi}(x_j))$ in market j when there are x_j units remaining at time period t. Note that with MXD, the pricing decision in market j only depends on the remaining capacity in market j. From time period $\tau + 1$ to T, MXD uses the same dynamic pricing strategy implied by (4.22) as FBD. Unlike FP or FBD, the total expected revenue achieved by MXD can not be computed exactly, hence we estimate it through simulation.

4.5.2 Experimental Setup

A hybrid problem instance is defined by the following parameters: Initial capacity C, number of time periods in the selling horizon T, number of markets n, capacity allocation vector $\boldsymbol{b}$, and ending time of separate markets τ. The demand process is characterized by arrival probabilities λ_i and the demand functions $d_i(\cdot), i = 1, \ldots, n$. In our experiments, we set $\tau = \frac{3}{4}T$ and use symmetric arrival probabilities $\lambda_i = 1/n, i = 1, \ldots, n$. We adopt linear demand functions where $d_i(p_i) = \alpha_i(1 - p_i/p_{max})$, where p_{max} is the maximum price at which the corresponding demand rate is 0. We use a common maximum price $p_{max} = 2$ for all the markets, and vary α_i in $[0, 1]$. Note that the price sensitivity of the demand model in market i is given by α_i/p_{max}, and the demand function vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ completely characterizes the demand models given fixed p_{max}. With this setting, when the controlled demand rate is d_{ft} at time period t in the joint market, the assumption that $d_i(d_{ft}) = d_i(p_i(d_{ft}))$ is convex and $r_f(d_{ft})$ is concave is satisfied.

We work with test problems with 2 or 5 markets, corresponding to $n = 2$ or $n = 5$. For each (C, T) combination, when $n = 2$, we vary $\boldsymbol{\alpha}$ in $\{(0.6, 0.6), (0.8, 0.4), (1.0, 0.2)\}$. Similarly, when $n = 5$, we vary $\boldsymbol{\alpha}$ in $\{(0.6, 0.6, 0.6, 0.6, 0.6), (0.8, 0.7, 0.6, 0.5, 0.4), (1.0, 0.8, 0.6, 0.4, 0.2)\}$. Hence for each (C, T) combination, we have 6 problem instances in total with different numbers of markets and different demand models.

Under this setup, the revenue function in market i is given by $r_i(d_i) =$

$\frac{p_{max}}{\alpha_i} d_i (\alpha_i - d_i)$, hence the unconstraint maximizer of the revenue function $r_i(d_i)$ is given by $d_i^0 = \alpha_i/2$. For a problem instance where initial capacity is C, the number of time periods is T and the demand function vector is $\boldsymbol{\alpha}$, if we set the price in each market to maximize the one period revenue in each time period, then the total expected demand is $T \sum_{i=1}^n \lambda_i d_i^0$. Hence we can define the load factor of the problem as $l(C, T, \boldsymbol{\alpha}) = \frac{T \sum_{i=1}^n \lambda_i d_i^0}{C} = \frac{T}{2C} \frac{\sum_{i=1}^n \alpha_i}{n}$, which is a measure of the tightness of the capacity for the problem instance. We want to evaluate the performance of different policies under different load factors. Note that given C and T, the load factor depends on $\frac{\sum_{i=1}^n \alpha_i}{n}$. For each (C, T) combination, due to the way we set up the demand vector $\boldsymbol{\alpha}$, all the 6 problem instances corresponding to $n = 2$ and $n = 3$ with different $\boldsymbol{\alpha}$'s share the same load factor.

We vary T among $\{80, 160, 240\}$. For $T = 80$, we vary C in $\{16, 12, 9, 7\}$ so the load factors are varied in the set $L = \{1.50, 2.00, 2.67, 3.43\}$. For each $C \in \{16, 12, 9, 7\}$, we solve a modified version of (NLP5) to generate the capacity allocation vector $\boldsymbol{b}$: we treat $\boldsymbol{b}$ as a decision variable in (NLP5) and add a constraint that $\sum_{i=1}^n b_i = C$. This modification does not change the concavity property of the nonlinear program. Then we round the capacity allocation implied by the optimal solution of this modified nonlinear program to integers, which is denoted by $\boldsymbol{b}(C)$, and use $\boldsymbol{b}(C)$ throughout the computation of all the benchmark policies for this problem instance. For $T = 160$ and $T = 240$, we vary C in $\{32, 24, 18, 14\}$ and $\{48, 36, 27, 21\}$ respectively, so that the load factors are always in the same set L. Note that in this way we are scaling T and $C's$ linearly in the same rate. We also scale the capacity allocation vector $\boldsymbol{b}$ in a linear fashion. More specifically, when $T = 160$ and $C \in \{32, 24, 18, 14\}$, we use a capacity allocation vector $2\,\boldsymbol{b}(C/2)$; when $T = 240$ and $C \in \{48, 36, 27, 21\}$, we use a capacity allocation vector $3\,\boldsymbol{b}(C/3)$. This scaling approach corresponds to the asymptotic regime in Section 4.4.3. Hence we can expect that both FP and FBD policies perform better as T increases from 80 to 240. According to our experimental setup, a problem instance is completely characterized by $(T, C, \boldsymbol{\alpha})$, or equivalently by $(T, l, \boldsymbol{\alpha})$, where l is the load factor. Hence in total we have $3 * 4 * 6 = 72$ problem instances.

4.5.3 Experimental Results

We give our computational results in Table 4.1 and 4.2. Table 4.1 shows the results for problem instances with 2 separate markets, while Table 4.2 shows results for problem instances with 5 separate markets. In each table, the first two columns give parameters $(\boldsymbol{\alpha}, T)$, while columns 4 to 7 in the second row indicate the load factors of the problem instances. For each problem instance, we report three statistics: the gap between the upper bound by UPB and the expected revenue achieved by MXD policy denoted by "UPB vs. MXD"; the gap between the expected revenue achieved by MXD policy and FP policy denoted by "MXD vs. FP"; and the gap between the expected revenue achieved by MXD policy and FBD policy denoted by "MXD vs. FBD". All the gaps are reported as a percentage. For each pair, a positive gap means that the revenue achieved by the former policy (or bound) is larger than that achieved by the latter policy. We also report the average gaps across each row and column of the table.

From Table 4.1 and Table 4.2 we can make several observations. First of all, among the three policies, MXD outperforms FP and FBD in every problem instance, while FBD achieves larger expected revenue than FP in every instance. On average, MXD offers 5.98% improvement over FP policy in expected rev-

Table 4.1: Revenue gaps for problem instances with 2 separate markets.

n=2			load factor l				
α	T	Measures	1.50	2.00	2.67	3.43	Average
(0.6, 0.6)	80	UPB vs. MXD	7.39%	7.92%	8.36%	8.34%	8.00%
		MXD vs. FP	2.63%	4.25%	6.77%	9.84%	5.87%
		MXD vs. FBD	0.63%	0.68%	0.45%	0.51%	0.57%
(0.6, 0.6)	160	UPB vs. MXD	4.26%	4.51%	4.58%	4.62%	4.49%
		MXD vs. FP	2.62%	3.76%	5.40%	7.15%	4.73%
		MXD vs. FBD	0.68%	0.67%	0.61%	0.53%	0.62%
(0.6, 0.6)	240	UPB vs. MXD	3.16%	3.36%	3.38%	3.33%	3.31%
		MXD vs. FP	2.38%	3.26%	4.58%	6.02%	4.06%
		MXD vs. FBD	0.50%	0.44%	0.39%	0.37%	0.43%
(0.8, 0.4)	80	UPB vs. MXD	8.32%	8.80%	9.81%	8.29%	8.81%
		MXD vs. FP	3.15%	4.91%	8.52%	9.57%	6.54%
		MXD vs. FBD	0.37%	0.60%	0.69%	0.49%	0.54%
(0.8, 0.4)	160	UPB vs. MXD	4.61%	4.82%	5.24%	5.14%	4.95%
		MXD vs. FP	2.87%	3.99%	6.28%	8.28%	5.36%
		MXD vs. FBD	0.44%	0.45%	0.59%	0.65%	0.53%
(0.8, 0.4)	240	UPB vs. MXD	3.54%	3.64%	3.71%	3.62%	3.63%
		MXD vs. FP	2.64%	3.64%	5.19%	6.64%	4.53%
		MXD vs. FBD	0.27%	0.31%	0.34%	0.35%	0.32%
(1, 0.2)	80	UPB vs. MXD	8.10%	9.50%	10.10%	9.59%	9.32%
		MXD vs. FP	3.16%	5.35%	8.62%	11.28%	7.10%
		MXD vs. FBD	0.12%	0.30%	0.53%	0.58%	0.38%
(1, 0.2)	160	UPB vs. MXD	4.85%	5.18%	4.78%	5.87%	5.17%
		MXD vs. FP	3.14%	4.36%	5.65%	9.36%	5.63%
		MXD vs. FBD	0.28%	0.27%	0.23%	0.47%	0.31%
(1, 0.2)	240	UPB vs. MXD	3.66%	3.71%	3.77%	3.87%	3.75%
		MXD vs. FP	2.86%	3.71%	5.35%	7.03%	4.74%
		MXD vs. FBD	0.16%	0.13%	0.21%	0.18%	0.17%
Average		UPB vs. MXD	5.32%	5.72%	5.97%	5.85%	5.71%
		MXD vs. FP	2.83%	4.14%	6.26%	8.35%	5.39%
		MXD vs. FBD	0.38%	0.43%	0.45%	0.46%	0.43%

enue, while the maximum improvement is as high as 14.56%. The improvement of MXD over FBD is marginal at 0.63% on average.

Second, for fixed $\boldsymbol{\alpha}$ and load factor, when T increases from 80 to 160 and then to 240, both “UPB vs. MXD” and “MXD vs. FP” decrease monotonically. In other words, as T increases, the gap between the deterministic upper bound and the revenue of MXD policy decreases; the gap between the expected revenue of MXD and FP also decreases. This is consistent with the asymptotic optimality result of FP and FBD policy in Section 4.4.3. Note that since MXD outperforms both FP and FBD uniformly, we conjecture that MXD policy is also asymptotically optimal, although we are not able to show this conjecture rigourously.

Third, for fixed T and $\boldsymbol{\alpha}$, the gap between the revenues of MXD and FP policy increases monotonically when the load factor increases. This observation makes intuitive sense: when the capacity becomes much tighter in the system, it becomes more important to respond to the fluctuation of the demand process. FP policy does not respond to demand fluctuation since fixed prices are offered no matter what. MXD is able to respond to demand fluctuation since for any market, the price offered by MXD at any time period depends on the remaining capacity at that time period. FBD stands somewhere in between FP and MXD since it only responds to demand fluctuation in the picked market.

Fourth, for fixed T and load factor, when the markets become more different with each other with respect to price sensitivities, the gap between the expected revenues of MXD and FP has an increasing trend. In other words, when the markets are more different with each other, the benefit of using MXD policy is more significant.

Table 4.2: Revenue gaps for problem instances with 5 separate markets.

n=5			load factor l				
α	T	Measures	1.50	2.00	2.67	3.43	Average
(0.6,0.6,0.6,0.6,0.6)	80	UPB vs. MXD	10.80%	11.82%	12.12%	12.65%	11.85%
		MXD vs. FP	2.76%	5.50%	9.52%	14.56%	8.08%
		MXD vs. FBD	0.81%	0.84%	1.10%	0.97%	0.93%
(0.6,0.6,0.6,0.6,0.6)	160	UPB vs. MXD	5.77%	6.28%	6.67%	6.72%	6.36%
		MXD vs. FP	2.88%	4.44%	6.86%	9.59%	5.94%
		MXD vs. FBD	1.05%	0.93%	0.74%	0.76%	0.87%
(0.6,0.6,0.6,0.6,0.6)	240	UPB vs. MXD	4.08%	4.40%	4.66%	4.69%	4.46%
		MXD vs. FP	2.52%	3.72%	5.48%	7.50%	4.80%
		MXD vs. FBD	0.77%	0.67%	0.49%	0.42%	0.59%
(0.8,0.7,0.6,0.5,0.4)	80	UPB vs. MXD	11.46%	11.99%	12.93%	12.55%	12.23%
		MXD vs. FP	3.16%	5.69%	10.34%	14.50%	8.42%
		MXD vs. FBD	0.86%	0.91%	1.15%	1.03%	0.99%
(0.8,0.7,0.6,0.5,0.4)	160	UPB vs. MXD	6.32%	6.97%	6.56%	6.67%	6.63%
		MXD vs. FP	3.13%	5.02%	6.67%	9.43%	6.06%
		MXD vs. FBD	1.08%	1.03%	0.70%	0.70%	0.88%
(0.8,0.7,0.6,0.5,0.4)	240	UPB vs. MXD	4.40%	5.04%	5.06%	4.75%	4.81%
		MXD vs. FP	2.69%	4.34%	6.30%	7.74%	5.27%
		MXD vs. FBD	0.69%	0.75%	0.67%	0.49%	0.65%
(1,0.8,0.6,0.4,0.2)	80	UPB vs. MXD	13.27%	12.59%	12.18%	12.47%	12.63%
		MXD vs. FP	4.02%	6.03%	9.48%	14.28%	8.45%
		MXD vs. FBD	1.02%	0.91%	1.00%	0.93%	0.97%
(1,0.8,0.6,0.4,0.2)	160	UPB vs. MXD	6.96%	7.34%	6.85%	7.25%	7.10%
		MXD vs. FP	3.41%	5.15%	7.18%	10.72%	6.62%
		MXD vs. FBD	1.06%	0.88%	0.71%	0.85%	0.87%
(1,0.8,0.6,0.4,0.2)	240	UPB vs. MXD	4.99%	5.41%	4.92%	4.97%	5.07%
		MXD vs. FP	3.02%	4.66%	6.05%	8.27%	5.50%
		MXD vs. FBD	0.81%	0.76%	0.55%	0.55%	0.67%
Average		UPB vs. MXD	7.56%	7.98%	7.99%	8.08%	7.90%
		MXD vs. FP	3.06%	4.95%	7.54%	10.73%	6.57%
		MXD vs. FBD	0.90%	0.85%	0.79%	0.74%	0.82%

Finally, if we compare Table 4.1 and 4.2, we observe that the gaps between the expected revenues of MXD and FP with 5 markets are uniformly larger than that with 2 markets. In particular, for problem instances with 2 markets, MXD offers 5.39% improvement on average over FP policy, while it offers 6.57% improvement on average when there are 5 markets. Besides, while the gaps between expected revenues of MXD and FBD policy are small in the instances shown, we do observe that as the number of markets increases, the benefit of MXD policy over FBD policy becomes more significant. In particular, although not shown in the tables, we make a note here that for problem instances with 11 markets, the gap between the revenues of MXD and FBD can be larger than 3%. Overall, the benefit of MXD over FP and FBD is more significant when we have more markets. This observation also make good intuitive sense, since MXD responds to demand fluctuations in every market, while FBD responds to demand fluctuations in one market and FP does not respond to demand fluctuations at all.

In summary, we observe that MXD policy is the best policy overall, and its benefits are more significant when the load factor is large, the the total number of markets is large, or different markets have drastic differences in their price sensitivities.

4.6 Conclusion

In this chapter, we study a problem where we sell a product in different markets with different price sensitivities. A strategic decision is whether we should adopt price discrimination by allocating capacity to different markets and charging different prices in different markets, or we should pool the markets together and charge a single price. If the former is chosen, we obtain total pricing flexibility while lose capacity flexibility by committing certain number of inventories to each market. If the latter is chosen, we lose pricing flexibility since a common price needs to be offered, however, we have total capacity flexibility since there is no market separation and we can use any remaining capacity to satisfy the demand from any of the markets.

Using consistent demand models, we characterize the region where the benefit of pricing flexibility outweighs the cost of capacity flexibility. The characterization also motivates a hybrid model where we adopt pricing discrimination at the early stage and switch to joint pricing towards the end of the selling horizon. The optimal dynamic prices in the hybrid model are intractable to solve due to the high dimensional state space. We develop heuristics based on single dimensional approximation for each market, whose boundary condition is based on a deterministic formulation. Computational experiments indicate that the heuristics show significant improvements over the fixed price policy, especially in problem instances where the capacity is tight, the number of markets is larger, or the price sensitivities of customers in different markets are more dispersed.

There are a couple of questions which are worth further investigation. First of all, as noted in Section 4.5.2, we use a modified version of (NLP5) to generate the capacity allocation vector, which is asymptotically optimal when the capacity and number of time periods scale up linearly in the same rate. It is of interest to see if $V_1(\boldsymbol{b})$ belongs to a class of discrete concave functions, in which case tractable methods could be devised to find the optimal capacity allocation among the different markets. Second, the hybrid model is mainly characterized by the deadline of separate pricing τ. We treat τ as a given parameter, while it actually could be a decision variable that the seller can control. Due to the intractability of the hybrid model given fixed τ, we do not see a tractable approach to find the optimal value of τ. However, it is an interesting problem to develop some tractable approximations which involve τ as a decision variable.

CHAPTER 5

CONCLUSION

In this dissertation we study three multi-dimensional problems in single-resource revenue management. The main contributions of the work lie in developing tractable approximations to the original intractable formulations and generating operationally appealing policies with good performance in theory and practice.

Chapter 2 deals with a problem faced by a firm selling a product under limited information about the probabilities of different failure types. The goal is to learn the failure probabilities as the sales take place and dynamically decide whether it is profitable to continue or stop selling the product. Our approach builds on a dynamic programming formulation with embedded optimal stopping and learning features. This dynamic programming formulation has a high-dimensional state variable, and we proposed two approximation methods to address the computational difficulties due to the high-dimensional state variable. The first approximation method focuses on each failure type individually and solves a sequence of dynamic programs with scalar state variables. The second approximation method is based on a deterministic formulation that ignores the benefits from learning the failure probabilities. The two approximation methods are complementary to each other in the sense that while the first approximation method provides upper bounds on the value functions, the second one provides lower bounds.

This problem opens up several possible directions for future research. First, our model in Chapter 2 assumes that the failure probabilities are unknown constants, but it is conceivable that the failure probabilities may depend on age or usage of a unit. It is of interests to extend the model to age-dependent failure probabilities. Second, our beta-binomial learning model assumes that different types of failures occur independently of each other, while Dirichlet-multinomial model ensures that each product can fail from one failure at a time, inducing negative correlations between numbers of failures of different types. In reality, there may be general correlations among the probabilities of different failure types and one may desire to adopt other learning models that allow general correlations. In order to continue future research in this direction, more theory on multi-dimensional stochastic orders is needed. Third, it is worthwhile to investigate the possibility that it may be better for the company not to serve all of the demand in a time period. By rationing supply in the early time periods, the company may be able to learn about the reliability of the product while controlling the risk of facing too many returns.

In Chapter 3, we study the pricing and capacity allocation problem in a dual channel hotel revenue management setting, where a fixed price is used in a conference market and dynamic pricing can be employed in a spot market. The two markets join together after a predetermined deadline. The decision maker needs to allocate the capacity between the two markets and fix the price in the conference market at the beginning of the selling horizon.

For the operational problem of making the pricing decisions in the spot market, we construct a single dimensional approximation to the exact two dimensional dynamic programming formulation. The idea is to modify the boundary condition of a classical single dimensional dynamic pricing recursion to incorporate the possible joining capacity from the conference market at a future time. The policy based on our single dimensional approximation is not only appealing from an operational perspective, it also shows robust performance in numerical experiments with respect to load factor and higher level tactical decisions. At the tactical level, based on a deterministic formulation of a relaxed problem, we develop a heuristic to make decisions on conference market pricing and capacity allocation simultaneously. The mixed policy by combining the tactical level heuristic and the operational level pricing policy based on single dimensional approximation is asymptotically optimal, and it shows satisfactory performance in numerical experiments.

There are a couple of directions to enrich the model. First, while we focus on the problem with single night stay, it is a natural extension to consider the problem with multiple nights stay. The multiple nights stay problem is significantly more complicated and usually studied through heuristics. Our mixed policy may prove useful in designing good heuristics. Second, in this chapter, we assume that the conference market and the spot market are separated perfectly. In particular, the demand function in each market only depends on the price offered in its own market, and all future conference market demands are lost once the reserved rooms for the conference are fully booked. It is of interest to relax this assumption to model the possibility that conference customers can still make bookings in the spot market.

In Chapter 4, we study a problem where we sell a product to different markets with heterogeneous price sensitivities. A strategic decision is whether we should adopt price discrimination by allocating capacity to different markets and charging different prices in different markets, in which case we gain pricing flexibility, or we should merge all markets and all capacities together, in which case we obtain capacity flexibility. We characterize the region where the benefit of pricing flexibility outweighs the benefit of capacity flexibility. For a hybrid model where we adopt price discrimination at the early stage and switch to joint pricing towards the end of the selling horizon, we develop tractable approximations based on a single-dimensional approximation for each market. Computational experiments indicate that the policy based on our single-dimensional approximation shows significant improvement over the fixed price policy based on a deterministic approximation, especially when the capacity is tight, the number of markets is large, or the price sensitivities of customers in different markets are more dispersed.

There are a couple questions which are worth further investigation. First, in Chapter 4 we were not able to show structural properties of the value function. It is of interest to see if $V_1(\boldsymbol{b})$ belongs to a class of discrete concave functions, in which case tractable methods could be devised to find the optimal capacity allocation vector $\boldsymbol{b}$ among the different markets. Second, the hybrid model is mainly characterized by the deadline τ. In Chapter 4 we treat τ as a given parameter, while it actually could be a decision variable that the seller can control. It is interesting to develop a model that involves τ as a decision variable.

APPENDIX A

APPENDIX FOR CHAPTER 2

A.1 Simplification of the Cost Function

In this section, we show that $C_t(\boldsymbol{\theta}_t) - S_t(\boldsymbol{\theta}_t) + S_{t+1}(\boldsymbol{\theta}_t) = K\,\sum_{i=1}^n c_i\,\theta_{it}\,D_t$. To see that this identity holds, we use the definitions of $C_t(\cdot)$ and $S_t(\cdot)$ to obtain

$$\begin{aligned}
&C_t(\boldsymbol{\theta}_t) - S_t(\boldsymbol{\theta}_t) + S_{t+1}(\boldsymbol{\theta}_t) \\
&= \sum_{i=1}^n \sum_{s=0}^t \mathbf{1}(t-s<K)\,D_s\,c_i - \sum_{i=1}^n \sum_{\ell=t}^\infty \sum_{s=0}^{t-1} c_i\,\mathbf{1}(\ell-s<K)\,D_s\,\theta_{it} \\
&\qquad\qquad + \sum_{i=1}^n \sum_{\ell=t+1}^\infty \sum_{s=0}^{t} c_i\,\mathbf{1}(\ell-s<K)\,D_s\,\theta_{it} \\
&= \sum_{i=1}^n \sum_{\ell=t}^\infty \sum_{s=0}^{t} c_i\,\mathbf{1}(\ell-s<K)\,D_s\,\theta_{it} - \sum_{i=1}^n \sum_{\ell=t}^\infty \sum_{s=0}^{t-1} c_i\,\mathbf{1}(\ell-s<K)\,D_s\,\theta_{it} \\
&= \sum_{i=1}^n \sum_{\ell=t}^\infty c_i\,\mathbf{1}(\ell-t<K)\,D_t\,\theta_{it} = K\sum_{i=1}^n c_i\,\theta_{it}\,D_t,
\end{aligned}$$

where the first equality uses the fact that $W_t = \sum_{s=0}^t \mathbf{1}(t-s<K)\,D_s$.

A.2 Proofs of Structural Properties

Proof of Proposition 2.4.2. Using the notation defined after Lemma 2.4.1, the random variable $Y_{it}(\theta_{it})$ can be written as $\mathsf{Binomial}(W_t, \mathsf{Beta}(\theta_{it}\,M_t, (1-\theta_{it})\,M_t))$. In this case, by using the second part of Lemma 2.4.1 and the discussion that follows this lemma, we observe that the family of random variables $\{Y_{it}(\theta_{it}) : \theta_{it} \in [0,1]\}$ is stochastically increasing and stochastically convex. Thus, the first part of Lemma 2.4.1 implies that $\{\lambda_t\,\theta_{it} + \frac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it}) : \theta_{it} \in [0,1]\}$ is a stochastically increasing and stochastically convex family of random variables. We are now ready to show the desired result by using induction over the time periods. The result trivially holds at time period $\tau+1$. Assuming that the result holds at time period $t+1$, we write the expectation on the right side of (2.2) as

$$\begin{aligned}&\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}\\ &\qquad = \mathbb{E}\{\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\,|\,Y_{jt}(\theta_{jt}) \text{ for all } j \in \{1,\ldots,n\}\setminus\{i\}\}\}.\end{aligned}$$

Since $V_{t+1}(\boldsymbol{\theta}_{t+1})$ is decreasing and convex in $\theta_{i,t+1}$ by the induction argument and the family of random variables $\{\lambda_t\,\theta_{it} + \frac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it}) : \theta_{it} \in [0,1]\}$ is stochastically increasing and stochastically convex, it follows that

$$\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\,|\,Y_{jt}(\theta_{jt}) \text{ for all } j \in \{1,\ldots,n\}\setminus\{i\}\}$$

is a decreasing and convex function of θ_{it} for any realization of the random variables $Y_{jt}(\theta_{jt})$ for $j \in \{1,\ldots,n\}\setminus\{i\}$. Noting that $\{Y_{it}(\theta_{it}) : i = 1,\ldots,n\}$ are independent of each other, taking expectation in the expression above, we get that $\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}$ is a decreasing and convex function of θ_{it}. Therefore, since $(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t$ is a decreasing and linear function of θ_{it}, $(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\}$ is decreasing and convex in θ_{it}. Noting that the pointwise maximum of two decreasing and convex functions is also decreasing and convex, the optimality equation in (2.2) implies that $V_t(\boldsymbol{\theta}_t)$ is decreasing and convex in θ_{it}. □

Proof of Proposition 2.4.4. Since it is optimal to continue at time period t when the state of the system is $\boldsymbol{\theta}_t$, it must be the case that $(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} > 0$. If no units fail at time period t, then the state at time period $t+1$ is $\boldsymbol{\theta}_{t+1} = \lambda_t\,\boldsymbol{\theta}_t$. To obtain a contradiction, we assume that $(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_{t+1})\,D_{t+1} + \mathbb{E}\{V_{t+2}(\lambda_{t+1}\,\boldsymbol{\theta}_{t+1} + \frac{1-\lambda_{t+1}}{W_{t+1}}\,\boldsymbol{Y}_{t+1}(\boldsymbol{\theta}_{t+1}))\} \le 0$ so that it is optimal to stop at time period $t+1$ when the state is $\boldsymbol{\theta}_{t+1} = \lambda_t\, \boldsymbol{\theta}_t$. The last inequality implies that $(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_{t+1})\, D_{t+1} \leq 0$ since we have $V_t(\cdot) \geq 0$ for all $t = 1, \ldots, \tau$ by the optimality equation in (2.2). In this case, we have

$$0 < (r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\, D_t + \mathbb{E}\{V_{t+1}(\lambda_t\, \boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\, \boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} \leq (r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\, D_t + \mathbb{E}\{V_{t+1}(\lambda_t\, \boldsymbol{\theta}_t)\}$$
$$= (r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\, D_t \leq (r - K\lambda_t\, \boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\, D_t = (r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_{t+1})\, D_t \leq 0,$$

which is a contradiction and this completes the proof. The second inequality in the chain of inequalities above follows by noting that $\lambda_t\, \boldsymbol{\theta}_t \leq \lambda_t\, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\, \boldsymbol{Y}_t(\boldsymbol{\theta}_t)$ and using the fact that the value functions are decreasing by Proposition 2.4.2. The first equality follows from the fact that $V_{t+1}(\lambda_t\, \boldsymbol{\theta}_t) = 0$ since we assume that it is optimal to stop at time period $t+1$ when the state is $\lambda_t\, \boldsymbol{\theta}_t$. □

A.3 Simplifying the Computation of Expectations Involving Beta-Binomial Random Variables

In this section, we show that by replacing the beta-binomial random variable $Y_{it}(\theta_{it})$ in the optimality equation (2.2) with a binomial random variable $Z_{it}(\theta_{it})$ with parameters (W_t, θ_{it}), we obtain lower bounds on the value functions. Using the vector $\boldsymbol{Z}_t(\boldsymbol{\theta}_t) = (Z_{1t}(\theta_{1t}), \ldots, Z_{nt}(\theta_{nt}))$, we consider the optimality equation

$$\bar{V}_t(\boldsymbol{\theta}_t) = \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\, D_t + \mathbb{E}\{\bar{V}_{t+1}(\lambda_t\, \boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\, \boldsymbol{Z}_t(\boldsymbol{\theta}_t))\}, 0\Big\}, \qquad \text{(A.1)}$$

with the boundary condition $\bar{V}_{\tau+1}(\cdot) = 0$, and the expectation is computed with respect to the binomial random variable $\boldsymbol{Z}_t(\boldsymbol{\theta}_t)$. The main result of this section is the following proposition, which shows that the value function $\bar{V}_t(\cdot)$ provides a lower bound on the original value function.

Proposition A.3.1. *For all $t = 1, \dots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$, we have $V_t(\boldsymbol{\theta}_t) \geq \bar{V}_t(\boldsymbol{\theta}_t)$.*

The proof of the above result makes use of the following general form of Jensen's inequality for componentwise convex functions.

Lemma A.3.2. *If $g(\cdot) : \Re^n \to \Re$ is a componentwise convex function and $\boldsymbol{X} = (X_1, \dots, X_n)$ is a random variable taking values in $\Re^n$ with independent components, then $\mathbb{E}\{g(\boldsymbol{X})\} \geq g(\mathbb{E}\{\boldsymbol{X}\})$.*

Proof. We prove the result by induction on n. When $n = 1$, we have a scalar convex function and the result trivially holds by the standard form of Jensen's inequality. Assuming that the result holds when we deal with functions that map $\Re^{n-1}$ to $\Re$, we have

$$\mathbb{E}\{g((X_1, \dots, X_n))\} = \mathbb{E}\{\mathbb{E}\{g((X_1, \dots, X_{n-1}, X_n)) \,|\, X_1, \dots, X_{n-1}\}\}.$$

Since $g(\cdot)$ is componentwise convex, $g((X_1, \dots, X_{n-1}, X_n))$ is a convex function of X_n and applying Jensen's inequality on the scalar convex function $g(X_1, \dots, X_{n-1}, \cdot)$, we obtain

$$\mathbb{E}\{g((X_1, \dots, X_{n-1}, X_n)) \,|\, X_1, \dots, X_{n-1}\} \geq g((X_1, \dots, X_{n-1}, \mathbb{E}\{X_n\})),$$

where we use the fact that the distribution of X_n conditional on $X_1, \dots, X_{n-1}$ is the same as the unconditional distribution of X_n. Viewing $g(\cdot, \dots, \cdot, \mathbb{E}\{X_n\})$ on the right side of the inequality above as a function that maps $\Re^{n-1}$ to $\Re$, by the induction assumption, we have

$$\mathbb{E}\{g((X_1, \dots, X_{n-1}, \mathbb{E}\{X_n\}))\} \geq g((\mathbb{E}\{X_1\}, \dots, \mathbb{E}\{X_{n-1}\}, \mathbb{E}\{X_n\})),$$

so that we obtain

$$\begin{aligned}\mathbb{E}\{g(\boldsymbol{X})\} &= \mathbb{E}\{\mathbb{E}\{g((X_1,\ldots,X_{n-1},X_n))\,|\,X_1,\ldots,X_{n-1}\}\} \\ &\geq \mathbb{E}\{g((X_1,\ldots,X_{n-1},\mathbb{E}\{X_n\}))\} \\ &\geq g((\mathbb{E}\{X_1\},\ldots,\mathbb{E}\{X_{n-1}\},\mathbb{E}\{X_n\})) = g(\mathbb{E}\{\boldsymbol{X}\}).\end{aligned}$$

□

Proof of Proposition A.3.1. We show the result by using induction over the time periods. The result trivially holds at time period $\tau+1$. Assuming that the result holds at time period $t+1$, we have

$$\begin{aligned}\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} &= \mathbb{E}\{\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\,|\,\boldsymbol{P}_t(\boldsymbol{\theta}_t)\}\} \\ &= \mathbb{E}\{G_{t+1}(\boldsymbol{P}_t(\boldsymbol{\theta}_t))\},\end{aligned} \tag{A.2}$$

where we use the vector $\boldsymbol{P}_t(\boldsymbol{\theta}_t) = (P_{1t}(\theta_{1t}),\ldots,P_{nt}(\theta_{nt}))$ and define the function $G_{t+1}(\cdot):[0,1]^n \to \Re$ as $G_{t+1}(\boldsymbol{p}) = \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\,|\,\boldsymbol{P}_t(\boldsymbol{\theta}_t) = \boldsymbol{p}\}$. Conditional on $P_{it}(\theta_{it}) = p_i$, the random variable $Y_{it}(\theta_{it})$ has a binomial distribution with parameters (W_t, p_i). Therefore, noting that the family of random variables $\{\mathsf{Binomial}(W_t, p_i) : p_i \in [0,1]\}$ is stochastically convex and $V_{t+1}(\cdot)$ is componentwise convex by Proposition 2.4.2, $G_{t+1}(\boldsymbol{p}) = \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\,|\,\boldsymbol{P}_t(\boldsymbol{\theta}_t) = \boldsymbol{p}\}$ is componentwise convex in $\boldsymbol{p}$. In this case, Jensen's inequality in Lemma A.3.2 implies that $\mathbb{E}\{G_{t+1}(\boldsymbol{P}_t(\boldsymbol{\theta}_t))\} \geq G_{t+1}(\mathbb{E}\{\boldsymbol{P}_t(\boldsymbol{\theta}_t)\}) = G_{t+1}(\boldsymbol{\theta}_t)$ and we continue the chain of equalities in (A.2) as

$$\begin{aligned}&\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} \\ \geq& G_{t+1}(\boldsymbol{\theta}_t) = \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\,|\,\boldsymbol{P}_t(\boldsymbol{\theta}_t) = \boldsymbol{\theta}_t\} \\ =& \mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Z}_t(\boldsymbol{\theta}_t))\} \geq \mathbb{E}\{\bar{V}_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{W_t}\,\boldsymbol{Z}_t(\boldsymbol{\theta}_t))\},\end{aligned}$$

where the second equality follows from the fact that conditional on $\boldsymbol{P}_t(\boldsymbol{\theta}_t) = \boldsymbol{\theta}_t$, the random variable $\boldsymbol{Y}_t(\boldsymbol{\theta}_t)$ has the same distribution as $\boldsymbol{Z}_t(\boldsymbol{\theta}_t)$ and the second inequality follows from the induction assumption. Therefore, we have $\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Y}_t(\boldsymbol{\theta}_t))\} \geq \mathbb{E}\{\bar{V}_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \frac{1-\lambda_t}{W_t}\,\boldsymbol{Z}_t(\boldsymbol{\theta}_t))\}$ and noting the optimality equations in (2.2) and (A.1), it follows that $V_t(\boldsymbol{\theta}_t) \geq \bar{V}_t(\boldsymbol{\theta}_t)$. $\square$

A.4 Convexity of the Upper Bound

In this section, we establish that $V^U_{it}(\theta_{it}\,|\,\boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$ and show how to compute a subgradient of $V^U_{it}(\theta_{it}\,|\,\boldsymbol{\rho})$ with respect to $\boldsymbol{\rho}$. We begin by using induction over the time periods to show that $V^U_{it}(\theta_{it}\,|\,\boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$. The result trivially holds at time period $\tau+1$. Assuming that the result holds at time period $t+1$, it follows that $V^U_{i,t+1}(\lambda_t\,\theta_{it} + \frac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$ for every realization of $Y_{it}(\theta_{it})$. Therefore, $(\rho_i - K c_i\,\theta_{it})\,D_t + \mathbb{E}\{V^U_{i,t+1}(\lambda_t\,\theta_{it} + \frac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}$ is a convex function of $\boldsymbol{\rho}$. Since the pointwise maximum of two convex functions is also convex, (2.3) implies that $V^U_{it}(\theta_{it}\,|\,\boldsymbol{\rho})$ is a convex function of $\boldsymbol{\rho}$.

We need to define some new notation to compute the subgradient of $V^U_{it}(\theta_{it}\,|\,\boldsymbol{\rho})$ with respect to $\boldsymbol{\rho}$. We let $\mathbf{1}^*_{it}(\cdot\,|\,\boldsymbol{\rho}) : [0,1] \to \{0,1\}$ be the decision function at time period t that we obtain by solving the optimality equation in (2.3). Then we have $\mathbf{1}^*_{it}(\theta_{it}\,|\,\boldsymbol{\rho}) = 1$ whenever $(\rho_i - K c_i\,\theta_{it})\,D_t + \mathbb{E}\{V^U_{i,t+1}(\lambda_t\,\theta_{it} + \frac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\} > 0$ and $\mathbf{1}^*_{it}(\theta_{it}\,|\,\boldsymbol{\rho}) = 0$ otherwise. In the optimality equation in (2.3), if we start with state θ_{it} at time period t and use $\Delta_{it}(\theta_{it}\,|\,\boldsymbol{\rho})$ to denote the total expected demand that we observe until we stop selling the product, then

$\Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho})$ satisfies the recursion

$$\Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) = \mathbf{1}^*_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) \left\{ D_t + \mathbb{E}\{\Delta_{i,t+1}(\lambda_t \, \theta_{it} + \tfrac{1-\lambda_t}{W_t} \, Y_{it}(\theta_{it}) \,|\, \boldsymbol{\rho})\} \right\}, \tag{A.3}$$

with the boundary condition $\Delta_{i,\tau+1}(\cdot \,|\, \boldsymbol{\rho}) = 0$. In the rest of this section, we use induction over the time periods to show that $V^U_{it}(\theta_{it} \,|\, \cdot)$ satisfies the subgradient inequality

$$V^U_{it}(\theta_{it} \,|\, \hat{\boldsymbol{\rho}}) \geq V^U_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) + \Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) \, [\hat{\rho}_i - \rho_i], \tag{A.4}$$

which implies that $\Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho})$ is a subgradient of $V^U_{it}(\theta_{it} \,|\, \boldsymbol{\rho})$ with respect to $\boldsymbol{\rho}$. If we start with state θ_{it} at time period t and follow the decision function $\mathbf{1}^*_{is}(\cdot \,|\, \boldsymbol{\rho})$ at time periods $s = t, t+1, \ldots, \tau$, then $\Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho})$ corresponds to the total expected demand that we observe until we stop selling the product. Therefore, we can use Monte Carlo simulation to estimate $\Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho})$. Alternatively, noting that the random variable $Y_{it}(\theta_{it})$ has a finite number of realizations, we can compute $\Delta_{it}(\theta_{it} \,|\, \boldsymbol{\rho})$ by computing $\Delta_{is}(\cdot \,|\, \boldsymbol{\rho})$ for a finite number of values at time periods $s = t+1, t+2, \ldots, \tau$.

The subgradient inequality in (A.4) trivially holds at time period $\tau + 1$. Assuming that this subgradient inequality holds at time period $t+1$, we write the optimality equation in (2.3) as

$$V^U_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) = \mathbf{1}^*_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) \left\{ (\rho_i - K c_i \, \theta_{it}) \, D_t + \mathbb{E}\{V^U_{i,t+1}(\lambda_t \, \theta_{it} + \tfrac{1-\lambda_t}{W_t} \, Y_{it}(\theta_{it}) \,|\, \boldsymbol{\rho})\} \right\}.$$

On the other hand, since we have $\mathbf{1}^*_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) \in \{0, 1\}$, if we solve the optimality equation in (2.3) after replacing $\boldsymbol{\rho}$ with $\hat{\boldsymbol{\rho}}$, then the value function $V^U_{it}(\theta_{it} \,|\, \hat{\boldsymbol{\rho}})$ satisfies

$$V^U_{it}(\theta_{it} \,|\, \hat{\boldsymbol{\rho}}) \geq \mathbf{1}^*_{it}(\theta_{it} \,|\, \boldsymbol{\rho}) \left\{ (\hat{\rho}_i - K c_i \, \theta_{it}) \, D_t + \mathbb{E}\{V^U_{i,t+1}(\lambda_t \, \theta_{it} + \tfrac{1-\lambda_t}{W_t} \, Y_{it}(\theta_{it}) \,|\, \hat{\boldsymbol{\rho}})\} \right\}.$$

Subtracting the last equality from the last inequality side by side, we obtain

$$V^U_{it}(\theta_{it}\,|\,\hat{\boldsymbol{\rho}}) \geq V^U_{it}(\theta_{it}\,|\,\boldsymbol{\rho}) + \mathbf{1}^*_{it}(\theta_{it}\,|\,\boldsymbol{\rho}) \Big\{ D_t\,[\hat{\rho}_i - \rho_i]$$
$$+\, \mathbb{E}\{V^U_{i,t+1}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\hat{\boldsymbol{\rho}})\} - \mathbb{E}\{V^U_{i,t+1}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}\Big\}. \quad \text{(A.5)}$$

The induction assumption implies that

$$\mathbb{E}\{V^U_{i,t+1}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\hat{\boldsymbol{\rho}})\} - \mathbb{E}\{V^U_{i,t+1}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}$$
$$\geq \mathbb{E}\{\Delta_{i,t+1}(\lambda_t\,\theta_{it} + \tfrac{1-\lambda_t}{W_t}\,Y_{it}(\theta_{it})\,|\,\boldsymbol{\rho})\}\,[\hat{\rho}_i - \rho_i], \quad \text{(A.6)}$$

in which case, the subgradient inequality in (A.4) follows by using (A.6) and (A.3) in (A.5).

A.5 Closed Form Expression for the Lower Bound

In this section, we use induction over the time periods to show that the value functions computed through the optimality equation in (2.5) are given by the closed form expression in (2.6). The result trivially holds at time period $\tau + 1$. Assuming that the result holds at time period $t + 1$, the optimality equation in (2.5) implies that

$$\begin{aligned} V^L_t(\boldsymbol{\theta}_t) &= \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,D_t + [r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \sum_{s=t+1}^{\tau} D_s, 0\Big\} \\ &= [r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \sum_{s=t}^{\tau} D_s, \end{aligned}$$

where the second equality follows by noting that the $\max$ operator returns a nonzero value if and only if $r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t > 0$.

A.6 Shrinking $\mathcal{C}_t(m)$ as a Function of m

In this section, our ultimate goal is to give a proof for Proposition 2.6.2. This proof requires establishing results for stochastic convex orders that do not appear in the earlier literature and we begin this section with two new results for stochastic convex orders. For two random variables X and Y, we say that X is greater than or equal to Y in stochastic convex order whenever $\mathbb{E}\{\phi(Y)\} \leq \mathbb{E}\{\phi(X)\}$ for any convex function $\phi(\cdot) : \Re \to \Re$. We denote this stochastic convex order between X and Y by using $Y \leq_{cx} X$. The following two lemmas show convex orders for the families of random variables $\{\mathsf{Binomial}(W, \theta)/W : W = 1, 2, \ldots\}$ and $\{\mathsf{Binomial}(mW, \mathsf{Beta}(\theta\, mM, (1-\theta)\, mM))/mW : m = 1, 2, \ldots\}$ for any fixed $\theta \in [0, 1]$. These lemmas are useful for proving Proposition 2.6.2, but they may also have independent interest.

Lemma A.6.1. *For any fixed $\theta \in [0, 1]$, letting $X(W) = \mathsf{Binomial}(W, \theta)/W$, we have $X(W + 1) \leq_{cx} X(W)$.*

Proof. Throughout the proof, we use $=_{st}$ to denote equality in distribution. By Theorem 3.A.4 in Shaked and Shanthikumar (2007), it suffices to construct a pair of random variables $\hat{X}(W + 1)$ and $\hat{X}(W)$ such that $\hat{X}(W + 1) =_{st} X(W + 1)$, $\hat{X}(W) =_{st} X(W)$ and $\{\hat{X}(W + 1), \hat{X}(W)\}$ is a martingale.

We let $\hat{X}(W + 1) = X(W + 1) = \mathsf{Binomial}(W + 1, \theta)/(W + 1)$, in which case, $\hat{X}(W+1)$ takes values in the set $\{0, 1/(W + 1), \ldots, W/(W + 1), 1\}$. We construct $\hat{X}(W)$ in the following way. If $\hat{X}(W+1) = 0$ or 1, then we set $\hat{X}(W) = \hat{X}(W+1)$. If, on the other hand, $\hat{X}(W + 1) = j/(W + 1)$ for some $j = 1, \ldots, W$, then we set $\hat{X}(W) = (j - 1)/W$ with probability $\frac{j/W - j/(W+1)}{1/W}$ and $\hat{X}(W) = j/W$ with probability $\frac{j/(W+1) - (j-1)/W}{1/W}$. One can check that these probabilities add up to one.

We proceed to show $\hat{X}(W) =_{st} X(W)$. To see this equivalence in distribution, we note that

$$
\begin{aligned}
&\mathbb{P}\{\hat{X}(W) = 0\} \\
=&\mathbb{P}\{\hat{X}(W) = 0 \,|\, \hat{X}(W+1) = 0\}\, \mathbb{P}\{\hat{X}(W+1) = 0\} \\
&\quad + \mathbb{P}\{\hat{X}(W) = 0 \,|\, \hat{X}(W+1) = 1/(W+1)\}\, \mathbb{P}\{\hat{X}(W+1) = 1/(W+1)\} \\
=&1\binom{W+1}{0}(1-\theta)^{W+1} + \frac{1/W - 1/(W+1)}{1/W}\binom{W+1}{1}\theta\,(1-\theta)^{W} \\
=&(1-\theta)^{W+1} + \frac{1}{W+1}\,(W+1)\,\theta\,(1-\theta)^{W} = (1-\theta)^{W}.
\end{aligned}
$$

Similarly, we can show that $\mathbb{P}\{\hat{X}(W) = 1\} = \theta^{W}$. On the other hand, for $j = 1, \ldots, W-1$, the definition of $\hat{X}(W)$ implies that

$$
\begin{aligned}
\mathbb{P}\{\hat{X}(W) = \tfrac{j}{W}\} &= \mathbb{P}\{\hat{X}(W) = \tfrac{j}{W} \,|\, \hat{X}(W+1) = \tfrac{j}{W+1}\}\, \mathbb{P}\{\hat{X}(W+1) = \tfrac{j}{W+1}\} \\
&\quad + \mathbb{P}\{\hat{X}(W) = \tfrac{j}{W} \,|\, \hat{X}(W+1) = \tfrac{j+1}{W+1}\}\, \mathbb{P}\{\hat{X}(W+1) = \tfrac{j+1}{W+1}\} \\
&= \frac{j/(W+1) - (j-1)/W}{1/W}\binom{W+1}{j}\theta^{j}\,(1-\theta)^{W+1-j} \\
&\quad + \frac{(j+1)/W - (j+1)/(W+1)}{1/W}\binom{W+1}{j+1}\theta^{j+1}\,(1-\theta)^{W-j} \\
&= \frac{W-j+1}{W+1}\left[\frac{(W+1)W\ldots(W-j+2)}{j!}\right]\theta^{j}\,(1-\theta)^{W+1-j} \\
&\quad + \frac{j+1}{W+1}\left[\frac{(W+1)W\ldots(W-j+1)}{(j+1)j!}\right]\theta^{j+1}\,(1-\theta)^{W-j} \\
&= \binom{W}{j}\theta^{j}\,(1-\theta)^{W-j}\,(1-\theta+\theta) = \binom{W}{j}\theta^{j}\,(1-\theta)^{W-j}.
\end{aligned}
$$

Therefore, we have $\hat{X}(W) =_{st} \mathsf{Binomial}(W,\theta)/W =_{st} X(W)$.

It remains to show that $\{\hat{X}(W+1), \hat{X}(W)\}$ is a martingale. Correspondingly, if $\hat{X}(W+1) = 0$ or 1, then we naturally have $\mathbb{E}\{\hat{X}(W)|\hat{X}(W+1)\} = \hat{X}(W+1)$ and the martingale equality holds. If, on the other hand, $\hat{X}(W+1) = j/(W+1)$

for some $j = 1, \ldots, W$, then we have

$$\begin{aligned}
&\mathbb{E}\{\hat{X}(W)|\hat{X}(W+1)\}\\
=&\frac{j-1}{W}\left[\frac{j/W - j/(W+1)}{1/W}\right] + \frac{j}{W}\left[\frac{j/(W+1) - (j-1)/W}{1/W}\right]\\
=&\frac{j}{W+1} = \hat{X}(W+1).
\end{aligned}$$

Therefore, $\{\hat{X}(W+1), \hat{X}(W)\}$ is indeed a martingale. $\square$

Corollary A.6.2. *For any fixed* $\theta \in [0,1]$*, let* $X(m) = \frac{Y^m(\theta)}{mW}$*, where* $Y^m(\theta) = \mathsf{Binomial}(mW, P^m(\theta))$ *and* $P^m(\theta) = \mathsf{Beta}(\theta\, mM, (1-\theta)\, mM)$*, then it holds that* $X(m+1) \leq_{cx} X(m)$ *for all* $m = 1, 2, \ldots$*.*

Proof. The result follows by noting that for any convex function $\phi(\cdot) : \Re \to \Re$, we have

$$\begin{aligned}
\mathbb{E}\Big\{\phi\Big(\frac{Y^m(\theta)}{mW}\Big)\Big\} &= \mathbb{E}\Big\{\phi\Big(\frac{\mathsf{Binomial}(mW, P^m(\theta))}{mW}\Big)\Big\}\\
&\geq \mathbb{E}\Big\{\phi\Big(\frac{\mathsf{Binomial}(mW, P^{m+1}(\theta))}{mW}\Big)\Big\}\\
&\geq \mathbb{E}\Big\{\phi\Big(\frac{\mathsf{Binomial}((m+1)W, P^{m+1}(\theta))}{(m+1)W}\Big)\Big\}\\
&= \mathbb{E}\Big\{\phi\Big(\frac{Y^{m+1}(\theta)}{(m+1)W}\Big)\Big\}. \qquad \text{(A.7)}
\end{aligned}$$

To see that the first inequality holds, for $p \in [0,1]$, we let $g(p) = \mathbb{E}\left\{\phi\left(\frac{Y^m(\theta)}{mW}\right) \mid P^m(\theta) = p\right\}$ so that the expression on the left side of the first inequality can be written as $\mathbb{E}\{g(P^m(\theta))\}$. Therefore, the first inequality can be written as $\mathbb{E}\{g(P^m(\theta))\} \geq \mathbb{E}\{g(P^{m+1}(\theta))\}$. To see that this last inequality holds, conditional on $P^m(\theta) = p$, $Y^m(\theta)$ has a binomial distribution with parameters (mW, p). Due to stochastic convexity of the binomial family $\{\mathsf{Binomial}(mW, p) : p \in [0,1]\}$, we get that $g(\cdot)$ is also a convex function. In this case, the first inequality in (A.7) follows from the monotone convergence property of the beta operator under convex functions shown by Adell, Badia, de la Cal, and Plo (1996). The second inequality in (A.7) follows by conditioning on $P^{m+1}(\theta)$ and applying the above lemma W times.

$\square$

Proof of Proposition 2.6.2. The proposition states that if $\boldsymbol{\theta}_t \in \mathcal{C}_t(m+1)$, then we have $\boldsymbol{\theta}_t \in \mathcal{C}_t(m)$. Noting the definition of $\mathcal{C}_t(m)$, it suffices to show that

$$\frac{1}{m} V_t(\boldsymbol{\theta}_t \,|\, m) \geq \frac{1}{m+1} V_t(\boldsymbol{\theta}_t \,|\, m+1)$$

for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$. We show this inequality by using induction over the time periods. The result trivially holds at time period $\tau + 1$. Assuming that the result holds at time period $t+1$, we have

$$\begin{aligned} &\frac{1}{m}\,\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{mW_t}\,\boldsymbol{Y}_t^m(\boldsymbol{\theta}_t) \,|\, m)\} \\ \geq &\frac{1}{m+1}\,\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{mW_t}\,\boldsymbol{Y}_t^m(\boldsymbol{\theta}_t) \,|\, m+1)\} \\ \geq &\frac{1}{m+1}\,\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{(m+1)W_t}\,\boldsymbol{Y}_t^{m+1}(\boldsymbol{\theta}_t) \,|\, m+1)\}. \end{aligned} \tag{A.8}$$

The first inequality above follows from the induction assumption. To see that the second inequality holds, we note that $V_{t+1}(\cdot \,|\, m+1)$ is a componentwise convex function by Proposition 2.4.2, in which case, we can apply Corollary A.6.2 component by component n times. To obtain the final result in the statement of Proposition 2.6.2, we observe that

$$\begin{aligned} &\frac{1}{m}\,V_t(\boldsymbol{\theta}_t \,|\, m) = \max\Big\{(r - K\boldsymbol{c}^\top\boldsymbol{\theta}_t)D_t + \frac{1}{m}\,\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{mW_t}\,\boldsymbol{Y}_t^m(\boldsymbol{\theta}_t) \,|\, m)\}, 0\Big\} \\ &\geq \max\Big\{(r - K\boldsymbol{c}^\top\boldsymbol{\theta}_t)D_t + \frac{1}{m+1}\,\mathbb{E}\{V_{t+1}(\lambda_t\,\boldsymbol{\theta}_t + \tfrac{1-\lambda_t}{(m+1)W_t}\,\boldsymbol{Y}_t^{m+1}(\boldsymbol{\theta}_t) \,|\, m+1)\}, 0\Big\} \\ &= \frac{1}{m+1} V_t(\boldsymbol{\theta}_t \,|\, m+1), \end{aligned}$$

where the two equalities follow from the optimality equation in (2.2) and the inequality uses (A.8). $\square$

A.7 Proof of Proposition 2.6.3

In this section, we give a proof for Proposition 2.6.3, which says that $V_t(\boldsymbol{\theta}_t \,|\, m)$ deviates from $V_t^L(\boldsymbol{\theta}_t \,|\, m)$ by a term that grows in the order of $\sqrt{m}$.

Proof. Since Proposition 2.6.1 shows the first inequality, we only focus on the second inequality. The proof uses induction over the time periods to show that $V_t(\boldsymbol{\theta}_t \,|\, m) \leq V_t^L(\boldsymbol{\theta}_t \,|\, m) + G_t(\boldsymbol{\theta}_t)\sqrt{m}$, where $G_t(\boldsymbol{\theta}_t)$ is a componentwise concave function of $\boldsymbol{\theta}_t$, which does not depend on m. The result trivially holds at time period $\tau + 1$. Assuming that the result holds at time period $t + 1$, we define the random variable $\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t)$ as $\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t) = \lambda_t \, \boldsymbol{\theta}_t + \frac{1-\lambda_t}{mW_t} \, \boldsymbol{Y}_t^m(\boldsymbol{\theta}_t)$. In this case, we obtain

$$
\begin{aligned}
\mathbb{E}\{V_{t+1}(\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t) \,|\, m)\} &\leq \mathbb{E}\{V_{t+1}^L(\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t) \,|\, m)\} + \mathbb{E}\{G_{t+1}(\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t))\} \sqrt{m} \\
&\leq \mathbb{E}\{V_{t+1}^L(\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t) \,|\, m)\} + G_{t+1}(\mathbb{E}\{\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t)\}) \sqrt{m} \\
&= \mathbb{E}\{[r - K\boldsymbol{c}^\top \boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t)]^+\} \sum_{s=t+1}^{\tau} mD_s + G_{t+1}(\boldsymbol{\theta}_t) \sqrt{m},
\end{aligned}
\tag{A.9}
$$

where the first inequality follows from the induction assumption, the second inequality follows by noting that $G_{t+1}(\cdot)$ is a componentwise concave function and using the general form of Jensen's inequality for componentwise convex functions that we derive in Appendix A.3 and the last equality follows by using the closed form expression for $V_{t+1}^L(\cdot \,|\, m)$ given in (2.6) and noting that $\mathbb{E}\{\boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t)\} = \boldsymbol{\theta}_t$.

We proceed to bound the expectation $\mathbb{E}\{[r - K\boldsymbol{c}^\top \boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t)]^+\}$ on the right side of (A.9). The variance of the random variable $K\boldsymbol{c}^\top \boldsymbol{\Theta}_{t+1}^m(\boldsymbol{\theta}_t)$ can be bounded by noting that

$$
\begin{aligned}
&\mathrm{Var}(K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t)) \\
&= K^2 \sum_{i=1}^{n} c_i^2 \,\mathrm{Var}(\boldsymbol{\Theta}^m_{i,t+1}(\theta_{it})) = K^2\,\frac{(1-\lambda_t)^2}{m^2 W_t^2} \sum_{i=1}^{n} c_i^2 \,\mathrm{Var}(Y^m_{it}(\theta_{it})) \qquad \text{(A.10)} \\
&= K^2\,\frac{(1-\lambda_t)^2}{m^2 W_t^2} \sum_{i=1}^{n} c_i^2 \frac{mW_t\,(mM_t\,\theta_{it})\,mM_t\,(1-\theta_{it})(mM_t+mW_t)}{(mM_t)^2\,(mM_t+1)} \\
&\le K^2\,\frac{(1-\lambda_t)^2(M_t+W_t)}{mW_t\,M_t} \sum_{i=1}^{n} c_i^2\,\theta_{it}\,(1-\theta_{it}),
\end{aligned}
$$

where the first equality follows because components of $\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t)$ are independent, the third equality uses the fact that $Y^m_{it}(\theta_{it})$ is a beta-binomial random variable with parameters $(mW_t, P^m_{it}(\theta_{it}))$ and $P^m_{it}(\theta_{it})$ is a beta random variable with parameters $(\theta_{it}\,mM_t, (1-\theta_{it})\,mM_t)$, and the inequality follows simply by noting that $mM_t + 1 \ge mM_t$. For a deterministic scalar z and a real-valued random variable Z with finite mean μ and finite variance σ^2, Gallego (1992) shows that $\mathbb{E}\{[z-Z]^+\} \le [\sqrt{\sigma^2+(z-\mu)^2}+(z-\mu)]/2 \le [z-\mu]^+ + \sigma/2$. Therefore, we can bound the expectation $\mathbb{E}\{[r-K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t)]^+\}$ by

$$
\begin{aligned}
&\mathbb{E}\{[r-K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t)]^+\} \\
&\le [r-K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\theta}_t]^+ + \frac{1}{2}\sqrt{K^2\,\frac{(1-\lambda_t)^2(M_t+W_t)}{mW_t\,M_t}\sum_{i=1}^{n} c_i^2\,\theta_{it}\,(1-\theta_{it})} \\
&= [r-K\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\theta}_t]^+ + \frac{L_t(\boldsymbol{\theta}_t)}{\sqrt{m}},
\end{aligned}
$$

where we let $L_t(\boldsymbol{\theta}_t) = \frac{K\,(1-\lambda_t)}{2}\sqrt{\frac{(M_t+W_t)}{W_t M_t}\sum_{i=1}^{n} c_i^2\,\theta_{it}\,(1-\theta_{it})}$. In this case, we can continue the chain of inequalities in (A.9) as

$$\begin{aligned}
&\mathbb{E}\{V_{t+1}(\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t) \,|\, m)\} \\
&\leq \Big[[r - K\,\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ + \frac{L_t(\boldsymbol{\theta}_t)}{\sqrt{m}}\Big] \sum_{s=t+1}^{\tau} mD_s + G_{t+1}(\boldsymbol{\theta}_t)\,\sqrt{m} \\
&= [r - K\,\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \sum_{s=t+1}^{\tau} mD_s + \sqrt{m}\,L_t(\boldsymbol{\theta}_t) \sum_{s=t+1}^{\tau} D_s + G_{t+1}(\boldsymbol{\theta}_t)\,\sqrt{m}. \\
&= [r - K\,\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \sum_{s=t+1}^{\tau} mD_s + G_t(\boldsymbol{\theta}_t)\,\sqrt{m},
\end{aligned}$$

where we let $G_t(\boldsymbol{\theta}_t) = L_t(\boldsymbol{\theta}_t)\sum_{s=t+1}^{\tau} D_s + G_{t+1}(\boldsymbol{\theta}_t)$. We note that $L_t(\cdot)$ is componentwise concave and since $G_{t+1}(\cdot)$ is componentwise concave by the induction assumption, it follows that $G_t(\cdot)$ is also componentwise concave. To finish the proof, we write the optimality equation in (2.2) as

$$\begin{aligned}
V_t(\boldsymbol{\theta}_t \,|\, m) &= \max\Big\{(r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t)\,mD_t + \mathbb{E}\{V_{t+1}(\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t) \,|\, m)\}, 0\Big\} \\
&\leq [r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+\,mD_t + \mathbb{E}\{V_{t+1}(\boldsymbol{\Theta}^m_{t+1}(\boldsymbol{\theta}_t) \,|\, m)\} \\
&\leq [r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+\,mD_t + [r - K\,\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \sum_{s=t+1}^{\tau} mD_s + G_t(\boldsymbol{\theta}_t)\,\sqrt{m} \\
&= V_t^L(\boldsymbol{\theta}_t \,|\, m) + G_t(\boldsymbol{\theta}_t)\,\sqrt{m},
\end{aligned}$$

where the first inequality follows by noting that $[r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t]^+ \geq r - K\boldsymbol{c}^\mathsf{T}\boldsymbol{\theta}_t$ and $V_{t+1}(\cdot) \geq 0$ and the second equality follows by noting (2.6). This completes the induction argument and we have $V_t(\boldsymbol{\theta}_t \,|\, m) \leq V_t^L(\boldsymbol{\theta}_t \,|\, m) + G_t(\boldsymbol{\theta}_t)\sqrt{m}$ for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$.

To obtain the result in the statement of the proposition, we let $\bar{\frac{1}{2}} = (\frac{1}{2}, \ldots, \frac{1}{2}) \in \Re^n$ and note that $L_t(\boldsymbol{\theta}_t) \leq L_t(\bar{\frac{1}{2}})$ for all $t = 1, \ldots, \tau$ and $\boldsymbol{\theta}_t \in [0,1]^n$. In this case, we obtain $G_t(\boldsymbol{\theta}_t) \leq G_t(\bar{\frac{1}{2}})$ and letting $\bar{G}_t = G_t(\bar{\frac{1}{2}})$ in the statement of the proposition suffices. □

APPENDIX B

APPENDIX FOR CHAPTER 3

B.1 Proof of Proposition 3.3.1

In this section we provide proofs for Proposition 3.3.1.

Proof of (i). We show this result by induction over the time periods. The result trivially holds at time $\tau+1$ since $V_{\tau+1}(x,y) = \Psi_{\tau+1}(x+y)$, which is a concave function. Assuming that the result holds at time period $t+1$, define $f^*_{ij} = \arg\max_{d_{st}\in[0,1]} \Big\{ r_s(d_{st}) - d_{st}\, \Delta_x V_{t+1}(x+i, y+j) \Big\}, i, j \in \{0,1,2\}$, we have

$$
\begin{aligned}
&\Delta_x V_t(x+2,y) - \Delta_x V_t(x+1,y) \\
&= V_t(x+2,y) - V_t(x+1,y) - V_t(x+1,y) + V_t(x,y) \\
&= \lambda_{st}[r_s(f^*_{20}) - f^*_{20}\,\Delta_x V_{t+1}(x+2,y)] + \lambda_{ct}[r_c(d_c) + d_c\,V_{t+1}(x+2,y-1)] \\
&\qquad\qquad + (1-\lambda_{ct}\,d_c)V_{t+1}(x+2,y) \\
&\quad - \{\lambda_{st}[r_s(f^*_{10}) - f^*_{10}\,\Delta_x V_{t+1}(x+1,y)] + \lambda_{ct}[r_c(d_c) + d_c\,V_{t+1}(x+1,y-1)] \\
&\qquad\qquad +(1-\lambda_{ct}\,d_c)V_{t+1}(x+1,y)\} \\
&\quad - \{\lambda_{st}[r_s(f^*_{10}) - f^*_{10}\,\Delta_x V_{t+1}(x+1,y)] + \lambda_{ct}[r_c(d_c) + d_c\,V_{t+1}(x+1,y-1)] \\
&\qquad\qquad +(1-\lambda_{ct}\,d_c)V_{t+1}(x+1,y)\} \\
&\quad + \lambda_{st}[r_s(f^*_{00}) - f^*_{00}\,\Delta_x V_{t+1}(x,y)] + \lambda_{ct}[r_c(d_c) + d_c\,V_{t+1}(x,y-1)] \\
&\qquad\qquad + (1-\lambda_{ct}\,d_c)V_{t+1}(x,y).
\end{aligned}
$$

By combining and rearranging terms, we can continue the equality as

$$
\begin{aligned}
&\Delta_x V_t(x+2,y) - \Delta_x V_t(x+1,y) \\
&= \lambda_{st}[r_s(f_{20}^*) - f_{20}^*\,\Delta_x V_{t+1}(x+2,y)] - \lambda_{st}[r_s(f_{10}^*) - f_{10}^*\,\Delta_x V_{t+1}(x+1,y)] \\
&\quad - \lambda_{st}[r_s(f_{10}^*) - f_{10}^*\,\Delta_x V_{t+1}(x+1,y)] + \lambda_{st}[r_s(f_{00}^*) - f_{00}^*\,\Delta_x V_{t+1}(x,y)] \\
&\quad + \lambda_{ct}\,d_c[\Delta_x V_{t+1}(x+2,y-1) - \Delta_x V_{t+1}(x+1,y-1)] \\
&\quad + (1-\lambda_{ct}\,d_c)[\Delta_x V_{t+1}(x+2,y) - \Delta_x V_{t+1}(x+1,y)].
\end{aligned}
$$

Due to the definition of f_{10}^*, the above equation is upper bounded by

$$
\begin{aligned}
&\lambda_{st}[r_s(f_{20}^*) - f_{20}^*\,\Delta_x V_{t+1}(x+2,y)] - \lambda_{st}[r_s(f_{20}^*) - f_{20}^*\,\Delta_x V_{t+1}(x+1,y)] \\
&\quad - \lambda_{st}[r_s(f_{00}^*) - f_{00}^*\,\Delta_x V_{t+1}(x+1,y)] + \lambda_{st}[r_s(f_{00}^*) - f_{00}^*\,\Delta_x V_{t+1}(x,y)] \\
&\quad + \lambda_{ct}\,d_c[\Delta_x V_{t+1}(x+2,y-1) - \Delta_x V_{t+1}(x+1,y-1)] \\
&\quad + (1-\lambda_{ct}\,d_c)[\Delta_x V_{t+1}(x+2,y) - \Delta_x V_{t+1}(x+1,y)] \\
&= -\lambda_{st}\,f_{20}^*[\Delta_x V_{t+1}(x+2,y) - \Delta_x V_{t+1}(x+1,y)] \\
&\quad + \lambda_{st}\,f_{00}^*[\Delta_x V_{t+1}(x+1,y) - \Delta_x V_{t+1}(x,y)] \\
&\quad + (1-\lambda_{ct}\,d_c)[\Delta_x V_{t+1}(x+2,y) - \Delta_x V_{t+1}(x+1,y)] \\
&\quad + \lambda_{ct}\,d_c[\Delta_x V_{t+1}(x+2,y-1) - \Delta_x V_{t+1}(x+1,y-1)] \\
&\le (1-\lambda_{ct}\,d_c - \lambda_{st}\,f_{20}^*)[\Delta_x V_{t+1}(x+2,y) - \Delta_x V_{t+1}(x+1,y)] \le 0.
\end{aligned}
$$

The first inequality follows from the induction assumption, while the last inequality holds because $(1-\lambda_{ct}\,d_c-\lambda_{st}\,f_{20}^*) \ge (1-\lambda_{ct}-\lambda_{st}) \ge 0$ and the induction assumption. □

Proof of (ii). We use induction argument over time periods to show the result.

At time period $\tau+1$, we have

$$
\begin{aligned}
& \Delta_x V_{\tau+1}(x+1,y+1) - \Delta_x V_{\tau+1}(x+1,y) \\
= \; & V_{\tau+1}(x+1,y+1) - V_{\tau+1}(x,y+1) - V_{\tau+1}(x+1,y) + V_{\tau+1}(x,y) \\
= \; & \Psi_{\tau+1}(x+y+2) - \Psi_{\tau+1}(x+y+1) - \Psi_{\tau+1}(x+y+1) + \Psi_{\tau+1}(x+y) \leq 0
\end{aligned}
$$

due to the concavity of $\Psi_{\tau+1}(\cdot)$. Assuming that the result holds at time period $t+1$, we have

$$
\begin{aligned}
& \Delta_x V_t(x+1,y+1) - \Delta_x V_t(x+1,y) \\
& = V_t(x+1,y+1) - V_t(x,y+1) - V_t(x+1,y) + V_t(x,y) \\
& = \lambda_{st}[r_s(f_{11}^*) - f_{11}^* \, \Delta_x V_{t+1}(x+1,y+1)] + \lambda_{ct}[r_c(d_c) + d_c \, V_{t+1}(x+1,y)] \\
& \qquad + (1-\lambda_{ct} \, d_c) V_{t+1}(x+1,y+1) \\
& \quad - \{\lambda_{st}[r_s(f_{01}^*) - f_{01}^* \, \Delta_x V_{t+1}(x,y+1)] + \lambda_{ct}[r_c(d_c) + d_c \, V_{t+1}(x,y)] \\
& \qquad + (1-\lambda_{ct} \, d_c) V_{t+1}(x,y+1)\} \\
& \quad - \{\lambda_{st}[r_s(f_{10}^*) - f_{10}^* \, \Delta_x V_{t+1}(x+1,y)] + \lambda_{ct}[r_c(d_c) + d_c \, V_{t+1}(x+1,y-1)] \\
& \qquad + (1-\lambda_{ct} \, d_c) V_{t+1}(x+1,y)\} \\
& \quad + \lambda_{st}[r_s(f_{00}^*) - f_{00}^* \, \Delta_x V_{t+1}(x,y)] + \lambda_{ct}[r_c(d_c) + d_c \, V_{t+1}(x,y-1)] \\
& \qquad + (1-\lambda_{ct} \, d_c) V_{t+1}(x,y) \\
& = \lambda_{st}[r_s(f_{11}^*) - f_{11}^* \, \Delta_x V_{t+1}(x+1,y+1)] - \lambda_{st}[r_s(f_{10}^*) - f_{10}^* \, \Delta_x V_{t+1}(x+1,y)] \\
& \quad - \lambda_{st}[r_s(f_{01}^*) - f_{01}^* \, \Delta_x V_{t+1}(x,y+1)] + \lambda_{st}[r_s(f_{00}^*) - f_{00}^* \, \Delta_x V_{t+1}(x,y)] \\
& \quad + \lambda_{ct} \, d_c[\Delta_x V_{t+1}(x+1,y) - \Delta_x V_{t+1}(x+1,y-1)] \\
& \qquad + (1-\lambda_{ct} \, d_c)[\Delta_x V_{t+1}(x+1,y+1) - \Delta_x V_{t+1}(x+1,y)]
\end{aligned}
$$

Due to the definition of f^*_{10} and f^*_{01}, the above equation is upper bounded by

$$
\begin{aligned}
&\leq \lambda_{st}[r_s(f^*_{11}) - f^*_{11}\,\Delta_x V_{t+1}(x+1,y+1)] - \lambda_{st}[r_s(f^*_{11}) - f^*_{11}\,\Delta_x V_{t+1}(x+1,y)] \\
&\quad - \lambda_{st}[r_s(f^*_{00}) - f^*_{00}\,\Delta_x V_{t+1}(x,y+1)] + \lambda_{st}[r_s(f^*_{00}) - f^*_{00}\,\Delta_x V_{t+1}(x,y)] \\
&\quad + \lambda_{ct}\,d_c[\Delta_x V_{t+1}(x+1,y) - \Delta_x V_{t+1}(x+1,y-1)] \\
&\quad + (1-\lambda_{ct}\,d_c)[\Delta_x V_{t+1}(x+1,y+1) - \Delta_x V_{t+1}(x+1,y)] \\
&= -\lambda_{st}\,f^*_{11}[\Delta_x V_{t+1}(x+1,y+1) - \Delta_x V_{t+1}(x+1,y)] \\
&\quad + \lambda_{st}\,f^*_{00}[\Delta_x V_{t+1}(x,y+1) - \Delta_x V_{t+1}(x,y)] \\
&\quad + (1-\lambda_{ct}\,d_c)[\Delta_x V_{t+1}(x+1,y+1) - \Delta_x V_{t+1}(x+1,y)] \\
&\quad + \lambda_{ct}\,d_c[\Delta_x V_{t+1}(x+1,y) - \Delta_x V_{t+1}(x+1,y-1)] \\
&\leq (1-\lambda_{ct}\,d_c - \lambda_{st}\,f^*_{11})[\Delta_x V_{t+1}(x+1,y+1) - \Delta_x V_{t+1}(x+1,y)] \leq 0,
\end{aligned}
$$

The first inequality follows from the induction assumption, while the last inequality holds because $(1-\lambda_{ct}\,d_c-\lambda_{st}\,f^*_{11}) \geq (1-\lambda_{ct}-\lambda_{st}) \geq 0$ and the induction assumption. □

Proof of (iii).

$$
\begin{aligned}
&\Delta_x V_t(x+1,y) - \Delta_x V_{t+1}(x+1,y) = V_t(x+1,y) - V_t(x,y) - \Delta_x V_{t+1}(x+1,y) \\
&= \lambda_{st}[r_s(f^*_{10}) - f^*_{10}\,\Delta_x V_{t+1}(x+1,y)] + \lambda_{ct}[r_c(d_c) + d_c\,V_{t+1}(x+1,y-1)] \\
&\qquad + (1-\lambda_{ct}\,d_c)V_{t+1}(x+1,y) \\
&\quad - \{\lambda_{st}[r_s(f^*_{00}) - f^*_{00}\,\Delta_x V_{t+1}(x,y)] + \lambda_{ct}[r_c(d_c) + d_c\,V_{t+1}(x,y-1)] \\
&\qquad + (1-\lambda_{ct}\,d_c)V_{t+1}(x,y)\} - \Delta_x V_{t+1}(x+1,y) \\
&= \lambda_{st}[r_s(f^*_{10}) - f^*_{10}\,\Delta_x V_{t+1}(x+1,y)] - \lambda_{st}[r_s(f^*_{00}) - f^*_{00}\,\Delta_x V_{t+1}(x,y)] \\
&\quad + \lambda_{ct}\,d_c[\Delta_x V_{t+1}(x+1,y-1) - \Delta_x V_{t+1}(x+1,y)]
\end{aligned}
$$

From Property (i) we have $\Delta_x V_{t+1}(x,y) \geq \Delta_x V_{t+1}(x+1,y)$; from Property (ii) we know $\Delta_x V_{t+1}(x+1,y-1) - \Delta_x V_{t+1}(x+1,y) \geq 0$, hence the above equation can be lower bounded by

$$
\begin{aligned}
&\geq \lambda_{st}[r_s(f_{10}^*) - f_{10}^*\, \Delta_x V_{t+1}(x+1, y)] - \lambda_{st}[r_s(f_{00}^*) - f_{00}^*\, \Delta_x V_{t+1}(x+1, y)] \\
&= \lambda_{st}[r_s(f_{10}^*) - r_s(f_{00}^*) - \Delta_x V_{t+1}(x+1, y)(f_{10}^* - f_{00}^*)] \\
&= \lambda_{st}[r_s(f_{10}^*) - r_s(f_{00}^*) - r_s'(f_{10}^*)\,(f_{10}^* - f_{00}^*)] \geq 0.
\end{aligned}
$$

The last equality holds since f_{10}^* is characterized by $r_s'(f_{10}^*) = \Delta_x V_{t+1}(x+1, y)$. The last inequality follows from the assumption that $r_s(\cdot)$ is a concave function.

□

B.2 Proof of Proposition 3.3.3

In this section we provide the proof for Proposition 3.3.3.

Proof. Let $\{\hat{d}_{st} : t = 1, \ldots \tau\}, \{\hat{d}_{ft} : t = \tau + 1, \ldots, T\}$ be a series of demand rates offered in the spot and final market under optimal policy from (3.1), and $\{\hat{p}_{st} : t = 1, \ldots \tau\}, \{\hat{p}_{ft} : t = \tau + 1, \ldots, T\}$ be the corresponding prices. Note that all of them are random variables due to the randomness of real demands. Let $\hat{S}_t = 1$ if we sell one room at time period t in the spot market and $\hat{S}_t = 0$ otherwise. For $t = 1, \ldots, \tau$, we have $\mathbb{E}(\hat{S}_t \,|\, \hat{p}_{st}) = \mathbb{E}(\hat{S}_t \,|\, \hat{d}_{st}) = \lambda_{st}\, \hat{d}_{st}$. Similarly, for $t = \tau + 1, \ldots, T$, we have $\mathbb{E}(\hat{S}_t \,|\, \hat{p}_{ft}) = \hat{d}_{ft}$. Total revenue is composed of three parts: revenue in the spot market from time period 1 to τ, revenue in the conference market from time period 1 to τ and revenue in the spot market from time period $\tau+1$ to T. Then we have

$$\begin{aligned}
V_1(C-b,b) &= \sum_{t=1}^{\tau} \mathbb{E}(\hat{S}_t\,\hat{p}_{st}) + \sum_{t=\tau+1}^{T} \mathbb{E}(\hat{S}_t\,\hat{p}_{ft}) + p_c\,\mathbb{E}\min\{b, D_c\} \\
&= \sum_{t=1}^{\tau} \mathbb{E}(\mathbb{E}(\hat{S}_t \,|\, \hat{p}_{st})\,\hat{p}_{st}) + \sum_{t=\tau+1}^{T} \mathbb{E}(\mathbb{E}(\hat{S}_t \,|\, \hat{p}_{ft})\,\hat{p}_{ft}) + p_c\,\mathbb{E}\min\{b, D_c\} \\
&= \sum_{t=1}^{\tau} \mathbb{E}(\lambda_{st}\,\hat{d}_{st}\,\hat{p}_{st}) + \sum_{t=\tau+1}^{T} \mathbb{E}(\,\hat{d}_{ft}\,\hat{p}_{ft}) + p_c\,\mathbb{E}\min\{b, D_c\} \\
&= \sum_{t=1}^{\tau} \lambda_{st}\mathbb{E}(r_s(\hat{d}_{st})) + \sum_{t=\tau+1}^{T} \mathbb{E}(r_f(\hat{d}_{ft})) + p_c\,\mathbb{E}\min\{b, D_c\} \\
&\le \sum_{t=1}^{\tau} \lambda_{st}\,r_s(\mathbb{E}(\hat{d}_{st})) + \sum_{t=\tau+1}^{T} r_f(\mathbb{E}(\hat{d}_{ft})) + p_c\,\mathbb{E}\min\{b, D_c\},
\end{aligned}$$

where the last inequality holds due to the concavity of $r_s(\cdot)$ and $r_f(\cdot)$. Define $\tilde{d}_{st} = \mathbb{E}(\hat{d}_{st}), t = 1, \ldots, \tau$, $\tilde{d}_{ft} = \mathbb{E}(\hat{d}_{ft}), t = \tau+1, \ldots, T$, then the above inequality can be written as

$$V_1(C-b,b) \le \sum_{t=1}^{\tau} \lambda_{st}\,r_s(\tilde{d}_{st}) + \sum_{t=\tau+1}^{T} r_f(\tilde{d}_{ft}) + p_c\,\mathbb{E}\min\{b, D_c\}. \tag{B.1}$$

Also for $t = 1, \ldots, \tau$, we have $\mathbb{E}(\hat{S}_t) = \mathbb{E}(\lambda_{st}\,\hat{d}_{st}) = \lambda_{st}\,\tilde{d}_{st}$, and for $t = \tau+1, \ldots, T$, we have $\mathbb{E}(\hat{S}_t) = \tilde{d}_{ft}$. Due to capacity constraint, we know

$$\begin{aligned}
\sum_{t=1}^{\tau} \hat{S}_t \qquad\qquad &\le C - b, \\
\sum_{t=1}^{\tau} \hat{S}_t + \sum_{t=\tau+1}^{T} \hat{S}_t &\le C - b + (b - D_c)^+.
\end{aligned}$$

Taking expectation for both sides, we get

$$\begin{aligned}
\sum_{t=1}^{\tau} \lambda_{st}\tilde{d}_{st} \qquad\qquad &\le C - b, \\
\sum_{t=1}^{\tau} \lambda_{st}\tilde{d}_{st} + \sum_{t=\tau+1}^{T} \tilde{d}_{ft} &\le C - b + \mathbb{E}(b - D_c)^+.
\end{aligned}$$

Comparing above inequalities with constraints in (NLP1), we see that $\tilde{d}_{st}, t = 1, \ldots, \tau$, $\tilde{d}_{ft}, t = \tau+1, \ldots, T$ is a feasible solution for (NLP1). The objective value is $\sum_{t=1}^{\tau} \lambda_{st} r_s(\tilde{d}_{st}) + \sum_{t=\tau+1}^{T} r_f(\tilde{d}_{ft})$, which is no less than $V_1(C-b,b) - p_c \mathbb{E}\min\{b, D_c\}$ as shown in (B.1). Thus $V_1(C-b,b) \leq Z_1 + p_c \mathbb{E}\min\{b, D_c\}$. The inequality $\Phi_1(C-b) \leq Z_1$ then follows immediately from Proposition 3.3.2. $\square$

B.3 Variance inequality

Lemma B.3.1. *For a deterministic scalar z and a nonnegative real random variable Z with finite mean μ and variance σ^2, we have Var$\{(z-Z)^+\} \leq \sigma^2$.*

Proof. From the definition of variance, we have

$$\text{Var}\{(z-Z)^+\} = \mathbb{E}\{[(z-Z)^+]^2\} - [\mathbb{E}(z-Z)^+]^2 \leq \mathbb{E}\{[(z-Z)^+]^2\} - [(z-\mu)^+]^2.$$

If $z \geq \mu$, we can continue the chain of inequality as

$$\text{Var}\{(z-Z)^+\} \leq \mathbb{E}\{(z-Z)^2\} - (z-\mu)^2 = \text{Var}(z-Z) = \text{Var}(Z) = \sigma^2.$$

If $z < \mu$, we can continue the chain of inequality as

$$\begin{aligned} \text{Var}\{(z-Z)^+\} \leq \mathbb{E}\{[(z-Z)^+]^2\} &= \int_0^z (z-x)^2 \, d\, F_Z(x) \leq \int_0^z (\mu-x)^2 \, d\, F_Z(x) \\ &\leq \int_0^\infty (\mu-x)^2 \, d\, F_Z(x) = \text{Var}(Z) = \sigma^2. \end{aligned}$$

$\square$

B.4 Proof of Lemma 3.4.3

We provide a proof for Lemma 3.4.3 in this section.

Proof. It suffices to show that the following equivalent result: *For any* $b_1 = 0, 1, \ldots, C-1$ *and* $\epsilon \in (0,1)$*,* $W(b_1+\epsilon)$ *is a linear function of* ϵ*.*

For the first two terms in $W(b)$, letting $b = b_1 + \epsilon$, then $C - b = (C - b_1 - 1) + (1-\epsilon)$, We have

$$\begin{aligned}\mathbb{E}\min\{b, D_c(\bar{d}_c)\} &= \mathbb{E}\min\{b_1+\epsilon, D_c(\bar{d}_c)\} \\ &= \sum_{k=0}^{b_1} k\,\text{Prob}\{D_c(\bar{d}_c)=k\} + (b_1+\epsilon)\text{Prob}\{D_c(\bar{d}_c) > b_1\},\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\min\{C-b, D_s(\bar{d}_s)\} = \mathbb{E}\min\{(C-b_1-1)+(1-\epsilon), D_s(\bar{d}_s)\} \\ &= \sum_{k=0}^{C-b_1-1} k\,\text{Prob}\{D_s(\bar{d}_s)=k\} + (C-b_1-\epsilon)\text{Prob}\{D_s(\bar{d}_s) > C-b_1-1\}.\end{aligned}$$

are both linear in ϵ. The third term

$$\begin{aligned}&\mathbb{E}\min\{(b-D_c(\bar{d}_c))^+ + (C-b-D_s(\bar{d}_s))^+, D_f(\bar{d}_f)\} \\ &= \sum_{k_c=0}^{\infty}\sum_{k_s=0}^{\infty}\text{Prob}\{D_c(\bar{d}_c)=k_c\}\text{Prob}\{D_s(\bar{d}_s)=k_s\} \\ &\qquad * \mathbb{E}\min\{(b-k_c)^+ + (C-b-k_s)^+, D_f(\bar{d}_f)\}.\end{aligned}$$

Letting $h(k_c, k_s, b) = \mathbb{E}\min\{(b-k_c)^+ + (C-b-k_s)^+, D_f(\bar{d}_f)\}$, it suffices to show that for any $k_c, k_s \in \mathbb{Z}_{\geq 0}$, $h(k_c, k_s, b)$ is linear in ϵ. We show this by discussing different cases.

(a) If $b \geq k_c$ and $C - b \geq k_s$, then we have

$$h(k_c, k_s, b) = \mathbb{E}\min\{b-k_c+C-b-k_s, D_f(\bar{d}_f)\} = \mathbb{E}\min\{C-k_c-k_s, D_f(\bar{d}_f)\},$$

which does not contain ϵ terms.

(b) If $b \geq k_c$ and $C - b < k_s$, then we have

$$h(k_c, k_s, b) = \mathbb{E}\min\{b - k_c, D_f(\bar{d}_f)\} = \mathbb{E}\min\{(b_1 - k_c) + \epsilon, D_f(\bar{d}_f)\},$$

which is a linear function of ϵ according to similar arguments for the first term in $W(b)$.

(c) If $b < k_c$ and $C - b \geq k_s$, then we have

$$\begin{aligned} h(k_c, k_s, b) &= \mathbb{E}\min\{C - b - k_s, D_f(\bar{d}_f)\} \\ &= \mathbb{E}\min\{(C - b_1 - k_s - 1) + (1 - \epsilon), D_f(\bar{d}_f)\}, \end{aligned}$$

which is a linear function of ϵ according to similar arguments for the second term in $W(b)$.

(d) If $b < k_c$ and $C - b < k_s$, then we have

$$h(k_c, k_s, b) = \mathbb{E}\min\{0, D_f(\bar{d}_f)\} = 0$$

Thus the result follows.

□

APPENDIX C

APPENDIX FOR CHAPTER 4

C.1 Proof of Proposition 4.4.1

In this section we provide a proof for Proposition 4.4.1, which says that Z_5 from (NLP5) provides an upper bound for the optimal total expected revenue of the stochastic hybrid model in Section 4.4.

Proof. Let $\{\hat{\boldsymbol{d}}_t : \, t = 1, \dots \tau\}, \{\hat{d}_{ft} : t = \tau + 1, \dots, T\}$ be the demand rates offered in the separate markets and joint market under optimal policy from (4.21), and $\{\hat{\boldsymbol{p}}_t : \, t = 1, \dots \tau\}, \{\hat{p}_{ft} : t = \tau + 1, \dots, T\}$ be the corresponding prices. Note that all of them are random variables due to the randomness of real demands.

For $t = 1, \dots \tau$, let $\hat{S}_{it} = 1$ if we sell one unit of product at time period t in market i and $\hat{S}_{it} = 0$ otherwise. Then we have $\mathbb{E}(\hat{S}_{it} \,|\, \hat{p}_{it}) = \mathbb{E}(\hat{S}_t \,|\, \hat{d}_{it}) = \lambda_i \, \hat{d}_{it}$. Similarly, for $t = \tau + 1, \dots, T$, let $\hat{J}_t = 1$ if we sell one unit of product at time period t in the joint market and $\hat{J}_t = 0$ otherwise. Then we have $\mathbb{E}(\hat{J}_t \,|\, \hat{p}_{ft}) = \sum_{i=1}^{n} \lambda_i d_i(\hat{d}_{ft})$.

Total revenue is composed of two parts: revenue in the separate markets from time period 1 to τ and revenue in the joint market from time period $\tau + 1$

to T. Then we have

$$\begin{aligned}
V_1(\boldsymbol{b}) &= \sum_{t=1}^{\tau}\sum_{i=1}^{n}\mathbb{E}(\hat{S}_{it}\,\hat{p}_{it}) + \sum_{t=\tau+1}^{T}\mathbb{E}(\hat{J}_t\,\hat{p}_{ft}) \qquad \text{(C.1)}\\
&= \sum_{t=1}^{\tau}\sum_{i=1}^{n}\mathbb{E}(\mathbb{E}(\hat{S}_{it}\,|\,\hat{p}_{it})\,\hat{p}_{it}) + \sum_{t=\tau+1}^{T}\mathbb{E}(\mathbb{E}(\hat{J}_t\,|\,\hat{p}_{ft})\,\hat{p}_{ft})\\
&= \sum_{t=1}^{\tau}\sum_{i=1}^{n}\mathbb{E}(\lambda_i\,\hat{d}_{it}\,\hat{p}_{it}) + \sum_{t=\tau+1}^{T}\mathbb{E}(\sum_{i=1}^{n}\lambda_i d_i(\hat{d}_{ft})\,\hat{p}_{ft})\\
&= \sum_{t=1}^{\tau}\sum_{i=1}^{n}\lambda_i\mathbb{E}(r_i(\hat{d}_{it})) + \sum_{t=\tau+1}^{T}\mathbb{E}(r_f(\hat{d}_{ft}))\\
&\le \sum_{t=1}^{\tau}\sum_{i=1}^{n}\lambda_i\,r_s(\mathbb{E}(\hat{d}_{it})) + \sum_{t=\tau+1}^{T}r_f(\mathbb{E}(\hat{d}_{ft})),
\end{aligned}$$

where the last inequality holds due to the concavity of $r_i(\cdot)$ and $r_f(\cdot)$. Define $\tilde{d}_{it} = \mathbb{E}(\hat{d}_{it}), t = 1, \ldots, \tau$, $\tilde{d}_{ft} = \mathbb{E}(\hat{d}_{ft}), t = \tau + 1, \ldots, T$, then the above inequality can be written as

$$V_1(\boldsymbol{b}) \le \sum_{t=1}^{\tau}\sum_{i=1}^{n}\lambda_i\,r_i(\tilde{d}_{it}) + \sum_{t=\tau+1}^{T}r_f(\tilde{d}_{ft}). \qquad \text{(C.2)}$$

Besides, for $t = 1, \ldots, \tau$, we have $\mathbb{E}(\hat{S}_{it}) = \mathbb{E}(\lambda_i\,\hat{d}_{it}) = \lambda_{st}\,\tilde{d}_{it}$, and for $t = \tau + 1, \ldots, T$, we have $\mathbb{E}(\hat{J}_t) = \mathbb{E}\{\sum_{i=1}^{n}\lambda_i d_i(\hat{d}_{ft})\} \ge \sum_{i=1}^{n}\lambda_i d_i(\tilde{d}_{ft})$ using Jensen's inequality. Due to capacity constraint, we know

$$\begin{aligned}
\sum_{t=1}^{\tau}\hat{S}_{it} \qquad\qquad\qquad &\le b_i,\\
\sum_{t=1}^{\tau}\sum_{i=1}^{n}\hat{S}_{it} + \sum_{t=\tau+1}^{T}\hat{J}_t &\le C.
\end{aligned}$$

Taking expectation for both sides and replacing $\mathbb{E}(\hat{J}_t)$ with a smaller quantity $\sum_{i=1}^{n}\lambda_i d_i(\tilde{d}_{ft})$, we get

$$\begin{aligned}
\sum_{t=1}^{\tau}\lambda_i\tilde{d}_{it} \qquad\qquad\qquad\qquad &\le b_i,\\
\sum_{t=1}^{\tau}\sum_{i=1}^{n}\lambda_i\tilde{d}_{it} + \sum_{t=\tau+1}^{T}\lambda_i d_i(\tilde{d}_{ft}) &\le C.
\end{aligned}$$

Comparing above inequalities with constraints in (NLP5), we see that $\{\tilde{d}_{it}, t = 1, \ldots, \tau\}$, $\{\tilde{d}_{ft}, t = \tau+1, \ldots, T\}$ is a feasible solution for (NLP5) with an objective value $\sum_{t=1}^{\tau} \sum_{i=1}^{n} \lambda_i r_i(\tilde{d}_{it}) + \sum_{t=\tau+1}^{T} r_f(\tilde{d}_{ft})$, which is no less than $V_1(\boldsymbol{b})$ as shown in (C.2). Thus $V_1(\boldsymbol{b}) \leq Z_5$. $\square$

BIBLIOGRAPHY

J.A. Adell, F.G. Badia, and J. de la Cal. Beta-type operators preserve shape properties. *Stochastic Processes and their Applications*, 48(1):1–8, 1993.

J.A. Adell, F.G. Badia, J. de la Cal, and Fernando Plo. On the property of monotonic convergence for beta operators. *Journal of Approximation Theory*, 84(1): 61–73, 1996.

D. Adelman and A.J. Mersereau. Relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 56(3):712–727, 2008.

Y. Akçay, H.P. Natarajan, and S.H. Xu. Joint dynamic pricing of multiple perishable products under consumer choice. *Management Science*, 56(8):1345–1361, 2010.

V.F. Araman and R. Caldentey. Dynamic pricing for non-perishable products with demand learning. *Operations Research*, 57(5):1169–1188, 2009.

A. Arlotto, S.E. Chick, and N. Gans. Optimal hiring and retention policies for heterogeneous workers who learn. Technical report, Wharton School of Business, 2011.

Y. Aviv and A. Pazgal. A partially observed Markov decision process for dynamic pricing. *Management Science*, 51(9):1400–1416, 2005.

D.A. Berry and B. Fristedt. *Bandit problems: Sequential allocation of experiments*. Springer, New York, NY, 1985.

D. Bertsimas and A.J. Mersereau. A learning approach for interactive marketing to a customer segment. *Operations Research*, 55(6):1120–1135, 2007.

G. Bitran and R. Caldentey. Commissioned paper: An overview of pricing models for revenue management. *Manufacturing & Service Operations Management*, 5(3):203–229, 2003.

R.N. Bradt, S.M. Johnson, and S. Karlin. On sequential designs for maximizing the sum of n observations. *The Annals of Mathematical Statistics*, 27(4):1060–1074, 1956.

D.B. Brown, J.E. Smith, and P. Sun. Information relaxations and duality in stochastic dynamic programs. *Operations Research*, 58(4):785–801, 2010. ISSN 0030-364X.

A.N. Burnetas and M.N. Katehakis. Dynamic allocation policies for the finite horizon one armed bandit problem. *Stochastic Analysis and Applications*, 16(5): 811–824, 1998.

A.N. Burnetas and M.N. Katehakis. Asymptotic Bayes Analysis for the Finite-Horizon One-Armed-Bandit Problem. *Probability in the Engineering and Informational Sciences*, 17(1):53–82, 2003.

F. Caro and J. Gallien. Dynamic assortment with demand learning for seasonal consumer goods. *Management Science*, 53(2):276, 2007.

L. Dong, P. Kouvelis, and Z. Tian. Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11(2): 317–339, 2009.

W. Elmaghraby and P. Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science*, 49(10):1287–1309, 2003.

A. Erdelyi and H. Topaloglu. A dynamic programming decomposition method for making overbooking decisions over an airline network. *INFORMS Journal on Computing*, 22(3):443–456, 2010a.

A. Erdelyi and H. Topaloglu. Using decomposition methods to solve pricing problems in network revenue management. *Journal of Revenue & Pricing Management*, 10(4):325–343, 2010b.

V.F. Farias and B. Van Roy. Dynamic pricing with a prior on market response. *Operations Research*, 58(1):16–29, 2010.

Y. Feng and G. Gallego. Optimal starting times for end-of-season sales and optimal stopping times for promotional fares. *Management Science*, 41(8):1371–1391, 1995.

Y. Feng and B. Xiao. A continuous-time yield management model with multiple prices and reversible price changes. *Management Science*, 46(5):644–657, 2000a.

Y. Feng and B. Xiao. Optimal policies of yield management with multiple predetermined prices. *Operations Research*, 48(2):332–343, 2000b.

G. Gallego. A minmax distribution free procedure for the (Q, R) inventory model. *Operations Research Letters*, 11(1):55–60, 1992.

G. Gallego and G. Van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994.

G. Gallego and G. Van Ryzin. A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research*, 45(1):24–41, 1997.

J.C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):148–177, 1979.

J.C. Gittins. *Multi-armed Bandit Allocation Indices*. John Wiley & Sons Inc, Chichester, West Sussex, UK, 1989.

A. Goldenshluger and A. Zeevi. Woodroofe's one-armed bandit problem revisited. *Annals of Applied Probability*, 19(4):1603–1633, 2009.

A. Goldenshluger and A. Zeevi. A note on performance limitations in bandit problems with side information. *IEEE Transactions on Information Theory*, 57 (3):1707–1713, 2011.

JM Harrison, NB Keskina, and A. Zeevi. Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. forthcoming in management science. 2011.

J.T. Hawkins. *A Lagrangian decomposition approach to weakly coupled dynamic optimization problems and its applications*. PhD thesis, MIT, 2003.

C. Maglaras and J. Meissner. Dynamic pricing strategies for multi-product revenue management problems. *Manufacturing & Service Operations Management*, 8(2):136–148, 2006.

D.A. Malueg and C.M. Snyder. Bounding the relative profitability of price discrimination. *International Journal of Industrial Organization*, 24(5):995–1011, 2006.

David A. Malueg. Bounding the welfare effects of third-degree price discrimination. *The American Economic Review*, 83(4):pp. 1011–1021, 1993.

J.I. McGill and G.J. Van Ryzin. Revenue management: Research overview and prospects. *Transportation science*, 33(2):233, 1999.

Mission Repair. Mission repair: Iphone 4s, 2012. URL http://www.missionrepair.com/iPhone_4S_Repair_and_Diagnosis_s/762.htm.

J. Robinson. *The economics of imperfect competition*. Macmillan, 1933.

R. Schmalensee. Output and welfare implications of monopolistic third-degree price discrimination. *The American Economic Review*, 71(1):242–247, 1981.

M. Shaked and J.G. Shanthikumar. *Stochastic orders*. Springer Science+ Business Media, LLC, 2007.

K.T. Talluri and G. Van Ryzin. *The theory and practice of revenue management*. Springer Science + Business Media, Inc., Boston, MA, 2005.

K. Teerapabolarn. A bound on the binomial approximation to the beta binomial distribution. *International Mathematical Forum*, 3(28):1355–1358, 2008.

Top Extended Vehicle Warranty Providers. A top ten list of automotive extended warranty companies, 2010. URL http://ezinearticles.com/?A-Top-Ten-List-of-Automotive-Extended-Warranty-Companies&id=3728026.

H. Topaloglu. Using Lagrangian Relaxation to Compute Capacity-Dependent Bid Prices in Network Revenue Management. *Operations Research*, 57(3):637–649, 2009.

H. Topaloglu and S. Kunnumkal. Computing time-dependent bid-prices in network revenue management problems. *Transportation Sci*, 44(1):38–62, 2010.

Xerox Corporation. U.S. and Canada on-site service agreement, 2011. URL http://www.office.xerox.com/latest/SERTC-16.PDF.

D. Zhang and W.L. Cooper. Pricing substitutable flights in airline revenue management. *European Journal of Operational Research*, 197(3):848–861, 2009.

W. Zhao and Y.S. Zheng. Optimal dynamic pricing for perishable assets with nonhomogeneous demand. *Management Science*, 46(3):375–388, 2000.