

Near-Optimal Data-Driven Approximation Schemes for Joint Pricing and Inventory Control Models

by

Hanzhang Qin

B.Eng., Industrial Engineering, B.S., Mathematics,
Tsinghua University (2015)

Submitted to the Department of Civil and Environmental Engineering and the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of

Master of Science in Transportation

and

Master of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

© Massachusetts Institute of Technology 2018. All rights reserved.

Signature redacted

Author.....

Department of Civil and Environmental Engineering
Department of Electrical Engineering and Computer Science

Signature redacted

May 11, 2018

Certified by.....

David Simchi-Levi

Professor of Civil and Environmental Engineering and Engineering Systems

Signature redacted

Thesis Supervisor

Certified by.....

John N. Tsitsiklis

Clarence J Lebel Professor of Electrical Engineering and Computer Science

Signature redacted

Thesis Reader

Accepted by.....

Jesse Kroll

Professor of Civil and Environmental Engineering Chair

Graduate Program Committee

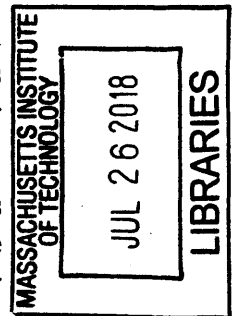
Accepted by.....

Signature redacted

Lesie A. Kolodziejski

Professor of Electrical Engineering and Computer Science

Chair, Department Committee on Graduate Students



ARCHIVES

Near-Optimal Data-Driven Approximation Schemes for Joint Pricing and Inventory Control Models

by

Hanzhang Qin

Submitted to the Department of Civil and Environmental Engineering
and the Department of Electrical Engineering and Computer Science
on May 11, 2018, in partial fulfillment of the
requirements for the degrees of
Master of Science in Transportation
and
Master of Science in Electrical Engineering and Computer Science

Abstract

The thesis studies the classical multi-period joint pricing and inventory control problem in a data-driven setting. In the problem, a retailer makes periodic decisions of the prices and inventory levels of an item that the retailer wishes to sell. The objective is to match the inventory level with a random demand that depends on the price in each period, while maximizing the expected profit over finite horizon. In reality, the demand functions or the distribution of the random noise are usually unavailable, whereas past demand data are relatively easy to collect. A novel data-driven nonparametric algorithm is proposed, which uses the past demand data to solve the joint pricing and inventory control problem, without assuming the parameters of the demand functions and the noise distributions are known. Explicit sample complexity bounds are given, on the number of data samples needed to guarantee a near-optimal profit. A simulation study suggests that the algorithm is efficient in practice.

Thesis Supervisor: David Simchi-Levi

Title: Professor of Civil and Environmental Engineering and Engineering Systems

Thesis Reader: John N. Tsitsiklis

Title: Clarence J Lebel Professor of Electrical Engineering and Computer Science

Acknowledgments

First and foremost, I would like to thank my thesis advisor, Professor David Simchi-Levi. This thesis could not be finished at all without David's supervision. As a top-tier scholar, David's thoughtful insights really helped me improve the quality of this work. Moreover, beyond his academic excellence, David is really more than merely an academic advisor. He also taught me the many important aspects of leading a good life. I could not imagine how my life at MIT would be without David as my advisor.

Then, I would like to say my sincere thanks to Professor John Tsitsiklis. During my first semester at MIT, I took John's probability theory course and later his dynamic programming course, which helped me develop the mathematics knowledge and technical skills that shape this whole thesis. John also provided critical comments on an early version of this work, which is precious to me while I was still entangled with some mathematical details.

Professor Rahul Mazumder is among the long list of people I want to thank as well. Our research collaboration started early in 2016, and I really renewed my knowledge about statistics and machine learning due to a large number of stimulating discussions with Rahul.

My research collaborator, labmate, roommate, friend, Li Wang, is also who I want to give special thanks to. I remember we took many same graduate courses together, had a lot of fun together, and the moments Li entertained the people around him. This thesis is also a result of the many days and nights of work of us, and in particular he contributed greatly on the simulation part.

I would also like to thank the Department of Civil and Environmental Engineering (CEE) that provided me with such a great environment where I studied as a student

in the Master of Transportation (MST) program; I appreciate the Department of Electrical Engineering and Computer Science (EECS) for giving me the chance as being a dual SM student, and I enjoyed the solid curriculum it offered.

I am so fortunate to have made many amazing friends during my time at MIT, whose names are too many to enumerate. Besides, I feel very lucky and am grateful to the CEE department and the Center for Computational Engineering (CCE) that allowed me to continue my academic adventure at MIT as a Ph.D. student. I will certainly try my best to continue chasing my dreams that I once dreamed since I was only a child.

Last but not least, I owe my deepest gratitude to Shaobo Hu, who has accompanied me during most of these three years, as well as to my grandmother Guojin Fei, my cousin Xihe Jin, my dad Wenxin Qin, and my mom Yan Jin, who have been always in support of me despite all of my ups and downs. This thesis is dedicated to them.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 13 |
| 1.1 | Related literature | 15 |
| 1.2 | Contributions and connections to the literature | 17 |
| 1.3 | Organization of the thesis | 18 |
| 2 | Joint Pricing and Inventory Control Model | 19 |
| 2.1 | Full-Information Problem | 23 |
| 2.2 | Data-Driven Problem | 27 |
| 3 | The Multi-Period Data-Driven Approximation | 31 |
| 3.1 | Empirical Revenue and Cost Functions | 31 |
| 3.2 | Functions in the Empirical Dynamic Program | 33 |
| 4 | Theoretical Algorithm Analysis for linear demand models | 37 |
| 4.1 | The Sample Complexity Bound | 37 |
| 4.1.1 | The First-Order Analysis | 38 |
| 4.1.2 | The Zero-Order Analysis | 42 |
| 4.2 | On the Computational Complexity | 44 |

| | | |
|----------|--|-----------|
| 5 | Extension to Joint Pricing and Capacitated Inventory Control Models under General Concave Demand Models | 47 |
| 6 | A Simulation Study | 53 |
| 7 | Conclusions and Future Directions | 57 |
| A | Tables | 59 |
| B | Figures | 61 |
| C | Proofs of Lemmas and Theorems | 65 |
| D | Additional Lemmas | 93 |

List of Figures

| | | |
|-----|---|----|
| 6-1 | The expected profit generated by the data-driven policy v.s. the full-information optimal policy, in the first period. | 56 |
| B-1 | Illustration of Fact 9. | 62 |
| B-2 | Illustration of Proof of Lemma 7. | 62 |
| B-3 | Data-driven $\hat{d}_t^*(y_t)$ v.s. full-information $d_t^*(y_t)$ from the problem in Chapter 6. | 63 |
| B-4 | Expected profit generated by the data-driven policy v.s. the full-information optimal policy from the problem in Chapter 6. | 64 |

List of Tables

| | | |
|-----|---|----|
| 6.1 | Main problem parameters | 54 |
| A.1 | Other problem parameters in Chapter 6 | 59 |

Chapter 1

Introduction

The joint pricing and inventory control problem is one of the core problems in business operations, and it succinctly characterizes the challenge of coordinating pricing and inventory decisions simultaneously. The joint consideration of pricing and inventory control provides companies with opportunities to correlate sales and operations activities, as companies tend to make independent decisions across departments. Combining pricing and inventory decisions helps companies improve their bottom lines. Higher prices compensate for low inventory levels and thus help avoid high inventory backlog volumes, whereas lower prices accelerate the consumption of high inventory levels and hence reduce inventory holding costs. However, this problem is difficult to solve both in theory and in practice, especially when the exact form of the random demand is unknown.

In this thesis, we study the multi-period joint pricing and inventory control problem in a data-driven setting. Specifically, we suppose a retailer, who sells a single item, makes periodic decisions on the item's pricing and inventory strategies, which directly affect revenues gained and costs incurred. The retailer's goal is to match

the inventory level with a random demand that depends on the price in each period, while maximizing the total profit (total revenue less total cost) over a finite planning horizon of multiple periods.

In the operations management research community, most previous studies on this problem assume the functional dependence between demand and price (demand functions) and the distributions of the random noise in demand (noise distributions) are known to the retailer. We call this specific problem the *full-information problem*. However, in practice, the demand functions and noise distributions are usually unavailable or too complex to work with, whereas past sales data are relatively easy to collect. Thus, we study this problem in a data-driven setting, which we call the *data-driven problem*. We give a practical solution to the data-driven problem, by proposing an algorithm that takes past data as inputs and returns pricing and inventory policies as outputs. The algorithm constructs approximate revenue functions by piecewise linear approximation, and estimates the noise distributions by Sample Average Approximation (SAA).

We prove that, for any accuracy level $\epsilon > 0$ and confidence level $0 < 1 - \beta < 1$, the empirical solution returned by our algorithm using a data set of size more than a certain bound, generates an expected total profit over the finite horizon that is at least $1 - \epsilon$ of the optimal expected profit with probability at least $1 - \beta$. We give the explicit forms of bounds on the numbers of data samples needed to guarantee a near-optimal profit, based on the properties of piecewise linear approximation and several concentration inequalities in the large deviations theory.

To the best of our knowledge, this paper is the first to present a nonparametric data-driven scheme to fully approximate the classic multi-period joint inventory and pricing control problem using purely past data. It is also the first to show the sample complexity of this problem in terms of the accuracy parameter ϵ , confidence

parameter β and number of periods T . Our results accord with the typical sample complexity for a general stochastic dynamic program/multi-stage stochastic program.

1.1 Related literature

The full-information joint pricing and inventory control problem is introduced by [15] in a single-period fashion. Since this first work, much research has been devoted to analyzing and solving the joint pricing and inventory control problem, with the assumption that the retailer has perfect information of the underlying demand. The review papers by [12], [7], [16], and [4] summarize the voluminous literature related to studies on various versions of the full-information problem. The underlying model in this paper is adapted from the classic single-item multi-period models developed by [8] and [3], with the assumption of zero fixed ordering cost.

Assuming data are observed on the fly as the retailer changes prices, [2] consider a data-driven version of the joint pricing and inventory control problem, with the focus on solving the problem in an “online learning” setting. Specifically, the retailer has no past data on the underlying demand and collects data only after setting an initial price. Based on the newly collected data, the retailer determines a new price and collects demand data that correspond to the new price. This process is repeated iteratively. We call the data-driven problem in this specific setting the *online problem*.

Compared to the online problem, the data-driven problem in this paper is different in the respect that all data are collected a priori. Using purely past data, our algorithm outputs a simple pricing and inventory decision policy that can be easily implemented, and there is no need to design complicated price experimentation after the policy is carried out. We call the data-driven problem, in which only

past data are used, the *offline problem*. For the offline problem, we aim to provide finite sample bounds, which guarantee near-optimal total profits over multiple periods. In contrast, the online problem focuses on the asymptotic analysis of the “exploration-exploitation” trade-off in a single period, as more data are received on the fly. Therefore, [2] prove an optimal regret rate of their online algorithm, instead of giving finite sample bounds for provable near-optimality.

Although research on the offline data-driven joint pricing and inventory control problem is absent in the current literature, the pure inventory control problem in the offline data-driven setting has been studied. [11] present a data-driven algorithm to the multi-period inventory control problem and show a polynomial sample complexity for provable near-optimality. [5] extend the inventory control model in [11] by introducing a maximum ordering capacity constraint in each period, and prove a polynomial sample complexity for the more generalized problem. Using a different analysis framework from that in [11], [5] establish a uniform approximation of the derivatives of the value functions in the problem, called the first-order analysis. Then, based on the first-order analysis, they obtain provable near-optimality of the function values, called the *zero-order analysis*. In this paper, we adapt this first-order and zero-order analysis framework in our development of the finite sample bounds for provably near-optimal profits in the joint pricing and inventory control problem.

Other work on the data-driven inventory control problem includes research by [9] and [1]. [9] shows similar sample complexity results based on the technique of K -approximation sets and functions, which is introduced by [10]. [1] uses data-driven confidence intervals of the optimal policies, instead of proving finite sample lower bounds to obtain near-optimal point estimates.

The joint pricing and inventory control problem is a special type of multi-stage stochastic programs, which are intractable in general, as discussed by [14]. For gen-

eral convex multi-stage stochastic programs, [13] shows that the sample complexity lower bound is exponential. Therefore, the goal in this paper is to develop finite sample complexity for a special type of multi-stage stochastic programs, which in general are very hard to approximate and require exponential sample lower bounds. Furthermore, the computational complexity of general multi-stage stochastic convex programs is known as #P-hard, as shown by [6]. In a data-driven setting, [2] prove that solving the SAA counterparts of the uncapacitated multi-period inventory control problem is NP-complete, which indicates that solving the SAA counterparts of the multi-period joint pricing and inventory control problem is also NP-complete.

1.2 Contributions and connections to the literature

Our work is mostly related to [11] and [5]. In specific, our analysis is quite close to [5], in which they also apply a first-order to zero-order approximation analysis. However, we point out several significant differences that blocks us to directly extend their approach to our work:

- (1) Unlike the pure data-driven inventory control problem that we have access to unbiased samples of the underlying demand distributions, in the data-driven joint pricing and inventory control problem we are unable to obtain unbiased samples directly after observing the demands, since the random demand model is a combination of random noise and an unknown function of price. This is also pointed out by [2].
- (2) In the pure data-driven inventory control problem, the point estimate of the underlying optimal base-stock policy is sufficient to describe a data-driven policy. However, in the data-driven joint pricing and inventory control problem,

we also need to estimate an implicit function defined uniformly on an interval (specified by the optimal list-price base-stock policy), which is clearly a much harder task in order for a uniform first-order approximation argument to work.

- (3) Instead of approximating a one-dimensional dynamic program as in the pure data-driven inventory control problem, the data-driven joint pricing and inventory control problem requires the approximation of a two-dimensional dynamic program.

1.3 Organization of the thesis

The rest of this thesis is organized as follows: Chapter 2 formulates the model and presents the full-information joint pricing and inventory control problem under linear demand models and the data-driven counterpart. Chapter 3 describes the multi-period data-driven approximation scheme. Chapter 4 contains the first-order analysis and the zero-order analysis of the data-driven approximation scheme, and a concrete data-driven algorithm and the corresponding running time upper bounds and sample complexity bounds. Chapter 5 extends the result in Chapter 4 by showing the same results (only trivially differ) hold for joint pricing and capacitated inventory control models under general concave demand functions as well. Chapter 6 shows results from numerical experiments to provide more insights into the performance of our algorithm in practice. Chapter 7 contains conclusions and discussions about future research.

Chapter 2

Joint Pricing and Inventory Control Model

In this chapter, we specify the basic joint pricing and inventory control model. We state the model assumptions, derive some structural results if we have full information of the model, and present the data-driven model if we have only past data.

We consider a retailer who sells a single item and makes periodic decisions on the item's pricing and inventory strategies. The pricing and inventory decisions affect revenues gained and costs incurred by the retailer, whose goal is to maximize the total profit (total revenue less total cost) over a finite planning horizon of T periods. The periods are numbered $t = 1, \dots, T$. At the beginning of each period, the retailer makes two decisions: the size of the inventory replenishment (if any) and the price of the item to be charged in that period. We assume all replenishment orders are satisfied with zero lead times. After the decisions are made and replenishment orders satisfied, demand, which depends on the price, is revealed. The retailer then uses inventory on hand to satisfy the demand, and backlogs all unsatisfied demand if the

demand is greater than the inventory.

We assume demands are independent over all periods, and in each period t , the demand is given by a stochastic demand function of the price for period t :

$$G_t(p_t, \eta_t) = D_t(p_t) + \eta_t,$$

where p_t is the price for period t , and η_t is a random noise with mean zero.

For each period t , we consider a finite interval $[p_t^{\min}, p_t^{\max}]$ as the set of feasible prices where

p_t^{\min} = lowest price allowed to be charged in period t , and

p_t^{\max} = highest price allowed to be charged in period t .

The expected demand function $D_t(p_t)$ is assumed to be real-valued and strictly decreasing in p_t . As a result, the inverse demand function $D_t^{-1}(\cdot)$ exists, and thus $p_t = D_t^{-1}(d_t)$, where d_t is the expected demand associated with price p_t , i.e., $d_t = D_t(p_t)$. In addition, d_t is confined in the finite interval, $[D_t(p_t^{\max}), D_t(p_t^{\min})]$, which is also denoted as $[d_t^{\min}, d_t^{\max}]$.

Assumption 1 *The expected demand function is defined to be linear: $D_t(p_t) = 2\kappa_t p_t + \theta_t$ and some prior knowledge $|\kappa_t|^{\min} \leq |\kappa_t| \leq |\kappa_t|^{\max}$, $|\theta_t| < |\theta_t|^{\max}$ is given. Moreover, we assume $\kappa_t < 0$ and $|\kappa_t|^{\min} > 0$ so that the expected demand function is strictly decreasing.*

In particular, we mainly consider the linear demand model in this thesis. However, the obtained result can be extended to more complex demand functions, for which the discussion will be deferred. We also assume some prior knowledge regarding

the value of the slope κ_t and the value of the intercept θ_t is known as a technical requirement. Indeed, this is not necessarily needed in our analysis for the linear demand model, but it will allow us to generalize the result for other demand models.

Assumption 2 *The random noise η_t is a continuous random variable with zero mean and a finite support $[\omega_t^{\min}, \omega_t^{\max}]$.*

We assume the random noise η_t is continuous and has a finite support $[\omega_t^{\min}, \omega_t^{\max}]$. Let F_t denote the cumulative distribution function (c.d.f.) of η_t . We also assume that each $F_t(\cdot)$ is Lipschitz-continuous with constant l_t .

Assumption 3 *The c.d.f. of η_t , F_t , is Lipschitz-continuous with constant l_t , i.e., $|F_t(\eta_t^1) - F_t(\eta_t^2)| \leq l_t |\eta_t^1 - \eta_t^2|$ for all η_t^1 and η_t^2 in $[\omega_t^{\min}, \omega_t^{\max}]$.*

We use x_t to denote the inventory level at the start of period t and y_t to denote the inventory level after the replenishment order is delivered. Therefore, $y_t - x_t$ is the size of the replenishment order, and we must have $y_t \geq x_t$. After the stochastic demand $G_t(p_t, \eta_t)$ is realized, there are three types of costs that can arise: (1) inventory ordering cost; (2) inventory holding cost when the inventory after replenishment is greater than the demand realized ($y_t > G_t(p_t, \eta_t)$); and (3) backlogging cost when there is excess demand ($G_t(p_t, \eta_t) > y_t$). For the inventory ordering cost, a unit ordering cost c_t is charged on each unit in the replenishment order of size $y_t - x_t$, and we assume $c_t < b_t$ otherwise the case is trivial (no ordering will occur). For each unit of the excess inventory or the excess demand, a unit holding cost h_t or a unit backlogging cost b_t is charged. Thus, for the ending inventory of $y_t - G_t(p_t, \eta_t)$, the cost is $h_t (y_t - G_t(p_t, \eta_t))^+ + b_t (-(y_t - G_t(p_t, \eta_t)))^+$. Here, the unit holding cost h_t and unit backlogging cost b_t are deterministic and known to the retailer in advance. The ending inventory $y_t - G_t(p_t, \eta_t)$ becomes the starting inventory for the next

period $t + 1$, i.e., $x_{t+1} = y_t - G_t(p_t, \eta_t)$. With expectation taken over η_t , the expected cost, as a function of y_t and p_t , is defined as

$$C_t(y_t, p_t) = \mathbb{E}_{\eta_t} [h_t (y_t - G_t(p_t, \eta_t))^+ + b_t (G_t(p_t, \eta_t) - y_t)^+] + c_t y_t.$$

We define the expected revenue function (of price p_t) as $R_t(p_t) = \mathbb{E}_{\eta_t} [p_t G_t(p_t, \eta_t)] = p_t D_t(p_t)$, since $\mathbb{E}_{\eta_t} G_t(p_t, \eta_t) = D_t(p_t)$. Because of the one-to-one correspondence between p_t and d_t , the expected revenue function can also be formulated in terms of d_t : $R_t(d_t) = d_t D_t^{-1}(d_t)$. We assume the expected revenue is strictly concave in expected demand d_t (but not necessarily in price p_t). Without loss of generality, We assume $c_t = 0$ since any nonzero c_t can be absorbed into h_t, b_t and $R_t(d_t)$. The same assumption is also made in [11], [5] and [2].

The expected profit in period t is the expected revenue less the expected cost and is defined as $g_t(y_t, p_t) = R_t(p_t) - C_t(y_t, p_t)$. We also assume that the starting inventory for period 1 is known and that leftover inventory after period T has no salvage value. The retailer's goal is to choose y_t and p_t for each period t , so that the total expected profit over the T periods is maximized, while conforming to the constraints that $y_t \geq x_t$ for every t .

In reality, the expected demand function $D_t(\cdot)$ and the distribution $F_t(\cdot)$ of η_t are rarely known to the retailer, but it is easy to collect past data of the model. In the next subsection, we establish some structural results of the full-information model when the retailer has access to $D_t(\cdot)$ and $F_t(\cdot)$. In the last subsection, we discuss the data-driven problem, in which the retailer has past data of the model but does not know $D_t(\cdot)$ or $F_t(\cdot)$.

2.1 Full-Information Problem

If the retailer has full information about the model, i.e., $D_t(\cdot)$ and $F_t(\cdot)$ are known, the expected profit maximization problem is the classic finite-horizon joint pricing and inventory control problem. This problem naturally bears a dynamic programming structure, with the Bellman equation recursively defined as

$$\begin{cases} V_{T+1}(x_{T+1}) = 0, \\ V_t(x_t) = \max_{p_t \in [p_t^{\min}, p_t^{\max}], y_t \geq x_t} \{g_t(y_t, p_t) + \mathbb{E}_{\eta_t}[V_{t+1}(y_t - G_t(p_t, \eta_t))]\}, \forall t = T, \dots, 1. \end{cases}$$

Using the dynamic programming structure, we can compute the optimal solution with backward induction, starting from period T to period 1. The optimal strategy in this problem when we do not put restriction on inventory levels is well-known as a “base-stock list-price” policy, which consists of a base inventory level and a list price, denoted respectively as S_t^* and P_t^* , in each period t . If the starting inventory x_t in period t is below the base inventory level S_t^* , then the base-stock list-price policy orders inventory up to the base inventory level and charges the list price as the unit price. That is, it sets $y_t = S_t^*$ and $p_t = P_t^*$. If the starting inventory x_t is above the base inventory level S_t^* , zero additional inventory is ordered and a price discount on the list price is offered, i.e., $y_t = x_t$ and $p_t \leq P_t^*$. In this case, the optimal price p_t is the price that maximizes the total expected profit over the interval of periods from t to T , if the retailer orders the inventory level up to y_t in period t . Thus, the optimal price is a function of y_t . However, this function is highly dependent on the problem itself, and it can rarely be explicitly defined.

Fortunately, if we shift our analysis of this problem from using y_t and p_t as the decision variables to using y_t and d_t , some useful structural results can be established.

Since $D_t(\cdot)$ is known, determining p_t is equivalent to determining its corresponding expected demand $d_t = D_t(p_t)$ (and vice versa). Therefore, we can formulate an equivalent problem in which the retailer periodically determines inventory levels y_t and expected demands d_t . If the retailer sets the expected demand as d_t , the stochastic demand is simply $d_t + \eta_t$. The expected revenue function in terms of d_t is defined as $R_t(d_t) = d_t D_t^{-1}(d_t)$, and the expected cost function in terms of y_t and d_t is defined as $C_t(y_t, d_t) = \mathbb{E}_{\eta_t} [h_t (y_t - d_t - \eta_t)^+ + b_t (d_t + \eta_t - y_t)^+]$.

Fact 1 $R_t(d_t)$ is strictly concave in d_t , and $C_t(y_t, d_t)$ is jointly convex in y_t and d_t .

Hence, by Fact 1, the expected profit function, which is defined as $g_t(y_t, d_t) = R_t(d_t) - C_t(y_t, d_t)$, is jointly concave in y_t and d_t (strictly in d_t).

The Bellman equation can be reformulated in terms of y_t and d_t :

$$\begin{cases} V_{T+1}(x_{T+1}) = 0, \\ V_t(x_t) = \max_{d_t \in [d_t^{\min}, d_t^{\max}], y_t \geq x_t} \{g_t(y_t, d_t) + \mathbb{E}_{\eta_t}[V_{t+1}(y_t - d_t - \eta_t)]\}, \forall t = T, \dots, 1. \end{cases}$$

The value function $V(x_t)$ is the optimal expected profit over the periods from t to T , assuming that the starting inventory level in period t is x_t and the retailer is going to make optimal decisions in periods from $t + 1$ to T . Let $U_t(y_t, d_t)$ be the expected profit in the interval $[t, T]$, if the retailer sets the inventory level to y_t and the expected demand to d_t in period t , assuming optimal decisions are going to be made in periods $[t + 1, T]$. Therefore, $U_t(y_t, d_t) = g_t(y_t, d_t) + \mathbb{E}_{\eta_t}[V_{t+1}(y_t - d_t - \eta_t)]$, and $V_t(x_t) = \max_{d_t \in [d_t^{\min}, d_t^{\max}], y_t \geq x_t} U_t(y_t, d_t)$. Since $g_t(\cdot, \cdot)$ is jointly concave, we have the following facts:

Fact 2 $U_t(y_t, d_t)$ is jointly concave in y_t and d_t , and, in particular, strictly concave in d_t .

Fact 3 $V_t(x_t)$ is concave and nonincreasing in x_t .

The optimal base-stock list-price policy, (S_t^*, P_t^*) , translates to the optimal base-stock list-demand policy, which is $(S_t^*, D_t^*) = (S_t^*, D_t(P_t^*))$. Therefore, if the starting inventory x_t is below S_t^* , the inventory level is raised to the base inventory level and the expected demand is set as the list demand, i.e., $y_t = S_t^*$ and $d_t = D_t^*$. If x_t is above S_t^* , no inventory is ordered ($y_t = x_t$) and the expected demand is set to the corresponding optimal demand, which depends on y_t . We define this function as $d_t^*(y_t) = \arg \max_{d_t \in [d_t^{\min}, d_t^{\max}]} U_t(y_t, d_t)$, which is the optimal demand (in period t) that maximizes the total expected profit over the periods from t to T , if the inventory level is set to y_t in period t . Let $W_t(y_t)$ denote the expected profit in the interval $[t, T]$, if the retailer sets the inventory level to y_t and the expected demand to $d_t^*(y_t)$ in period t , assuming the optimal base-stock list-demand policy is followed in periods $[t+1, T]$. Hence, $W_t(y_t) = U_t(y_t, d_t^*(y_t))$ and $V_t(x_t) = \max_{y_t \geq x_t} W_t(y_t)$. The optimal base stock S_t^* is defined as a maximizer of $W_t(y_t)$, and $D_t^* = d_t^*(S_t^*)$. Therefore, the optimal base-stock list-demand policy, (S_t^*, D_t^*) , can also be characterized by the equations:

$$V_t(x_t) = \begin{cases} W_t(S_t^*) = U_t(S_t^*, D_t^*), & x_t \leq S_t^*, \\ W_t(x_t) = U_t(x_t, d_t^*(x_t)), & x_t > S_t^*. \end{cases}$$

Fact 4 $W_t(y_t)$ is concave in y_t .

Fact 5 $y_t - d_t^*(y_t)$ is nondecreasing in y_t .

As we assume that η_t is continuous in Assumption 2, the partial derivatives of $C_t(y_t, d_t)$ with respect to y_t and d_t exist, and $\frac{\partial}{\partial y_t} C_t(y_t, d_t) = -\frac{\partial}{\partial d_t} C_t(y_t, d_t)$. In addition, since $R_t'(d_t)$ exists, it can be shown that $\frac{\partial}{\partial y_t} U_t(y_t, d_t)$, $\frac{\partial}{\partial d_t} U_t(y_t, d_t)$, $V_t'(x_t)$ and $W_t'(y_t)$ all exist. Since $W_t(y_t)$ is concave in y_t (Fact 4), any y_t in the set

$\{y_t : W_t'(y_t) = 0\}$ is a maximizer of $W_t(y_t)$, and among all the maximizers, we choose the largest maximizer to be the optimal base stock, i.e., $S_t^* = \max\{y_t : W_t'(y_t) = 0\}$.

Suppose the starting inventory x_t for period t is confined in $[x_t^{\min}, x_t^{\max}]$. Since $W_t(y_t)$ is concave, increasing y_t beyond $\max\{x_t^{\max}, S_t^*\}$ only reduces the total expected profit. Thus, the range of optimal inventory levels y_t is bounded in a finite interval, if the retailer follows the optimal base-stock list-price policy. As a result, there is no point in considering an infinite range of inventory levels, and, therefore, we limit y_t in a finite range of values that are of interest. We denote it as $[y_t^{\min}, y_t^{\max}]$. Since there is no starting inventory for period 1, i.e., $[x_1^{\min}, x_1^{\max}] = \{0\}$, and η_t has a finite support $[\omega_t^{\min}, \omega_t^{\max}]$ (Assumption 2), it is easy to show by induction that $[x_t^{\min}, x_t^{\max}]$ is finite in all periods, and hence limiting y_t in the finite range $[y_t^{\min}, y_t^{\max}]$ is always valid.

Because, for any fixed y_t , the function $U_t(y_t, d_t)$ is strictly concave in d_t (Fact 2), it has a unique global maximizer $d_t^\#(y_t)$, which satisfies the condition that $\frac{\partial}{\partial d_t} U_t(y_t, d_t^\#(y_t)) = 0$. Since $d_t^*(y_t)$ is the constrained maximizer of $U_t(y_t, d_t)$ over the interval $[d_t^{\min}, d_t^{\max}]$, $d_t^*(y_t)$ is the same as $d_t^\#(y_t)$ if $d_t^\#(y_t) \in [d_t^{\min}, d_t^{\max}]$, and $d_t^*(y_t)$ takes value of d_t^{\min} or d_t^{\max} otherwise.

Assumption 4 *The range of the possible expected demand levels in period t , $[d_t^{\min}, d_t^{\max}]$, contains $d_t^\#(y_t)$ for all $y_t \in [y_t^{\min}, y_t^{\max}]$, where $[y_t^{\min}, y_t^{\max}]$ is the range of inventory levels of interest.*

Assumption 4 guarantees that, for any the inventory level y_t we might consider (y_t in $[y_t^{\min}, y_t^{\max}]$), $d_t^*(y_t) = d_t^\#(y_t)$, and therefore $d_t^*(y_t)$ always satisfies the first order optimality condition: $\frac{\partial}{\partial d_t} U_t(y_t, d_t^*(y_t)) = 0$, from which we develop the two following lemmas:

Lemma 1 $0 \leq d_t^{*'}(y_t) \leq 1$ for any y_t in $[y_t^{\min}, y_t^{\max}]$.

Lemma 2 $V'_t(\cdot)$ is Lipschitz-continuous with constant $|\kappa_t|$ over the interval $[x_t^{\min}, x_t^{\max}]$.

The structural results of the full-information model are crucial in the analysis of the performance of the data-driven approximation algorithm. In order to make a $(1 - \epsilon)$ -approximation meaningful, we make the last assumption about the full-information problem:

Assumption 5 For any y_t, d_t in the range of the possible expected demand level and inventory level $y_t \in [y_t^{\min}, y_t^{\max}]$, $d_t \in [d_t^{\min}, d_t^{\max}]$, the expected profit is positive, i.e. $U_t(y_t, d_t) > 0$.

This assumption guarantees that the underlying problem we try to solve is meaningful, in the sense that there exist feasible $(1 - \epsilon)$ -approximations to this problem. Note for any given problem that has bounded negative expected profit, we can always add a bounded positive constant to make it satisfy Assumption 5, so this assumption is without loss of generality.

2.2 Data-Driven Problem

The embedded joint pricing and inventory control model in the data-driven problem is the same as that in the full-information problem. However, in the data-driven problem, the retailer has no access to $D_t(\cdot)$ or $F_t(\cdot)$, and therefore cannot compute the optimal base-stock list-price policies. Instead, the retailer has past data of the model, and wants to use algorithms, which take the data as inputs and compute the inventory and pricing decisions that can generate “near-optimal” profits. We measure the “near-optimality” by the relative loss of profit between the profit generated if the retailer uses a data-driven algorithm and the profit generated if the optimal base-stock list-price policy is followed.

In each period t , we have a series of data in the form of price-demand pairs. There are $K_t + 1$ distinct prices (denoted as $p_t^0, \dots, p_t^{K_t}$) in the range $[p_t^{\min}, p_t^{\max}]$, and for each price p_t^i , there are N_t demand samples (denoted as $d_t^{i1}, \dots, d_t^{iN_t}$):

$$K_t + 1 \left\{ \begin{array}{l} \overbrace{(d_t^{01}, \dots, d_t^{0j}, \dots, d_t^{0N_t})}^{N_t} \\ \vdots \\ p_t^i : (d_t^{i1}, \dots, d_t^{ij}, \dots, d_t^{iN_t}) \\ \vdots \\ p_t^{K_t} : (d_t^{K_t 1}, \dots, d_t^{K_t j}, \dots, d_t^{K_t N_t}) \end{array} \right. .$$

To ensure the entire range $[p_t^{\min}, p_t^{\max}]$ is covered, we assume that $p_t^0 = p_t^{\min}$ and $p_t^{K_t} = p_t^{\max}$. However, we do not make assumptions on the distribution of the prices within (p_t^{\min}, p_t^{\max}) .

For each price p_t^i , we use d_t^i to denote its corresponding expected demand, i.e., $d_t^i = D_t(p_t^i)$. Then, for each price-demand pair, (p_t^i, d_t^i) , the sample demand d_t^{ij} is the sum of the expected demand corresponding to p_t^i and one realization of the random noise η_t , i.e., $d_t^{ij} = d_t^i + \eta_t^{ij}$. The sample revenue r_t^{ij} is the product of p_t^i and d_t^{ij} , and therefore $r_t^{ij} = p_t^i d_t^{ij} = p_t^i d_t^i + p_t^i \eta_t^{ij} = R_t(d_t^i) + p_t^i \eta_t^{ij}$, where $R_t(\cdot)$ is the expected revenue (in terms of d_t).

We use \bar{d}_t^i to denote the sample mean demand corresponding to p_t^i , i.e., $\bar{d}_t^i = (\sum_{j=1}^{N_t} d_t^{ij}) / N_t = d_t^i + (\sum_{j=1}^{N_t} \eta_t^{ij}) / N_t$. Similarly, the sample mean revenue \bar{r}_t^i is defined as $\bar{r}_t^i = (\sum_{j=1}^{N_t} r_t^{ij}) / N_t = R_t(d_t^i) + p_t^i (\sum_{j=1}^{N_t} \eta_t^{ij}) / N_t$. Let Δ_t^i denote $(\sum_{j=1}^{N_t} \eta_t^{ij}) / N_t$. We define $\hat{\eta}_t^{ij} = d_t^{ij} - \bar{d}_t^i = \eta_t^{ij} - \Delta_t^i$ as the biased samples of η_t , where Δ_t^i is the bias term.

Since we are now only considering the linear demand model, we just assume

$K_t = 1$, which is clearly sufficient for our purpose (in the case we have $K_t > 1$, we can simply ignore the price-demand pairs with indices larger than 1).

Remark 1 *Since $D_t(\cdot)$ is unknown, the retailer cannot directly observe the values of $D_t(p_t^i)$ and the true samples η_t^{ij} , from price-demand pairs (p_t^i, d_t^{ij}) . An intuitive remedy is to compute the sample means of the demand, and then compute the (biased) samples of η_t . However, because of the bias terms, this method leaves two problems:*

1. *The sample mean demand or sample mean revenue data are not the true function values of $D_t(\cdot)$ or $R_t(\cdot)$, which we may want to approximate in the data-driven algorithm;*
2. *The values of the expected cost can be biased, if the expectation is taken over the biased samples of η_t .*

Fact 6 *For any $\alpha > 0$, $\mathbb{P}[|\Delta_t^i| \leq \alpha] \geq 1 - 2 \exp\left(-\frac{2N_t\alpha^2}{(\omega_t^{\max} - \omega_t^{\min})^2}\right)$.*

However, Fact 6, as a direct result of Hoeffding's inequality (Lemma 9 in Appendix D) applied to the sample mean of η_t , guarantees that the bias term Δ_t^i is small with high probability if N_t is large.

Chapter 3

The Multi-Period Data-Driven Approximation

In this chapter, we describe a data-driven approximation scheme, which takes a data set in the form specified in Section 2.2 and constructs a data-driven dynamic program, that mimics the behavior of the full-information dynamic program.

3.1 Empirical Revenue and Cost Functions

We use $(p_t^0, \bar{d}_t^0), (p_t^1, \bar{d}_t^1)$ to first construct $\hat{D}_t(p_t)$. In particular, we compute

$$\hat{\kappa}_t = \frac{\bar{d}_t^1 - \bar{d}_t^0}{2(p_t^1 - p_t^0)},$$
$$\hat{\theta}_t = \bar{d}_t^0 - 2\hat{\kappa}_t p_t^0.$$

If we obtain $\hat{\kappa}_t > -|\kappa_t^{\min}|$ then set $\hat{\kappa}_t = -|\kappa_t^{\min}|$ (when $\hat{\kappa}_t < -|\kappa_t^{\max}|$, set $\hat{\kappa}_t = -|\kappa_t^{\max}|$), and if $|\hat{\theta}_t| > |\theta_t^{\max}|$ then we set $|\hat{\theta}_t| = |\theta_t^{\max}|$ (do not change the sign). So

we have $\hat{D}_t(p_t) = 2\hat{\kappa}_t p_t + \hat{\theta}_t$ and $\hat{D}_t^{-1}(d_t) = \frac{1}{2\hat{\kappa}_t} d_t - \frac{\hat{\theta}_t}{2\hat{\kappa}_t}$. Then, the empirical revenue function is defined as

$$\hat{R}_t(d_t) = d_t \hat{D}_t^{-1}(d_t).$$

The empirical cost function $\hat{C}_t(y_t, d_t)$ is defined using Sample Average Approximation (SAA) with a set of the biased samples $(\hat{\eta}_t^{0j})$ of size N_t . We use $\mathbb{E}_{\hat{\eta}_t}[f(\hat{\eta}_t)]$ to denote $\frac{1}{N_t} \sum_{j=1}^{N_t} f(\hat{\eta}_t^{0j})$. Therefore, $\hat{C}_t(y_t, d_t)$ is defined as

$$\begin{aligned} \hat{C}_t(y_t, d_t) &= \frac{1}{N_t} \sum_{j=1}^{N_t} \left(h_t (y_t - d_t - \hat{\eta}_t^{0j})^+ + b_t (d_t + \hat{\eta}_t^{0j} - y_t)^+ \right) \\ &= \mathbb{E}_{\hat{\eta}_t} [h_t (y_t - d_t - \hat{\eta}_t)^+ + b_t (d_t + \hat{\eta}_t - y_t)^+]. \end{aligned}$$

Remark 2 *Although we have $(K_t + 1)N_t$ biased samples in the data set, we decide to use only one set of size N_t , $(\hat{\eta}_t^{0j})$, to construct the empirical cost function by SAA. If we chose to use all the $(K_t + 1)N_t$ samples, we had to make sure all the samples are close to the true sample values with high probability. However, the bound on the probability deteriorates for large K_t in our analysis, which somewhat contradicts the idea that a larger data set generates better performance lower bound over a smaller data set. As a result, we drop the dependence on K_t when constructing $\hat{C}_t(y_t, d_t)$, and this solves the problem nicely.*

The function $\hat{C}_t(y_t, d_t)$ is not differentiable on its breakpoints. Therefore, we use $\frac{\partial_-}{\partial y_t} \hat{C}_t(y_t, d_t)$ and $\frac{\partial_+}{\partial y_t} \hat{C}_t(y_t, d_t)$ to denote the left and right partial derivatives of $\hat{C}_t(y_t, d_t)$. Note \hat{C}_t can also be represented as a univariate function of $y_t - d_t$, with left and right derivative $\hat{C}_t^l(y_t - d_t)$ and $\hat{C}_t^r(y_t - d_t)$.

3.2 Functions in the Empirical Dynamic Program

Similar to the functions $V_t(x_t)$, $U_t(y_t, d_t)$ and $W_t(y_t)$ in the full-information model, the empirical functions $\hat{V}_t(x_t)$, $\hat{U}_t(y_t, d_t)$ and $\hat{W}_t(y_t)$ are defined recursively, from period T to period 1:

1. Define $\hat{V}_{T+1}(x_t) = 0$. Let $t = T$.
2. Define $\hat{U}_t(y_t, d_t) = \hat{R}_t(d_t) - \hat{C}_t(y_t, d_t) + \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}(y_t - d_t - \hat{\eta}_t)]$.
3. Define $\hat{d}_t^*(y_t) \in \arg \max_{d_t \in [d_t^{\min}, d_t^{\max}]} \hat{U}_t(d_t, y_t)$.
4. Define $\hat{W}_t(y_t) = \hat{U}_t(y_t, \hat{d}_t^*(y_t))$.
5. Define empirical base stock and list demand: \hat{S}_t^* is the largest number \hat{S}_t^* is the largest number such that $\hat{S}_t^* \in \arg \max_{y_t \in [y_t^{\min}, y_t^{\max}]} \hat{W}_t(y_t)$. Define $\hat{D}_t^* = \hat{d}_t^*(\hat{S}_t^*)$.
6. Define $\hat{V}_t(x_t) = \begin{cases} \hat{W}_t(\hat{S}_t^*) = \hat{U}_t(\hat{S}_t^*, \hat{D}_t^*), & x_t \leq \hat{S}_t^*, \\ \hat{W}_t(x_t) = \hat{U}_t(x_t, \hat{d}_t^*(x_t)), & x_t > \hat{S}_t^*. \end{cases}$
7. Repeat 2. to 6. for $t = T - 1, \dots, 1$.

Using the computed empirical base-stock list-demand policy $(\hat{S}_t^*, \hat{D}_t^*)$ (along with the function $\hat{d}_t^*(y_t)$), the retailer can make pricing and inventory decisions in each period without knowing the underlying demand function or distribution of the random noise. Moreover, we point out the empirical dynamic program mimics the behavior of the full-information dynamic program and preserves concavity.

Fact 7 For all $t = 1, \dots, T$, $\hat{V}_t(x_t)$ is concave in x_t and $\hat{U}_t(y_t, d_t)$ is jointly concave in y_t, d_t .

We briefly mention the proof of this result, since it is identical to the proof for the full-information problem (i.e., Fact 2 and 3). In the base case, the joint concavity of $\hat{U}_T(y_T, d_T)$ is easy to verify. Then the concavity of $\hat{V}_T(x_T)$ will be proved by showing $\hat{W}_T(y_T) = \hat{U}_T(y_T, \hat{d}_T^*(y_T))$ is concave. Consider $y'_T < y''_T$ both belong to $[y_T^{\min}, y_T^{\max}]$ and an arbitrary constant $0 < \lambda < 1$. Then we have $\hat{U}_T(\lambda y'_T + (1 - \lambda)y''_T, \hat{d}_T^*(\lambda y'_T + (1 - \lambda)y''_T)) \geq \hat{U}_T(\lambda y'_T + (1 - \lambda)y''_T, \lambda \hat{d}_T^*(y'_T) + (1 - \lambda)\hat{d}_T^*(y''_T))$ from the definition of $\hat{d}_T^*(y_T)$ as maximizer. Then it is easy to check $\hat{U}_T(\lambda y'_T + (1 - \lambda)y''_T, \lambda \hat{d}_T^*(y'_T) + (1 - \lambda)\hat{d}_T^*(y''_T)) \geq \lambda \hat{U}_T(\hat{d}_T^*(y'_T)) + (1 - \lambda)\hat{U}_T(\hat{d}_T^*(y''_T))$ using the concavity of \hat{R}_T and \hat{C}_T (regarded as an univariate function of $y_T - d_T$.) Then we complete the proof by induction, assuming $\hat{U}_{t+1}(y_{t+1}, d_{t+1})$ is jointly concave and $\hat{V}_{t+1}(x_{t+1})$ is concave. Again, the joint concavity of $\hat{U}_t(y_t, d_t)$ is straightforward since $\hat{V}_{t+1}(x_{t+1})$ is assumed to be concave. Then, following the same argument as in the base case we can show $\hat{V}_t(x_t)$ is concave.

Fact 8 Suppose $N_t \geq \frac{B_t \log(18/\beta)}{\kappa_t^2}$ for all $t = 1, \dots, T$, if we exploit a stronger version of Assumption 4, such that in each period t the interval $[d_t^{\min} + (T - s + 1)/|\kappa_s|, d_t^{\max} - (T - s + 1)/|\kappa_s|]$, contains $d_t^\#(y_t)$ for all $y_t \in [y_t^{\min}, y_t^{\max}]$, then the interval $[d_t^{\min}, d_t^{\max}]$ contains a point $d_t^\$(y_t)$ such that $0 \in \left[\frac{\partial_+}{\partial d_t} \hat{U}_t(y_t, d_t^\$(y_t)), \frac{\partial_-}{\partial d_t} \hat{U}_t(y_t, d_t^\$(y_t)) \right]$ with probability at least $1 - \beta$.

For technical issues, we also require a stronger version of Assumption 4, such that in each period t the interval $[d_t^{\min}, d_t^{\max}]$ also contains a point $d_t^\$(y_t)$ such that $0 \in \left[\frac{\partial_+}{\partial d_t} \hat{U}_t(y_t, d_t^\$(y_t)), \frac{\partial_-}{\partial d_t} \hat{U}_t(y_t, d_t^\$(y_t)) \right]$ with high probability, as stated by Fact 8. This fact is due to Lemma 7 and the induction argument in the proof of Theorem 3. So a more rigorous argument should actually require us to combine the proof of this fact into the proof of Theorem 3, but we feel the separate statement of the fact makes the induction proof less involved and gives better clarity. We will use this

stronger version of Assumption 4 throughout the rest of the thesis.

Fact 9 *Given the stronger version of Assumption 4, with probability at least $1 - \beta$, $\hat{d}_t^*(y_t)$ is a nondecreasing piecewise linear function with left derivative less than or equal to 1 (equal to 1 when $\left| \frac{\partial}{\partial d_t} \hat{U}_t(\hat{d}_t^*(y_t), y_t) \right| > 0$).*

Given Fact 8, we point out a procedure of computing the value of $\hat{d}_t(y_t^*)$. We illustrate Fact 9 via Figure B-1. Consider the range $[d_t^{\min}, d_t^{\max}]$ large enough so that computing $\hat{d}_t^*(y_t)$ is equivalent to minimizing the difference between $\hat{R}'_t(d_t)$ and $\hat{f}_t^l(y_t - d_t)$, for any fixed $y_t \in [y_t^{\min}, y_t^{\max}]$. Define $\hat{f}_t^l(y_t - d_t) = -\hat{C}_t(y_t - d_t) + \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}(y_t - d_t - \hat{\eta}_t)]$. Recall from Fact 7 we have shown $f_t(y_t - d_t)$ is convex in d_t and $\hat{R}_t(d_t)$ is concave in d_t . We take $y_1 < y_2 < y_3$ all belonging to $[y_t^{\min}, y_t^{\max}]$ and note increasing y_t^1 to y_t^2 is equivalent to “shifting” $\hat{f}_t^l(y_t - d_t)$ to the right by $y_t^2 - y_t^1$, as an univariate function of d_t . It is worth pointing out that $\hat{f}_t^l(y_t - d_t)$ is a nondecreasing piecewise-linear function in d_t , by an easy induction argument.

When $\hat{R}'_t(d_t)$ intersects with $\hat{f}_t^l(y_t - d_t)$, then clearly the intersection point produces $\hat{d}_t^*(y_t)$ that sets the difference to zero. And this also shows $\hat{d}_t^*(y_t)$ remains linear for the part $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, \hat{d}_t^*(y_t)) = 0$ (e.g., $\hat{d}_t^*(y_t^1), \hat{d}_t^*(y_t^2)$ in Figure B-1, and it is easy to see $\hat{d}_t^*(y_t^2) - \hat{d}_t^*(y_t^1) \leq y_t^2 - y_t^1$). On the other hand, when the two function “intersect” at some vertical part of $\hat{f}_t^l(y_t - d_t)$ (so they do not actually intersect), by no means can we find $d_t \in [d_t^{\min}, d_t^{\max}]$ such that the difference becomes zero. However, it is still easy to verify the intersection point produces $\hat{d}_t^*(y_t)$ the difference. Also, it is immediate to see in such case $\hat{d}_t^*(y_t)$ changes exactly as y_t changes. For example, in Figure 1 consider $y_t^3 = y_t^2 + \delta_t$ for some $\delta_t > 0$, then we have $\hat{d}_t^*(y_t^3) = \hat{d}_t^*(y_t^2) + \delta_t$.

In the next section, we analyze the performance of our data-driven algorithm and provide a lower bound on the profit the retailer would obtain, if the empirical

base-stock list-demand policy was followed, compared to the profit if the optimal policy was followed.

Chapter 4

Theoretical Algorithm Analysis for linear demand models

In this chapter, we conduct theoretical analysis on the sample complexity bound on the number of samples N_t needed, to achieve a near-optimal data-driven policy. We also discuss how the exact algorithm is and the importance of sparsification in order to achieve efficient implementation.

4.1 The Sample Complexity Bound

Our proof is conducted from first proving the first order uniform approximation bounds, i.e., for all $t, d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}]$, $\left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) \right|$ and $\left| W_t'(y_t) - \hat{W}_t'(y_t) \right|$ are uniformly bounded with high probability, and the difference can be made arbitrarily small given we have large enough N_t (c.f. Theorem 3); then, we use the first order approximation bounds to show zero-order approximation bounds, that is, for all $y_t \in [y_t^{\min}, y_t^{\max}]$, $U_t(\hat{d}_t^*(y_t), y_t) \geq (1 - \epsilon_t^U) U_t(d_t^*(y_t), y_t)$,

$W_t(\hat{S}_t^*) \geq (1 - \epsilon_t^W)W_t(S_t^*)$ with high probability, and the positive constant $\epsilon_t^U, \epsilon_t^W$ can be made arbitrarily small given large enough N_t (c.f. Lemma 8). Then, we complete the proof by an induction argument (c.f., Theorem 4 and 5).

4.1.1 The First-Order Analysis

Lemma 3 gives the first-order approximation bounds on $\left| R_t'(d_t) - \hat{R}_t'(d_t) \right|$, that is to say, given sufficiently large N_t s, we have for all $t = 1, \dots, T$, the derivative function of $\hat{R}_t(d_t)$ approximates the derivative function of $R_t(d_t)$ uniformly on the interval $[d_t^{\min}, d_t^{\max}]$ with high probability.

Lemma 3 *Consider the empirical revenue function $\hat{R}_t(d_t)$ constructed using a data set, in which $N_t \geq \max \{ (2d_t^{\max} + |\theta_t|^{\max})^2, 16(p_t^0)^2(|\kappa_t|^{\max})^2, 4(p_t^1 - p_t^0)^2(|\kappa_t|^{\max})^2 \} \frac{2(\omega^{\max} - \omega^{\min})^2 \log(10/\beta)}{(p_t^1 - p_t^0)^2(|\kappa_t|^{\min})^4 \alpha^2}$. Then the derivative of $\hat{R}_t(d_t)$ uniformly approximates the derivative of $R_t(d_t)$ with high probability. That is, for any $\alpha > 0$,*

$$\mathbb{P} \left[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}] : \left| R_t'(d_t) - \hat{R}_t'(d_t) \right| \leq \alpha \right] \geq 1 - \beta.$$

The proof of Lemma 3 is provided in Appendix C.

Lemma 4 and Lemma 5 give the first-order approximation bounds on $\left| \frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial}{\partial y_t} \hat{C}_t(y_t, d_t) \right|$ and $\left| \frac{\partial}{\partial d_t} C_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{C}_t(y_t, d_t) \right|$, that is to say, given sufficiently large N_t s, we have for all $t = 1, \dots, T$, the left derivative function of $\hat{C}_t(y_t, d_t)$ approximates the derivative function of $C_t(y_t, d_t)$ uniformly on the interval $[d_t^{\min}, d_t^{\max}]$ and interval $[y_t^{\min}, y_t^{\max}]$ with high probability.

Lemma 4 *Consider the empirical cost function $\hat{C}_t(y_t, d_t)$ constructed using a data set, in which $N_t \geq \frac{2(2 \log 2 - \log \beta)(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2}{\alpha^2}$. Then the left partial derivative of $\hat{C}_t(y_t, d_t)$ with respect to y_t uniformly approximates the partial derivative of $C_t(y_t, d_t)$*

with respect to y_t with high probability. That is, for any $\alpha > 0$,

$$\mathbb{P} \left[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}] : \left| \frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial}{\partial y_t} \hat{C}_t(y_t, d_t) \right| \leq \alpha \right] \geq 1 - \beta.$$

The same result holds also for the right partial derivative of $\hat{C}_t(y_t, d_t)$ with respect to y_t , $\frac{\partial_+}{\partial y_t} \hat{C}_t(y_t, d_t)$.

The proof of Lemma 4 is provided in Appendix C.

Since $\frac{\partial}{\partial d_t} C_t(y_t, d_t) = -\frac{\partial}{\partial y_t} C_t(y_t, d_t)$ and $\frac{\partial}{\partial d_t} \hat{C}_t(y_t, d_t) = -\frac{\partial}{\partial y_t} \hat{C}_t(y_t, d_t)$, we have the following lemma:

Lemma 5 Consider the empirical cost function $\hat{C}_t(y_t, d_t)$ constructed using a data set, in which $N_t \geq \frac{2(2 \log 2 - \log \beta)(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2}{\alpha^2}$. Then, for any $\alpha > 0$,

$$\mathbb{P} \left[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}] : \left| \frac{\partial}{\partial d_t} C_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{C}_t(y_t, d_t) \right| \leq \alpha \right] \geq 1 - \beta.$$

The same result holds also for the right partial derivative of $\hat{C}_t(y_t, d_t)$ with respect to d_t , $\frac{\partial_+}{\partial d_t} \hat{C}_t(y_t, d_t)$.

Lemma 5 is a direct result of Lemma 4, and therefore its proof is omitted.

Lemma 6 gives the first-order approximation bounds on $|\mathbb{E}_{\eta_t}[V'_t(y_t - d_t - \eta_t)] - \mathbb{E}_{\hat{\eta}_t}[V'_t(y_t - d_t - \hat{\eta}_t)]|$, that is to say, given sufficiently large N_t s, we have for all $t = 1, \dots, T$, the function $\mathbb{E}_{\hat{\eta}_t}[V'_t(y_t - d_t - \hat{\eta}_t)]$ approximates the function $\mathbb{E}_{\eta_t}[V'_t(y_t - d_t - \eta_t)]$ uniformly on the interval $[d_t^{\min}, d_t^{\max}]$ and interval $[y_t^{\min}, y_t^{\max}]$ with high probability.

Lemma 6 Consider the true value function $V_t(y_t, d_t)$ and the biased samples $\hat{\eta}_t$ from a data set, in which $N_t \geq \frac{2(2 \log 2 - \log \beta)(\omega_t^{\max} - \omega_t^{\min})^2 \kappa_t^2}{\alpha^2}$. Then, for any $\alpha > 0$,

$$\mathbb{P} \left[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}] : \left| \mathbb{E}_{\eta_t}[V'_t(y_t - d_t - \eta_t)] - \mathbb{E}_{\hat{\eta}_t}[V'_t(y_t - d_t - \hat{\eta}_t)] \right| \leq \alpha \right] \geq 1 - \beta.$$

The proof of Lemma 6 is provided in Appendix C.

For convenience, we define

$$B_t = \max \left\{ \left\{ (2d_t^{\max} + |\theta_t|^{\max})^2, 16(p_t^0)^2(|\kappa_t|^{\max})^2, 4(p_t^1 - p_t^0)^2(|\kappa_t|^{\max})^2 \right\} \frac{18(\omega_t^{\max} - \omega_t^{\min})^2}{(p_t^1 - p_t^0)^2(|\kappa_t|^{\min})^4}, \right. \\ \left. 18(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2, 18(\omega_t^{\max} - \omega_t^{\min})^2 \kappa_t^2 \right\}.$$

Then, Theorem 1 states that given the knowledge $\hat{V}_{t+1}^l(x_t)$ uniformly approximates $V'_{t+1}(x_t)$ on the interval $[x_t^{\min}, x_t^{\max}]$, and given sufficiently large N_t s, we have for all $t = 1, \dots, T$, $\frac{\partial}{\partial} \hat{U}_t(y_t, d_t)$ uniformly approximates $\frac{\partial}{\partial d_t} U_t(y_t, d_t)$ on the interval $[d_t^{\min}, d_t^{\max}]$ and interval $[y_t^{\min}, y_t^{\max}]$ with high probability. The proof is mainly by piecing together the results of Lemma 3, Lemma 4, Lemma 5 and Lemma 6.

Theorem 1 *Suppose we are given that the empirical left derivative $\hat{V}_{t+1}^l(x_t)$ uniformly approximates $V'_{t+1}(x_t)$. That is, $|\hat{V}_{t+1}^l(x_t) - V'_{t+1}(x_t)| \leq \gamma$, for some $\gamma > 0$. Consider the empirical function $\hat{U}_t(y_t, d_t)$ constructed using a data set, in which $N_t \geq \frac{B_t \log(18/\beta)}{\alpha^2}$. Then the left partial derivative of $\hat{U}_t(y_t, d_t)$ with respect to d_t uniformly approximates the partial derivative of $U_t(y_t, d_t)$ with respect to d_t with high probability. That is, for any $\alpha > 0$, $\mathbb{P}[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}] : \left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \leq \alpha + \gamma] \geq 1 - \beta$. The same result holds also for the right partial derivative of $\hat{U}_t(y_t, d_t)$ with respect to d_t , $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t)$.*

The proof of Theorem 1 is provided in Appendix C.

Lemma 7 states that given the knowledge $\hat{V}_{t+1}^l(x_t)$ uniformly approximates $V'_{t+1}(x_t)$ on the interval $[x_t^{\min}, x_t^{\max}]$, and given sufficiently large N_t s, we have for all $t = 1, \dots, T$, $\hat{W}_t^l(y_t)$ uniformly approximates $W'_t(y_t)$ on the interval $[y_t^{\min}, y_t^{\max}]$ with high probability.

Lemma 7 *Suppose we are given that the empirical left derivative $\hat{V}_{t+1}^l(x_t)$ uniformly approximates $V'_{t+1}(x_t)$. That is, $\left| \hat{V}_{t+1}^l(x_t) - V'_{t+1}(x_t) \right| \leq \gamma$, for some $\gamma > 0$. Consider the empirical function $\hat{W}_t(y_t)$ constructed using a data set, in which $N_t \geq \frac{16B_t \log(18/\beta)}{9\alpha^2}$. Then the left derivative of $\hat{W}_t(y_t)$ uniformly approximates the derivative of $W_t(y_t)$ with high probability. That is, for any $\alpha > 0$,*

$$\mathbb{P} \left[\text{for all } y_t \in [y_t^{\min}, y_t^{\max}] : \left| \hat{W}_t^l(y_t) - W_t'(y_t) \right| \leq \alpha + \gamma \right] \geq 1 - \beta.$$

The same result holds also for the right derivative of $\hat{W}_t(y_t)$, $\hat{W}_t^r(y_t)$.

The proof of Lemma 7 is provided in Appendix C.

Theorem 2 states a key induction step in the first-order analysis: given the knowledge $\hat{V}_{t+1}^l(x_t)$ uniformly approximates $V'_{t+1}(x_t)$ on the interval $[x_t^{\min}, x_t^{\max}]$, and given sufficiently large N_t s, we have for all $t = 1, \dots, T$, $\hat{V}_t^l(x_t)$ uniformly approximates $V'_t(x_t)$ on the interval $[x_t^{\min}, x_t^{\max}]$ with high probability. The proof crucially uses the results of Theorem 1 and Lemma 7.

Theorem 2 *Suppose we are given that the empirical left derivative $\hat{V}_{t+1}^l(x_t)$ uniformly approximates $V'_{t+1}(x_t)$. That is, $\left| \hat{V}_{t+1}^l(x_t) - V'_{t+1}(x_t) \right| \leq \gamma$, for some $\gamma > 0$. Consider the empirical function $\hat{V}_t(y_t)$ constructed using a data set, in which $N_t \geq \frac{16B_t \log(18/\beta)}{9\alpha^2}$. Then the left derivative of $\hat{V}_t(y_t)$ uniformly approximates the derivative of $V_t(y_t)$ with high probability. That is, for any $\alpha > 0$,*

$$\mathbb{P} \left[\text{for all } x_t \in [x_t^{\min}, x_t^{\max}] : \left| \hat{V}_t^l(x_t) - V'_t(x_t) \right| \leq \alpha + \gamma \right] \geq 1 - \beta.$$

The same result holds also for the right derivative of $\hat{V}_t(x_t)$, $\hat{V}_t^r(x_t)$.

The proof of Theorem 2 is provided in Appendix C.

Theorem 3 summarizes the main result of the first order analysis. That is, given sufficiently large N_t s, we have for all $t = 1, \dots, T$, the function $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t)$ uniformly

approximates the function $\frac{\partial}{\partial d_t} U_t(y_t, d_t)$ on the interval $[d_t^{\min}, d_t^{\max}]$ and the interval $[y_t^{\min}, y_t^{\max}]$ with high probability; the function $\hat{W}_t^l(y_t)$ uniformly approximates the function $W_t^l(y_t)$ on the interval $[y_t^{\min}, y_t^{\max}]$ with high probability.

Theorem 3 *Consider the empirical function $\hat{U}_t(y_t, d_t)$ and $\hat{W}_t(y_t)$ constructed using a data set, in which $N_t \geq \frac{16B_t \log(18T/\beta)}{9\alpha_t^2}$. Then the left partial derivative of $\hat{U}_t(y_t, d_t)$ with respect to d_t uniformly approximates the partial derivative of $U_t(y_t, d_t)$ with respect to d_t with high probability. That is, for any $\alpha_t, \dots, \alpha_T > 0$,*

$\mathbb{P} \left[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}] : \left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \leq \sum_{s=t}^T \alpha_s \right] \geq 1 - \beta$. The same result holds also for the right partial derivative of $\hat{U}_t(y_t, d_t)$ with respect to d_t , $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t)$. Moreover, the left derivative of $\hat{W}_t(y_t)$ uniformly approximates the derivative of $W_t(y_t)$ with high probability. That is, for any $\alpha_t, \dots, \alpha_T > 0$, $\mathbb{P} \left[\text{for all } y_t \in [y_t^{\min}, y_t^{\max}] : \left| \hat{W}_t^l(y_t) - W_t^l(y_t) \right| \leq \sum_{s=t}^T \alpha_s \right] \geq 1 - \beta$. The same results also hold for the right partial derivative of $\hat{U}_t(y_t, d_t)$ with respect to d_t , $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t)$ and the right derivative of $\hat{W}_t(y_t)$, $\hat{W}_t^r(y_t)$.

The proof of Theorem 3 is provided in Appendix C.

4.1.2 The Zero-Order Analysis

For convenience, define the constants

$$\lambda_t^U = \min_{y_t \in [y_t^{\min}, y_t^{\max}]} \min \left\{ \frac{U_t(d_t^*(y_t), y_t) - U_t(d_t^{\min}, y_t)}{d_t^*(y_t) - d_t^{\min}}, \frac{U_t(d_t^{\max}, y_t) - U_t(d_t^*(y_t), y_t)}{d_t^{\max} - d_t^*(y_t)} \right\},$$

$\lambda_t^W = \min \left\{ \frac{W_t(S_t^*) - W_t(y_t^{\min})}{S_t^* - y_t^{\min}}, \frac{W_t(y_t^{\max}) - W_t(S_t^*)}{y_t^{\max} - S_t^*} \right\}$. Then, Lemma 8 transforms the first order approximation bounds derived by Theorem 3 to zero order approximation bounds. That is to say, given sufficiently large N_t s, we have for all $t = 1, \dots, T$, $W_t(\hat{S}_t^*) \geq (1 - \epsilon_t^W) W_t(S_t^*)$, and $U_t(\hat{d}_t^*(y_t), y_t) \geq (1 - \epsilon_t^U) U_t(d_t^*(y_t), y_t)$ for all $y_t \in [y_t^{\min}, y_t^{\max}]$ with high probability, for some positive constant $\epsilon_t^W, \epsilon_t^U$.

Lemma 8 Consider the empirical base stock \hat{S}_t^* and the empirical expected demand function $\hat{d}_t^*(y_t)$ constructed using a data set, in which $N_t \geq \frac{16B_t \log(18T/\beta)}{9\alpha^2}$, $\forall t = 1, \dots, T$. Then the empirical base stock level \hat{S}_t^* is $(1 - \epsilon_t^W)$ -optimal with high probability. That is, $\mathbb{P} \left[W_t(\hat{S}_t^*) \geq (1 - \epsilon_t^W) W_t(S_t^*) \right] \geq 1 - \beta$, where $\epsilon_t^W = \frac{2}{\lambda_t^W} \sum_{s=t}^T \alpha_s$. Moreover, the empirical expected demand function $\hat{d}_t^*(y_t)$ is $(1 - \epsilon_t^U)$ -optimal with high probability. That is,

$$\mathbb{P} \left[\forall y_t \in [y_t^{\min}, y_t^{\max}] : U_t(\hat{d}_t^*(y_t), y_t) \geq (1 - \epsilon_t^U) U_t(d_t^*(y_t), y_t) \right] \geq 1 - \beta, \text{ where } \epsilon_t^U = \frac{2}{\lambda_t^U} \sum_{s=t}^T \alpha_s.$$

To state the main result of the zero-order analysis, we let

$\mu_t = ((S_t, d_t(\cdot)), (S_{t+1}, d_{t+1}(\cdot)), \dots, (S_T, d_T(\cdot)))$ denote a policy for time periods from t to T . In each time period $i \in [t, T]$, the base stock S_i and demand function $d_i^*(\cdot)$ are used. Let μ_t^* denote to optimal policy: $\mu_t^* = ((S_t^*, d_t^*(\cdot)), (S_{t+1}^*, d_{t+1}^*(\cdot)), \dots, (S_T^*, d_T^*(\cdot)))$. Let $\hat{\mu}_t$ denote to the empirical optimal policy using a given data set, i.e. $\hat{\mu}_t = ((\hat{S}_t^*, \hat{d}_t^*(\cdot)), (\hat{S}_{t+1}^*, \hat{d}_{t+1}^*(\cdot)), \dots, (\hat{S}_T^*, \hat{d}_T^*(\cdot)))$. Define function $\pi_t(x_t; \mu_t)$ as the total expected profit from time period t to T , if the policy μ_t is used and there are x_t units of starting inventory in time period t .

Then, Theorem 4 states a key induction step that given the policy $\hat{\mu}_{t+1}$ is $\left(1 - \sum_{s=t+1}^T \epsilon_s\right)$ -optimal, and given sufficiently large N_s , we have for all $t = 1, \dots, T$, the policy $\hat{\mu}_t$ is $(1 - \epsilon_t)$ -optimal with high probability, for some $\epsilon_t, \dots, \epsilon_T > 0$.

Theorem 4 Suppose we are given that the empirical optimal policy $\hat{\mu}_{t+1}$ is $\left(1 - \sum_{s=t+1}^T \epsilon_s\right)$ -optimal. That is, $\pi_{t+1}(x_{t+1}; \hat{\mu}_{t+1}) \geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi_{t+1}(x_{t+1}; \mu_{t+1}^*)$, for some $\epsilon_{t+1}, \dots, \epsilon_T > 0$. Consider the empirical optimal solution $\hat{\mu}_t$ constructed using a data set, in which $N_s \geq \frac{16B_s \log(18T/\beta)}{9\alpha^2}$, $\forall s = t, \dots, T$. Then the the empirical

optimal solution $\hat{\mu}_t$ is $(1 - \epsilon)$ -optimal with high probability. That is,

$$\mathbb{P} \left[\text{for all } x_t \in [x_t^{\min}, x_t^{\max}] : \pi_t(x_t; \hat{\mu}_t) \geq (1 - \epsilon)\pi_t(x_t; \mu_t^*) \right] \geq 1 - \beta, \text{ where } \epsilon = \epsilon_{t+1} + \epsilon_t^U + \epsilon_t^W.$$

The proof of Theorem 4 is provided in Appendix C.

Finally, Theorem 5 states the main result of the section, that is a sufficient sample complexity bound on N_t in order to achieve $(1 - \epsilon)$ -optimality for some $\epsilon > 0$ with high probability. The proof is mainly by completing the induction argument of Theorem 4.

Theorem 5 *Consider the data-driven problem with T time periods. To achieve $(1 - \epsilon)$ -optimality with probability at least $1 - \beta$, the sufficient size of the data set needed is that*

$$N_t \geq \frac{256B_t \min_{t=1, \dots, T} \min\{\lambda_t^U, \lambda_t^W\} T^4 \log(18T/\beta)}{9\epsilon^2}, \forall t = 1, \dots, T.$$

The proof of Theorem 5 is provided in Appendix C.

4.2 On the Computational Complexity

We formally state our data-driven algorithm as follows. Notice a sparsification step is included in for each t . The sparsification parameter ζ is set to be $\zeta = \alpha_1 = \frac{\min_{t=1, \dots, T} \min\{\lambda_t^U, \lambda_t^W\} \epsilon}{4T^2}$.

Obviously, without the sparsification step, the computational complexity of the algorithm will be exponential in T , since it can be verified that the number of break-points of the function $V_t^l(x_t)$ is in the order of $\prod_{s=t}^T N_s$, which is exponential in T even when all N_s s are polynomial in T . After using the sparsification step, the computational complexity of the algorithm becomes polynomial in T , as stated by Fact 10.

Input: For all $t = 1, \dots, T$, independent samples d_t^{ij} are drawn,

$$\forall i = 0, \dots, K_t, j = 1, \dots, N_t.$$

Define $\hat{V}_{T+1}(x_t) = 0$ for all x_t .

for $t=T, \dots, 1$ **do**

Construct $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) = \hat{R}'_t(d_t) + \hat{C}_t^l(y_t - d_t) - \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t)]$.

Construct $\hat{f}_t^l(y_t - d_t) = -\hat{C}_t^l(y_t - d_t) + \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t)]$, by merging the breakpoints of $\hat{C}_t^l(y_t - d_t)$ and $\mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t)]$ and calculating the corresponding function values via binary searches. Note $\hat{f}_t^l(y_t - d_t)$ is a piecewise-linear function with finitely many breakpoints.

Construct $\hat{d}_t^*(y_t)$ as the intersection point of $\hat{R}'_t(d_t)$ and $\hat{f}_t^l(y_t - d_t)$, for all $y_t \in [y_t^{\min}, y_t^{\max}]$.

Construct $\hat{W}_t^l(y_t) = \hat{R}'_t(\hat{d}_t^*(y_t))$.

Compute empirical base stock and list demand: using binary search to find \hat{S}_t^* so \hat{S}_t^* is the largest number such that $\hat{W}_t^l(\hat{S}_t^*) \geq 0$; set $\hat{D}_t^* = \hat{d}_t^*(\hat{S}_t^*)$.

Construct $\hat{V}_t^l(x_t) = \begin{cases} 0 & , x_t \leq \hat{S}_t^* \\ \hat{W}_t^l(x_t) & , x_t > \hat{S}_t^* \end{cases}$.

For all $x_t \in [x_t^{\min}, x_t^{\max}]$, update $\hat{V}_t(x_t) \leftarrow \zeta \lfloor \hat{V}_t(x_t) / \zeta \rfloor$.

end

Return $\hat{S}_t^*, \hat{d}_t^*(y_t)$ for all $t = 1, \dots, T, y_t \in [y_t^{\min}, y_t^{\max}]$.

Algorithm 1: The data-driven algorithm for linear demand models without capacity constraints.

Fact 10 *The running time of Algorithm 1 is polynomial in*

$$O(\max\{N_1, \dots, N_T, T, \sum_{s=t+1}^T (h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max}) / \zeta\}).$$

We briefly mention the proof of Fact 10. Fix some $t = 1, \dots, T - 1$, the sparsification procedure ensures the number of breakpoints of $\hat{V}_{t+1}^l(x_t)$ is $O(\sum_{s=t+1}^T (h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max}) / \zeta)$, thus the number of breakpoints of $\hat{f}_t^l(y_t - d_t)$ is $O(N_t(\sum_{s=t+1}^T (h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max}) / \zeta) + N_t) = O(N_t(\sum_{s=t+1}^T (h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max}) / \zeta))$, the sum of the number of breakpoints of $\hat{C}_t^l(y_t, d_t)$ and the number of breakpoints of $\frac{1}{N_t} \sum_{j=1}^{N_t} \hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t^{0j})$. Sorting the breakpoints of $\hat{f}_t^l(y_t -$

d_t) requires $O(N_t(\sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta) \log(N_t(\sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta)))$ operations, and computing the corresponding function values in $O(N_t N_t \log((\sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta) + N_t(\sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta) \log(N_t))) = O(N_t^2 \log(N_t \sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta))$.

Thus, the construction of $\hat{d}_t(y_t)$ and $\hat{W}_t^l(\hat{d}_t^*(y_t))$ requires $O(N_t^2 \log(N_t \sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta))$ operations. The computation of \hat{S}_t^* requires $O(\log(N_t(\sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta)))$ operations. The construction of $\hat{V}_t^l(x_t)$ then requires $O(N_t^2 \log(N_t \sum_{s=t+1}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta))$ operations. After the specification step, the number of breakpoints of $\hat{V}_t^l(x_t)$ is at most $O(\sum_{s=t}^T(h_t + b_t + 1/|\kappa_t|^{\min} d_t^{\max} + |\theta_t|^{\max})/\zeta)$. Complete this induction on t from $T - 1$ to 1 we obtain the fact.

It is not hard to observe, regarding the analysis in Section 4.1, the sparsification step deteriorates the first order approximation bounds in Theorem 2 by at most ζ , i.e., given $|\hat{V}_{t+1}^l(x_t) - V_{t+1}'(x_t)| \leq \gamma$, we know $|\hat{V}_t^l(x_t) - V_t'(x_t)| \leq \alpha + \zeta + \gamma$ with probability at least $1 - \beta$, given $N_t \geq \frac{16B_t \log(18/\beta)}{9\alpha^2}$. Thus, it is immediate to obtain Theorem 3 holds with all α_s s replaced by $\alpha_s + \zeta$. Given our choice of $\zeta = \alpha_1$, we know the main result (c.f. Theorem 5) still holds with the same condition on N_t but a slightly deteriorated factor on ϵ (ϵ replaced by 2ϵ), as stated by the following corollary. The corollary, along with Fact 10, suggests that we actually have an FPTAS (fully polynomial-time approximation scheme) for the full information problem.

Corollary 1 *Given the same choice of N_t as in Theorem 5, Algorithm 1 returns a policy that is $(1 - 2\epsilon)$ -optimal with probability at least $1 - \beta$.*

Remark 3 *Lemma 4.1, 4.2 in [5] are similar to Fact 10 and Corollary 1.*

Chapter 5

Extension to Joint Pricing and Capacitated Inventory Control Models under General Concave Demand Models

In this chapter, we extend our previous results to the joint pricing and capacitated inventory control models under general concave demand models, that is to say, the model is modified as follows:

- (1) A capacity constraint on the ordering inventory level is imposed, in each time period t , the ordering quantity $y_t - x_t$ must be less than or equal to I_t , the maximum quantity of the item we can order. That is, $0 \leq y_t - x_t \leq I_t$, for all $t = 1, \dots, T$.
- (2) Regarding the stochastic demand function $G_t(p_t, \eta_t) = D_t(p_t) + \eta_t$, it is not assumed $D_t(p_t)$ is a linear function in p_t as in Assumption 1, but only assumed

to be a strictly decreasing function. Moreover, the expected revenue function $R_t(d_t) = d_t D_t^{-1}(d_t)$ is assumed to be strictly concave in d_t , and smooth enough, its second derivative is uniformly bounded, i.e., $\max_{d_t \in d_t^{\min}, d_t^{\max}} |R_t''(d_t)|$ is finite. Note $R_t(d_t)$ is not necessarily a quadratic function.

Then, suppose we construct the empirical revenue function $\hat{R}_t(d_t)$ in some way, using $N_t \geq N_{R_t}(\alpha, \beta, K_t)$, $K_t \geq K_{R_t}(\alpha, \beta)$, where $N_{R_t}(\alpha, \beta, K_t)$, $K_{R_t}(\alpha, \beta)$ are positive finite constants defined by α, β, K_t , such that the left derivative of $\hat{R}_t(d_t)$ uniformly approximates the derivative of $R_t(d_t)$ with high probability. That is, $\mathbb{P}[|\hat{R}_t'(d_t) - R_t'(d_t)| \leq \alpha] \geq 1 - \beta$, for all $t = 1, \dots, T$. Moreover, we assume $\hat{R}_t'(d_t)$ is also strictly decreasing and continuous in d_t .

Recall the optimal policy of solving the full-information dynamic program is a modified base-stock list-price policy, which is stated as follows. In period t , given the starting inventory level x_t , and the base-stock levels S_t^*, \dots, S_T^* , the list-price levels L_t^*, \dots, L_T^* , and the optimal demand functions $d_t^*(y_t), \dots, d_T^*(y_T)$, the order-up-to level y_t is determined as

$$y_t = \begin{cases} x_t + I_t, & \text{if } x_t \in [x_t^{\min}, S_t^* - I_t), \\ S_t^*, & \text{if } x_t \in [S_t^* - I_t, S_t^*), \\ x_t, & \text{if } x_t \in [S_t^*, x_t^{\max}]. \end{cases}$$

The expected demand level is determined as

$$d_t = \begin{cases} d_t^*(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, S_t^* - I_t), \\ L_t^*, & \text{if } x_t \in [S_t^* - I_t, S_t^*), \\ d_t^*(x_t), & \text{if } x_t \in [S_t^*, x_t^{\max}]. \end{cases}$$

Thus, given $W_t(y_t) = U_t(y_t, d_t^*(y_t))$, the value function in each period t is defined

as

$$V_t(x_t) = \begin{cases} W_t(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, S_t^* - I_t), \\ W_t(I_t^*), & \text{if } x_t \in [S_t^* - I_t, S_t^*), \\ W_t(x_t), & \text{if } x_t \in [S_t^*, x_t^{\max}]. \end{cases}$$

Like before, our data-driven policy mimics the full-information optimal policy, i.e., in period t , given the starting inventory level x_t , and the data-driven base-stock levels $\hat{S}_t^*, \dots, \hat{S}_T^*$, the data-driven list-price levels $\hat{L}_t^*, \dots, \hat{L}_T^*$, and the data-driven demand functions $\hat{d}_t^*(y_t), \dots, \hat{d}_T^*(y_T)$, the order-up-to level y_t is determined as

$$y_t = \begin{cases} x_t + I_t, & \text{if } x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ \hat{S}_t^*, & \text{if } x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ x_t, & \text{if } x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$$

The expected demand level is determined as

$$d_t = \begin{cases} \hat{d}_t^*(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ \hat{L}_t^*, & \text{if } x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ \hat{d}_t^*(x_t), & \text{if } x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$$

Thus, given $\hat{W}_t(y_t) = \hat{U}_t(y_t, \hat{d}_t^*(y_t))$, the value function in each period t is defined

as

$$\hat{V}_t(x_t) = \begin{cases} \hat{W}_t(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ \hat{W}_t(I_t^*), & \text{if } x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ \hat{W}_t(x_t), & \text{if } x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$$

Algorithm 2 describes concretely the data-driven problem for solving the joint pricing and capacitated inventory control problem. Then it is immediate to see Fact 10 still holds (the running time is also polynomial in K_1, \dots, K_T), as long as the

running time of constructing $\hat{R}_t(d_t)$ is polynomial in N_t, K_t .

Input: For all $t = 1, \dots, T$, independent samples d_t^{ij} are drawn,
 $\forall i = 0, \dots, K_t, j = 1, \dots, N_t$.

Define $\hat{V}_{T+1}(x_t) = 0$ for all x_t .

for $t=T, \dots, 1$ **do**

Construct $\frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) = \hat{R}_t^l(d_t) + \hat{C}_t^l(y_t - d_t) - \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t)]$.

Construct $\hat{f}_t^l(y_t - d_t) = -\hat{C}_t^l(y_t - d_t) + \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t)]$, by merging the breakpoints of $\hat{C}_t^l(y_t - d_t)$ and $\mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - d_t - \hat{\eta}_t)]$ and calculating the corresponding function values via binary searches. Note $\hat{f}_t^l(y_t - d_t)$ is a piecewise-linear function with finitely many breakpoints.

Construct $\hat{d}_t^*(y_t)$ as the intersection point of $\hat{R}_t^l(d_t)$ and $\hat{f}_t^l(y_t - d_t)$, for all $y_t \in [y_t^{\min}, y_t^{\max}]$.

Construct $\hat{W}_t^l(y_t) = \hat{R}_t^l(\hat{d}_t^*(y_t))$.

Compute empirical base stock and list demand: using binary search to find \hat{S}_t^* so \hat{S}_t^* is the largest number such that $\hat{W}_t^l(\hat{S}_t^*) \geq 0$; set $\hat{D}_t^* = \hat{d}_t^*(\hat{S}_t^*)$.

Construct $\hat{V}_t^l(x_t) = \begin{cases} \hat{W}_t^l(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ 0, & \text{if } x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ \hat{W}_t^l(x_t), & \text{if } x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$

For all $x_t \in [x_t^{\min}, x_t^{\max}]$, update $\hat{V}_t(x_t) \leftarrow \zeta \lfloor \hat{V}_t^l(x_t) / \zeta \rfloor$.

end

Return $\hat{S}_t^*, \hat{d}_t^*(y_t)$ for all $t = 1, \dots, T, y_t \in [y_t^{\min}, y_t^{\max}]$.

Algorithm 2: The data-driven algorithm for general concave demand models with capacity constraints.

Theorem 6 *Suppose we construct the empirical revenue function $\hat{R}_t(d_t)$ using a data set, in which $N_t \geq N_{R_t}(\alpha, \beta, K_t)$, $K_t \geq K_{R_t}(\alpha, \beta)$, such that the left derivative of $\hat{R}_t(d_t)$ uniformly approximates the derivative of $R_t(d_t)$ with high probability. That is, $\mathbb{P}[|\hat{R}_t^l(d_t) - R_t^l(d_t)| \leq \alpha] \geq 1 - \beta$, for all $t = 1, \dots, T$. Moreover, we assume $\hat{R}_t^l(d_t)$ is also strictly decreasing and continuous in d_t , which means $\hat{R}_t(d_t)$ is a strictly concave*

function. Then, given for all $t = 1, \dots, T$,

$$N_t \geq \max \{N_{R_t}(\alpha_t/3, 5\beta/(9T), K_t),$$

$$2(2 \log 2 - \log(9T/(2\beta))) (\omega_t^{\max} - \omega_t^{\min})^2 \max \{l_t^2 (h_t + b_t)^2, \max_{d_t \in [d_t^{\min}, d_t^{\max}]} |R_t''(d_t)|\} / \alpha_t^2 \},$$

$$K_t \geq K_{R_t}(\alpha_t/3, 5\beta/(9T)),$$

where $\alpha_t = \min_{t=1, \dots, T} \min \{\lambda_t^U, \lambda_t^W\} \epsilon / (4T^2)$, Algorithm 2 returns a $(1 - 2\epsilon)$ -optimal policy with probability at least $1 - \beta$.

The proof of Theorem 6 is provided in Appendix C. From Theorem 6, we know given $N_{R_t}(\alpha_t/3, 5\beta/(9T), K_{R_t}(\alpha_t/3, 5\beta/(9T)))$ are polynomial in $T, 1/\epsilon$, as α_t is set as $\min_{t=1, \dots, T} \min \{\lambda_t^U, \lambda_t^W\} \epsilon / (4T^2)$, we still have a FPTAS for the full information problem.

Chapter 6

A Simulation Study

In this chapter, we conduct a simulation study, in which a cell phone retailer faces the joint inventory and pricing control problem. After gathering the past data samples, the retailer specifies $D_t(p_t)$ to be a linear function in p_t . The retailer introduces a new version of the phone in October each year. Moreover, the retailer observes the demand patterns can be mainly split into five seasons: from October to December ($t = 1, 2, 3$) is the holiday season, when the demand for the new phone is very high; from next January to next March ($t = 4, 5, 6$) is the post-holiday season, when the demand for the new phone is very low; from next April to next May ($t = 7, 8$) is the normal season, when the demand is also in a normal level; from June to August ($t = 9, 10, 11$) is the summer season, when the demand is pretty high; Finally, next September ($t = 12$) is the last month of the selling horizon, since next October another new version of the phone will be released, so a very large holding cost is imposed in this period.

The retailer wants to make a whole-year monthly plan of jointly controlling the price and the inventory level, based on the gathered seasonal demand samples. Sup-

Table 6.1: Main problem parameters

| Period t | Season | c_t | h_t | b_t | κ_t | θ_t | p_t^0 | p_t^1 |
|---------------|--------------|-------|-------|-------|------------|------------|---------|---------|
| Oct, Nov, Dec | holiday | 100 | 2 | 225 | -18 | 9000 | 200 | 250 |
| Jan, Feb, Mar | post-holiday | 100 | 2 | 175 | -10 | 3000 | 150 | 200 |
| Apr, May | normal | 100 | 2 | 200 | -7.5 | 3000 | 175 | 225 |
| Jun, Jul, Aug | summer | 100 | 2 | 175 | -17 | 5000 | 150 | 200 |
| Sep | normal | 100 | 50 | 100 | -7.5 | 3000 | 175 | 225 |

pose in each past season, two distinct prices were offered, p_t^0 , p_t^1 , and the underlying random noise η_t follows uniform distributions on the supports $[-1500, 1500]$, $[-400, 400]$, $[-500, 500]$, $[-1000, 1000]$, $[-500, 500]$ in the five seasons respectively. The retailer has collected $2N_t$ number of demand samples, related to the two prices respectively. Moreover, we assume the inventory levels are bounded within the intervals $[3000, 7500]$, $[1000, 4650]$, $[800, 3600]$, $[1500, 3700]$, $[800, 3600]$ in the five seasons respectively (see the detailed parameter descriptions in Table A.1). The other (main) problem parameters are listed in Table 6.1.

We conduct the numerical experiment in which the number of price-demand samples N_t is 1500, 400, 500, 1000, 500 in the each five seasons. We examine the expected profit generated from the data-driven policy, against the expected profit generated from the optimal policy, on the entire interval $x_1 \in [6000, 7500]$. Then, our simulation result shows

$$\max_{x_1 \in [6000, 7500]} \frac{\pi_1(x_1; \mu_1^*) - \pi_1(x_1; \hat{\mu}_1^*)}{\pi_1(x_1; \mu_1^*)} \leq \frac{3001024 - 3000080}{3001024} = 0.31\%,$$

which shows the data-driven policy is very near-optimal. The plot of the expected profits generated by the two different policies is given by Figure 6-1. And, all plots of the optimal demand function $d_t^*(y_t)$, $\hat{d}_t^*(y_t)$, and the profit functions $\pi_t(x_t; \mu_t^*)$,

$\pi_t(x_t; \hat{\mu}_t^*)$ in all $t = 1, \dots, 12$ periods are provided in Appendix B.

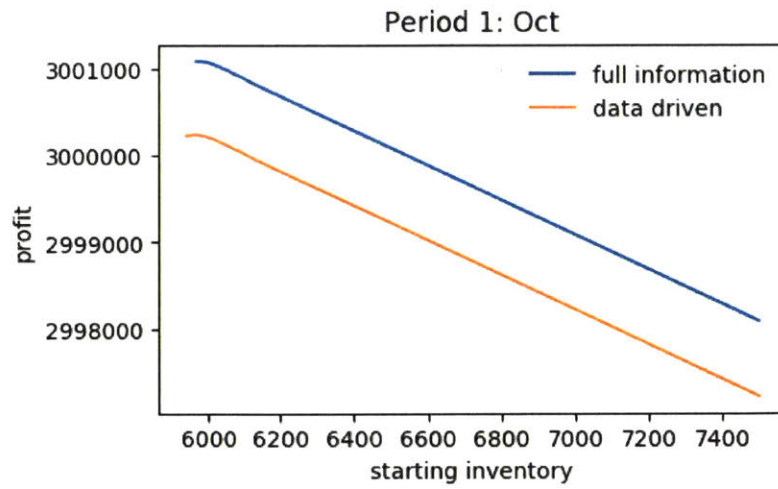


Figure 6-1: The expected profit generated by the data-driven policy v.s. the full-information optimal policy, in the first period.

Chapter 7

Conclusions and Future Directions

The thesis presents a sampling-based scheme for computing the solution of the data-driven joint pricing and inventory control problem. It is proved that with sample complexity and computational complexity both polynomial in T and $1/\epsilon$, our algorithm is able to achieve $(1 - 2\epsilon)$ -optimality with high probability. Moreover, our numerical simulation suggests the potential good practical performance of the algorithm.

We also point out some future directions of our work. First, our theoretical analysis and algorithm implicitly assumes perfect control on the expected demand level d_t , which might not be very practical since often in practice we only have control ability of the price p_t instead of d_t . However, such an extension is not trivial in our model since our analysis relies on the joint concavity induced by using d_t instead of p_t as our optimization variable, especially when $R_t(d_t)$ might be a general concave function.

Second, we only consider the additive stochastic demand models, but in the joint pricing and inventory control literature often the multiplicative stochastic demand

models are also of interest: $G_t(p_t, \eta_t) = \eta_t D_t(p_t)$, where η_t is a random variable with mean 1. Thus it would also be interesting to extend our work to such scenario.

Third, our result in Chapter 5, though works for general concave demand functions, does not specify how to deal with the fully nonparametric scenario. In such a scenario, no parametric assumption of $D_t(p_t)$ or $R_t(p_t)$ is known in priori, so we need to construct $\hat{R}_t(d_t)$ in a fully nonparametric way, perhaps just using a piecewise-linear fit. However, we still need to ensure simultaneously the strict concavity of $\hat{R}_t(d_t)$ and the uniform bound on $|\hat{R}'_t(d_t) - R'_t(d_t)|$ for our analysis to work, which is not so clear to us so far.

Appendix A

Tables

Table A.1: Other problem parameters in Chapter 6

| Period t | Season | ω_t^{\min} | ω_t^{\max} | y_t^{\min} | y_t^{\max} | p_t^{\min} | p_t^{\max} |
|---------------|--------------|-------------------|-------------------|--------------|--------------|--------------|--------------|
| Oct, Nov, Dec | holiday | -1500 | 1500 | 3000 | 7500 | 150 | 250 |
| Jan, Feb, Mar | post-holiday | -400 | 400 | 1000 | 4650 | 150 | 250 |
| Apr, May | normal | -500 | 500 | 800 | 3600 | 150 | 250 |
| Jun, Jul, Aug | summer | -1000 | 1000 | 1500 | 3700 | 150 | 250 |
| Sep | normal | -500 | 500 | 800 | 3600 | 150 | 250 |

Appendix B

Figures

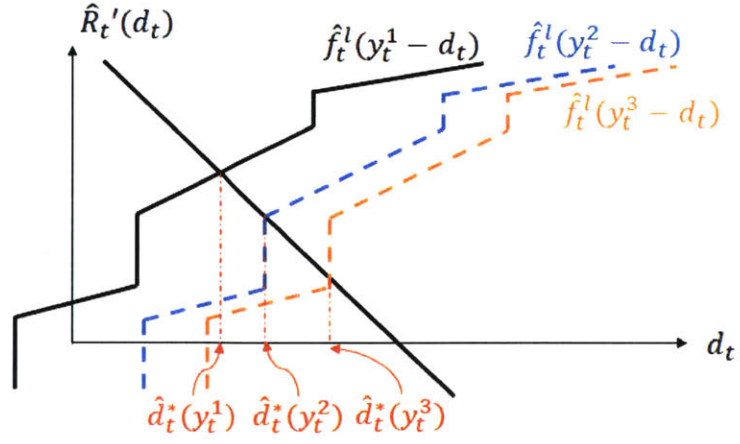


Figure B-1: Illustration of Fact 9.

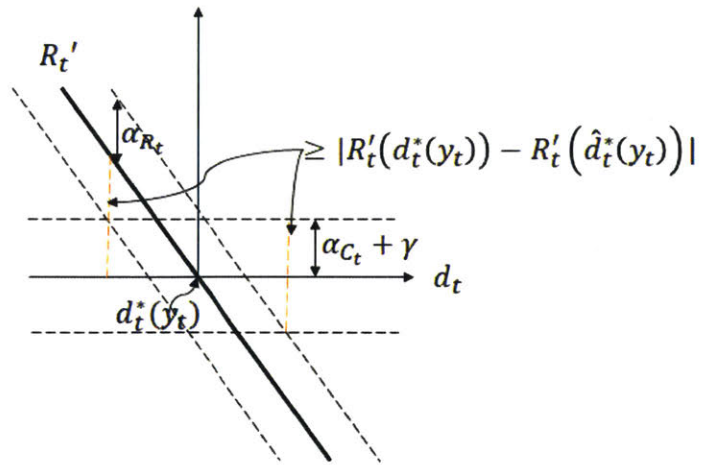


Figure B-2: Illustration of Proof of Lemma 7.

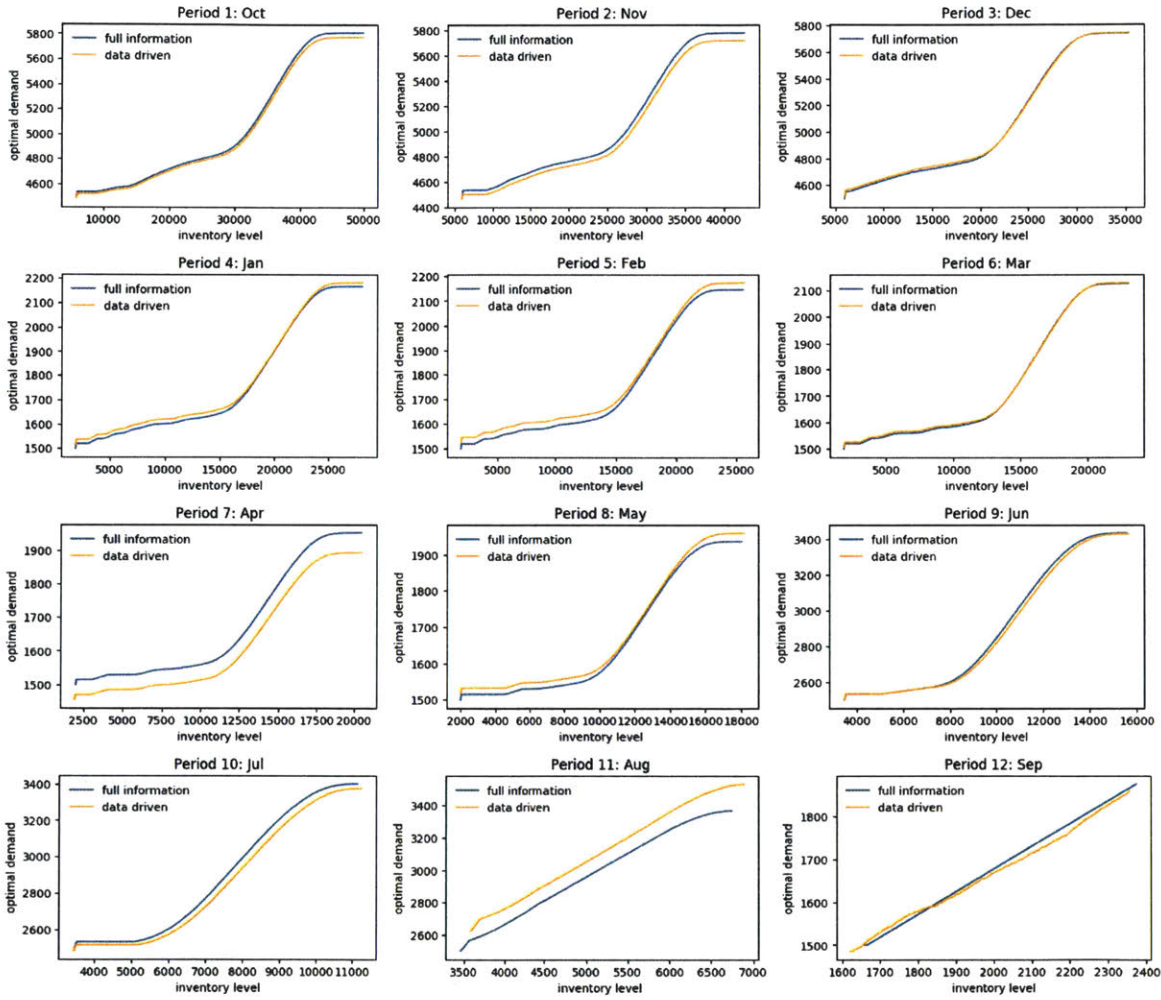


Figure B-3: Data-driven $\hat{d}_t^*(y_t)$ v.s. full-information $d_t^*(y_t)$ from the problem in Chapter 6.

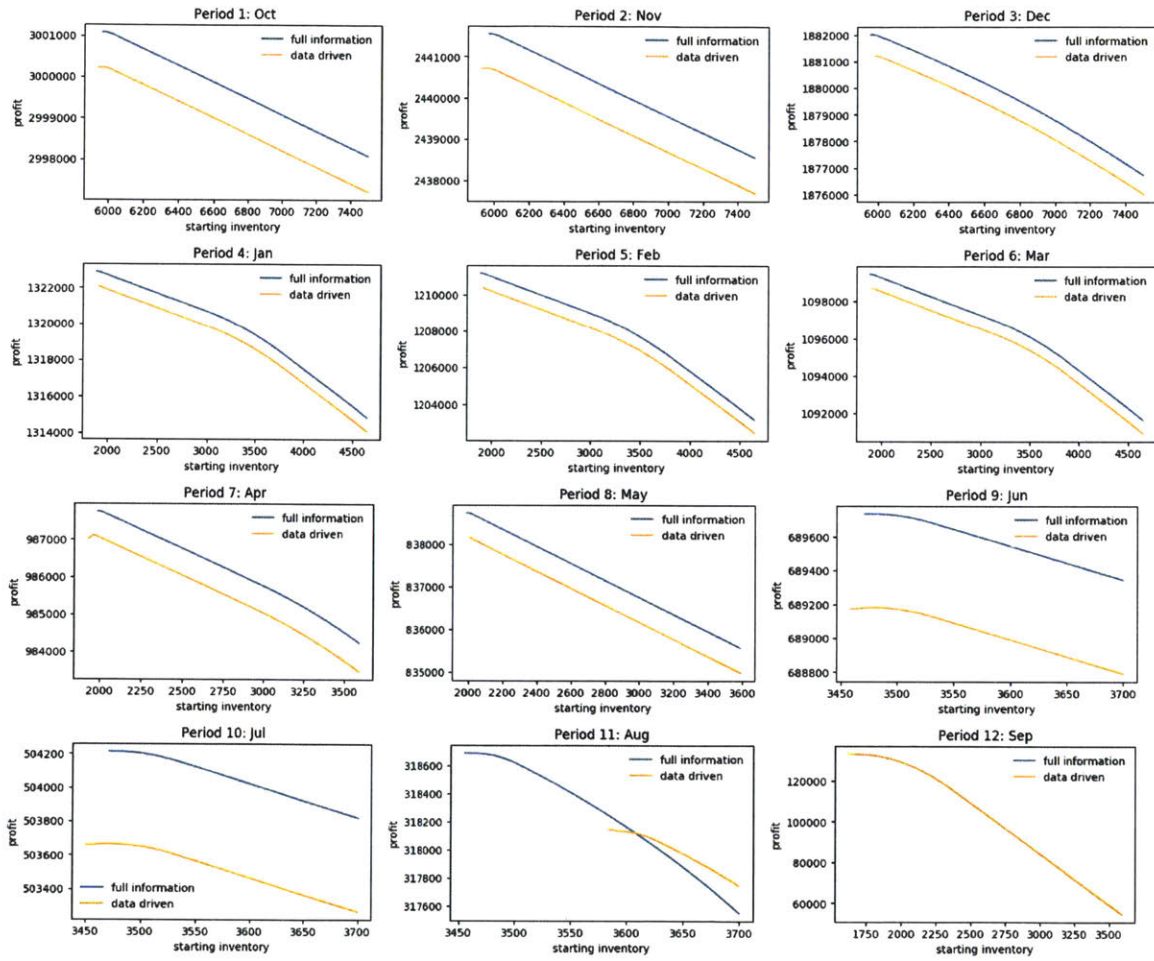


Figure B-4: Expected profit generated by the data-driven policy v.s. the full-information optimal policy from the problem in Chapter 6.

Appendix C

Proofs of Lemmas and Theorems

Proof of Lemma 1. For any $y_t \in [y_t^{\min}, y_t^{\max}]$, $d_t^*(y_t)$ is the unique value that satisfies $\frac{\partial}{\partial d_t} U_t(y_t, d_t^*(y_t)) = 0$. Hence, the following two sets are equivalent:

$$\{(y_t, d_t^*(y_t)) : y_t \in [y_t^{\min}, y_t^{\max}]\} = \left\{ (y_t, d_t) : y_t \in [y_t^{\min}, y_t^{\max}], \frac{\partial}{\partial d_t} U_t(y_t, d_t) = 0 \right\}.$$

Since $U_t(y_t, d_t) = g_t(y_t, d_t) + \mathbb{E}_{\eta_t}[V_{t+1}(y_t - d_t - \eta_t)]$, we define

$$f_t(y_t, d_t) = \frac{\partial}{\partial d_t} U_t(y_t, d_t) = R_t'(d_t) - \frac{\partial}{\partial d_t} C_t(y_t, d_t) - \mathbb{E}_{\eta_t}[V_{t+1}'(y_t - d_t - \eta_t)].$$

Then, we have

$$\begin{aligned} \frac{\partial}{\partial y_t} f_t(y_t, d_t) &= -\frac{\partial^2}{\partial y_t \partial d_t} C_t(y_t, d_t) - \mathbb{E}_{\eta_t}[V_{t+1}''(y_t - d_t - \eta_t)], \\ \frac{\partial}{\partial d_t} f_t(y_t, d_t) &= R_t''(d_t) - \frac{\partial^2}{\partial d_t^2} C_t(y_t, d_t) + \mathbb{E}_{\eta_t}[V_{t+1}''(y_t - d_t - \eta_t)]. \end{aligned}$$

Since $C_t(y_t, d_t) = \mathbb{E}_{\eta_t}[h_t(y_t - d_t - \eta_t)^+ + b_t(d_t + \eta_t - y_t)^+]$,

$\frac{\partial^2}{\partial d_t^2} C_t(y_t, d_t) = -\frac{\partial^2}{\partial y_t \partial d_t} C_t(y_t, d_t) \geq 0$. Since $R_t(d_t)$ is strictly concave, $R_t''(d_t) < 0$. By

Fact 3, $V_t(x_t)$ is concave, so $\mathbb{E}_{\eta_t}[V''_{t+1}(y_t - d_t - \eta_t)] \leq 0$.

Therefore, for all $y_t \in [y_t^{\min}, y_t^{\max}]$ and $d_t \in [d_t^{\min}, d_t^{\max}]$,

$$-\frac{\frac{\partial}{\partial y_t} f_t(y_t, d_t)}{\frac{\partial}{\partial d_t} f_t(y_t, d_t)} = \frac{\frac{\partial^2}{\partial d_t^2} C_t(y_t, d_t) - \mathbb{E}_{\eta_t}[V''_{t+1}(y_t - d_t - \eta_t)]}{-R'_t(d_t) + \left(\frac{\partial^2}{\partial d_t^2} C_t(y_t, d_t) - \mathbb{E}_{\eta_t}[V''_{t+1}(y_t - d_t - \eta_t)] \right)},$$

which is between 0 and 1.

Thus, using the implicit function theorem, we have $d_t^{*'}(y_t) = -\frac{\frac{\partial}{\partial y_t} f_t(y_t, d_t^*(y_t))}{\frac{\partial}{\partial d_t} f_t(y_t, d_t^*(y_t))}$. Therefore, we conclude that $0 \leq d_t^{*'}(y_t) \leq 1$ for all y_t in $[y_t^{\min}, y_t^{\max}]$.

Proof of Lemma 2. To prove the Lipschitz-continuity of $V'_t(x_t)$ is equivalent to show a finite upper bound for $|V''_t(x_t)|$. Recall that

$$V_t(x_t) = \begin{cases} U_t(S_t^*, D_t^*), & x_t \leq S_t^*, \\ U_t(x_t, d_t^*(x_t)), & x_t > S_t^*. \end{cases}$$

Therefore, $V'_t(x_t)$ is

$$V'_t(x_t) = \begin{cases} 0, & x_t \leq S_t^*, \\ \frac{d}{dx_t} U_t(x_t, d_t^*(x_t)), & x_t > S_t^*. \end{cases}$$

Recall that, if $x_t > S_t^*$, $U_t(x_t, d_t^*(x_t)) = R_t(d_t^*(x_t)) - C_t(x_t, d_t^*(x_t)) + \mathbb{E}_{\eta_t}[V_{t+1}(x_t - d_t^*(x_t) - \eta_t)]$, and also $d_t^*(x_t)$ satisfies $\frac{\partial}{\partial d_t} U_t(x_t, d_t^*(x_t)) = 0$. Hence, we have

$$\begin{aligned}
\frac{d}{dx_t} U_t(x_t, d_t^*(x_t)) &= R_t'(d_t^*(x_t)) d_t^{*'}(x_t) \\
&\quad - \left(\frac{\partial}{\partial x_t} C_t(x_t, d_t^*(x_t)) - \mathbb{E}_{\eta_t}[V_{t+1}'(x_t - d_t^*(x_t) - \eta_t)] \right) (1 - d_t^{*'}(x_t)) \\
&= R_t'(d_t^*(x_t)) \\
&\quad + \left(R_t'(d_t^*(x_t)) + \frac{\partial}{\partial d_t} C_t(x_t, d_t^*(x_t)) - \mathbb{E}_{\eta_t}[V_{t+1}'(x_t - d_t^*(x_t) - \eta_t)] \right) (d_t^{*'}(x_t) - 1) \\
&= R_t'(d_t^*(x_t)) + \left(\frac{\partial}{\partial d_t} U_t(x_t, d_t^*(x_t)) \right) (d_t^{*'}(x_t) - 1) \\
&= R_t'(d_t^*(x_t)).
\end{aligned}$$

Therefore, $V_t''(x_t) = \begin{cases} 0, & x_t \leq S_t^*, \\ R_t''(d_t^*(x_t)) d_t^{*'}(x_t), & x_t > S_t^*, \end{cases}$ and thus

$|V_t''(x_t)| \leq |R_t''(d_t^*(x_t))| |d_t^{*'}(x_t)|$. From Lemma 1, we have $0 \leq d_t^{*'}(x_t) \leq 1$, and therefore $|d_t^{*'}(x_t)| \leq 1$. Since for all $x_t \in [x_t^{\min}, x_t^{\max}]$, $|R_t''(d_t^*(x_t))| = |\kappa_t|$, we conclude that $|V_t''(x_t)| \leq |\kappa_t|$, and thus $V_t'(x)$ is Lipschitz-continuous with constant $|\kappa_t|$.

Proof of Lemma 3 From definition, we have

$$\hat{\kappa}_t = \frac{D_t(p_t^1) + \Delta_t^1 - (D_t(p_t^0) + \Delta_t^0)}{2(p_t^1 - p_t^0)} = \kappa_t + \frac{\Delta_t^1 - \Delta_t^0}{2(p_t^1 - p_t^0)}.$$

Therefore, from Fact 6 we derive for any $\alpha_1 > 0$,

$$\mathbb{P}[|\kappa_t - \hat{\kappa}_t| \leq \alpha_1] \geq \mathbb{P}[|\Delta_t^0|, |\Delta_t^1| \leq (p_t^1 - p_t^0)\alpha_1] \geq 1 - 4 \exp\left(-\frac{2(p_t^1 - p_t^0)^2 N_t \alpha_1^2}{(\omega^{\max} - \omega^{\min})^2}\right).$$

Then, from Assumption 1 we know

$$\left| \frac{1}{\kappa_t} - \frac{1}{\hat{\kappa}_t} \right| \leq \frac{|\kappa_t - \hat{\kappa}_t|}{(|\kappa_t|^{\min})^2}.$$

On the other hand,

$$\hat{\theta}_t = D_t(p_t^0) + \Delta_t^0 - 2\hat{\kappa}_t p_t^0 = \theta_t + \Delta_t^0 + 2p_t^0(\kappa_t - \hat{\kappa}_t).$$

So we have for any $\alpha_2 > 0$,

$$\begin{aligned} \mathbb{P} \left[|\theta_t - \hat{\theta}_t| \leq \alpha_2 \right] &\geq \mathbb{P} \left[|\kappa_t - \hat{\kappa}_t| \leq \frac{\alpha_2}{4p_t^0}, |\Delta_t^0| \leq \frac{\alpha_2}{2} \right] \\ &\geq 1 - 4 \exp \left(-\frac{(p_t^1 - p_t^0)^2 N_t \alpha_2^2}{8(\omega_t^{\max} - \omega_t^{\min})^2 (p_t^0)^2} \right) - 2 \exp \left(-\frac{N_t \alpha_2^2}{2(\omega_t^{\max} - \omega_t^{\min})^2} \right). \end{aligned}$$

Again from Assumption 1,

$$\left| \frac{\theta_t}{\kappa_t} - \frac{\hat{\theta}_t}{\hat{\kappa}_t} \right| = \left| \frac{\hat{\kappa}_t(\theta_t - \hat{\theta}_t) + \hat{\theta}_t(\hat{\kappa}_t - \kappa_t)}{\kappa_t \hat{\kappa}_t} \right| \leq \frac{|\theta_t|^{\max} |\kappa_t - \hat{\kappa}_t| + |\kappa_t|^{\max} |\theta_t - \hat{\theta}_t|}{(|\kappa_t|^{\min})^2}.$$

Finally, note

$$\left| R'_t(d_t) - \hat{R}'_t(d_t) \right| = \left| 2 \left(\frac{1}{\kappa_t} - \frac{1}{\hat{\kappa}_t} \right) d_t + \left(\frac{\theta_t}{\kappa_t} - \frac{\hat{\theta}_t}{\hat{\kappa}_t} \right) \right| \leq 2d_t^{\max} \left| \frac{1}{\kappa_t} - \frac{1}{\hat{\kappa}_t} \right| + \left| \frac{\theta_t}{\kappa_t} - \frac{\hat{\theta}_t}{\hat{\kappa}_t} \right|.$$

So by setting

$$\begin{aligned} \alpha_1 &= \frac{(|\kappa_t|^{\min})^2}{2(2d_t^{\max} + |\theta_t|^{\max})} \alpha, \\ \alpha_2 &= \frac{(|\kappa_t|^{\min})^2}{2|\kappa_t|^{\max}} \alpha, \end{aligned}$$

and

$$\begin{aligned} 4 \exp\left(-\frac{2(p_t^1 - p_t^0)^2 N_t \alpha_1^2}{(\omega_t^{\max} - \omega_t^{\min})^2}\right) &\leq \frac{2\beta}{5}, \\ 4 \exp\left(-\frac{(p_t^1 - p_t^0)^2 N_t \alpha_2^2}{8(\omega_t^{\max} - \omega_t^{\min})^2 (p_t^0)^2}\right) &\leq \frac{2\beta}{5}, \\ 2 \exp\left(-\frac{N_t \alpha_2^2}{2(\omega_t^{\max} - \omega_t^{\min})^2}\right) &\leq \frac{\beta}{5}, \end{aligned}$$

we obtain the desired result.

Proof of Lemma 4. Consider any $d_t \in [d_t^{\min}, d_t^{\max}]$ and $y_t \in [y_t^{\min}, y_t^{\max}]$.

First we show the bound for the right derivative. Notice the explicit expressions for the derivatives of the true newsvendor cost function and its data-driven counterpart are:

$$\begin{aligned} \frac{\partial}{\partial y_t} C_t(y_t, d_t) &= -b_t + (h_t + b_t) \mathbb{P}[\eta_t \leq y_t - d_t], \\ \frac{\partial_+}{\partial y_t} \hat{C}_t(y_t, d_t) &= -b_t + (h_t + b_t) \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\hat{\eta}_t^{0j} \leq y_t - d_t]. \end{aligned}$$

If $h_t + b_t = 0$, then the two derivatives are the same and the lemma is trivially true.

Assume $h_t + b_t > 0$. Recall that we want to show that, with probability at least $1 - \beta$,

$$\left| \frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial_+}{\partial y_t} \hat{C}_t(y_t, d_t) \right| \leq \alpha,$$

which is equivalent to

$$\left| \mathbb{P}[\eta_t \leq y_t - d_t] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\hat{\eta}_t^{0j} \leq y_t - d_t] \right| \leq \frac{\alpha}{h_t + b_t}.$$

By definition, $\hat{\eta}_t^{0j} = \eta_t^{0j} + \Delta_t^0$. Therefore, we have

$$\begin{aligned}
& \left| \mathbb{P}[\eta_t \leq y_t - d_t] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\hat{\eta}_t^{0j} \leq y_t - d_t] \right| \\
&= \left| \mathbb{P}[\eta_t \leq y_t - d_t] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\eta_t^{0j} \leq y_t - d_t - \Delta_t^0] \right| \\
&= \left| \mathbb{P}[\eta_t \leq y_t - d_t] - \mathbb{P}[\eta_t \leq y_t - d_t - \Delta_t^0] \right| \\
&\quad + \left| \mathbb{P}[\eta_t \leq y_t - d_t - \Delta_t^0] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\eta_t^{0j} \leq y_t - d_t - \Delta_t^0] \right| \\
&\leq \left| \mathbb{P}[\eta_t \leq y_t - d_t] - \mathbb{P}[\eta_t \leq y_t - d_t - \Delta_t^0] \right| \\
&\quad + \left| \mathbb{P}[\eta_t \leq y_t - d_t - \Delta_t^0] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\eta_t^{0j} \leq y_t - d_t - \Delta_t^0] \right|.
\end{aligned}$$

By the Lipschitz continuity of the c.d.f. of η_t with constant B_t^1 in Assumption 3, we have

$$\left| \mathbb{P}[\eta_t \leq y_t - d_t] - \mathbb{P}[\eta_t \leq y_t - d_t - \Delta_t^0] \right| \leq B_t^1 |\Delta_t^0|.$$

By Massart's inequality (Lemma 10 in Appendix D), we have, for any $\alpha_t > 0$,

$$\mathbb{P} \left[\left| \mathbb{P}[\eta_t \leq y_t - d_t - \Delta_t^0] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\eta_t^{0j} \leq y_t - d_t - \Delta_t^0] \right| \leq \alpha_t \right] \geq 1 - 2 \exp(-2N_t \alpha_t^2).$$

Hence, we have

$$\mathbb{P} \left[\left| \mathbb{P}[\eta_t \leq y_t - d_t] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\hat{\eta}_t^{0j} \leq y_t - d_t] \right| \leq B_t^1 |\Delta_t^0| + \alpha_t \right] \geq 1 - 2 \exp(-2N_t \alpha_t^2).$$

From Fact 6, we have, for any $\alpha_{\Delta_t} > 0$, that

$$\mathbb{P}[|\Delta_t^0| \geq \alpha_{\Delta_t}] \leq 2 \exp\left(-\frac{2N_t \alpha_{\Delta_t}^2}{(\omega_t^{\max} - \omega_t^{\min})^2}\right).$$

Thus, if we let $\alpha_t = \frac{\alpha}{2(h_t + b_t)}$ and $\alpha_{\Delta_t} = \frac{\alpha}{2B_t^1(h_t + b_t)}$, we have

$$\begin{aligned} & \mathbb{P}\left[\left|\mathbb{P}[\eta_t \leq y_t - d_t] - \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{1}[\hat{\eta}_t^{0j} \leq y_t - d_t]\right| \leq \frac{\alpha}{h_t + b_t}\right] \\ & \geq 1 - 2 \exp\left(-\frac{N_t \alpha^2}{2(h_t + b_t)^2}\right) - 2 \exp\left(-\frac{N_t \alpha^2}{2(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2}\right). \end{aligned}$$

Use the fact that $(\omega_t^{\max} - \omega_t^{\min}) B_t^1 \geq 1$, we have

$$\mathbb{P}\left[\left|\frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial}{\partial y_t} \hat{C}_t(y_t, d_t)\right| \leq \alpha\right] \geq 1 - 4 \exp\left(-\frac{N_t \alpha^2}{2(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2}\right).$$

Since $N_t \geq \frac{2(2 \log 2 - \log \beta)(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2}{\alpha^2}$, we have

$$\mathbb{P}\left[\left|\frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial}{\partial y_t} \hat{C}_t(y_t, d_t)\right| \leq \alpha\right] \geq 1 - \beta.$$

So we prove the result for the right derivative. Then we point out the reason why the left derivative preserves the same probabilistic error bound. Note for any $d_t \in [d_t^{\min}, d_t^{\max}]$, $C_t(\cdot, d_t)$ and $\hat{C}_t(\cdot, d_t)$ are both convex, since they are the mathematical expectation (under different distributions) of the convex piecewise-linear newsvendor cost function $h_t(y_t - d_t)^+ + b_t(d_t - y_t)^+$.

We know for any univariate convex function $f : \mathbb{R} \rightarrow \mathbb{R}$, $\lim_{z \rightarrow y} f^r(z) = f^l(y)$, where y is any real number. Moreover, we have $\left|\frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial}{\partial y_t} \hat{C}_t(y_t, d_t)\right| \leq 2(b_t + h_t)$ so

by taking a nonincreasing sequence that converges to y_t and applying the dominated convergence theorem to the left-hand side of the inequality $\left| \frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial_+}{\partial y_t} \hat{C}_t(y_t, d_t) \right| \leq \alpha$, we obtain

$$\mathbb{P} \left[\left| \frac{\partial}{\partial y_t} C_t(y_t, d_t) - \frac{\partial_-}{\partial y_t} \hat{C}_t(y_t, d_t) \right| \leq \alpha \right] \geq 1 - \beta.$$

as well given the assumption on N_t .

Proof of Lemma 6. Consider any $d_t \in [d_t^{\min}, d_t^{\max}]$ and $y_t \in [y_t^{\min}, y_t^{\max}]$. Let $w_t = y_t - d_t$.

Recall that

$$\mathbb{E}_{\hat{\eta}_t} [V'_t(w_t - \hat{\eta}_t)] = \frac{1}{N_t} \sum_{j=1}^{N_t} V'_t(w_t - \hat{\eta}_t^{0j}).$$

Since $\hat{\eta}_t^{0j} = \eta_t^{0j} + \Delta_t^0$. Therefore, we have

$$\mathbb{E}_{\hat{\eta}_t} [V'_t(w_t - \hat{\eta}_t)] = \frac{1}{N_t} \sum_{j=1}^{N_t} V'_t(w_t - \eta_t^{0j} - \Delta_t^0).$$

By triangular inequality,

$$\begin{aligned}
& \left| \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t)] - \frac{1}{N_t} \sum_{j=1}^{N_t} V'_t(w_t - \eta_t^{0j} - \Delta_t^0) \right| \\
&= \left| \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t)] - \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t - \Delta_t^0)] \right. \\
& \quad \left. + \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t - \Delta_t^0)] - \frac{1}{N_t} \sum_{j=1}^{N_t} V'_t(w_t - \eta_t^{0j} + \Delta_t^0) \right| \\
&\leq \underbrace{\left| \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t)] - \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t + \Delta_t^0)] \right|}_{(a)} \\
& \quad + \underbrace{\left| \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t + \Delta_t^0)] - \frac{1}{N_t} \sum_{j=1}^{N_t} V'_t(w_t - \eta_t^{0j} + \Delta_t^0) \right|}_{(b)}.
\end{aligned}$$

Since $V'_t(\cdot)$ is Lipschitz-continuous with constant κ_t (Lemma 2), we have $(a) \leq \kappa_t |\Delta_t^0|$.

From Fact 6, we have, for any $\alpha_a > 0$, that

$$\mathbb{P}[(a) \leq \alpha_a] \geq 1 - 2 \exp\left(-\frac{2N_t \alpha_a^2}{\kappa_t^2 (\omega_t^{\max} - \omega_t^{\min})^2}\right).$$

For (b) , we define a new function

$$v_t(\eta_t^{01}, \dots, \eta_t^{0N_t}) = \frac{1}{N_t} \sum_{j=1}^{N_t} V'_t(w_t - \eta_t^{0j} + \Delta_t^0).$$

We view $\eta_t^{01}, \dots, \eta_t^{0N_t}$ as i.i.d. random variables that have the same distribution as η_t . It is not hard to see that $\mathbb{E}v_t = \frac{1}{N_t} \sum_{j=1}^{N_t} \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t + \Delta_t^0)] = \mathbb{E}_{\eta_t} [V'_t(w_t - \eta_t + \Delta_t^0)]$. So now (b) becomes $|\mathbb{E}v_t - v_t|$.

Due to the Lipschitz continuity of $V'_t(x)$ and the finiteness of the support of η_t

stated in Assumption 2, we first show the following condition: for any realization $\eta_t^{01}, \dots, \eta_t^{0N_t}$ and for any $\eta_t^{0i'}$,

$$\begin{aligned} & |v_t(\eta_t^{01}, \dots, \eta_t^{0i}, \dots, \eta_t^{0N_t}) - v_t(\eta_t^{01}, \dots, \eta_t^{0i'}, \dots, \eta_t^{0N_t})| \\ & \leq \frac{1}{N_t} |V'_t(w_t - \eta_t^{0i} + \Delta_t^0) - V'_t(w_t - \eta_t^{0i'} + \Delta_t^0)| \leq \frac{1}{N_t} \kappa_t |\eta_t - \hat{\eta}_t^i| \leq \frac{1}{N_t} \kappa_t (\omega_t^{\max} - \omega_t^{\min}). \end{aligned}$$

Then, by McDiarmid's Inequality (Lemma 11 in Appendix D), we have, for any α_b ,

$$\mathbb{P}[(b) \leq \alpha_b] = \mathbb{P}[|\mathbb{E}\hat{g} - \hat{g}| \leq \alpha_b] \geq 1 - 2 \exp\left(-\frac{2N_t\alpha_b^2}{\kappa_t^2 (\omega_t^{\max} - \omega_t^{\min})^2}\right).$$

Thus, if we let $\alpha_a = \alpha_b = \frac{\alpha}{2}$, we have

$$\mathbb{P}[(a) + (b) \leq \alpha] \geq 1 - 4 \exp\left(-\frac{N_t\alpha^2}{2\kappa_t^2 (\omega_t^{\max} - \omega_t^{\min})^2}\right).$$

Since $N_t \geq \frac{2(2 \log 2 - \log \beta)(\omega_t^{\max} - \omega_t^{\min})^2 \kappa_t^2}{\alpha^2}$, we have

$$\mathbb{P}[|\mathbb{E}_{\eta_t}[V'_t(y_t - d_t - \eta_t)] - \mathbb{E}_{\hat{\eta}_t}[V'_t(y_t - d_t - \hat{\eta}_t)]| \leq \alpha] \geq 1 - \beta.$$

Proof of Theorem 1. Consider any $d_t \in [d_t^{\min}, d_t^{\max}]$ and $y_t \in [y_t^{\min}, y_t^{\max}]$.

Notice the explicit expressions for the two terms are

$$\begin{aligned} \frac{\partial}{\partial d_t} U_t(y_t, d_t) &= R'_t(d_t) - \frac{\partial}{\partial d_t} C_t(y_t, d_t) - \mathbb{E}_{\eta_t}[V'_t(y_t - d_t - \eta_t)], \\ \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) &= \hat{R}'_t(d_t) - \frac{\partial}{\partial d_t} \hat{C}_t(y_t, d_t) - \mathbb{E}_{\hat{\eta}_t}[\hat{V}'_t(y_t - d_t - \hat{\eta}_t)]. \end{aligned}$$

Therefore, we have

$$\begin{aligned}
& \left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial_-}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \\
& \leq \underbrace{\left| R'_t(d_t) - \hat{R}'_t(d_t) \right|}_{(a)} + \underbrace{\left| \frac{\partial}{\partial d_t} C_t(y_t, d_t) - \frac{\partial_-}{\partial d_t} \hat{C}_t(y_t, d_t) \right|}_{(b)} \\
& + \underbrace{\left| \mathbb{E}_{\eta_t} [V'_t(y_t - d_t - \eta_t)] - \mathbb{E}_{\hat{\eta}_t} [\hat{V}'_t(y_t - d_t - \hat{\eta}_t)] \right|}_{(c)}.
\end{aligned}$$

By triangular inequality, we have

$$\begin{aligned}
(c) & \leq \underbrace{\left| \mathbb{E}_{\eta_t} [V'_t(y_t - d_t - \eta_t)] - \mathbb{E}_{\hat{\eta}_t} [V'_t(y_t - d_t - \hat{\eta}_t)] \right|}_{(d)} \\
& + \underbrace{\left| \mathbb{E}_{\hat{\eta}_t} [V'_t(y_t - d_t - \hat{\eta}_t)] - \mathbb{E}_{\hat{\eta}_t} [\hat{V}'_t(y_t - d_t - \hat{\eta}_t)] \right|}_{(e)}.
\end{aligned}$$

By Lemma 3, we have $\mathbb{P}[(a) \leq \alpha_a] \geq 1 - \beta_a$,

if $N_t \geq \{(2d_t^{\max} + |\theta_t|^{\max})^2, 16(p_t^0)^2(|\kappa_t|^{\max})^2, 4(p_t^1 - p_t^0)^2(|\kappa_t|^{\max})^2\} \frac{2(\omega^{\max} - \omega^{\min})^2 \log(10/\beta)}{(p_t^1 - p_t^0)^2 (|\kappa_t|^{\min})^4 \alpha_a^2}$.

By Lemma 5, we have $\mathbb{P}[(b) \leq \alpha_b] \geq 1 - \beta_b$, if $N_t \geq \frac{2(2 \log 2 - \log \beta_b)(\omega_t^{\max} - \omega_t^{\min})^2 l_t^2 (h_t + b_t)^2}{\alpha_b^2}$.

By Lemma 6, we have $\mathbb{P}[(d) \leq \alpha_d] \geq 1 - \beta_d$, if $N_t \geq \frac{2(2 \log 2 - \log \beta_d)(\omega_t^{\max} - \omega_t^{\min})^2 \kappa_t^2}{\alpha_d^2}$.

By the given assumption that $\left| \hat{V}'_{t+1}(x_t) - V'_{t+1}(x_t) \right| \leq \gamma$, we have $(e) \leq \gamma$.

By setting $\alpha_a = \alpha_b = \alpha_d = \alpha/3$ and $\beta_a = 5\beta/9, \beta_b = \beta_d = 2\beta/9$ and using the bounds on δ_t^{\max} and N_t , we get $\mathbb{P}\left[\left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial_-}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \leq \alpha + \gamma\right] \geq 1 - \beta$.

Since from Lemma 5 we also have the same probabilistic bound of

$\left| \frac{\partial}{\partial d_t} C_t(y_t, d_t) - \frac{\partial_+}{\partial d_t} \hat{C}_t(y_t, d_t) \right|$ as (b), so clearly the same result holds for the probabilistic bound of $\left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial_+}{\partial d_t} \hat{U}_t(y_t, d_t) \right|$.

Proof of Lemma 7. We first prove the result for the left derivatives. We know from Theorem 1 that given $N_t \geq \frac{B_t^1 \log(18/\beta_t)}{\alpha_1^2}$ we have for all $d_t \in [d_t^{\min}, d_t^{\max}]$ and $y_t \in [y_t^{\min}, y_t^{\max}]$, $\mathbb{P} \left[\left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \leq \alpha_1 + \gamma \right] \geq 1 - \beta$, $\mathbb{P} \left[\left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \leq \alpha_1 + \gamma \right] \geq 1 - \beta$. And we can define properly $\alpha_1 = \alpha_{R_t} + \alpha_{C_t}$ as in the proof of Theorem 7. From Fact 7, we know $\hat{U}_t(y_t, d_t)$ is jointly concave so that we must have $0 \in \left[\frac{\partial}{\partial d_t} \hat{U}_t(y_t, \hat{d}_t^*(y_t)), \frac{\partial}{\partial d_t} \hat{U}_t(y_t, \hat{d}_t^*(y_t)) \right]$ from the definition of $\hat{d}_t^*(y_t)$ for any $y_t \in [y_t^{\min}, y_t^{\max}]$. So this implies

$$\mathbb{P} \left[\left| \frac{\partial}{\partial d_t} U_t(y_t, \hat{d}_t^*(y_t)) \right| \leq \alpha_1 + \gamma \right] \geq 1 - \beta.$$

Since we know from definition of $d_t^*(y_t)$ that

$$\frac{\partial}{\partial d_t} U_t(y_t, d_t^*(y_t)) = 0,$$

and using the fact $U_t(y_t, d_t^*(y_t))$ is strictly concave in d_t so $d_t^*(y_t)$ is uniquely defined, we know

$$|R'_t(d_t^*(y_t)) - R'_t(\hat{d}_t^*(y_t))| \leq \alpha_1 + \gamma,$$

with probability at least $1 - \beta$. This step can be verified by observing that in the extreme case such that $f_t^l(y_t - d - t)$ is "stretched out" (think of it remains constant for all $d_t \in [d_t^{\min}, d_t^{\max}]$, which is the way to make $|d_t^*(y_t) - \hat{d}_t^*(y_t)|$ the biggest given a fixed y_t , see Figure B-2), then, since $\hat{R}'_t(d_t)$ and $f_t^l(y_t - d_t)$ can only intersect within the parallelogram constructed by the dash lines with high probability, given assumed conditions. Then it is obvious to see (just need to upper bound the length of the orange line segment in Figure B-2)

Recall from the proof of Lemma 2 that $W'_t(y_t) = R'_t(d_t^*(y_t))$. Moreover, we have

$$\begin{aligned}\hat{W}_t^l(y_t) &= \hat{R}'_t(\hat{d}_t^*(y_t))(\hat{d}_t^*)^l(y_t) \\ &\quad - \left(\frac{\partial}{\partial y_t} \hat{C}_t(y_t, \hat{d}_t^*(y_t)) - \mathbb{E}_{\hat{\eta}_t}[\hat{V}_{t+1}^l(y_t - \hat{d}_t^*(y_t) - \hat{\eta}_t)] \right) (1 - (\hat{d}_t^*)^l(y_t)) \\ &= \hat{R}'_t(\hat{d}_t^*(y_t)) + \left(\frac{\partial}{\partial d_t} \hat{U}_t^l(y_t, \hat{d}_t^*(y_t)) \right) ((\hat{d}_t^*)^l(y_t) - 1).\end{aligned}$$

From Fact 9 we know the term $\left(\frac{\partial}{\partial d_t} \hat{U}_t^l(y_t, \hat{d}_t^*(y_t)) \right) ((\hat{d}_t^*)^l(y_t) - 1)$ equals to zero. Therefore, with probability 1, we have $\hat{W}_t^l(y_t) = \hat{R}'_t(\hat{d}_t^*(y_t))$.

Hence,

$$\begin{aligned}\left| W'_t(y_t) - \hat{W}_t^l(y_t) \right| &= \left| R'_t(d_t^*(y_t)) - \hat{R}'_t(\hat{d}_t^*(y_t)) \right| \\ &\leq \left| R'_t(d_t^*(y_t)) - R'_t(\hat{d}_t^*(y_t)) \right| + \left| R'_t(\hat{d}_t^*(y_t)) - \hat{R}'_t(\hat{d}_t^*(y_t)) \right|.\end{aligned}$$

By Lemma 3, we have $\left| R'_t(d_t^*(y_t)) - R'_t(\hat{d}_t^*(y_t)) \right| \leq \alpha_1/3$.

Hence, with probability at least $1 - \beta$, $\left| W'_t(y_t) - \hat{W}_t^l(y_t) \right| \leq (1 + \frac{1}{3})\alpha_1 + \gamma$.

By our choice of the bounds of N_t , we have $\frac{4}{3}\alpha_1 = \alpha$.

Since $W'_t(y_t), \hat{W}_t^l(y_t)$ are both concave we can repeat the argument at the end of the proof of Lemma 4 to show the same result holds for the bound of $\left| W'_t(y_t) - \hat{W}_t^r(y_t) \right|$.

Proof of Theorem 2. We first prove the result for the left derivatives. Recall from definition we have for all $x_t \in [x_t^{\min}, x_t^{\max}]$,

$$V'_t(x_t) = \begin{cases} 0, & x_t \leq S_t^*, \\ W'_t(x_t), & x_t > S_t^*, \end{cases}$$

$$\hat{V}_t^l(x_t) = \begin{cases} 0, & x_t \leq \hat{S}_t^*, \\ \hat{W}_t^l(x_t), & x_t > \hat{S}_t^*. \end{cases}$$

We consider the following four cases:

(1) $x_t \leq \min(S_t^*, \hat{S}_t^*)$.

$$\left| V_t'(x_t) - \hat{V}_t^l(x_t) \right| = |0 - 0| = 0.$$

(2) $x_t \geq \max(S_t^*, \hat{S}_t^*)$.

By Lemma 7, we have, with probability at least $1 - \beta$,

$$\left| V_t'(x_t) - \hat{V}_t^l(x_t) \right| = \left| W_t'(x_t) - \hat{W}_t^l(x_t) \right| \leq \alpha + \gamma.$$

(3) $S_t^* < x_t < \hat{S}_t^*$.

$$\left| V_t'(x_t) - \hat{V}_t^l(x_t) \right| = |V_t'(x_t) - 0| = |W_t'(x_t)|.$$

Since $W_t'(x_t) \leq 0$ and $W_t'(\cdot)$ is nonincreasing, $|W_t'(x_t)| = -W_t'(x_t) \leq -W_t'(\hat{S}_t^*)$.

By concavity of $\hat{W}_t(y_t)$ and the definition of \hat{S}_t^* , we have $\hat{W}_t^l(\hat{S}_t^*) \geq 0$.

$$\text{Thus, } \left| V_t'(x_t) - \hat{V}_t^l(x_t) \right| \leq \hat{W}_t^l(\hat{S}_t^*) - W_t'(\hat{S}_t^*) \leq \left| \hat{W}_t^l(\hat{S}_t^*) - W_t'(\hat{S}_t^*) \right|.$$

By Lemma 7, we have, with probability at least $1 - \beta$,

$$\left| V_t'(x_t) - \hat{V}_t^l(x_t) \right| \leq \alpha + \gamma.$$

(4) $\hat{S}_t^* < x_t < S_t^*$.

$$\left| V_t'(x_t) - \hat{V}_t^l(x_t) \right| = \left| 0 - \hat{V}_t^l(x_t) \right| = \left| \hat{W}_t^l(x_t) \right|.$$

By concavity of $\hat{W}_t(y_t)$ and the definition of \hat{S}_t^* , we have $\hat{W}_t^l(x_t) < 0$. Thus,

$$\left| \hat{W}_t^l(x_t) \right| = -\hat{W}_t^l(x_t)$$

Since $W'_t(x_t) \geq 0$, $\left|V'_t(x_t) - \hat{V}_t^l(x_t)\right| \leq W'_t(x_t) - \hat{W}_t^l(x_t) \leq \left|W'_t(x_t) - \hat{W}_t^l(x_t)\right|$.

By Lemma 7, we have, with probability at least $1 - \beta$,

$$\left|V'_t(x_t) - \hat{V}_t^l(x_t)\right| \leq \alpha + \gamma.$$

Again, since both $V_t(x_t)$ and $\hat{V}_t(x_t)$ are concave, following the argument at the end of the proof of Lemma 4 produces $\mathbb{P}\left[\left|V'_t(x_t) - \hat{V}_t^r(x_t)\right| \leq \alpha + \gamma\right] \geq 1 - \beta$.

Proof of Theorem 3. We just provide the proof for the left derivatives, since the proof for the right derivatives are identical. Define the probabilistic events

$$\begin{aligned} A_t &= \left\{ \forall s = t, \dots, T : \left| \frac{\partial_-}{\partial d_t} \hat{U}_t(y_t, d_t) - \frac{\partial}{\partial d_t} U_t(y_t, d_t) \right| \leq \sum_{r=s}^T \alpha_r, \right. \\ &\quad \left. \forall y_t \in [y_t^{\min}, y_t^{\max}], d_t \in [d_t^{\min}, d_t^{\max}] \right\}, \\ B_t &= \left\{ \forall s = t, \dots, T : \left| \hat{W}_t^l(y_t) - W'_t(y_t) \right| \leq \sum_{r=s}^T \alpha_r, \forall y_t \in [y_t^{\min}, y_t^{\max}] \right\}, \\ C_t &= \left\{ \forall s = t, \dots, T : \left| \hat{V}_t^l(x_t) - V'_t(x_t) \right| \leq \sum_{r=s}^T \alpha_r, \forall x_t \in [x_t^{\min}, x_t^{\max}] \right\}. \end{aligned}$$

Given our assumption on the bound of N_t , by Theorem 1, $\mathbb{P}[A_t | C_{t+1}] \geq 1 - \beta/T$.

By Lemma 7, $\mathbb{P}[B_t | C_{t+1}] \geq 1 - \beta/T$.

By Theorem 2, $\mathbb{P}[C_t | C_{t+1}] \geq 1 - \beta/T$.

So we have

$$\begin{aligned} & \mathbb{P} \left[\forall t = 1, \dots, T, \forall x \in [x_t^{\min}, x_t^{\max}] : |\hat{V}_t^l(x) - V_t'(x)| \leq \sum_{s=t}^T \alpha_s \right] \\ & \geq \mathbb{P} \left[\bigcap_{s=1}^T C_s \right] = \prod_{t=1}^T \mathbb{P}[C_t | C_{t+1}] \geq 1 - \beta. \end{aligned}$$

And this shows

$$\begin{aligned} & \mathbb{P} \left[\text{for all } d_t \in [d_t^{\min}, d_t^{\max}], y_t \in [y_t^{\min}, y_t^{\max}] : \left| \frac{\partial}{\partial d_t} U_t(y_t, d_t) - \frac{\partial_-}{\partial d_t} \hat{U}_t(y_t, d_t) \right| \leq \sum_{s=t}^T \alpha_s \right] \\ & \geq 1 - \beta, \\ & \mathbb{P} \left[\text{for all } y_t \in [y_t^{\min}, y_t^{\max}] : \left| \hat{W}_t^l(y_t) - W_t'(y_t) \right| \leq \sum_{s=t}^T \alpha_s \right] \geq 1 - \beta. \end{aligned}$$

Proof of Lemma 8. We just provide the proof for $W_t(\hat{S}_t^*)$ since the proof for $U_t(\hat{d}_t^*(y_t), y_t)$ is identical.

Notice by concavity of $W_t(\cdot)$, $W_t(S_t^*) \leq W_t(\hat{S}_t^*) + |W'(\hat{S}_t^*)| |\hat{S}_t^* - S_t^*|$.

By Theorem 3 and our assumption on the bound of N_t , we have $\left| \hat{W}_t^l(\hat{S}_t^*) - W_t'(\hat{S}_t^*) \right| \leq \sum_{s=t}^T \alpha_s$, $\left| \hat{W}_t^r(\hat{S}_t^*) - W_t'(\hat{S}_t^*) \right| \leq \sum_{s=t}^T \alpha_s$, with probability at least $1 - \beta$.

By the definition of \hat{S}_t^* and concavity of $\hat{W}_t(\hat{S}_t^*)$, we know that $0 \in [\hat{W}_t^r(\hat{S}_t^*), \hat{W}_t^l(\hat{S}_t^*)]$.

Therefore, we have $\left| W_t'(\hat{S}_t^*) \right| \leq \sum_{s=t}^T \alpha_s$.

Hence, $W_t(\hat{S}_t^*) \geq W_t(S_t^*) - \sum_{s=t}^T \alpha_s |\hat{S}_t^* - S_t^*|$.

Since we have defined

$$\lambda_t^W = \min \left\{ \frac{W_t(S_t^*) - W_t(y_t^{\min})}{S_t^* - y_t^{\min}}, \frac{W_t(y_t^{\max}) - W_t(S_t^*)}{y_t^{\max} - S_t^*} \right\},$$

we have

$$\begin{aligned} W_t(S_t^*) &\geq W_t(y_t^{\min}) + \lambda_t^W (S_t^* - y_t^{\min}), \\ W_t(S_t^*) &\geq W_t(y_t^{\max}) + \lambda_t^W (y_t^{\max} - S_t^*). \end{aligned}$$

By Assumption 5, $W_t(\cdot)$ is nonnegative within $[y_t^{\min}, y_t^{\max}]$.

Combining these two equations, we get

$$\begin{aligned} W_t(S_t^*) &\geq \frac{W_t(y_t^{\min}) + W_t(y_t^{\max})}{2} + \frac{\lambda_t^W}{2} (S_t^* - y_t^{\min}) + \frac{\lambda_t^W}{2} (y_t^{\max} - S_t^*) \\ &\geq \frac{W_t(y_t^{\min}) + W_t(y_t^{\max})}{2} + \frac{\lambda_t^W}{2} (y_t^{\max} - y_t^{\min}) \\ &= \frac{\lambda_t^W}{2} (y_t^{\max} - y_t^{\min}). \end{aligned}$$

Furthermore, since $|\hat{S}_t^* - S_t^*| \leq (y_t^{\max} - y_t^{\min})$, we derive that

$$W_t(\hat{S}_t^*) \geq W_t(S_t^*) - \sum_{s=t}^T \alpha_s (y_t^{\max} - y_t^{\min}) \geq W_t(S_t^*) - \frac{2}{\lambda_t^W} \sum_{s=t}^T \alpha_s W_t(S_t^*) = (1 - \epsilon) W_t(S_t^*).$$

Proof of Theorem 4. Note from definition of \hat{S}_t^* , we have

$$\pi_t(x_t; \hat{\mu}_t) = \begin{cases} R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \hat{\mu}_{t+1})], & x_t \leq \hat{S}_t^*, \\ R_t(\hat{d}_t^*(x_t)) - C_t(x_t, \hat{d}_t^*(x_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(x_t - \hat{d}_t^*(x_t) - \eta_t; \hat{\mu}_{t+1})], & x_t > \hat{S}_t^*. \end{cases}$$

By assumption, we have $\pi_{t+1}(x_{t+1}; \hat{\mu}_{t+1}) \geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi_{t+1}(x_{t+1}; \mu_{t+1}^*)$.

We consider the all four cases:

(1) $x_t \leq \min(S_t^*, \hat{S}_t^*)$.

$$\begin{aligned}
\pi_t(x_t; \hat{\mu}_t) &= R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \hat{\mu}_{t+1})] \\
&\geq R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) \\
&\quad + \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \mu_{t+1}^*)] \\
&\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \left(R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*))\right) \\
&\quad + \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \left(\mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \mu_{t+1}^*)]\right) \\
&= \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)).
\end{aligned}$$

By Lemma 8, we have $U_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) \geq (1 - \epsilon_t^U) U_t(\hat{S}_t^*, d_t^*(\hat{S}_t^*)) = (1 - \epsilon_t^U) W_t(\hat{S}_t^*)$.

Again by Lemma 8, we have $W_t(\hat{S}_t^*) \geq (1 - \epsilon_t^W) W_t(S_t^*) = (1 - \epsilon_t^W) \pi_t(x_t; \mu_t^*)$.

Therefore, we have $\pi_t(x_t; \hat{\mu}_t) \geq \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U - \epsilon_t^W\right) \pi_t(x_t; \mu_t^*)$.

(2) $x_t \geq \max(S_t^*, \hat{S}_t^*)$.

$$\begin{aligned}
\pi_t(x_t; \hat{\mu}_t) &= R_t(\hat{d}_t^*(x_t)) - C_t(x_t, \hat{d}_t^*(x_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(x_t - \hat{d}_t^*(x_t) - \eta_t; \hat{\mu}_{t+1})] \\
&\geq R_t(\hat{d}_t^*(x_t)) - C_t(x_t, \hat{d}_t^*(x_t)) \\
&\quad + \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \mathbb{E}_{\eta_t}[\pi_{t+1}(x_t - \hat{d}_t^*(x_t) - \eta_t; \mu_{t+1}^*)] \\
&\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \left(R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*))\right) \\
&\quad + \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \left(\mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \mu_{t+1}^*)]\right) \\
&= \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, \hat{d}_t^*(x_t)).
\end{aligned}$$

By Lemma 8, we have $U_t(x_t, \hat{d}_t^*(x_t)) \geq (1 - \epsilon_t^U) U_t(x_t, d_t^*(x_t)) = (1 - \epsilon_t^U) \pi_t(x_t; \mu_t^*)$.

Therefore, we have $\pi_t(x_t; \hat{\mu}_t) \geq (1 - \epsilon_{t+1} - \epsilon_t^U) \pi_t(x_t; \mu_t^*)$.

(3) $S_t^* < x_t < \hat{S}_t^*$.

$$\begin{aligned}
\pi_t(x_t; \hat{\mu}_t) &= R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \hat{\mu}_{t+1})] \\
&\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) \\
&\geq \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U\right) U_t(\hat{S}_t^*, d_t^*(\hat{S}_t^*)) = (1 - \epsilon_{t+1} - \epsilon_t^U) W_t(\hat{S}_t^*) \\
&\geq \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U - \epsilon_t^W\right) W_t(S_t^*).
\end{aligned}$$

Since $S_t^* < x_t$, we have $\pi_t(x_t; \mu_t^*) = W_t(x_t) \leq W_t(S_t^*)$.

Therefore, we have $\pi_t(x_t; \hat{\mu}_t) \geq \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U - \epsilon_t^W\right) \pi_t(x_t; \mu_t^*)$.

$$(4) \quad \hat{S}_t^* < x_t < S_t^*.$$

$$\begin{aligned} \pi_t(x_t; \hat{\mu}_t) &= R_t(\hat{d}_t^*(x_t)) - C_t(x_t, \hat{d}_t^*(x_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(x_t - \hat{d}_t^*(x_t) - \eta_t; \hat{\mu}_{t+1})] \\ &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, \hat{d}_t^*(x_t)) \\ &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U\right) U_t(x_t, d_t^*(x_t)) \\ &= \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U\right) W_t(x_t). \end{aligned}$$

Since $\hat{S}_t^* < x_t < S_t^*$, we have $W_t(x_t) \geq W_t(\hat{S}_t^*)$.

By Lemma 8, we have $W_t(\hat{S}_t^*) \geq (1 - \epsilon_t^W)W_t(S_t^*) = (1 - \epsilon_t^W)\pi_t(x_t; \mu_t^*)$.

Therefore, we have $\pi_t(x_t; \hat{\mu}_t) \geq \left(1 - \sum_{s=t+1}^T \epsilon_s - \epsilon_t^U - \epsilon_t^W\right) \pi_t(x_t; \mu_t^*)$.

So we complete the induction step.

Proof of Theorem 5. This is a direct result by setting

$$\alpha_t = \frac{\min_{t=1, \dots, T} \min\{\lambda_t^U, \lambda_t^W\} \epsilon}{4T^2}, \forall t = 1, \dots, T,$$

in Theorem 4 and given the choice of N_t for all t , completing the induction from $t = T$ to 1.

Proof of Theorem 6. The proof proceeds similar to the analysis in Section 4.1 and 4.2. We observe the statement of Lemma 3 can be replaced by the given condition on $\hat{R}_t(d_t)$. A key observation is that after imposing the capacity constraints, Fact 1, 2, 3, 4, 5, 7 still hold, so that Fact 9 and Lemma 7 still hold. Note the first-order approximation bounds on $\left| \frac{\partial}{\partial d_t} \hat{U}_t(y_t, d_t) - \frac{\partial}{\partial d_t} U_t(y_t, d_t) \right|$ are irrelevant to I_t , so by proving the claim that the statement of Theorem 2 and Theorem 4 remain unchanged after introducing I_t , we complete the proof.

Modified Proof of Theorem 2. We just need to modify the proof the left derivatives. Recall from definition we have for all $x_t \in [x_t^{\min}, x_t^{\max}]$,

$$V_t'(x_t) = \begin{cases} W_t'(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, S_t^* - I_t), \\ 0, & \text{if } x_t \in [S_t^* - I_t, S_t^*), \\ W_t'(x_t), & \text{if } x_t \in [S_t^*, x_t^{\max}]. \end{cases}$$

$$\hat{V}_t^l(x_t) = \begin{cases} \hat{W}_t^l(x_t + I_t), & \text{if } x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ 0, & \text{if } x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ \hat{W}_t^l(x_t), & \text{if } x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$$

We consider the following four cases: 1. $\hat{S}_t^* \leq S_t^* - I_t$; 2. $S_t^* - I_t < \hat{S}_t^* \leq S_t^*$; 3.

$$\hat{S}_t^* - I_t \leq S_t^* < \hat{S}_t^*; \text{ 4. } S_t^* < \hat{S}_t^* - I_t.$$

$$1. \hat{S}_t^* \leq S_t^* - I_t.$$

$$(1) \ x_t \leq \hat{S}_t^* - I_t. \text{ We know } x_t \leq S_t^* - I_t \text{ as well. We have } V_t'(x_t) = W_t'(x_t + I_t), \\ \hat{V}_t^l(x_t) = \hat{W}_t^l(x_t + I_t). \text{ So } |V_t'(x_t) - \hat{V}_t^l(x_t)| = |W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t)| \leq \\ \alpha + \gamma.$$

$$(2) \ \hat{S}_t^* - I_t < x_t \leq \hat{S}_t^*. \text{ We have } V_t'(x_t) = W_t'(x_t + I_t), \hat{V}_t^l(x_t) = 0. \text{ Since } \\ W_t'(x_t + I_t) \geq W_t'(S_t^*) = 0 \text{ (concavity of } W_t(\cdot)), \hat{W}_t^l(x_t + I_t) < 0 \text{ (} x_t + I_t > \hat{S}_t^* \\ \text{and the definition of } \hat{S}_t^*), \text{ we have } |V_t'(x_t) - \hat{V}_t^l(x_t)| = V_t'(x_t) - \hat{V}_t^l(x_t) = \\ W_t'(x_t + I_t) - 0 \leq W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t) = |W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t)| \leq \\ \alpha + \gamma.$$

$$(3) \ \hat{S}_t^* < x_t \leq S_t^*. \text{ We have } V_t'(x_t) = W_t'(x_t + I_t), \hat{V}_t^l(x_t) = \hat{W}_t^l(x_t). \text{ Since } \\ W_t'(x_t + I_t) \geq W_t'(S_t^*) = 0 \text{ (concavity of } W_t(\cdot)), \hat{W}_t^l(x_t + I_t) \leq \hat{W}_t^l(x_t) < 0 \\ \text{(concavity of } \hat{W}_t(\cdot) \text{ and the definition of } \hat{S}_t^*), \text{ we have } |V_t'(x_t) - \hat{V}_t^l(x_t)| = \\ V_t'(x_t) - \hat{V}_t^l(x_t) = W_t'(x_t + I_t) - \hat{W}_t^l(x_t) \leq W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t) = \\ |W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t)| \leq \alpha + \gamma.$$

$$(4) \ S_t^* - I_t < x_t \leq S_t^*. \text{ We have } V_t'(x_t) = 0, \hat{V}_t^l(x_t) = \hat{W}_t^l(x_t). \text{ Since } W_t'(x_t) \geq \\ W_t'(S_t^*) = 0 \text{ (concavity of } W_t(\cdot)), \hat{W}_t^l(x_t) < 0 \text{ (definition of } \hat{S}_t^*), \text{ we have } \\ |V_t'(x_t) - \hat{V}_t^l(x_t)| = |0 - \hat{W}_t^l(x_t)| = -\hat{W}_t^l(x_t) \leq W_t'(x_t) - \hat{W}_t^l(x_t) = |W_t'(x_t) - \\ \hat{W}_t^l(x_t)| \leq \alpha + \gamma.$$

$$(5) \ x_t > S_t^*. \text{ We have } V_t'(x_t) = W_t'(x_t), \hat{V}_t^l(x_t) = \hat{W}_t^l(x_t). \text{ So } |V_t'(x_t) - \\ \hat{V}_t^l(x_t)| = |W_t'(x_t) - \hat{W}_t^l(x_t)| \leq \alpha + \gamma.$$

$$2. S_t^* - I_t < \hat{S}_t^* \leq S_t^*.$$

$$(1) \ x_t \leq \hat{S}_t^* - I_t. \text{ We have } V_t'(x_t) = W_t'(x_t + I_t), \hat{V}_t^l(x_t) = \hat{W}_t^l(x_t + I_t). \text{ So } \\ |V_t'(x_t) - \hat{V}_t^l(x_t)| = |W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t)| \leq \alpha + \gamma.$$

- (2) $\hat{S}_t^* - I_t < x_t \leq S_t^* - I_t$. We have $V_t'(x_t) = W_t'(x_t + I_t)$, $\hat{V}_t^l(x_t) = 0$. Since $W_t'(x_t + I_t) \geq W_t'(S_t^*) = 0$ ($x_t + I_t \leq S_t^*$ and concavity of $W_t(\cdot)$), $\hat{W}_t^l(x_t + I_t) < 0$ ($x_t + I_t > \hat{S}_t^*$ and the definition of \hat{S}_t^*), we have $|V_t'(x_t) - \hat{V}_t^l(x_t)| = |W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t)| \leq |W_t'(x_t + I_t) - \hat{W}_t^l(x_t + I_t)| \leq \alpha + \gamma$.
- (3) $S_t^* - I_t < x_t \leq \hat{S}_t^*$. We have $V_t'(x_t) = 0$, $\hat{V}_t^l(x_t) = 0$. So $|V_t'(x_t) - \hat{V}_t^l(x_t)| = 0 \leq \alpha + \gamma$.
- (4) $\hat{S}_t^* < x_t \leq S_t^*$. We have $V_t'(x_t) = 0$, $\hat{V}_t^l(x_t) = \hat{W}_t^l(x_t)$. Since $W_t'(x_t) \geq W_t'(S_t^*) = 0$ (concavity of $W_t(\cdot)$), $\hat{W}_t^l(x_t) < 0$ (definition of \hat{S}_t^*), we have $|V_t'(x_t) - \hat{V}_t^l(x_t)| = |0 - \hat{W}_t^l(x_t)| = -\hat{W}_t^l(x_t) \leq W_t'(x_t) - \hat{W}_t^l(x_t) = |W_t'(x_t) - \hat{W}_t^l(x_t)| \leq \alpha + \gamma$.
- (5) $x_t > S_t^*$. We have $V_t'(x_t) = W_t'(x_t)$, $\hat{V}_t^l(x_t) = \hat{W}_t^l(x_t)$. So $|V_t'(x_t) - \hat{V}_t^l(x_t)| = |W_t'(x_t) - \hat{W}_t^l(x_t)| \leq \alpha + \gamma$.
3. $\hat{S}_t^* - I_t \leq S_t^* < \hat{S}_t^*$. The same proof follows as in 2. by interchanging S_t^* and \hat{S}_t^* .
4. $S_t^* < \hat{S}_t^* - I_t$. The same proof follows as in 1. by interchanging S_t^* and \hat{S}_t^* .

Thus we complete the proof.

Modified Proof of Theorem 4. Note from definition of S_t^* and \hat{S}_t^* , we have

$$\pi_t(x_t; \mu_t^*) = \begin{cases} U_t(x_t + I_t, d_t^*(x_t + I_t)), & x_t \in [x_t^{\min}, S_t^* - I_t), \\ U_t(S_t^*, d_t^*(S_t^*)), & x_t \in [S_t^* - I_t, S_t^*), \\ U_t(x_t, d_t^*(x_t)), & x_t \in [S_t^*, x_t^{\max}]. \end{cases}$$

$$\pi_t(x_t; \hat{\mu}_t) = \begin{cases} R_t(\hat{d}_t^*(\bar{x}_t)) - C_t(x_t + I_t, \hat{d}_t^*(\bar{x}_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(x_t + I_t - \hat{d}_t^*(\bar{x}_t) - \eta_t; \hat{\mu}_{t+1})], & x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ R_t(\hat{d}_t^*(\hat{S}_t^*)) - C_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(\hat{S}_t^* - \hat{d}_t^*(\hat{S}_t^*) - \eta_t; \hat{\mu}_{t+1})], & x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ R_t(\hat{d}_t^*(x_t)) - C_t(x_t, \hat{d}_t^*(x_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(x_t - \hat{d}_t^*(x_t) - \eta_t; \hat{\mu}_{t+1})], & x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$$

Note we use the notation $\bar{x}_t = x_t + I_t$. By assumption, we have $\pi_{t+1}(x_{t+1}; \hat{\mu}_{t+1}) \geq (1 - \sum_{s=t+1}^T \epsilon_s) \pi_{t+1}(x_{t+1}; \mu_{t+1}^*)$. Thus, we know for all $y_t \in [y_t^{\min}, y_t^{\max}]$, we have

$$\begin{aligned} & R_t(\hat{d}_t^*(y_t)) - C_t(y_t, \hat{d}_t^*(y_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(y_t - \hat{d}_t^*(y_t) - \eta_t; \hat{\mu}_{t+1})] \\ & \geq R_t(\hat{d}_t^*(y_t)) - C_t(y_t, \hat{d}_t^*(y_t)) + \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \mathbb{E}_{\eta_t}[\pi_{t+1}(y_t - \hat{d}_t^*(y_t) - \eta_t; \mu_{t+1}^*)] \\ & \geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \left(R_t(\hat{d}_t^*(y_t)) - C_t(y_t, \hat{d}_t^*(y_t)) + \mathbb{E}_{\eta_t}[\pi_{t+1}(y_t - \hat{d}_t^*(y_t) - \eta_t; \mu_{t+1}^*)]\right) \\ & = \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(y_t, \hat{d}_t^*(y_t)). \end{aligned}$$

Therefore, we have

$$\pi(x_t; \hat{\mu}_t) \geq \begin{cases} \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, \hat{d}_t^*(x_t + I_t)), & x_t \in [x_t^{\min}, \hat{S}_t^* - I_t), \\ \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)), & x_t \in [\hat{S}_t^* - I_t, \hat{S}_t^*), \\ \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, \hat{d}_t^*(x_t)), & x_t \in [\hat{S}_t^*, x_t^{\max}]. \end{cases}$$

We consider the following two cases: 1. $S_t^* \leq \hat{S}_t^*$; 2. $S_t^* > \hat{S}_t^*$.

1. $S_t^* \leq \hat{S}_t^*$.

(1) $x_t \leq S_t^* - I_t$. Note $x_t \leq \hat{S}_t^* - I_t$ is also implied. Then, by Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, \hat{d}_t^*(x_t + I_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, d_t^*(x_t + I_t)) \\
&= (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

(2) $S_t^* - I_t < x_t \leq \hat{S}_t^* - I_t$. Note $x_t + I_t \geq S_t^*$. Then, by Lemma 8 and concavity of $U_t(\cdot, d_t^*(\cdot))$,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, \hat{d}_t^*(x_t + I_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, d_t^*(x_t + I_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(S_t^*, d_t^*(S_t^*)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

(3) $\hat{S}_t^* - I_t < x_t \leq \hat{S}_t^*$. By Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, d_t^*(\hat{S}_t^*)) \\
&\geq (1 - \epsilon_t^U - \epsilon_t^W) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(S_t^*, d_t^*(S_t^*)) \\
&\geq (1 - \epsilon_t^U - \epsilon_t^W) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

(4) $x_t > \hat{S}_t^*$. Note $x_t > S_t^*$ is also implied. Then, by Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, \hat{d}_t^*(x_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, d_t^*(x_t)) \\
&= (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

2. $S_t^* > \hat{S}_t^*$.

(1) $x_t \leq \hat{S}_t^* - I_t$. It is also implied $x_t \leq S_t^* - I_t$. Then, by Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, \hat{d}_t^*(x_t + I_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t + I_t, d_t^*(x_t + I_t)) \\
&= (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

(2) $\hat{S}_t^* - I_t < x_t \leq \hat{S}_t^*$. By Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, \hat{d}_t^*(\hat{S}_t^*)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(\hat{S}_t^*, d_t^*(\hat{S}_t^*)) \\
&\geq (1 - \epsilon_t^U - \epsilon_t^W) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(S_t^*, d_t^*(S_t^*)) \\
&\geq (1 - \epsilon_t^U - \epsilon_t^W) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

(3) $\hat{S}_t^* < x_t \leq S_t^*$. By concavity of $U_t(\cdot, d_t^*(\cdot))$ and Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, \hat{d}_t^*(x_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, d_t^*(x_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(S_t^*, d_t^*(S_t^*)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

(4) $x_t > S_t^*$. Note $x_t > \hat{S}_t^*$ is also implied. Then, by Lemma 8,

$$\begin{aligned}
\pi(x_t; \hat{\mu}_t) &\geq \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, \hat{d}_t^*(x_t)) \\
&\geq (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) U_t(x_t, d_t^*(x_t)) \\
&= (1 - \epsilon_t^U) \left(1 - \sum_{s=t+1}^T \epsilon_s\right) \pi(x_t; \mu_t^*).
\end{aligned}$$

So we complete the induction step.

Appendix D

Additional Lemmas

Lemma 9 (*Hoeffding's inequality*) Suppose the variables X_1, \dots, X_n are independent, and X_i has mean μ_i and finite support $[a_i, b_i]$. Then for all $t \geq 0$, we have

$$\mathbb{P} \left[\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i \leq -t \right] \leq e^{-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}}.$$

$$\mathbb{P} \left[\left| \sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i \right| \geq t \right] \leq 2e^{-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}}.$$

Lemma 10 (*Massart's inequality*) Suppose the variables X_1, \dots, X_n are independent and identically distributed with cumulative distribution function $F(\cdot)$. Let F_n denote the associated empirical distribution function defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n 1[X_i \leq x], x \in \mathbb{R}.$$

Then for all $t \geq 0$,

$$\mathbb{P} \left[\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \geq t \right] \leq 2e^{-2nt^2}.$$

Lemma 11 (*McDiarmid's inequality*) Suppose the variables X_1, \dots, X_n are independent taking values in the set \mathcal{X} . Further, let $f : \mathcal{X}^n \rightarrow \mathbb{R}$ be a function of X_1, \dots, X_n that satisfies for all $x_1 \in \mathcal{X}, \dots, x_n \in \mathcal{X}, \forall i, x'_i \in \mathcal{X}$,

$$|f(x_1, \dots, x_i, \dots, x_n) - f(x_1, \dots, x'_i, \dots, x_n)| \leq c_i.$$

Then for all $t > 0$,

$$\mathbb{P}[|f - \mathbb{E}[f]| \geq t] \leq 2e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}.$$

Bibliography

- [1] Gah-Yi Ban. Confidence intervals for data-driven inventory policies with demand censoring. 2018.
- [2] Boxiao Chen, Xiuli Chao, and Hyun-Soo Ahn. Coordinating pricing and inventory replenishment with nonparametric demand learning. 2015.
- [3] Xin Chen and David Simchi-Levi. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, 52(6):887–896, 2004.
- [4] Xin Chen and David Simchi-Levi. Pricing and inventory management, 2012.
- [5] Wang Chi Cheung and David Simchi-Levi. Sampling-based approximation schemes for capacitated stochastic inventory control models. 2015.
- [6] Martin Dyer and Leen Stougie. Computational complexity of stochastic programming problems. *mathematical programming*, 106(3):423–432, 2006.
- [7] Wedad Elmaghraby and Pinar Keskinocak. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management science*, 49(10):1287–1309, 2003.
- [8] Awi Federgruen and Aliza Heching. Combined pricing and inventory control under uncertainty. *Operations research*, 47(3):454–475, 1999.
- [9] Nir Halman. Provably near-optimal approximation schemes for sample-based dynamic programs with emphasis on stochastic inventory control models, 2015.
- [10] Nir Halman, Diego Klabjan, Mohamed Mostagir, Jim Orlin, and David Simchi-Levi. A fully polynomial-time approximation scheme for single-item stochastic inventory control with discrete demand. *Mathematics of Operations Research*, 34(3):674–685, 2009.

- [11] Retsef Levi, Robin O Roundy, and David B Shmoys. Provably near-optimal sampling-based policies for stochastic inventory control models. *Mathematics of Operations Research*, 32(4):821–839, 2007.
- [12] Nicholas C Petruzzi and Maqbool Dada. Pricing and the newsvendor problem: A review with extensions. *Operations research*, 47(2):183–194, 1999.
- [13] MMCR Reaiche. A note on sample complexity of multistage stochastic programs. *Operations Research Letters*, 44(4):430–435, 2016.
- [14] Alexander Shapiro and Arkadi Nemirovski. On complexity of stochastic programming problems. In *Continuous optimization*, pages 111–146. Springer, 2005.
- [15] Thomson M Whitin. Inventory control and price theory. *Management science*, 2(1):61–68, 1955.
- [16] Candace A Yano and Stephen M Gilbert. Coordinated pricing and production/procurement decisions: A review. In *Managing Business Interfaces*, pages 65–103. Springer, 2005.