# Data Driven Operations: From Algorithm Development to Experimental Design

by

Jinglong Zhao

Submitted to the Institute for Data, Systems, and Society
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Social and Engineering Systems and Statistics
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2021



Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Institute for Data, Systems, and Society
May 13, 2021

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
David Simchi-Levi
Professor of Engineering Systems
Thesis Supervisor

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Alberto Abadie
Professor of Economics
Thesis Committee Member

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Iavor Bojinov
Assistant Professor of Business Administration
Thesis Committee Member

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
John Tsitsiklis
Clarence J. Lebel Professor of Electrical Engineering
Thesis Committee Member

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Fotini Christia
Chair, Social and Engineering Systems Doctoral Program

**Data Driven Operations: From Algorithm Development to Experimental Design**

by

Jinglong Zhao

Submitted to the Institute for Data, Systems, and Society
on May 13, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Social and Engineering Systems and Statistics

## Abstract

Digital innovation has gained increasing attention in today's world. The explosion of data that is generated through modern marketplaces provides new opportunities to use data-driven tools to understand and optimize the marketplace operations. This dissertation studies various problems around the following two pillars of data-driven operations: optimization and econometrics.

In the first module we focus on optimization, in which we consider dynamic resource allocation problems under zero adaptivity. Dynamic resource allocation problems are omnipresent in modern business operations. In the revenue management setting, there are unreplenishable resources to allocate to heterogeneous consumer demands, immediately and irrevocably upon their arrivals. In such settings, zero adaptivity refers to a policy whose actions are independent of the remaining resources. Traditional revenue management literature has mainly focused on fully adaptive policies; and there is a gap between the provable effectiveness of adaptive policies in theory, and the applicability of non-adaptive policies in practice. We show that under different models of demand uncertainty, carefully designed non-adaptive policies may provably perform almost as well as the best fully adaptive counterparts.

In the second module we focus on econometrics, in which we consider experimental design problems. Experimental design is a widely adopted approach for firms to evaluate the effectiveness of their initiatives, by comparing the standard offering to a new initiative. Such a task is often challenging due to interference, both over time and across units. Traditional experimental design methods suffer from large variances of the estimators when accounting for interference; and practitioners have recognized that insufficient precision may lead to unreliable inference. We build the theoretical foundations to use optimization approach to maximize the precision when designing experiments.

Finally, we conclude with discussions of the limitations of the models and methods we have considered. We also provide practical suggestions to applied researchers and data scientists.

Thesis Supervisor: David Simchi-Levi
Title: Professor of Engineering Systems

Thesis Committee Member: Alberto Abadie
Title: Professor of Economics

Thesis Committee Member: Iavor Bojinov
Title: Assistant Professor of Business Administration

Thesis Committee Member: John Tsitsiklis
Title: Clarence J. Lebel Professor of Electrical Engineering

# Acknowledgments

I have been extremely fortunate to have learned from a number of extraordinary researchers prior to and during my doctoral study. In 2021, I was able to graduate and pursue my childhood dream of becoming an academic. This would have not been possible without the countless people who have supported me over the years.

First and foremost, I would like to thank my advisor, Professor David Simchi-Levi, who has been a great mentor to me. I met him when I was an undergraduate student at Tsinghua University, and have discussed research with him even before my doctoral study at MIT. Over the years, David has never failed to amaze me with his shiny research ideas. He has guided my research both at a high level, and at a tactical level, and has always been extremely to-the-point during our conversations. Beyond research, I have received generous funding support and numerous industry experiences through the consulting projects as well as tremendous suggestions and help in my final-year academic job market search. Moreover, I have learned from David's wisdom during my doctoral study. I always look up to him as an emulated figure himself whose academic career keeps growing through focus, perseverance, and relentless efforts, who has never failed to motivate and inspire me.

Next, I would like to thank my thesis committee members. Professor Alberto Abadie has taught two of my econometrics courses that significantly re-shaped the research flavor of my doctoral study. Alberto has opened up the door for me to peek the glowing new world, has guided me hand-by-hand by answering my silly questions, and has encouraged me to pursue any direction in the field. I have received incredible supports both within and beyond research, that have come to me like spring breezes. Professor Iavor Bojinov is one of my coauthors. I wish I could have started our collaboration earlier during my doctoral study. Beyond co-authorship, I have received detailed suggestions from him as someone who has just went through the process of academic job market search, which is very valuable to me. Professor John Tsitsiklis has taught my linear optimization and dynamic optimization courses that laid the solid mathematical foundation of my doctoral study. John has planted in my heart a seed of curiosity, and has motivated me to always seek mathematical elegance and simplicity. I have had many interactions and discussions with him over the years, and I always admire him as a real scholar.

My two undergraduate advisors have accompanied me over the years. They have always been playing fatherly role to me even during my doctoral study. Without their consistent supports I would not have been where I stand today. Professor Xin Chen has kindly trained me on all the details in conducting high-quality research. In the earlier years he has dived into each single proof with me on the fascinating concepts of discrete convexity, and has enlightened me with his research philosophy. He never fails to answer any single question I have, both within and beyond the scope of research, through his deep knowledge and great patience. Professor Zuo-Jun (Max) Shen has gently taught me all the basics about operations research. I always remember the earlier years when I carefully digested the thick introductory books and greedily gulped down all the knowledge by closely following his guidance and suggestions. He has constantly given me various suggestions during my doctoral study, both research-wise and life-wise, to ensure that I am on a right path. The training and teaching I have received from my two undergraduate advisors have a profound impact on me that will persist through my life.

I really appreciate the support from my home department IDSS. I would like to thank the faculties at IDSS, especially Professors Alberto Abadie, Fotini Christia, Munther Dahleh, Dean Eckles, Stephen Graves, Ali Jadbabaie, Sandy Pentland, Noelle Selin, Devavrat Shah, John Tsitsiklis, Tauhid Zaman, and Jinhua Zhao, for creating the IDSS doctoral program, for creating a collaborative atmosphere, and most importantly, for making IDSS supportive. I would also like to thank the administrative staff at IDSS for their help, especially Laura Dorson, Elizabeth Milnes, and Kim Strampel. I would like to thank my peer cohorts at IDSS, who have shared a lot of discussions, gossips, laughters, and memories with me: Paolo, Marie, Pin-Yi, Sarah, Bernardo, Andy, Cate, Chin-Chia, Eaman, Bomin, Yan, Hanwei, Sirui, Minghao, Manon, Arnab, Ian, Yi, Amir, Max, Erin, Manxi, Mengying, Shenbo, Yuanhan, Yunzong, Qi, Leon, Yuan, Peter, and Feng. I am proud to be the first year of doctoral students at IDSS. IDSS is not only my home department, but also my home.

At MIT, I was fortunate to sit in the classroom and learn from a number of leading experts: Professors Alberto Abadie, Dimitri Bertsimas, Guy Bresler, Victor Chernozhukov, Glenn Ellison, David Gamarnik, Michel Goemans, Negin Golrezaei, David Karger, Peter Kempthorne, Mihalis Markakis, Whitney Newey, Bart van Parys, Parag Pathak, Yury Polyanskiy, Sasha Rakhlin, Philippe Rigollet, Robert Townsend, John Tsitsiklis, Michael Whinston, Muhamet Yildiz, and Karen Zheng. I have also learned by closely following the books by Professors Joshua Angrist, Dimitri Bertsekas, Dimitri Bertsimas, Rick Durrett, Guido Imbens, Erol Pekoz, Jorn-Steffen Pischke, Sheldon Ross, Donald Rubin, Alexander Schrijver, and John Tsitsiklis. I would love to thank all of them for delivering the knowledge to me.

During my doctoral study, I grew from working with a number of talented researchers in both academia and industry, including my collaborators (Iavor Bojinov, Will Ma, David Simchi-Levi, and Yunzong Xu), AB InBev project team (Felipe Aragao, Tina Gui, Brent Hayden, Benjamin Lee, Hang Li, Hanwei Li, Will Ma), Coppel project team (Otto Alvarado, Jinzhi Bu, José Enrique, Kirby Ledvina, Li Wang, Michelle Wu), Zalando project team (Mehrdad Baghery, Niklas Forck, Silvia Jakob, Emmanuelle Lassiaz, Julius Monello, Gustavo Rocha, Michelle Tran, Chonghuan Wang, Michelle Wu), and Mango project team (David Bofill, Judit Carbo, Michelle Wu), and my labmates (in chronological order: Yehua Wei, Kris Johnson, He Wang, Wang Chi Cheung, Mila Nambiar, Clark Pixton, Zach Owen, Will Ma, Louis Chen, Michael Beeler, Yan Zhao, Xiao Fang, Peter Zhang, Yiqun Hu, Elaheh Fata, Rui Sun, Hanzhang Qin, Li Wang, Hanwei Li, Zhenzhen Yan, Xinshang Wang, Xiaoyue Gong, Ruihao Zhu, Kirby Ledvina, Yunzong Xu, Sabrina Zhai, Jinzhi Bu, Chonghuan Wang, Feng Zhu, Ali Shameli, Prem Talwai, and Menglong Li). I am especially thankful for the support of Janet Kerrigan who has been helping me over the years in many aspects. She has done most of the background work that has shielded me from more stress.

This thesis is supported through generous funding from AB InBev, Coppel, Zalando SE, Mango, and the Accenture-MIT partnership in Business Analytics.

Life at MIT would not have been the same without the accompany of my family and friends, who have tolerated me, have believed in me, and have offered their minds and hands to me without asking for anything in return. I owe my deepest gratitude to them.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Traditionally, the field of revenue management is primarily concerned about the interface with the market with the objective of increasing revenues, in the face of market uncertainty (Talluri and Van Ryzin 2006). Over the past score years, with the aid of the ever-growing computing power, a new type of business model has emerged that places data and algorithms at the center of its operations, in the hope of quantifying the uncertainty and making proper decisions under uncertainty. Many traditional firms have recognized its value and have begun their own digital transformation.

Urged by such digital transformation, this dissertation studies various problems around the following two pillars of data-driven operations: optimization and econometrics. This dissertation draws techniques from these two fields to showcase the possibility of increasing business revenues. The two modules of this thesis are dedicated to one topic each.

In the first module we focus on optimization, in which we consider dynamic resource allocation problems under zero adaptivity. Dynamic resource allocation problems are omnipresent in modern business operations. In the revenue management setting, there are unreplenishable resources to allocate to heterogeneous consumer demands, immediately and irrevocably upon their arrivals. In such settings, zero adaptivity refers to a policy whose actions are independent of the remaining resources. Traditional revenue management literature has mainly focused on fully adaptive policies; and there is a gap between the provable effectiveness of adaptive policies in theory, and the applicability of non-adaptive policies in practice. We show that under different models of demand uncertainty, carefully designed non-adaptive policies may provably perform almost as well as the best fully adaptive counterparts.

In the second module we focus on econometrics, in which we consider experimental design problems. Experimental design is a widely adopted approach for firms to evaluate the effectiveness of their initiatives, by comparing the standard offering to a new initiative. Such a task is often challenging due to interference, both over time and across units. Traditional experimental design methods suffer from large variances of the estimators when accounting for interference; and practitioners have recognized that insufficient precision may lead to unreliable inference. We build the theoretical foundations to use optimization approach to maximize the precision when designing experiments.

Finally, we conclude with discussions of the limitations of the models and methods we have considered. We also provide practical suggestions to applied researchers and data scientists.

## 1.1 First Module: Dynamic Resource Allocation under Zero Adaptivity

This module consists of Chapters 2– 4.

Chapter 2 is motivated by the operational problem in a large consumer packaged goods (CPG) company. While the company appreciates the advantages of dynamic pricing, they deem it operationally much easier to plan out a static price calendar in advance. In this Chapter, we investigate the efficacy of static control policies for revenue management problems whose optimal solution is inherently dynamic. In these problems, a firm has limited inventory to sell over a finite time horizon, over which heterogeneous customers stochastically arrive. We consider both pricing and assortment controls, and derive simple static policies in the form of a price calendar or a planned sequence of assortments, respectively. In the assortment planning problem, we also differentiate between the static vs. dynamic substitution models of customer demand. We show that our policies are within 1-1/e (approximately 0.63) of the optimum under stationary (IID) demand, and 1/2 of the optimum under non-stationary demand, with both guarantees approaching 1 if the starting inventories are large. We adapt the technique of prophet inequalities from optimal stopping theory to pricing and assortment problems, and our results are relative to the linear programming relaxation. Under the special case of IID single-item pricing, our results improve the understanding of irregular and discrete demand curves, by showing that a static calendar can be $(1-1/e)$-approximate if the prices are sorted high-to-low. Finally, we demonstrate on both data from the CPG company and synthetic data from the literature that our simple price and assortment calendars are effective.

Chapter 3 focuses on the network revenue management problem and stochastic packing problem under zero adaptivity, which extends the dynamic pricing problem and joint assortment-and-pricing problem discussed in Chapter 2. In these problems, a firm operates by selling a number of products, which consume limited inventory, over a finite time horizon, over which homogeneous customers stochastically arrive. We specifically focus on the revenue loss incurred due to the constraint of how many changes in actions are allowed. We show that the techniques from Chapter 2 would lead to a small revenue loss that scales sublinearly with the scale of the system, using no more than a certain number of changes in actions. We also show that, such a threshold of total number of changes in actions is critical, in the sense that if we cannot make enough changes then a linear revenue loss is inevitable.

Chapter 4 focuses on an online knapsack problem where the items arrive sequentially and must be either immediately packed into the knapsack or irrevocably discarded. Each item has a different size and the objective is to maximize the total size of items packed. We focus on the class of randomized algorithms which initially draw a threshold from some distribution, and then pack every fitting item whose size is at least that threshold. Threshold policies satisfy many desiderata including simplicity, fairness, and incentive-alignment. We derive two optimal threshold distributions, the first of which implies a competitive ratio of 0.432 relative to the optimal offline packing, and the second of which implies a competitive ratio of 0.428 relative to the optimal fractional packing. We also consider the generalization to multiple knapsacks, where an arriving item has a different size in each knapsack and must be placed in at most one. We derive a randomized threshold algorithm for this problem which is 0.214-competitive. We also show that any randomized algorithm for this problem cannot be more than 0.461-competitive, providing the first upper bound strictly less than 0.5. This online knapsack problem finds applications in many areas, like supply chain ordering, online advertising, and healthcare scheduling, refugee integration, and crowdsourcing. We show how our optimal threshold distributions can be naturally implemented in the warehouses for a Latin American chain department store. We run simulations on their large-scale order data, which demonstrate the efficacy of our proposed algorithms.

## 1.2 Second Module: Experimental Design under Interference and Heterogeneity

This module consists of Chapter 5, which focuses on the design and analysis of switchback experiments. Switchback experiments, where a firm sequentially exposes an experimental unit to a random treatment, are among the most prevalent design used in the technology sector, with applications ranging from ride-hailing platforms to online marketplaces. Although practitioners have widely adopted this technique, the derivation of the optimal design has been elusive, hindering practitioners from drawing valid causal conclusions with enough statistical power. We address this limitation by deriving the optimal design of switchback experiments under a range of different assumptions on the order of the carryover effect —- the length of time a treatment persists in impacting the outcome. We cast the optimal experimental design problem as a minimax discrete optimization problem, identify the worst-case adversarial strategy, establish structural results, and solve the reduced problem via a continuous relaxation. For switchback experiments conducted under the optimal design, we provide two approaches for performing inference. The first provides exact randomization based $p$-values, and the second uses a new finite population central limit theorem to conduct conservative hypothesis tests and build confidence intervals. We further provide theoretical results when the order of the carryover effect is misspecified and provide a data-driven procedure to identify the order of the carryover effect. We conduct extensive simulations to study the empirical properties of our results and conclude with practical suggestions.

# Chapter 2

# Dynamic Pricing under a Static Calendar

## 2.1 Introduction

We consider the following general revenue management problem. A firm has finite inventory of multiple items to sell over a finite time horizon. The starting inventory is unreplenishable and exogenously given, having been determined by supply chain constraints or a higher-level managerial decision. The firm can control its sales through sequential decisions in the form of accepting/rejecting customer requests, pricing, or adjusting the assortment of items offered. Its objective is to maximize the cumulative revenue earned before the time horizon or inventory runs out.

We consider the setting in which customer demand is *distributionally-known* and *independent* over the time horizon; this can be estimated from, e.g., the historical sales data of our partner consumer packaged goods (CPG) company. The literature has also considered other settings, where an unknown IID demand distribution (Besbes and Zeevi 2009, 2012, Agrawal et al. 2017) or an evolving demand process correlated across time (Araman and Caldentey 2009, Ciocan and Farias 2012, Ahn et al. 2019) must be dynamically learned, or where demand is adversarial (Ball and Queyranne 2009, Eren and Maglaras 2010). In our setting, the firm's decision at one point in time has no impact on its estimate of the demand at another point in time, which is supported by our data (see Section 2.1.4 and Section 2.4.1 for further discussion). Instead, the time periods are linked by the inventory constraints, and the firm must trade off between *revenue-centric* decisions that maximize expected revenue irrespective of inventory consumption and *inventory-centric* decisions that maximize the yield from the remaining inventory.

Revenue-centric decisions tend to be myopic and maximize the sales volumes of the most popular items, while inventory-centric decisions tend to be conservative and charge higher prices or prioritize selling highly stocked items. Intuitively, the optimal control policy would make revenue-centric decisions when the overall remaining inventory is plentiful for the remaining time horizon, and inventory-centric decisions when the overall remaining inventory is scarce relative to the remaining time horizon.

However, not all companies have the infrastructure to query the state of the inventory in real-time or adjust their decisions instantaneously. In fact, in the case of our partner CPG company, prices must be negotiated with the brick-and-mortar retailers that sell their products. As a result, a price calendar for the year is planned in advance. On the one hand, this allows the CPG company's management to estimate its promotional budget, make production plans, and coordinate logistics; on the other hand, this allows the retailer to make advertisements, estimate marketing budgets, and lay out shelf space and price labels accordingly.

Motivated by this problem, we analyze the performance of *static* policies, which must plan out all of the firm's decisions (in this case, the price for each week) at the start of the time horizon (in this case, one year), in revenue management problems that are intrinsically *dynamic*, where the optimal control would adapt based on the inventory that remains for the time horizon. If an item's inventory runs out before the end of the time horizon, then its shelf/catalog price is still marked according to the calendar, but no sales of that item can be realized, since its shelf at the brick-and-mortar retailer would be empty. We show that our static policies are effective on data provided by the CPG company. They are also structurally very simple and have performance guarantees comparable to their dynamic counterparts.

### 2.1.1 Models Considered

We consider the time horizon to consist of a discrete number of time periods. This does not lose generality, since a continuous time horizon can be modeled by the limiting case in which the time periods are arbitrarily granular. Similarly, we model each item as having discrete "price points" at which it could be sold. This allows us to both approximate a continuous price range and capture situations where fixed price points have been predetermined by market standards. As is common for many retailers, our CPG company typically chooses prices that end in \$.99 (e.g. \$15.99, \$16.99, \$17.99, \$19.99). Due to menu costs (Mankiw 1985, Stamatopoulos et al. 2017), such a price ladder is rarely changed.

We will separately consider the following two demand models because the design of effective policies differs significantly between them.

1. **Stationary** (Section 2.2.3): the demand distribution for a specific decision, e.g. the purchase probability $q_t(p)$ associated with price $p$, is identical for all time periods $t$.

2. **Non-stationary** (Sections 2.2.4 and 2.2.5): the demand distribution for any decision can vary arbitrarily over time (but still independent across $t$).

We will also consider two types of decisions made by the firm.

1. **Pricing (for a Single Item)**: There is a single item with a discrete starting inventory. We are given, for each time period $t$ and each feasible price $p$, the probability $q_t(p)$ of earning a sale if price $p$ is offered during period $t$. The goal is to plan the price to offer during each period $t$, with no sales occurring if inventory has stocked out.

2. **Assortment (and Pricing)**: There are multiple items each with a discrete starting inventory. We are given, for each time period $t$ and each assortment $S$ of items that could be offered (as well as corresponding prices), the probability of selling each item in $S$ during period $t$. The goal is to plan out the assortment

of items (and prices) to offer during each period $t$, with no sales occurring if the customer chooses an item that has stocked out.

If the assortment problem includes pricing, then it captures the pricing problem with a single item.

Our results also generalize to the **fractional-demand** setting, where the demand distribution given for each period $t$ and price $p$ is over the continuous interval [0,1] (after normalizing), and the sales in the period equal the minimum of the realized demand and remaining inventory. The generalization to $[0, 1]$-demand gives us considerable modeling power. The dynamic pricing literature (e.g. see Gallego and Van Ryzin (1994), Talluri and Van Ryzin (2006), den Boer (2015), Bitran and Caldentey (2003), Elmaghraby and Keskinocak (2003)) has focused on the case of Bernoulli demand because the firm can control the price with arbitrary granularity and hence ensure that at most one sale occurs during any "time period". However, in the case of our CPG company, they can only control prices at the week level, during which the demand distribution can range anywhere from a few hundred to a few thousand units. We will apply the generalization to [0,1]-demand on the data provided by the CPG company in Section 2.4.1.

### 2.1.2 Differences between Our Static Policies and Existing Policies

In this paper, we use the term "static" to describe a policy that prescribes a *deterministic* pricing and/or assortment decision for each period $t$ at the very start of the time horizon. Therefore, the decisions of the static policy must be independent of the sales that end up being realized. Should an item that the calendar planned to offer be out of stock, we distinguish between two models for how customers behave.

1. **Static Substitution**: customers still see the same marked prices (and assortments), but if a customer would have chosen an out-of-stock item, then no sales are realized.

2. **Dynamic Substitution**: customers only see the calendar-planned items with remaining inventory at the time. They never choose an out-of-stock item, and may or may not substitute to another in-stock item.

This distinction is irrelevant for the dynamic policies previously studied in the literature, since they can be changed on-the-fly to never offer an out-of-stock item.

Our static policies are based on deterministic linear programs (see Section 2.2.1 for details), which can be formulated for a given problem instance (items, inventory, prices, and demand distributions) in advance, and hence be used to derive static policies. At a high level, the LPs use deterministic values to approximate the random execution of a policy, and we can use its optimal solution as a "guide" in designing actual policies.

Such an LP was first used for the single-item pricing problem under stationary demand in Gallego and Van Ryzin (1994), who show that the LP will suggest a single price to offer, and hence a static policy. A recent paper by Chen et al. (2018) also proposes a similar single-price policy in the face of strategic customers, that achieves the same $1-1/e$ guarantee. However, this single-price policy requires the critical assumption that the demand, as a function over a continuous price range, is *regular*[1]. In the general setting with irregular demand or a demand function over discrete price points, the LP will suggest two prices, in which case Gallego and Van Ryzin (1994) develop a policy that adaptively switches between them.

By contrast, we show that it is always better to switch from the higher suggested price to the lower suggested price, and furthermore, we show that a static *switching point* can be computed in advance based on the LP. The original dynamic pricing policy of Gallego and Van Ryzin (1994) allows the two prices to be offered in either order, but we show that if the policy must be static, then only the high-to-low ordering of prices is effective.

[1]The regularity assumptions (Assumptions 7.1, 7.2 and 7.3 in Talluri and Van Ryzin (2006)) require the demand function (as a function of price) to be strictly decreasing and continuously differentiable and require the revenue function (as a function of demand) to be concave. However, in practice, demand functions are usually not regular. See Figure 5.2 from Talluri and Van Ryzin (2006).

Moving to non-stationary demand, we can no longer directly follow the LP solution. In fact, we may want to modify certain decisions suggested by the LP to ensure that sufficient inventory is "reserved" for higher-revenue time periods (see Example 2.1 in Section 2.2.4). To accomplish this, we introduce a bid price $c_i$ for each item $i$, which can be interpreted as the opportunity cost of a unit of item $i$'s inventory. Policies based on bid prices are common in revenue management, and bid prices which vary with the time $t$ can be derived using the approximate dynamic programming techniques in Adelman (2007), Rusmevichientong et al. (2020). By contrast, our bid prices $c_i$ are time-invariant and reflect the aggregate value of item $i$ over the non-stationary time horizon. This may be easier for managers to interpret, and also shows managers that an aggregate forecast of demand over the time horizon is sufficient for determining effective bid prices, if we translate those bid prices into a policy appropriately.

Our static policy is to take the LP solution, remove from the suggested assortments all instances where an item $i$ is offered at a price less than $c_i$, and then follow a derandomized version of the modified solution. We essentially treat $c_i$ as an acceptance threshold. Our policy is similar to those of Wang et al. (2015), Gallego et al. (2016), in that it imitates the LP solution and independently determines for each item $i$ when to discard it from the assortment. However, our discarding rule is static and based on our fixed time-invariant bid prices $c_i$, whereas their discarding rule is dynamic and based on the realized inventory levels.

### 2.1.3 Performance Guarantees and Analytical Techniques

We establish performance guarantees for our static policies which, in many cases, improve existing guarantees even for dynamic policies. All of our guarantees are ratios relative to the optimal LP objective value, which is an upper bound on the performance of any static or dynamic policy. Generally, these LPs are useful because they portray a relaxation of the optimal policy, and hence an optimal LP solution can be used as a "guide" in designing a policy for the corresponding problem. In this paper, we will focus on converting the LP solution into a static policy.

Table 2.1: Lower bounds on the performance of static and dynamic policies.

| | Dynamic Policies | **Static Policies** |
|---|---|---|
| Stationary Demand | $1-1/e \xrightarrow[b\to\infty]{\text{w/ error rate } O(1/\sqrt{b})} 1$ | $1-1/e \xrightarrow[b\to\infty]{\text{w/ error rate } O(1/\sqrt{b})} 1$ |
| Single-item Pricing/Assignment | [Gallego and Van Ryzin (1994)] | [**Theorem 2.2**] |
| Assortment (and Pricing) | [Liu and Van Ryzin (2008); **Theorem 2.2**] | |
| Non-stationary Demand | 1/2 | 1/2 |
| Single-item Pricing/Assignment | [Wang et al. (2015)] | [Rusmevichientong et al. (2020); **Theorem 2.6**] |
| Assortment (and Pricing) | [Gallego et al. (2016)] | |
| Non-stationary Demand | $\xrightarrow[b\to\infty]{\text{w/ error rate } O(1/\sqrt{b})} 1$ | $\xrightarrow[b\to\infty]{\text{w/ error rate } O(\sqrt{\log b/b})} 1$ |
| Single-item Pricing/Assignment | [Wang et al. (2015)] | [Hajiaghayi et al. (2007); **Theorem 2.8**] |
| Assortment (and Pricing) | [Gallego et al. (2016)] | [**Theorem 2.8**] |

Note: Our new results are bolded. $b$ refers to the amount of starting inventory (or the smallest starting inventory, if there are multiple items).

Our results are outlined in Table 2.1. The baseline performance ratio is $1-1/e$ for stationary demand and $1/2$ for non-stationary demand. That is, our static policies always earn at least 50% of the optimum in expectation, with the ratio improving to $\approx 63.2\%$ if the given demand distributions are stationary. Both of these ratios are tight. The ratios also increase to 100% as $b$, the starting inventory level when demand has been normalized to lie in [0,1] (or in the assortment setting, the minimum starting inventory among the items), increases to $\infty$.

In the stationary-demand pricing problem, Gallego and Van Ryzin (1994) derived both the lower bound of $1-1/e$ and an asymptotic-optimality result. However, their policy is in general dynamic, unless the demand function is *regular* over a continuous interval – the concavity assumption allows for a single price in the LP. By contrast, we show that the same results can be obtained using our high-to-low static policy, regardless of demand regularity. Additionally, in our analysis, we derive the tightest possible bound for every value of $b$ and $T$ (the number of time periods), which allows us to establish asymptotic optimality in only $b$ (instead of scaling both $T\to\infty$ and $b\to\infty$).

In the stationary-demand assortment problem, we analyze the policies originally proposed by Liu and Van Ryzin (2008) and obtain the same bounds as above that are tight in both $b$ and $T$. To our knowledge, this type of result, which includes the baseline lower bound of $1-1/e$ when starting inventory is 1, has been previously unknown[2] for the assortment problem. Asymptotic optimality was previously derived by Liu and Van Ryzin (2008) when both $T \to \infty$ and $b \to \infty$.

Moving to non-stationary demand, the lower bound of 1/2 which improves to 1 as $b \to \infty$ has been previously established using dynamic policies, in the assignment problem of Wang et al. (2015) and the more general assortment problem of Gallego et al. (2016). We establish the same bounds using static policies, with an extremely simple analysis based on *prophet inequalities* from optimal stopping theory. However, our convergence rate of $1 - O(\sqrt{(\log b)/b})$ is worse than the rate of $1 - O(1/\sqrt{b})$ achievable with their dynamic policies.

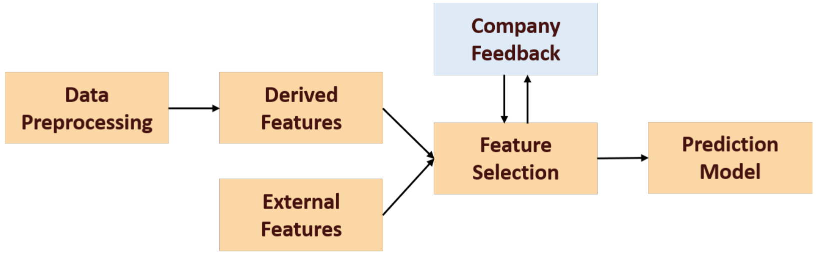
We should mention that the lower bound of 1/2 for static policies under non-stationary demand was also recently established by Rusmevichientong et al. (2020). Their bound and analysis differ from ours in that theirs are relative to the optimal dynamic policy instead of the deterministic LP relaxation. One benefit of using the LP is that it directly extends to the fractional-demand setting, since the LP does not change when demand can take any value in [0,1], which is our application of interest with the CPG company. By contrast, their framework is designed for a very general setting where resources can be reused after a random amount of time. We numerically compare the performance of their policy in Section 2.4.2.

### 2.1.4 Application on Data from CPG Company

We use aggregated weekly sales data from a CPG company to validate our model, and test the performance of our proposed policies. We use random forest to build prediction models that suggest demand distributions (normalized to lie in [0,1], possibly fractional numbers) under different prices. Then we take these distributions as inputs, and numerically compare the performance of our policies to some basic benchmarks.

[2] The results in Golrezaei et al. (2014) imply performance guarantees for our problem, but their ratios are smaller than ours, since they are designed to hold under the more general setting of *adversarial* demand. Under this demand model, they only obtain a $(1 - 1/e)$-guarantee under the additional assumptions that each item has a single price, and that starting inventories are asymptotically large.

Figure 2-1: Work flow: from data to prediction model




Working together with the CPG company, we used the work flow depicted in Figure 2-1 to build our demand model. The average out-of-sample percent error in its sales predictions is 19.41%. It is worth highlighting the features selected by the random forest: the tagged price, external competitor prices, and some external features such as seasonality. However, neither internal competitor prices (the prices of other SKUs of the CPG company) nor historical prices were selected. This observation validates our model in the following two aspects: internal competitor prices not being selected suggests that we can separately optimize the price calendar for each item; historical prices not being selected suggests that demand can be modeled as independent over time. The latter aspect is also validated by a stream of empirical literature on the "pantry effect" (Ailawadi and Neslin 1998, Bell et al. 1999), which observes for various consumable goods that if customers attempt to stockpile it when the price is low, then they will untimately consume it more quickly; hence, the low price did not necessarily cannibalize future demand.

Optimizing the price calendar based on our demand model, we find that for scenarios where the starting inventory is of moderate size compared to the total expected demand (i.e., for SKUs that were initially neither overstocked nor understocked), our static policies outperform basic LP-based static policies by 5% under stationarity and 1% under non-stationarity. Furthermore, our static policies lose at most 1% under stationarity and 4% under non-stationarity, compared to the optimal dynamic policies.

Both our theoretical guarantees and computational experiments suggest that static calendars perform nearly as well as their dynamic counterparts. Managers will not lose much from planning a sequence of prices / assortments in advance.

Further details about our demand modeling and calendar optimization with data from the CPG company can be found in Sections 2.4.1 and A.1.

### 2.1.5 Related Work

Our 1/2 guarantee for the general assortment and pricing problem under non-stationary demand (and small inventory) is motivated by *prophet inequalities*, which provide an elegant method for bounding the performance of online vs. offline algorithms (see Samuel-Cahn et al. (1984), Kleinberg and Weinberg (2012)). The basic idea is to compute a threshold price for each item, based on the offline solution, such that either we are satisfied if an item sells out at its threshold price; or, if it does not, we are still satisfied from having had the opportunity to offer that item to every customer. Using prophet inequalities, there are very general results for maximizing welfare in online combinatorial auction settings (Feldman et al. 2014, Dütting et al. 2017).

However, to the best of our knowledge, it is important for these techniques that the objective is *welfare*, where the decision maker earns a reward equal to the sum of *revenue* and *customer surplus* generated. When the objective is revenue alone, elegant connections have been made for the single-parameter domain (Chawla et al. 2010, Correa et al. 2019), which hold under multiple settings including different arrival orderings. Our work is the first to make the connection to revenue maximization for assortment optimization under substitutable choice models, which is essentially a multi-parameter domain. The result is a simple and elegant 1/2 guarantee for assortment and pricing relative to the well-studied linear programming based upper bound, which holds without any assumptions on inventories being large or customers having identical willingness-to-pay distribution. The fact that our thresholds and guarantee are relative to the upper bound is also novel. The aforementioned literature has focused on comparing against the expected value of a "prophet" who knows the realized valuations in advance (a weaker benchmark than the LP).

We also compare our model in which demand is stochastic and distributionally known, to the models in which demand is completely unknown, or adversarial. One of the benefits of the adversarial model is that it does not rely on correct "forecasts" of demand over time. However, the drawback is that the resulting algorithm does not make use of forecasted demand information. Starting with the online booking problem of Ball and Queyranne (2009) in which the decision concerns whether to accept or reject each customer, single-item pricing (Eren and Maglaras 2010, Ma et al. 2018), assortment optimization (Golrezaei et al. 2014), and joint assortment and pricing (Ma and Simchi-Levi 2017) have all been studied under the adversarial demand model.

The guarantees relative to the optimum are worse than ours, because we have more demand information—in fact, in all of the settings except Golrezaei et al. (2014), the papers resort to *instance-dependent* competitive ratios because a universal non-zero guarantee is impossible when inventories can be depleted at multiple potential prices. These papers have also tested the empirical performance of their algorithms on datasets and found that using a *hybrid* strategy (as proposed by Mahdian et al. (2007)) which employs the algorithms from both the adversarial and stochastic demand models performs best. This suggests that our algorithmic improvements in the stochastic demand model are valuable even if the given demand distributions are not 100% correct.

### 2.1.6 Outline

In Section 2.2.1 we define our basic problems and state the assumptions. In Section 2.2.2 we discuss some generalizations and their required assumptions. In Sections 2.2.3 – 2.2.5, we introduce randomized policies in stationary demand, non-stationary demand, and non-stationary demand with large inventory, respectively. Then, in Sections 2.3.1 and 2.3.2, we introduce general sampling-based de-randomization methods, and structural de-randomization methods, respectively. These de-randomization methods yield deterministic calendars that (i) have the same theoretical guarantees, (ii) significantly improve computational performance, and (iii) are much easier for companies to accept. Finally, in Sections 2.4.1 and 2.4.2, we conduct numerical experiments using real data provided by the CPG company, and synthetic data from the literature. We also introduce how we estimate the demands from the data in Section 2.4.1.

## 2.2 Problem Definitions and Performance Guarantees via Randomized Static Policies

### 2.2.1 Problem Definitions

Let $\mathbb{N}$ and $\mathbb{N}_0$ denote the positive and non-negative integers, respectively. For any positive integer $n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$.

A firm has $n \in \mathbb{N}$ items to sell over a finite time horizon of $T \in \mathbb{N}$ time periods. Each item $i$ is endowed with $b_i \in \mathbb{N}$ units of starting inventory, which is unreplenishable. We assume that $b_i \leq T$, which does not lose generality since at most one unit of any item $i$ can be sold during any time period. Let $\underline{b}$ denote $\min_{i \in [n]} b_i$.

The firm can offer each item at one of $m \in \mathbb{N}$ prices, $p_1, \ldots, p_m$, which are positive real numbers. We will refer to each item-price combination $(i, j) \in [n] \times [m]$ as a *product*, in which case the general assortment and pricing problem can be described as offering a set of products to each customer.

We let $\mathcal{S}$ be any *downward-closed*[3] family, which can be used to capture both physical constraints such as shelf-size limitations and business constraints whereby certain items cannot be offered at certain prices (or the same item cannot be simultaneously offered at multiple prices in the form of different products). We allow for a constraint on the sets of products that can be feasibly offered, imposing that they must lie in some family $\mathcal{S}$ of subsets of $\{(i, j) : i \in [n], j \in [m]\}$. We will refer to elements in $\mathcal{S}$ as *assortments*.

A *static policy* is a calendar that must be fixed at the start, prescribing the assortments $S_1, \ldots, S_T \in \mathcal{S}$ to offer over the time horizon. In this section, we allow this calendar to be determined in a *random* fashion at the start. After this calendar has been fixed, sequentially over time $t = 1, \ldots, T$, customer $t$ arrives and chooses to

[3] A family of subsets $\mathcal{S}$ is downward-closed if for any $S \in \mathcal{S}$ and $S' \subseteq S$, we also have $S' \in \mathcal{S}$.

purchase at most one product from assortment $S_t$. For the situation in which some of the products $(i, j)$ in the planned assortment $S_t$ have had their items $i$ stock out before time $t$, we distinguish between two models for how customers choose:

1. Customer $t$ always sees all of the products in $S_t$, and if her first choice from $S_t$ has stocked out, no sales are realized (*static substitution*);

2. Customer $t$ only sees the products in $S_t$ that are still in stock[4], and chooses her favorite product from this selection (*dynamic substitution*).

We note that *dynamic policies*, as traditionally studied in the literature, do not need to distinguish between static vs. dynamic substitution, since they can decide the assortments $S_t$ on-the-fly to never offer an out-of-stock item (Rusmevichientong et al. 2020). By contrast, for static policies, both models can be justified. Static substitution occurs in parking systems, where customers are often shown "phantom" parking spots, only to drive there and discover that the spot is occupied (Owen and Simchi-Levi 2017). On the other hand, dynamic substitution occurs if customers switch to a different product when they see that the shelf for their favorite product is empty (Anupindi et al. 1998, Mahajan and Van Ryzin 2001, Honhon et al. 2010, Goyal et al. 2016). One factor that possibly reduces the prevalence of dynamic substitution is brand loyalty, under which customers commit to a favorite brand at the supermarket and make no purchase if that brand has stocked out (Jacoby and Kyner 1973, Amine 1998, Roehm et al. 2002). For product categories where brand loyalty is less common, dynamic substitution is more common.

In either case, for all $t \in [T]$, $S \in \mathcal{S}$, and $(i, j) \in S$, we let $q_t(i, j, S)$ be the probability of customer $t$ choosing product $(i, j)$ when she *sees* assortment $S$. Note that $\sum_{(i,j)\in S} q_t(i, j, S) \leq 1$ for all $t$ and $S$, where $1 - \sum_{(i,j)\in S} q_t(i, j, S)$ denotes the probability of customer $t$ purchasing nothing when she sees assortment $S$. If $(i, j) \notin S$, then $q_t(i, j, S) = 0$. If the choice probabilities $q_t(i, j, S)$ are equal across $t = 1, \ldots, T$, for all assortments $S$ and $(i, j) \in S$, then we say that demand is *stationary*. If so, we omit the subscript $t$ and refer to the choice probabilities as $q(i, j, S)$.

[4]Since $\mathcal{S}$ is downward-closed, it is guaranteed that this selection of products seen by the customer is still feasible. Furthermore, $\emptyset \in \mathcal{S}$, the empty set is always allowable.

### Choice-based deterministic linear programs

Any (static or dynamic) policy for the assortment problem can be captured by the following LP: let $x_t(S), \forall t \in [T], \forall S \in \mathcal{S}$ represent the probability of offering assortment $S$ at time $t$.

$$J^{CDLP-N} = \max \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} x_t(S) \sum_{(i,j) \in S} p_j q_t(i,j,S) \tag{2.1}$$

$$\text{s.t.} \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} x_t(S) \sum_{j:(i,j) \in S} q_t(i,j,S) \leq b_i \quad \forall\ i = 1, \ldots, n \tag{2.2}$$

$$\sum_{S \in \mathcal{S}} x_t(S) = 1 \quad \forall\ t = 1, \ldots, T \tag{2.3}$$

$$x_t(S) \geq 0 \quad \forall\ t = 1, \ldots, T;\ \forall\ S \in \mathcal{S} \tag{2.4}$$

Constraints (2.2) ensure that total units sold will not exceed the initial inventory, in expectation; constraints (2.3) ensure that only one price can be chosen in each time period. Note that we can assume equality in constraint (2.3) because $\mathcal{S}$ (being downward-closed) always contains the empty assortment $\emptyset$; hence, we can increase $x_t(\emptyset)$ until equality is achieved.

When demand is stationary, we let $x(S), \forall S \in \mathcal{S}$ represent the probability of offering assortment $S$ at any given time period. Constraints (2.3) are then equivalent to the single constraint (2.7).

$$J^{CDLP-S} = \max\ T \cdot \sum_{S \in \mathcal{S}} x(S) \sum_{(i,j) \in S} p_j q(i,j,S) \tag{2.5}$$

$$\text{s.t. } T \cdot \sum_{S \in \mathcal{S}} x(S) \sum_{j:(i,j) \in S} q(i,j,S) \leq b_i \qquad \forall\ i = 1, \ldots, n \tag{2.6}$$

$$\sum_{S \in \mathcal{S}} x(S) = 1 \tag{2.7}$$

$$x(S) \geq 0 \qquad \forall\ S \in \mathcal{S} \tag{2.8}$$

We derive performance guarantees for our policies, which are based on the deterministic LPs, relative to the optimal objective values of those LPs. This also provides a performance guarantee relative to the revenue of any dynamic policy, which is upper-bounded by the LP objective value—this is a well-known type of result in revenue management.

**Lemma 2.1** (Gallego and Van Ryzin (1994), Gallego et al. (2004))**.** *The expected revenue of any (static or dynamic) policy for the assortment problem is upper-bounded by the optimal objective value of CDLP-N from* (2.1) *when demand is non-stationary, and CDLP-S from* (2.5) *when demand is stationary. Analogously, the expected revenue of any policy for the single-item pricing problem is upper-bounded by the optimal objective value of DLP-S from* (2.14) *in Section 2.3.2.*

Hereinafter, we will always use the LP objective value as our optimum and denote it using $\mathsf{OPT}_{\mathrm{LP}}$, where the distinction between the LPs will be clear from the context.

The following mild assumption is required for some of our results. It originated from Golrezaei et al. (2014) and has been nearly omnipresent in the subsequent literature on inventory-constrained assortment optimization (Gallego et al. 2016, Chen et al. 2016, Ma and Simchi-Levi 2017, Rusmevichientong et al. 2020, Ma et al. 2020b, Cheung et al. 2018).

**Assumption 2.1.** *For all $t \in [T]$, $S \in \mathcal{S}$, $S' \subseteq S$, and $(i,j) \in S'$, we have $q_t(i,j,S') \geq q_t(i,j,S)$.*

This assumption states that the probability of selling a product $(i,j)$ can only be improved if it is offered as part of a smaller assortment $S'$ instead of a larger assortment $S$. The condition on the choice probabilities is often called *substitutability.* It is implied by any random-utility choice model (e.g., the multinomial logit choice model used in our computational study in Section 2.4.2), which treats the products as *substitutes.*

### 2.2.2 Generalized Model with Multi-Consumption and Fractional Consumption

Before stating our results, we describe a generalized version of the model from Section 2.2.1 that allows for multiple products, as well as fractional amounts of a product, to be consumed in the same time period. This generalization is natural if we interpret each $t \in [T]$ as a larger-scale epoch (e.g., one week in the case of our CPG company) instead of the choice made by a single customer.

**Multiple purchases**

We first consider the generalization where a customer can choose multiple products (but still demands exactly 1 unit of each product chosen). This requires the following modifications to the model from Section 2.2.1.

We are now given the joint distribution for the *set* of products demanded, when any assortment $S \in \mathcal{S}$ is seen by any customer $t$. We assume that the set demanded never contains two different products $(i, j_1), (i, j_2)$ corresponding to the same item $i$, so that even when there is only one unit of $i$ in stock, there is no ambiguity about whether $(i, j_1)$ or $(i, j_2)$ is purchased.

Note that this assumption can be enforced by restricting $\mathcal{S}$ to not contain any assortment that simultaneously offers different products $(i, j_1), (i, j_2)$ corresponding to the same $i$. Such a restriction is natural if the only difference between products $(i, j_1)$ and $(i, j_2)$ is in price ($p_{j_1}$ vs. $p_{j_2}$), in which case it is nonsensical to mark item $i$ at multiple prices.

The demand is said to be stationary if the distribution of the set of products demanded is identical across $t = 1, \ldots, T$ for all assortments $S \in \mathcal{S}$. The definitions of Assumption 2.1 and the CDLPs remain unchanged if $q_t(i, j, S)$ now represents the *marginal* probability of product $(i, j)$ being demanded when assortment $S$ is seen by customer $t$.

**Fractional demand consumption.**

The further generalization where demand can be fractional (for multiple products) requires the following further modifications to the model.

We are now given the joint distribution for the *quantity* of each product demanded, when any assortment $S \in \mathcal{S}$ is seen by any $t$. We assume that this is a joint distribution over $[0, 1]^{|S|}$, where we have normalized the sales of any item within a time period to lie in $[0, 1]$ (and scaled its prices accordingly). We now allow the starting inventories $b_1, \ldots, b_n$ to be any real numbers that are at least 1. We can assume a lower bound of 1 because any item $i$ with $b_i < 1$ can have its demand scaled up by $1/b_i$ so that the maximum possible sales during a time period is 1.

When demand can be fractional, we assume that the joint distribution of demand only depends on the assortment shown, not the exact quantity of each product available. As a result, *we assume static substitution when demand can be fractional.* Under static substitution, for each $t \in [T]$, $S \in \mathcal{S}$, and $(i, j) \in S$, let $F_{t,(i,j,S)}(\cdot)$ be the CDF function for the quantity of product $(i, j)$ being demanded, should assortment $S$ be offered during time $t$. $F_{t,(i,j,S)}(\cdot)$ is given and known. As before, demand is said to be stationary if the joint demand distribution is identical across time. In this case, $F_{t,(i,j,S)}(\cdot) = F_{(i,j,S)}(\cdot), \forall t \in [T], \forall S \in \mathcal{S}$. The definitions of Assumption 2.1 and the CDLPs remain unchanged if $q_t(i, j, S) = \mathbb{E}_{Q \sim F_{t,(i,j,S)}}[Q]$ now represents the *expected quantity* of product $(i, j)$ demanded when assortment $S$ is shown at time $t$.

We state one final assumption that is required for our result under stationary demand, *only when demand can be fractional.* This assumption is again very mild, in that it automatically holds for {0,1}-demand, which is the case studied in all of the existing literature. We have to add it as a technical assumption in the setting of [0,1] demands. Whereas we do not make any assumptions (e.g. concavity) on the demand distributions themselves, this assumption concerns the relationships between the different CDF's for the different prices.

**Assumption 2.2.** *Let $F_{(i,j,S)}$ denote the marginal CDF for the quantity of product $(i, j)$ demanded when assortment $S$ is shown at time $t$. For all items $i$, feasible assortments $S, S' \in \mathcal{S}$, and prices $j, j'$ with $p_j > p_{j'}$, assume that for all $c \in [0, 1]$,*

$$\frac{\mathbb{E}_{Q \sim F_{(i,j,S)}}[\min\{c, Q\}]}{\mathbb{E}_{Q \sim F_{(i,j,S)}}[Q]} \geq \frac{\mathbb{E}_{Q' \sim F_{(i,j',S')}}[\min\{c, Q'\}]}{\mathbb{E}_{Q' \sim F_{(i,j',S')}}[Q']}.$$

The intuitive explanation of Assumption 2.2 is that for any amount of remaining inventory $c$ for item $i$, the fraction of un-truncated demand sold $\frac{\mathbb{E}_{Q \sim F_{(i,j,S)}}[\min\{c,Q\}]}{\mathbb{E}_{Q \sim F_{(i,j,S)}}[Q]}$ is greater at the higher price $j$. Assumption 2.2 can be seen as a weaker version of a stochastic dominance assumption on the *hazard rates* of the distributions $F_{(i,j,S)}(x), F_{(i,j',S')}(x)$. We provide examples and detailed discussions in Section A.2 in Appendix A.

### 2.2.3 Stationary Demand

#### Statement of results

Our assortment policy probabilistically follows the LP solution, without specifically re-ordering the decisions portrayed in the LP.

---

**Algorithm 1** Assortment (and pricing) policy when demand is stationary

1: Solve CDLP-S, and let $\{x^*(S) : S \in \mathcal{S}\}$ denote an optimal solution.
2: Independently for each time $t$, set the assortment $S_t$ to be $S$ with probability $x^*(S)$, for all $S \in \mathcal{S}$.

---

This policy that probabilistically imitates the LP was originally studied by Gallego et al. (2004), Liu and Van Ryzin (2008), where it was shown to be empirically effective and asymptotically optimal. We now derive the first provable guarantees for it in the non-asymptotic setting, as well as a tight characterization of how the guarantee depends on both $T$ and $\underline{b}$.

**Theorem 2.2.** *Under the static substitution model (with Assumption 2.2 needed if demand is fractional), for the assortment (and pricing) problem under stationary demand, if there are $T$ time periods and $\underline{b} = \min_{i\in[n]} b_i$, then Algorithm 1 earns expected revenue of at least*

$$\frac{\mathbb{E}[\min\{\mathrm{Bin}(T, \underline{b}/T), \underline{b}\}]}{\underline{b}} \cdot \mathsf{OPT}_{\mathrm{LP}}, \tag{2.9}$$

*where* $\mathrm{Bin}(T, \underline{b}/T)$ *denotes a Binomial random variable consisting of $T$ trials of probability $\underline{b}/T$.*

*If we let $\Delta^{APX}$ denote the term $\frac{\mathbb{E}[\min\{\mathrm{Bin}(T,\underline{b}/T),\underline{b}\}]}{\underline{b}}$ from expression (2.9), then*

$$\Delta^{APX} \geq 1 - \frac{\underline{b}^{\underline{b}}}{\underline{b}!} e^{-\underline{b}}, \tag{2.10}$$

*which states that* $\Delta^{APX} = 1 - O(1/\sqrt{\underline{b}})$*, and increases from* $1 - 1/e$ *to 1 as* $\underline{b} \to \infty$ *(regardless of* $T$*).*

In Section 2.2.3 we sketch our proof technique for Theorem 2.2, and in Section 2.2.3 we show that our approximation guarantee of $\Delta^{APX} = \frac{\mathbb{E}[\min\{\text{Bin}(T,\underline{b}/T),\underline{b}\}]}{\underline{b}}$ is tight for every value of $T$ and $\underline{b}$. But first, we demonstrate why Theorem 2.2 does not hold generally under the dynamic substitution model, and identify a special case when it does hold.

**Proposition 2.3.** *Under the dynamic substitution model and Assumption 2.1 (substitutability), there is an instance of the assortment and pricing problem for which the expected revenue of Algorithm 1 is strictly less than* $(1 - 1/e) \cdot \mathsf{OPT}_{\text{LP}}$*.*

The counterexample for Proposition 2.3 is detailed in Section A.7 in Appendix A. Nonetheless, the counterexample requires both multiple prices (i.e., item 1 that can be sold at multiple prices) and multiple items (i.e., a second item that "shields" the first item from being sold at the lower price) to exist. Theorem 2.2 holds in both of the canonical cases of:

1. Single item, multiple prices (because with a single item, dynamic and static substitution are equivalent);

2. Multiple items, single price per item (this is the pure assortment problem without pricing, as stated next in Proposition 2.4).

**Proposition 2.4.** *Under the dynamic substitution model and Assumption 2.1 (substitutability), if each item has only one single price (pure assortment problem without pricing), then Algorithm 1 earns expected revenue of at least* $\Delta^{APX} \cdot \mathsf{OPT}_{\text{LP}}$*.*

We prove Proposition 2.4 and Theorem 2.2 together in Section A.5 in Appendix A. Moreover, for the joint assortment and pricing problem under dynamic substitution, a static calendar can still earn $\frac{1}{2} \cdot \mathsf{OPT}_{\text{LP}}$, even when demand is non-stationary, as we will show via Algorithm 2 in Section 2.2.4.

**Two-step proof sketch of Theorem 2.2**

The proof can be divided into two steps, which we will illustrate using the following example. Consider a problem instance with a single item, time periods $T = 3$ and starting inventory $b = 2$. Suppose we have two prices. The higher price $p_{\mathsf{H}} = 2$ earns a sale with probability 1/3; the lower price $p_{\mathsf{L}} = 1$ earns a sale with probability 1, i.e. deterministically. The optimal LP solution from (2.5) – (2.8) suggests to offer a higher price $\mathsf{H}$ for 1.5 time periods, and a lower price $\mathsf{L}$ for 1.5 time periods.

Let $\mathbb{E}[\mathsf{Rev}(0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L})]$ denote the expected revenue of a randomized policy that offers $\mathsf{H}$ and $\mathsf{L}$ each with probability one half in each period. Suppose, for the purpose of analysis, that there existed a virtual price $p_{\mathsf{C}} = (\mathbb{E}[Q_{\mathsf{H}}]p_{\mathsf{H}} + \mathbb{E}[Q_{\mathsf{L}}]p_{\mathsf{L}})/(\mathbb{E}[Q_{\mathsf{H}}] + \mathbb{E}[Q_{\mathsf{L}}])$ with CDF $F_{\mathsf{C}}(x) = 0.5F_{\mathsf{H}}(x) + 0.5F_{\mathsf{L}}(x), \forall x \in [0, 1]$. Note that $\mathbb{E}[Q_{\mathsf{C}}] = 2/3$. We then establish the following sequence of two inequalities:

$$\mathsf{OPT}_{\mathrm{LP}} \cdot \frac{\mathbb{E}[\min\{\mathrm{Bin}(3, \frac{2}{3}), 2\}]}{2} \leq \mathbb{E}[\mathsf{Rev}(\mathsf{C}; \mathsf{C}; \mathsf{C})] \tag{2.11}$$

$$\leq \mathbb{E}[\mathsf{Rev}(0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L})] \tag{2.12}$$

Inequality (2.11) relates the LP optimum to the expected revenue of a virtual calendar that always offers $p_{\mathsf{C}}$. We interpret the LHS as the expectation of some Binomial random variable truncated by initial inventory and the RHS as the expectation of an identical-mean, smaller-variance random variable that is also truncated by initial inventory. Although this virtual calendar cannot actually be offered (because the price $p_{\mathsf{C}}$ never exists), it can bridge our analysis.

Inequality (2.12) is true under Assumption 2.2. If the demand is never truncated by the amount of remaining inventory, then offering the virtual price $p_{\mathsf{C}}$ is equivalent to randomly choosing prices $p_{\mathsf{H}}$ and $p_{\mathsf{L}}$ each with probability one-half. However, if there is truncation, then Assumption 2.2 guarantees that the revenue from randomly choosing between the real prices $p_{\mathsf{H}}$ and $p_{\mathsf{L}}$ cannot be less.

The formal, general proof of Theorem 2.2 is deferred to Section A.5 in Appendix A.

### Tightness of results

We now show that the ratio produced in expression (2.9), which is dependent on $\underline{b}$ and $T$, is tight. The proof of Proposition 2.5 can be found in Section A.6 in Appendix A.

**Proposition 2.5.** *For any positive integers $T$ and $b$, there exists an instance of the stationary-demand single-item pricing problem with $T$ time periods and $b$ starting inventory, for which the expected revenue of any policy is upper-bounded by expression (2.9).*

## 2.2.4 Non-stationary Demand with Small Inventory

In this section, we present our results for non-stationary demand with small inventory. Our results for non-stationary demand in the asymptotic regime will be discussed in Section 2.2.5.

### Statement of results

In contrast to stationary demand, under the more general setting of non-stationary demand, following the LP solution may be undesirable, because it may be beneficial to "reserve" inventory for the highest-revenue time periods. The following Example 2.1 demonstrates this idea.

**Example 2.1.** Let there be $T = 2$ periods and $b = 1$ unit of initial inventory. Let $\epsilon \in (0, 1)$ be some small positive number. Let there be two prices: $p_1 = 1/\epsilon^2, p_2 = 1$. Let random demands be Bernoulli random variables. During day 1, the purchase probability of offering the higher price $p_1$ is 0; and the purchase probability of offering the lower price $p_2$ is $1 - \epsilon$. During day 2, the purchase probability of offering both prices is $\epsilon$. DLP-N suggests that we offer $p_2$ in the first period, then $p_1$ in the second

| Prices | Period 1 | Period 2 |
|---|---|---|
| $p_1$ | $0$ | $\epsilon$ |
| $p_2$ | $1-\epsilon$ | $\epsilon$ |

period. The objective value of DLP-N is $(1-\epsilon) + \epsilon \cdot \frac{1}{\epsilon^2} = 1 - \epsilon + \frac{1}{\epsilon}$. By simply using the DLP-N solution as a calendar, the expected revenue is $(1-\epsilon)+\epsilon\cdot\epsilon\cdot\frac{1}{\epsilon^2}=2-\epsilon$. We can pick $\epsilon$ to be arbitrarily small and thus directly using LP can be arbitrarily bad. ∎

Nonetheless, we can still use the LP as a guide for our reservation policies.

---

**Algorithm 2** Assortment (and pricing) policy when demand is non-stationary

1: Solve CDLP-N, and let $\{x_t^*(S) : t \in [T], S \in \mathcal{S}\}$ denote an optimal solution.

2: For each item $i$, let $r_i^* = \sum_{t=1}^{T}\sum_{S\in\mathcal{S}} x_t^*(S) \sum_{j:(i,j)\in S} p_j q_t(i,j,S)$ be the contribution from item $i$ to the optimal objective value (note that $\mathsf{OPT}_{\text{LP}} = \sum_{i=1}^{n} r_i^*$).

3: Independently for each time $t$, first randomly select a $\tilde{S}_t$ to be equal to each $S \in \mathcal{S}$ with probability $x_t^*(S)$, which is a proper probability distribution by constraint (2.3). If $\sum_S x_t^*(S) < 1$, then select $\tilde{S}_t$ to be the empty set $\emptyset$ with the remaining probability, where $\emptyset \in \mathcal{S}$ is guaranteed by the downward-closed statement in Assumption 2.1.

4: Define a discarding rule $D : \mathcal{S} \to \mathcal{S}$ to be

$$D(\tilde{S}_t) = \left\{ (i,j) \in \tilde{S}_t \,\middle|\, p_j > \frac{r_i^*}{2b_i} \right\}. \tag{2.13}$$

After $\tilde{S}_t$ has been selected, set the final assortment to be $S_t = D(\tilde{S}_t) \subseteq \tilde{S}_t$, which is a feasible assortment to offer since $\mathcal{S}$ is downward-closed.

---

Our assortment policy under non-stationary demand uses each cost $r_i^*/(2b_i)$ as an acceptance threshold. We remove from the planned assortments all products of item $i$ being offered at prices below their thresholds. It is probable that the final $S_t$ is an empty set $\emptyset$, even if $\tilde{S}_t$ is not empty, because we discard all the products from $\tilde{S}_t$.

**Theorem 2.6.** *Under Assumption 2.1 (substitutability), for the assortment (and pricing) problem where demand may be non-stationary, Algorithm 2 earns expected revenue of at least* $\mathsf{OPT}_{\text{LP}}/2$.

### Proof sketch of Theorem 2.6

By finding the assortment suggested by expression (2.13), each unit of item $i$ sold earns at least one-half of the per-inventory revenue of the corresponding $r_i^*$, which is its contribution to the LP objective. Thus, if inventory runs out during the horizon, then we have earned in total at least one-half of the LP upper bound. If inventory never runs out, then the algorithm extracts the full "opportunity" from each time period which also results in at least one-half of the LP upper bound. In other words, setting one-half of the per-inventory revenue as an acceptance threshold is neither too high nor too low, and results in a "win-win" situation. This argument is based on the classical *prophet inequalities* from Krengel and Sucheston (1977), Samuel-Cahn et al. (1984), where we have modified their argument for optimal stopping to the pricing and assortment settings.

We outline two key steps here, and defer the details of our proof to Section A.8 in Appendix A.

1. To evaluate Algorithm 2, we take out $\frac{r_i^*}{2b_i}$ revenue earned from each period for each product $(i, j)$. Since after the discarding rule, the prices should be no less than the threshold, i.e. $p_j \geq \frac{r_i^*}{2b_i}$. Thus, this difference should always be non-negative. That is,

$$\mathsf{Rev} \geq \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} A_t(S) \sum_{(i,j) \in D(S)} (p_j - \frac{r_i^*}{2b_i}) \min\{B_{t-1}(i), Q_t(i, j, D(S))\} + \sum_{i=1}^{n} \frac{r_i^*}{2b_i}(b_i - B_T(i)),$$

where $A_t(S)$ is an indicator if assortment $\tilde{S}_t = S$ was selected in period $t$, before the discarding rule from Algorithm 2 was applied; the infimum between $B_{t-1}(i)$, the (random) remaining inventory of item $i$ at the end of period $t-1$, and $Q_t(i, j, D(S))$, the (random) quantity of product $(i, j)$ demanded, is the actual inventory of item $i$ sold in period $t$.

2. We relate the first triple summation term to CDLP-N, the deterministic linear program.

$$\sum_{t=1}^{T} \sum_{S \in \mathcal{S}} A_t(S) \sum_{(i,j) \in D(S)} (p_j - \frac{r_i^*}{2b_i}) \min\{B_{t-1}(i), Q_t(i, j, D(S))\} \geq \sum_{i=1}^{n} \frac{\mathbb{E}[B_T(i)]}{b_i} \left( r_i^* - \frac{r_i^*}{2b_i} \cdot b_i \right).$$

Then after canceling and re-arranging terms we prove the desired result.

### Tightness of results

We now show that the ratio in Theorem 2.6 is tight. It suffices to find an instance in the single-item pricing problem to show that the general result of assortments (and pricing) problem is tight. The proof of Proposition 2.7 can be found in Section A.9 in Appendix A.

**Proposition 2.7.** *There exists an instance of the non-stationary demand single-item pricing problem for which the expected revenue of any policy is upper-bounded by* $\mathsf{OPT}_{\mathrm{LP}}/2$.

## 2.2.5 Non-stationary Demand with Large Inventory

We present alternative policies for non-stationary demand that conduct "reservation" to a lesser degree than in Algorithm 2. Our policies have better performance if starting inventory is large, where the law of large numbers reduces the necessity of reservation, even under non-stationary demand.

We propose a different asymptotic regime from the literature Gallego and Van Ryzin (1997), Talluri and Van Ryzin (1998), Cooper (2002), to name a few. This is because traditional scaling requires $T$ and $b_i, \forall i \in [n]$ to scale up linearly, and under non-stationarity it is unclear how to scale the system. Instead of letting all $T$ and $b_i, \forall i \in [n]$ to scale up linearly, we allow for arbitrary dependence among $T$ and $b_i, \forall i \in [n]$. This asymptotic regime is more of theoretical interests, and is sometimes used in the theoretical CS literature. Note that in practice, the number of initial inventory of different items may be significantly different, which might require some non-trivial normalization to fit into the standard asymptotic regime.

### Statement of results

In Algorithm 3, $\delta$ can be interpreted as the "reservation" probability, which decreases to zero as initial inventory increases. We reserve inventory by offering the

**Algorithm 3** Assortment (and pricing) policy when demand is non-stationary and inventory is large

1: Solve CDLP-N, and let $\{x_t^*(S) : t \in [T], S \in \mathcal{S}\}$ denote an optimal solution.
2: For each time $t$, offer each assortment $S \in \mathcal{S}$ with probability $x_t^*(S) \cdot (1-\delta)$, and offer $\emptyset \in \mathcal{S}$ with probability $\delta$, where $\delta = \sqrt{\frac{3\log(\underline{b})}{\underline{b}}}$.

empty set, which is always available. Note that Theorem 2.8 requires no assumption in the static substitution model.

**Theorem 2.8.** *Under either the static substitution model or under Assumption 2.1 (substitutability), for the assortment (and pricing) policy where demand may be non-stationary, if $\underline{b} \geq 6$, then Algorithm 3 earns revenue that is at least*

$$\left(1 - \sqrt{\frac{3\log(\underline{b})}{\underline{b}}}\right) \mathsf{OPT}_{\mathrm{LP}}$$

*in expectation. In particular, Algorithm 3 is asymptotically optimal as the starting inventories approach infinity.*

**Proof sketch of Theorem 2.8**

Algorithm 3 scales the LP solution by a factor of $1-\delta$, where $\delta$ is a small "reservation" probability. $\delta$ is selected to balance two factors. First, it is small enough such that if we never stock out, then earning $(1-\delta) \cdot \mathsf{OPT}_{\mathrm{LP}}$ is an asymptotically optimal ratio. On the other hand, $\delta$ is large enough such that we stock out with probability at most $1/b$. This intuition is motivated by a tutorial of Anupam Gupta (Gupta 2009), where they introduced the original work of Hajiaghayi et al. (2007). We improve the bounds in the original paper, so that our bound only depends on $b$, but not on $T$. We also generalize to assortment and pricing problems with fractional-demand consumptions.

We outline two key steps here and defer the details of our proof to Section A.10 in Appendix A. The intuition is as follows: conditioning on the event that "inventory never runs out", the expected revenue is at least $1-\delta$ fraction of the LP objective. Then, we show that this event happens with high probability.

Table 2.2: A summary of de-randomization methods

| | Single-item Pricing/Assignment | Assortment (and Pricing) |
|---|---|---|
| Stationary Demand | **Theorem 2.11 (or 2.9)** | **Theorem 2.9** |
| Non-Stationary Demand | **Theorem 2.15 (or 2.9)** | **Theorem 2.9** |
| Non-Stationary Demand (Large Inventory) | **Theorem 2.9** | **Theorem 2.9** |

1. Lower bound the expected revenue by a multiplicative factor of the LP objective, i.e. $\mathbb{E}[\mathsf{Rev}] \geq \Pr[B_T > 0](1-\delta)\mathsf{OPT}_{\mathrm{LP}}$, where $\delta$ is as defined in Algorithm 3.

2. Using concentration of inequality, lower bound the probability that inventory never runs out, i.e.

$$\begin{aligned}\Pr[B_T > 0] &\geq 1 - \Pr\left[\sum_{t=1}^{T}\{Q_t - \mathbb{E}[Q_t]\} \geq \delta b\right] \\ &\geq 1 - \exp\left(-\frac{(\delta b)^2}{2\mathrm{Var}(\sum_{t=1}^{T} Q_t) + 2/3\delta b}\right) \geq 1 - \exp\left(-\frac{\delta^2 b}{2}\right) = 1 - \frac{1}{b}.\end{aligned}$$

## 2.3 De-randomization Methods

In this section we introduce de-randomization methods, which yield deterministic calendars that (i) have the same theoretical guarantees, (ii) significantly improve computational performance (See Section 2.4.2), and (iii) are much easier to accept in practice.

Specifically, we first introduce a general de-randomization method in Section 2.3.1 that applies to any randomized static policy for joint assortment and pricing that we proposed in Section 2.2. Then we introduce two specialized de-randomization methods for Algorithms 1 and 2 when there is only a single item. These methods take advantage of the structural properties in the single-item pricing problem. The de-randomization methods are summarized in Table 2.2.

### 2.3.1 General De-randomization Methods

In this section, we introduce a general simulation-based de-randomization method that achieves the same guarantee as any policy suggested by Algorithms 1–3 does. We consider the general assortment (and pricing) problem under non-stationary demand, which captures single-item pricing and stationary demand as special cases.

Any policy as suggested by Algorithms 1–3 that *independently* chooses the assortment in each time period, implies a distribution over static calendars, which can be characterized by the following vector:

$$\boldsymbol{z} = (z_t(S) : t \in [T], S \in \mathcal{S}) \in [0,1]^{T \cdot |\mathcal{S}|},$$

where $z_t(S) = \Pr\{S_t = S\}$ is the probability that we offer $S$ in period $t$. Note there might exist $t$ such that $z_t(\emptyset) > 0$ because our policies could possibly suggest offering nothing in some periods. We will use the following example to illustrate our de-randomization procedure.

**Example 2.2.** Consider a three-period problem with three options $\mathcal{S} = \{S_A, S_B, \emptyset\}$. Note that $\emptyset$ is always available. Suppose that our randomized policy (possibly from Algorithm 2) is characterized by

$$\begin{aligned} \boldsymbol{z} = (z_1(S_A) = 0.5, \quad & z_1(S_B) = 0.5, \quad & z_1(\emptyset) = 0; \\ z_2(S_A) = 0.5, \quad & z_2(S_B) = 0, \quad & z_2(\emptyset) = 0.5; \\ z_3(S_A) = 0, \quad & z_3(S_B) = 1, \quad & z_3(\emptyset) = 0) \end{aligned}$$

This policy implies a distribution over static calendars such that it takes $(S_A, S_A, S_B)$ with prob. 1/4, $(S_A, \emptyset, S_B)$ with prob. 1/4, $(S_B, S_A, S_B)$ with prob. 1/4, and $(S_B, \emptyset, S_B)$ with prob. 1/4 (because the assortments for $t = 1$ and $t = 2$ are drawn independently). The idea of our de-randomization method is to select one of them that garners the same expected revenue as the distribution of calendars does. ∎

In general, this distribution over static calendars has a finite but exponentially large support; computing the expected revenue of each calendar using brute force is impossible. Instead, our method identifies the best assortment to offer iteratively over $t = 1, \ldots, T$, using simulation. Our method requires a simulator $\nu(\boldsymbol{z}, \xi) : [0,1]^{T \cdot |\mathcal{S}|} \times \Xi \to \mathbb{R}$, whose source of randomness (e.g. the random seed) is characterized by

$\xi \in \Xi$. In each single run of the simulator, it randomly generates (i) a calendar of assortments $S_1, S_2, ..., S_T$, based on the probabilities suggested by $\boldsymbol{z}$, and (ii) a sequence of demands based on the choice models and the assortments on the calendar. Finally, the simulator calculates revenue based on the simulated assortments and demands. The simulator generates revenue from a bounded interval $[0, (b_1 + \cdots + b_n)p_{\max}]$, where $p_{\max} = \max_{j\in[m]} p_j$ is given.

Now that we have the simulator, if we query this simulator $K$ times, then we obtain an estimator $\hat{\mu}_K(\boldsymbol{z}) = \frac{1}{K}\sum_{k=1}^{K} \nu(\boldsymbol{z}, \xi_k)$ of the expected revenue of policy $\boldsymbol{z}$. We can select $K$ to be a large number such that $\hat{\mu}_K(\boldsymbol{z})$ is close to $\mathbb{E}[\mathsf{Rev}(\boldsymbol{z})]$ via a concentration inequality. We specify our de-randomization method in Algorithm 4.

---

**Algorithm 4** Simulation-based de-randomization method for Algorithms 1–3

1: Initialize $\boldsymbol{z}$ to be the distribution over calendars suggested by any of Algorithms 1–3.

2: Fix $K = \lceil \frac{T^2(b_1+\cdots+b_n)^2 p_{\max}^2}{\mathsf{OPT}_{\mathrm{LP}}^2} \cdot \frac{1}{\epsilon^2} \cdot (\log n + \log T)\rceil$, the number of samples to average in each period.

3: **for** $t = 1, 2, ..., T$ **do**

4: For each assortment $S \in \mathcal{S}$ with $z_t(S) > 0$, estimate the revenue of the randomized calendar $\boldsymbol{z}$ with the assortment $S_t$ to offer at time $t$ deterministically set to $S$. Let $\hat{\mu}_K(\boldsymbol{z}\,|S_t = S)$ denote the estimate realized (taking the average of $K$ simulation runs), for each $S \in \mathcal{S}$. Select

$$\hat{S}_t \in \underset{S\in\{S\in\mathcal{S}|z_t(S)>0.\}}{\arg\max} \hat{\mu}_K(\boldsymbol{z}|S_t = S).$$

5: Update $\boldsymbol{z}$ such that $z_t(\hat{S}_t) = 1$, and $z_t(S) = 0$ for all $S \neq \hat{S}_t$.

6: Offer assortment $\hat{S}_t$ at time $t$.

7: **end for**

---

**Example 2.3** (Example 2.2 Continued)**.** Let $(0.5S_A, 0.5S_B; 0.5S_A, 0.5\emptyset; S_B)$ denote a randomized policy that offers $S_A$ and $S_B$ each with probability one half in the first period, offers $S_A$ and $\emptyset$ each with half-probability in the second period, and finally offers $S_B$ in the third period. Algorithm 4 first finds the better one between $(S_A; 0.5S_A, 0.5\emptyset; S_B)$ and $(S_B; 0.5S_A, 0.5\emptyset; S_B)$. If the latter is better, in the second iteration it finds the better one between $(S_B; S_A; S_B)$ and $(S_B; \emptyset; S_B)$. ■

This idea of iterative de-randomization, when the support of the randomized solution is exponentially sized, has commonly appeared in the computer science literature (Motwani and Raghavan 1995). However, the need for a simulator to evaluate the assortments at each iteration and the analysis of how many samples are needed to lose at most $\epsilon$ in the final de-randomized solution are new to our paper, to the best of our knowledge.

We prove the following result, the proof of which is deferred to Section A.11 in Appendix A.

**Theorem 2.9.** *If any policy from Algorithms 1 – 3 earns expected revenue of at least* $\alpha \cdot \mathsf{OPT}_{\mathrm{LP}}$*, then the static calendar suggested by Algorithm 4 earns expected revenue of at least* $(\alpha - \epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}}$*.*

*The time complexity of Algorithm 4 is* $O(\frac{(b_1+\cdots+b_n)^2 p_{\max}^2}{\mathsf{OPT}_{\mathrm{LP}}^2} \cdot \frac{1}{\epsilon^2} n^2 T^4 (\log n + \log T))$*.*

## 2.3.2 Single-Item Stationary Demand

We now introduce a specific de-randomization method for Algorithm 1 (the algorithm for stationary demand) in the special case of single-item pricing. While the generic de-randomization method from Section 2.3.1 will also suffice, the one presented here contains additional structural insights about the de-randomized calendar. Note that with a single item, dynamic and static substitution are equivalent.

We establish a structural property in Section 2.3.2 which shows that sorting the static calendar in order of high-to-low prices is dominating. We show that sorting in the opposite order (low-to-high) earns strictly less expected revenue, in Example 2.4 in Section 2.3.2. We also show that using only one price (without regularity assumptions) earns strictly less expected revenue, in Proposition 2.14 in Section 2.3.2.

In the single-item pricing problem, we have $n = 1$, and we will omit index $i$. To be able to handle [0,1]-demand (instead of just {0,1}-demand), we require the following assumption.

**Assumption 2.3.** *For any* $j, j' \in [m]$*, either* $F_j(x) \geq F_{j'}(x)$*, or* $F_j(x) \leq F_{j'}(x)$*, for all* $x \in [0, 1]$*.*

When demand is stationary, we have the following LP. For simplicity, we omit the infinite price, and we put an inequality instead of an equality in (2.16).

$$J^{DLP-S} = \max T \cdot \sum_{j=1}^{m} p_j q_j x_j \tag{2.14}$$

$$\text{s.t.} T \cdot \sum_{j=1}^{m} q_j x_j \leq b \tag{2.15}$$

$$\sum_{j=1}^{m} x_j \leq 1 \tag{2.16}$$

$$x_j \geq 0 \qquad \forall\, j = 1, \ldots, m \tag{2.17}$$

The LP for pricing under stationary demand has the following structure.

**Lemma 2.10** (Gallego and Van Ryzin (1994))**.** *The DLP-S defined by (2.14), (2.15), (2.16), and (2.17) has a basic optimal solution $(x_j^*)_{j=1}^m$ with at most two non-zeros in its support, which we will denote using $x_{\mathsf{H}}^*$ ("Higher price") and $x_{\mathsf{L}}^*$ ("Lower price"), with $p_{\mathsf{H}} \geq p_{\mathsf{L}}$*

Based on the above LP and the optimal structure from Lemma 2.10, we devise the following policy. Our policy offers the prices in a *high-to-low* order, with a static switching point. Intuitively, the high-to-low ordering is desirable, because should we stock out early from higher-than-expected demand realizations, we would rather lose low-priced sales at the end.

---

**Algorithm 5** Single-item pricing policy when demand is stationary

1: Solve DLP-S, and let $p_{\mathsf{H}}, p_{\mathsf{L}}, x_{\mathsf{H}}^*, x_{\mathsf{L}}^*$ correspond to an optimal solution as described in Lemma 2.10. Denote $s_{\mathsf{H}} = T \cdot x_{\mathsf{H}}^*/(x_{\mathsf{H}}^* + x_{\mathsf{L}}^*)$.

2: Set the price to be $p_{\mathsf{H}}$ for $t = 1, \ldots, s^*$ and $p_{\mathsf{L}}$ for $t = s^* + 1, \ldots, T$. Here $s^*$, the duration for which the higher price is offered, is either $\lfloor s_{\mathsf{H}}^* \rfloor$ or $\lceil s_{\mathsf{H}}^* \rceil$.

---

**Theorem 2.11.** *Under Assumption 2.3, for the single-item pricing problem under stationary demand with $b$ units of inventory to sell over $T$ periods, Algorithm 5 earns* *expected revenue at least*

$$\frac{\mathbb{E}[\min\{\mathrm{Bin}(T, b/T), b\}]}{b} \cdot \mathsf{OPT}_{\mathrm{LP}},$$

We prove Theorem 2.11 in the next section.

**Structural property: monotonicity**

We begin by quickly establishing a structural property, Lemma 2.12, as a warm-up to the proof of Lemma 2.13, which is the key to the de-randomization in the single-item stationary demand setting. Let $v_t \in [m]$ denote the price index for time $t$ in a calendar, and $v_t^* \in [m]$ denote the optimal price index in a revenue-maximizing calendar. We use $\boldsymbol{v}$ to describe the calendar, a vector of price indices. The structural property states the following:

**Lemma 2.12.** *In any calendar $\boldsymbol{v}$, if two consecutive price indices $v_t, v_{t+1}$ are such that $p_{v_t} < p_{v_{t+1}}$, then indices $v_t$ and $v_{t+1}$ can be exchanged in the calendar without decreasing its expected revenue.*

The proof of Lemma 2.12 is deferred to Section A.12 in Appendix A. From this Lemma, we know that there exists an optimal static calendar, the prices of which are non-increasing over time.

Now we strengthen the monotonicity property in Lemma 2.12. Consider a problem instance with $T = 3$ time periods and starting inventory $b = 2$. Suppose we have two prices. The higher price of 2 earns a sale with probability 1/3; the lower price of 1 earns a sale with probability 1, i.e. deterministically. The optimal LP solution (according to Lemma 2.10) suggests offering a higher price index $\mathsf{H}$ for 1.5 time periods, and a lower price index $\mathsf{L}$ for 1.5 time periods.

We let $\mathbb{E}[\mathsf{Rev}(\mathsf{H}; 0.5\mathsf{H}, 0.5\mathsf{L}; \mathsf{L})]$ denote the expected revenue of a randomized policy that offers $\mathsf{H}$ in the first period, offers $\mathsf{H}$ and $\mathsf{L}$ each with probability one half in the second period, and offers $\mathsf{L}$ in the third period. Similarly, we also define the following $\mathbb{E}[\mathsf{Rev}(0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L})]$.

The structural property states that if there is a positive probability that one policy offers a lower price before a higher price, then this policy can be improved. For example,

$$\mathbb{E}[\mathsf{Rev}(0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L})] \leq \mathbb{E}[\mathsf{Rev}(\mathsf{H}; 0.5\mathsf{H}, 0.5\mathsf{L}; \mathsf{L})], \tag{2.18}$$

There is a positive probability that the policy $(0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}, 0.5\mathsf{L})$ does so, since there is already a $1/4$ chance that it offers $\mathsf{L}$ in period 1 and $\mathsf{H}$ in period 2. However, the randomized policy $(\mathsf{H}; 0.5\mathsf{H}, 0.5\mathsf{L}; \mathsf{L})$ could only lead to the calendars $(\mathsf{H}; \mathsf{H}; \mathsf{L})$ or $(\mathsf{H}; \mathsf{L}; \mathsf{L})$; in either case it always offers higher prices before lower prices. The conclusion of inequality (2.18) is that the first policy can be changed to the second policy without reducing revenue; note that the total expected number of periods that both $\mathsf{H}$ and $\mathsf{L}$ are offered is still the same (1.5 periods each).

We now formalize inequality (2.18). Let $\{x\} = x - \lfloor x \rfloor$ be the fractional part of a real number $x$.

**Lemma 2.13.** *Consider the following two policies:*

1. *A policy that offers in each period the same probabilistic mixture of two prices, i.e. a probability $\alpha$ of offering the higher price and a probability $1-\alpha$ of offering the lower price;*

2. *A policy that starts by deterministically offering the higher price for $\lfloor \alpha \cdot T \rfloor$ periods, then in the next period offers the higher price with probability $\{\alpha \cdot T\}$ and the lower price with probability $1 - \{\alpha \cdot T\}$, and finally switches to offering the lower price in the last $\lceil (1-\alpha) \cdot T \rceil - 1$ periods.*

*The expected revenue of the second policy is no less than the expected revenue of the first policy.*

The proof of Lemma 2.13 can be found in Section A.12.

*Proof.* Proof of Theorem 2.11. Observe that Algorithm 1 suggests the first calendar in Lemma 2.13; Algorithm 5 suggests the second calendar in Lemma 2.13. Theorem 2.11 holds by invoking Theorem 2.2. $\square$

**Switching from high to low is necessary**

We show that switching from a higher price to a lower price is necessary, in the sense that if we switch from a lower price to a higher price, we may fail to achieve the bound by expression (2.9).

**Example 2.4.** Let there be $T = 2$ periods and $b = 1$ unit of initial inventory. Let there be two prices: $p_1 = 8, p_2 = 1$. The corresponding purchase probabilities are $q_1 = 0.1, q_2 = 0.9$. The LP suggests that we offer both $p_1$ and $p_2$ for exactly one period. The LP objective is $\mathsf{OPT}_{\text{LP}} = p_1 q_1 \cdot 1 + p_2 q_2 \cdot 1 = 1.7$

We calculate the bound in expression (2.9): it suggests a $\mathbb{E}[\min\{\text{Bin}(T, b), b\}]/b =$ 75% guarantee.

If we offer $p_2$ in period 1 and then $p_1$ in period 2, this earns an expected revenue of $p_2 q_2 + (1 - q_2) p_1 q_1 = 0.98$, which is $0.98/1.7 \approx 57.6\%$ of the LP upper bound.

If we offer $p_1$ in period 1 and then $p_2$ in period 2, this earns an expected revenue of $p_1 q_1 + (1 - q_1) p_2 q_2 = 1.61$, which is $1.61/1.7 \approx 94.7\%$ of the LP upper bound. ∎

This example demonstrates that switching from a lower price to a higher price is worse than the bound by expression (2.9), and it is even worse than the $1 - 1/e \approx 63.2\%$ ratio. On the other hand, switching from a higher price to a lower price performs much better.

We also show that two prices are needed to obtain our results in Theorem 2.2. This is because we do not assume regularity assumptions. We state Proposition 2.14 here and defer its proof to Section A.13 in Appendix A.

**Proposition 2.14.** *There exists an instance of the stationary-demand single-item pricing problem for which the expected revenue of any single price policy is strictly smaller than expression (2.9).*

### 2.3.3 Single-Item Non-Stationary Demand

Analogous to Section 2.3.2, we now introduce a specific de-randomization method for Algorithm 2 (the algorithm for non-stationary demand) in the special case of single-item pricing.

When demand is non-stationary, we have the following LP. Let $x_{tj}, \forall t \in [T], j \in [m]$ be the probability that we offer price $j$ in time $t$.

$$
\begin{aligned}
J^{DLP-N} = \max \sum_{t=1}^{T} \sum_{j=1}^{m} & p_j q_{tj} x_{tj} \\
\text{s.t.} \sum_{t=1}^{T} \sum_{j=1}^{m} & q_{tj} x_{tj} \leq b \\
\sum_{j=1}^{m} & x_{tj} \leq 1 \qquad \forall\, t = 1, \ldots, T \\
& x_{tj} \geq 0 \qquad \forall\, t = 1, \ldots, T; \ \forall\, j = 1, \ldots, m
\end{aligned}
$$

**Algorithm 6** Single-item pricing policy when demand is non-stationary

1: Solve DLP-N, and let $r^*$ denote the optimal objective value.
2: For each time $t$, set the price to be $p_{j_t}$, where

$$
j_t \in \arg\max_j (p_j - \frac{r^*}{2b}) q_{tj}. \tag{2.19}
$$

In (2.19), $r^*/b$ can be interpreted as the *per-inventory revenue of the LP*. Algorithm 6 guarantees to sell inventory for at least half of this value, since at each time $t$, it maximizes the expected profit with a *bid price* (opportunity cost) of $r^*/(2b)$. The intuition is that when there is only one item, we can treat the threshold as a bid price and maximize with respect to it to obtain a deterministic calendar. This de-randomization is because we use maximization instead of a discarding rule.

**Theorem 2.15.** *For the single-item pricing problem where demand may be non-stationary, Algorithm 6 earns expected revenue at least* $\mathsf{OPT}_{\mathrm{LP}}/2$.

We defer the proof of Theorem 2.15 to Section A.14 in Appendix A.

## 2.4 Computational Study

### 2.4.1 Computational Study: Using Real Data from A CPG Company

We first describe the business model. Then in Section 2.4.1, we explain how we develop the prediction model from data. We discuss the details of feature selection in Section 2.4.1 and justify the motivation of our dynamic pricing model. In Section 2.4.1, we explain how we adapt the prediction model such that its output is consistent with managerial intuition and statistically effective. Finally, in Section 2.4.1, we explain the numerical performance under our proposed policies.

At the end of each year, the CPG company requires a price calendar to be planned for the next year. This calendar contains 52 weekly prices for each SKU. The CPG company then brings this calendar to its channels (e.g. supermarkets) to negotiate the price-to-customers (PTC). We assume that they are the same, since the CPG company has full bargaining power. After the calendar is delivered to channels, the channels decide their yearly advertising strategy, produce flyers, and make price tags. These are the reasons (e.g., long lead time on flyers) why we need to plan a calendar in advance. Customers will not see the prices until the channels release their prices, so there is no anticipatory behavior.

#### The random forest model

In this section we explain in detail how we develop the prediction model from the data. We will follow the workflow shown in Figure 2-1 from Section 2.1.4.

We begin with weekly sales data in the past 3 years. After cleaning the missing data, we select SKUs that generated 90% of the revenue in the past three years and eliminate the rest. We also eliminate SKUs that were newly introduced in the most recent year. Some SKUs are already grouped together by the company. They are similar brands sold at similar pack sizes. The company requires that all SKUs in the same group be sold at the same price. There are 52 distinct groups in total. We build

Table 2.3: Different combinations of features, and the resulting out-of-sample error rates

| | | | | | | |
|---|---|---|---|---|---|---|
| Tagged price | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Seasonal industry trend (after moving average) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Total number of stores in the district | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Festivals and sports events | | ✓ | ✓ | ✓ | ✓ | |
| External competitor prices | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Internal competitor prices (within brand) | | ✓ | | ✓ | | ✓ |
| History prices | | ✓ | ✓ | | | |
| Average error rate | 32.21 | 21.07 | 21.09 | 19.47 | **19.41** | 19.80 |

Note: In the last row, the average error rate is taken over both time periods and different SKUs.

group-specific prediction models with the same combination of features, i.e., all SKUs use the feature "tagged price", but it refers to a different tagged price for each SKU.

We derive a list of features from the data that will be used to predict demand at each time step. These features include the price that this group is tagged at, its internal competitor prices, its external competitor prices, and its history prices. The internal competitor prices are the prices of the brands owned by the same company. The external competitor prices are the prices of its true competitors, owned by its rival companies. The features of history prices are take from the past week to the past 3 weeks, as 3 different features.

The external features include industry seasonal trend (after applying moving average), total number of stores in the district, festivals and sports events. The first two features are provided by the company, and the rest are obtained by scripting from the Internet. We create dummy variables for festivals and sports events to characterize categorical data.

We tested a few algorithms and finally choose to use random forest (Liaw et al. (2002), Ferreira et al. (2016)) as the prediction model. In Section 2.4.1 we will discuss an important challenge associated with random forest prediction and how it is addressed. We aggregate all the features together, then simultaneously perform feature selection and parameter tuning, by using a 5-fold cross-validation. Finally, the average prediction error is reported as 19.41%. This demonstrates a very good prediction model, compared with the number in Ferreira et al. (2016).

### Feature selection

In this section we explain how we select features. We validated our model based on the selected features, because this provides indications of the existence (or non-existence) of different sources of cannibalization effects. For example, if we identified a temporal cannibalization effect, which states that a promotion given today will take away future sales, then our model should address it. Fortunately, there are no such complications, as supported by the data.

In cross-validation, we evaluated each feature combination based on its median absolute percentage error (MdAPE) on the validation set. During this procedure, we engaged in rounds of discussions with the company to ensure that the features selected are interpretable. There are some sub-optimal combinations that the company believed would make more practical sense, and we followed their advice.

These features were both approved by the CPG company's management as consistent with their expedience and also resulted in the lowest out-of-sample prediction errors — see Table 2.3 for the reported error rates. Each column depicts a combination of features, and the corresponding numbers are prediction errors under this feature combination. The first column serves as a benchmark. We omit some trivial duplicates of the same feature, but note that some rows represent many features, e.g., festivals and sports events.

The features that were ultimately selected include: the tagged price, external competitor prices[5], and some external features. Note that this list includes neither internal competitor prices (from other SKUs of the CPG company) nor historical prices. We validate our model with the following two observations, which suggest that our model captures the real retail dynamics:

Cross-product cannibalization is not significant. A probable reason is that we have already grouped similar brands sold at similar pack sizes. This suggests that we can employ single-item calendar pricing without considering the joint optimization of

[5] We do not know the true competitor prices, but we can use ARIMA (Hyndman et al. 2007, 2020), a time series model, to predict competitor prices; by substituting true competitor prices with predicted competitor prices we find the prediction accuracy evaluated on the testing set remains almost unchanged. Thus, we use predicted competitor prices instead of true competitor prices.

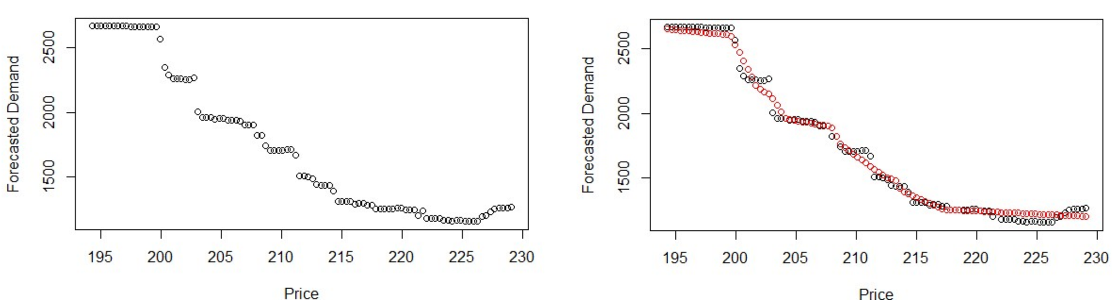
simultaneously deciding all SKUs' calendars.

Inter-temporal cannibalization is also not significant. This suggests that demands are not correlated across time, in alignment with our theoretical model. A good explanation of this observation is due to the "pantry effect" (Ailawadi and Neslin 1998, Bell et al. 1999). Even if customers stockpile a product during promotions, this accelerates their consumption rate. Thus, they ultimately purchase as much of the product in subsequent time periods as they otherwise would have, especially for products such as carbonated beverages and ice cream. Therefore, we model demands as independent across time.

### Monotone demand curve

If we focus on the price-demand relationship, we observe that the direct output of random forest yields a non-monotone prediction. In Figure 2-2, the black dots show the predicted price-demand curve for one SKU in one week. There are occasions when the predicted demand has positive price elasticity, e.g. when price is between 225 and 230. Since we are selling consumer packaged goods, there is no conspicuous leisure (Veblen 2017), and price elasticity should be negative. A non-decreasing demand curve is not acceptable to the CPG company from a managerial perspective. Theoretically, a non-decreasing demand curve might also (but not necessarily) violate Assumption 2.3.

Figure 2-2: Predicted price-demand relationship, before and after curve fitting




Note: The prices are in industry units.

To solve this problem, we introduce curve fitting. First uniformly draw 100 samples $(P_i, D_i)_{i=1}^{100}$ from the demand curve. Then fit a piecewise linear function, written as

$$f_\Theta(x) = \sum_{i=1}^{d-1} \mathbb{1}_{\{x_i<x<x_{i+1}\}} \left( \frac{y_{i+1}-y_i}{x_{i+1}-x_i}(x-x_i)+y_i \right) + \mathbb{1}_{\{x<x_1\}} \left(a_l(x-x_1)+y_1\right) + \mathbb{1}_{\{x>x_d\}} \left(a_r(x-x_d)+y_d\right)$$

parameterized by $\Theta = \{x_1, ..., x_d, y_1, ..., y_d, a_l, a_r\} \in \mathrm{R}_+^{2d+2}$, where $d$ is the number of breakpoints, and $\mathbb{1}_{\{x>a\}}$ are indicator functions equal to one if $x > a$, zero if $x \leq a$. We arbitrarily selected $d$ to be 10. Finally, we minimize the mean squared error over these 100 sample points, with shape constraints enforcing a monotonic decreasing function.

$$\begin{aligned} \min_{\Theta \in \mathcal{R}^{2d+2}} \quad & \sum_{i=1}^{100} (Y_i - f_\Theta(X_i))^2 \\ s.t. \quad & \min_j X_j \leq x_1 < x_2 < ... < x_d \leq \max_j X_j \\ & \max_j Y_j \geq y_1 > y_2 > ... > y_d \geq \min_j Y_j \\ & a_l, a_r < 0 \end{aligned}$$

We solve the above program using heuristics. The curve-fitting output is depicted in Figure 2-2 as the red dots. The accuracy is slightly improved from 19.41% to 18.66%. We do not view this improvement as tremendous, but we have built a model that is in greater agreement with managerial suggestions. Finally, a monotone demand curve[6] satisfies Assumption 2.3, as addressed in Section 2.3.2.

**Computational performance of policies**

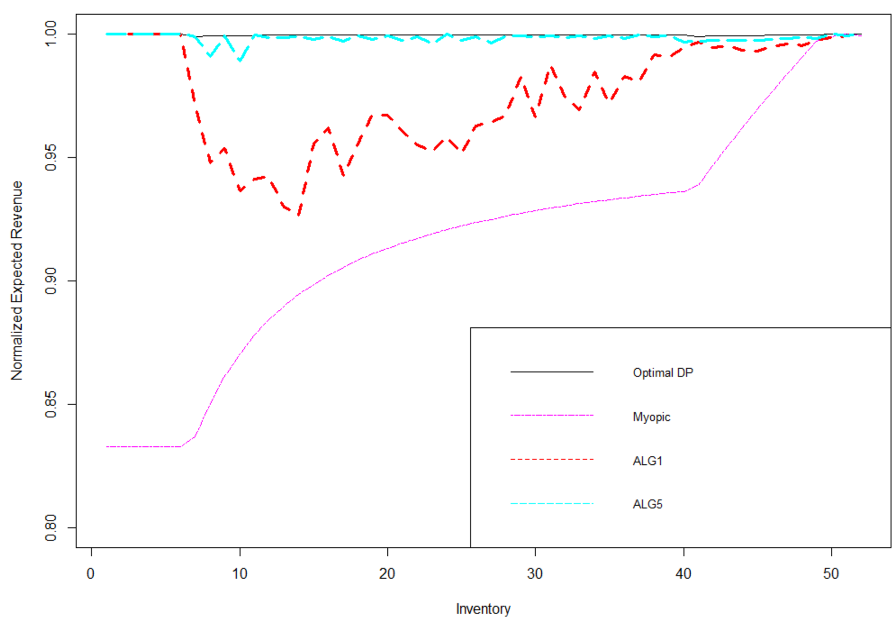
In this section, we take distributions obtained from the above sections as inputs and compare the performance of our policies to selected benchmarks. We fix the feasible price set for each SKU to be the prices from its historical data. The planning horizon is one year, 52 weeks. We normalize demands to take $[0, 1]$ values by dividing the

[6] To be precise, we derive several scenarios for a specific SKU at a specific price to obtain a discrete distribution, and this suggests a monotone demand curve under each scenario.

predicted demands by the highest predicted demand. We consider different scenarios in which the starting inventory ranges from 1 unit to 52 units and analyze both stationary and non-stationary demand models.

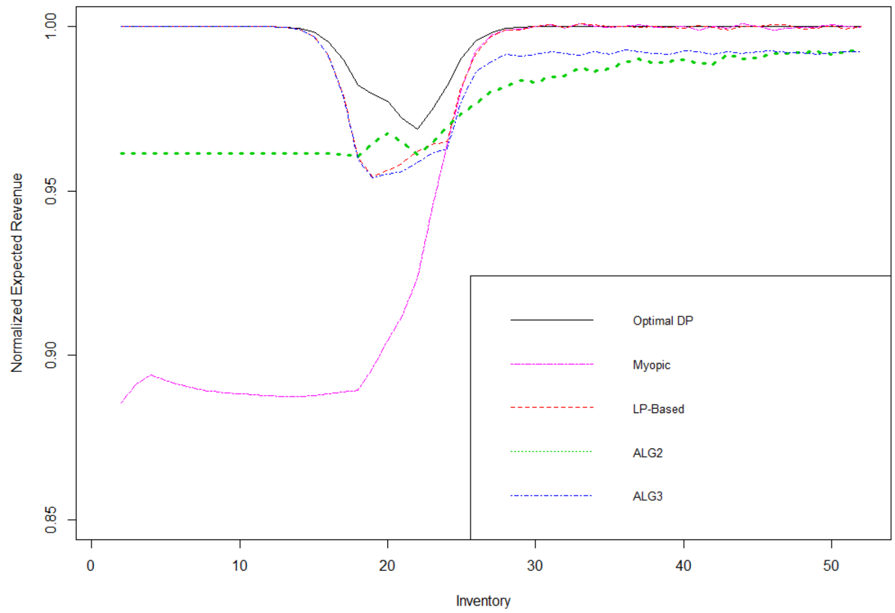
We compute the expected revenue from our proposed policies (**ALG1** from Algorithm 1 under stationarity, **ALG2** from Algorithm 2 under non-stationarity, **ALG3** from Algorithm 3 under non-stationarity, and **ALG5** from Algorithm 5 the deterministic policy under stationarity), the LP upper bound, the **Optimal DP** for the policy solving the optimal dynamic program, the **Myopic** policy as one benchmark, and an **LP-Based** randomized policy as another benchmark. The results are shown in Figures 2-3 and 2-4, where we have divided all numbers by the corresponding LP upper bound, meaning that the performance ratio is always between 0 and 1, with higher ratios indicating better performance.

Figure 2-3: Computational performance of polocies under stationary demand




For scenarios in which the starting inventory is of moderate size compared to the total expected demand (i.e. for SKUs that were initially neither overstocked nor understocked), our static policies outperform basic LP-based policies by 5% under stationarity and 1% under non-stationarity. Furthermore, our static policies lose at

Figure 2-4: Computational performance of polocies under non-stationary demand




most 1% under stationarity and 4% under non-stationarity, compared to the optimal dynamic policies.

Note that in practice, it is rare for the initial inventory level to be very small or very large, since it would have been pre-optimized[7] to sell out exactly. When inventory is of moderate size, our policies outperform the existing benchmarks under both stationary and non-stationary demand settings.

In fact, if we consider the prediction model, the expected demand is approximately $0.2 \sim 0.6$ on each day under different prices, which corresponds to the dip when inventory $b$ is around $10 \sim 30$. If we divide them by the time horizon of $T = 52$ weeks, we see that the expected units sold per week $b/T$ roughly meets the expected (normalized) demand of $0.2 \sim 0.6$. This is the region where the pricing problem is non-trivial in theory, and most common in practice. When the inventory level is such that the problem falls into degenerate cases, all the curves are close to the LP upper bound. This moderate inventory size corresponds to the moderate load scaling factor of $0.6 \sim 1.4$, which is the ratio between initial inventory and mean demand in the

[7]Pre-optimized by some higher level managers, Talluri and Van Ryzin (2006).

admission control problems originated from Zhang and Cooper (2005).

### 2.4.2 Computational Study: Using Synthetic Data from Literature

In this section, we study the joint assortment and pricing problem, using synthetic data that are commonly adopted in the choice-based deterministic linear program literature, such as Zhang and Cooper (2005), Liu and Van Ryzin (2008), Gallego et al. (2016).

We closely follow their numerical setup. Let there be three items, each of which can be consumed fractionally. The items have initial inventory proportional to $\boldsymbol{b} = (3, 5, 4)$. We will later normalize initial inventory by a load scaling factor $\alpha \in \{0.6, 0.8, 1.0, 1.2, 1.4\}$. Each item has two prices to be offered, high and low. We fix the low prices $\boldsymbol{p}_{\mathrm{L}} = (400, 500, 300)$ and let the high price to change from $\boldsymbol{p}_{\mathrm{H}} \in \{(800, 1000, 600), (8000, 10000, 6000)\}$. We also denote these prices as $\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3$ and $\mathrm{H}_1, \mathrm{H}_2, \mathrm{H}_3$.

We specify a choice model to be adopted in our computational study. It is an adaption from a mixture of MNL models (see, e.g., Li et al. (2018), McFadden and Train (2000)). However, we interpret the choice probabilities as purchasing market shares, which are between [0,1]. The choice model can be explicitly written as follows

$$Q_t(i, j, S) = A_t^{\mathrm{L}} \frac{\sum_{(l, \mathrm{L}_l) \in S} \mathbb{1}_{(l, \mathrm{L}_l) = (i, j)} \cdot v_{l, \mathrm{L}_l}}{v_\emptyset^L + \sum_{(l, \mathrm{L}_l) \in S} v_{l, \mathrm{L}_l}} + A_t^{\mathrm{H}} \frac{\sum_{(l, \mathrm{H}_l) \in S} \mathbb{1}_{(l, \mathrm{H}_l) = (i, j)} \cdot v_{l, \mathrm{H}_l}}{v_\emptyset^H + \sum_{(l, \mathrm{H}_l) \in S} v_{l, \mathrm{H}_l}},$$

where we adopt the fashion that $\frac{0}{0} = 0$. The randomness in this model comes from the coefficients in the front of each single MNL model. We let $A_t^{\mathrm{L}}$ and $A_t^{\mathrm{H}}$ be Bernoulli random variables. The mean values $\mathbb{E}[A_t^{\mathrm{L}}]$ and $\mathbb{E}[A_t^{\mathrm{H}}]$ are given for any $t \in [T]$.

In our computational study, we distinguish between a stationary setting and a non-stationary setting. In both settings, $T = 20$. In the stationary setting, $\mathbb{E}[A_t^{\mathrm{L}}] = 0.3, \mathbb{E}[A_t^{\mathrm{H}}] = 0.2, \forall t \in [T]$. In the non-stationary setting, $\mathbb{E}[A_t^{\mathrm{L}}] = 0.8, \mathbb{E}[A_t^{\mathrm{H}}] = 0, \forall t \leq 12; \mathbb{E}[A_t^{\mathrm{L}}] = 0.2, \mathbb{E}[A_t^{\mathrm{H}}] = 0.2, \forall t \geq 13$. We also specify the attractiveness vectors as

$\boldsymbol{v}_{\mathrm{L}} = (5,1,10), \boldsymbol{v}_{\mathrm{H}} = (5,10,1)$. We test four different no-purchase vectors to be $(v_{\emptyset}^{\mathrm{L}}, v_{\emptyset}^{\mathrm{H}}) \in \{(0,0),(1,5),(5,10),(10,20)\}$.

In Tables 2.4–2.6, **LP UB** stands for the corresponding DLP upper bounds; **Myopic** stands for the myopic policy that offers the assortment that gives the highest expected revenue, regardless of inventory; **LP-Sol** stands for the CDLP benchmark, which first solves the LP, then directly uses the optimal solution to implement a (randomized) policy; **ALG2** and **ALG3** stand for the policy suggested by Algorithm 2 and Algorithm 3, respectively; **RST17** stands for the static policy suggested by Rusmevichientong et al. (2020), by considering the extension in their Section 5.2; and **DeRLP**, **DeR2**, and **DeR3** stand for our de-randomization method from Algorithm 4 applied to the LP-Sol, ALG2, and ALG3, respectively. All of the percentages are relative to the LP upper bound.

Table 2.4: Computational performance in the stationary setting

| | LP UB | Myopic | LP-Sol | RST17 | DeRLP | LP-UB | Myopic | LP-Sol | RST17 | DeRLP |
|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | | (0,0) | | | | | (1,5) | | |
| 0.6 | 4300.0 | 67.76% | 74.51% | 81.79% | **82.19%** | 3800.0 | 71.15% | 80.15% | **86.62%** | 85.50% |
| 0.8 | 5200.0 | 71.28% | 80.27% | 83.55% | **83.90%** | 4266.7 | 75.93% | 87.41% | 88.23% | **90.11%** |
| 1 | 6050.0 | 71.95% | 82.77% | 81.27% | **83.22%** | 4566.7 | 78.22% | **92.74%** | 89.74% | 92.53% |
| 1.2 | 6100.0 | 79.70% | 90.54% | 89.17% | **90.88%** | 4586.7 | 83.13% | 95.04% | 93.77% | **95.87%** |
| 1.4 | 6150.0 | 84.42% | 93.46% | 92.13% | **94.45%** | 4606.7 | 86.61% | 95.26% | 95.85% | **97.63%** |
| $\alpha$ | | | (5,10) | | | | | (10,20) | | |
| 0.6 | 3200.0 | **91.67%** | 87.92% | 89.54% | 91.23% | 2468.9 | **94.46%** | 92.80% | 91.05% | 90.81% |
| 0.8 | 3466.7 | 94.28% | 94.20% | 90.97% | **93.32%** | 2533.3 | **97.47%** | 97.33% | 94.78% | 97.33% |
| 1 | 3500.0 | 97.37% | 97.46% | 95.29% | **97.47%** | 2533.3 | 99.29% | **99.35%** | 98.23% | **99.35%** |
| 1.2 | 3500.0 | **99.14%** | 99.13% | 98.00% | 99.09% | 2533.3 | 99.76% | 99.81% | 99.41% | **99.87%** |
| 1.4 | 3500.0 | 99.86% | 99.81% | 99.28% | **99.89%** | 2533.3 | **100.06%** | 99.92% | 99.94% | 99.96% |

In the stationary setting, we observe from Table 2.4 that Myopic performs the best when inventory is too much, which is not surprising. LP-Sol and RST17 have similar performance. The de-randomization method from Algorithm 4 uniformly improves $(7.68\% \sim -1.99\%)$ on the randomized policy in most scenarios. It also has better performance than RST17 in most scenarios.

Moving to non-stationary setting, we observe from Table 2.5 that when the price difference (between $\boldsymbol{p}_{\mathrm{L}}$ and $\boldsymbol{p}_{\mathrm{H}}$) is small, our ALG2 is identical to CDLP and performs well $(81.37\% \sim 99.87\%)$. This is because in many cases, our virtual cost is not large enough to discard any products from the assortment. ALG2 is among the best algorithms in many of the simulation scenarios $(v_{\emptyset}^{L}, v_{\emptyset}^{L}) \in \{(1,5),(5,10),(10,20)\}$.

Table 2.5: Computational performance in the non-stationary setting when the price difference is small

| | LP UB | Myopic | LP-Sol | ALG2 | ALG3 | RST17 | DeRLP | DeR2 | DeR3 |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | | | | (0,0) | | | | |
| 0.6 | 3936.0 | 53.76% | 81.42% | 81.37% | 79.29% | **84.13%** | **83.47%** | **83.40%** | **83.51%** |
| 0.8 | 4981.3 | 54.43% | 84.65% | 84.61% | 82.34% | **86.26%** | **86.80%** | **86.82%** | 85.39% |
| 1 | 6026.7 | 54.50% | **86.39%** | **86.41%** | 84.23% | 84.16% | **87.34%** | **87.59%** | **87.31%** |
| 1.2 | 6304.0 | 60.83% | 87.17% | 87.22% | 84.94% | **90.08%** | **89.72%** | **89.67%** | **89.60%** |
| 1.4 | 6581.3 | 66.59% | 87.33% | 87.38% | 85.37% | **90.58%** | **89.43%** | **89.37%** | **89.42%** |
| $\alpha$ | | | | | (1,5) | | | | |
| 0.6 | 3696.0 | 65.96% | 84.81% | 84.75% | 82.85% | **87.85%** | **86.03%** | **85.93%** | **85.93%** |
| 0.8 | 4396.3 | 62.98% | **91.80%** | **91.84%** | 88.75% | 87.00% | 91.67% | **92.28%** | **92.23%** |
| 1 | 4535.0 | 66.91% | **95.15%** | **95.16%** | 91.14% | 92.00% | **95.95%** | **96.02%** | **96.10%** |
| 1.2 | 4673.7 | 70.39% | **94.91%** | **94.93%** | 91.38% | 94.69% | **96.12%** | **96.11%** | **96.11%** |
| 1.4 | 4765.1 | 74.26% | **97.00%** | **96.87%** | 92.86% | 95.20% | **97.13%** | **97.21%** | **97.22%** |
| $\alpha$ | | | | | (5,10) | | | | |
| 0.6 | 2862.7 | 84.17% | **94.29%** | **94.16%** | 90.32% | 85.83% | **95.22%** | **95.19%** | **95.13%** |
| 0.8 | 3250.2 | 90.20% | **94.56%** | **94.47%** | 90.77% | 91.11% | **95.50%** | **95.54%** | **95.45%** |
| 1 | 3633.9 | 90.88% | **95.24%** | **95.17%** | 91.67% | 93.20% | **95.17%** | **95.14%** | **95.16%** |
| 1.2 | 3696.0 | 95.20% | **97.20%** | **97.09%** | 92.92% | **97.16%** | **97.50%** | **97.56%** | **97.55%** |
| 1.4 | 3730.3 | **97.96%** | **97.97%** | **97.92%** | 93.55% | **98.16%** | **97.97%** | **97.85%** | **97.87%** |
| $\alpha$ | | | | | (10,20) | | | | |
| 0.6 | 2364.1 | 92.34% | 91.70% | 91.80% | 88.57% | **92.83%** | **93.37%** | **93.37%** | **93.43%** |
| 0.8 | 2755.7 | 93.34% | 94.63% | **94.67%** | 91.17% | 93.46% | **95.07%** | **95.11%** | **95.08%** |
| 1 | 2878.3 | **96.58%** | **96.82%** | **96.83%** | 92.70% | **96.52%** | **96.74%** | **96.65%** | **96.79%** |
| 1.2 | 2910.8 | **98.94%** | **98.97%** | **99.00%** | 94.40% | **99.05%** | **98.98%** | **99.07%** | **98.99%** |
| 1.4 | 2910.8 | **99.92%** | **99.87%** | **99.87%** | 94.95% | **99.88%** | **99.87%** | **99.90%** | **99.89%** |

RST17 also performs well ($84.13\% \sim 99.88\%$) and is the best algorithm in some of the simulation scenarios $(v_\emptyset^L, v_\emptyset^L) \in \{(0,0),(10,20)\}$. The de-randomization method from Algorithm 4 also uniformly improves on the corresponding randomized policy in almost all the scenarios and has better performance than RST17.

In Table 2.6, when the price difference is large, our algorithm performs uniformly well ($84.80\% \sim 99.87\%$), while CDLP does not ($67.26\% \sim 100.02\%$). Note that when $\alpha$ is large and $(v_\emptyset^L, v_\emptyset^L)$ are large, this corresponds to scenarios when inventory is too much. Myopic performs the best when inventory is too high, which is not surprising. For the rest of the scenarios, ALG2 and RST17 have the best performance, and sometimes one has better performance than the other. While it is difficult to say which algorithm performs the best, we remark that RST17 requires more information than ALG2, because RST17 needs to know the exact order in which customers $t$ arrive to solve the DP, while ALG2 only needs to know the universe of customers to solve the LP. Again, the de-randomization method from Algorithm 4 uniformly improves on the corresponding randomized policy in almost all the scenarios and has better performance than RST17.

Table 2.6: Computational performance in the non-stationary setting when the price difference is large

| | LP UB | Myopic | LP-Sol | ALG2 | ALG3 | RST17 | DeRLP | DeR2 | DeR3 |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | | | | | (0,0) | | | | |
| 0.6 | 48034.0 | 11.54% | 70.38% | **92.04%** | 69.82% | 87.43% | 84.06% | **90.88%** | **90.49%** |
| 0.8 | 50745.3 | 13.99% | 71.22% | **91.36%** | 71.16% | 87.03% | 86.35% | **91.99%** | 89.14% |
| 1 | 53456.7 | 16.15% | 71.47% | **89.96%** | 72.49% | 85.30% | 74.28% | **89.91%** | 88.36% |
| 1.2 | 54176.0 | 18.91% | 69.99% | 87.80% | 71.65% | **92.07%** | 82.56% | 87.56% | **88.78%** |
| 1.4 | 54895.3 | 22.73% | 69.25% | 84.80% | 71.21% | **92.41%** | 74.46% | 84.25% | **91.03%** |
| $\alpha$ | | | | | (1,5) | | | | |
| 0.6 | 36064.0 | 70.34% | 67.26% | 85.89% | 68.89% | **87.00%** | 71.36% | 85.82% | **87.97%** |
| 0.8 | 37835.7 | 70.06% | **84.14%** | 83.66% | 82.85% | **84.41%** | 87.79% | 83.59% | **92.26%** |
| 1 | 38195.4 | 71.51% | **87.89%** | 84.81% | 85.54% | 85.50% | 91.34% | 84.36% | **93.34%** |
| 1.2 | 38555.0 | 72.62% | **88.64%** | 86.90% | 86.62% | 86.30% | **94.42%** | 86.22% | 93.61% |
| 1.4 | 38831.1 | 73.61% | **88.88%** | 87.01% | 87.21% | 87.44% | 91.94% | 87.12% | **94.16%** |
| $\alpha$ | | | | | (5,10) | | | | |
| 0.6 | 28569.7 | 76.22% | 88.32% | 88.41% | 85.89% | **89.11%** | 90.74% | 89.03% | **93.34%** |
| 0.8 | 29574.9 | 87.57% | 88.32% | **88.39%** | 86.34% | **88.44%** | 94.87% | 88.65% | **96.43%** |
| 1 | 30580.1 | **96.76%** | 88.94% | 87.70% | 87.03% | 87.55% | **95.67%** | 87.58% | 95.06% |
| 1.2 | 30758.5 | 98.52% | **99.06%** | 89.08% | 94.47% | 89.45% | **98.79%** | 89.46% | 92.86% |
| 1.4 | 30855.2 | 98.95% | **99.30%** | **99.26%** | 94.26% | 90.88% | 99.25% | **99.41%** | 97.83% |
| $\alpha$ | | | | | (10,20) | | | | |
| 0.6 | 21379.9 | 83.74% | **88.02%** | 86.44% | 85.86% | **87.89%** | **94.24%** | 86.18% | 91.16% |
| 0.8 | 22275.8 | **97.58%** | 92.11% | 86.74% | 90.23% | 87.44% | 93.75% | 87.05% | **95.77%** |
| 1 | 22594.0 | **98.92%** | 98.81% | 91.26% | 94.00% | 89.01% | **99.02%** | 91.38% | 95.73% |
| 1.2 | 22676.2 | **99.80%** | 99.61% | 99.56% | 95.03% | 90.80% | **99.71%** | **99.77%** | 95.90% |
| 1.4 | 22676.2 | **99.96%** | **100.02%** | 99.87% | 95.20% | 92.55% | **100.00%** | **99.79%** | **100.13%** |

## 2.5 Conclusions

We proposed and analyzed a calendar pricing problem that a consumer packaged goods company favors given its operational convenience. We considered both single-item pricing and assortment (and pricing) controls. We showed that our policies are within 1-1/e (approximately 0.63) of the optimum under stationary demand and 1/2 of the optimum under non-stationary demand, with both guarantees approaching 1 if the starting inventory is large. Our techniques to analyze the best-possible performance guarantees are of theoretical interest per se. Finally, we fitted the real problem faced by the CPG company into the fractional demand setting of our model and demonstrated using data provided by the CPG company that our simple price calendars are effective. We also tested our simple policies and literature benchmarks on synthetic data, using the same numerical setup as in the literature.

# Chapter 3

# Network Revenue Management and Stochastic Packing under a Static Calendar

## 3.1 Introduction

As an extension to the previous Chapter, in this Chapter we focus on the network revenue management problem and the stochastic packing problem under a static calendar. The NRM problem is a (full-information) stochastic control problem which originates from the airline industry (Gallego and Van Ryzin 1997, Talluri and Van Ryzin 1998), and has been extensively studied in the revenue management literature (Jasin 2014, Adelman 2007, Topaloglu 2009, Ma et al. 2020b) with diverse applications. The stochastic packing problem is a more general form of the network revenue management problem.

**NRM Setup.** Let there be discrete, finite time horizon with $T$ periods. Time starts from period 1 and ends in period $T$. Let there be $n$ different products generated by $d$ different resources, each resource endowed with finite initial inventory $B_i$, $\forall i \in [d]$. Let $A = (a_{ij})_{i\in[d],j\in[n]}$ be the consumption matrix. Each entry $a_{ij} \in \mathbb{R}_+$ stands for the amount of inventory $i \in [d]$ used, if one unit of product $j \in [n]$ is sold. Each column

$A^j$ stands for product $j$'s "unit consumption vector" of different resources, and we assume each $A^j$ contains at least one nonzero entry. Let $A_i$ denote the $i$-th row of $A$. Let $a_{\mathsf{max}} = \max_{i,j} a_{ij}$ to be some bounded constant.

In each period $t$, a decision maker can post prices for the $n$ products by selecting a price vector from a finite set of $K$ price vectors $\{\boldsymbol{p}_1, ..., \boldsymbol{p}_K\}$, which we denote using $z_t \in [K]$. A price vector is $\boldsymbol{p}_k = (p_{1,k}, ..., p_{n,k})$, and $p_{j,k} \in [0, p_{\mathsf{max}}]$ is the price for product $j$ under $\boldsymbol{p}_k$. This captures situations where fixed price points have been pre-determined by market standards, *e.g.*, a common menu of prices that end in \$9.99: \$69.99, \$79.99, \$99.99. It also aligns with the stochastic packing setup, as we will introduce shortly below.

Given price (vector) $\boldsymbol{p}_k$, the demand for each product $j \in [n]$ is a distributionally known, bounded random variable, $Q_{j,k} := Q_j(\boldsymbol{p}_k) \in [0, 1]$, the distribution of which is known at the beginning of the entire horizon. Let $q_{j,k} := \mathbb{E}[Q_{j,k}]$ denote the mean demand for product $j$ under price $\boldsymbol{p}_k$, and $\mathbb{Q} = (Q_{j,k})_{j\in[n],k\in[K]}$, $\boldsymbol{q} = (q_{j,k})_{j\in[n],k\in[K]}$. For each unit of demand generated for product $j \in [n]$ under price vector $\boldsymbol{p}_k$, the decision maker generates $p_{j,k}$ units of revenue by depleting $a_{ij}$ units of each inventory $i \in [d]$. If no demand is generated, all the remaining inventory is carried over into the next period. The selling process stops immediately when the total cumulative demand of any resource exceeds its initial inventory; see Section 3.1.1 for more discussions. We use $\mathcal{I} = (T, \boldsymbol{B}, K, d, n, \boldsymbol{p}, A; \mathbb{Q})$ to stand for a $\mathsf{NRM}$ problem instance.

The objective of the decision maker is to maximize the expected total cumulative revenue (collected before exhausting the resources) over $T$ periods. The performance is measured by the *regret*, which is defined as the worst-case expected revenue loss compared with the DLP objective (whihc we will present in Section 3.2). We assume $T$ and $(B_i)_{i\in[d]}$ are in the same comparable order, and are much larger than any one of $(K, d, n, p_{\mathsf{max}}, a_{\mathsf{max}})$. see Section 3.1.1 for more discussions.

**<u>SP Setup.</u>** Similar to the $\mathsf{NRM}$ problem, let there be discrete, finite time horizon with $T$ periods. Time starts from period 1 and ends in period $T$. Unlike $\mathsf{NRM}$, there is no "product" nor "consumption matrix" in $\mathsf{SP}$. Let there be $d$ different resources, each endowed with finite initial capacity $B_i$, $\forall i \in [d]$.

In each period $t$, the decision maker pulls one arm from a finite set of $K$ distinct arms, which we denote using $z_t \in [K]$. Each time an arm $k \in [K]$ is pulled, a random reward $R_k \in [0, R_{\mathsf{max}}]$ is received at a random cost $C_{i,k} \in [0, C_{\mathsf{max}}]$ of each resource $i$, which we denote using the cost vector $\boldsymbol{C}_k \in [0, C_{\mathsf{max}}]^d$. The distributions of both the random reward and the random cost vector are fixed and distributionally known to the decision maker, and are known at the beginning of the entire horizon. The decision maker stops at the earlier time when one or more resource constraint is violated, or when the time horizon ends. We use $\mathcal{I} = (T, \boldsymbol{B}, K, d; \boldsymbol{R}, \boldsymbol{C})$ to stand for one instance of the problem.

### 3.1.1 Related Modeling Components

We survey the related modeling components that have appeared in the literature, including the stopping criterion and the regime for regret analysis.

#### Stopping Criterion.

At each point in time, as long as the remaining inventory for any resource is zero, the selling horizon stops. This stopping criterion is standard in the blind network revenue management and bandits with knapsacks literature when the distributional information is unknown and has to be sequentially learned, see Besbes and Zeevi (2012), Badanidiyuru et al. (2013). We refer to this stopping criterion as the "ungenerous" stopping criterion.

There is a second stopping criterion that is common in the revenue management literature when the stochastic distribution is known. This setup assumes time horizon never stops. Even if some resources are stocked-out, the decision maker continues to generate revenue from products that does not use the stocked-out resources. Either the admissible policy eliminates the possibility to sell stocked-out resources Gallego and Van Ryzin (1997), Rusmevichientong et al. (2020), or the realized demand in each period is simply the minimum between the remaining inventory and the generated demand Ma et al. (2020a). We refer to this stopping criterion as the "generous" stopping criterion.

Following each trajectory of randomness, the ungenerous stopping criterion stops earlier than the generous stopping criterion, hence the regret is larger.

**Regime for Regret Analysis.**

We aim to derive finite-time bounds on the regret of policies in terms of the number of time periods $T$. Following the literature, we assume $n, p_{\mathsf{max}}, a_{\mathsf{max}}, R_{\mathsf{max}}, C_{\mathsf{max}} = O(1)$. We adopt the following regret analysis regime: $T$ and $(B_i)_{i\in[d]}$ are in the same comparable order, and are much larger than $K$ and $d$. In other words, $B_i = \Theta(T)$ for all $i \in [d]$, while $K, d = o(T)$. This regime is standard in the network revenue management literature (sometimes stated in a different "asymptotic" fashion), *e.g.*, Gallego and Van Ryzin (1997), Liu and Van Ryzin (2008), Besbes and Zeevi (2012), Jasin (2014), Bumpensanti and Wang (2018), Ferreira et al. (2018), Chen et al. (2019), Chen and Shi (2019).

### 3.1.2 Overview of Results

Our techniques and results in this section are standard, yet provide a clean characterization of the revenue loss. In this section we present:

- It is well known that the Deterministic Linear Programs (DLP's) provide upper bounds on the revenue generated from any admissible policy. We define the DLP's in Section 3.2.

- We present lower and upper bounds on the regret in Sections 3.3 and 3.4, respectively. Combining both results, we show an intrinsic gap between linear regret and sublinear regret, depending on the switching budget. See Figure 3-1. When the DLP is non-degenerate, the number of resource constraints $d$ is the minimum required number of switches to achieve a sublinear regret; when the switching budget is strictly below $d$, a linear regret is inevitable.

- Algorithmically, the use of a discounting factor $\gamma$ and the action allocation rule in Section 3.4 borrows ideas from Section 2.2.5 in Chapter 2. And to prove the lower bound as shown in Section 3.3, the construction of the hard instances in Theorem 3.1 is novel.

For any problem instance $\mathcal{I} = (T, \boldsymbol{B}, K, d, n, \boldsymbol{p}, A; \boldsymbol{Q})$ or any $\mathcal{I} = (T, \boldsymbol{B}, K, d; \boldsymbol{R}, \boldsymbol{C})$, we adopt the general notation $\pi : \mathbb{R}^d \times [s] \times [T] \to \Delta([K])$ to denote any policy with the full information about stochastic distributions, which suggests a (possibly randomized) price vector to use given the remaining inventory, remaining switching budget, and the remaining periods. For any $s \in \mathbb{N}$, let $\Pi[s]$ be the set of policies that changes prices for no more than $s$ times on this problem instance $\mathcal{I}$. For any $s, s' \in \mathbb{N}$ such that $s \leq s'$, we know that $\Pi[s] \subseteq \Pi[s']$. Let $\Pi := \lim_{s \to +\infty} \Pi[s]$ be the set of policies with an infinite switching budget (which is the set of all possible policies). Let $\mathsf{Rev}(\pi)$ be the expected revenue that policy $\pi$ generates on this problem instance $\mathcal{I}$. Let $\pi^*[s] \in \arg\max_{\pi \in \Pi[s]} \mathsf{Rev}(\pi)$ be one of the optimal dynamic policies with switching budget $s$.

## 3.2 The Deterministic Linear Programs

For any problem instance $\mathcal{I} = (T, \boldsymbol{B}, K, d, n, \boldsymbol{p}, A; \boldsymbol{Q})$, the literature have extensively studied the following deterministic linear program (DLP) in the $\mathsf{NRM}$ setup. See Gallego and Van Ryzin (1997), Cooper (2002), Maglaras and Meissner (2006), Liu and Van Ryzin (2008).

$$\mathsf{J}^{\mathsf{DLP}} = \max_{(x_1,\ldots,x_K)} \sum_{k\in[K]} \sum_{j\in[n]} p_{j,k}\; q_{j,k}\; x_k \tag{3.1}$$

$$\text{s.t.} \sum_{k\in[K]} \sum_{j\in[n]} a_{ij}\; q_{j,k}\; x_k \leq B_i \qquad \forall\, i \in [d] \tag{3.2}$$

$$\sum_{k\in[K]} x_k \leq T \tag{3.3}$$

$$x_k \geq 0 \qquad \forall\, k \in [K] \tag{3.4}$$

It is well known that in the $\mathsf{NRM}$ setup, the above DLP serves as an upper bound on the expected revenue of any policy, even an optimal policy with an infinite switching budget (i.e., $\pi^*[\infty]$). It is well-known that the gap between the expected revenue obtained by the optimal policy and the DLP upper bound is bounded by $O(\sqrt{T})$ for all instances, i.e., $\mathsf{Rev}(\pi^*[\infty]) = \mathsf{J}^{\mathsf{DLP}} - O(\sqrt{T})$.

More generally, for any $\mathcal{I} = (T, \boldsymbol{B}, K, d; \boldsymbol{R}, \boldsymbol{C})$, the literature have studied the following generalization of the DLP, which we refer to as DLP-G.

$$\mathsf{J}^{\mathsf{DLP-G}} = \max_{(x_1,\dots,x_K)} \sum_{k\in[K]} r_k \;\; x_k \tag{3.5}$$

$$\text{s.t.} \sum_{k\in[K]} c_{i,k} x_k \le B_i \qquad \forall\, i \in [d] \tag{3.6}$$

$$\sum_{k\in[K]} x_k \le T \tag{3.7}$$

$$x_k \ge 0 \qquad \forall\, k \in [K] \tag{3.8}$$

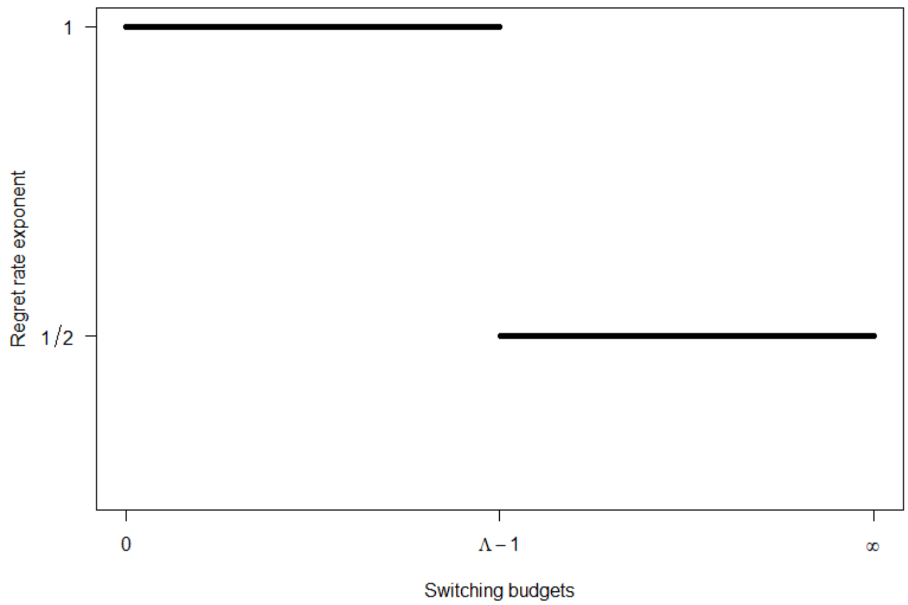
Since such a linear program is a packing LP, this generalization of the $\mathsf{NRM}$ problem is also referred to as the Stochastic Packing ($\mathsf{SP}$) problem.

Let the set of optimal solutions to the DLP be

$$X^* = \arg\max_{\boldsymbol{x}\in\mathbb{R}^K} \{(3.1)\,|(3.2), (3.3), (3.4) \text{ are satisfied}\}.$$

With a little abuse of notations, let the set of optimal solutions to the DLP-G be the same notation $X^* = \arg\max_{\boldsymbol{x}\in\mathbb{R}^K} \{(3.5)\,|(3.6), (3.7), (3.8) \text{ are satisfied}\}$. The distinction between DLP and DLP-G should be clear from the context. Let $\Lambda = \min\{\|\boldsymbol{x}\|_0\,|\boldsymbol{x} \in X^*\}$ be the least number of non-zero variables of any optimal solution. Let $\mathcal{X} = \arg\min\{\|\boldsymbol{x}\|_0\,|\boldsymbol{x} \in X^*\}$ be the set of such solutions. For any $\boldsymbol{x}^* \in \mathcal{X}$, let $\mathcal{Z}(\boldsymbol{x}^*) = \{k \in [K]\,|\boldsymbol{x}^* \neq 0\} \subseteq [K]$ be the subset of dimensions that are non-zero in $\boldsymbol{x}^*$. Note that $\Lambda$ is an instance-dependent quantity such that $\Lambda \le d+1$, where $d+1$ is the number of all constraints (resource constraints and time constraint) in the linear program. When DLP (or DLP-G) is non-degenerate, then equality holds

Figure 3-1: The intrinsic gap on the optimal regret $\mathsf{Rev}(\pi^*[\infty]) - \mathsf{Rev}(\pi^*[s])$ in the distributionally-known setup




Note: while the $\Theta(T)$ bound is tight for all $s < \Lambda - 1$, the $\widetilde{O}(\sqrt{T})$ bound shown for $s \geq \Lambda - 1$ is not necessarily tight for all $s \geq \Lambda - 1$; characterizing the exact rate of $\mathsf{Rev}(\pi^*[\infty]) - \mathsf{Rev}(\pi^*[s])$ for every $s \geq \Lambda - 1$ is an interesting future direction.

and $\Lambda = d + 1$.

In Sections 3.3 and 3.4 we show that for any problem instance, the instance-dependent quantity $(\Lambda - 1)$ is a critical switching budget — greater or equal to which the regret is in the order of $\widetilde{O}(\sqrt{T})$, and below which the regret is in the order of $\Theta(T)$. Combining the above results, we show that there is an intrinsic gap on the regret, if one more critical price change is allowed. See Figure 3-1.

## 3.3 Lower Bounds

In this section we show that when the switching budget is below $\Lambda - 1$ (at most $\Lambda - 2$), then a linear regret rate is inevitable. Recall that $\Pi[\Lambda - 2]$ stands for the family of admissible policies that make no more than $\Lambda - 2$ changes.

**Theorem 3.1.** *For any problem instance $\mathcal{I} = (T, \boldsymbol{B}, K, d, n, \boldsymbol{p}, A; \boldsymbol{Q})$ under the* NRM *setup, or any problem instance $\mathcal{I} = (T, \boldsymbol{B}, K, d; \boldsymbol{C}, \boldsymbol{R})$ under the* SP *setup, there is* *an associated* $\Lambda$ *number (defined in Section 3.2). Any policy* $\pi \in \Pi[\Lambda - 2]$ *earns an expected revenue (or reward):*

$$\mathsf{Rev}(\pi) \leq \mathsf{J}^{\mathsf{DLP}} - c \cdot T \qquad \textit{or} \qquad \mathsf{Rev}(\pi) \leq \mathsf{J}^{\mathsf{DLP-G}} - c \cdot T$$

*where* $c > 0$ *is some (distribution-dependent) constant independent of* $T$*. Combined with the known fact that* $\mathsf{Rev}(\pi^*[\infty]) \geq \mathsf{J}^{\mathsf{DLP}} - O(\sqrt{T})$ *or* $\mathsf{Rev}(\pi^*[\infty]) \geq \mathsf{J}^{\mathsf{DLP-G}} - O(\sqrt{T})$*, it holds that*

$$\mathsf{Rev}(\pi) \leq \mathsf{Rev}(\pi^*[\infty]) - \Omega(T),$$

*i.e., the regret scales linearly with* $(T, \boldsymbol{B})$ *when other parameters are fixed.*

We outline three key steps here and defer the details of our proof to Appendix B.1. We first identify a clean event, such that the realized demands are close to the expected demands that the LP suggests. This clean event happens with high probability $(1 - \frac{2}{T^3})$. In the second step, conditioning on such event, the maximum amount of revenue we generate is no more than $O(\sqrt{T})$ compared to what the LP suggests; and the minimum amount of inventory demanded is no less than $O(\sqrt{T})$ compared to what the LP suggests, resulting in no more than $O(\sqrt{T})$ of realized revenue. In the third step, we show that the regret from insufficient price changes scales in the order of $\Omega(T)$, which dominates the $O(\sqrt{T})$ amount revenue due to randomness. Such clean event analysis, originating from the online learning literature to prove upper bounds (Badanidiyuru et al. 2013, Lattimore and Szepesvári 2018, Slivkins 2019), was recently used in Arlotto and Gurvich (2019) to prove lower bounds.

The lower bound established in Theorem 3.1 is a *per-instance* lower bound, as it holds for every single problem instance. Such a result is much stronger than the worst-case type lower bounds that are widely considered in the revenue management literature.

## 3.4 Upper Bounds

In this section we show that when the switching budget is greater or equal to $\Lambda - 1$, then the regret is $\widetilde{O}(\sqrt{T})$. Such a sub-linear guarantee is achieved by tweaking the well-known static control policy in the network revenue management literature.

The static control policy (Gallego and Van Ryzin 1997, Cooper 2002, Maglaras and Meissner 2006, Liu and Van Ryzin 2008) achieves a similar $O(\sqrt{T})$ regret in a similar setup, when the selling horizon stops immediately when the total cumulative demand of any resource exceeds its initial inventory. In our setup the stopping criterion is different[1], which requires slightly different techniques. The static control policy was also used in Besbes and Zeevi (2012) to prove a similar result in the NRM setup under the same stopping criterion. However, their proof technique critically requires the maximum price to be bounded, which does not generalize to the SP setup.

We tweak the static control policy, so that with high probability the selling horizon never stops earlier than the last period $T$. See Algorithm 7. This is achieved by selecting the value of $\gamma$ in the first step of Algorithm 7. Similar ideas have been used in Hajiaghayi et al. (2007), Ma et al. (2020a), Balseiro et al. (2019) to prove asymptotic results in different setups.

We explain the third step permutation. Suppose $\mathcal{Z}(\boldsymbol{x}^*) = \{1, 3, 4\}$. In this case, $\Lambda = 3$ and there are 6 permutations. There are 6 possible policies as suggested in Algorithm 7. While some of these policies may have better empirical performance than others, they all achieve $\widetilde{O}(\sqrt{T})$ regret.

**Theorem 3.2.** *Any policy $\pi$ as defined in Algorithm 7 satisfies $\pi \in \Pi[\Lambda - 1]$ and earns an expected revenue (or reward):*

$$\mathsf{Rev}(\pi) \geq \mathsf{J}^{\mathsf{DLP}} - c\sqrt{T \log T} \geq \mathsf{Rev}(\pi^*[\infty]) - c\sqrt{T \log T}$$
$$\textit{or} \;\; \mathsf{Rev}(\pi) \geq \mathsf{J}^{\mathsf{DLP-G}} - c\sqrt{T \log T} \geq \mathsf{Rev}(\pi^*[\infty]) - c\sqrt{T \log T},$$

[1]Following each trajectory of randomness, the ungenerous stopping criterion stops earlier than the generous criterion, hence the regret is larger. As a result, our upper bound under the ungenerous stopping criterion is not a direct implication of Gallego and Van Ryzin (1997).

**Algorithm 7** Tweaked LP Policy

**Input:** $\mathcal{I} = (T, \boldsymbol{B}, K, d, n, \boldsymbol{p}, A; \boldsymbol{Q})$ in the $\mathsf{NRM}$ setup,
or $\mathcal{I} = (T, \boldsymbol{B}, K, d; \boldsymbol{R}, \boldsymbol{C})$ in the $\mathsf{SP}$ setup.

**Policy:**

1: Define $\gamma = 1 - 2\frac{a_{\mathsf{max}}}{B_{\mathsf{min}}}\sqrt{nT\log T}$ in the NRM setup; or $\gamma = 1 - \frac{2C_{\mathsf{max}}}{B_{\mathsf{min}}}\sqrt{T\log T}$ in the $\mathsf{SP}$ setup.

2: Solve the DLP as defined by (3.1), (3.2), (3.3), and (3.4). Find an optimal solution with the least number of non-zero variables, $\boldsymbol{x}^* \in \mathcal{X}$.

3: Arbitrarily choose any permutation $\sigma : [\Lambda] \to \mathcal{Z}(\boldsymbol{x}^*)$ from all $(\Lambda)!$ possibilities.

4: Execute, depending on the problem setup:

In the $\mathsf{NRM}$ setup, set the price vector to be $\boldsymbol{p}_{\sigma(1)}$ for the first $\gamma \cdot x^*_{\sigma(1)}$ periods, then $\boldsymbol{p}_{\sigma(2)}$ for the next $\gamma \cdot x^*_{\sigma(2)}$ periods, ..., and finally $\boldsymbol{p}_{\sigma(\Lambda)}$ for the last $T - \gamma \cdot \sum_{l=1}^{\Lambda-1} x^*_{\sigma(l)}$ periods (we assume that $x_k, \forall k \in [K]$ are integers, because rounding issues incur a regret of at most $(d \cdot \max_k \boldsymbol{p}_k^{\mathsf{T}} \cdot \bar{q}_k)$, which is negligible compared with $\sqrt{T}$).

In the $\mathsf{SP}$ setup, pull arm $\sigma(1)$ for the first $\gamma \cdot x^*_{\sigma(1)}$ periods, then $\sigma(2)$ for the next $\gamma \cdot x^*_{\sigma(2)}$ periods, ..., and finally $\sigma(\Lambda)$ for the last $T - \gamma \cdot \sum_{l=1}^{\Lambda-1} x^*_{\sigma(l)}$ periods.

*where $c > 0$ is some constant independent of $T, K, d, \mathbb{Q}$.*

The above upper bound uniformly holds for all instances. We outline two key steps here and defer the details of our proof to Appendix B.2. In the first step, we show that with high probability, the selling horizon never stops earlier than the last period $T$. Second, conditioning on this high probability event, the expected revenue is at least $\gamma$ fraction of the LP objective. Combining the two steps together we know that the total regret is upper bounded by $(1-\gamma)\mathsf{J}^{\mathsf{DLP}}$ or $(1-\gamma)\mathsf{J}^{\mathsf{DLP-G}}$.

# Chapter 4

# Online Knapsack Using a Static Threshold

## 4.1 Introduction

Consider the following problem. There is a knapsack of size 1 and an unknown sequence of items with sizes at most 1. The items arrive one-by-one, and each item must be irrevocably either packed into the knapsack or discarded upon arrival. An item can be packed only if its size does not exceed the remaining knapsack capacity. The goal is to maximize the sum of sizes of packed items, i.e. maximize the total capacity filled.

The decision of whether to accept each item into the knapsack is made by an *online algorithm*, which does not know the sizes of future items, nor the number of future items. Meanwhile, for any sequence of items, one could consider its *optimal offline packing* knowing the entire sequence in advance. For $c \leq 1$, a fixed (but possibly randomized) online algorithm is said to be *c-competitive* if on any sequence, its (expected) capacity packed is at least $c$ times the optimal offline packing. We are interested in the highest-possible value of $c$, which is called the *competitive ratio*.

For this problem, randomization is necessary to achieve any non-trivial competitive ratio. Indeed, a deterministic algorithm, when faced with an initial item of a small size $\varepsilon > 0$, must either accept or reject. If it accepts, then it achieves a poor ratio when the item is followed by an item of size 1, packing size $\varepsilon$ when the optimal packing has size 1. On the other hand, if it rejects, then it achieves a poor ratio when the sequence ends after the first item, packing 0 when the optimum is $\varepsilon$.

With randomization, a simple idea, originally due to Han et al. (2015), yields a $\frac{1}{2}$-competitive algorithm. First, a fair coin is flipped. If Heads, then the algorithm greedily packs any item that fits. If Tails, then the algorithm rejects all items until the first one that the greedy policy *would not have fit*, and starts to greedily accept items from that item (including that item). In expectation, this algorithm is $\frac{1}{2}$-competitive, since either the greedy packing (which the algorithm mimics half the time) is optimal, or the algorithm's sum of capacity packed under the two outcomes exceeds 1, and hence its expected packing exceeds $\frac{1}{2}$, while the optimum is at most 1. Furthermore, it follows from the example above that $\frac{1}{2}$ is the competitive ratio for the class of all randomized algorithms.

### 4.1.1 Motivation for this Paper: Threshold Policies

In this paper, we derive the competitive ratio for a subclass of algorithms: (random) threshold algorithms. Threshold algorithms initially draw a threshold $\tau \in [0, 1]$, and then accept every item of size at least $\tau$ that fits, but never change the threshold throughout the entire horizon after its initial draw. When $\tau = 0$, the algorithm mimics the greedy algorithm.

Threshold algorithms constitute a natural subclass of algorithms with many benefits, as we outline below.

- **Simplicity**: First, threshold algorithms are logistically easy to implement, making non-adaptive accept/reject decisions that do not require recording the history of past items. We can use them to derive a random-threshold algorithm for a generalization of our problem to multiple knapsacks, as we discuss in Section 4.1.3.

- **Applicability**: Second, a threshold algorithm is characterized by a CDF $F$ for the threshold $\tau$, which has a simple interpretation for how to implement this randomized algorithm in practice. We implement simulations on an industry partner's data, as we discuss in Section 4.1.4.

- **Incentive-compatibility**: Most importantly, threshold algorithms treat identical items equally, in a first-come-first-serve order, which implies that the items truthfully represent the demands for the knapsack.

  We elaborate on this incentive issue. In many applications, an arriving "item" corresponds to an order placed by a customer. Under a threshold algorithm, a customer is incentivized to place a *single* order for her desired amount, *immediately* upon her arrival. Indeed, there is no benefit to waiting since orders exceeding $\tau$ are first-come-first-served; moreover, should the customer's order get rejected, there is no benefit to trying again later. Therefore, the sequence of items observed truthfully represents the desires of the customers, in order.

  By contrast, in the randomized algorithm described above (Han et al. 2015), a customer can easily manipulate the system. For example, if the first customer sees her order get rejected (because the coin in the algorithm landed Tails), then she can repeatedly try placing the same order again. After generating sufficiently many "fake" orders, the greedy policy cannot fit all the orders, and hence the algorithm will accept her order.

- **Fairness**: As a final note, our random-threshold algorithms "treat similar individuals similarly", which is the definition of fairness proposed in Dwork et al. (2012). In particular, the probability of our random-threshold accepting an item (which fits) is dependent on only the size of that item, and moreover, this probability changes smoothly with respect to a change in size. Another notion of fairness in sequential decision-making was introduced by Gupta and Kamble (2019). Our random-threshold algorithms also satisfy their definition.

### 4.1.2 Techniques for Analyzing Threshold Policies

The analysis for the subclass of random threshold algorithms also becomes move involved, since a small change in $\tau$ could have a *ripple effect* on the items that fit and hence the items that are packed by the policy. We derive the best-possible CDF's for threshold $\tau$, and hence the tight competitive ratio for threshold algorithms, under two different definitions of the optimal offline packing:

1. A $\frac{3}{7} \approx 0.428$-competitive random-threshold distribution, relative to the optimal *fractional* packing;

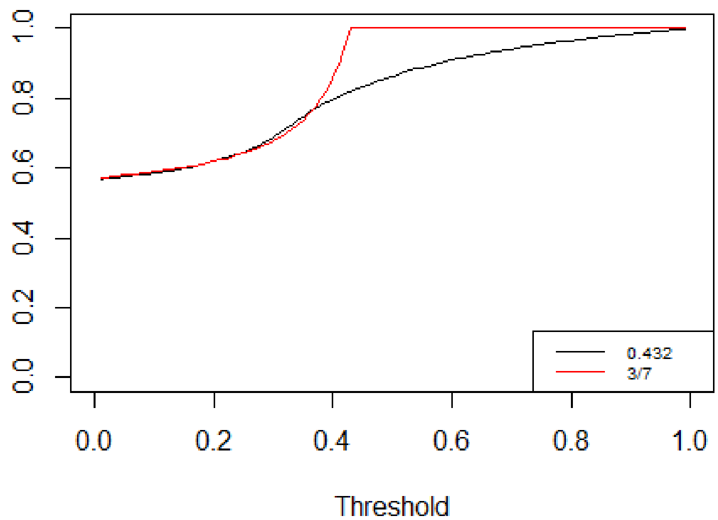
2. A $\approx 0.432$-competitive random-threshold distribution, relative to the optimal *integer* packing.

While both optima know the set of items in advance, the difference between them is that a fractional packing can potentially "truncate" items to obtain a perfect packing of size 1. Interestingly, there is a *separation* between the competitive ratio relative to the stronger, fractional optimum and that relative to the weaker, integer optimum. Moreover, there is a surprising difference between the two random-threshold algorithms used to achieve these competitive ratios (see Figure 4-1). In particular, the threshold for the $\frac{3}{7}$-competitive algorithm never exceeds $\frac{3}{7}$, while the threshold for the 0.432-competitive algorithm has positive support on all of [0,1]. By contrast, for arbitrary randomized algorithms, it follows from Han et al. (2015) that both competitive ratios are $\frac{1}{2}$.

We now describe our techniques for establishing results 1–2 above. We first start with the following randomized algorithm, which is neither of the two algorithms described above. This algorithm flips an initial coin. With probability $2/3$, the algorithm greedily accepts any item which fits in the knapsack. With probability $1/3$, the algorithm accepts only the first item to have size at least $1/2$ (if such an item exists).

We claim that this simple algorithm yields a constant competitiveness guarantee of $1/3$. To see why, first note that if the greedy policy can fit all the items, then it is optimal, and since the algorithm is greedy with probability $2/3$, it would be at least

Figure 4-1: CDF functions of the thresholds from two random threshold algorithms




2/3-competitive. Therefore, suppose that the greedy policy cannot fit some items, and consider two cases. If the sequence contains no items of size at least $1/2$, then the greedy policy must have packed size greater than $1/2$ by the time it could not fit an item, and hence the algorithm packs expected size at least $2/3 \times 1/2 = 1/3$. In the other case, let $m$ denote the size of the *first* item to have size at least $1/2$. When the algorithm is not greedy, it packs size $m$; and when it is greedy, it packs size at least $\min\{m, 1-m\}$, which equals $1-m$ because $m \geq 1/2$. In expectation, the algorithm packs size at least

$$\frac{1}{3}m + \frac{2}{3}(1-m) = \frac{2}{3} - \frac{1}{3}m \geq \frac{1}{3}.$$

Since the algorithm in both cases packs size at least $1/3$, and the optimal offline packing cannot exceed 1, this completes the claim that the algorithm is $1/3$-competitive.

Now, note that the previous algorithm effectively sets a random threshold whose distribution is 0 with probability $2/3$, and $1/2$ with probability $1/3$. To improve upon it, we consider an arbitrary distribution for the threshold $\tau$ given by the CDF $F(x) = \Pr(\tau \leq x)$, and generalize the above analysis. We now let $m$ denote the size of the *smallest* item which the greedy policy does not fit. In the case where $m < 1/2$, we use similar arguments as above to deduce that the algorithm packs expected size at least

$$F(0)(1-m) + (F(m) - F(0)) \min\{m, 1-m\}. \tag{4.1}$$

However, the other case where $m \geq 1/2$ is more challenging, because when $m = 1$ both terms in (4.1) equal 0. To refine the analysis, we define $q$ to be the *maximum* number such that, at the time of arrival of the item of size $m$, it would not fit even if we could "magically discard" every accepted item of size less than $q$. By the maximality of $m$, there must exist an item of size $q$. After carefully analyzing the cases (including the one where $q > m$), we show that the algorithm's expected packing size is minimized in the case where it equals $q$ when the threshold is at most $q$, and $1 - q + \varepsilon$ (for an arbitrarily small $\varepsilon > 0$) when the threshold is greater than $q$. Therefore, it is lower-bounded by

$$F(q)q + (1 - F(q))(1 - q). \tag{4.2}$$

Finally, we solve for the maximum $c$ at which there exists a threshold distribution $F$ such that both expressions (4.1) and (4.2) exceed $c$ (for all $m$ and $q$). This turns out to be $c = 3/7 \approx 0.428$, and since the optimal *fractional* packing cannot exceed 1, the corresponding random-threshold algorithm is 0.428-competitive relative to the stronger, fractional optimum, as shown in **Theorem 4.2**. This competitiveness is tight relative to the stronger optimum, as shown in **Theorem 4.3**.

In **Theorem 4.4**, we improve the competitiveness to 0.432 relative to the weaker, integer optimum. The previous analysis with expressions (4.1) and (4.2) is no longer tight, because it merely lower-bounded the algorithm's expected packing size without considering the consequences on the optimal integer packing. To improve upon the previous distribution, we perturb it to have a positive mass on all of [0,1] (instead of never setting a threshold above 3/7, as shown in Figure 4-1). Intuitively, this prevents the adversary from making the optimal packing size always 1 by appending a size-1 item to the end of any sequence, because if he did, then there will always be a positive probability that the algorithm sets a threshold high enough to get the size-1 item. In fact, this perturbed threshold distribution, which yields a 0.432-competitive algorithm, is best-possible for threshold algorithms relative to the optimal integer packing, as shown in **Theorem 4.5**.

### 4.1.3 Generalization to Static Policies for Multiple Knapsacks

We generalize to an assignment problem for multiple knapsacks, defined as follows. In this setting, there are multiple knapsacks, with potentially different capacities. An arriving item takes up a potentially different size in each knapsack, and must be irrevocably assigned to a knapsack where it fits, or outright rejected. The objective is to maximize the sum of capacities packed across all the knapsacks.

For this problem, we derive a randomized $\frac{3}{14} \approx 0.214$-competitive algorithm using a static policy which, again, does not require recording the history of past items. In particular, each item is first routed to the knapsack where it takes the greatest size (without knowing whether it would get accepted), and then an independent threshold policy at that knapsack controls whether to accept the item. This greedy routing policy is a simple implementation of threshold policies with multiple knapsacks, and is shown to be $\frac{1}{2} \times \frac{3}{7} = \frac{3}{14}$-competitive relative to the stronger optimum in **Theorem 4.7**, assuming that each knapsack's threshold is chosen randomly according to the CDF which is $\frac{3}{7}$-competitive for a single knapsack. Interestingly, the CDF which improves the competitiveness to 0.432 for a single knapsack, relative to the weaker optimum, does not appear to translate to a competitiveness result for multiple knapsacks. Also, for multiple knapsacks we derive an upper bound of 0.461 on the competitive ratio for arbitrary randomized algorithms, in **Theorem 4.8**. This shows that the tight $\frac{1}{2}$-competitiveness for a single knapsack does not hold with multiple knapsacks, even if one could go beyond static policies.

To our knowledge, we are the first to study our generalized online assignment problem with multiple unit-density knapsacks, and derive constant-factor competitiveness results. However, our problem and results are closely related to two problems from the literature, as we outline below.

- **AdWords**: Our only difference from the AdWords problem studied in online advertising is that items cannot be "truncated". That it, in AdWords, an item can be assigned to a knapsack where it exceeds the remaining capacity, and have its size "reduced" to fit the knapsack. With this truncation allowed, Mehta et al. (2005) derive a $\frac{1}{2}$-competitive algorithm in general, and a best-possible $(1 - \frac{1}{e})$-competitive algorithm under the "small bids" assumption that initial capacities ("advertiser budgets") are large compared to the sizes an item can take ("bids").

  The competitiveness guarantee in our setting relative to the fractional optimum can only be worse than the $\frac{1}{2}$-competitiveness for AdWords, since the offline optimum is the same (being able to truncate using fractions), while the online algorithm is restricted from truncating. Our 0.214-competitive result can be seen as a weaker guarantee which holds for weaker online algorithms.

- **Appointment Scheduling in Healthcare**: Our only difference from the appointment scheduling problem of Stein et al. (2018) is that the sequence of items are completely unknown, instead of drawn independently from known distributions. That is, our arrival sequence can be seen as "adversarial" instead of "stochastic". Stein et al. (2018) derive a 0.321-competitive algorithm in the stochastic setting, where the definition of competitiveness takes an expectation over the arrival sequence when evaluating both the online algorithm and the offline optimum.

  The adversarial competitive ratio can only be worse than the stochastic one. Our 0.214-competitive result can be seen as a weaker guarantee which holds in a more general setting.

Our aforementioned results, including those for a single knapsack, are summarized in Table 4.1. Note that our Theorem 4.7 also implies a $\frac{1}{4}$-competitiveness guarantee for general multi-knapsack algorithms.

Table 4.1: Summary of lower and upper bounds on the competitive ratio, for different classes of algorithms, relative to different optima, in different settings

| | Relative to Stronger Optimum | Relative to Weaker Optimum |
|---|---|---|
| Single Knapsack | | |
| random-threshold algorithms | [0.428 (**Thm. 4.2**), 0.428 (**Thm. 4.3**)] | [0.432 (**Thm. 4.4**), 0.432 (**Thm. 4.5**)] |
| arbitrary algorithms | [0.5 (Han et al. 2015), $\longrightarrow$] | [$\longleftarrow$, 0.5 (Han et al. 2015)] |
| Multiple Knapsacks | | |
| static algorithms | [0.214 (**Thm. 4.7**), $\longrightarrow$] | [$\longleftarrow$, $\downarrow$] |
| arbitrary algorithms | [0.25 (**Thm. 4.7**), $\longrightarrow$] | [$\longleftarrow$, 0.461 (**Thm. 4.8**)] |
| AdWords (truncation allowed) | [$0.5^a$ (Mehta et al. 2005), $\longrightarrow$] | [$\longleftarrow$, 0.632 (Karp et al. 1990)] |
| Scheduling (stochastic arrivals) | [0.321 (Stein et al. 2018), $\longrightarrow$] | [$\longleftarrow$, 0.5 (Stein et al. 2018)] |

Note: Results from this paper are **bolded**. An arrow indicates that the best-known lower (resp. upper) bound is implied by that from a more restricted (resp. less restricted) setting, pointing in the direction of that setting. Note that our paper is the only one to establish a *separation* between the competitiveness relative to the two different optima.

[a] Improves to $1 - 1/e \approx 0.632$ under the small bids assumption.

### 4.1.4 Simulations Using Supply Chain Data of A Latin American Chain Department Store

We now describe how our optimal random-threshold distributions for a single knapsack can be implemented across the supply chain of our industry partner, a Latin American chain department store. They sell 974 SKU's in the young women's fashion category. There are 21 warehouses, and every SKU is stored in a subset of different warehouses. Every (SKU, warehouse)-pair faces a stream of orders, each for a specific number of units. Orders cannot be split or redirected to a different warehouse, so order sizes greater than the available inventory must be rejected. Therefore, our industry partner faces the same accept/reject problem on order sizes, and has the same goal of maximizing total inventory fulfilled, equal to the sum of sizes of accepted orders.

The data we observe is the sizes of all orders *accepted* by a greedy First-Come-First-Serve (FCFS) policy. The sum of all observed order sizes for each of our (SKU, warehouse)-pairs is then at most the starting inventory, since the order sizes that cannot be fulfilled have been censored. To create non-trivial instances, we re-scale the starting inventory amounts (which we know) for each SKU at each warehouse by a factor $\alpha \in [0, 1]$, and test the performance of different accept/reject policies over different scaling factors $\alpha$.

To implement our random-threshold policies, we take 21 evenly-spaced percentiles of the threshold distribution $F$, that is, we take the 21 thresholds defined by $F^{-1}(0)$, $F^{-1}(0.05)$, ..., $F^{-1}(0.95)$, $F^{-1}(1)$. Then we assign them (i.e. randomly permute them) over the 21 warehouses, making accept/reject decisions at each warehouse based on the assigned threshold (scaled by the starting inventory). We believe this assignment of percentiles to warehouses is how our threshold distribution $F$ would be implemented in practice. We then average the fulfillment ratios over the warehouses to determine the performance for a specific SKU. We take an outer average over many independent random permutations of warehouses to define a final performance ratio for each of the 974 SKU's.

We find that the greedy FCFS policy has the best average-case performance ratio, even when the scaling factor $\alpha$ for initial inventory is small. While this is discouraging, we believe that the way in which order sizes are censored in our data favors FCFS, since large orders cannot come at the end. Nonetheless, for any $\alpha$, if we look at the worst-case SKU, then our random-threshold policy has the best performance. Indeed, the way in which it distributes different thresholds over the warehouses provides a form of "hedging" for each SKU, and our random-threshold policy being robust to the worst case is consistent with it having the best competitive ratio. This robustness is not achieved by FCFS, the algorithm of Han et al. (2015), or even any deterministic-threshold algorithm.

### 4.1.5 Other Related Work and Applications

To the best of our knowledge, we are the first to use threshold policies to study the competitive ratios of randomized algorithms for this foundational unit-density[1] online knapsack problem. Without the unit-density assumption, the non-existence of any constant competitive ratio guarantee $c > 0$, even for randomized algorithms on a single knapsack, was first established in Marchetti-Spaccamela and Vercellis (1995). Tight *instance-dependent* competitive ratios (where the guarantee $c$ can depend on

[1]In our problem, since the objective is capacity packed, the reward from packing each item is equal to its size, and hence the term "unit-density".

parameters based on the sequence of items) have also been established in Zhou et al. (2008). For a thorough discussion of recent results across many variants online knapsack, we refer to Cygan et al. (2016).

There is also a rich literature which studies the stochastic online knapsack problem, that assumes the arrival sequences are drawn from a given distribution. There are papers on the optimal policies on a single knapsack; see Kleywegt and Papastavrou (1998), Papastavrou et al. (1996) when the order of arriving items is fixed. When the items can be inserted in any order but their sizes are stochastic, the concept of "adaptivity gap" between adaptive and non-adaptive algorithms was proposed in Dean et al. (2008). Using their language, the adaptivity gap for our problem on a single knapsack is within 7%; see Table 4.1. Variants where the arrival sequence could be (partially) learned over time are studied in Modaresi et al. (2019), Hwang et al. (2018).

We briefly mention some other applications areas (other than Adwords and healthcare scheduling) where online knapsack/assignment problems with *unit density* and *no truncation* arise.

**Refugee Integration**: Bansak et al. (2018) have studied a refugee integration problem. This is a real problem in many developed democracies, where "refugees face challenges integrating into host societies". Refugees in non-splittable groups (e.g. families) arrive in an online fashion, and democracies assign refugees across resettlement locations subject to capacity constraints. If a group of refugees do not have a suitable settlement location, they temporarily stay in the refugee camps. The objective is to maximize the number of assigned refugees.

**Crowdsourcing**: Ho and Vaughan (2012) have studied a problem in online crowdsourcing, where a requester asks workers that arrive online to finish his / her tasks, and cannot split tasks into two. Each worker spends some time to finish the assigned work. The objective is to maximize the total benefit that the requester obtains from the completed work, given time constraints. In a variant (Assadi et al. 2015), each worker picks a subset of tasks, along with task-specific bid numbers. The requester has to assign no more than one task to each worker, by paying the worker the on the bid. The objective of the requester is to either maximize the number of assigned tasks to workers, while not violating the budget constraint.

### 4.1.6 Roadmap

In Section 4.2 we introduce the model and notations. In Section 4.3 we introduce our results on a single knapsack. Section 4.3.1 introduces the 3/7 competitive algorithm relative to the optimal fractional packing, and Section 4.3.2 introduces the 0.432 competitive algorithm relative to the optimal integer packing. Then in Section 4.4 we introduce our results on multiple knapsacks. Section 4.4.1 introduces the 0.214 competitive algorithm, and Section 4.4.2 introduces the impossibility result for a 0.461 competitive algorithm. Finally in Section 4.5, we conduct computational study using real data from a Latin American chain department Store, and show the efficacy of threshold algorithms.

## 4.2 Definition of Problems, Notations

In this paper we denote $[T] = \{1, 2, ..., T\}$, for any positive integer $T$. Let the capacity of the knapsack be 1. Let the entire set of items be indexed by $t \in [T]$, the sequence of its arrival. For any $t \in [T], s_t$ refers to the size of item $t$. The entire sequence of item sizes is then $S = (s_1, s_2, ..., s_T)$. For any $A \subseteq [T]$, a subset of indices, let $\mathsf{size}(A) = \sum_{t \in A} s_t$ be the total size of items in $A$.

Suppose there is a clairvoyant decision maker who knows the entire sequence in advance. This decision maker is going to take the *optimal* actions (accept / reject) over the process. Let this policy be $\mathsf{OPT}$. Note that $\mathsf{OPT}$ does not necessarily guarantee to fill all the capacity of the knapsack, but it must be upper bounded by 1.

For any specific sequence of $S$, let $\mathsf{ALG}(S)$ denote the total amount filled by $\mathsf{ALG}$ on this instance *in expectation*, where expectation is taken over the randomness of the algorithm. Here $\mathsf{ALG}$ is any generic algorithm, where in the following sections we will specify which algorithm it is by using slightly different notations for each algorithm. Let $\mathsf{OPT}(S)$ denote the total amount filled by $\mathsf{OPT}$ on this sequence. Note that

$\mathsf{OPT}(S) = \max_{A \subseteq [T]:\mathsf{size}(A) \leq 1} \mathsf{size}(A)$. We will also refer to a stronger optimum $\mathsf{OPT}^+$ which is not only clairvoyant, but allowed to truncate items at will, with

$$\mathsf{OPT}^+(S) = \min\{s_1 + \cdots + s_T, 1\} \tag{4.3}$$

It is self-evident that $\mathsf{OPT}(S) \leq \mathsf{OPT}^+(S)$ for any $S$. We also use $\mathsf{ALG}, \mathsf{OPT}$ and $\mathsf{OPT}^+$ for $\mathsf{ALG}(S), \mathsf{OPT}(S)$, and $\mathsf{OPT}^+(S)$, respectively, if the sequence $S$ is clear from the context.

Under any policy, we say that an item $s_t$ is *rejected* because it fails to meet the admission critrion of this policy, e.g. failure to exceed the threshold of a threshold policy. If an item is rejected under a policy, we say that this policy rejects this item.

Under any policy, we say that an item $s_t$ is *blocked* at the moment it arrives, if the remaining capacity of the knapsack is not enough for $s_t$ to fit in. An item is said to be blocked regardless of the fact if it would have been rejected by the policy. If an item is blocked under a policy, we say that this policy blocks this item.

The focus of this paper is on randomized (non-adaptive) threshold algorithms.

We define *threshold* algorithms as follows: Let $\mathsf{THR}(\tau), \forall \tau \in [0, 1]$ be a threshold algorithm that accepts any item whose size is greater or equal to $\tau$, as long as it can fit into the knapsack. A $\mathsf{THR}(0)$ policy is also referred to as a greedy policy, $\mathsf{Greedy}$: accept any item regardless of its size, as long as it can fit into the knapsack. We will interchangeably use $\mathsf{THR}(0)$ and $\mathsf{Greedy}$ for the same policy.

We say that threshold algorithms are *non-adaptive*, because the decision of whether to accept an item (assuming it fits) is dependent on only the item's size, and not the past items observed. Note that a threshold algorithm can be *randomized*, in which case $\tau$ is chosen from a probability distribution at the start and then fixed over time.

## 4.3 A Single Knapsack

We first start with a single knapsack. We introduce the algorithm from Han et al. (2015) here.

**Definition 4.1.** [Algorithm $\mathsf{TwoBins}$, Han et al. (2015)]

1. Randomly flip a fair coin, with probability 1/2 on each side.
2. If Heads, apply $\mathsf{Greedy}$.
3. If Tails, reject everything until the first one that $\mathsf{Greedy}$ would have not fit. Then apply $\mathsf{Greedy}$ (including this item).

Note that when Tails, the algorithm has to adapt on what are the items that have been rejected. So unlike our algorithms, this is an *adaptive* algorithm. Nonetheless, it provides the best possible competitive ratio.

**Proposition 4.1** (Theorem 1, Han et al. (2015))**.** *The algorithm from Definition 4.1 is* $1/2$ *competitive, and it is tight for any algorithm, i.e.* $\forall S$,

$$\inf_{S} \frac{\mathsf{TwoBin}(S)}{\mathsf{OPT}^{+}(S)} = \frac{1}{2}.$$

The proof is very simple, as we outlined in the Introduction. Either greedy is optimal, in which case the algorithm is optimal half the time, or the sum of the algorithm's packing under Heads and Tails exceeds 1, in which case the algorithm's expected packing must be at least half of the optimum.

Next, in Section 4.3.1, we prove a tight 3/7 competitive ratio relative to the optimal fractional packing, in the family of non-adaptive threshold algorithms, by lower bounding the performance of our proposed algorithm and loosely upper bounding the optimal fractional packing by 1. In Section 4.3.2, we prove a tight 0.432 competitive ratio relative to the optimal integer packing, in the family of non-adaptive threshold algorithms, by lower bounding the performance of our proposed algorithm and upper bounding the exact optimal integer packing at the same time.

### 4.3.1 A 0.428 Competitive Algorithm Relative to the Optimal Fractional Packing

We propose a randomized threshold policy, $\mathsf{ALG}_{\mathsf{N4.2}}$, and prove it is 3/7-competitive.

**Definition 4.2.** Let $\mathsf{ALG}_{\mathsf{N4.2}}$ be a randomized threshold policy that runs as follows,

1. At the beginning of the entire process, randomly draw $\tau$ from a distribution whose cumulative distribution function (CDF) is given by

$$F_{N4.2}(x) = \begin{cases} \dfrac{4/7 - x}{1 - 2x}, & x \in [0, 3/7] \\ 1, & x \in (3/7, 1] \end{cases} \tag{4.4}$$

2. We apply $\mathsf{THR}(\tau)$ policy throughout the process.

Notice that $F_{N4.2}(0) = 4/7$. This is the point mass we put on $\tau = 0$. This means that with probability $4/7$, we will perform $\mathsf{Greedy}$.

It is easy to check that our desired algorithm does not know how many items are there in total, not does it know the sizes of the items.

Now we state and prove our first result.

**Theorem 4.2.**

$$\inf_S \frac{\mathsf{ALG}_{\mathsf{N4.2}}(S)}{\mathsf{OPT}^+(S)} \geq \frac{3}{7}$$

*Proof.* Proof of Theorem 4.2. For any instance of arrival sequence $S$, we will show $\frac{\mathsf{ALG}_{\mathsf{N4.2}}(S)}{\mathsf{OPT}^+(S)} \geq 3/7$.

First of all, $\mathsf{Greedy}$ always accepts something. Denote the set of items accepted by $\mathsf{Greedy}$ as $G$. Denote $|G| = g$. If $G = [T]$ then $\mathsf{Greedy}$ is optimal. In this case

$$\frac{\mathsf{ALG}_{\mathsf{N4.2}}}{\mathsf{OPT}^+} \geq \Pr(\tau = 0) \cdot 1 + \Pr(\tau > 0) \cdot 0 \geq F(0) = 4/7 \geq 3/7.$$

If $G \subsetneq [T]$, let $M = [T] \backslash G$ denote the set of items blocked by $\mathsf{Greedy}$. Since $\mathsf{Greedy}$ always accepts an item as long as it can fill in, any item blocked by $\mathsf{Greedy}$ must exceed the remaining space of the knapsack, at the moment it is blocked. We also know that $G \cup M = [T]$, $G \cap M = \phi$.

Let $m$ be the smallest size in $M$, i.e. $m = \min_{t \in M} s_t$. Define index $t_m$ for the smallest item, or the first smallest item, if there are multiple smallest items.

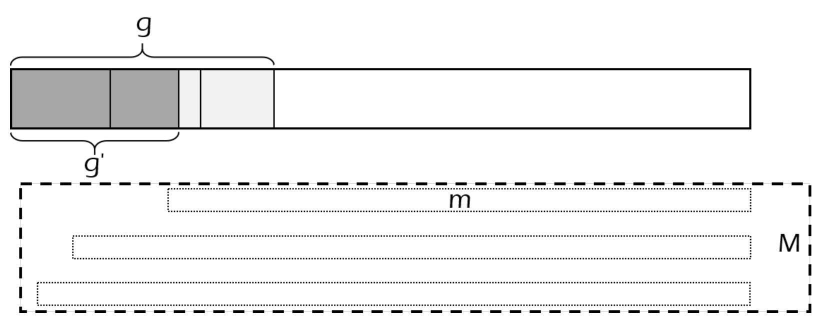
$$t_m = \min \{t \in [T] \,|\, s_t = m\}\,. \tag{4.5}$$

Denote $G'$ as the set of items accepted by $\mathsf{Greedy}$, at the moment item $t_m$ is blocked. Let $g' = \mathsf{size}(G')$. See Figure 4-2. A straightforward, but useful information about $m$ is:

$$g' + m > 1, \tag{4.6}$$

because $m$ is blocked by $\mathsf{Greedy}$. We wish to understand when we can pack an item of size at least $m$, by selecting a proper threshold $\tau$.

Figure 4-2: Illustration of the items that $\mathsf{Greedy}$ accepts, and blocks




We distinguish two cases: $m \geq 1/2$ and $m < 1/2$.

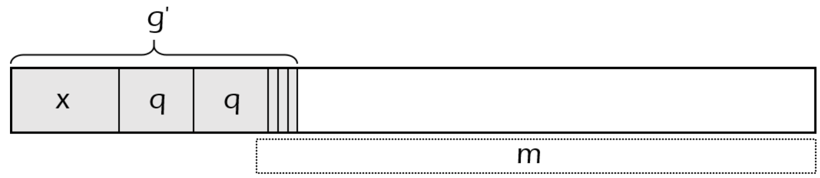
**Case 1**: $m \geq 1/2$.
Let $S^{\mathsf{THR}}(\tau)$ be the set of items that have sizes at least $\tau$, i.e. $S^{\mathsf{THR}}(\tau) = \{t \in S \,|\, s_t \geq \tau\}$. Now define

$$\begin{aligned} q = \max \quad & \tau \\ s.t. \quad & m + \mathsf{size}(S^{\mathsf{THR}}(\tau) \cap G') > 1 \end{aligned} \tag{4.7}$$

This means that if we adopt a $\mathsf{THR}(q)$ policy, then the size $m$ item must be blocked (possibly it will also be rejected, due to $q > m$, which leads to the discussion in Case 1.1).

Now consider the items in $S^{\mathsf{THR}}(q) \cap G'$. These items have sizes at least $q$. We count how many size $q$ items are there, and let $n$ be the number of size $q$ items. Denote

Figure 4-3: Illustration of Case 1 (and specifically, Case 1.2)




the total size of the remaining items be $x$. We know that $\mathsf{size}(S^{\mathsf{THR}}(q) \cap G') = nq + x$. See Figure 4-3.

We make the following observations:

1. There must exist some item from $G'$ that is of size $q$, i.e.

$$\exists t_q \in G' \subseteq [T], s.t.\ s_{t_q} = q. \tag{4.8}$$

   This is because otherwise we can select the smallest item size in $G'$ that is also larger than $q$. This item size satisfies (4.7), and violates the maximum property of $q$.

2. Size $m$ items can not fit in together with all the items in $S^{\mathsf{THR}}(q) \cap G'$, i.e.

$$nq + x + m > 1 \tag{4.9}$$

   This is because $\mathsf{size}(S^{\mathsf{THR}}(q) \cap G') = nq + x$. This is implied by (4.7).

3. A size $m$ item can fit in together with items $S^{\mathsf{THR}}(\tau) \cap G', \forall \tau > q$, i.e.

$$x + m \leq 1 \tag{4.10}$$

   This is because otherwise we could further increase $q$ to $\hat{q}$, so that $\hat{q}$ still satisfy equation (4.7). Define $\hat{q} = \min_{\varepsilon > 0, t \in S^{\mathsf{THR}}(q+\varepsilon) \cap G'} s_t$. We know (i) $\hat{q} > q$; (ii) $\mathsf{size}(S^{\mathsf{THR}}(\hat{q}) \cap G') + m > 1$. So $\hat{q}$ violates the maximum property of $q$.

We further distinguish two cases: $q > m$, and $q \leq m$.

**Case 1.1**: $q > m$.

In this case, if we adopt $\mathsf{Greedy}$ then we can get as much as $g$.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (0, q]$ then we can get no less than $q$. This is because due to (4.8) there must exist some item $t_q \in G'$ of size $q$. We either accept it, in which case we immediately earn $q$, or we have blocked it because we admitted some item $z \in [T]$ from $M$ and consumed too much space. But $\mathsf{Greedy}$ blocks item $z$ earlier than it accepts item $t_q$, which means that $s_z \geq s_{t_q} = q$. So in either case we earn $q$.

We have the following:

$$\begin{aligned}
\mathsf{ALG}_{\mathsf{N4.2}} &\geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq q) \cdot q \\
&= F_{N4.2}(0) \cdot g + (F_{N4.2}(q) - F_{N4.2}(0)) \cdot q \\
&\geq F_{N4.2}(0) \cdot (1 - 2q) + F_{N4.2}(q) \cdot q \\
&= 4/7 \cdot (1 - 2q) + 1 \cdot q \\
&= 4/7 - 1/7 \cdot q \\
&\geq 3/7
\end{aligned}$$

where the second inequality is because $g \geq g' > 1 - m$ (due to (4.6)) and $1 - m > 1 - q$ (Case 1.1: $q > m$); second equality is because $q > m \geq 1/2$ and the way we defined $F_{N4.2}(\cdot)$ in (4.4) so $F_{N4.2}(q) = 1$; last inequality is because $q \leq 1$.

Since $\mathsf{OPT}^+ \leq 1$, we have $\frac{\mathsf{ALG}_{\mathsf{N4.2}}}{\mathsf{OPT}^+} \geq \frac{3}{7}$.

**Case 1.2**: $q \leq m$.

In this case, if we adopt $\mathsf{Greedy}$ then we can get as much as $g$. This is the definition of $g$.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (0, q]$ then we get no less than $q$. This is because due to (4.8) there must exist some item $t_q \in G'$ of size $q$. We either accept it, in which case we immediately earn $q$, or we have blocked it because we admitted some item $z \in [T]$ from $M$ and consumed too much space. But $\mathsf{Greedy}$ blocks item $z$ earlier than it accepts item $t_q$, which means that $s_z \geq s_{t_q} = q$. So in either case we earn $q$.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (q, m]$ then we get no less than $m$. This is because due to (4.10), any item in $S^{\mathsf{THR}}(\tau) \cap G'$ will not block item $t_m$ (from expression (4.5));

and $\tau \le m$ so we will not reject item $t_m$. We either accept item $t_m$, in which case we immediately earn $m$, or we have blocked it because we admitted some item $z \in [T]$ from M and consumed too much space. But $m$ is smallest item size in $M$, which means that $s_z \ge m$. So in either case we earn $m$.

We have the following:

$$
\begin{aligned}
\mathsf{ALG}_{\mathsf{N4.2}} &\ge \Pr(\tau = 0)\cdot g + \Pr(0 < \tau \le q)\cdot q + \Pr(q < \tau \le m)\cdot m \\
&= F_{N4.2}(0)\cdot g + (F_{N4.2}(q) - F_{N4.2}(0))\cdot q + (F_{N4.2}(m) - F_{N4.2}(q))\cdot m \\
&\ge F_{N4.2}(0)\cdot (nq+x) + (F_{N4.2}(q) - F_{N4.2}(0))\cdot q + \\
&\qquad (F_{N4.2}(m) - F_{N4.2}(q))\cdot (1-(nq+x)) \\
&= (F_{N4.2}(q) - F_{N4.2}(0))\cdot q + 1 - F_{N4.2}(q) + (F_{N4.2}(q) - 3/7)\cdot (nq+x) \\
&\ge (F_{N4.2}(q) - F_{N4.2}(0))\cdot q + 1 - F_{N4.2}(q) + (F_{N4.2}(q) - 3/7)\cdot q \\
&= F_{N4.2}(q)\cdot (2q-1) + 1 - q
\end{aligned}
$$

where the second inequality is because $g \ge g' \ge nq + x$ and $m > 1 - (nq + x)$ (due to (4.9)); second equality is because $m \ge 1/2$ and the way we defined $F_{N4.2}(\cdot)$ in (4.4) so $F_{N4.2}(m) = 1$; the last inequality is because $F_{N4.2}(q) \ge F_{N4.2}(0) = 4/7 > 3/7$, so the coefficient in front of $nq + x$ is positive.

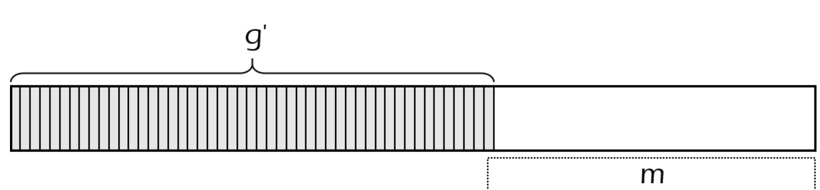
Now we plug in the expression of $F_{N4.2}(q)$ as defined in (4.4). If $q \le 3/7$ then $\mathsf{ALG}_{\mathsf{N4.2}} \ge \frac{4/7-q}{1-2q}\cdot(2q-1)+1-q = 3/7$; If $q > 3/7$ then $\mathsf{ALG}_{\mathsf{N4.2}} \ge q > 3/7$. So in either case we have shown $\mathsf{ALG}_{\mathsf{N4.2}} \ge 3/7$.

Since $\mathsf{OPT}^+ \le 1$, we have $\frac{\mathsf{ALG}_{\mathsf{N4.2}}}{\mathsf{OPT}^+} \ge \frac{3}{7}$.

**Case 2**: $m < 1/2$.

In this case, a crude analysis is enough. See Figure 4-4.

Figure 4-4: Illustration of Case 2



If we adopt $\mathsf{Greedy}$ then we can get as much as $g$. This is because $g$ is defined this way.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (0, m]$ then we either get $m$, or $m$ is blocked, in which case we must have already earned at least $1-m$ to block $m$.

We have the following:

$$\begin{aligned}
\mathsf{ALG}_{\mathsf{N4.2}} &\geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq m) \cdot \min\{m, 1-m\} \\
&\geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq m) \cdot m \\
&= F_{N4.2}(0) \cdot g + (F_{N4.2}(m) - F_{N4.2}(0)) \cdot m \\
&\geq F_{N4.2}(0) \cdot (1-m) + (F_{N4.2}(m) - F_{N4.2}(0)) \cdot m \\
&= F_{N4.2}(m) \cdot m + 4/7 \cdot (1-2m)
\end{aligned}$$

where the second inequality is because $m < 1/2$; the last inequality is because $g \geq g' > 1-m$ (due to (4.6)).

Now we plug in the expression of $F_{N4.2}(m)$ as defined in (4.4). If $m > 3/7$ then $\mathsf{ALG}_{\mathsf{N4.2}} \geq 4/7 - 1/7 \cdot m \geq 3/7$, because $m < 1/2 \leq 1$; If $m \leq 3/7$ then

$$\begin{aligned}
\mathsf{ALG}_{\mathsf{N4.2}} \geq \frac{4/7 - m}{1-2m} \cdot m + \frac{4}{7} \cdot (1-2m) = \frac{9}{28} \cdot (1-2m) + \frac{1}{28} \cdot \frac{1}{1-2m} + \frac{3}{14} \\
\geq 2\sqrt{\frac{9}{28} \cdot \frac{1}{28}} + \frac{3}{14} = \frac{3}{7}.
\end{aligned}$$

So in either case we have $\mathsf{ALG}_{\mathsf{N4.2}} \geq 3/7$.

Since $\mathsf{OPT}^+ \leq 1$, we have $\frac{\mathsf{ALG}_{\mathsf{N4.2}}}{\mathsf{OPT}^+} \geq \frac{3}{7}$.

In all, we have enumerated all the possible cases, to find $\dfrac{\mathsf{ALG}_{\mathsf{N4.2}}}{\mathsf{OPT}^+} \geq \dfrac{3}{7}$ always holds. $\square$

**Tightness proof of the $3/7$ competitive algorithm.**

In this section we show that the guarantee of $\inf_S \frac{\mathsf{ALG}_{\mathsf{N4.2}}(S)}{\mathsf{OPT}^+(S)} \geq \frac{3}{7}$ from Definition 4.2 is best-possible, relative to $\mathsf{OPT}^+$, among all randomized threshold policies. To do this, we invoke the minimax theorem of Yao (1977), which says that it suffices to construct a distribution $\mathcal{S}$ over sequences $S$ for which

$$\sup_{\mathsf{ALG}:\mathsf{ALG}=\mathsf{THR}(\tau),\tau\in[0,1]} \frac{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{ALG}(S)]}{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{OPT}^+(S)]} \le \frac{3}{7}.$$

In particular, we only need to establish that $\frac{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{ALG}(S)]}{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{OPT}^+(S)]} \le \frac{3}{7}$ for *deterministic* threshold policies specified by a $\tau \in [0,1]$.

**Theorem 4.3.** *There exists a distribution $\mathcal{S}$ over arrival sequences $S$ such that for any $\tau \in [0,1]$, the algorithm $\mathsf{ALG} = \mathsf{THR}(\tau)$ has $\frac{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{ALG}(S)]}{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{OPT}^+(S)]} \le \frac{3}{7}$.*

*Proof.* Proof of Theorem 4.3. Prove by construction. Let the random arrival sequence be $S$:

$$S = \begin{cases} (1/3, 2/3+\varepsilon), & \text{with prob. } 3/7; \\ (\underbrace{\varepsilon, \varepsilon, ..., \varepsilon}_{2/3+\varepsilon}, 1/3), & \text{with prob. } 3/7; \\ (\varepsilon, 1), & \text{with prob. } 1/7; \end{cases} \tag{4.11}$$

Following each realization of $S$, $\mathsf{OPT}^+(S) = 1$. So we have $\mathbb{E}_S[\mathsf{OPT}^+(S)] = 1$.

For any $\mathsf{ALG} = \mathsf{THR}(\tau), \tau \in [0,1]$, we enumerate all the potential values of $\tau$ in the following.

**Case 1**: $0 \le \tau \le \varepsilon$. In this case,

$$\mathbb{E}_S[\mathsf{ALG}(S)] = \frac{1}{3}\cdot\frac{3}{7} + \left(\frac{2}{3}+\varepsilon\right)\cdot\frac{3}{7} + \varepsilon\cdot\frac{1}{7} = \frac{3}{7} + \frac{4}{7}\cdot\varepsilon$$

**Case 2**: $\varepsilon < \tau \le 1/3$. In this case,

$$\mathbb{E}_S[\mathsf{ALG}(S)] = \frac{1}{3}\cdot\frac{3}{7} + \frac{1}{3}\cdot\frac{3}{7} + 1\cdot\frac{1}{7} = \frac{3}{7}$$

**Case 3**: $1/3 < \tau \le 2/3+\varepsilon$. In this case,

$$\mathbb{E}_S[\mathsf{ALG}(S)] = \left(\frac{2}{3}+\varepsilon\right)\cdot\frac{3}{7} + 0\cdot\frac{3}{7} + 1\cdot\frac{1}{7} = \frac{3}{7} + \frac{3}{7}\cdot\varepsilon$$

**Case 4**: $2/3+\varepsilon<\tau\le 1$. In this case,

$$\mathbb{E}_S[\mathsf{ALG}(S)] = 1\cdot\frac{1}{7}=\frac{1}{7}$$

In all, we have enumerated all the values that a threshold can take. In all cases, the performance of the threshold $\mathsf{THR}(\tau)$ policy has an expected performance of no more than $3/7+4/7\cdot\varepsilon$. But $\mathbb{E}_S[\mathsf{OPT}^+(S)]=1$. By taking $\varepsilon\to 0^+$ we finish the proof. $\square$

### 4.3.2 A 0.432 Competitive Algorithm Relative to the Optimal Integer Packing

In this section we are going to introduce a threshold policy that achieves the best-possible competitive ratio in the non-adaptive threshold family. In Section 4.3.2 we will show it is best-possible.

We first define some parameters that are going to be useful in the following analysis.

Let $H:[3/7,1/2)\times(0,1/2)\to\mathrm{R}_+$ be a bivariate real function defined as follows:

$$H(c,x)=\frac{1-2c}{x}-\frac{(1-2c)\ln(1-x)}{1-2x}-(1-c)$$

Now fix $c$ to be any number between $[3/7,1/2)$. Define $q_{N4.3}$ to be the only local minimizer on the second coordinate of $H(c,x)$, between $(0,1/2)$ – it can be implicitly given as the only solution between $[3/7,1/2)$, such that

$$2q_{N4.3}^3-7q_{N4.3}^2+5q_{N4.3}-1-2(1-q_{N4.3})q_{N4.3}^2\ln(1-q_{N4.3})=0,$$

or, approximately,

$$q_{N4.3}\approx 0.31847.$$

Define $c_{N4.3}$ to be the only solution between $[3/7, 1/2)$, such that

$$H(c_{N4.3}, q_{N4.3}) = \frac{1-2c_{N4.3}}{q_{N4.3}} - \frac{(1-2c_{N4.3})\ln(1-q_{N4.3})}{1-2q_{N4.3}} - (1-c_{N4.3}) = 0, \quad (4.12)$$

or, approximately,

$$c_{N4.3} \approx 0.43236.$$

We can check the following inequality: $\forall q \in (0, 1/2)$,

$$H(c_{N4.3}, q) \geq H(c_{N4.3}, q_{N4.3}) = 0 \quad (4.13)$$

We propose another randomized threshold policy, $\mathsf{ALG}_{\mathsf{N4.3}}$, using another random threshold. It gives us an improved 0.432 competitive guarantee.

**Definition 4.3.** Let $\mathsf{ALG}_{\mathsf{N4.3}}$ be a randomized threshold policy that runs as follows,

1. At the beginning of the entire process, randomly draw $\tau$ from a distribution whose CDF is given by

$$F_{N4.3}(x) = \begin{cases} (1-c_{N4.3}) - \dfrac{(1-2c_{N4.3})\ln(1-x)}{1-2x}, & x \in [0, q_{N4.3}] \\ 2(1-c_{N4.3}) - \dfrac{1-2c_{N4.3}}{x}, & x \in (q_{N4.3}, 1] \end{cases} \quad (4.14)$$

2. We apply $\mathsf{THR}(\tau)$ policy throughout the process.

Notice that $F_{N4.3}(0) = 1 - c_{N4.3}$. This is the point mass we put on $\tau = 0$. This means that with probability $1 - c_{N4.3} \approx 0.568$, we will perform $\mathsf{Greedy}$.

We state our main result here.

**Theorem 4.4.**

$$\inf_S \frac{\mathsf{ALG}_{\mathsf{N4.3}}(S)}{\mathsf{OPT}(S)} \geq c_{N4.3} \approx 0.432$$

The proof idea is the same as in Theorem 4.2, but in order to improve it, we are more careful in upper bounding the performance of $\mathsf{OPT}$. To compare to the proof of Theorem 4.2, Case 1.2 will be different. The proof details are deferred to Section C.1.

**Tightness proof of the 0.432 competitive algorithm.**

In this section we show that the guarantee of $\inf_S \frac{\mathsf{ALG}(S)}{\mathsf{OPT}(S)} \geq c_{N4.3}$ from Definition 4.3 is best-possible among all randomized threshold policies. As in Section 4.3.1, we invoke the minimax theorem of Yao (1977), which says that it suffices to construct a distribution $\mathcal{S}$ over sequences $S$ for which

$$\sup_{\mathsf{ALG}:\mathsf{ALG}=\mathsf{THR}(\tau),\tau\in[0,1]} \frac{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{ALG}(S)]}{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{OPT}(S)]} \leq c_{N4.3}.$$

**Theorem 4.5.** *There exists a distribution $\mathcal{S}$ over arrival sequences $S$ such that for any $\tau \in [0,1]$, the algorithm $\mathsf{ALG} = \mathsf{THR}(\tau)$ has $\frac{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{ALG}(S)]}{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{OPT}(S)]} \leq c_{N4.3} \approx 0.432$.*

*Proof.* Proof of Theorem 4.5. Prove by construction. Let the random arrival sequence be $S$:

$$S = \begin{cases} (\underbrace{\varepsilon, \varepsilon, ..., \varepsilon}_{1-q+\varepsilon}, q), & \text{where } q \in [1-q_{N4.3}, 1] \text{ conforms } u(\cdot); \\ (q_{N4.3}, 1-q_{N4.3}+\varepsilon, 1-q_{N4.3}+2\varepsilon, ..., 1), & \text{with prob. } x; \\ (\underbrace{\varepsilon, \varepsilon, ..., \varepsilon}_{1-q_{N4.3}+\varepsilon}, q_{N4.3}), & \text{with prob. } y; \\ (\varepsilon, 1), & \text{with prob. } z; \end{cases} \tag{4.15}$$

where

$$x = \frac{1-2c_{N4.3}}{1-2q_{N4.3}} \approx 0.37112; \qquad y = \frac{1-2c_{N4.3}}{q_{N4.3}} \approx 0.42309;$$
$$z = c_{N4.3} - x \approx 0.00954; \qquad u(q) = \frac{x}{q}.$$

We can verify that

$$x+y+z+\int_{1-q_{N4.3}}^{1} \frac{x}{q}\,\mathrm{d}q = x + \frac{1-2c_{N4.3}}{q_{N4.3}} + (c_{N4.3}-x) - \frac{1-2c_{N4.3}}{1-2q_{N4.3}} \cdot \ln(1-q_{N4.3}) = 1,$$

by plugging the expressions into the equation and using $H(c_{N4.3}, q_{N4.3}) = 0$ from (4.12). This equation shows that our construction conforms a legitimate probability measure.

Following each realization of $S$, $\mathsf{OPT}(S) = 1$. So we have $\mathbb{E}_S[\mathsf{OPT}(S)] = 1$.

For any $\mathsf{ALG} = \mathsf{THR}(\tau), \tau \in [0, 1]$, we enumerate all the potential values of $\tau$ in the following.

**Case 1**: $0 \le \tau \le \varepsilon$. In this case,

$$\begin{aligned}
\mathbb{E}_S[\mathsf{ALG}(S)] &= q_{N4.3} \cdot x + (1 - q_{N4.3} + \varepsilon) \cdot y + \varepsilon \cdot z + \int_{1-q_{N4.3}}^{1} (1 - q + \varepsilon) \cdot u(q) \mathrm{d}q \\
&= q_{N4.3} \cdot x + (1 - q_{N4.3}) \cdot \frac{1 - 2c_{N4.3}}{q_{N4.3}} - \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} \cdot \ln(1 - q_{N4.3}) \\
&\qquad - x \cdot q_{N4.3} + \varepsilon \cdot (1 - x) \\
&= \frac{1 - 2c_{N4.3}}{q_{N4.3}} - \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} \cdot \ln(1 - q_{N4.3}) - (1 - 2c_{N4.3}) + \varepsilon \cdot (1 - x) \\
&= c_{N4.3} + \varepsilon \cdot (1 - x)
\end{aligned}$$

where the last equality is due to (4.12).

**Case 2**: $\varepsilon < \tau \le q_{N4.3}$. In this case,

$$\begin{aligned}
\mathbb{E}_S[\mathsf{ALG}(S)] &= q_{N4.3} \cdot x + q_{N4.3} \cdot y + 1 \cdot z + \int_{1-q_{N4.3}}^{1} q \cdot u(q) \mathrm{d}q \\
&= q_{N4.3} \cdot \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} + q_{N4.3} \cdot \frac{1 - 2c_{N4.3}}{q_{N4.3}} + c_{N4.3} - \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} \\
&\qquad + q_{N4.3} \cdot \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} \\
&= c_{N4.3} + (1 - 2c_{N4.3}) \left( 2\frac{q_{N4.3}}{1 - 2q_{N4.3}} + 1 - \frac{1}{1 - 2q_{N4.3}} \right) \\
&= c_{N4.3}
\end{aligned}$$

**Case 3**: $q_{N4.3} < \tau \le 1 - q_{N4.3} + \varepsilon$. In this case,

$$\begin{aligned}
\mathbb{E}_S[\mathsf{ALG}(S)] &= (1 - q_{N4.3} + \varepsilon) \cdot x + 1 \cdot z + \int_{1-q_{N4.3}}^{1} q \cdot u(q) \mathrm{d}q \\
&= (1 - q_{N4.3}) \cdot \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} + c_{N4.3} - \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} + q_{N4.3} \cdot \frac{1 - 2c_{N4.3}}{1 - 2q_{N4.3}} + \varepsilon \cdot x \\
&= c_{N4.3} + \varepsilon \cdot x
\end{aligned}$$

**Case 4**: $1 - q_{N4.3} + \varepsilon < \tau \leq 1$. In this case,

$$\begin{aligned}\mathbb{E}_S[\mathsf{ALG}(S)] &= \tau \cdot x + 1 \cdot z + \int_\tau^1 q \cdot u(q)\mathrm{d}q \\ &= \tau \cdot x + c_{N4.3} - x + (1-\tau) \cdot x \\ &= c_{N4.3}\end{aligned}$$

In all, we have enumerated all the values that a threshold can take. In all cases, the performance of the threshold $\mathsf{THR}(\tau)$ policy has an expected performance of no more than $c_{N4.3} + \varepsilon \cdot (1-x)$. But $\mathbb{E}_S[\mathsf{OPT}(S)] = 1$. By taking $\varepsilon \to 0^+$ we finish the proof. □

## 4.4 Multiple Knapsacks

In this section we generalize our results to multiple knapsacks. We define the problem here, then in Section 4.4.1 we introduce the 0.214 competitive algorithm, and in Section 4.4.2 we introduce the impossibility result for a 0.461 competitive algorithm.

We manage $N$ divisible knapsacks indexed as $j \in [N]$, each having size $B_j, \forall j \in [N]$. In each period of time, one item $t \in [T]$ arrives with an associated vector of $N$ sizes $(s_{t1}, s_{t2}, ..., s_{tN}) \in (0,1]^N$. The sizes are revealed upon arrival, and each item must immediately be either entirely accepted by one knapsack, in which case $s_{tj}$ amount is filled up in knapsack $j$, or entirely rejected (there is no partial fulfillment). The objective is to maximize the sum of sizes of accepted items from all knapsacks, i.e. maximize the space in the knapsacks filled.

We compare the algorithm's performance relative to the space filled by an optimal offline packing, who knows the entire sequence of items in advance. This generalization can be seen as a modification of the AdWords budgeted allocation problem as in Mehta et al. (2005), where we do not allow the partial allocation of any queries that go over budget.

### 4.4.1 A 0.214 Competitive Algorithm

We first overview the AdWords problem originally proposed in Mehta et al. (2005). The language we use are from the tutorial Mehta et al. (2013). In each period of time, one item $s_t$ arrives with an associated vector of $N$ sizes $(s_{t1}, s_{t2}, ..., s_{tN}) \in (0,1]^N$. Suppose that, at this moment, some $b_j$ amount of space has been filled in each knapsack $j \in [N]$. If we assign the item to knapsack $j$, then $\min\{1 - b_j, s_{tj}\}$ amount of stock from knapsack $j$ will be filled – we allow for **truncation** in the AdWords problem. For this AdWords problem, the following greedy algorithm is well-known.

**Definition 4.4** (Algorithm 8, Mehta et al. (2013))**.** When item $t$ arrives, find $\tilde{j} \in \arg\max_{j \in [N]} s_{tj}$, and fit the item to knapsack $\tilde{j}$.

We make the following comments.

1. This greedy algorithm is irrespective to how much each knapsack has been filled. It is possible that the algorithm routes one item to a full knapsack, and completely wastes it.

2. This greedy algorithm is non-adaptive, in the sense that it routes items to knapsacks only based on the current item sizes, but not on the status of the knapsacks, nor the historically accepted / rejected item sizes (as long as there is remaining capacity).

It is well known that the greedy algorithm defined above achieves a competitive ratio of 1/2. For any instance $S$, let $\mathsf{ALG}_{\mathsf{AW}}(S)$ denote the total amount filled by the greedy algorithm from Definition 4.4, which is allowed to truncate. Let $\mathsf{OPT}_{\mathsf{AW}}(S)$ denote the total amount filled by a clairvoyant decision maker, which is also allowed to truncate.

**Proposition 4.6** (Theorem 5.1, Mehta et al. (2013))**.** *The greedy algorithm from Definition 4.4 is* $1/2$ *competitive for the AdWords problem, i.e.* $\forall S$,

$$\mathsf{ALG}_{\mathsf{AW}}(S) \geq \frac{1}{2}\mathsf{OPT}_{\mathsf{AW}}(S).$$

Now we return to our multiple knapsack problem without truncation.

**Definition 4.5.** We define our proposed algorithm, which essentially combines the two algorithms from Definitions 4.2 and 4.4.

1. For each item $t$, find $\tilde{j} \in \arg\max_{j\in[N]} s_{tj}$, and route item $t$ to knapsack $\tilde{j}$.
2. Adopt a single-knapsack policy from Definition 4.2, to decide if we accept item $t$ or not.
3. Actually accept item $t$ by matching it to knapsack $\tilde{j}$, if both we accept item $t$, and it fits.

We do not prescribe the correlations between the thresholds for each knapsack. They can be arbitrarily correlated, and they can also be independent. In other words, we use the algorithm from Definition 4.4 to route items to knapsacks, and then use our algorithms to decide if we actually accept it.

For any instance $S$, let $\mathsf{ALG}(S)$ denote the total amount filled by the combined algorithm from Definition 4.5, which is **not** allowed to truncate. Let $\mathsf{OPT}(S)$ denote the total amount filled by a clairvoyant decision maker, which is also **not** allowed to truncate.

**Theorem 4.7.** *The combined algorithm from Definition 4.5 is $\frac{1}{2}c$ competitive, i.e.*

$$\inf_S \frac{\mathsf{ALG}(S)}{\mathsf{OPT}(S)} \geq \frac{1}{2}c,$$

*where $c$ is the performance guarantee of any algorithm on a single-knapsack relative to $\mathsf{OPT}^+$ as defined in* (4.3)*.*

*In particular, if we adopt the algorithm from Definition 4.5, then $c = 3/7$; if we adopt the algorithm from Definition 4.1, then $c = 1/2$.*

*Proof.* Proof of Theorem 4.7 For any knapsack $j$, let $I_j$ be the set of items routed to it in Step 1 of Definition 4.5 ($I_j$ includes items that are later discarded by the threshold of knapsack $j$). Note that $I_j$ does not depend on the adoption of single-knapsack algorithms from Step 2.

Denote $\mathsf{OPT}_j^+ = \min\{\sum_{i \in I_j} s_{ij}, B_j\}, \forall j \in [N]$. It is obvious that $\mathsf{ALG}_{\mathsf{AW}} = \sum_{j \in [N]} \mathsf{OPT}_j^+$., due to the allowance of truncation in $\mathsf{ALG}_{\mathsf{AW}}$.

From Definition 4.2, we earn at least $c \cdot (\sum_{i \in I_j} s_{ij}) = c \cdot \mathsf{OPT}_j^+$, in expectation. So that

$$\mathsf{ALG} \geq c \cdot \mathsf{OPT}_j^+ \geq c \cdot (\frac{1}{2}\mathsf{OPT}_{\mathsf{AW}}) \geq \frac{c}{2}\mathsf{OPT}$$

where the first inequality is because on each knapsack $\mathsf{ALG}$ earns at least $c$ fraction of what $\mathsf{ALG}_{\mathsf{AW}}$ does; the second inequality is from Proposition 4.6; and the third inequality is simply the fact that $\mathsf{OPT}_{\mathsf{AW}} \geq \mathsf{OPT}$, because any optimal assignment when truncation is not allowed is a feasible solution to the problem when truncation is allowed. $\square$

### 4.4.2 An upper bound for multiple knapsacks strictly less than 0.5

Fix a small $\varepsilon > 0$. There are $N$ knapsacks with sizes $B_1 = \cdots = B_N = 1$. The arrival sequence $S$ deterministically starts with $N$ items each of which take size $\varepsilon$ in a particular knapsack, and size 0 in all other knapsacks. Specifically,

$$s_{t,t'} = \begin{cases} \varepsilon, & \text{if } t = t' \\ 0, & \text{otherwise} \end{cases} \qquad\qquad \forall\ t, t' \in [N].$$

After item $N$, with probability $\alpha$, the arrival sequence terminates; with probability $1-\alpha$, there are $N$ more items whose sizes adhere to a “upper-triangular graph”, defined as follows. A permutation $\pi : [N] \to [N]$ is chosen uniformly at random among all $N!$ possibilities. Item $N+1$ takes size 1 in all knapsacks. Item $N+2$ takes size 1 in all knapsacks, except knapsack $\pi(1)$, where it takes size 0. Item $N+3$ takes size 1 in all knapsacks, except knapsacks $\pi(1)$ and $\pi(2)$, where it takes size 0. This construction is repeated until item $2N$, which takes size 1 only in knapsack $\pi(N)$. Formally,

$$s_{N+t,t'} = \begin{cases} 1, & \text{if } t' \notin \{\pi(1), \dots, \pi(t-1)\} \\ 0, & \text{if } t' \in \{\pi(1), \dots, \pi(t-1)\} \end{cases} \qquad \forall\, t, t' \in [N].$$

The optimal solution matches items $1, \dots, N$ to their corresponding knapsacks if the arrival sequence terminates after item $N$, and rejects items $1, \dots, N$ otherwise, matching items $N+1, \dots, 2N$ to knapsacks $\pi(1), \dots, \pi(N)$, respectively, instead. Therefore

$$\mathbb{E}_S[\mathsf{OPT}(S)] = (\alpha)N\varepsilon + (1-\alpha)N. \tag{4.16}$$

Meanwhile, the algorithm does not know whether the arrival sequence will terminate after item $N$, nor does it know $\pi$. In the first phase, any algorithm can be captured by how many $\varepsilon$'s it accepts, which we denote using $e \in \{0, 1, ..., N\}$. Should the second phase occur, by the symmetry of the random permutation, an algorithm cannot do better than placing an arriving item arbitrarily into an empty knapsack (where it will take size 1) whenever possible. Therefore, the expected reward of any algorithm when the second phase does occur is completely determined by $e$.

We now formalize this construction to derive an upper bound on the competitive ratio.

**Theorem 4.8.** *There exists a distribution $\mathcal{S}$ over arrival sequences $S$ such that for any (adaptive or non-adaptive) algorithm* $\mathsf{ALG}$*, we have* $\frac{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{ALG}(S)]}{\mathbb{E}_{S\sim\mathcal{S}}[\mathsf{OPT}(S)]} \le \frac{35}{76} \approx 0.461$. *By Yao's minimax theorem, the competitive ratio cannot be greater than 35/76.*

*Proof.* Proof of Theorem 4.8. Consider the example we described above, where $N = 4$, and $\alpha = 1 - \frac{12}{7}\varepsilon$. By equation (4.16) above,

$$\mathbb{E}_S[\mathsf{OPT}(S)] = (1 - \frac{12}{7}\varepsilon)N\varepsilon + (\frac{12}{7}\varepsilon)N = 76/7\varepsilon - 48/7\varepsilon^2.$$

Now we analyze the maximum possible value of $\mathbb{E}_S[\mathsf{ALG}(S)]$. As discussed before, any algorithm is characterized by $e \in \{0, 1, \dots, 4\}$, which is the number of size-$\varepsilon$ items accepted.

**Case 1**: $e = 0$. With probability $\alpha$, the arrival sequence terminates with 0 accepted; with probability $1 - \alpha$, there are $N$ more items. We enumerate all the 24 possibilities, to find there are 6 cases that a deterministic algorithm accepts 2 of them, 17 cases that accepts 3, and 1 case that accepts 4. In expectation we fill 67/24 into the knapsacks. In this case

$$\frac{\mathbb{E}_S[\mathsf{ALG}(S)]}{\mathbb{E}_S[\mathsf{OPT}(S)]} = \frac{12/7\varepsilon \cdot 67/24}{76/7\varepsilon - 48/7\varepsilon^2} = \frac{67}{152 - 96\varepsilon}.$$

**Case 2**: $e = 1$. With probability $\alpha$, the arrival sequence terminates with $\varepsilon$ accepted; with probability $1-\alpha$, there are $N$ more items. Out of all the 24 possibilities, there are 16 cases that a deterministic algorithm accepts 2 of them, and 8 cases that accepts 3. In expectation we fill 56/24 into the knapsacks. In this case

$$\frac{\mathbb{E}_S[\mathsf{ALG}(S)]}{\mathbb{E}_S[\mathsf{OPT}(S)]} = \frac{1 \cdot \varepsilon + 12/7\varepsilon \cdot 56/24}{76/7\varepsilon - 48/7\varepsilon^2} = \frac{35}{76 - 48\varepsilon}.$$

**Case 3**: $e = 2$. With probability $\alpha$, the arrival sequence terminates with $2\varepsilon$ accepted; with probability $1-\alpha$, there are $N$ more items. Out of all the 24 possibilities, there are 6 cases that a deterministic algorithm accepts 1 of them, and 18 cases that accepts 2. In expectation we fill 42/24 into the knapsacks. In this case

$$\frac{\mathbb{E}_S[\mathsf{ALG}(S)]}{\mathbb{E}_S[\mathsf{OPT}(S)]} = \frac{2 \cdot \varepsilon + 12/7\varepsilon \cdot 42/24}{76/7\varepsilon - 48/7\varepsilon^2} = \frac{35}{76 - 48\varepsilon}.$$

**Case 4**: $e = 3$. With probability $\alpha$, the arrival sequence terminates with $3\varepsilon$ accepted; with probability $1 - \alpha$, there are $N$ more items. The algorithm must be able to fill one item into the unfilled knapsack in the first round of phase two. In this case

$$\frac{\mathbb{E}_S[\mathsf{ALG}(S)]}{\mathbb{E}_S[\mathsf{OPT}(S)]} = \frac{3 \cdot \varepsilon + 12/7\varepsilon \cdot 1}{76/7\varepsilon - 48/7\varepsilon^2} = \frac{33}{76 - 48\varepsilon}.$$

**Case 5**: $e = 4$. With probability $\alpha$, the arrival sequence terminates with $3\varepsilon$

accepted; with probability $1-\alpha$, there are $N$ more items. But the algorithm cannot fill in any because all the knapsacks are all occupied with $\varepsilon$'s. In this case

$$\frac{\mathbb{E}_S[\mathsf{ALG}(S)]}{\mathbb{E}_S[\mathsf{OPT}(S)]} = \frac{4 \cdot \varepsilon + 12/7\varepsilon \cdot 0}{76/7\varepsilon - 48/7\varepsilon^2} = \frac{28}{76 - 48\varepsilon}.$$

In all cases, any policy has an expected performance of no more than $\frac{35}{76-48\varepsilon}$. By taking $\varepsilon \to 0^+$ we finish the proof. $\square$

## 4.5 Computational Study: Using Real Data from A Latin American Chain Department Store

We use supply chain data from a Latin American chain department store, to computationally study the performance of our algorithms. The supply chain data contains 974 SKU's, and their associated order quantities from different local stores to a total of 21 regional warehouses.

The category that we focus on is young women's fashion products. Since fashion products are highly unpredictable in its sales, we adopt the lens of competitive analysis, which is natural when there is no knowledge about future arrival sequences. There is typically only 1 selling season, and the selling season typically lasts for 3-6 months. At the beginning of the selling season, there is an initial stock placed in each regional warehouse. There is no inventory replenishment throughout this process. Orders to a specific warehouse cannot be split or redirected to a different warehouse, because there is a specific warehouse which serves each local store, so order sizes greater than the available inventory must be rejected. Therefore, our industry partner faces the same accept/reject problem on order sizes, and has the same goal of maximizing total inventory fulfilled, equal to the sum of sizes of accepted orders.

To give a concrete example, here is an arrival sequence in the winter between year 2015 and 2016, for one SKU of women purse.

$$S = (7, 18, 80, 41, 1, 30, 12, 17). \tag{4.17}$$

This selling season spans the Revolution Day[2], Christmas, and New Year. And the order quantities are in commercial units.

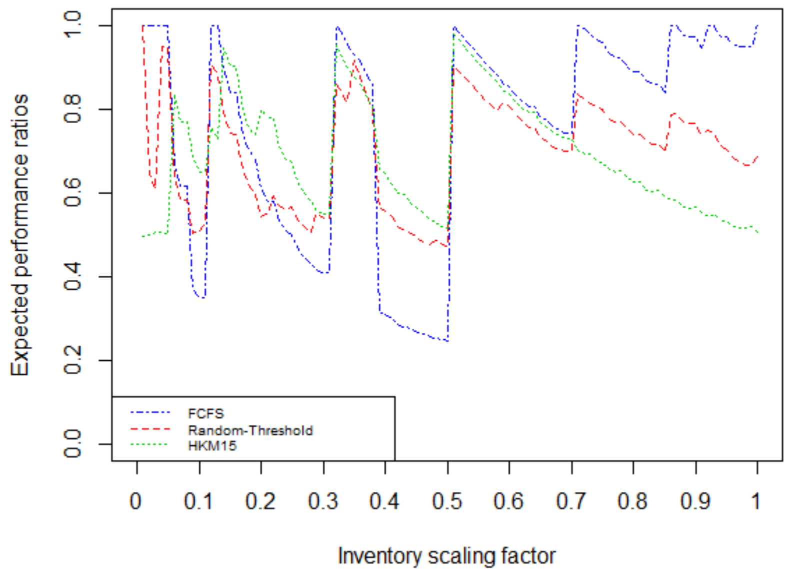
Sequence $S$ is an observed sequence of order sizes, accepted by greedy (FCFS) in the real world supply chain. The sum of all the order sizes is smaller than the starting inventory, which, in this case, is 208 units. Any orders which could not have been fulfilled are *censored* from the data. As a result, we create non-trivial instances by re-scaling the starting inventory amounts (which we know) for each SKU at each warehouse by a factor $\alpha \in [0, 1]$, and then test the performance of different accept/reject policies over different scaling factors $\alpha$. This is a limitation of our computational study. Nonetheless, we believe that this censoring only favors FCFS in our computational study, because large orders rarely appear in the end of a sequence (which would cause FCFS to perform poorly). Such an experiment setup to vary the initial inventory level is very common in the revenue management literature; see Zhang and Cooper (2005), Liu and Van Ryzin (2008).

We compute the expected revenue from our proposed random threshold algorithm from Definition 4.2, named **Random-Threshold**; the (deterministic) revenue from first-come-first-serve policy, named **FCFS**; and the expected revenue from the algorithm suggested in Han et al. (2015), named **HKM15**. The results are shown in Figure 4-5, where we have divided all the numbers by its corresponding offline optimal integer packing. The offline optimal packing serves as an upper bound, so that the performance ratio is always between 0 and 1, with higher ratios indicating better performance.

When inventory level is either very small or very large, first-come-first-serve achieves near-optimal performance. This is not surprising, because FCFS always tries to accept an order if possible: when inventory is large then FCFS could almost accept everything except for the ones that arrive late. In our specific example shown in (4.17), the late items are fairly small – and this is why FCFS is near-optimal. On the other hand, when inventory is small then anything that FCFS successfully fits into the knapsack is already very large, relative to the small capacity. So FCFS has

[2] A Mexican black Friday

Figure 4-5: Computational performance using a real arrival sequence. The sequence is as shown in Equation (4.17).




a near-optimal performance when inventory is very small. For scenarios where the starting inventory is of moderate size (i.e. for SKU's that were neither overstocked nor understocked initially), our proposed algorithm has a relatively smaller variance.

### 4.5.1 Average and Worst-case Performance over all SKU's

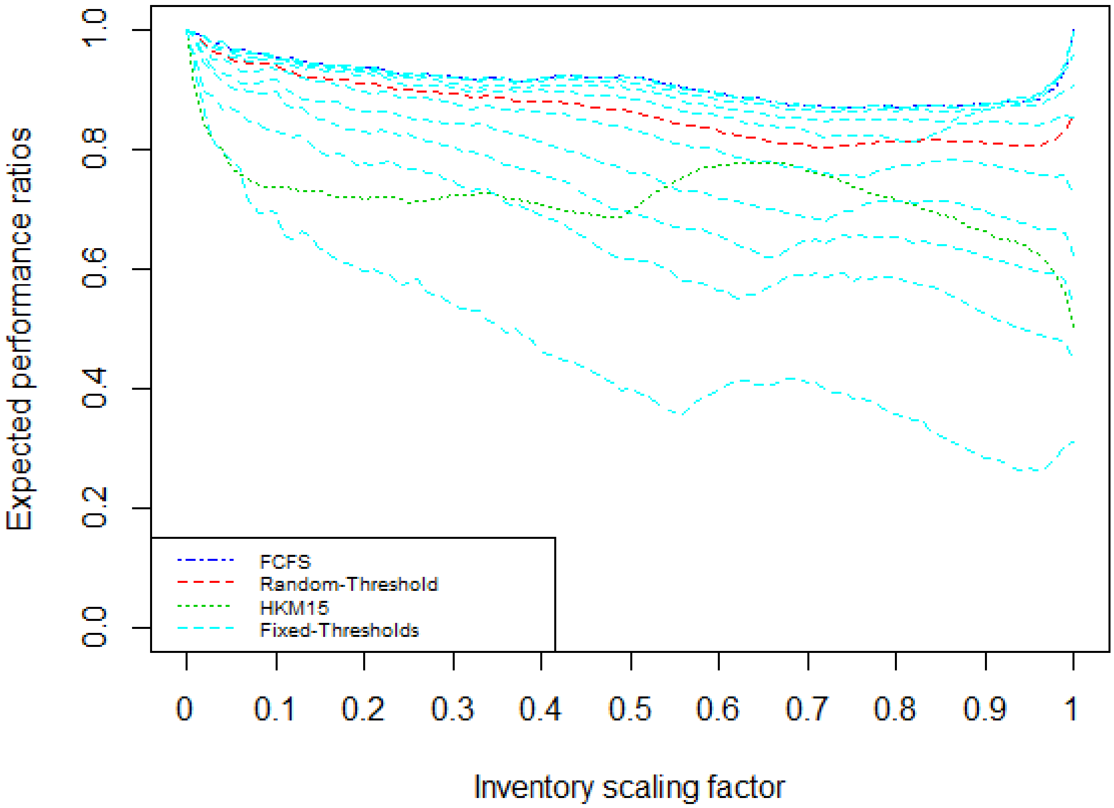
The earlier results shown for a specific SKU was used to illustrate the experimental setup. Now we show aggregate experimental results over all the SKU's.

There are 974 different SKU's, carried in 21 different warehouses over the country. For any fixed scaling factor, we first compute the average performance over different SKU's.

We compute our **Random-Threshold** algorithm in the following manner. We take 21 evenly-spaced percentiles[3] of the threshold distribution $F$, that is, we take the 21 thresholds defined by $F^{-1}(0), F^{-1}(0.05), ..., F^{-1}(0.95), F^{-1}(1)$. These are 21 natural thresholds to be implemented over the 21 warehouses. Then we randomly permute them and assign them over the 21 warehouses, making accept/reject decisions at each warehouse based on the assigned threshold (scaled by the starting inventory). We

[3]Note that some SKU's are only stored in a subset of them, say, only 6 warehouses. And we take 6 evenly-spaced percentiles.

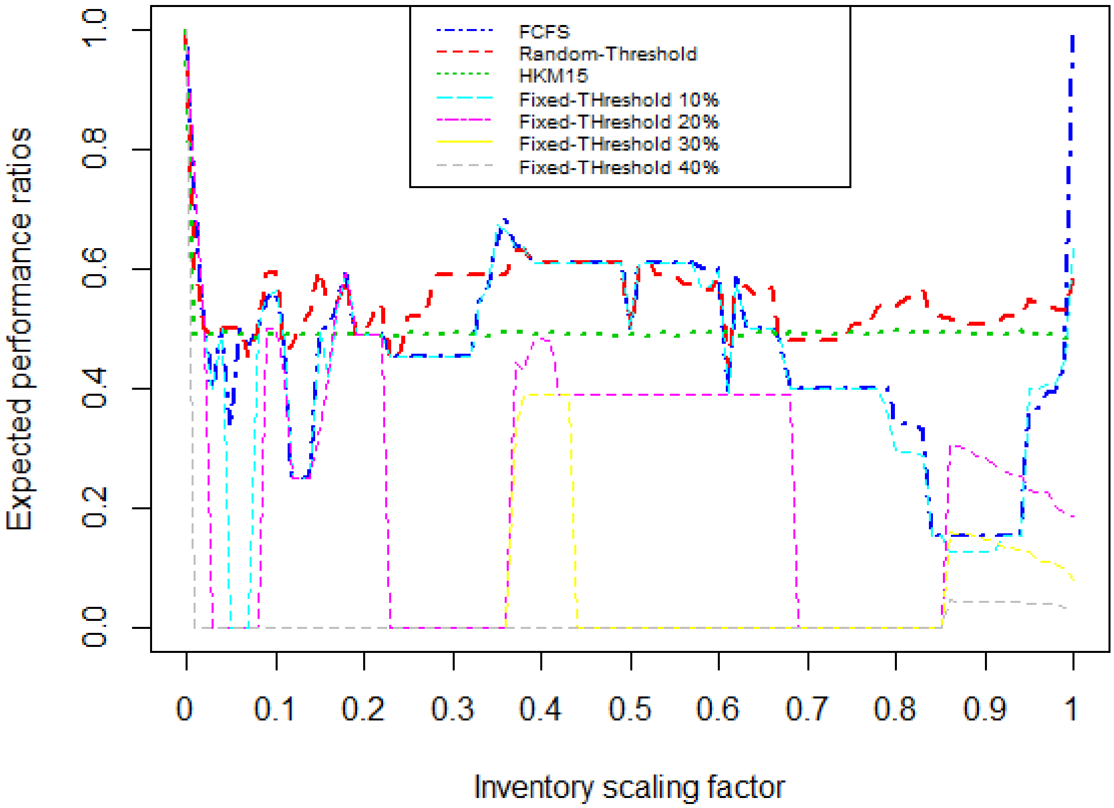
Figure 4-6: Computational of the average case performance using real arrival sequences over the country. The 10 non-overlapping lines from top to bottom correspond to the ten fixed thresholds, $3\%, 5\%, 10\%, ..., 80\%$, respectively.




then average the fulfillment ratios over the warehouses to determine the performance for a specific SKU. Finally we take an outer average over many independent random permutations of warehouses to define a final performance ratio for each SKU.

We compare our performance against the (deterministic) revenue from the greedy first-come-first-serve policy, named **FCFS**; the expected revenue from the algorithm suggested in Han et al. (2015), named **HKM15**; and the deterministic revenues of 10 fixed-threshold algorithms, named **Fixed-Thresholds**. The fixed thresholds are re-scaled to be $3\%, 5\%, 10\%, 15\%, 20\%, 30\%, 40\%, 50\%, 60\%, 80\%$ of the initial capacity. The results are shown in Figure 4-6, where we have divided all the numbers by its corresponding offline optimal integer packing. The performance of the Fixed-Thresholds algorithms have a decreasing performance ratio with respect to their thresholds. FCFS, if we interpret it as a threshold policy with threshold equal to zero, has the best performance. Interestingly, we integrate the CDF function and

Figure 4-7: Computational of the worst case performance using real arrival sequences over the country. Some curves from Fixed-Thresholds are removed, for clarity.




find the expected threshold suggested by our Random-Threshold policy to be roughly 28% – and the performance of our Random-Threshold policy coincides to be between the 20% and 30% curves of the Fixed-Thresholds policies.

Again, we see that when inventory is either very small or very large, FCFS and the Fixed-Thresholds whose thresholds are small, they all achieve near-optimal performance. We find that the FCFS policy has the best average-case performance ratio. And the higher thresholds we increase for the Fixed-Thresholds, the worse performance it yields.

While this is discouraging, we believe that the way in which order sizes are censored in our data favors FCFS, since large orders cannot come at the end. Moreover, the gap between FCFS and our Random-Threshold algorithm is always smaller than 7%. Also, note that our Random-Threshold algorithm *always* outperforms HKM15. The gap between these two algorithms is between 5%–27%.

Finally, to illustrate the benefit of our Random-Threshold algorithm, we also display the *worst-case* performance of each algorithm over the SKU's, for each scaling factor. The computational results are shown in Figure 4-7.

Again, we see that when inventory is either very small or very large, FCFS, HKM15, our Random-Threshold, and the Fixed-Threshold algorithms whose thresholds are small all achieve near-optimal performance. However, our Random-Threshold algorithm has the best worst-case performance for a large fraction of scaling factors. This shows that setting different thresholds at different warehouses, according to the distributions we derived, indeed provides the best baseline guarantee on the fulfillment of any SKU.

We see that as we increase the thresholds, the Fixed-Threshold policies tend to have worse performance. Note that the performance of FCFS or any Fixed-Threshold policy does not have any guarantee – FCFS and the Fixed-Threshold policies whose thresholds are less or equal to 5% can have a performance as bad as 15%; and the Fixed-Threshold policies whose thresholds are at least 10% can have a worst-case performance guarantee of 0. By contrast, the worst-case performance of our algorithm over all of the scaling factors is 44% (close to the theoretical guarantee of 43%). Meanwhile, the performance for HKM15 always equals its theoretical guarantee of 50%.

## 4.6 Conclusion

In this paper, we study an online knapsack problem and its competitive analysis. We focus on a particular class of random threshold algorithms, that initially draw a random threshold and never change the threshold throughout the entire horizon. This class of algorithms benefit us from simplicity, applicability, and incentive-compatibility, which are important in real-world applications.

We start from the single knapsack problem, and study its generalization to the multiple knapsack problems. We provide constant factor competitive results that are tight on a single knapsack, relative to two different offline optimal packing optimum. Numerical experiments suggest that the performance is better than merely our theoretical guarantees.

# Chapter 5

# Design and Analysis of Switchback Experiments

## 5.1 Introduction

Academic scholars have appreciated the benefits that experimentation brings to firms for many decades (March 1991, Sitkin 1992, Sarasvathy 2001, Thomke 2001, Kohavi and Thomke 2017, Sun et al. 2018, Xiong et al. 2019). However, widespread adoption of the practice has only taken off in the last decade, partly fueled by the rapid cost reductions achieved by firms in the technology sector (Kohavi et al. 2007, 2009, Azevedo et al. 2019, Kohavi et al. 2020). Most large firms now possess internal tools for experimentation, and a growing number of smaller and more conventional companies are purchasing the capabilities from third-party sellers that offer full-stack integration (Thomke 2020). These tools typically allow simple "A/B" tests that compare the standard offering "A" to a new or improved version "B". The comparisons are made across a range of different business outcomes, and the tests are usually conducted for at least a week (Kohavi et al. 2020). This simple practice has provided tremendous value to firms (Koning et al. 2019).

Some firms and authors, however, have recognized the limitations of these simple A/B tests (Gupta et al. 2019, Bojinov et al. 2020b). Technology firms have identified two common challenges in conducting A/B tests. The first challenge is in handling interference, the scenario where the assignment of one subject impacts another. Many online platforms and retail marketplaces often observe varying levels of interference when conducting experiments. See Chamandy (2016), Cui et al. (2017), Kastelman and Ramesh (2018), Farronato et al. (2018), Glynn et al. (2020), Holtz et al. (2020) for online platforms (*e.g.*, Airbnb, DoorDash, Lyft, Uber), and Caro and Gallien (2012), Ferreira et al. (2016), Cui et al. (2019), Ma et al. (2020a) for retail markets (*e.g.*, Amazon, AB InBev, Rue la la, Zara). The second challenge is in estimating heterogeneous (or personalized) effects. See Nie et al. (2018), Deshpande et al. (2018), Hadad et al. (2019).

In this paper, we simultaneously tackle both of these challenges by developing a theoretical framework for the optimal design and analysis of switchback experiments under the minimal amount of assumptions. In switchback experiments, we sequentially expose a unit to a random treatment, measure its response, and repeat the procedure for a fixed period of time (Robins 1986, Bojinov and Shephard 2019). By administering alternate treatments to the same unit, we can directly estimate an individual level causal effect and alleviate the challenges posed by interference.

The minimal amount of assumptions gains our framework validity compared to the literature. Some literature assumes specific outcome models under interference. Wager and Xu (2019), Johari et al. (2020) both work on experimental design for two-sided online platforms, by assuming that the interference can be captured via mean-field approximation. Glynn et al. (2020) assumes an underlying Markov Chain model and formulates the experimental design problem as estimating the difference between two steady state reward distributions. Some other literature directly models the interference through a network, *e.g.* Li et al. (2015), Athey et al. (2018), Eckles et al. (2016), Sussman and Airoldi (2017), Basse et al. (2019a), Puelz et al. (2019). In such models, a treatment assigned to one node of the network creates a "spillover effect," which impacts the outcomes of the neighboring nodes. All of the above methods make specific assumptions on the outcome models. If these assumptions hold, the above methods correctly identify the causal effects (or the model parameters); if these assumptions do not hold, the estimates are likely biased.

In this paper, no specific outcome models are assumed. Instead, we make general assumptions about the existence of the carryover effects. The carryover effects refer to the persistence of past interventions in impacting the future outcomes; and the order of carryover effects refers to the duration of time periods of such persistence. In this paper, we establish results on the optimal design of switchback experiments under different assumptions of the order of the carryover effects; we also propose a data-driven procedure to estimate the order of the carryover effects.

**Applications**. There are two classes of applications where switchback experiments are widely used in practice. The first arises when units interfere with each other either through a network or some more complicated unknown structure. For example, consider a ride-hailing platform that wants to test a new fare pricing algorithm's effectiveness in a large city (Farronato et al. 2018). Administering the test version to a subset of drivers can impact their behavior, which, in turn, could change the behavior of drivers that are receiving the old version. Directly comparing the revenue generated by the drivers across the two groups will likely provide a biased estimate of what would happen if everyone were assigned to the new version compared to the old. Instead, practitioners treat the city as a single aggregated unit and use a switchback experiment to estimate the intervention's effectiveness, thereby alleviating the problem caused by interference. A similar issue often arises in revenue management when, for example, a retailer wants to test the effectiveness of a new promotion planning algorithm (Ferreira et al. 2016). Administering the new version to a subset of stock keeping units (SKU's) cannibalizes the sales from the other SKU's. Again comparing the generated revenue across the two groups is unlikely to provide an accurate measure of the promotion's effectiveness. Instead, practitioners can treat all the SKU's as a single aggregated unit and use a switchback experiment to obtain accurate estimates of the promotion's effectiveness.

The second application arises when we have a limited number of experimental units, and we believe the effects are likely to be heterogeneous. For example, Bojinov and Shephard (2019) used switchback experiments to make causal claims about the relative effectiveness of algorithms compared with humans at executing large financial trades across a range of financial markets. More generally, psychologists and biostatisticians regularly use switchback experiments whenever studying the effectiveness of an intervention on a single unit, *e.g.*, Lillie et al. (2011) and Boruvka et al. (2018).

**Main Contributions**. There are three significant challenges to using switchback experiments. The first is that causal estimators from switchback experiments have large variances as the precision is a function of the total number of assignments. The second is that past interventions are likely to impact future outcomes; this is often referred to as a carryover effect. Typically, many authors assume that there are no carryover effects (Chamberlain 1982, Athey and Imbens 2018, Imai and Kim 2019), although some recent work has relaxed this assumption (Robins 1986, Sobel 2012, Bojinov et al. 2020a). The third is that standard super population inference — where researchers either assume a model for the outcome, or that the units are sampled from an infinitely large population — requires unrealistic assumptions that fail to capture the problem's personalized nature (Bojinov and Shephard 2019).

This paper's main contributions are to address these three challenges and present a framework that allows firms and researchers to run reliable switchback experiments. First, we derive optimal designs for switchback experiments, ensuring that we can select a design that leads to the lowest variance among the most popular class assignment mechanisms. The designs are optimal in the sense that we search for both the optimal randomization points, as well as the optimal randomization probabilities, which, together, capture a general class of randomization mechanisms. Second, we assume the presence of a carryover effect and show that our estimation and inference are valid both when the order of the carryover effect is correctly specified and misspecified, the later leading to a minor increase in the variance. For practitioners, we also propose a method to identify the order of the carryover effect by running a series of carefully designed switchback experiments. Finally, we take a purely design-based perspective on uncertainty; that is, we treat the outcomes as unknown but fixed (or equivalently, we condition on the set of potential outcomes) and assume that the assignment mechanism is the only source of randomness (Fisher et al. 1937, Kempthorne 1955, Rubin 1980, Abadie et al. 2020). The main benefit of a design-based perspective is that the inference, and in turn the causal conclusions, do not depend on our ability to correctly specify a model describing the phenomena we are studying, ensuring that our findings are wholly non-parametric and robust to model misspecification (Imbens and Rubin 2015, Chapter 5).

**Roadmap**. The paper is structured as follows. In Section 5.2 we define the notations, the assumptions, and the assignment mechanism that we focus on, which we will refer to as the regular switchback experiments. In Section 5.3, we discuss how to design an effective regular switchback experiment under the minimax rule. The design is optimal with respect to (i) the optimal treatment assignment probability, and (ii) the randomization frequency and randomization points. We cast the design problem as a minimax discrete optimization problem, identify the worst-case adversarial strategy, establish structural results, and then explicitly find the optimal design. In Section 5.4, we discuss how to perform inference and conduct statistical testing based on the results obtained from an optimally designed switchback experiment. We propose an exact test for sharp null hypotheses, and an asymptotic test for testing the average treatment effect. We also discuss how to make inference when the carryover effect is misspecified, and how to conduct hypothesis testing to identify the true order of the carryover effect. In Section 5.5, we run simulations to test the correctness and effectiveness of our proposed theoretical results under various simulation setups. In Section 5.6, we give empirical illustrations on how to conduct a switchback experiment in practice and conclude with limitations which may lead to future research directions. All technical proofs are in the Appendix.

## 5.2 Notations, Assumptions, and Regular Switchback Experiments

### 5.2.1 Assignment Paths and Potential Outcomes

We focus our discussion on a single experimental unit. For example, this unit could be a ride-hailing platform testing the effectiveness of a new fare pricing algorithm in a city. At each time point $t \in [T] = \{1, 2, ..., T\}$, we assign the unit to receive an intervention $W_t \in \{0, 1\}$. For example, one experimental period could be one to two hours for a ride-hailing platform and $T$ could be two weeks, i.e., $T = 336$ when one period is one hour. In some applications, external factors determine the time horizon $T$, *e.g.*, a typical experimental duration for a ride-hailing platform is a few weeks; however, when $T$ is not pre-determined, Section 5.6 provides details for how to choose an appropriate $T$. Throughout most of this paper, with the exception being the derivation of our asymptotic results, we consider $T$ to be a known, fixed constant.

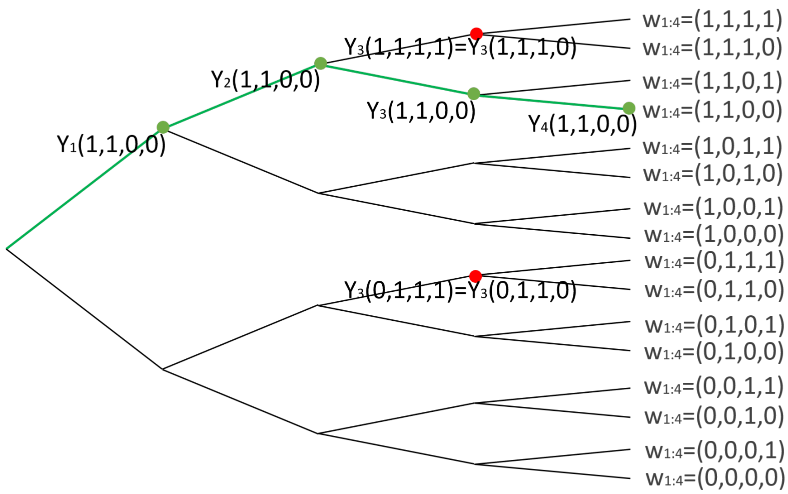
Following convention, we say that the unit is assigned to treatment if $W_t = 1$ and control when $W_t = 0$; in A/B testing terminology, "A" is control and "B" is treatment. For example, Chamandy (2016) studied how a new surge-pricing subsidy (the treatment) compared to the current setup without the subsidy (the control). The assignment path is then the collection of assignments and is denoted using a vector notation whose dimensions are specified in the subscript, $\boldsymbol{W}_{1:T} = (W_1, W_2, ..., W_T) \in \{0, 1\}^T$. We adopt the convention that $\boldsymbol{W}_{1:T}$ stands for a random assignment path, while $\boldsymbol{w}_{1:T}$ stands for one realization.

After administering the assigned intervention, we observe a corresponding outcome. For example, this could be the average ride-matching rate during each two hour experimental period. Following the extended potential outcomes framework, at time $t \in [T]$, we posit that for each possible assignment path $\boldsymbol{w}_{1:T}$ there exists a corresponding potential outcome denoted by $Y_t(\boldsymbol{w}_{1:T})$; the set of all potential outcomes are collected in $\mathbb{Y} = \{Y_t(\boldsymbol{w}_{1:T})\}_{t \in [T], \boldsymbol{w}_{1:T} \in \{0,1\}^T}$ with support $\mathbb{Y} \in \mathcal{Y}$.

**Example 5.1.** When $T = 4$, there are 16 assignment paths as shown in Figure 5-1. Associated with each assignment path $\boldsymbol{w}_{1:4}$ are four potential outcomes $Y_1(\boldsymbol{w}_{1:4})$, $Y_2(\boldsymbol{w}_{1:4})$, $Y_3(\boldsymbol{w}_{1:4})$, and $Y_4(\boldsymbol{w}_{1:4})$. ∎

Throughout this paper, we do not directly model the potential outcomes or impose a parametric relationship with the assignment path; instead, we treat them as unknown but fixed quantities, or, equivalently, we implicitly condition on $\mathbb{Y}$. Our setup does not preclude the possibility that the potential outcomes were generated

Figure 5-1: Illustrator of assignment paths and potential outcomes when $T = 4$. The green path stands for one assignment path $\boldsymbol{w}_{1:4} = (1, 1, 0, 0)$. Following the green path there are four potential outcomes. The two red dots each stands for two potential outcomes that are equal under Assumption 5.1. And the potential outcomes at the two red dots are equal if Assumption 5.2 is further assumed.




through a dynamic process; however, it allows us to be completely agnostic to the data generating process, making our causal claims more objective. To make inference possible, we rely on the variation introduced by the random assignment path; this is commonly referred to as finite-sample or design-based perspective and has a long history going back to Neyman (1923), Fisher et al. (1937), Kempthorne (1955), Rubin (1980). Unlike traditional sampling-based inference, the design-based approach does not require a hypothetical population from which to sample experimental units, see the textbook reference Imbens and Rubin (2015) and a recent paper Abadie et al. (2020) for recent reviews.

We make two assumptions that limit the dependence of the potential outcomes on assignment paths. Below let $\{t : t'\} = \{t, t + 1, ..., t'\}$, for any $t < t' \in [T]$.

**Assumption 5.1** (Non-anticipating Potential Outcomes)**.** *For any* $t \in [T]$*,* $\boldsymbol{w}_{1:t} \in \{0, 1\}^t$*, and for any* $\boldsymbol{w}'_{t+1:T}, \boldsymbol{w}''_{t+1:T} \in \{0, 1\}^{T-t}$*,*

$$Y_t(\boldsymbol{w}_{1:t}, \boldsymbol{w}'_{t+1:T}) = Y_t(\boldsymbol{w}_{1:t}, \boldsymbol{w}''_{t+1:T}).$$

Assumption 5.1 states that the potential outcomes at time $t$ do not depend on future treatments (Bojinov and Shephard 2019, Basse et al. 2019b, Rambachan and Shephard 2019). Since we control the assignment mechanism instead of letting the experimental units to administer future assignments (*e.g.*, at a ride-hailing platform, a passenger does not know the price in the next hour), the design ensures that this assumption is satisfied.

**Example 5.2** (Example 5.1 Continued)**.** Under Assumption 5.1, $Y_3(1,1,1,1) = Y_3(1,1,1,0)$. In Figure 5-1 the red dot at $Y_3(1,1,1)$ stands for both $Y_3(1,1,1,1)$ and $Y_3(1,1,1,0)$. ∎

**Assumption 5.2** ($m$-Carryover Effects)**.** *There exists a fixed and given $m$, such that for any $t \in \{m+1, m+2, ..., T\}, \boldsymbol{w}_{t-m:T} \in \{0,1\}^{T-t+m+1}$, and for any $\boldsymbol{w}'_{1:t-m-1}$, $\boldsymbol{w}''_{1:t-m-1} \in \{0,1\}^{t-m-1}$,*

$$Y_t(\boldsymbol{w}'_{1:t-m-1}, \boldsymbol{w}_{t-m:T}) = Y_t(\boldsymbol{w}''_{1:t-m-1}, \boldsymbol{w}_{t-m:T}).$$

Assumption 5.2 restricts the order of the carryover effect (Laird et al. 1992, Senn and Lambrou 1998, Bojinov and Shephard 2019, Basse et al. 2019b). The validity of Assumption 5.2 depends on the setting and requires practitioners to use their domain knowledge to choose an appropriate $m$. Examples arise in ride-hailing, in which the effect of surge pricing on a ride-hailing platform typically dissipates after one or two hours, depending on the city size (Garg and Nazerzadeh 2019). Moreover, we propose in Section 5.4.4 a hypothesis testing method to select appropriate $m$ using an experimental approach.

Assumptions 5.1 and 5.2 allow us to simplify notation. For any $t \in \{m+1, ..., T\}$ and any two assignment paths $\boldsymbol{w}_{1:T}, \boldsymbol{w}'_{1:T} \in \{0,1\}^{m+1}$, whenever $\boldsymbol{w}_{t-m:t} = \boldsymbol{w}'_{t-m:t}$ this leads to

$$Y_t(\boldsymbol{w}_{1:T}) = Y_t(\boldsymbol{w}'_{1:T}).$$

In the remainder of this paper, we will write $Y_t(\boldsymbol{w}_{t-m:t}) := Y_t(\boldsymbol{w}_{1:T})$ to emphasize the dependence on treatments $\boldsymbol{w}_{t-m:t}$. For example, the potential outcomes at the two red dots in Figure 5-1 are equal, i.e., $Y_3(1,1) := Y_3(1,1,1,1) = Y_3(1,1,1,0) = Y_3(0,1,1,1) = Y_3(0,1,1,0)$

### 5.2.2 Causal Effects

In the potential outcomes approach to causal inference, any comparison of potential outcomes has a causal interpretation. In this paper, we focus on a special set of causal estimands that measure the relative effectiveness of persistently assigning a unit to treatment as opposed to control. For any $p \in \{0, 1, ..., T-1\}$, let $\mathbf{1}_{p+1} = (1, 1, ..., 1)$ be a vector of $(p+1)$ ones; let $\mathbf{0}_{p+1} = (0, 0, ..., 0)$ be a vector of $(p+1)$ zeros. Define the average lag-$p$ causal effect of consecutive treatments on the outcome, for any $p \in \{0, 1, ..., T-1\}$,

$$\tau_p(\mathbb{Y}) = \frac{1}{T-p} \sum_{t=p+1}^{T} [Y_t(\mathbf{1}_{p+1}) - Y_t(\mathbf{0}_{p+1})]. \tag{5.1}$$

This estimand captures the effects of permanently deploying a new policy, and has been widely studied in the longitudinal experiments since the early work of Robins (1986).

*Remark* 5.1. Although we focus on an average causal effect, all of our results and analysis trivially extend to the total causal effect, which does not normalize, *i.e.*, $(T-p)\tau_p(\mathbb{Y})$. The optimal design as we will show in Section 5.3 will remain unchanged.

It is worth noting that $p$ reflects the experimental designer's knowledge of the order of the carryover effect. See discussion below Assumption 5.2. Such a knowledge is sometimes correct, which we refer to as the perfect knowledge case ($p = m$ ); it is sometimes incorrect, which we refer to as the "misspecified" $m$ case[1] ($p \neq m$). In this section we focus on the $p = m$ case; Section 5.4.3 discusses the $p \neq m$ case. Section 5.4.4 introduces a procedure to identify $m$.

The challenge of causal inference on switchback experiments is that we only observe one assignment path. In other words, for each period $t$, we observe at most

[1]Some authors specifically focus on $p < m$, particularly when $m$ is of the same order as $T$ (Bojinov and Shephard 2019).

either $Y_t(\mathbf{1}_{p+1})$ or $Y_t(\mathbf{0}_{p+1})$ (and sometimes neither). After conducting a switchback experiment, the observed data contains $\boldsymbol{w}_{1:T}^{\mathsf{obs}}$ the realized assignment path, and $Y_t^{\mathsf{obs}} = Y_t(\boldsymbol{w}_{1:T}^{\mathsf{obs}})$ the observed outcome at time $t$ under the realized assignment path $\boldsymbol{w}_{1:T}^{\mathsf{obs}}$. To link the observed and potential outcomes, we assume there is only one version of the treatment[2], and that there is no non-compliance.

### 5.2.3 Regular Switchback Experiments

The design of switchback experiment induces a probabilistic distribution over assignment paths $\boldsymbol{w}_{1:T} \in \{0,1\}^T$. Formally, a design of switchback experiment is any $\eta : \{0,1\}^T \to [0,1]$ such that

$$\sum_{\boldsymbol{w}_{1:T} \in \{0,1\}^T} \eta(\boldsymbol{w}_{1:T}) = 1, \qquad\qquad \eta(\boldsymbol{w}_{1:T}) \geq 0, \ \forall \ \boldsymbol{w}_{1:T} \in \{0,1\}^T.$$

Explicitly, $\eta(\cdot)$ is the underlying discrete distribution of the random assignment path $\boldsymbol{W}_{1:T}$.

In this paper, we narrow our scope to the family of regular switchback experiments. This family of experiments are parameterized by $\mathbb{T}$ and $\mathbb{Q}$, defined as

$$\mathbb{T} = \{t_0 = 1 < t_1 < t_2 < ... < t_K\} \subseteq [T],$$

where $K < T$ is a positive integer, and $\mathbb{T}$ contains a total of $K+1$ integers, which is a subset of all the time indices; and

$$\mathbb{Q} = (q_0, q_1, ..., q_K) \in (0,1)^{K+1} := \mathcal{Q},$$

where $\mathbb{Q}$ is a vector of $K+1$ real numbers between $(0,1)$. For the ease of notations also denote $t_{K+1} = T+1$, though our time horizon is only $T$ periods.

*Definition* 5.1 (Regular Switchback Experiments). For any $\mathbb{T} = \{t_0 = 1 < t_1 < ... <$

[2]When combined with non-interference if there were multiple units, this is known as the stable unit treatment value assumption (Rubin 1980).

$t_K\} \subseteq [T]$, and any $\mathbb{Q} = (q_0, q_1, ..., q_K) \in (0,1)^{K+1}$, a regular switchback experiment $(\mathbb{T}, \mathbb{Q})$ administers a probabilistic treatment at any time $t$, given by:

$$\Pr(W_t = 1) = q_k, \quad \text{if } \ t_k \leq t \leq t_{k+1} - 1 \tag{5.2}$$

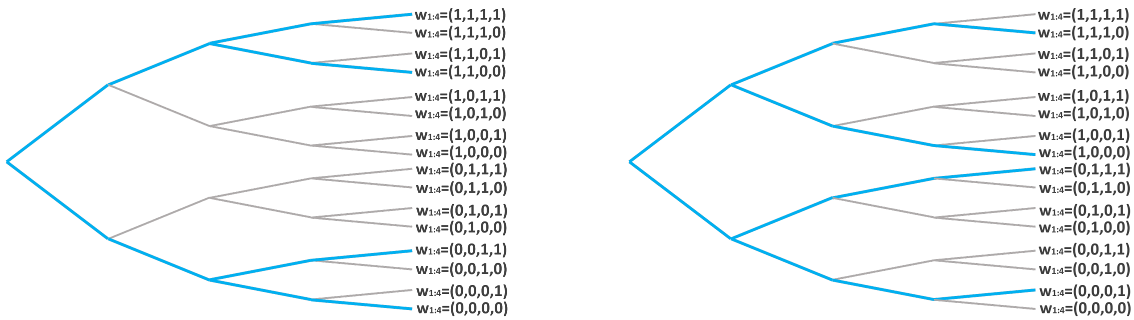
In words, the experimental designer jointly decides on a collection of randomization points, which consists of flipping biased coins at each period $t \in \{t_0, ..., t_K\}$, as well as a collection of randomization probabilities behind the biased coins, $(q_0, ..., q_K)$. If the resulting flip at period $t_k$ is heads, then the experimental designer assigns the unit to treatment during periods $(t_k, t_k + 1, ..., t_{k+1} - 1)$; otherwise, if tails, assigns the unit to control during periods $(t_k, t_k + 1, ..., t_{k+1} - 1)$.

*Example* 5.3. When $T = 4$, $\mathbb{T} = \{t_0 = 1, t_1 = 3\}, \mathbb{Q} = (q_0, q_1) = (1/2, 1/2)$ corresponds to the following design: with probability 1/4, $\boldsymbol{W}_{1:4} = (1,1,1,1)$; with probability 1/4, $\boldsymbol{W}_{1:4} = (1,1,0,0)$; with probability 1/4, $\boldsymbol{W}_{1:4} = (0,0,1,1)$; with probability 1/4, $\boldsymbol{W}_{1:4} = (0,0,0,0)$. See Figure 5-2 (left figure) for the four assignment paths that are in the support of the discrete probability distribution. ∎

*Example* 5.4. Not all switchback experiments are regular. For example, when $T = 4$: with probability 1/4, $\boldsymbol{W}_{1:4} = (1,1,1,0)$; with probability 1/4, $\boldsymbol{W}_{1:4} = (1,0,0,0)$; with probability 1/4, $\boldsymbol{W}_{1:4} = (0,1,1,1)$; with probability 1/4, $\boldsymbol{W}_{1:4} = (0,0,0,1)$. See Figure 5-2 (right figure) for the four assignment paths that are in the support of the discrete probability distribution. ∎

It has been widely acknowledged in the literature that a good design should be the one that flips fair coins, i.e., $q_k = 1/2, \forall\ k \in \{0, 1, ..., K\}$. The reason behind fair coin flips reflects experimental designer's limited assumption on the outcome model, and the inherent symmetry in the potential outcomes. The literature has either shown that fair coin flips are optimal, *e.g.*, Wu (1981), Li (1983), Basse et al. (2019b) where they make mild assumptions on permutation invariance; or has explicitly made assumptions that the coins flips be fair, *e.g.*, Bai (2019), Harshaw et al. (2019). In Section 5.3, we show that fair coin flipping is indeed optimal, under a mild assumption

Figure 5-2: Two designs of switchback experiments




Note: The blue lines stand for the possible treatment assignments that a design could administer. Left: regular switchback experiment (Example 5.3); Right: irregular switchback experiment (Example 5.4).

that is different from permutation invariance.

Since most firms design the entire experiment before the experiment is launched, the treatment assignments are typically not updated based on the observed outcomes; therefore, we do not consider adaptively changing the treatment assignments. We briefly outline adaptive experimental designs as future extensions in Section 5.6.

For any regular switchback experiment $(\mathbb{T}, \mathbb{Q})$, we may also use $\mathbb{T}$ to refer to the same experiment when $\mathbb{Q}$ is clear from the context. We denote the underlying discrete probability distribution using $\eta_{\mathbb{T},\mathbb{Q}}(\cdot)$. For any $\mathbb{T}$ and $\mathbb{Q}$, the discrete probability distribution has a total of $2^{K+1}$ many supports. The assignment path is random, and follows the discrete probability distribution $\eta_{\mathbb{T},\mathbb{Q}}(\cdot)$:

$$\eta_{\mathbb{T},\mathbb{Q}}(\boldsymbol{w}_{1:T}) = \begin{cases} \displaystyle\prod_{k=0}^{K} \frac{\mathbb{1}\{w_{t_k}=1\}}{q_{t_k}} \cdot \frac{\mathbb{1}\{w_{t_k}=0\}}{\bar{q}_{t_k}}, & \text{if } \forall\ k \in \{0,1,...,K\}, \\ & \quad w_{t_k} = w_{t_k+1} = ... = w_{t_{k+1}-1}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

In the remainder of this paper, unless explicitly noted, all probabilities and expectations are taken with respect to this discrete probability distribution $\eta_{\mathbb{T},\mathbb{Q}}(\cdot)$.

### 5.2.4 Estimation

Now that $\eta_{\mathbb{T},\mathbb{Q}}(\cdot)$ is determined, following any realization of the assignment path $\boldsymbol{w}_{1:T}$, we use the Horvitz-Thompson estimator to estimate the causal effect:

$$\hat{\tau}_p(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y}) = \frac{1}{T-p} \sum_{t=p+1}^{T} \left\{ Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{w}_{t-p:t} = \mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1})} - Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{w}_{t-p:t} = \mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1})} \right\}. \tag{5.4}$$

We emphasize that the estimator $\hat{\tau}_p(\cdot,\cdot,\cdot)$ depends on (i) the probability distribution that the assignment path is sampled from, (ii) the realization of the assignment path, and (iii) the set of potential outcomes.

*Example* 5.5. Suppose $T = 4, p = m = 1$. Suppose the assignments are probabilistic and $\Pr(W_t = 1) = \Pr(W_t = 0) = 1/2, \forall t \in [4]$. With probability 1/16 the green assignment path as in Figure 5-1 is administered, $\boldsymbol{W}_{1:4} = (1, 1, 0, 0)$. The estimator is then $\hat{\tau}_1 = \frac{1}{3}\{4Y_2(1,1) + 0 - 4Y_4(0,0)\}$. ∎

Since the assignment path $\boldsymbol{W}_{1:T}$ is random, this Horvitz-Thompson estimator is also random. Moreover, when the assignment path satisfies a regular switchback, the probabilities in the denominator are known. As we will show in Theorem 5.4, under the optimal design, these probabilities will be multiplicatives of 1/2, allowing us to avoid the known stability issues of the Horvitz-Thompson estimator when the probabilities are extreme (either close to 0 or close to 1). It is well-known that the Horvitz-Thompson estimator is unbiased if the treatment and control probabilities are both non-zero.

*Proposition* 5.2 (Unbiasedness of the Horvitz-Thompson Estimator). *In a regular switchback experiment, under Assumptions 5.1 and 5.2, the Horvitz-Thompson estimator is unbiased for the average lag-p causal effect of consecutive treatments on outcome, i.e.,*

$$\mathbb{E}[\hat{\tau}_p(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{W}_{1:T}, \mathbb{Y})] = \tau_p(\mathbb{Y}).$$

The expectation $\mathbb{E}[\cdot]$ is taken with respect to the random assignment $\boldsymbol{W}_{1:T} \sim$

$\eta_{\mathbb{T},\mathbb{Q}}(\cdot)$. when it is obvious we will compress the subscript in the expectation writing $\mathbb{E}[\cdot]$ to mean $\mathbb{E}_{\boldsymbol{W}_{1:T}\sim\eta_{\mathbb{T},\mathbb{Q}}}[\cdot]$. The proof to Proposition 5.2 is standard, by checking the expectations. We defer its proof to Section D.2 in the Appendix.

### 5.2.5 Evaluation of Experiments: the Decision-Theoretic Framework

To evaluate the quality of a design of experiment, we adopt the decision-theoretic framework (Berger 2013, Bickel and Doksum 2015). When the random design is $\eta_{\mathbb{T},\mathbb{Q}}(\cdot)$, for any realization of the assignment path $\boldsymbol{w}_{1:T}$ and any set of potential outcomes $\mathbb{Y}$, we define the loss function

$$L(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y}) = \left(\hat{\tau}_p(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y}) - \tau_p(\mathbb{Y})\right)^2$$

and the risk function

$$r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) = \sum_{\boldsymbol{w}_{1:T}\in\{0,1\}^T} \eta_{\mathbb{T},\mathbb{Q}}(\boldsymbol{w}_{1:T}) \cdot \left(\hat{\tau}_p(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y}) - \tau_p(\mathbb{Y})\right)^2 \tag{5.5}$$

Such a risk function quantifies the expected squared difference between our estimand and estimator. Since the estimator is unbiased, the risk function also has a second interpretation: the variance of the estimator. A design with a lower risk is also a design whose estimator has a lower variance.

*Example* 5.6 (Examples 5.3 and 5.5 Revisited). Suppose $T = 4$ and $p = m = 1$. As in Example 5.3, $\mathbb{T} = \{1, 3\}$. With probability 1/4, $\boldsymbol{W}_{1:4} = (1, 1, 0, 0)$, $\hat{\tau}_1(\mathbb{T}) = \frac{1}{3}\{2Y_2(1,1) - 2Y_4(0,0)\}$, $L(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y}) = \frac{1}{9}\{Y_2(1,1) + Y_2(0,0) - Y_3(1,1) + Y_3(0,0) - Y_4(1,1) - Y_4(0,0)\}^2$. As in Example 5.5, $\tilde{\mathbb{T}} = \{1, 2, 3, 4\}$. With probability 1/16, $\boldsymbol{W}_{1:4} = (1, 1, 0, 0)$, $\hat{\tau}_1(\tilde{\mathbb{T}}) = \frac{1}{3}\{4Y_2(1,1) - 4Y_4(0,0)\}$, $L(\eta_{\tilde{\mathbb{T}},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y}) = \frac{1}{9}\{3Y_2(1,1) + Y_2(0,0) - Y_3(1,1) + Y_3(0,0) - Y_4(1,1) - 3Y_4(0,0)\}^2$. ∎

Example 5.6 suggests that, even if the two realizations of the assignment path are the same and the potential outcomes are the same, since the probability distributions

$\eta_{\mathbb{T},\mathbb{Q}}$ and $\eta_{\tilde{\mathbb{T}},\mathbb{Q}}$ are distinct, the corresponding estimators $\hat{\tau}_1(\mathbb{T})$ and $\hat{\tau}_1(\tilde{\mathbb{T}})$ could be different, and the corresponding loss functions $L(\eta_{\mathbb{T},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y})$ and $L(\eta_{\tilde{\mathbb{T}},\mathbb{Q}}, \boldsymbol{w}_{1:T}, \mathbb{Y})$ could also be different. This observation suggests that there exists some design $\eta_{\mathbb{T}^*}$ that has a small risk. In the next section we find such a design when $m$ is correctly specified.

## 5.3 Design of Regular Switchback Experiments under Minimax Rule

The goal of this section is to find the optimal design of regular switchback experiments, i.e., to select the optimal randomization points and the optimal randomization probabilities. Throughout this section we assume $m$ is known and we set $p = m$.

We formalize our experimental design problem through the minimax framework. The minimax decision rule (Berger 2013, Wu 1981, Li 1983) finds an optimal design of experiment such that the worst-case risk against an adversarial selection of potential outcomes is minimized,

$$\min_{\mathbb{T}\in[T],\mathbb{Q}\in\mathcal{Q}} \max_{\mathbb{Y}\in\mathcal{Y}} r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) = \min_{\mathbb{T}\in[T],\mathbb{Q}\in\mathcal{Q}} \max_{\mathbb{Y}\in\mathcal{Y}} \sum_{\boldsymbol{w}_{1:T}\in\{0,1\}^T} \eta_{\mathbb{T},\mathbb{Q}}(\boldsymbol{w}_{1:T}) \cdot \left(\hat{\tau}_p(\boldsymbol{w}_{1:T}, \mathbb{Y}) - \tau_p(\mathbb{Y})\right)^2 . \tag{5.6}$$

One compelling reason to adopt the minimax framework, as commented in the seminal work of Wu (1981), is that "the experimenter's information about the model is never perfect. When a model is proposed, there is always the possibility that the 'true' model deviates from the assumed model." Instead of finding the best possible design by imposing a model, we try to derive the best possible design for the worse possible set of potential outcomes.

To overcome minimaxity and to lay out the foundation for inference, we impose an additional assumption on the support of the potential outcome. Since the potential outcomes are unknown but fixed, we assume that their absolute values are bounded from above, and that bound is attainable at every time period.

*Assumption* 5.3 (Bounded Potential Outcomes). *The potential outcomes are bounded by some constant, i.e.,* $\exists\ B > 0, s.t.\ \forall\ t \in [T],\ \forall\ \boldsymbol{w} \in \{0,1\}^T,\ |Y_t(\boldsymbol{w})| \leq B$, *or, equivalently,* $\mathbb{Y} \in \mathcal{Y} = [-B, B]^T$.

Assumption 5.3 is often satisfied since it assumes that the potential outcomes are bounded by the same (possibly a large) constant, (*e.g.*, the ride-matching rate from each experimental period is always a finite quantity) and that the extreme could possibly occur at any point in time (*e.g.*, the maximum ride-matching rate could be observed at any time). In particular, knowledge about the magnitude of $B$ is not required, and, as we show below, the optimal design does not depend on $B$.

The reason to make Assumption 5.3 is two fold. First, for optimization purposes, Assumption 5.3 reflects the inherent symmetry in the potential outcomes under both treatment and control, which is in the same spirit as the permutation invariance assumption (Wu 1981, Li 1983, Basse et al. 2019b). It is such symmetry that ensures the optimality of fair coin flipping. See Theorem 5.4 below. Second, for inferential purposes, Assumption 5.3 ensures that the variance of the estimator is well-behaved, which is commonly assumed in the finite-sample inference literature (Aronow et al. 2017, Chin 2018, Bojinov et al. 2019, Bojinov and Shephard 2019, Li et al. 2020, Han et al. 2021). It is the well-behaved variance that lays the foundation of our limiting distribution in Theorem 5.7.

To solve the minimax problem (5.6), we start by focusing on the inner maximization part. We characterize the worst-case potential outcomes by identifying two dominating strategies for the adversarial selection of potential outcomes. Denote $\mathbb{Y}^+ = \{Y_t(\mathbf{1}_{m+1}) = Y_t(\mathbf{0}_{m+1}) = B\}_{t \in \{m+1:T\}}$ and $\mathbb{Y}^- = \{Y_t(\mathbf{1}_{m+1}) = Y_t(\mathbf{0}_{m+1}) = -B\}_{t \in \{m+1:T\}}$

*Lemma* 5.3. *Under Assumptions 5.1–5.3,* $\mathbb{Y}^+$ *and* $\mathbb{Y}^-$ *are the only two dominating strategies for the adversarial selection of potential outcomes. That is, for any* $\mathbb{T} \subseteq [T]$ *and for any* $\mathbb{Y} \in \mathcal{Y}$,

$$r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}^+) \geq r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}); \qquad\qquad r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}^-) \geq r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}).$$

*Moreover, for any* $\mathbb{Y} \in \mathcal{Y}$ *such that* $\mathbb{Y} \neq \mathbb{Y}^+, \mathbb{Y} \neq \mathbb{Y}^-$, *the above two inequalities are* *strict.*

The proof of Lemma 5.3 can be found in Section D.3.3. Lemma 5.3 simplifies the minimax problem in (5.6), as it allows us to replace $\mathbb{Y}$ by $\mathbb{Y}^* = \mathbb{Y}^+$ or $\mathbb{Y}^* = \mathbb{Y}^-$, and reduce the minimax problem (5.6) into a minimization problem

$$\min_{\mathbb{T}\in[T],\mathbb{Q}\in\mathcal{Q}} r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}^*).$$

Next we solve this minimization problem by first finding the optimal $\mathbb{Q}$ values.

*Theorem* 5.4 (Optimality of Fair Coin Flipping). *Under Assumptions 5.1–5.3, any optimal design of experiment* $(\mathbb{T}, \mathbb{Q})$ *must satisfy* $q_0 = q_1 = ... = q_K = 1/2$.

The proof of Theorem 5.4 can be found in Section D.3.4. Theorem 5.4 suggests that the optimal randomization probabilities should be 1/2. So we can restrict our scope to only finding the experiments induced by fair coin flipping, and focus on the trade-off behind the number and timing of the randomization points.

The trade-off lies between having too many randomization points (corresponding to large $K$) and too few randomization points (corresponding to small $K$). Intuitively, too many decreases the probability of observing consecutive treatments $\mathbf{1}_{m+1}$ or controls $\mathbf{0}_{m+1}$, which, in turn, decreases the amount of useful data. On the other hand, too few decreases the number of independent observations and reduces our ability to produce reliable results. Both of these scenarios reduce our ability to draw valid causal claims. This is the objective of Theorem 5.5.

*Theorem* 5.5 (Optimal Design). *Under Assumptions 5.1–5.3, the optimal solution to the design of regular switchback experiment as we have introduced in* (5.6) *is equivalent to the optimal solution to the following subset selection problem.*

$$\min_{\mathbb{T}\subset[T]} \left\{ 4\sum_{k=0}^{K}(t_{k+1}-t_k)^2 + 8m(t_K - t_1) + 4m^2K - 4m^2 + 4\sum_{k=1}^{K-1}[(m - t_{k+1} + t_k)^+]^2 \right\} \tag{5.7}$$

Table 5.1: An example of the optimal design $\mathbb{T}^* = \{1, 5, 7, 9\}$ when $T = 12$ and $p = m = 2$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbb{T}^*$ | ✓ | – | – | – | ✓ | – | ✓ | – | ✓ | – | – | – |

Note: Each checkmark beneath a time period $t$ indicates that $t$ is a randomization point.

*In particular, when* $m = 0$ *then* $\mathbb{T}^* = \{1, 2, 3, ..., T\}$*; when* $m > 0$*, and if there exists* $n \geq 4 \in \mathbb{N}$*, s.t.* $T = nm$*, then* $\mathbb{T}^* = \{1, 2m+1, 3m+1, ..., (n-2)m+1\}$*.*

The proof of Theorem 5.5 is deferred to Section D.3.6 in the appendix. Theorem 5.5 presents the optimal design in a class of perfect cases when the time horizon splits into several equal-length epochs. See Figure 5.1 for an example. In practice, when possible, we recommend selecting $T$ that satisfies the condition in Theorem 5.5. See Section 5.6 for a discussion. For other imperfect cases when $T$ is not divisible by $m$, we can also solve (5.7) and find the optimal design. However, we do not present closed-form solutions to such subset selection problem due to integrality issues. Technical discussions about the optimal design in such imperfect cases are deferred to Section D.3.6 in the Appendix.

There are two important implications of Theorem 5.5. First, the optimal randomization frequency depends on the physical duration of the carryover effect, regardless of the granularity of one single experimental period. This observation suggests that practitioners may set each period to be almost as long as the order of the carryover effect, which sheds some light on the selection of granularity when practitioners design the experiment. See Example 5.7. Second, an important special case arises when experimental designers believe there is very little carryover effect, in which case the optimal designs are almost the same. This observation suggest a layer of robustness. See Example 5.8.

*Example* 5.7 (Two Granularity Levels)*.* In the ride-sharing application, suppose the firm has two options to treat one single time period either as 0.5 hour or 1 hour; and suppose the carryover effect lasts for 2 hours. When one single experimental period corresponds to 0.5 hour, the carryover effect lasts for $m = 4$ periods. When one single experimental period corresponds to 1 hour, the carryover effect lasts for $m = 2$ periods. From Theorem 5.5, the optimal design exhibits an optimal structure that randomizes once every $m$ periods (except for the first and last epoch, which lasts for $2m$ time periods each). In both cases, the optimal design would randomize once every two hours. ∎

*Example* 5.8 (Little Carryover Effect)*.* For example, Theorem 5.5 suggests that the optimal design when $m = 0$ is $\mathbb{T}^* = \{1, 2, 3, ..., T\}$, and when $m = 1$ is $\mathbb{T}^* = \{1, 3, 4, ..., T - 1\}$. This suggests that the minimax optimal design in the absence of a carryover effect is robust to the existence of a short carryover effect.

## 5.4 Inference and Statistical Testing

After designing and running the experiment, we obtain two time series. The first is the observed assignment path $\boldsymbol{w}_{1:T}^{\mathsf{obs}}$, and the second is the corresponding observed outcomes $\boldsymbol{Y}_{1:T}^{\mathsf{obs}}$. See Figure 5-3. To draw inference from this data we propose two methods, an exact randomization based test and a finite population conservative test that establishes asymptotic result.

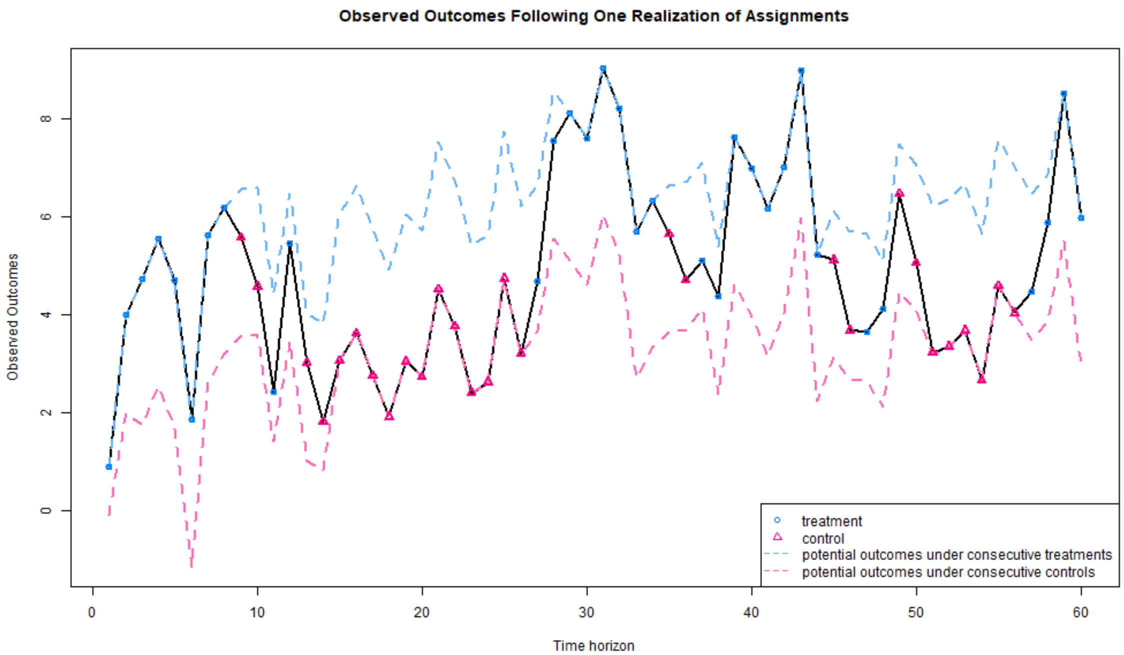
In Sections 5.4.1 and 5.4.2, we assume perfect knowledge of $m$, i.e., $p = m$, and we will write $\tau_m$ and $\hat{\tau}_m$ to stand for $\tau_p$ and $\hat{\tau}_p$, respectively. We discuss in Section 5.4.3 the case when $p \neq m$ and show that our inference methods are still valid. To conclude this section, we provide in Section 5.4.4 a data-driven procedure to identify a possible value for the carryover effect by running multiple experiments. Such a procedure relaxes Assumption 5.2 and is of great practical relevance.

### 5.4.1 Exact Inference

We propose an exact non-parametric test for the sharp null of no effect at every time point (Fisher et al. 1937, Rubin 1980):

$$H_0 : Y_t(\boldsymbol{w}_{t-m:t}) - Y_t(\boldsymbol{w}'_{t-m:t}) = 0 \quad \text{for all } \boldsymbol{w}_{t-m:t}, \boldsymbol{w}'_{t-m:t}, \quad t \in \{m + 1 : T\}. \quad (5.8)$$

Figure 5-3: Illustrator of the observed assignment path $\boldsymbol{w}_{1:T}^{\mathsf{obs}}$ (blue and red dots) and the observed outcomes $\boldsymbol{Y}_{p+1:T}^{\mathsf{obs}}$ (black curve)




Note: The dashed lines are the potential outcomes under consecutive treatments / controls.

The sharp null hypothesis implies that $Y_t(\boldsymbol{w}_{t-m:t}^{\mathsf{obs}}) = Y_t(\boldsymbol{w}_{t-m:t})$ for all $\boldsymbol{w}_{t-m:t} \in \{0,1\}^t$. That is, regardless of the assignment path $\boldsymbol{w}_{t-m:t}$ we would have observed the same outcomes.

We can conduct exact tests by using the known assignment mechanism to simulate new assignment paths; see Algorithm 8 for details. The test depends on the observation that, under the sharp null hypothesis of no treatment effect (5.8), any assignment path $\boldsymbol{w}_{1:T}^{[i]}$ leads to the same observed outcomes. In particular, in Step 3, we assume the observed outcomes remain unchanged. Thus all treatment paths lead to the same observed outcomes $Y_{m+1:T}^{\mathsf{obs}}$. To obtain a confidence interval, we propose inverting a sequence of exact hypothesis tests to identify the region outside of which (5.8) is violated at the prespecified nominal level (Imbens and Rubin 2015, Chapter 5). In practice, obtaining confidence intervals through this approach is somewhat challenging; instead, we refer the reader to the subsequent section that provides a less computationally intensive approach.

**Algorithm 8** Algorithm for performing a sharp-null hypothesis test

**Require:** Fix $I$, total number of samples drawn.

1: **for** i in $1:I$ **do**

2: Sample a new assignment path $\boldsymbol{w}_{1:T}^{[i]}$ according to the assignment mechanism.

3: Hold $Y_{p+1:T}^{\mathsf{obs}}$ unchanged. Compute $\hat{\tau}^{[i]}$ according to (5.4),

$$\hat{\tau}^{[i]} = \frac{1}{T-m} \sum_{t=m+1}^{T} \left\{ Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{w}_{t-m:t}^{[i]} = \mathbf{1}_{m+1}\}}{\Pr(\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1})} - Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{w}_{t-m:t}^{[i]} = \mathbf{0}_{m+1}\}}{\Pr(\boldsymbol{W}_{t-m:t} = \mathbf{0}_{m+1})} \right\}.$$

4: **end for**

5: Compute $\hat{p}_{\mathsf{F}} = I^{-1} \sum_{i=1}^{I} \mathbb{1}\left\{ \left|\hat{\tau}^{[i]}\right| > |\hat{\tau}| \right\}$

6: **return** $\hat{p}_{\mathsf{F}}$, the estimated $p$-value. For large $I$, this is exact.

### 5.4.2 Asymptotic Inference

We now introduce a conservative test for the null of no average treatment effect:

$$H_0 : \tau_m = \frac{1}{T-m} \sum_{t=m+1}^{T} [Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{0}_{m+1})] = 0. \tag{5.9}$$

To test such a null, we derive a finite population central limit theorem to approximate the distribution of the Horvitz-Thompson estimator.

Assume $n = T/m \geq 4$ is an integer, then under the optimal design as shown in Theorems 5.4 and 5.5, the assignment path is determined by the realizations at $W_1, W_{2m+1}, ..., W_{(n-2)m+1}$. To make the dependence on randomization clear, we introduce the following notations. For any $k \in \{0, 1, ..., n-2\}$, let $\bar{Y}_k(\mathbf{1}_{m+1}) = \sum_{t=(k+1)m+1}^{(k+2)m} Y_t(\mathbf{1}_{m+1})$ and $\bar{Y}_k(\mathbf{0}_{m+1}) = \sum_{t=(k+1)m+1}^{(k+2)m} Y_t(\mathbf{0}_{m+1})$. Moreover, for any $k \in \{0, 1, ..., n-2\}$, let $\bar{Y}_k^{\mathsf{obs}} = \sum_{t=(k+1)m+1}^{(k+2)m} Y_t^{\mathsf{obs}}$ be the sum of the observed outcomes.

*Lemma* 5.6 (Variance of the Horvitz-Thompson Estimator Under the Optimal Design). *Under Assumptions 5.1–5.3 and under the optimal design as shown in Theorems 5.4 and 5.5, if $n = T/m \geq 4$ is an integer, then the variance of the Horvitz-*

*Thompson estimator,* $\mathsf{Var}(\hat{\tau}_m)$*, is*

$$\begin{aligned}\mathsf{Var}(\hat{\tau}_m) = \frac{1}{(T-m)^2} \Bigg\{ & \bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2 + 2\bar{Y}_0(\mathbf{1}_{m+1})\bar{Y}_0(\mathbf{0}_{m+1}) \\ & + \sum_{k=1}^{n-3} \left[3\bar{Y}_k(\mathbf{1}_{m+1})^2 + 3\bar{Y}_k(\mathbf{0}_{m+1})^2 + 2\bar{Y}_k(\mathbf{1}_{m+1})\bar{Y}_k(\mathbf{0}_{m+1})\right] \\ & + \bar{Y}_{n-2}(\mathbf{1}_{m+1})^2 + \bar{Y}_{n-2}(\mathbf{0}_{m+1})^2 + 2\bar{Y}_{n-2}(\mathbf{1}_{m+1})\bar{Y}_{n-2}(\mathbf{0}_{m+1}) \\ & + \sum_{k=0}^{n-3} 2\left[\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1})\right] \cdot \left[\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1})\right] \Bigg\}\end{aligned} \tag{5.10}$$

Lemma 5.6 provides the variance of the Horvitz-Thompson estimator under the optimal design. Since we never observe all the potential outcomes, most of the cross-product terms in (5.10) can not be directly estimated. Instead, we provide the following upper bound to (5.10) and propose an unbiased estimator.

*Corollary* 5.6.1. *Under the conditions in Lemma 5.6, there exists an upper bound for the variance of the Horvitz-Thompson estimator,* $\mathsf{Var}(\hat{\tau}_m) \leq \mathsf{Var}^{\mathsf{U}}(\hat{\tau}_m)$*, which can be estimated by* $\hat{\sigma}^2_{\mathsf{U}}$*, defined as:*

$$\hat{\sigma}^2_{\mathsf{U}} = \frac{1}{(T-m)^2} \left\{ 8(\bar{Y}_0^{\mathsf{obs}})^2 + \sum_{k=1}^{n-3} 32(\bar{Y}_k^{\mathsf{obs}})^2 \mathbb{1}\{W_{km+1} = W_{(k+1)m+1}\} + 8(\bar{Y}_{n-2}^{\mathsf{obs}})^2 \right\}.$$

*Moreover,* $\hat{\sigma}^2_{\mathsf{U}}$ *is unbiased, i.e.,* $\mathbb{E}[\hat{\sigma}^2_{\mathsf{U}}] = \mathsf{Var}^{\mathsf{U}}(\hat{\tau}_m)$.

Corollary 5.6.1 provides the foundation to make conservative inference. We make the following technical assumption for the asymptotic normal distribution to hold.

*Assumption* 5.4 (Non-negligible Variance). *Assume that the randomization distribution has a non-negligible variance, i.e.,*

$$\mathsf{Var}(\hat{\tau}_m) \geq \Omega(n^{-1}). \tag{5.11}$$

*In particular, one sufficient condition for* (5.11) *is to assume that all the potential out-*

*comes are positive, i.e., there exists some constant* $b > 0$*, such that* $\forall t \in [T], \forall \boldsymbol{w}_{1:T} \in \{0,1\}^T$*,* $Y_t(\boldsymbol{w}_{1:T}) \geq b$*.*

Intuitively, the key to most central limit theorems is that all the variables roughly have variances of the same order. In other words, there cannot be a small number of variables that compromise the majority of the variance. Since under Assumption 5.3 the potential outcomes are bounded, each variable contributes to the total variance of order $O(n^{-2})$. Assumption 5.4 suggests that the total variance is large enough, such that it cannot come from only a few of the time periods.

*Theorem* 5.7 (Asymptotic Normality). *Let* $m$ *be fixed. For any* $n \geq 4 \in \mathbb{N}$*, define an* $n$*-replica experiment such that there are* $T = nm$ *time periods. We take the optimal design as in Theorem 5.5 whose randomization points are at* $\mathbb{T}^* = \{1, 2m+1, 3m+1, ..., (n-2)m+1\}$*. Under Assumptions 5.1–5.2, and under Assumption 5.4, the limiting distribution of the Horvitz-Thompson estimator in the* $n$*-replica experiment has an asymptotic normal distribution. That is, let* $\mathsf{Var}(\hat{\tau}_m)$ *be defined in Lemma 5.6. As* $n \to +\infty$*,*

$$\frac{\hat{\tau}_m - \tau_m}{\sqrt{\mathsf{Var}(\hat{\tau}_m)}} \xrightarrow{D} \mathcal{N}(0,1).$$

Theorem 5.7 is in the spirit of the finite population central limit theorems as in Li and Ding (2017), Aronow et al. (2017), Chin (2018), Bojinov et al. (2019, 2020a), Han et al. (2021). Note that, Theorem 5.7 does not require $\mathsf{Var}(\hat{\tau}_m)$ to converge as $n \to +\infty$.

To conduct inference, we replace $\mathsf{Var}(\hat{\tau}_m)$ by $\hat{\sigma}^2_{\mathsf{U}}$ as provided in Corollary 5.6.1. Define the test statistic to be $z = |\hat{\tau}_m| / \sqrt{\hat{\sigma}^2_{\mathsf{U}}}$. When the alternative hypothesis is two-sided, the estimated $p$-value is given by $\hat{p}_{\mathsf{N}} = 2 - 2\Phi(z)$, where $\Phi$ is the CDF of a standard normal distribution.

The proofs of Lemma 5.6, Corollary 5.6.1, and Theorem 5.7 are deferred to Sections D.4.2, D.4.3, and D.4.4 in the Appendix, respectively.

### 5.4.3 Inference under Misspecified $m$

Up to now, we assumed that we knew the order of the carryover effect $m$, and set $p = m$. In practice, we may not know the exact value of the carryover effect, and we have to select $p$ either based on domain knowledge or the procedure we recommend in Section 5.4.4. In this section, we consider what happens when $p \neq m$ and show that the estimation and inference are still valid and meaningful, although the design from Theorem 5.5 is no longer optimal.

Below we distinguish two cases: $p > m$ and $p < m$. When $p > m$, due to Assumption 5.2, $Y_t(\mathbf{1}_{p+1}) = Y_t(\mathbf{1}_{m+1}), \forall t \in \{p+1 : T\}$, and the lag-$p$ causal effect is essentially the lag-$m$ causal effect. So all the estimation and inference results still hold.

However, when $p < m$, the Horvitz-Thompson estimator (5.4) will be biased for the causal estimand. See Section D.4.5 for more discussions. When $p < m$, the exact inference procedure as in Section 5.4.1 remains valid. For the asymptotic inference procedure, a similar result to Theorem 5.7 still holds when $m$ is misspecified, as we state in Corollary 5.7.1. The only difference is that when $p < m$, the asymptotic normal distribution will not be centered around the causal estimand as we defined in (5.1), but some quantity that we will discuss in Section D.4.5. The proof is deferred to Section D.4.7 in the Appendix.

*Corollary* 5.7.1 (Asymptotic Normality when $m$ is Misspecified). *For any $n \geq 4 \in \mathbb{N}$, define an $n$-replica experiment such that there are $T = np$ time periods. Take the optimal design as in Theorem 5.5 whose randomization points are at $\mathbb{T}^* = \{1, 2p+1, 3p+1, ..., (n-2)p+1\}$. We have the following two observations.*

*i When $p > m$, under Assumptions 5.1–5.2, the variance of the Horvitz-Thompson estimator, $\mathsf{Var}(\hat{\tau}_p)$, is explicitly given by* (5.10)*.*

*ii Furthermore, no matter if $p > m$ or $p < m$, under Assumptions 5.1–5.3 and assume $\mathsf{Var}(\hat{\tau}_p) \geq \Omega(n^{-1})$, the limiting distribution of the Horvitz-Thompson estimator in the $n$-replica experiment has an asymptotic normal distribution.*

*That is, as $n \to +\infty$,*

$$\frac{\hat{\tau}_p - \tau_p}{\sqrt{\mathsf{Var}(\hat{\tau}_p)}} \xrightarrow{D} \mathcal{N}(0,1).$$

Corollary 5.7.1, together with Theorem 5.7, is the key to identification of $m$, the order of the carryover effect. In Section 5.4.4 we provide a procedure to identify $m$.

### 5.4.4 Identifying the Order of the Carryover Effect

We borrow Theorem 5.7 and Corollary 5.7.1 to define a hypothesis testing procedure, which, combined with a searching method, estimates the order of the carryover effect.

To build intuition, suppose we have access to two comparable experimental units. The two experimental units could be two separate units or two time epochs on one experimental unit such that the two epochs are far enough such that the carryover effect from one does not affect the outcomes of the other. Suppose, on the first experimental unit, we design an optimal experiment under $p = p_1$ and on the second unit, we use $p = p_2$; without loss of generality let $p_1 < p_2$.

After running the experiment and collecting the results, consider the following two statistics. For the first unit, we calculate $\hat{\tau}_{p_1}$, the sampling average, and $\hat{\sigma}^2_{p_1}$, the conservative sampling variance as suggested by Corollary 5.6.1. For the second unit, we calculate $\hat{\tau}_{p_2}$ and $\hat{\sigma}^2_{p_2}$.

Define a procedure that tests the following null hypothesis:

$$H_0: \ m \leq p_1 \tag{5.12}$$

Under the null hypothesis (5.12), $\tau_{p_1} = \tau_{p_2} = \tau_m$, and so both $\hat{\tau}_{p_1}$ and $\hat{\tau}_{p_2}$ are unbiased estimators of $\tau_m$. Furthermore, given that the two estimators both conform asymptotic normal distributions, and that the two experimental units are independent, the difference between the two estimators should be an asymptotic normal distribution centered around zero, i.e., $(\hat{\tau}_{p_1} - \hat{\tau}_{p_2})/\sqrt{\mathsf{Var}(\tau_{p_1}) + \mathsf{Var}(\tau_{p_2})} \xrightarrow{D} \mathcal{N}(0,1)$. To test the null hypothesis (5.12), define the test statistic to be $z = |\hat{\tau}_{p_1} - \hat{\tau}_{p_2}| / \sqrt{\hat{\sigma}^2_{p_1} + \hat{\sigma}^2_{p_2}}$. The estimated $p$-value is given by $\hat{p} = 2 - 2\Phi(z)$, where $\Phi$ is the CDF of a standard normal distribution.

The above procedure enables us to test the null hypothesis (5.12). We can combine such a procedure with any searching method to identify $m$.

## 5.5 Simulation Study

There are five goals for this simulation study. First, to show that the optimal design in Theorem 5.5 has the smallest risk compared against two benchmarks. There are two dimensions for our comparison: the worst-case risk and the risk under a specific outcome model. Second, to verify the asymptotic normal distribution under a non-asymptotic setup, and to study the quality of the upper bound proposed in Corollary 5.6.1. Third, to understand the rejection rate and its dependence on the length of time horizon. Fourth, to study the performance of the optimal design under a misspecified $m$, and to compare the difference of the two inference methods proposed in Section 5.4. Fifth, to study the performance of the hypothesis testing procedure as proposed in Section 5.4.4, which identifies $m$ the length of the carryover effect.

We start with a simple linear additive carryover effect model which originates from Oman and Seiden (1988), Hedayat et al. (1978), Jones and Kenward (2014).

$$Y_t(\boldsymbol{w}_{1:t}) = \mu + \alpha_t + \delta^{(1)} w_t + \delta^{(2)} w_{t-1} + ... + \delta^{(t)} w_1 + \epsilon_t \tag{5.13}$$

where $\mu$ is a fixed effect; $\alpha_t$ is a fixed effect associated to period $t$; $\delta^{(1)}, \delta^{(2)}, ..., \delta^{(t)}$ are non-stochastic coefficients; $w_t, w_{t-1}, ..., w_1$ are the treatment indicators; $\epsilon_t$ is the random noise in period $t$. We will run many simulations based on this model. For a more detailed discussion of the flexibility of the potential outcome framework, see Section D.5.1 in the Appendix.

### 5.5.1 Comparison of the Risk Functions

#### Simulation setup

We consider two setups. The first setup is for the worst-case risk. We consider $T = 120, p = m = 2$ where $m$ is correctly identified, and $Y_t(\mathbf{1}_3) = Y_t(\mathbf{0}_3) = 10$. We compare three different designs of switchback experiments. The first one is our proposed optimal design as in Theorem 5.5, such that $\mathbb{T}^* = \{1, 5, 7, ..., 117\}$. The second one is the most common and naive switchback experiment, which independently assign treatment/control in every period with half-half probability. It is parameterized by $\mathbb{T}^{\mathsf{H1}} = \{1, 2, 3, ..., 120\}$. The third one is the "intuitive" experiment discussed in Table 5.1, which divides the time horizon into several epochs each with length $m+1 = 3$. It is parameterized by $\mathbb{T}^{\mathsf{H2}} = \{1, 4, 7, ..., 118\}$.

Second, we run simulations based on the outcome model as in (5.13). Similar to the first setup, we consider again $T = 120, p = m = 2$ where $m$ is correctly identified. For the outcome model, we consider $\mu = 0$, $\alpha_t = \log(t)$, and $\epsilon_t \sim N(0, 1)$ are i.i.d. standard normal distributions. For any $t > 3$, let $\delta^{(t)} = 0$. We will vary the values of $\delta^{(1)}, \delta^{(2)}, \delta^{(3)} \in \{1, 2\}$ and conduct experiments under $2^3 = 8$ different scenarios. Again we compare the same three different designs of switchback experiments. $\mathbb{T}^* = \{1, 5, 7, ..., 117\}, \mathbb{T}^{\mathsf{H1}} = \{1, 2, 3, ..., 120\}$, and $\mathbb{T}^{\mathsf{H2}} = \{1, 4, 7, ..., 118\}$.

We simulate one assignment path at a time, and conduct an experiment following this assignment path. Since the outcome model is prescribed, we can calculate both the causal estimand and and the observed outcomes (along the simulated assignment path). Then, we calculate the Horvitz-Thompson estimator based on the simulated assignment path and the simulated observed outcomes. With both the estimand and estimator, we can calculate the loss function. We repeat the above procedure enough (100000) times to obtain an accurate approximation of the risk function.

#### Simulation results

First, we calculate the worst-case risk functions via simulations. Notice that, when $p = m = 2$, we could explicitly calculate the worst-case risk functions under the three different designs of switchback experiments $\mathbb{T}^*, \mathbb{T}^{\mathsf{H1}}$, and $\mathbb{T}^{\mathsf{H2}}$. Even though we can explicitly calculate them via the following expression (See Lemma D.9 in the Appendix for details),

$$\frac{B^2}{(T-m)^2}\left\{4\sum_{k=1}^{K+1}(t_k-t_{k-1})^2+8m(t_K-t_1)+4m^2K-4m^2+4\sum_{k=2}^{K}[(m-t_k+t_{k-1})^+]^2\right\}, \tag{5.14}$$

we still use the simulation to confirm this result. See Table 5.2 for our simulation results.

The causal effect is $\tau_2 = 0$ because $Y_t(\mathbf{1}_3) = Y_t(\mathbf{0}_3) = 10$. The simulated estimator is $\mathbb{E}[\hat{\tau}_2^*] = -0.0291$ for our proposed optimal design, and $\mathbb{E}[\hat{\tau}_2^{\mathsf{H1}}] = 0.0104$ and $\mathbb{E}[\hat{\tau}_2^{\mathsf{H2}}] = -0.0478$ for the two benchmarks, respectively. The risk function is $r(\eta_{\mathbb{T}^*}) = 26.78$ for our proposed optimal design, and $r(\eta_{\mathbb{T}^{\mathsf{H1}}}) = 33.67$ and $r(\eta_{\mathbb{T}^{\mathsf{H1}}}) = 27.85$ for the two benchmarks, respectively. Such simulation results suggest that our proposed optimal design have the smallest risk, under the worst case outcome model. In the last three columns are the risk functions of the three designs, all suggested by expression (5.14). The risk functions calculated from theory take values that are very close to the risk functions calculated from expression (5.14), which verifies our theory.

Table 5.2: Simulation results for the worst-case risk function

| $\tau_2$ | $\mathbb{E}[\hat{\tau}_2^*]$ | $\mathbb{E}[\hat{\tau}_2^{\mathsf{H1}}]$ | $\mathbb{E}[\hat{\tau}_2^{\mathsf{H2}}]$ | $r(\eta_{\mathbb{T}^*})$ | $r(\eta_{\mathbb{T}^{\mathsf{H1}}})$ | $r(\eta_{\mathbb{T}^{\mathsf{H2}}})$ | $\tilde{r}(\eta_{\mathbb{T}^*})$ | $\tilde{r}(\eta_{\mathbb{T}^{\mathsf{H1}}})$ | $\tilde{r}(\eta_{\mathbb{T}^{\mathsf{H2}}})$ |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.0250 | 0.0200 | 0.0059 | 26.78 | 33.67 | 27.85 | 26.67 | 33.96 | 27.81 |

Note: The optimal design $\mathbb{T}^*$ as suggested by Theorem 5.5 yields the smallest risk, both in theory and confirmed by simulations.

Second, we calculate the risk functions based on the outcome model in (5.13). See Table 5.3. As we vary the values of $\delta^{(1)}$, $\delta^{(2)}$ and $\delta^{(3)}$, the total lag-2 causal effect is being changed. All three estimators are able to reflect the change as the estimand changes. The risk function can be simulated and we see that the risk function associated with the first benchmark $\mathbb{T}^{\mathsf{H1}}$ is $28\% \sim 32\%$ larger than the optimal design; and the second benchmark $\mathbb{T}^{\mathsf{H2}}$ is $1\% \sim 2\%$ larger. Such simulation results suggest again that our proposed optimal design have the smallest risk. Moreover, as $r(\eta_{\mathbb{T}^{\mathsf{H2}}})$ is close

Table 5.3: Simulation results for the risk function based on the outcome model in (5.13)

| $\delta^{(1)}$ | $\delta^{(2)}$ | $\delta^{(3)}$ | $\tau_2$ | $\mathbb{E}[\hat{\tau}_2^*]$ | $\mathbb{E}[\hat{\tau}_2^{\mathsf{H1}}]$ | $\mathbb{E}[\hat{\tau}_2^{\mathsf{H2}}]$ | $r(\eta_{\mathbb{T}^*})$ | $r(\eta_{\mathbb{T}^{\mathsf{H1}}})$ | $r(\eta_{\mathbb{T}^{\mathsf{H2}}})$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 3 | 3.016 | 3.012 | 3.002 | 7.96 | 10.22 | 8.11 |
| 1 | 1 | 2 | 4 | 4.018 | 4.013 | 4.002 | 9.57 | 12.39 | 9.74 |
| 1 | 2 | 1 | 4 | 4.018 | 4.013 | 4.002 | 9.57 | 12.39 | 9.74 |
| 2 | 1 | 1 | 4 | 4.018 | 4.013 | 4.002 | 9.57 | 12.39 | 9.74 |
| 1 | 2 | 2 | 5 | 5.020 | 5.015 | 5.003 | 11.34 | 14.81 | 11.52 |
| 2 | 1 | 2 | 5 | 5.020 | 5.015 | 5.003 | 11.34 | 14.81 | 11.52 |
| 2 | 2 | 1 | 5 | 5.020 | 5.015 | 5.003 | 11.34 | 14.81 | 11.52 |
| 2 | 2 | 2 | 6 | 6.022 | 6.016 | 6.003 | 13.28 | 17.48 | 13.47 |

Note: For each row, the random seed that generates the simulation setup is fixed. The optimal design $\mathbb{T}^*$ as suggested in Theorem 5.5, though solved from a minimax program, still yields the smallest risk for the outcome model in (5.13). A few rows are redundant because our switchback experiment, combining with the causal estimand (5.1), is only able to measure the total additive treatment effect. We cannot distinguish the source of the additive treatment effects, i.e., we are unable to distinguish $\delta^{(1)}$, $\delta^{(2)}$, and $\delta^{(3)}$.

to $r(\eta_{\mathbb{T}^*})$ and both are much smaller than $r(\eta_{\mathbb{T}^{\mathsf{H1}}})$, our results suggest that when $m$ is unknown, it is better to select $p$ to be slightly larger than the true $m$ as opposed to significantly smaller.

As the magnitude of treatment effects increase, the associated risk functions also increase. The relative difference between risk functions of $r(\eta_{\mathbb{T}^{\mathsf{H1}}})$ and $r(\eta_{\mathbb{T}^*})$ increases, while the relative difference between $r(\eta_{\mathbb{T}^{\mathsf{H1}}})$ and $r(\eta_{\mathbb{T}^*})$ decreases. This coincides with the intuitions discussed in Section 5.3.

### 5.5.2 Asymptotic Normality

#### Simulation setup

We run simulations based on the outcome model in (5.13), with $T = 120$ and $m = 2$. We will consider three cases: (i) $m$ is correctly specified so $p = 2$; (ii) $p = 3$, and we estimate lag-3 causal estimand as in (5.1); (iii) $p = 1$, and we pretend as if we estimated the lag-1 causal estimand. However, as the lag-1 causal estimand is not well defined, we instead estimate a different quantity, which we refer to as the "$m$-misspecified lag-$p$ causal estimand" (See details and definition in (D.10)).

For the outcome model, we consider $\mu = 0$, $\alpha_t = \log(t)$, and $\epsilon_t \sim N(0,1)$ are i.i.d. standard normal distributions. For any $t > 3$, let $\delta^{(t)} = 0$. For simplicity, let $\delta^{(1)} = \delta^{(2)} = \delta^{(3)} = \delta$. We vary $\delta \in \{1, 2, 3\}$ and conduct experiments under 3 different scenarios. We simulate one assignment path at a time, and conduct experiments following this assignment path. Since the outcome model is prescribed, we calculate the observed outcomes based on the simulated assignment path. Then we calculate the Horvitz-Thompson estimator, and the conservative estimator of the randomization variance (Corollary 5.6.1), based on the simulated assignment path and the simulated observed outcomes. On the other hand, the lag-$p$ causal estimand is easy to calculate once the outcome model is prescribed. Yet the $m$-misspecified lag-$p$ causal estimand has to be calculated in conjunction with the simulated assignment path. By repeating the above procedure enough (100000) times we obtain a distribution of the estimator.

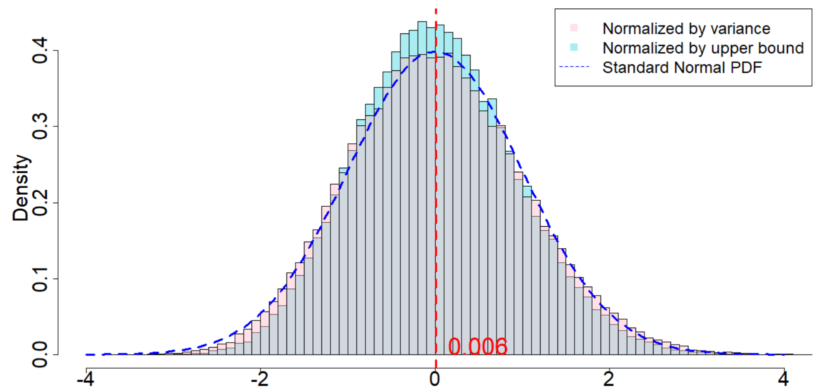
**Simulation results**

In Figure 5-4, the dotted dark blue line is the Probability Density Function of the standard normal distribution. The pink histogram corresponds to the distribution induced by $\frac{\hat{\tau}_p - \tau_p}{\sqrt{\mathsf{Var}(\hat{\tau}_p)}}$, which is the estimator (after re-centering at zero) normalized by the square root of the true randomization variance[3]. Such a distribution, as suggested by Theorem 5.7, converges to a standard normal distribution when $T$ is large. Comparing to the dotted dark blue line, Figure 5-4 suggests that Theorem 5.7 approximately holds for moderate values of $T$. The light blue histogram corresponds to the distribution induced by $\frac{\hat{\tau}_p - \tau_p}{\sqrt{\mathbb{E}[\hat{\sigma}_U^2]}}$, which is the estimator (after re-centering at zero) normalized by the expectation of the conservative upper bound of the randomization variance. Since we replace the true variance by the conservative upper bound, the shape of the distribution is more concentrated around zero, as we see from the "taller" histogram. The red vertical line is the expected value of the randomization distribution for the pink histogram. The cases of $\delta = 1$ and $\delta = 2$ are similar, and the

[3]We numerically find such variance $\mathsf{Var}(\hat{\tau}_p)$, and the expectation of the conservative upper bound $\mathbb{E}[\hat{\sigma}_U^2]$

cases of overestimated $m$ and underestimated $m$ are also similar. We discuss them in Section D.5.2 in the Appendix.

Figure 5-4: Approximate normality of the randomization distribution when $m = 2, p = 2, \delta = 3$.




For all the nine cases ($p \in \{1, 2, 3\}$ and $\delta \in \{1, 2, 3\}$), see Table 5.4 for the expected values and the variances of the randomization distributions, as well as the conservative estimator of the randomization variances. Note that the three cases all have the same underlying outcome model. It is the different knowledge of $m$ that leads to three different designs of experiments.

Table 5.4: Simulation results for the randomization distribution

| | | $\tau_p$ | $\tau_p^{[m]}$ | $\mathbb{E}[\hat{\tau}_p]$ | $\mathsf{Var}(\hat{\tau}_p)$ | $\mathbb{E}[\hat{\sigma}^2_{\mathsf{U}}]$ |
|---|---|---|---|---|---|---|
| $m = 2, p = 2$ | $\delta = 1$ | 3 | – | 3.016 | 7.96 | 8.48 |
| | $\delta = 2$ | 6 | – | 6.022 | 13.28 | 15.16 |
| | $\delta = 3$ | 9 | – | 9.028 | 20.10 | 24.25 |
| $m = 2, p = 3$ | $\delta = 1$ | 3 | – | 3.006 | 11.92 | 12.67 |
| | $\delta = 2$ | 6 | – | 6.009 | 19.89 | 22.70 |
| | $\delta = 3$ | 9 | – | 9.012 | 30.10 | 36.32 |
| $m = 2, p = 1$ | $\delta = 1$ | – | 2 | 2.016 | 4.00 | 4.13 |
| | $\delta = 2$ | – | 4 | 4.026 | 6.69 | 7.06 |
| | $\delta = 3$ | – | 6 | 6.037 | 10.14 | 10.92 |

Note: The randomization distribution is unbiased in all 9 cases (when $p < m$ it is unbiased for the $m$-misspecified average lag-1 causal effect). The conservative estimation of the variance upper bound from Corollary 5.6.1 is close to the true variance.

From Table 5.4, we make the following two observations. **(i) Unbiasedness of the Horvitz-Thompson estimator**. When $m$ is correctly specified, $\mathbb{R}[\hat{\tau}_p]$ is

very close to $\tau_p$, verifying the unbiasedness of the estimator. When $m = 2, p = 3$, the estimand remains unchanged, and the estimator remains unbiased. But the variance of the estimator is larger. When $m = 2, p = 1$, the estimand is the $m$-misspecified estimand, and the estimator is unbiased for this $m$-misspecified estimand.
**(ii) Quality of Corollary 5.6.1 and 5.7.1**. As we increase $\delta$, the variances of the randomization distributions also increase. The conservative estimators of the randomization variances are very close to the true variances, which suggests that Corollary 5.6.1 and 5.7.1 approximate the true variances quite well.

#### Robustness check

In this section we run simulations under almost the same setup as introduced in Section 5.5.2, with the only difference that we select each $\epsilon_t$ to be an i.i.d. Student's t-distribution with 1 degree of freedom. The purpose of this section is to verify our theory when $\epsilon_t$ are drawn from heavy tailed distributions.

When $m = 2, p = 2, \delta = 1$, as we can see from Figure 5-5, the randomization distribution is significantly different from a standard normal distribution. This is because $T = 120$ is too small. Alternatively, we increase $T = 1200$ to see that the randomization distribution behaves like a normal distribution. In other words, when $\epsilon_t$ noises are heavy tailed, our Theorem 5.7 has a slower convergence rate to a normal distribution. We conduct extensive simulation study under other parameters, as we will show in Section D.5.2 in the Appendix.

### 5.5.3 Rejection Rates

#### Simulation setup

In this simulation, we run multiple simulations based on the outcome model as in (5.13). We vary $T \in \{120, 240, ..., 1200\}$. We consider $p = m = 2$ where $m$ is correctly specified. Similar to Section 5.5.2, we consider the same parameterization and conduct experiments under 3 different scenarios $\delta \in \{1, 2, 3\}$.

We simulate one assignment path at a time, and conduct experiments following

Figure 5-5: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 2, \delta = 1, T = 120$.

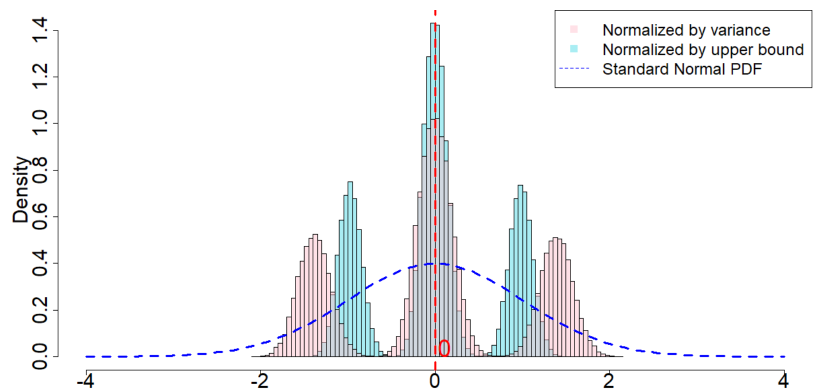
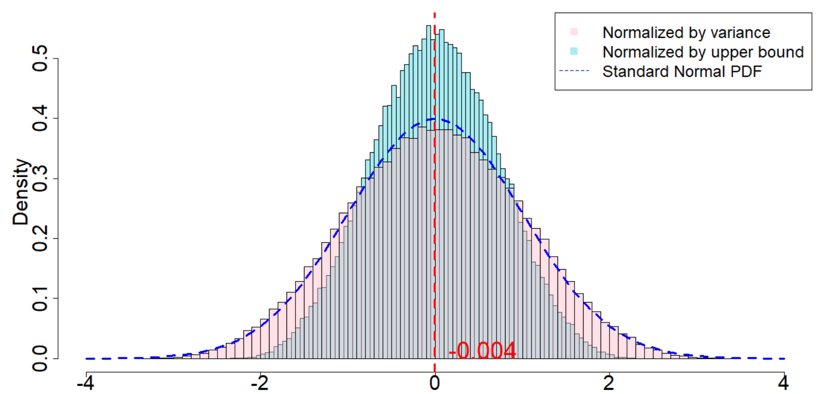



Figure 5-6: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 2, \delta = 1, T = 1200$.

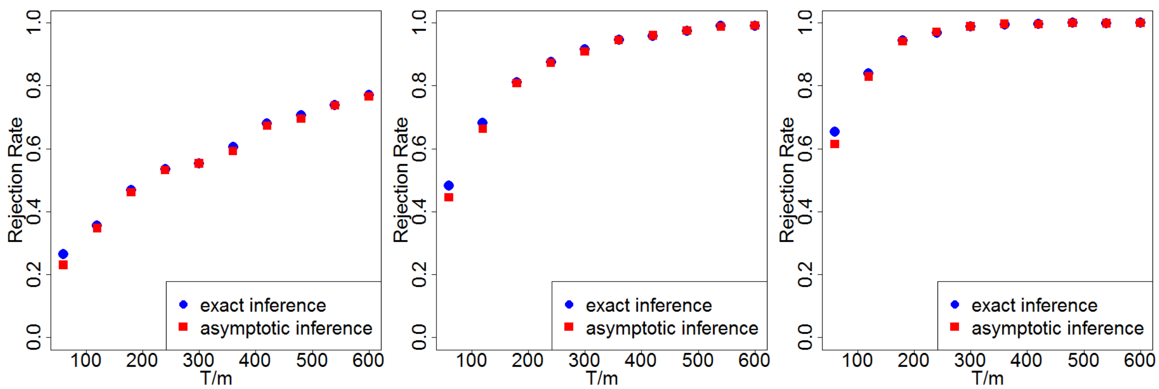



this assignment path. We first calculate the observed outcomes and the Horvitz-Thompson estimator. Then we conduct the two inference methods as proposed in Section 5.4, and obtain two estimated $p$-values. For the asymptotic inference method, we plug in $\hat{\sigma}_{\mathsf{U}}^2$, the conservative upper bound of the variance. We reject the corresponding null hypothesis when the $p$-value is smaller than 0.1 (In Section D.5.3 we run additional simulations by replacing such 0.1 threshold by 0.05 and 0.01). By repeating the above procedure enough (in this simulation, 1000) times we obtain the frequency of a null hypothesis being rejected, which we refer to as the rejection rate.

### Simulation results

We calculate the rejection rates via simulations and then plot Figure 5-7. The blue dots are rejection rates under exact inference; the red dots are under asymptotic inference. In all the simulations, $\delta \neq 0, \tau_p \neq 0$. So, ideally, we would wish to reject both the Fisher's null hypothesis (5.8) and the Neyman's null hypothesis (5.9).

Figure 5-7: Rejection rates and their dependence on $T/m$. Left: $\delta = 1$; Middle: $\delta = 2$; Right: $\delta = 3$




From Figure 5-7 we make the following three observations. **(i) Dependence on** $T/m$. The rejection rates increase as the length of the horizon increases – more specifically, as $T/m$ the total number of epochs increases. In practice, when firms have to capability to choose the length of $T$, they can refer to Figure 5-7 to choose $T$ properly. Also see discussion in Section 5.6. **(ii) Between two inference methods**. In all three cases, the rejection rate from testing a sharp null hypothesis (5.8) is slightly higher than that from testing the Neyman's null (5.9). This coincides with our intuition that a sharp null is more likely to be rejected. We discuss this in Section 5.5.4 together with the associated $p$-values. **(iii) Dependence on the signal-to-noise ratio**. The rejection rates all increase as $\delta$ increases from 1 to 3 (while holding the noise from the model fixed). This suggests that when the treatment effect is relatively larger, we do not require a long experimental horizon to achieve a desired rejection rate.

### 5.5.4 Estimation under a Misspecified $m$

#### Simulation setup

We run simulations based on the outcome model as in (5.13). We consider $T = 120, m = 2$. We consider three cases: (i) $m$ correctly specified so $p = 2$; (ii) $p = 3$, and we estimate the lag-3 causal estimand as in (5.1); (iii) $p = 1$, and we pretend as if we estimated the lag-1 causal estimand. However, the lag-1 causal estimand is not well defined. Instead, we estimate the 2-misspecified lag-1 causal estimand as in (D.10).

For the outcome model, we consider the same parameterization as in Section 5.5.2, and conduct experiments under 3 different scenarios $\delta \in \{1, 2, 3\}$.

We only simulate one assignment path. Since the outcome model is prescribed, we calculate the observed outcomes. There is only one time series of such observed outcomes. We calculate the Horvitz-Thompson estimator based on the simulated assignment path and the simulated observed outcomes. We calculate the lag-$p$ causal estimand directly, and also the $m$-misspecified lag-$p$ causal estimand in conjunction with the simulated assignment path. Finally, we perform the two inference methods from Section 5.4, and report their associated estimated $p$-values. For the asymptotic inference method we plug in $\hat{\sigma}^2_{\mathsf{U}}$ the conservative upper bound of the variance. We choose $I = 100000$ to be the number of samples drawn in the exact inference method as shown in Algorithm 8.

#### Simulation results

Notice this is only one experiment under one simulated experimental setup from one simulated assignment path. So the estimators $\hat{\tau}_p$ we derive are different from $\tau_p$. But they are still following the true causal effects which they estimate. See Table 5.5.

From Table 5.5 we see that both our estimator and the estimated variance are well defined in all the cases when $p = m$, $p > m$, and $p < m$. In each case, as $\delta$ increases from 1 to 3, the associated $p$-values exhibit decreasing trends, suggesting a stronger rejection rate against the null hypothesis. Moreover, the $p$-values suggested

Table 5.5: Simulation results for correctly specified $m$ case, and two misspecified $m$ cases

| | | $\tau_p$ | $\tau_p^{[m]}$ | $\hat{\tau}_p$ | $\hat{\sigma}^2_{\mathsf{U}2}$ | $\hat{p}_{\mathsf{F}}$ | $\hat{p}_{\mathsf{N}}$ |
|---|---|---|---|---|---|---|---|
| $m=2, p=2$ | $\delta=1$ | 3 | – | 1.35 | 8.81 | 0.626 | 0.648 |
| | $\delta=2$ | 6 | – | 4.30 | 15.16 | 0.231 | 0.269 |
| | $\delta=3$ | 9 | – | 7.25 | 23.88 | 0.101 | 0.138 |
| $m=2, p=3$ | $\delta=1$ | 3 | – | 1.77 | 14.26 | 0.606 | 0.639 |
| | $\delta=2$ | 6 | – | 5.00 | 24.69 | 0.262 | 0.314 |
| | $\delta=3$ | 9 | – | 8.23 | 39.00 | 0.136 | 0.188 |
| $m=2, p=1$ | $\delta=1$ | – | 2 | -1.03 | 3.87 | 0.590 | 0.599 |
| | $\delta=2$ | – | 4 | 0.41 | 6.28 | 0.866 | 0.870 |
| | $\delta=3$ | – | 6 | 1.86 | 9.47 | 0.530 | 0.547 |

Note: The simulation setup for the three $\delta = 1$ cases is the same; so are the $\delta = 2$ cases and $\delta = 3$ cases. The estimated $p$-values $\hat{p}_{\mathsf{F}}$ derived from the exact inference are slightly smaller than the $p$-values $\hat{p}_{\mathsf{N}}$ derived from the asymptotic inference.

by the exact inference are always slightly smaller than the $p$-values suggested by the asymptotic inference. This coincides with our intuition that: (i) the exact inference method possesses a stronger null hypothesis (5.8) which implies the null hypothesis of (5.9); (ii) in the asymptotic inference we replaced the true randomization variance by its conservative upper bound, which further leads to a larger $p$-value.

### 5.5.5 Estimation of $m$

We run simulations based on the outcome model as in (5.13), to test the performance of the procedure described in Section 5.4.4. In this section we only focus on $\delta = 3$. Suppose we have narrowed down the range of the order of the carryover effect to be $m \leq 3$. In the first round, we use our procedure to test a null hypothesis $m \leq 2$. Then we would observe row 3 and 6 from Table 5.5, with $\hat{\tau}_2 = 7.25, \hat{\sigma}^2_2 = 23.88; \hat{\tau}_3 = 8.23, \hat{\sigma}^2_3 = 39.00$. So the estimated $p$-value for the null hypothesis $m \leq 2$ is estimated to be $\hat{p} = 0.902$, which is too large to reject the null hypothesis. In the second round, we consult the procedure to test a null hypothesis $m \leq 1$. Then we would observe row 3 and 9 from Table 5.5, with $\hat{\tau}_1 = 1.86, \hat{\sigma}^2_3 = 9.47; \hat{\tau}_2 = 7.25, \hat{\sigma}^2_2 = 23.88$. The estimated $p$-value for the null hypothesis $m \leq 1$ is estimated to be $\hat{p} = 0.350$. This is still rather large, yet a significant difference from 0.902.

We conduct a few more numerical simulations with different time periods. The setup is the same as in Section 5.5.4, except that $T$ takes values in $T \in \{210, 1020, 2010\}$[4]. When $T = 210$, in the first round the estimated $p$-value for the null hypothesis $m \leq 2$ is estimated to be $\hat{p} = 0.956$; in the second round the estimated $p$-value for the null hypothesis $m \leq 1$ is estimated to be $\hat{p} = 0.182$. When $T = 1020$, in the first round the estimated $p$-value for the null hypothesis $m \leq 2$ is estimated to be $\hat{p} = 0.869$; in the second round the estimated $p$-value for the null hypothesis $m \leq 1$ is estimated to be $\hat{p} = 0.163$. When $T = 2010$, in the first round the estimated $p$-value for the null hypothesis $m \leq 2$ is estimated to be $\hat{p} = 0.760$; in the second round the estimated $p$-value for the null hypothesis $m \leq 1$ is estimated to be $\hat{p} = 0.037$. In practice, we suggest increasing the horizon's length to a degree such that $T/p > 100$.

## 5.6 Practical Implications, Limitations, and Concluding Remarks

When a firm decides to use a switchback experiment for the evaluation of a new product or initiative, they have to make multiple decisions to ensure that the results are reliable, practical, and replicable. First, the firm must determine an appropriate outcome(s) that adequately captures the relative effectiveness of the change. In practice, this requires substantive domain knowledge combined with an understanding of the likely impact of the change; see Kohavi et al. (2020) for an in-depth discussion of metric definition strategies.

Second, as part of the design of the experiment, the firm often have control over the granularity of one single experimental period. As we have shown in Example 5.7, as long as each time period is smaller than the length of the carryover effect and the length of the carryover effect is divisible by the length of one time unit, the selection of granularity makes no difference to the optimal design and analysis of switchback experiments. On the other hand, setting each period's length longer than

[4]The values of $T$ were selected such that they were both divisible by both 2 and 3, the possible values of the carryover effect.

the carryover effect will lead to a loss in precisions. Consider an extreme case where the carryover effect is 1 minute, while each period is selected to be an hour. Had we set each period to be a minute, we would have collected order of magnitude more useful data. Hence, we suggest that each period's length be smaller than the carryover effect duration.

Third, the firm must form a knowledge and decide an appropriate value $p$ for the order of the carryover effect $m$. This, again, would require substantive domain knowledge. When domain knowledge is not accurate, or when the firm would like to reaffirm the domain knowledge, we have discussed in Section 5.4.4 a procedure to estimate the likely order of the carryover effect through hypothesis testing. we do encourage empirical researchers who apply our method to use domain knowledge to narrow down $m$ first, before using the procedure in Section 5.4.4 to identify $m$. This is because, in theory, when $m$ is relatively large compared to $T$, this procedure could accept (not reject) the null hypothesis simply due to insufficient statistical power. So this procedure could require too many testing periods. And in practice, each hypothesis testing to identify (5.12) needs to consume experimental resources at the scale of $T/m > 100$ to distinguish two candidate values, which could be luxurious when the resource is scarce.

Fourth, when the firm have control over the horizon of the experiment, the firm usually set $p = m$ and control the overall duration of the experiment $n = T/p = T/m$. We suggest choosing $n$ by referring to the rejection rate curve, as shown in Section 5.5.3; intuitively, this procedure resembles a typical power analysis. We begin with selecting our inference method, as described in Section 5.4. We then use our domain knowledge to estimate the expected signal-to-noise ratio; this could be done by looking at historical experiments or through dummy experiments. Finally, we choose the desired rejection rate and find out the length of the horizon required.

Finally, using the above four three steps, the firm decide the collection of randomization points and samples the assignment path from the appropriate randomization distribution. This final step has already been discussed at length, as we showed in Section 5.3 the optimal design is obtained from Theorems 5.4 and 5.5. In cases when the time horizon is pre-determined and when $T/p$ is not an integer, our optimization formulation as shown in Theorem 5.5 can always be used to find an optimal solution without discarding any periods. Just in the "imperfect cases" we do not have closed-form solutions. Our suggestion is that if the experimental designer wishes not to discard any periods, then solve the optimal solution (using any commercial software); if the experimental designer wishes not to solve an optimization problem, then discard a few periods and consult the explicit solution suggested in Theorem 5.5.

After designing the experiment, the firm can use the data collected from the test to draw causal conclusions about the new innovation's performance using the two inferential methods as discussed in Section 5.4.

# Chapter 6

# Conclusions

As mentioned in Chapter 1, it is the market uncertainty that has made the revenue management problems challenging. In this dissertation, we have showcased how minimax optimization, and, more specifically, competitive analysis, could guide operational decisions. Such a solution concept is powerful not only in the traditional revenue management problem, but also in statistical decision making problems. We conclude this dissertation by pointing out the limitations, possible variations, and open questions associated with all the models.

In Chapter 2, there are three limitations and possible variations. First, if we were introducing new products to the market or if there were very little data to generate accurate demand distributions, then we should treat demand distributions as unknown, instead of distributionally-known. One interesting question is to address the unknown demand distributions. The second possible variant is to consider strategic customers. If we observed significant inter-temporal cannibalization, it would be interesting to develop a strategic customer model. The literature (Gul et al. 1986, Chen et al. 2018) suggests that commitment power gains an advantage in the presence of strategic customers. Calendar pricing is naturally such a commitment policy. A third possible variant is to consider oligopoly pricing. If we observed cross-SKU cannibalization, then it would be worthwhile to consider how the incentives of different agents align with one another and even with the retailers that carry all competitors' products.

In Chapter 4, there are two unsolved questions. First, what does the best non-adaptive policy look like? Or are there structural properties concerning the optimal policy? Second, a potentially easier question to asnwer is in the network revenue management setting, where the static control policy is a well studied policy that suggests a combination of actions to execute. Using the actions suggested by the deterministic LP, it remains open to understand which permutation of the actions might lead to superior performance.

In Chapter 4, there are two possible variations. First, we could extend our model to consider multiple consumptions over different knapsacks. Suppose we are managing a manufacturing plant that requires different resources to make products, instead of a warehouse managing a single stock. One arriving item could potentially require more than one single resource to be produced. For each unit of resource consumed, there is an associated revenue / cost. Our goal is to maximize the total revenue / minimize the total cost throughout the horizon. Second, we could also extend our model to reusable products. Suppose there is a fixed total amount of cloud computing resources whose capacity is 1, and an unknown sequence of tasks with sizes at most 1. These resources immediately become available after the usage time. If one unit of resource is occupied for one period of time, a constant amount of revenue is generated. Our goal is to maximize the total revenue generated throughout the horizon.

In Chapter 5, there are three limitations and open questions. First, when $m$, the order of the carryover effect is as large as comparable to $T$ the horizon's length, our method, though still unbiased in theory, incurs a large variance that typically prohibits the firm from making meaningful inference. This is because our method is general and requires the minimum amount of modeling assumptions. If we have strong domain knowledge about the outcome model, we can incorporate them to improve the design. Second, our method only considers flipping independent coins before the experiment even begins. We do not consider adaptively changing the coin flip probabilities, as it requires further assumptions about the outcome model, *e.g.*, some time-homogeneity of the data generating process. Third, in this paper we have only considered the estimand as in (5.1), which is motivated when firms want to decide whether to permanently adopt a policy. If the primary focus is on some other general causal estimands, our results do not directly apply. It remains open to derive new results for other estimands, using a similar strategy that we have employed.

# Appendix A

# Appendix to Chapter 2

## A.1 Building the Random Forest Model

In this section we explain in detail how we built the prediction model from the data. We follow the workflow shown in Figure 2-1.

We begin with weekly sales data in the past 3 years. After cleaning the missing data, we select SKUs that generated 90% of the revenue in the past three years and eliminate the rest. We also eliminate SKUs that were newly introduced in the most recent year. Some SKUs are already grouped together by the company. They are similar brands sold at similar pack sizes. The company requires that all SKUs in the same group be sold at the same price. There are 52 distinct groups in total. We build group-specific prediction models with the same combination of features, i.e., all SKUs use the feature "tagged price", but it refers to a different tagged price for each SKU.

We derive a list of features from the data that will be used to predict demand at each time step. These features include the price that this group is tagged at, its internal competitor prices, its external competitor prices, and its history prices. The internal competitor prices are the prices of the brands owned by the same company. The external competitor prices are the prices of its true competitors, owned by its rival companies. The features of history prices are take from the past week to the past 3 weeks, as 3 different features.

The external features include industry seasonal trend (after applying moving average), total number of stores in the district, festivals and sports events. The first two features are provided by the company, and the rest are obtained by scripting from the Internet. We create dummy variables for festivals and sports events to characterize categorical data.

We tested a few algorithms and finally choose to use random forest (Liaw et al. (2002), Ferreira et al. (2016)) as the prediction model. Random forest provides us the flexibility to use piecewise constant functions to approximate any true demand function, possibly nonlinear functions. Random forest provides us better performance than simpler models such as linear regression. On the other hand, it preserves some interpretability of the features, compared to more advanced methods such as neural networks.

Then we aggregate all the features together and simultaneously perform feature selection and parameter tuning by using a 5-fold cross-validation. We use stepwise backward selection to select features. In the cross-validation, we evaluate each combination based on its performance on the validation set.

During this procedure, we engaged in rounds of discussions with the company to ensure that the features selected are interpretable. There are some sub-optimal combinations that the company believed would make more practical sense, and we followed their advice. These features were both approved by the CPG company's management as consistent with their expedience and also resulted in the lowest out-of-sample prediction errors — see Table 2.3 for the reported error rates. Each column depicts a combination of features, and the corresponding numbers are prediction errors under this feature combination. The first column serves as a benchmark. We omit some trivial duplicates of the same feature, but note that some rows represent many features, e.g., festivals and sports events.

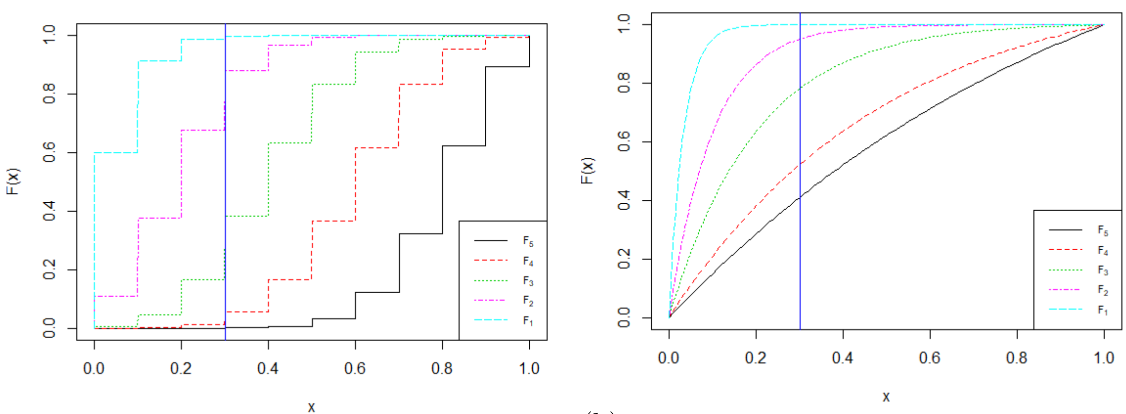
The average prediction error is reported as 19.41%.

## A.2 Further Justification of Assumptions 2.2 and 2.3

In this section we provide further explanation of Assumptions 2.2 and 2.3. Throughout this section we assume that the prices are sorted in decreasing order, i.e. $p_1 > p_2 > ... > p_m$. This is without any loss of generality.

### A.2.1 Two Examples

We provide two examples here to illustrate our assumptions.




(a) Example A.1: CDF for normalized binomials

(b) Example A.2: CDF for truncated exponentials

*Example* A.1. *Normalized binomial distributions.* Let us restrict ourselves to normalized binomial distributions that have the same number of coin flips, i.e. $\text{Bin}(N, \beta_j)/N$, where $N$ denotes the total number of coin flips, and $\beta_j$ denotes the probability of head-ups. We normalize it by $N$ so that this is a proper distribution with bounded support within $[0, 1]$.

Assumption 2.3 is naturally satisfied. From Lemma A.6 we know that $\forall j < j', \beta_j \leq \beta_{j'}$ ensures Assumption 2.2 to hold.

*Example* A.2. *Truncated exponential distributions.* Let us restrict ourselves to truncated exponential distributions with bounded support on $[0, 1]$, whose CDF can be written as

$$F(x) = \frac{1 - e^{-\lambda x}}{1 - e^{-\lambda}}, \forall x \in [0, 1].$$

Again Assumption 2.3 is naturally satisfied.

Assumption 2.2 states that $\forall j < j', c \in [0,1], \frac{\mathbb{E}[Q_j]}{\mathbb{E}[Q_{j'}]} \leq \frac{\mathbb{E}[\min\{c,Q_j\}]}{\mathbb{E}[\min\{c,Q_{j'}\}]}$. Notice that $\mathbb{E}[X] = \frac{1}{\lambda}, \mathbb{E}[\min\{c, X\}] = \frac{1-e^{-\lambda c}}{\lambda(1-e^{-\lambda})}$. So we have

$$\frac{\mathbb{E}[Q_j]}{\mathbb{E}[Q_{j'}]} \leq \frac{\mathbb{E}[\min\{c, Q_j\}]}{\mathbb{E}[\min\{c, Q_{j'}\}]} \iff \frac{\lambda_{j'}}{\lambda_j} \leq \frac{\lambda_{j'}}{\lambda_j} \frac{\frac{1-e^{-\lambda_j c}}{\lambda_j(1-e^{-\lambda_j})}}{\frac{1-e^{-\lambda_{j'} c}}{\lambda_{j'}(1-e^{-\lambda_{j'}})}} \iff \lambda_j \geq \lambda_{j'}$$

So Assumption 2.2 holds if and only if $\forall j < j', \lambda_j \geq \lambda_{j'}$.

### A.2.2 Necessity of Assumption 2.2

Notice that inequality (2.12) holds if and only if Assumption 2.2 holds. So if Assumption 2.2 does not hold then inequality (2.12) and Theorem 2.2 break down.

### A.2.3 Necessity of Assumption 2.3 Through Examples

We show an example that does not satisfy Assumption 2.3 and breaks Lemma 2.13.

*Example* A.3. Let there be $T = 2$ periods and $b = 1$ unit of initial inventory. Let $\epsilon \in (0, 1)$ be some small positive number. Let there be two prices: $p_1 = 1 + \epsilon, p_2 = 1$. Demand at the higher price $p_1$ is deterministically $1/2 - \epsilon$; and demand at the lower price $p_2$ is 1 with probability $1/2$, and $2\epsilon$ with probability $1/2$.

DLP-S suggests that we offer both prices $p_1$ and $p_2$ for one period, since that uses up the $b = 1$ inventory exactly in expectation. Indeed, if we ignore the $\epsilon$ terms, $\mathsf{Rev}(\mathsf{H};\mathsf{L}) \approx 3/4$; and $\mathsf{Rev}(0.5\mathsf{H}, 0.5\mathsf{L}; 0.5\mathsf{H}0.5\mathsf{L}) \approx 13/16 > 3/4$. So Inequality (2.18) and Lemma 2.13 break down.

## A.3 Inequalities Involving Truncations

*Lemma* A.1. *$\forall c, x, y \geq 0$,*

$$\min\{c, x\} + \min\{c, y\} \geq \min\{c, x + y\}$$

*Proof.* Proof. We prove by discussing all the possibilities. If $c \leq \min\{x, y\}$, then following from $c \geq 0$ we know that $\min\{c, x\} + \min\{c, y\} = c + c \geq c = \min\{c, x + y\}$.

If $\min\{x, y\} \leq c \leq \max\{x, y\}$, then following from $x, y \geq 0$ we know that $\min\{c, x\} + \min\{c, y\} = \min\{x, y\} + c \geq c = \min\{c, x + y\}$.

If $c \geq \max\{x, y\}$ then we know that $\min\{c, x\} + \min\{c, y\} = x + y \geq \min\{c, x + y\}$. □

*Lemma* A.2. *For any $p > 0, c \geq 0, x \geq w \geq 0, z \geq y \geq 0$, if $J$ is differentiable, $J(0) = 0, 0 \leq J'(u) \leq p$, and $J$ is a concave function, then the following holds:*

$$\begin{aligned}
0 \leq & -p\min\{c, w + y\} + p\min\{c, w + z\} \\
& + p\min\{c, x + y\} - p\min\{c, x + z\} \\
& - J((c - w - y)^+) + J((c - w - z)^+) \\
& + J((c - x - y)^+) - J((c - x - z)^+)
\end{aligned}$$

*Proof.* Proof. We prove by enumerating all the possibilities. If $c \leq w + y$, then $0 \leq 0$ we are done.

If $w+y \leq c \leq \min\{w+z, x+y\}$, then it suffices to show that $0 \leq p(c-w-y) - J(c-w-y)$, which is proved by $J(c-w-y) = \int_0^{c-w-y} J'(u)\mathrm{d}u \leq \int_0^{c-w-y} p\mathrm{d}u = p(c-w-y)$.

If $\min\{w+z, x+y\} \leq c \leq \max\{w+z, x+y\}$, without loss of generality we assume $w + z \leq x + y$. So it suffices to show that $0 \leq p(z - y) + J(c - w - z) - J(c - w - y)$, which is proved by $J(c - w - y) = \int_0^{c-w-y} J'(u)\mathrm{d}u \leq \int_0^{c-w-z} J'(u)\mathrm{d}u + \int_{c-w-z}^{c-w-y} p\mathrm{d}u = J(c - w - z) + p(z - y)$.

If $\max\{w + z, x + y\} \leq c \leq x + z$, it suffices to show that $J(c - w - z) + J(c - x - y) + p(x + z - c) \geq J(c - w - y)$. Since $J$ is concave, $J'$ is non-increasing.

$$\begin{aligned}
& J(c - w - z) + J(c - x - y) + p(x + z - c) \\
= & \int_0^{c-w-z} J'(u)\mathrm{d}u + \int_0^{c-x-y} J'(u)\mathrm{d}u + \int_{2c-w-x-y-z}^{c-w-y} p\mathrm{d}u \\
\geq & \int_0^{c-w-z} J'(u)\mathrm{d}u + \int_0^{c-x-y} J'(u + (c - w - z))\mathrm{d}u + \int_{2c-w-x-y-z}^{c-w-y} p\mathrm{d}u
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^{c-w-z} J'(u)\mathrm{d}u + \int_{c-w-z}^{2c-w-x-y-z} J'(u)\mathrm{d}u + \int_{2c-w-x-y-z}^{c-w-y} p\mathrm{d}u \\
&\geq \int_0^{c-w-z} J'(u)\mathrm{d}u + \int_{c-w-z}^{2c-w-x-y-z} J'(u)\mathrm{d}u + \int_{2c-w-x-y-z}^{c-w-y} J'(u)\mathrm{d}u \\
&= J(c-w-y)
\end{aligned}
$$

where the first inequality is due to concavity of $J$; second inequality due to $J'(u) \leq p$.

Finally if $c \geq x+z$, it suffices to show that $0 \leq -J(c-w-y) + J(c-w-z) + J(c-x-y) - J(c-x-z)$, which is due to concavity of $J$. □

## A.4 Lemmas for the proof of Theorem 2.2

These Lemmas are of independent interests. We state them here, and prove them one by one.

*Lemma* A.3. *Let $c \in \mathrm{R}^+$ be any positive real number, $T \in \mathrm{N}$ any positive integer, and $p \in [0,1]$ be any positive fractional number. Let $\{X_t\}, t = 1,2,...,T$ be i.i.d. random variables with bounded support over $[0,1]$, such that $\mathbb{E}[X_t] = p, \forall t = 1,2,...,T$. Let $\{Y_t\}, t = 1,2,...,T$ be i.i.d. Bernoulli random variables, such that with probability $p$, $Y_t = 1, \forall t = 1,2,...,T$. Then we have*

$$
\mathbb{E}\left[\min\{c, \sum_{t=1}^{T} X_t\}\right] \geq \mathbb{E}\left[\min\{c, \sum_{t=1}^{T} Y_t\}\right].
$$

*Lemma* A.4. *Suppose $a_i, b_i > 0, \forall i \in [n]$, and $\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq ... \geq \frac{a_n}{b_n}$; suppose $\beta_1 \geq \beta_2 \geq ... \geq \beta_n \geq 0$. Then we have*

$$
\frac{\sum_{i\in[n]} \beta_i a_i}{\sum_{i\in[n]} \beta_i b_i} \geq \frac{\sum_{i\in[n]} a_i}{\sum_{i\in[n]} b_i}
$$

*Lemma* A.5. *Let $c \in \mathrm{R}^+$ be any positive real number, $T \in \mathrm{N}$ any positive integer, the following function $f : (0,1] \to \mathrm{R}^+$ is non-increasing in $x$.*

$$
f(x) = \frac{\mathbb{E}\left[\min\{c, \mathrm{Bin}(T,x)\}\right]}{Tx}.
$$

*Lemma* A.6. *Let $T \in \mathrm{N}$ be any positive integer. For any positive real numbers $x, y$ such that $T \geq x \geq y > 0$, we have*

$$\frac{\mathbb{E}[\min\{\mathrm{Bin}(T, x/T), x\}]}{x} \geq \frac{\mathbb{E}[\min\{\mathrm{Bin}(T, y/T), y\}]}{y}.$$

### A.4.1 Proof of Lemma A.3

*Proof.* Proof of Lemma A.3. Let $F_X(\cdot)$ and $F_Y(\cdot)$ denote the CDF of $X_1$ and $Y_1$, respectively. First we show $\forall c \geq 0$,

$$\mathbb{E}\left[\min\{c, X_1\}\right] \geq \mathbb{E}\left[\min\{c, Y_1\}\right]. \tag{A.1}$$

This is obvious when $c \geq 1$, and now we focus on $c < 1$ case.

$$\begin{aligned}
\mathbb{E}\left[\min\{c, X_1\}\right] =& \int_{[0,1]} \min\{c, x\} \mathrm{d}F_X(x) \\
=& \int_{[0,1]} x \mathrm{d}F_X(x) - \int_{[c,1]} (x - c) \mathrm{d}F_X(x) \\
=& \mathbb{E}[X_1] - \left\{ \int_{[c,1]} x \mathrm{d}F_X(x) - c(1 - F_X(c)) \right\} \\
=& \mathbb{E}[X_1] - \left\{ 1 - cF_X(c) - \int_{[c,1]} F_X(x - 0^+) \mathrm{d}x - c(1 - F_X(c)) \right\} \\
=& \mathbb{E}[X_1] - (1 - c) + \int_{[c,1]} F_X(x) \mathrm{d}x
\end{aligned}$$

where the fourth equality is due to integration by part, as a corollary of Fubini's theorem. Due to similar analysis, $\mathbb{E}\left[\min\{c, Y_1\}\right] = \mathbb{E}[Y_1] - (1 - c) + \int_{[c,1]} F_Y(x) \mathrm{d}x$. So it suffices to show $\int_{[c,1]} F_X(x) \mathrm{d}x \geq \int_{[c,1]} F_Y(x) \mathrm{d}x$.

Note that $\int_{[0,1]} F_X(x) \mathrm{d}x = \int_{[0,1]} F_Y(x) \mathrm{d}x$, that $F_Y(x) = p, \forall x \in [0, 1)$ is a constant, and that $F_X(x)$ is non-decreasing over $x \in [0, 1)$. Denote $x_0$ to be the smallest number from $x_0 = \arg\min_{x \in [0,1)} \{x \,|F_X(x) \geq p, \lim_{u \to x^-} F_X(u) \leq p\}$. Since $F_X(\cdot)$ is right-continuous, $\lim_{u \to x^+} F_X(u) = F_X(x)$.

We distinguish the following two cases. When $c \geq x_0$, $\forall x \in [c, 1], F_X(x) \geq F_Y(x)$. So we know $\int_{[c,1]} F_X(x) \mathrm{d}x \geq \int_{[c,1]} F_Y(x) \mathrm{d}x$. When $c \leq x_0$, $\forall x \in [0, c), F_X(x) \leq$

$F_Y(x)$. So we know $\int_{[c,1]} F_X(x)\mathrm{d}x = \int_{[0,1]} F_X(x)\mathrm{d}x - \int_{[0,c)} F_X(x)\mathrm{d}x \geq \int_{[0,1]} F_Y(x)\mathrm{d}x - \int_{[0,c)} F_Y(x)\mathrm{d}x = \int_{[c,1]} F_Y(x)\mathrm{d}x$. In both cases, we have shown that $\forall c \geq 0, \mathbb{E}[\min\{c, X_1\}] \geq \mathbb{E}[\min\{c, Y_1\}]$.

Then we prove the desired conclusion by pairwise switching $X_t$ into $Y_t$. Fix $\tau$. For any realization of random variables except $X_\tau, Y_\tau$, i.e., for any realization of $X_t = x_t, Y_t = y_t, \forall t \neq \tau$, we have the following:

$$
\begin{aligned}
&\mathbb{E}\left[\min\{c, \sum_{t=1}^{\tau} X_t + \sum_{t=\tau+1}^{T} Y_t\}\right] \\
=&\mathbb{E}\left[\min\{c, \sum_{t=1}^{\tau-1} x_t + \sum_{t=\tau+1}^{T} y_t + X_\tau\}\right] \\
=&\mathbb{E}\left[\min\left\{\left(c - \sum_{t=1}^{\tau-1} x_t - \sum_{t=\tau+1}^{T} y_t\right)^+, X_\tau\right\}\right] + \min\left\{c, \sum_{t=1}^{\tau-1} x_t + \sum_{t=\tau+1}^{T} y_t\right\} \\
\geq&\mathbb{E}\left[\min\left\{\left(c - \sum_{t=1}^{\tau-1} x_t - \sum_{t=\tau+1}^{T} y_t\right)^+, Y_\tau\right\}\right] + \min\left\{c, \sum_{t=1}^{\tau-1} x_t + \sum_{t=\tau+1}^{T} y_t\right\} \\
=&\mathbb{E}\left[\min\{c, \sum_{t=1}^{\tau-1} X_t + \sum_{t=\tau}^{T} Y_t\}\right]
\end{aligned}
$$

where the inequality is due to (A.1).

Repeatedly applying the above inequality, we have

$$
\mathbb{E}\left[\min\{c, \sum_{t=1}^{T} X_t\}\right] \geq \mathbb{E}\left[\min\{c, \sum_{t=1}^{T} Y_t\}\right], \forall c \geq 0.
$$

□

### A.4.2 Proof of Lemma A.4

*Proof.* Proof of Lemma A.4. Since $\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq ... \geq \frac{a_n}{b_n}$, we have $\forall i \in [n-1]$,

$$
\frac{\sum_{j\in[i]} a_j}{\sum_{j\in[i]} b_j} \geq \frac{\sum_{j\in[n]} a_j}{\sum_{j\in[n]} b_j}.
$$

Then we plug it into the following fraction, and use the fact that $\beta_i \geq \beta_{i+1}, \forall i \in [n-1]$:

$$\begin{aligned}
\frac{\sum_{i\in[n]} \beta_i a_i}{\sum_{i\in[n]} \beta_i b_i} =& \frac{\sum_{i\in[n-1]}(\beta_i - \beta_{i+1}) \cdot \sum_{j\in[i]} a_j + \beta_n \cdot \sum_{j\in[n]} a_j}{\sum_{i\in[n-1]}(\beta_i - \beta_{i+1}) \cdot \sum_{j\in[i]} b_j + \beta_n \cdot \sum_{j\in[n]} b_j} \\
\geq& \frac{\sum_{i\in[n-1]}(\beta_i - \beta_{i+1}) \cdot \sum_{j\in[n]} a_j + \beta_n \cdot \sum_{j\in[n]} a_j}{\sum_{i\in[n-1]}(\beta_i - \beta_{i+1}) \cdot \sum_{j\in[n]} b_j + \beta_n \cdot \sum_{j\in[n]} b_j} \\
=& \frac{\sum_{i\in[n]} a_i}{\sum_{i\in[n]} b_i}
\end{aligned}$$

□

### A.4.3 Proof of Lemma A.5

Similar to Lemma A.4, we can show the following

*Lemma* A.7. *Suppose $a_i, b_i > 0, \forall i \in [n]$, and $\frac{a_0}{b_0} \geq \frac{a_1}{b_1} \geq ... \geq \frac{a_n}{b_n}$; suppose $0 \leq \beta_0 \leq \beta_1 \leq ... \leq \beta_n$. Then we have*

$$\frac{\sum_{i=0}^n a_i}{\sum_{i=0}^n b_i} \geq \frac{\sum_{i=0}^n \beta_i a_i}{\sum_{i=0}^n \beta_i b_i}.$$

The proof is the same as the proof of Lemma A.4.

*Proof.* Proof of Lemma A.5. Observe that Binomial distribution is a discrete distribution. It only suffices to prove Lemma A.5 in the case when $c \in \mathrm{N}$ is an integer.

Let $C_T^l$ be $T$ choose $l$.

Take any $x, y \in [0, 1]$ such that $x < y$. Notice that Binomial distributions $\mathrm{Bin}(T, x)$ and $\mathrm{Bin}(T, y)$ only have finite supports over $\{0, 1, ..., T\}$. It is trivial when $c \geq T$ because the truncation does not take effect and both fractions equal to 1. Take any $z \in \{0, 1, ..., T-1\}$. Since both enumerators are linear in $c \in (z, z+1)$, it only suffices to check for $c \in \{0, 1, ..., T-1\}$, i.e. when $c$ is an integer.

First notice that $\frac{x(1-y)}{y(1-x)} \leq 1$. $\forall l \in \{0, 1, ..., T-1\}$,

$$\left(\frac{x(1-y)}{y(1-x)}\right)^l \geq \left(\frac{x(1-y)}{y(1-x)}\right)^{l+1}.$$

Both multiply by $\left(\frac{1-x}{1-y}\right)^T$ we have $\forall l \in \{0, 1, ..., T-1\}$,

$$\frac{x^l(1-x)^{T-l}}{y^l(1-y)^{T-l}} \geq \frac{x^{l+1}(1-x)^{T-l-1}}{y^{l+1}(1-y)^{T-l-1}}.$$

Further multiply by some constants to both enumerators and denominators the inequality still holds.

$$\frac{C_T^l x^l(1-x)^{T-l} \cdot l}{C_T^l y^l(1-y)^{T-l} \cdot l} \geq \frac{C_T^{l+1} x^{l+1}(1-x)^{T-l-1} \cdot (l+1)}{C_T^{l+1} y^{l+1}(1-y)^{T-l-1} \cdot (l+1)}.$$

$\forall l \in \{0, 1, ..., T\}$, let $a_l = C_T^l x^l (1-x)^{T-l} \cdot l; b_l = C_T^l y^l (1-y)^{T-l} \cdot l$.

$\forall i \in \{0, 1, ..., c\}$, let $\beta_i = 0$; $\forall i \in \{c+1, c+2, ..., T\}$, $\beta_i = (i-c)/i$. It is easy to verify that $0 \leq \beta_0 \leq \beta_1 \leq ... \leq T$.

Invoking Lemma A.7, we have

$$\frac{\sum_{l=0}^{T} C_T^l x^l(1-x)^{T-l} \cdot l}{\sum_{l=0}^{T} C_T^l y^l(1-y)^{T-l} \cdot l} \geq \frac{\sum_{l=c}^{T} C_T^l x^l(1-x)^{T-l} \cdot (l-c)}{\sum_{l=c}^{T} C_T^l y^l(1-y)^{T-l} \cdot (l-c)}.$$

Re-arranging terms,

$$\frac{\sum_{l=0}^{c-1} C_T^l x^l(1-x)^{T-l} \cdot l + \sum_{l=c}^{T} C_T^l x^l(1-x)^{T-l} \cdot c}{\sum_{l=0}^{T} C_T^l x^l(1-x)^{T-l} \cdot l} \geq \frac{\sum_{l=0}^{c-1} C_T^l y^l(1-y)^{T-l} \cdot l + \sum_{l=c}^{T} C_T^l y^l(1-y)^{T-l} \cdot c}{\sum_{l=0}^{T} C_T^l y^l(1-y)^{T-l} \cdot l}.$$

Equivalently,

$$\frac{\mathbb{E}\left[\min\{c, \mathrm{Bin}(T, x)\}\right]}{Tx} \geq \frac{\mathbb{E}\left[\min\{c, \mathrm{Bin}(T, y)\}\right]}{Ty}$$

$\square$

### A.4.4 Proof of Lemma A.6

*Proof.* Proof of Lemma A.6. It suffices to prove the following

$$\mathbb{E}[\min\{\frac{y}{x} \cdot \mathrm{Bin}(T, \frac{x}{T}), y\}] \geq \mathbb{E}[\min\{\mathrm{Bin}(T, \frac{y}{T}), y\}]$$

For any $t \in [T]$, denote $X_t = \frac{y}{x} \cdot \mathrm{Ber}(T, \frac{x}{T})$ as a Bernoulli random variable such that $\mathbb{E}[X_t] = y/T$, and $X_t \in [0,1]$ has bounded support between $[0,1]$; For any $t \in [T]$, denote $Y_t = \mathrm{Ber}(T, \frac{y}{T})$ as a Bernoulli random variable such that with probability $y/T$, $Y_t = 1$. Pick $c = y$ to be a positive real number.

From Lemma A.3 we have

$$\mathbb{E}[\min\{\sum_{t \in [T]} X_t, y\}] \geq \mathbb{E}[\min\{\sum_{t \in [T]} Y_t, y\}],$$

which finishes the proof. □

## A.5 Proof of Theorem 2.2 and Proposition 2.4

We write the most general proof precisely by combining Theorem 2.2 and Proposition 2.4.

*Theorem* A.8. *Under one of the following three conditions:*

*(i) the static substitution model with integral demand;*

*(ii) the static substitution model with fractional demand and Assumption 2.2;*

*(iii) the dynamic substitution model with integral demand and Assumptions 2.1 (substitutability), and when one item has only one single price (pure assortment problem without pricing);*

*for the assortment (and pricing) problem under stationary demand, if there are $T$ time periods and $\underline{b} = \min_{i \in [n]} b_i$, then Algorithm 1 earns expected revenue of at least*

$$\frac{\mathbb{E}[\min\{\mathrm{Bin}(T, \underline{b}/T), \underline{b}\}]}{\underline{b}} \cdot \mathsf{OPT}_{\mathrm{LP}}, \tag{A.2}$$

*where* $\mathrm{Bin}(T, \underline{b}/T)$ *denotes a Binomial random variable consisting of* $T$ *trials of probability* $\underline{b}/T$.

*If we let* $\Delta^{APX}$ *denote the term* $\frac{\mathbb{E}[\min\{\mathrm{Bin}(T,\underline{b}/T),\underline{b}\}]}{\underline{b}}$ *from expression (2.9), then*

$$\Delta^{APX} \geq 1 - \frac{\underline{b}^{\underline{b}}}{\underline{b}!}e^{-\underline{b}}, \tag{A.3}$$

*which states that* $\Delta^{APX} = 1 - O(1/\sqrt{\underline{b}})$, *and increases from* $1 - 1/e$ *to 1 as* $\underline{b} \to \infty$ *(regardless of* $T$*).*

It is easy to see that Theorem 2.2 corresponds to (i) and (ii) of Theorem A.8, and Proposition 2.4 corresponds to (iii) of Theorem A.8 Now we prove Theorem A.8.

*Proof.* Proof of Theorem A.8. This proof consists of two steps. In the first step, we lower bound the performance of Algorithm 1, which is a randomized policy, by the performance of a virtual calendar. We define the choice model of this virtual calendar to have *only* static substitution, yet it is a lower bound to the performance of our Algorithm 1, under both static and dynamic substitution. In the second step, we lower bound the performance of this virtual calendar by $\Delta^{APX} \cdot \mathsf{OPT}_{\mathrm{LP}}$. Under all three conditions as stated in Theorem 2.2, the virtual calendar is the same. We state this virtual calendar in its most general form as in Step 0, and illustrate how the three conditions simplifies to this most general form. Under three conditions, the first step that lower bounds the performance of Algorithm 1 to the performance of this virtual calendar may be different, as we shall see in Step 1. Under all three conditions, the second step is the same, as we prove in Step 2.

Now we introduce the following random variables, which depict a run of our assortment policy. Let $S_t$ be the assortment that we select to offer in period $t$. Let $B_t(i)$ be the remaining inventory of item $i$ at the *end* of time $t$. We have $B_0(i) = b_i$. Under all three conditions, let $R_t(i, j)$ be the amount of sales that a customer chooses product $(i, j)$ during period $t$. We will always specify the distribution of $R_t(i, j)$, by using a conditional probability. For example, we will use $\mathbb{E}[R_t(i, j)\,|S_t = S, \boldsymbol{B}_t = \boldsymbol{B}]$, for the expected sales that a customer chooses product $(i, j)$ during period $t$, when we plan to offer assortment $S$, and when the remaining inventory level for each resource is $\boldsymbol{B} = (B_1, B_2, ..., B_n)$.

Under condition (i), under static substitution and when demand is integral, $R_t(i,j) \in \{0,1\}$ is binary. Conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{N}_0^n$, $R_t(i,j) = \mathbb{1}_{\{B_i>0\}}Q$, where $Q$ is a binary random variable, which takes 1 with probability $q(i,j,S)$. Under condition (ii) under static substitution and when demand is fractional, $R_t(i,j) \in [0,1]$ is continuous. Conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{R}_+^n$, $R_t(i,j) = \min\{B_i, Q\}$, where $Q$ is a random variable whose CDF is $F_{(i,j,S)}(\cdot)$. Under condition (iii) under dynamic substitution and when demand is integral, $R_t(i,j) \in \{0,1\}$ is binary. Conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{N}_0^n$, $R_t(i,j)$ takes 1 with probability $q(i,j,\breve{S})$. Here we define $\breve{S} = \{(i,j) \in S \,|\, B_i > 0\}$ to be a function of $S$.

Under condition (iii), Assumption 2.1 suggests that $q(i,j,\breve{S}) \geq q(i,j,S), \forall (i,j) \in \breve{S}$, because $\breve{S} \subseteq S$. The demand that originally would have chosen the stocked out items would go to their substitutes (as well as leaving, in which case the inequality takes equality). On the other hand, $q(i,j,\breve{S}) = 0, \forall (i,j) \notin \breve{S}$. The demand for any stocked out item is zero.

We can use indicator variables to write the above inequalities in a compact form

$$q(i,j,\breve{S}) \geq \mathbb{1}_{\{B_{t-1}(i)>0\}} q(i,j,S) \tag{A.4}$$

For any period $t$, given the remaining inventory from the last period to be $\boldsymbol{B}_{t-1}$, conditional on any $S \in \mathcal{S}$, the remaining inventory updates in the following fashion,

$$B_t(i) = B_{t-1}(i) - R_t(i,j), \forall i.$$

Note that no item can be offered multiple times at different prices in one assortment. Also note that we have defined $R_t(i,j)$ as the amount of sales, so $R_t(i,j)$ can never go beyond $B_{t-1}(i)$.

**Step 0 Statement of the virtual calendar.**

Consider Algorithm 1 that offers each assortment randomly. Define the set of items $I_{\boldsymbol{x}^*(S)} = \{i \in [n] \,|\, \exists j \in [m], \exists S \in \mathcal{S}, s.t. (i,j) \in S, x^*(S) > 0\}$. These are the items that are relevant to the probablistic offering of assortments from Algorithm 1. In other words, $I_{\boldsymbol{x}^*(S)}$ is the set of items such that there is a positive probability that Algorithm 1 suggests an assortment that contains item $i$.

Associated with each item in $I_{\boldsymbol{x}^*(S)}$, there is a unique price $p_{\mathsf{C},i}$. For any $i \in I_{\boldsymbol{x}^*(S)}$, define

$$p_{\mathsf{C},i} = \frac{\sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} p_j q(i,j,S)}{\sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} q(i,j,S)}.$$

We have $\sum_{S \in \mathcal{S}} x^*(S) = 1$, and $x^*(S) \geq 0, \forall S \in \mathcal{S}$ due to constraints (2.3) and (2.4). Notice that, here we only use $q(i,j,S)$.

Under conditions (i) and (ii), $p_{\mathsf{C},i}$ cannot be simplified. Under condition (iii), since each product has only one single price, we can define $j^i, \forall i \in [n]$ to be the price index that item $i$ can be offered. The price offered is simply $p_{\mathsf{C},i} = p_{j^i}$.

Now in each period suppose we had an option to offer a deterministic assortment $S_{\mathsf{C}}$ that consists of the products $S_{\mathsf{C}} = \left\{(i, p_{\mathsf{C},i}) \middle| \forall i \in I_{\boldsymbol{x}^*(S)} \right\}$.

Associated with each product $(i, p_{\mathsf{C},i}), \forall i \in I_{\boldsymbol{x}^*(S)}$, we prescribe a choice model. Under conditions (i) and (ii), we use the following notation. Let $Q(i, p_{\mathsf{C},i}, S_{\mathsf{C}}), \forall i \in I_{\boldsymbol{x}^*(S)}$ be a *random variable* for the quantity that customers attempt to purchase product $(i, p_{\mathsf{C},i})$, should assortment $S_{\mathsf{C}}$ be *offered*. Here we directly define the choice model to be under static substitution, for this virtual calendar. Nonetheless, as we will show in Step 1, the performance of this virtual calendar is a lower bound to the performance of our Algorithm 1, under both static and dynamic substitutions. We define the CDF function of $Q(i, p_{\mathsf{C},i}, S_{\mathsf{C}})$ to be $F_{(i, p_{\mathsf{C},i}, S_{\mathsf{C}})}(\cdot) = \sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} F_{(i,j,S)}(\cdot)$. Notice that $j : (i,j) \in S$ selects only one price, because no item can be offered multiple times at different prices in one assortment.

Under condition (iii), since we only consider dynamic substitution with integral demand, we can simplify the notations. Let $q(i, j^i, S_{\mathsf{C}}), \forall i \in I_{\boldsymbol{x}^*(S)}$ be the probability that product $(i, j^i)$ is demanded, should assortment $S_{\mathsf{C}}$ be offered. And assortment $S_{\mathsf{C}}$ is under static substitution – $q(i, j^i, S_{\mathsf{C}})$ is unchanged even if some items from the assortment stocks out. Denote $q(i, j^i, S_{\mathsf{C}}) = \sum_{S \in \mathcal{S}} x^*(S) q(i, j^i, S)$.

We wish to show that the expected revenue earned from this deterministic assortment is a lower bound to the probablistic offering of assortments from Algorithm 1. Denote $J_{\mathsf{C}}(\boldsymbol{b},t)$ as the *expected* revenue earned from always offering the deterministic assortment $S_{\mathsf{C}}$, if at the beginning of period $t$ we are endowed with $\boldsymbol{b}$ units of inventory—this corresponds to the expression in line (2.11). The expectation has only one source of randomness, which comes from the random demand. Denote $\mathsf{Rev}(\boldsymbol{b},t)$ as the *expected* revenue earned by the policy from Algorithm 1, if at the beginning of period $t$ we are endowed with $\boldsymbol{b}$ units of inventory—this corresponds to the expression in line (2.12). The expectation has two sources of randomness, which come from both the random demand, and the randomization from Algorithm 1.

**Step 1 Lower bounding the performance of Algorithm 1.**

In this step, we distinguish the following two cases: static substitution and dynamic substitution. Under static substitution, we talk about conditions (i) and (ii); under dynamic substitution, we talk about condition (iii).

**Case 1** Static substitution

This proof unifies conditions (i) and (ii). Note that condition (i) is integral, Bernoulli demand, which naturally satisfies Assumption 2.2.

Assumption 2.2 suggests that $\forall i \in [n], \forall S, S' \in \mathcal{S}$, for all $j, j' \in [m]$ such that $p_j > p_{j'}$, we have $\forall c \in [0,1]$,

$$\frac{x^*(S)\mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c,Q\}]}{x^*(S)\mathbb{E}_{Q\sim F_{(i,j,S)}}[Q]} \geq \frac{x^*(S')\mathbb{E}_{Q\sim F_{(i,j',S')}}[\min\{c,Q\}]}{x^*(S')\mathbb{E}_{Q\sim F_{(i,j',S')}}[Q]}.$$

From Lemma A.4, if we treat $p_j \geq p_{j'}$ as $\beta$'s, then we have $\forall i \in [n], \forall c \in [0,1]$,

$$\frac{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c,Q\}]}{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{(i,j,S)}}[Q]} \geq \frac{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c,Q\}]}{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} \mathbb{E}_{Q\sim F_{(i,j,S)}}[Q]}$$

which simplifies to

$$\frac{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c,Q\}]}{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c,Q\}]} \geq p_{\mathsf{C},i}$$

$$= \frac{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{(i,j,S)}}[Q]}{\sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} \mathbb{E}_{Q\sim F_{(i,j,S)}}[Q]} \tag{A.5}$$

Now we prove by backward induction on $t$. In the last period $T, \forall \boldsymbol{c} \geq \mathbf{0}$,

$$\begin{aligned}\mathsf{Rev}(\boldsymbol{c}, T) &= \sum_{i\in[n]} \sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c_i, Q\}] \\ &\geq \sum_{i\in[n]} p_{\mathsf{C},i} \sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c_i, Q\}] = J_{\mathsf{C}}(\boldsymbol{c}, T)\end{aligned} \tag{A.6}$$

The first equality is because each item can only be offered at one price in each assortment. So when truncation happens, there is no ambiguity which price of demand is lost. And the inequality is due to (A.5).

To continue the induction, if we can show $\mathsf{Rev}(\boldsymbol{c}, t+1) \geq J_{\mathsf{C}}(\boldsymbol{c}, t+1), \forall \boldsymbol{c} \geq \mathbf{0}$, then we can show:

$$\begin{aligned}&\mathsf{Rev}(\boldsymbol{c}, t) \\ =&\sum_{i\in[n]} \sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c_i, Q\}] + \sum_{S\in\mathcal{S}} x^*(S) \mathbb{E}_{\boldsymbol{Q}}[\mathsf{Rev}(\max\{\mathbf{0}, \boldsymbol{c} - \boldsymbol{Q}\}, t+1)] \\ \geq&\sum_{i\in[n]} p_{\mathsf{C},i} \sum_{S\in\mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} \mathbb{E}_{Q\sim F_{(i,j,S)}}[\min\{c_i, Q\}] + \sum_{S\in\mathcal{S}} x^*(S) \mathbb{E}_{\boldsymbol{Q}}[J_{\mathsf{C}}(\max\{\mathbf{0}, \boldsymbol{c} - \boldsymbol{Q}\}, t+1)] \\ =&J_{\mathsf{C}}(\boldsymbol{c}, t)\end{aligned}$$

where we use $\boldsymbol{Q} = (Q_1, ..., Q_n)$ as a vector form to stand for the joint distribution. Specifically, $Q_i \sim F_{(i,j^i,S)}, \forall i \in [n]$ specifies the distribution of demand associated with $i \in [n]$, and $j^i : (i,j) \in S$ specifies the price that is uniquely determined.

Now we explain this block of inequalities. The inequality is due to (A.5) and due to induction hypothesis. By induction on $t$ we show $\mathsf{Rev}(\boldsymbol{b}, t) \geq J_{\mathsf{C}}(\boldsymbol{b}, t)$. That is, the expected revenue earned from deterministic assortment $S_{\mathsf{C}}$ is a lower bound to the probabilistic offering of assortments from Algorithm 1.

**Case 2** Dynamic substitution

This proof is for condition (iii), for integral demand. We prove by backward induction on $t$. In the last period $T, \forall \boldsymbol{c} \geq \mathbf{0}$,

$$\begin{aligned}
\mathsf{Rev}(\boldsymbol{c}, T) &= \sum_{i \in [n]} \sum_{S \in \mathcal{S}} x^*(S) p_{j^i} q(i, j^i, \breve{S}) \\
&\geq \sum_{i \in [n]} p_{j^i} \mathbb{1}_{\{c_i > 0\}} \sum_{S \in \mathcal{S}} x^*(S) q(i, j^i, S) \\
&= \sum_{i \in [n]} p_{j^i} \mathbb{1}_{\{c_i > 0\}} q(i, j^i, S_{\mathsf{C}}) = J_{\mathsf{C}}(\boldsymbol{c}, T)
\end{aligned}$$

The inequality is due to (A.4).

To continue the induction, if we can show $\mathsf{Rev}(c, t+1) \geq J_{\mathsf{C}}(c, t+1), \forall c \geq 0$, then we can show:

$$\begin{aligned}
&\mathsf{Rev}(\boldsymbol{c}, t) \\
=& \sum_{i \in [n]} \sum_{S \in \mathcal{S}} x^*(S) p_{j^i} q(i, j^i, \breve{S}) + \sum_{S \in \mathcal{S}} x^*(S) \left\{ \sum_{i \in [n]} q(i, j^i, \breve{S}) \cdot \mathsf{Rev}(\max\{\mathbf{0}, \boldsymbol{c} - \boldsymbol{e}_i\}, t+1) \right. \\
& \qquad \left. + (1 - \sum_{i \in [n]} q(i, j^i, \breve{S})) \cdot \mathsf{Rev}(\boldsymbol{c}, t+1) \right\} \\
\geq& \sum_{i \in [n]} \sum_{S \in \mathcal{S}} x^*(S) p_{j^i} q(i, j^i, \breve{S}) + \sum_{S \in \mathcal{S}} x^*(S) \left\{ \sum_{i \in [n]} q(i, j^i, \breve{S}) \cdot J_{\mathsf{C}}(\max\{\mathbf{0}, \boldsymbol{c} - \boldsymbol{e}_i\}, t+1) \right. \\
& \qquad \left. + (1 - \sum_{i \in [n]} q(i, j^i, \breve{S})) \cdot J_{\mathsf{C}}(\boldsymbol{c}, t+1) \right\} \\
=& \sum_{S \in \mathcal{S}} x^*(S) \left\{ \sum_{i \in [n]} q(i, j^i, \breve{S}) \left\{ p_{j^i} + J_{\mathsf{C}}(\max\{\mathbf{0}, \boldsymbol{c} - \boldsymbol{e}_i\}, t+1) - J_{\mathsf{C}}(\boldsymbol{c}, t+1) \right\} \right\} + J_{\mathsf{C}}(\boldsymbol{c}, t+1) \\
\geq& \sum_{S \in \mathcal{S}} x^*(S) \left\{ \sum_{i \in [n]} q(i, j^i, S) \left\{ p_{j^i} + J_{\mathsf{C}}(\max\{\mathbf{0}, \boldsymbol{c} - \boldsymbol{e}_i\}, t+1) - J_{\mathsf{C}}(\boldsymbol{c}, t+1) \right\} \right\} + J_{\mathsf{C}}(\boldsymbol{c}, t+1) \\
=& J_{\mathsf{C}}(\boldsymbol{c}, t)
\end{aligned}$$

where the first inequality is due to induction hypothesis; the second equality is taking out $J_{\mathsf{C}}(\boldsymbol{c}, t+1)$ and re-arranging terms; the second inequality is because the marginal revenue of one extra unit of resource $i$ is bounded by $p_{j^i}$, and because $q(i, j^i, \breve{S}) \geq$

$q(i, j^i, S)$. By induction on $t$ we show $\mathsf{Rev}(\boldsymbol{b}, t) \geq J_{\mathsf{C}}(\boldsymbol{b}, t)$. That is, the expected revenue earned from deterministic assortment $S_{\mathsf{C}}$ is a lower bound to the probablistic offering of assortments from Algorithm 1.

**Step 2 Lower bounding the performance of the virtual calendar.**

Now we further lower bound the expected revenue earned from deterministic assortment $S_{\mathsf{C}}$. Since we have defined the choice model of the virtual calendar to be under static substitution, in all the remaining proof, we will only use notations like $Q(i, j, S)$, for the random quantity that customers attempt to purchase product $(i, j)$, should assortment $S$ be offered, no matter if any of the items from the assortment is stocked out. We do this because now we are under static substitution. The total quantity of demands attempting to consume inventory $i$ is

$$Q_i := \sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} Q(i, j, S) \tag{A.7}$$

Note that $Q_i$ does not depend on $t$. By the independence of both policy decisions and customer decisions across time, the total sales of inventory $i$ is a sum of $T$ trials of independent random variables $Q_i$, truncated by the starting inventory $b_i$. To summarize, the expected consumption of item $i$, regardless of inventory availability, is $\mathbb{E}[\min\{\sum_{t \in [T]} Q_i(\xi_t), b_i\}]$, with $Q_i$ defined as in (A.7), for all $i \in [n]$. We have used $\xi_t$ to emphasize the randomness in each trial.

Now denote $\rho_i := \sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} q(i, j, S)$. Due to Lemma A.3, we have $\mathbb{E}[\min\{\sum_{t \in [T]} Q_i(\xi_t), b_i\}] \geq \mathbb{E}[\min\{\text{Bin}(T, \rho_i), b_i\}]$. Summing over all $i \in [n]$, the expected revenue of the policy is

$$\begin{aligned}
&\sum_{i=1}^{n} \mathbb{E}[\min\{\text{Bin}(T, \rho_i), b_i\}] \cdot p_{\mathsf{C},i} \\
=&\sum_{i=1}^{n} \frac{\mathbb{E}[\min\{\text{Bin}(T, \rho_i), b_i\}]}{T\rho_i} \sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} p_j q(i, j, S) \\
\geq&\sum_{i=1}^{n} \frac{\mathbb{E}[\min\{\text{Bin}(T, b_i/T), b_i\}]}{T \cdot (b_i/T)} \sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j) \in S} p_j q(i, j, S)
\end{aligned}$$

$$\geq \frac{\mathbb{E}[\min\{\text{Bin}(T, \underline{b}/T), \underline{b}\}]}{\underline{b}} \sum_{i=1}^{n} \sum_{S \in \mathcal{S}} x^*(S) \sum_{j:(i,j)\in S} p_j q(i,j,S).$$

The first inequality follows from Lemma A.5, because $\rho_i \leq b_i/T$ for all $i$, since the LP solution $x^*(S)$ satisfies constraints (2.2). The second inequality is due to Lemma A.6.

□

*Proposition* A.9. *For all $b \leq T$, $\Delta^{APX} = \mathbb{E}[\min\{\text{Bin}(T, b/T), b\}]/b$ is non-increasing in $T$, and*

$$\lim_{T\to+\infty} \Delta^{APX} = \left(1 - \frac{\lfloor b\rfloor^{\lfloor b\rfloor}}{\lfloor b\rfloor!} e^{-\lfloor b\rfloor}\right) \geq 1 - 1/e$$

*Proof.* Proof of Proposition A.9. Since $T \in [N]$, and $T \geq b \geq \lfloor b\rfloor > 0$, from Lemma A.6 we have

$$\Delta^{APX} = \frac{\mathbb{E}[\min\{\text{Bin}(T, b/T), b\}]}{b} \geq \frac{\mathbb{E}[\min\{\text{Bin}(T, \lfloor b\rfloor/T), \lfloor b\rfloor\}]}{\lfloor b\rfloor}.$$

Since we have normalized $b$ such that $b \geq 1$, we know that $\lfloor b\rfloor \geq 1$. Now we wish to prove Proposition A.9 when $b \in [N]$ is any positive integer, i.e. we wish to show that for any positive integer $b \leq T$,

$$\lim_{T\to+\infty} \Delta^{APX} = \left(1 - \frac{b^b}{b!} e^{-b}\right) \geq 1 - 1/e$$

Denote $a = b/T$. We prove the first equality by telescoping. Denote $C_n^m = \frac{n!}{m!(n-m)!}$ as $n$ choose $m$.

$$\begin{aligned}
b\Delta^{APX} =& E\left[\min\left\{Bin(T, b/T), b\right\}\right] \\
=& Ta - \sum_{i=b+1}^{T} C_T^i (a^i (1-a)^{T-i})(i-b) \\
=& Ta - \sum_{i=b+1}^{T} \frac{T!}{(i-1)!(T-i)!} a^i (1-a)^{T-i} + \sum_{i=b+1}^{T} \frac{bT!}{i!(T-i)!} a^i (1-a)^{T-i} \\
=& Ta - \frac{T!}{b!(T-b-1)!} a^{b+1} (1-a)^{T-b-1} + \frac{bT!}{T!0!} a^T (1-a)^0
\end{aligned}$$

$$
\begin{aligned}
& - \sum_{i=b+1}^{T-1} \frac{T!}{i!(T-i-1)!} a^{i+1}(1-a)^{T-i-1} + \sum_{i=b+1}^{T-1} \frac{bT!}{i!(T-i)!} a^i (1-a)^{T-i} \\
=& Ta - \frac{T!}{b!(T-b-1)!} a^{b+1}(1-a)^{T-b-1} + ba^T \\
& + \sum_{i=b+1}^{T-1} \frac{T!}{i!(T-i)!} a^i (1-a)^{T-i-1} \underbrace{(b(1-a) - a(T-i))}_{=a(i-b)+(b-aT)\geq 0} \\
\geq & b\left(1 + a^T - C_{T-1}^b a^b (1-a)^{T-b-1}\right) \\
=& b\left(1 + a^T - C_T^b a^b (1-a)^{T-b}\right)
\end{aligned}
$$

where the fourth equality follows from telescoping. Then we can take $T \to +\infty$ and use Stirling's formula:

$$
\begin{aligned}
\lim_{T\to+\infty} \Delta^{APX} =& \lim_{T\to+\infty} 1 + a^T - C_T^b a^b (1-a)^{T-b} \\
=& 1 + \lim_{T\to+\infty} \left(\frac{b}{T}\right)^T - \lim_{T\to+\infty} \frac{T! b^b (T-b)^{T-b}}{b!(T-b)! T^b T^{T-b}} \\
=& 1 + 0 - \lim_{T\to+\infty} \frac{\sqrt{2\pi T}\frac{T^T}{e^T}(T-b)^{T-b}}{\sqrt{2\pi(T-b)}\frac{(T-b)^{T-b}}{e^{T-b}} T^T} \frac{b^b}{b!} \\
=& 1 - \frac{b^b}{b!} e^{-b}
\end{aligned}
$$

This term is increasing in $b$. So it obtains minimum when $b = 1$:

$$
\lim_{T\to+\infty} \Delta^{APX} = 1 - \frac{b^b}{b!} e^{-b} \geq 1 - \frac{1}{e}
$$

□

## A.6 Tightness of Theorem 2.2: Proof of Proposition 2.5

*Proof.* Proof of Proposition 2.5. Construct the following instance. There is only one price option, i.e. $m = 1$. So there only exists one calendar to sell at this single price everyday. For any given $b$ and $T$, the only price option has a purchase probability of

$b/T$, and earns 1 unit revenue.

The LP upper bound suggests a total of $b$ units of revenue. And the only calendar earns $\mathbb{E}\left[\min\left\{Bin(T, b/T), b\right\}\right]$ units of revenue. So the expected revenue of the only policy is exactly expression (2.9), which finishes the tightness proof. $\square$

## A.7 Necessity of Assumptions for Theorem 2.2: Proof of Proposition 2.3

*Proof.* Proof of Proposition 2.3. Consider the following problem with $T = 5$ periods. There are 3 products, $A$, $B$, and $C$. Products $A$ and $B$ use the first resource, and product $C$ uses the second resource. Both resources have initial inventory 1. Product $A$ is sold at $p_{\mathsf{H}} = 1$, and products $B$ and $C$ are sold at $p_{\mathsf{L}} = \epsilon$.

The choice model is a distribution of ordinal preferences: it takes $C \succ B \succ \emptyset$ with probability $1/5$, $A \succ \emptyset$ with probability $1/5 - \epsilon$, $B \succ \emptyset$ with probability $\epsilon$, and $\emptyset$ with probability $3/5$. Since this choice model is prescribed by a distribution of ordinal preferences, it satisfies Assumption 2.1, the substitutability assumption. Since consumptions from this choice model are binary, it satisfies Assumption 2.2.

The optimal solution from the LP is to offer assortment $\{A, B, C\}$ in all 5 periods. And the LP objective value is $1 + O(\epsilon)$. With some calculation, the actual expected revenue is $1959/3125 + O(\epsilon)$. Taking $\epsilon \to 0^+$ we have $1959/3125 \approx 0.6269 < 0.6321 \approx 1 - 1/e$. $\square$

## A.8 Proof of Theorem 2.6

*Proof.* Proof of Theorem 2.6. Denote the following random variables, which depict a run of our assortment policy from Algorithm 2. Let $A_t(S)$ be the indicator random variable for $\tilde{S}_t = S$, where $\tilde{S}_t$ was the assortment selected before discarding in (2.13) was applied. Let $B_t(i)$ be the remaining inventory of item $i$ at the *end* of time $t$. Defined for all $i \in [n]$ and $t = 0, \ldots, T$, where $B_0(i) = b_i$ for all $i$.

Under either static or dynamic substitution, let $R_t(i, j)$ be the amount of sales that a customer chooses product $(i,j)$ during period $t$. We will always specify the distribution of $R_t(i,j)$, by using a conditional probability. For example, we will use $\mathbb{E}[R_t(i,j)\,|S_t = S, \boldsymbol{B}_t = \boldsymbol{B}]$, for the expected sales that a customer chooses product $(i,j)$ during period $t$, when we plan to offer assortment $S$, and when the remaining inventory level for each resource is $\boldsymbol{B} = (B_1, B_2, ..., B_n)$.

Under static substitution, conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{R}^n_+$, $R_t(i,j) = \min\{B_i, Q\}$, where $Q$ is a random variable whose CDF is $F_{t,(i,j,S)}(\cdot)$. Under dynamic substitution, conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{N}^n_0$, $R_t(i,j)$ takes 1 with probability $q_t(i,j,\breve{S})$. Here we define $\breve{S} = \{(i,j) \in S\,|B_i > 0\}$ to be a function of $S$.

Under dynamic substitution, Assumption 2.1 suggests that $q_t(i,j,\breve{S}) \geq q_t(i,j,S)$, $\forall (i,j) \in \breve{S}$, because $\breve{S} \subseteq S$. The demand that originally would have chosen the stocked out items would go to their substitutes (as well as leaving, in which case the inequality takes equality). On the other hand, $q_t(i,j,\breve{S}) = 0, \forall (i,j) \notin \breve{S}$. The demand for any stocked out item is zero. We can use indicator variables to write the above inequalities in a compact form $q_t(i,j,\breve{S}) \geq \min\{B_{t-1}(i), q_t(i,j,S)\}$.

In all, we have

$$\mathbb{E}[R_t(i,j)\,|S_t = S, \boldsymbol{B}_t = \boldsymbol{B}] \geq \mathbb{E}_{Q\sim F_{t,(i,j,S)}}[\min\{B_{t-1}(i), Q\}] \tag{A.8}$$

where $F_{t,(i,j,S)}$ may prescribe a Bernoulli distribution, e.g. under dynamic substitution.

For any period $t$, given the remaining inventory from the last period to be $\boldsymbol{B}_{t-1}$, conditional on any $S \in \mathcal{S}$, the remaining inventory updates in the following fashion,

$$B_t(i) = B_{t-1}(i) - R_t(i,j), \forall i.$$

Note that no item can be offered multiple times at different prices in one assortment. Also note that we have defined $R_t(i,j)$ as the amount of sales, so $R_t(i,j)$ can never go beyond $B_{t-1}(i)$.

Following each sample path, we let $\mathsf{Rev}$ denote the revenue earned by the policy suggested in Algorithm 2.

$$
\begin{aligned}
&\mathbb{E}[\mathsf{Rev}] \\
=&\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{(i,j)\in D(S)}p_j\mathbb{E}[R_t(i,j)\,|S_t=D(S),\boldsymbol{B}_t=\boldsymbol{B}_{t-1}] \\
\geq&\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{(i,j)\in D(S)}p_j\mathbb{E}_{Q\sim F_{t,(i,j,D(S))}}[\min\{B_{t-1}(i),Q\}] \\
=&\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{(i,j)\in D(S)}(p_j-\frac{r_i^*}{2b_i})\mathbb{E}_{Q\sim F_{t,(i,j,D(S))}}[\min\{B_{t-1}(i),Q\}] \\
&+\sum_{i=1}^{n}\frac{r_i^*}{2b_i}\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in D(S)}\mathbb{E}_{Q\sim F_{t,(i,j,D(S))}}[\min\{B_{t-1}(i),Q\}] \\
=&\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{(i,j)\in D(S)}(p_j-\frac{r_i^*}{2b_i})\mathbb{E}_{Q\sim F_{t,(i,j,D(S))}}[\min\{B_{t-1}(i),Q\}]+\sum_{i=1}^{n}\frac{r_i^*}{2b_i}(b_i-\mathbb{E}_{\boldsymbol{Q}}[B_T(i)])
\end{aligned}
$$

where the first inequality is due to (A.8); the second equality is due to our discarding rule from (2.13); the last equality is counting how much inventory has been sold throughout the horizon, where we implicitly use $\boldsymbol{Q}$ to stand for the randomness from all periods.

We further observe that, $\forall t\in[T], S\in\mathcal{S}, (i,j)\in S$,

$$
\begin{aligned}
&\mathbb{E}_{Q\sim F_{t,(i,j,S)};B_{t-1}(i)}[\min\{B_{t-1}(i),Q_t(i,j,S)\}] \\
=&\mathbb{E}_{Q\sim F_{t,(i,j,S)};B_{t-1}(i)}\left[\frac{B_{t-1}(i)\cdot Q}{\max\{B_{t-1}(i),Q\}}\right] \\
\geq&\mathbb{E}_{Q\sim F_{t,(i,j,S)},B_{t-1}(i)}\left[\frac{B_{t-1}(i)\cdot Q}{b_i}\right] \\
=&\frac{\mathbb{E}_{B_{t-1}(i)}[B_{t-1}(i)]}{b_i}\cdot\mathbb{E}_{Q\sim F_{t,(i,j,S)}}[Q] \\
\geq&\frac{\mathbb{E}[B_T(i)]}{b_i}\cdot q_t(i,j,S)
\end{aligned}
\tag{A.9}
$$

where the first inequality is due to $B_{t-1}(i)\leq b_i$ inventory is no more than initial inventory level, and $Q\leq b_i,\forall t\in[T]$ demand is smaller than initial inventory; the second equality is because $B_{t-1}(i)$ is independent to any consumption that occurs at time $t$; the last inequality is because $B_t(i)$ is non-increasing in $t$, and evaluating the expectation of $\mathbb{E}_{Q\sim F_{t,(i,j,S)}}[Q] = q_t(i,j,S)$.

Finally, the expectation of the our policy's revenue can be decomposed as

$$
\begin{aligned}
&\mathbb{E}[\mathsf{Rev}] \\
&= \sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}\mathbb{E}[A_t(S)]\sum_{j:(i,j)\in D(S)}(p_j - \frac{r_i^*}{2b_i})\mathbb{E}_{Q\sim F_{t,(i,j,D(S))}}[\min\{B_{t-1}(i), Q\}] + \sum_{i=1}^{n}\frac{r_i^*}{2b_i}(b_i - \mathbb{E}[B_T(i)]) \\
&\geq \sum_{i=1}^{n}\frac{\mathbb{E}[B_T(i)]}{b_i}\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}x_t^*(S)\sum_{j:(i,j)\in D(S)}(p_j - \frac{r_i^*}{2b_i})q_t(i,j,D(S)) + \sum_{i=1}^{n}\frac{r_i^*}{2b_i}(b_i - \mathbb{E}[B_T(i)]) \\
&\geq \sum_{i=1}^{n}\frac{\mathbb{E}[B_T(i)]}{b_i}\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}x_t^*(S)\sum_{j:(i,j)\in S}(p_j - \frac{r_i^*}{2b_i})q_t(i,j,S) + \sum_{i=1}^{n}\frac{r_i^*}{2b_i}(b_i - \mathbb{E}[B_T(i)]) \\
&\geq \sum_{i=1}^{n}\frac{\mathbb{E}[B_T(i)]}{b_i}\left(r_i^* - \frac{r_i^*}{2b_i}\cdot b_i\right) + \sum_{i=1}^{n}\frac{r_i^*}{2b_i}(b_i - \mathbb{E}[B_T(i)]) \\
&= \sum_{i=1}^{n}\frac{r_i^*}{2}
\end{aligned}
$$

where the first equality is because the selection of set $\tilde{S}_t$ is independent to the remaining inventory $B_{t-1}(i)$, and also independent to the consumption $Q \sim F_{t,(i,j,D(S))}$ in period $t$; the first inequality is because of (A.9) and evaluating the expectation of $\mathbb{E}[A_t(S)] = x_t^*(S)$; the second inequality is because (i) we include non-positive terms into the summation, and (ii) $q_t(i,j,D(S)) \geq q_t(i,j,S)$ since $D(S) \subseteq S$, due to Assumption 2.1. the third inequality is due to the definition of $r_i^*$ in the policy, and the fact that the inventory constraint (2.2) in the LP is satisfied. Thus we complete the proof of the theorem. □

## A.9 Tighness of Theorem 2.6: Proof of Proposition 2.7

*Proof.* Proof of Proposition 2.7. We use the problem instance suggested by Example 2.1. Observe that any revenue-maximizing policy offers $p_1$ in Day 2. So the only decision to make is in Day 1. Since offering $p_1$ in Day 1 sells nothing, the outside option of selling nothing on Day 1 is simply the same as offering $p_1$ in Day 1.

Now we parameterize any revenue-maximizing policy by $x$, its probability in offering $p_1$ in Day 1; and $1-x$ is the probability in offering $p_2$ in Day 1. In expectation, this policy earns $\mathbb{E}[\mathsf{Rev}] = x \cdot (0 + \epsilon \cdot 1/\epsilon) + (1-x) \cdot ((1-\epsilon) \cdot 1 + \epsilon \cdot 1/\epsilon) = 1$ unit of revenue. But the LP objective is $2 - \epsilon$. So taking $\epsilon \to 0^+$, the expected revenue of any policy is upper-bounded by $\mathsf{OPT}_{\mathrm{LP}}/2$. $\square$

## A.10 Proof of Theorem 2.8

*Proof.* Proof of Theorem 2.8. Denote $\hat{\delta} = \sqrt{\frac{2\log(\underline{b})}{\underline{b}}}$. And note that $\delta b_i - 1 \geq \hat{\delta} b_i, \forall i \in [n]$, because $\underline{b} \geq 6$.

Denote the following random variables, which depict a run of our assortment policy from Algorithm 2. Let $A_t(S)$ be the indicator random variable for $\tilde{S}_t = S$, where $\tilde{S}_t$ was the assortment selected before discarding in (2.13) was applied. Let $B_t(i)$ be the remaining inventory of item $i$ at the *end* of time $t$. Defined for all $i \in [n]$ and $t = 0, \ldots, T$, where $B_0(i) = b_i$ for all $i$.

Under either static or dynamic substitution, let $R_t(i,j)$ be the amount of sales that a customer chooses product $(i,j)$ during period $t$. We will always specify the distribution of $R_t(i,j)$, by using a conditional probability. For example, we will use $\mathbb{E}[R_t(i,j) \,|\, S_t = S, \boldsymbol{B}_t = \boldsymbol{B}]$, for the expected sales that a customer chooses product $(i,j)$ during period $t$, when we plan to offer assortment $S$, and when the remaining inventory level for each resource is $\boldsymbol{B} = (B_1, B_2, ..., B_n)$.

Under static substitution, conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{R}_+^n$, $R_t(i,j) = \min\{B_i, Q\}$, where $Q$ is a random variable whose CDF is $F_{t,(i,j,S)}(\cdot)$. Under dynamic substitution, conditional on any $S \in \mathcal{S}, \boldsymbol{B} \in \mathbb{N}_0^n$, $R_t(i,j)$ takes 1 with probability $q_t(i,j,\breve{S})$. Here we define $\breve{S} = \{(i,j) \in S \,|\, B_i > 0\}$ to be a function of $S$.

Under dynamic substitution, Assumption 2.1 suggests that $q_t(i,j,\breve{S}) \geq q_t(i,j,S)$, $\forall (i,j) \in \breve{S}$, because $\breve{S} \subseteq S$. The demand that originally would have chosen the stocked out items would go to their substitutes (as well as leaving, in which case the inequality takes equality). On the other hand, $q_t(i,j,\breve{S}) = 0, \forall (i,j) \notin \breve{S}$. The demand for any stocked out item is zero. We can use indicator variables to write the above inequalities in a compact form $q_t(i,j,\breve{S}) \geq \min\{B_{t-1}(i), q_t(i,j,S)\}$.

In all, we have

$$\mathbb{E}[R_t(i,j)\,|S_t = S, \boldsymbol{B}_t = \boldsymbol{B}] \geq \mathbb{E}_{Q\sim F_{t,(i,j,S)}}[\min\{B_{t-1}(i), Q\}] \tag{A.10}$$

where $F_{t,(i,j,S)}$ may prescribe a Bernoulli distribution, e.g. under dynamic substitution.

For any period $t$, given the remaining inventory from the last period to be $\boldsymbol{B}_{t-1}$, conditional on any $S \in \mathcal{S}$, the remaining inventory updates in the following fashion,

$$B_t(i) = B_{t-1}(i) - R_t(i,j), \forall i.$$

Note that no item can be offered multiple times at different prices in one assortment. Also note that we have defined $R_t(i,j)$ as the amount of sales, so $R_t(i,j)$ can never go beyond $B_{t-1}(i)$.

Following each sample path, we let $\mathsf{Rev}$ denote the revenue earned by the policy suggested in Algorithm 3.

$$\begin{aligned}
\mathbb{E}[\mathsf{Rev}] &= \sum_{t=1}^{T}\sum_{S\in\mathcal{S}} A_t(S) \sum_{(i,j)\in S} p_j \mathbb{E}[R_t(i,j)\,|S_t = S, \boldsymbol{B}_t = \boldsymbol{B}_{t-1}] \\
&\geq \sum_{t=1}^{T}\sum_{S\in\mathcal{S}} A_t(S) \sum_{(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{t,(i,j,S)}}[\min\{B_{t-1}(i), Q\}]
\end{aligned}$$

Under static substitution, this inequality takes equality, and requires no assumption; under dynamic substitution, this inequality is true due to Assumption 2.1.

Taking expectation we have the following:

$$\begin{aligned}
&\mathbb{E}[\mathsf{Rev}] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{A.11}) \\
&= \sum_{t=1}^{T}\sum_{S\in\mathcal{S}} \mathbb{E}[A_t(S)] \sum_{(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{t,(i,j,S)}}[\min\{B_{t-1}(i), Q\}] \\
&\geq \sum_{t=1}^{T}\sum_{S\in\mathcal{S}} \mathbb{E}[A_t(S)] \sum_{(i,j)\in S} p_j \mathbb{E}_{Q\sim F_{t,(i,j,S)}}[\min\{B_{t-1}(i), Q\}\,|B_{t-1}(i) > 1] \cdot \Pr\{B_{t-1}(i) > 1\}
\end{aligned}$$

$$= \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} \mathbb{E}[A_t(S)] \sum_{(i,j) \in S} p_j \mathbb{E}_{Q \sim F_{t,(i,j,S)}}[Q \,|B_{t-1}(i) > 1] \cdot \Pr\{B_{t-1}(i) > 1\} \tag{A.12}$$
$$\geq \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} \mathbb{E}[A_t(S)] \sum_{(i,j) \in S} p_j \mathbb{E}_{Q \sim F_{t,(i,j,S)}}[Q] \cdot \Pr\{B_T(i) > 1\}$$
$$= \sum_{i=1}^{n} \Pr\{B_T(i) > 1\} \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} (1-\delta) x_t^*(S) \sum_{j:(i,j) \in S} p_j q_t(i,j,S)$$

where the first equality is because the random variables $A_t(S), B_{t-1}(i)$ and $Q \sim F_{t,(i,j,S)}$ are independent; the first inequality is re-writing the expectation by a conditional expectation, while ignoring the happening of other events "$B_{t-1}(i) \leq 1$"; the second equality is because conditioning on $B_{t-1}(i) > 1$, there must not be truncated demands during periods $t \in [T]$; the second inequality is because random variables $B_{t-1}(i)$ and $Q \sim F_{t,(i,j,S)}$ are independent, and that $B_t(i)$ is non-increasing in $t, \forall t \in [T]$.

We first lower bound $\Pr\{B_T(i) > 1\} = \Pr\{\sum_{t=1}^{T} \sum_{S \in \mathcal{S}} A_t(S) \sum_{j:(i,j) \in S} Q_t(i,j,S) < b_i - 1\}$, the probability that inventory $i$ never runs out. Conditioning on the event that inventory never runs out, we know that dynamic substitution will never happen. In all the remaining proof, we will only use notations like $Q(i,j,S)$, for the random quantity that customers attempt to purchase product $(i,j)$, should assortment $S$ be offered, no matter if any of the items from the assortment is stocked out. Note that the expected amount of inventory sold is strictly less than $b_i - 1$:

$$\mathbb{E}[\sum_{t=1}^{T} \sum_{S \in \mathcal{S}} A_t(S) \sum_{j:(i,j) \in S} Q_t(i,j,S)] =$$
$$(1-\delta) \sum_{t=1}^{T} \sum_{S \in \mathcal{S}} x_t^*(S) \sum_{j:(i,j) \in S} q_t(i,j,S) \leq (1-\delta) b_i < b_i - 1, \tag{A.13}$$

where the first inequality is due to constraint (2.2); the second inequality due to $\underline{b} \geq 6$. Since strict inequality holds, we can lower bound $\Pr\{B_T(i) > 0\}$ as follows.

$$\Pr\{B_T(i) > 0\}$$

$$
\begin{aligned}
=&1-\Pr\left\{\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S)\geq b_i\right\}\\
\geq&1-\Pr\left\{\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S)-\mathbb{E}[\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S)]\geq \delta b_i\right\}\\
\geq&1-\exp\left(-\frac{(\hat{\delta}b_i)^2}{2\mathsf{Var}(\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S))+2/3\hat{\delta}b_i}\right)\\
\geq&1-\exp\left(-\frac{\hat{\delta}^2 b_i}{2}\right)\\
\geq&1-\frac{1}{\underline{b}},
\end{aligned}
\tag{A.14}
$$

where the first inequality is due to (A.13); second inequality is Bernstein Inequality, where $\forall t, \left|\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S)-\mathbb{E}[\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S)]\right|\leq 1$ are zero-mean random variables and almost surely bounded by 1, and because $\delta b_i - 1 \geq \hat{\delta}b_i, \forall i\in[n]$; third inequality is because

$$
\begin{aligned}
&\mathsf{Var}(\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}A_t(S)\sum_{j:(i,j)\in S}Q_t(i,j,S))\\
\leq&\sum_{t=1}^{T}\left(\sum_{S\in\mathcal{S}}(1-\delta)x_t^*(S)\sum_{j:(i,j)\in S}q_t(i,j,S)\cdot\left(1-\sum_{S\in\mathcal{S}}(1-\delta)x_t^*(S)\sum_{j:(i,j)\in S}q_t(i,j,S)\right)\right)\\
\leq&\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}(1-\delta)x_t^*(S)\sum_{j:(i,j)\in S}q_t(i,j,S)\\
\leq&(1-\delta)b_i\leq(1-\hat{\delta})b_i.
\end{aligned}
$$

This is because if one random variable with bounded support over $[0,1]$ has the same mean as a Bernoulli random variable, its variance should be smaller than that of the Bernoulli random variable.

Finally putting (A.14) into (A.12) we have

$$
\mathbb{E}[\mathsf{Rev}]\geq\sum_{i=1}^{n}\Pr\{B_T(i)>0\}\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}(1-\delta)x_t^*(S)\sum_{j:(i,j)\in S}p_j q_t(i,j,S)
$$

$$
\begin{aligned}
&\geq (1-\frac{1}{\underline{b}})(1-\delta)\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{S\in\mathcal{S}}x_t^*(S)\sum_{j:(i,j)\in S}p_j q_t(i,j,S)\\
&=(1-\frac{1}{\underline{b}})(1-\sqrt{\frac{3\log(\underline{b})}{\underline{b}}})\mathsf{OPT}_{\mathrm{LP}} = (1-\sqrt{\frac{3\log(\underline{b})}{\underline{b}}}+o(\sqrt{\frac{\log(\underline{b})}{\underline{b}}}))\mathsf{OPT}_{\mathrm{LP}}
\end{aligned}
$$

which finishes the proof. By taking $b \to \infty$ we see the calendar is asymptotically optimal. □

## A.11 Proof of Theorem 2.9

*Proof.* Proof of Theorem 2.9. We introduce the following notations. Denote

$$\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t, S_t = S\,)$$

Denote "$\tilde{\boldsymbol{z}}\,\Big|S_\tau = \tilde{S}_\tau, \forall \tau \in \Gamma \subseteq \mathcal{S}$" to be a tweaked vector from $\boldsymbol{z}$, such that $\tilde{z}_\tau(\mathsf{X}) = \mathbb{1}_{\mathsf{X}=\tilde{S}_\tau}, \forall \tau \in \Gamma, \forall \mathsf{X} \in \mathcal{S}; \tilde{z}_\varrho(X) = z_\varrho(X), \forall \varrho \notin \Gamma, \forall X \in \mathcal{S}$. For example, $\tilde{\boldsymbol{z}}\,|S_1 = S$ is defined by $\tilde{z}_1(\mathsf{X}) = \mathbb{1}_{\mathsf{X}=S}, \forall \mathsf{X} \in \mathcal{S}; \tilde{z}_\varrho(X) = z_\varrho(X), \forall \varrho \geq 2, X \in \mathcal{S}$. Similarly, denote $\mathbb{E}[\mathsf{Rev}\,\Big|S_\tau = \tilde{S}_\tau, \forall \tau \in \Gamma\,]$ to be the expected revenue earned from a probabilistic offering of $\tilde{\boldsymbol{z}}\,\Big|S_\tau = \tilde{S}_\tau, \forall \tau \in \Gamma$.

Notice that, rigorously, $S_t(\boldsymbol{z})$ should be a random assortment based on the simulation results that are random. So rigorously, we should use $\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t]]$, because the conditioned event is a random event based on the simulation results. We take the outer expectation over the simulation randomness $\boldsymbol{\xi}$. In the inner expectation where we do not epecifically designate the source of randomness, the expectation is taken to find the expected revenue.

In each iteration of Algorithm 4, denote also $S_t^*(\boldsymbol{z})$ to be any element from

$$S_t^*(\boldsymbol{z}) \in \underset{S\in\{S\in\mathcal{S}|z_t(S)>0\}}{\arg\max}\ \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t, S_t = S],$$

which is the true best assortment to select in this iteration, if we were given a perfect oracle to query the expected revenue of a policy (instead of a simulator).

In each iteration, denote $O_t = \{S \in \mathcal{S} \,|\, z_t(S) > 0\}$ to be the set of candidate assortments to choose from. Let $o_t = |O_t|$. In each iteration, we can upper bound the sampling error incurred due to selecting the best empirical assortment, instead of (possibly) the true best assortment. $\forall t \in [T]$

$$
\begin{aligned}
&\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S_t^*(\boldsymbol{z})] - \mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t]] \\
=&\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S_t^*(\boldsymbol{z})] - \mathbb{E}[\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S_t^*(\boldsymbol{z}))] + \\
&\mathbb{E}[\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S_t^*(\boldsymbol{z}))] - \mathbb{E}_{\boldsymbol{\xi}}[\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t)] + \\
&\mathbb{E}_{\boldsymbol{\xi}}[\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t)] - \mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t]] \\
\le&\mathbb{E}_{\boldsymbol{\xi}}[\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t)] - \mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t]] + 0 + 0 \\
\le&\mathbb{E}_{\boldsymbol{\xi}}[\max_{S \in O_t} \{\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S) - \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S]\}]
\end{aligned}
$$

where the equality is how we decompose the difference by three differences; the first inequality is because the second difference is non-negative, since we are selecting the optimizer for the empirical performance, and because the first difference is zero, since given any fixed assortment, the empirical estimation is an unbiased estimation.

Now we further bound the sampling error. $\forall h > 0$,

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\xi}}[\max_{S \in O_t} \{\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S) \\
&\qquad\qquad\qquad -\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S]\}] \\
=&\frac{1}{h} \log \exp\left(h \cdot \mathbb{E}_{\boldsymbol{\xi}}[\max_{S \in O_t} \{\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S)\right. \\
&\qquad\qquad\qquad \left. -\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S]\}]\right) \\
\le&\frac{1}{h} \log \mathbb{E}_{\boldsymbol{\xi}}\left[\exp\left(h \cdot \max_{S \in O_t} \{\hat{\mu}_K(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S)\right.\right. \\
&\qquad\qquad\qquad \left.\left. -\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S]\}\right)\right] \\
=&\frac{1}{h} \log \mathbb{E}_{\boldsymbol{\xi}}\left[\max_{S \in O_t}\left\{\exp\left(\frac{h}{K} \cdot (\sum_{k=1}^{K} \nu(\boldsymbol{z}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S\,, \xi_k)\right.\right.\right. \\
&\qquad\qquad\qquad \left.\left.\left. - \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \le t-1, S_t = S])\right)\right\}\right]
\end{aligned}
$$

$$
\begin{aligned}
\leq & \frac{1}{h} \log \mathbb{E}_{\boldsymbol{\xi}} \left[ \sum_{S \in O_t} \exp \left( \frac{h}{K} \cdot ( \sum_{k=1}^{K} \nu(\boldsymbol{z} \,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t-1, S_t = S\ , \xi_k) \right.\right. \\
& \qquad\qquad\qquad\qquad \left.\left. - \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t-1, S_t = S]) \right) \right] \\
= & \frac{1}{h} \log \sum_{S \in O_t} \prod_{k=1}^{K} \mathbb{E}_{\boldsymbol{\xi}} \left[ \exp \left( \frac{h}{K} \cdot (\nu(\boldsymbol{z} \,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t-1, S_t = S\ , \xi_k) \right.\right. \\
& \qquad\qquad\qquad\qquad \left.\left. - \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t-1, S_t = S]) \right) \right] \\
\leq & \frac{1}{h} \log \sum_{S \in O_t} \prod_{k=1}^{K} \exp \left( \frac{h^2}{K^2} \cdot \frac{(2(b_1 + \cdots + b_n) p_{\max})^2}{8} \right) \\
= & \frac{\log o_t}{h} + \frac{h(b_1 + \cdots + b_n)^2 p_{\max}^2}{2K}
\end{aligned}
$$

where the first equality is re-writing the same expression; the first inequality is due to Jensen's inequality, because $\forall h > 0, f(x) = \exp(h \cdot x)$ is a convex function in $x$; the second equality is from the definition of $\hat{\mu}_K(\cdot)$; the second inequality is because all the exponentials are positive, and we are replacing the maximum with their sum; the third equality is due to linearity of expectations and due to the independence of different simulations under randomness $\xi_k$; the third inequality is due to Hoeffding's Lemma, where each random simulation yields a number bounded by $(b_1+\cdots+b_n)p_{\max}$.

The above bound holds for any $h > 0$. If we pick $h = \sqrt{\frac{2K \log o_t}{(b_1+\cdots+b_n)^2 p_{\max}^2}}$, then the above term can be simplified as

$$
\begin{aligned}
& \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t-1, S_t = S_t^*(\boldsymbol{z})] - \mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t]] \\
\leq & \sqrt{\frac{\log o_t}{2K}} (b_1 + \cdots + b_n) p_{\max} \\
= & \sqrt{\frac{\log o_t}{2(\log n + \log T)}} \cdot \epsilon \cdot \mathsf{OPT}_{\mathrm{LP}}.
\end{aligned} \tag{A.15}
$$

Finally we conclude the proof by induction on $t$. In the first iteration of Algorithm 4, we can re-write $\mathbb{E}[\mathsf{Rev}]$ as follows,

$$
\mathbb{E}[\mathsf{Rev}] = \sum_{S \in \{S \in \mathcal{S} | z_1(S) > 0\}} \mathbb{E}[\mathsf{Rev}\,|S_1 = S] \cdot z_1(S) \geq \alpha \cdot \mathsf{OPT}_{\mathrm{LP}}, \tag{A.16}
$$

where the inequality holds due to Theorems 2.2 – 2.8. Since we have selected the $S$'s such that $z_1(S) > 0$, and that $\sum_{S\in\{S\in\mathcal{S}|z_1(S)>0\}} z_1(S) = 1$. So

$$\mathbb{E}[\mathsf{Rev}\,|S_1 = S_1^*(\boldsymbol{z})] = \max_{S\in\{S\in\mathcal{S}|z_1(S)>0\}} \mathbb{E}[\mathsf{Rev}\,|S_1 = S] \geq \alpha \cdot \mathsf{OPT}_{\mathrm{LP}},$$

because otherwise the summation in (A.16) is strictly smaller than $\alpha{\cdot}\mathsf{OPT}_{\mathrm{LP}}$. Then we can plug in inequality (A.15), so that $\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_1 = S_1(\boldsymbol{z})]] \geq (\alpha - \sqrt{\frac{\log o_1}{2(\log n+\log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}}$

Suppose we have shown that

$$\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t]] \geq (\alpha - \sum_{\tau=1}^{t} \sqrt{\frac{\log o_\tau}{2(\log n + \log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}}.$$

In the $(t+1)^{\text{th}}$ iteration of Algorithm 4, we can re-write the expected revenue as follows,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t]]\\
=&\sum_{S\in\{S\in\mathcal{S}|z_{t+1}(S)>0\}} \mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t, S_{t+1} = S] \cdot z_{t+1}(S)\\
\geq&(\alpha - \sum_{\tau=1}^{t} \sqrt{\frac{\log o_\tau}{2(\log n + \log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}},
\end{aligned}$$

Similarly, we have

$$\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,\big|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t, S_{t+1} = S_{t+1}^*(\boldsymbol{z})]] \geq (\alpha - \sum_{\tau=1}^{t} \sqrt{\frac{\log o_\tau}{2(\log n + \log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}},$$

and then using inequality (A.15) we have $\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq t+1]] \geq (\alpha - \sum_{\tau=1}^{t+1} \sqrt{\frac{\log o_\tau}{2(\log n+\log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}}$.

By induction, we have

$$\mathbb{E}_{\boldsymbol{\xi}}[\mathbb{E}[\mathsf{Rev}\,|S_\tau = S_\tau(\boldsymbol{z}), \forall \tau \leq T]] \geq (\alpha - \sum_{\tau=1}^{T} \sqrt{\frac{\log o_\tau}{2(\log n + \log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}}$$

$$
\begin{aligned}
&\geq (\alpha - \sqrt{\frac{\log\left(\sum_{\tau=1}^{T} o_\tau\right)}{2(\log n + \log T)}}\epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}} \\
&\geq (\alpha - \epsilon) \cdot \mathsf{OPT}_{\mathrm{LP}}
\end{aligned}
$$

where the first inequality is due to concavity of $\sqrt{\log x}$; the second inequality is because $\sum_{\tau=1}^{T} o_\tau \leq 2(\log n + \log T)$. In the context of non-stationary arrivals, $\sum_{\tau=1}^{T} o_\tau$ is the total number of non-zero variables from CDLP, can be upper bounded by the total number of constraints, which is $n + T$. In the context of stationary arrivals, $\sum_{\tau=1}^{T} o_\tau$ has bounded supports of $n + 1$ in each of the $T$ periods – so it could be bounded by $(n + 1) \cdot T$.

Finally, we analyze the time complexity. In each iteration, we enumerate over all the assortments $\{S \in \mathcal{S} \,|\, z_t(S) > 0\}$ to find the empirical maximizer $\hat{S}_t$, which involves $K$ queries to the simulator $\hat{\mu}(\cdot)$. There are at most $n+1$ non-zero variables to enumerate. So the total number of queries are no more than $K \cdot (n+1) \cdot T = O(KnT)$. In each query of the simulator $\hat{\mu}(\cdot)$, it takes no more than $(n + 1)T$ operations to generate a sequence of $T$ assortments. On the other hand, since we know the CDF of the demand in each period, it takes only $O(1)$ operations to generate a random demand. So the total number of arithmetic operations for our de-randomization method is $O(Kn^2T^2)$. As suggested by Algorithm 3, $K$ is also polynomial in $n$, $T$, and $1/\epsilon$. □

## A.12 Lemmas for the Proof of Theorem 2.11

In this section we prove Lemmas 2.12 and 2.13.

### A.12.1 Proof of Lemma 2.12

*Proof.* Proof of Lemma 2.12. Let $(x)^+ = \max\{x, 0\}$ denote the maximum of $x$ and 0. Let $Q_{v_t}$ denote the random demand if we offer price $v_t$ on day $t$, which follows a distribution of $F_{v_t}(\cdot)$.

Similar to the proof to Lemma 2.12, the idea is to exchange a pair of two consecutive prices. Given any calendar $\boldsymbol{v} = (v_1, v_2, ..., v_T)$, if there exists $t \in [T-1]$, such that $p_{v_t} < p_{v_{t+1}}$, we compare to another calendar:

$$\boldsymbol{v}^* = (v_1, v_2, ..., v_{t-1}, v_{t+1}, v_t, v_{t+2}, ..., v_T).$$

Since we only exchange this pair of two prices, the expected revenue change comes only from these two periods: Before $t$ the expected revenue is trivially not changed. Since under both calendars, the distribution of total units of inventory consumed during $t$ and $t+1$ are exactly the same, we know that the distribution of initial inventory at the end of $t+1$ will also be the same. So after period $t+1$ the revenue will not be changed, neither.

Let $A(c,t)$ denote the event "$c$ units of inventory left at the beginning of period $t$", where $c \in [0,b], t \in [T]$. Let $\mathsf{Rev}(\boldsymbol{v})$ be a random variable for the expected revenue from all time periods under $\boldsymbol{v}$. Let $\mathsf{Rev}_{t,t+1}(\boldsymbol{v})$ be a random variable for the expected revenue from $t$ and $t+1$ under $\boldsymbol{v}$. Then for any $c \in [0,b]$, conditioning on $A(c,t)$ we have:

$$\begin{aligned}
\mathbb{E}[\mathsf{Rev}_{t,t+1}(\boldsymbol{v})\,|A(c,t)] =& p_{v_t}\mathbb{E}\left[\min\{c, Q_{v_t}\}\right] + p_{v_{t+1}}\mathbb{E}\left[\min\{(c-Q_{v_t})^+, Q_{v_{t+1}}\}\right] \\
=& (p_{v_t} - p_{v_{t+1}})\mathbb{E}\left[\min\{c, Q_{v_t}\}\right] + p_{v_{t+1}}\mathbb{E}\left[\min\{c, Q_{v_t} + Q_{v_{t+1}}\}\right]
\end{aligned}$$

where the second equation holds because $\min\{(c-Q_{v_t})^+, Q_{v_{t+1}}\} = \min\{c, Q_{v_t} + Q_{v_{t+1}}\} - \min\{c, Q_{v_t}\}$, and linearity of expectations.

Now let us compare the expected revenue from two calendars, still conditioning on $A(c,t)$:

$$\begin{aligned}
&\mathbb{E}[\mathsf{Rev}(\boldsymbol{v}^*)\,|A(c,t)] - \mathbb{E}[\mathsf{Rev}(\boldsymbol{v})\,|A(c,t)] \\
=& (p_{v_{t+1}} - p_{v_t})\mathbb{E}\left[\min\{c, Q_{v_{t+1}}\}\right] + p_{v_t}\mathbb{E}\left[\min\{c, Q_{v_{t+1}} + Q_{v_t}\}\right] \\
&- (p_{v_t} - p_{v_{t+1}})\mathbb{E}\left[\min\{c, Q_{v_t}\}\right] - p_{v_{t+1}}\mathbb{E}\left[\min\{c, Q_{v_t} + Q_{v_{t+1}}\}\right] \\
=& (p_{v_{t+1}} - p_{v_t})(\mathbb{E}\left[\min\{c, Q_{v_t}\}\right] + \mathbb{E}\left[\min\{c, Q_{v_{t+1}}\}\right] - \mathbb{E}\left[\min\{c, Q_{v_t} + Q_{v_{t+1}}\}\right]) \\
\geq& 0
\end{aligned}$$

where the last inequality is due to the fact that $\min\{c, x\} + \min\{c, y\} \geq \min\{c, x + y\}, \forall c, x, y \geq 0$, which is proved in the appendix as Lemma A.1. Notice that this is true for all $c \in [0, b]$.

By integrating over $c \in [0, b]$ we have $\mathbb{E}[\mathsf{Rev}(\boldsymbol{v}^*)] \geq \mathbb{E}[\mathsf{Rev}(\boldsymbol{v})]$, which finishes the proof. □

*Corollary* A.9.1. *There exists an optimal static calendar whose prices are non-increasing over time, i.e.* $p_{v_t^*} \geq p_{v_{t+1}^*}, \forall t \in [T-1]$.

*Proof.* Proof of Corollary A.9.1. Directly follows from Lemma 2.12. We can start from any calendar and use a finite number (no more than $T!$) of exchange operations to achieve the optimal non-decreasing structure. □

### A.12.2 Bridge from Lemma 2.12 to Lemma 2.13

*Lemma* A.10. *Under [0,1]-demand, in any two-price randomized policy* $\boldsymbol{v}$*, if two consecutive probabilities* $v_t$*,* $v_{t+1}$ *are such that* $v_t < v_{t+1}$*, then probabilities* $v_t$*,* $v_{t+1}$ *can be exchanged in the calendar without decreasing its expected revenue.*

*Proof.* Proof of Lemma A.10. Let $(x)^+ = \max\{x, 0\}$ denote the maximum of $x$ and 0. Let $Q_{\mathsf{H},t}$ and $Q_{\mathsf{L},t}$ denote the random demand if we offer the higher price and the lower price on day $t$, which follows a distribution of $F_H(\cdot)$ and $F_L(\cdot)$, respectively.

The idea is to exchange a pair of two consecutive prices. Given any calendar $\boldsymbol{v} = (v_1, v_2, ..., v_T)$, if there exists $t \in [T-1]$, such that $v_t < v_{t+1}$, we compare to another calendar:

$$\boldsymbol{v}^* = (v_1, v_2, ..., v_{t-1}, v_{t+1}, v_t, v_{t+2}, ..., v_T).$$

Since we only exchange this pair of two prices, the expected revenue before $t$ is trivially not changed.

Let $J_{\boldsymbol{v},t}(c)$ denote the expected revenue we would earn under calendar $\boldsymbol{v}$ if we were endowed with $c$ units of inventory at the beginning of period $t$. Its expectation is taken over future demand randomness. Let $A(c, t)$ denote the event "$c$ units of inventory at the beginning of period $t$", where $c \in [0,b], t \in [T]$. Then for any $c \in [0,b]$, conditioning on $A(c,t)$ we have:

$$
\begin{aligned}
J_{\boldsymbol{v},t}(c) =& v_t \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] p_{\mathsf{H}} + v_t v_{t+1} \mathbb{E}\left[\min\{(c-Q_{\mathsf{H},t})^+, Q_{\mathsf{H},t+1}\}\right] p_{\mathsf{H}} \\
&\qquad + v_t(1-v_{t+1})\mathbb{E}\left[\min\{(c-Q_{\mathsf{H},t})^+, Q_{\mathsf{L},t+1}\}\right] p_{\mathsf{L}} \\
&+ v_t v_{t+1}\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{H},t+1})^+)\right] \\
&+ v_t(1-v_{t+1})\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{L},t+1})^+)\right] \\
&+ (1-v_t)\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] p_{\mathsf{L}} + (1-v_t)v_{t+1}\mathbb{E}\left[\min\{(c-Q_{\mathsf{L},t})^+, Q_{\mathsf{H},t+1}\}\right] p_{\mathsf{H}} \\
&\qquad + (1-v_t)(1-v_{t+1})\mathbb{E}\left[\min\{(c-Q_{\mathsf{L},t})^+, Q_{\mathsf{L},t+1}\}\right] p_{\mathsf{L}} \\
&+ (1-v_t)v_{t+1}\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{H},t+1})^+)\right] \\
&+ (1-v_t)(1-v_{t+1})\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{L},t+1})^+)\right]
\end{aligned}
$$

Plugging in $\min\{(c-x)^+, y\} = \min\{c, x+y\} - \min\{c, x\}, \forall c, x, y \geq 0$ we have:

$$
\begin{aligned}
J_{\boldsymbol{v},t}(c) =& v_t \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] p_{\mathsf{H}} + v_t v_{t+1}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}+Q_{\mathsf{H},t+1}\}\right] p_{\mathsf{H}} \\
&\qquad + v_t(1-v_{t+1})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}+Q_{\mathsf{L},t+1}\}\right] p_{\mathsf{L}} \\
&- v_t v_{t+1}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] p_{\mathsf{H}} - v_t(1-v_{t+1})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] p_{\mathsf{L}} \\
&+ v_t v_{t+1}\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{H},t+1})^+)\right] \\
&+ v_t(1-v_{t+1})\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{L},t+1})^+)\right] \\
&+ (1-v_t)\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] p_{\mathsf{L}} + (1-v_t)v_{t+1}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}+Q_{\mathsf{H},t+1}\}\right] p_{\mathsf{H}} \\
&\qquad + (1-v_t)(1-v_{t+1})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}+Q_{\mathsf{L},t+1}\}\right] p_{\mathsf{L}} \\
&- (1-v_t)v_{t+1}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] p_{\mathsf{H}} - (1-v_t)(1-v_{t+1})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] p_{\mathsf{L}} \\
&+ (1-v_t)v_{t+1}\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{H},t+1})^+)\right] \\
&+ (1-v_t)(1-v_{t+1})\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{L},t+1})^+)\right]
\end{aligned}
$$

Merging similar expressions we have the following:

$$
J_{\boldsymbol{v},t}(c) = v_t(1-v_{t+1})(p_{\mathsf{H}}-p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right]
$$

$$
\begin{aligned}
&- (1 - v_t)v_{t+1}(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] \\
&+ v_t v_{t+1} p_{\mathsf{H}} \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ v_t(1 - v_{t+1}) p_{\mathsf{L}} \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ (1 - v_t)v_{t+1} p_{\mathsf{H}} \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ (1 - v_t)(1 - v_{t+1}) p_{\mathsf{L}} \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ v_t v_{t+1} \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ v_t(1 - v_{t+1}) \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
&+ (1 - v_t)v_{t+1} \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ (1 - v_t)(1 - v_{t+1}) \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{L},t+1})^+)\right]
\end{aligned} \tag{A.17}
$$

Similarly we have $J_{\boldsymbol{v}^*,t}(c)$, the expected revenue under calendar $\boldsymbol{v}^*$, as following:

$$
\begin{aligned}
J_{\boldsymbol{v}^*,t}(c) =& v_{t+1}(1 - v_t)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] \\
&- (1 - v_{t+1})v_t(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] \\
&+ v_{t+1} v_t p_{\mathsf{H}} \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ v_{t+1}(1 - v_t) p_{\mathsf{L}} \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ (1 - v_{t+1})v_t p_{\mathsf{H}} \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ (1 - v_{t+1})(1 - v_t) p_{\mathsf{L}} \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ v_{t+1} v_t \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ v_{t+1}(1 - v_t) \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
&+ (1 - v_{t+1})v_t \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ (1 - v_{t+1})(1 - v_t) \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{L},t+1})^+)\right]
\end{aligned}
$$

The equality holds because calendars $\boldsymbol{v}$ and $\boldsymbol{v}^*$ are the same from period $t+2$.

On the other hand, we know that $\forall c, x, y \geq 0, \min\{c, x+y\} \geq 0$ and $J_{\boldsymbol{v},t+2}((c - x - y)^+) \geq 0$. Due to Fubini's theorem, we can exchange the double integration:

$$
\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] = \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \tag{A.18}
$$

$$\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{L},t+1})^{+})\right]=\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{H},t+1})^{+})\right] \tag{A.19}$$

Then we calculate their difference, conditioning on $A(c,t)$:

$$\begin{aligned}
&J_{\boldsymbol{v}^*,t}(c)-J_{\boldsymbol{v},t}(c)\\
=&(v_{t+1}-v_t)(p_{\mathsf{H}}-p_{\mathsf{L}})\left(\mathbb{E}\left[\min\{c,Q_{\mathsf{H},t}\}\right]+\mathbb{E}\left[\min\{c,Q_{\mathsf{L},t}\}\right]-\mathbb{E}\left[\min\{c,Q_{\mathsf{H},t}+Q_{\mathsf{L},t+1}\}\right]\right)\\
=&(v_{t+1}-v_t)(p_{\mathsf{H}}-p_{\mathsf{L}})\left(\mathbb{E}\left[\min\{c,Q_{\mathsf{H},t}\}\right]+\mathbb{E}\left[\min\{c,Q_{\mathsf{L},t+1}\}\right]-\mathbb{E}\left[\min\{c,Q_{\mathsf{H},t}+Q_{\mathsf{L},t+1}\}\right]\right)\\
\geq&0
\end{aligned}$$

where the inequality holds because three components are all greater or equal to zero, the third component due to Lemma A.1. Notice that this is true for all $c\in[0,b]$.

By integrating over $c\in[0,b]$ we have $\mathbb{E}[\mathsf{Rev}(\boldsymbol{v}^*)]\geq\mathbb{E}[\mathsf{Rev}(\boldsymbol{v})]$, which finishes the proof. □

*Corollary* A.10.1. *There exists an optimal two-price randomized policy whose probabilities are non-increasing over time, i.e.* $v_t\geq v_{t+1},\forall t\in[T-1]$.

*Proof.* Proof of Corollary A.10.1. Directly follows from Lemma 2.12. We can start from any calendar and use a finite number (no more than $T!$) of exchange operations to achieve the optimal non-increasing structure. □

Notice that Lemma A.10 and Corollary A.10.1 are assumption-free.

### A.12.3 Proof of Lemma 2.13

*Lemma* A.11. *Under Assumption 2.3, in any two-price randomized policy* $\boldsymbol{v}$*, if two consecutive probabilities* $v_t$*,* $v_{t+1}$ *are such that* $v_t\geq v_{t+1},v_t<1,v_{t+1}>0$*, then the last pair of probabilities indexed by* $t=\sup_{\tau\in[T-1]}\{v_\tau<1,v_{\tau+1}>0,v_\tau\geq v_{\tau+1}\}$ *can be changed from* $(v_t,v_{t+1})$ *to* $(v_t+\delta,v_{t+1}-\delta)$ *where* $0\leq\delta\leq\max\{1-v_t,v_{t+1}\}$*, without decreasing its expected revenue.*

*Proof.* Proof. Let $(x)^+=\max\{x,0\}$ denote the maximum of $x$ and 0. Let $Q_{\mathsf{H},t}$ and $Q_{\mathsf{L},t}$ denote the random demand if we offer the higher price and the lower price on day $t$, which follows a distribution of $F_H(\cdot)$ and $F_L(\cdot)$, respectively.

Again our idea is to modify a pair of consecutive probabilities, to achieve a greater revenue. Given any calendar $\boldsymbol{v} = (v_1, v_2, ..., v_n)$, where $v_1 \geq v_2 \geq ... \geq v_T$. If $\exists \tau \in [T-1]$, such that $v_\tau < 1, v_{\tau+1} > 0, v_\tau \geq v_{\tau+1}$, we compare to another calendar:

$$\boldsymbol{v}^* = (v_1, v_2, ..., v_{t-1}, v_t + \delta, v_{t+1} - \delta, v_{t+2}, ..., v_T)$$

where $t = \sup_{\tau \in [T-1]} \{v_\tau < 1, v_{\tau+1} > 0, v_\tau \geq v_{\tau+1}\}$ and $0 \leq \delta \leq \min\{1 - v_t, v_{t+1}\}$. Notice that $t$ is the largest element in this set $\{v_\tau < 1, v_{\tau+1} > 0, v_\tau \geq v_{\tau+1}\}$, which indicates that $\forall \tau \geq t+2, v_\tau = 0$. Since we are changing probabilities from period $t$, the expected revenue does not change before period $t$.

Let $J_{\boldsymbol{v},t}(c)$ denote the expected revenue we would earn under calendar $\boldsymbol{v}$ if we were endowed with $c$ units of inventory at the beginning of period $t$. Its expectation is taken over future demand randomness.

Let $A(c,t)$ denote the event "$c$ units of inventory at the beginning of period $t$", where $c \in [0,b], t \in [T]$. Then for any $c \in [0,b]$, conditioning on $A(c,t)$ we can expand the expression of the expected revenue. We proceed from equation (A.17):

$$\begin{aligned}
J_{\boldsymbol{v}^*,t}(c) =& (v_t + \delta)(1 - v_{t+1} + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] \\
& - (1 - v_t - \delta)(v_{t+1} - \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] \\
& + (v_t + \delta)(v_{t+1} - \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
& + (v_t + \delta)(1 - v_{t+1} + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
& + (1 - v_t - \delta)(v_{t+1} - \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
& + (1 - v_t - \delta)(1 - v_{t+1} + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
& + (v_t + \delta)(v_{t+1} - \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
& + (v_t + \delta)(1 - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
& + (1 - v_t - \delta)(v_{t+1} - \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right] \\
& + (1 - v_t - \delta)(1 - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{L},t+1})^+)\right] \qquad \text{(A.20)}
\end{aligned}$$

where equality holds because calendars $\boldsymbol{v}$ and $\boldsymbol{v}^*$ are the same from period $t+2$.

Using (A.18) and (A.19), we subtract (A.20) and (A.17) to calculate their difference, conditioning on $A(c,t)$:

$$
\begin{aligned}
J_{\boldsymbol{v}^*,t}(c) - J_{\boldsymbol{v},t}(c) =& \delta(1 - v_{t+1} + v_t + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] \\
&+ \delta(1 - v_t + v_{t+1} - \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] \\
&+ \delta(v_{t+1} - v_t - \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ \delta(1 - v_{t+1} + v_t + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ \delta(-1 + v_t - v_{t+1} + \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ \delta(-v_t + v_{t+1} - \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ \delta(v_{t+1} - v_t - \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ \delta(-v_{t+1} + v_t + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
&+ \delta(v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ \delta(-v_t + v_{t+1} - \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{L},t+1})^+)\right]
\end{aligned}
$$

Merging similar expressions we have the following:

$$
\begin{aligned}
&\frac{J_{\boldsymbol{v}^*,t}(c) - J_{\boldsymbol{v},t}(c)}{\delta} \\
=&(p_{\mathsf{H}} - p_{\mathsf{L}})\left\{\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] + \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] - \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right]\right\} \\
&+ (v_t - v_{t+1} + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] \\
&- (v_t - v_{t+1} + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] \\
&- (v_t - v_{t+1} + \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ (v_t - v_{t+1} + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ (v_t - v_{t+1} + \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
&- (v_t - v_{t+1} + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
&- (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
&+ (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right]
\end{aligned}
$$

$$
\begin{aligned}
&- (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{L},t+1})^+)\right] \\
\geq &(v_t - v_{t+1} + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\left\{\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] + \mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right]\right. \\
&\qquad\qquad\qquad \left. - \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right]\right\} \\
&+ (v_t - v_{t+1} + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] \\
&- (v_t - v_{t+1} + \delta)(p_{\mathsf{H}} - p_{\mathsf{L}})\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}\}\right] \\
&- (v_t - v_{t+1} + \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ (v_t - v_{t+1} + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ (v_t - v_{t+1} + \delta)p_{\mathsf{H}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
&- (v_t - v_{t+1} + \delta)p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
&- (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
&+ (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&- (v_t - v_{t+1} + \delta)\mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{L},t+1})^+)\right]
\end{aligned}
$$

where the inequality holds because $\delta \leq \min\{1 - v_t, v_{t+1}\}$ so $v_t - v_{t+1} + \delta \leq 1$, and because of Lemma A.1.

Further merging similar expressions, while using (A.18) and (A.19) we have the following:

$$
\begin{aligned}
\frac{J_{\boldsymbol{v}^*,t}(c) - J_{\boldsymbol{v},t}(c)}{\delta(v_t - v_{t+1} + \delta)} \geq &(p_{\mathsf{H}} - p_{\mathsf{L}})\left\{2\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}\}\right] - \mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right]\right\} \\
&- p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{H},t+1}\}\right] \\
&+ p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t} + Q_{\mathsf{L},t+1}\}\right] \\
&+ p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{H},t+1}\}\right] \\
&- p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t} + Q_{\mathsf{L},t+1}\}\right] \\
&- \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{H},t+1})^+)\right] \\
&+ \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{H},t} - Q_{\mathsf{L},t+1})^+)\right] \\
&+ \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c - Q_{\mathsf{L},t} - Q_{\mathsf{H},t+1})^+)\right]
\end{aligned}
$$

$$
\begin{aligned}
&- \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{L},t+1})^+)\right] \\
\geq &- p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}+Q_{\mathsf{H},t+1}\}\right] \\
&+ p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{H},t}+Q_{\mathsf{L},t+1}\}\right] \\
&+ p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}+Q_{\mathsf{H},t+1}\}\right] \\
&- p_{\mathsf{L}}\mathbb{E}\left[\min\{c, Q_{\mathsf{L},t}+Q_{\mathsf{L},t+1}\}\right] \\
&- \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{H},t+1})^+)\right] \\
&+ \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{H},t}-Q_{\mathsf{L},t+1})^+)\right] \\
&+ \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{H},t+1})^+)\right] \\
&- \mathbb{E}\left[J_{\boldsymbol{v},t+2}((c-Q_{\mathsf{L},t}-Q_{\mathsf{L},t+1})^+)\right]
\end{aligned}
$$

where the second inequality is due to Lemma A.1.

Now let us introduce a coupling argument: suppose $U_\tau, \tau = t, t+1$ are two uniform distributions over $[0,1]$, and let $Q_{\mathsf{H},\tau} = F_{\mathsf{H}}^{-1}(U_\tau), Q_{\mathsf{L},\tau} = F_{\mathsf{L}}^{-1}(U_\tau), \tau = t, t+1$. Since $F_{\mathsf{H}}$ and $F_{\mathsf{L}}$ are CDF's of distributions, they are monotone increasing, thus the inverse function exists. And for any specific realization of $U_\tau, \tau = t, t+1$, we know that $Q_{\mathsf{H},\tau} \leq Q_{\mathsf{L},\tau}$ or $Q_{\mathsf{H},\tau} \geq Q_{\mathsf{L},\tau}$, due to Assumption 2.3.

Now we only need to understand $J_{\boldsymbol{v},t+2}(c)$ as a function of $c \in [0,b]$. Apparently $J_{\boldsymbol{v},t+2}(0) = 0$. Let us denote the distribution of random demands in the last $T-t+1$ periods as $X$, whose CDF denoted as $\hat{F}$. $X$ is a non-negative random variable. We sell at the lower price from period $t+2$, and so the expected revenue in the last $T-t+1$ periods is calculated as $J_{\boldsymbol{v},t+2}(c) = p_{\mathsf{L}}\mathbb{E}[\min\{c, X\}] = p_{\mathsf{L}} \int_{[0,b]} \min\{c,u\}\mathrm{d}\hat{F}(u)$. Due to Fubini's theorem, we can perform integration by part, and have:

$$
\begin{aligned}
J_{\boldsymbol{v},t+2}(c) =& p_{\mathsf{L}} \int_{[0,b]} \min\{c,u\}\mathrm{d}\hat{F}(u) \\
=& p_{\mathsf{L}} \int_{[0,c]} u\mathrm{d}\hat{F}(u) + p_{\mathsf{L}} \int_{[c,b]} c\mathrm{d}\hat{F}(u) \\
=& p_{\mathsf{L}}(c\hat{F}(c) - \int_{[0,c]} \hat{F}(u)\mathrm{d}u) + p_{\mathsf{L}}(c - c\hat{F}(c)) \\
=& p_{\mathsf{L}}(c - \int_{[0,c]} \hat{F}(u)\mathrm{d}u)
\end{aligned}
$$

This is a differentiable function with respect to $c$. Taking derivative we have: $J'_{\boldsymbol{v},t+2}(c) = p_{\mathsf{L}}(1-\hat{F}(c)) \in [0, p_{\mathsf{L}}]$. Since $\hat{F}(c)$ is a CDF, thus non-decreasing, we know that $J'_{\boldsymbol{v},t+2}(c)$ is non-increasing. So $J_{\boldsymbol{v},t+2}(c)$ is concave. Due to Lemma A.2 we finish our proof. □

Now let us prove Lemma 2.13.

*Proof.* Proof of Lemma 2.13. Let us start from the second calendar described in Lemma 2.12. We will show that by finitely applying Corollary A.10.1 and Lemma A.11 we can obtain the first calendar described in Lemma 2.12, which earns a greater revenue.

First deduct from Corollary A.10.1 that the randomized policy is non-increasing. Then by Lemma A.11 we can change the last pair of probabilities with non-increasing order, i.e. $t = \sup_{\tau\in[T-1]} \{v_\tau < 1, v_{\tau+1} > 0, v_\tau \geq v_{\tau+1}\}$, from $(v_t, v_{t+1})$ to $(v_t+\delta, v_{t+1}-\delta)$ with $\delta = \max\{1-v_t, v_{t+1}\} \geq 0$ to achieve a larger revenue. After such change, either $v_t = 1$ or $v_{t+1} = 0$ so we know that the total number of fractional (strictly greater than 0 and smaller than 1) periods have decreased by 1. By repeatedly applying Corollary A.10.1 and Lemma A.11 we will obtain the optimal randomized policy as described. □

## A.13 Necessity of Two Prices in Single-Item Pricing Problem: Proof of Proposition 2.14

*Proof.* Proof of Proposition 2.14. Construct the following instance. Let there be $T$ periods and $b = 1$ unit of initial inventory. Let $T$ be some large number. There are two prices $p_1 = 3, p_2 = 2$. Let random demand be Bernoulli random variables. The purchase probability of offering the higher price $p_1$ is $q_1 = 1/2T$; of offering the lower price $p_2$ is $q_2 = 1$.

The LP upper bound suggests a total of $\frac{4T-3}{2T-1} \approx 2$ units of revenue. Only offering the higher price $p_1$ suggests a total of $3 \cdot \mathbb{E}[\min\{\mathrm{Bin}(T, 1/2T), 1\}] = 3 \cdot (1-1/2T)^T \approx 1 \cdot (1-e^{-1}) \approx 1.180$ units of revenue. Only offering the lower price $p_2$ suggests a total of 1 unit of revenue. $\max\{1.180, 1\} < 2 \cdot (1 - 1/e) \approx 1.264$. So the expected revenue of either single-price policy is strictly smaller than expression (2.9), which finishes the tightness proof. $\square$

## A.14 Proof of Theorem 2.15

We prove Theorem 2.15 in the general [0,1]-demand setting; the {0,1}-demand setting is a special case of it.

*Proof.* Proof of Theorem 2.15. Let $j_t, \forall t \in [T]$ denote the prices selected from expression (2.19). Denote $B_t$ to be the remaining inventory at the *end* of time $t$, with $B_0 = 0$. Denote $Q_t$ to be the inventory at time $t$ that customer *would have* demanded if price $p_{j_t}$ is offered in time period $t$, which can take any value in $[0, 1]$. Then $\min\{Q_t, B_{t-1}\}$ is the actual sales at time $t$.

We let $\mathbb{E}[\mathsf{Rev}]$ denote the expected revenue earned by the deterministic calendar suggested in Algorithm 6, which can be written as:

$$\begin{aligned}
\mathbb{E}[\mathsf{Rev}] &= \sum_{t=1}^{T} p_{j_t} \mathbb{E}[\min\{Q_t, B_{t-1}\}] \\
&= \sum_{t=1}^{T} (p_{j_t} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b}) \mathbb{E}[\min\{Q_t, B_{t-1}\}] + \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b} \sum_{t=1}^{T} \mathbb{E}[\min\{Q_t, B_{t-1}\}] \\
&= \sum_{t=1}^{T} (p_{j_t} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b}) \mathbb{E}[\min\{Q_t, B_{t-1}\}] + \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b} (b - \mathbb{E}[B_T])
\end{aligned}$$

Now, note that the inventory level $B_t$ is decreasing in $t$ and that each $(p_{a_t} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b})$ term is non-negative, so we can bound each $B_{t-1}$ from below by $B_T$. The following can then be derived:

$$\begin{aligned}
\mathbb{E}[\mathsf{Rev}] = \sum_{t=1}^{T} (p_{j_t} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b}) \mathbb{E}[\min\{Q_t, B_{t-1}\} \cdot \frac{Q_t B_{t-1}}{\min\{Q_t, B_{t-1}\} \max\{Q_t, B_{t-1}\}}] \\
+ \frac{\mathsf{OPT}_{\mathrm{LP}}}{2} (1 - \frac{\mathbb{E}[B_T]}{b})
\end{aligned}$$

$$
\begin{aligned}
&\geq \sum_{t=1}^{T}(p_{j_t} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b})\mathbb{E}[\frac{Q_t B_{t-1}}{b}] + \frac{\mathsf{OPT}_{\mathrm{LP}}}{2}(1 - \frac{\mathbb{E}[B_T]}{b}) \\
&\geq \frac{\mathbb{E}[B_T]}{b}\sum_{t=1}^{T}(p_{j_t} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b})q_{tj_t} + \frac{\mathsf{OPT}_{\mathrm{LP}}}{2}(1 - \frac{\mathbb{E}[B_T]}{b}) \\
&\geq \frac{\mathbb{E}[B_T]}{b}\sum_{t=1}^{T}\sum_{j=1}^{J}(p_j - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b})q_{tj}x^*_{tj} + \frac{\mathsf{OPT}_{\mathrm{LP}}}{2}(1 - \frac{\mathbb{E}[B_T]}{b}) \\
&\geq \frac{\mathbb{E}[B_T]}{b}\Big(\mathsf{OPT}_{\mathrm{LP}} - \frac{\mathsf{OPT}_{\mathrm{LP}}}{2b}\cdot b\Big) + \frac{\mathsf{OPT}_{\mathrm{LP}}}{2}(1 - \frac{\mathbb{E}[B_T]}{b}) \\
&= \frac{\mathsf{OPT}_{\mathrm{LP}}}{2}
\end{aligned}
$$

where the first inequality is because $\forall t \in [T], Q_t \leq b$ and $B_t \leq b$; second inequality is because $B_t$ is decreasing in $t$; third inequality follows from the optimality of $j_t$ in expression (2.19); and the fourth inequality follows from constraint (2.15) in DLP-N. □

# Appendix B

# Appendix to Chapter 3

## B.1 Proof of Theorem 3.1.

*Proof.* Proof of Theorem 3.1 under $\mathsf{NRM}$ setup. For any problem instance $\mathcal{I} = (T, \boldsymbol{B}, K, d, n, \boldsymbol{p}, A; \boldsymbol{Q})$. Any policy $\pi \in \Pi[\mathsf{\Lambda} - 2]$ only selects no more than $(\mathsf{\Lambda} - 1)$ many price vectors. For any $k \in [K]$, let $\tau_k$ be the total number of periods that price $\boldsymbol{p}_k$ is offered during the selling horizon, under policy $\pi$. Notice that $\tau_k$ is a random variable, i.e., it is determined by the random trajectory of demand realization and action selection. Since $\pi \in \Pi[\mathsf{\Lambda} - 2]$, we know that for any realization of the random vector $(\tau_1, \ldots, \tau_K)$, it has at most $\Lambda - 1$ non-zero components.

Now denote $Y_{j,k}$ as the random amount of product $j$ sold, during the $\tau_k$ periods that price vector $k$ is offered. Here $\tau_k$ is a random amount, so we cannot directly use Hoeffding inequality to connect $Y_{j,k}$ with $\tau_k q_{j,k}$. But we can adapt the trick from Chapter 1.3 of Slivkins (2019). Suppose there was a tape of length $T$ for each product $j \in [n]$ and each price vector $k \in [K]$, with each cell independently sampled from the distribution of $Q_{j,k}$. This tape serves as a coupling of the random demand: in each period $t$ if price vector $k$ is offered, we simply generate a demand of each product $j \in [n]$ from the $t^{\text{th}}$ cell of the tape associated with product $j$ and price vector $k$. Let $Y_{j,k}(t)$ denote the random amount of product $j$ sold, during the first $t$ periods that the price vector $k$ is offered. Now we can use Hoeffding inequality on each reward tape:

$$\forall k, \forall j, \forall t, \Pr\left(|Y_{j,k}(t) - tq_{j,k}| \leq \sqrt{3t \log T}\right) \geq 1 - \frac{2}{T^6}.$$

Denote the following "clean event" $E$:

$$\forall k, \forall j, \forall t, |Y_{j,k}(t) - tq_{j,k}| \leq \sqrt{3t \log T}.$$

Using a union bound we have:

$$\Pr\left(\forall k, \forall j, \forall t, |Y_{j,k} - tq_{j,k}| \leq \sqrt{3t \log T}\right) \geq 1 - \frac{2}{T^3}$$

because $K, n$ are both less than $T$, and each arm cannot be pulled longer than $T$ periods. The happening of such event implies that

$$\forall j, \forall k, |Y_{j,k} - \tau_k q_{j,k}| \leq \sqrt{3\tau_k \log T},$$

i.e., the realized demands are close to the expected demands, suggesting that we can use LP to approximately bound the revenue generated by any policy $\pi \in \Pi[\Lambda - 2]$.

Specifically, we use the following arguments. If we focus on the usage of any price vector indexed by $k \in [K]$, the total revenue is $\sum_{j \in [n]} Y_{j,k} p_{j,k}$. Thus, conditional on $E$, the total revenue generated by policy $\pi$ during the entire horizon can be upper bounded by

$$\begin{aligned}
\sum_{k \in [K]} \sum_{j \in [n]} Y_{j,k} p_{j,k} &\leq \sum_{k \in [K]} \sum_{j \in [n]} (q_{j,k}\tau_k + \sqrt{3T \log T}) p_{j,k} \\
&\leq (\sum_{k \in [K]} \sum_{j \in [n]} q_{j,k}\tau_k p_{j,k}) + ndp_{\mathsf{max}}\sqrt{3T \log T},
\end{aligned}$$

where the last inequality follows from $\pi \in \Pi[\Lambda - 2]$ and $\Lambda \leq d + 1$. On the other hand, the consumption of resource $i$ must not violate the resource constraints.

$$\sum_{k \in [K]} \sum_{j \in [n]} Y_{j,k} a_{ij} \leq B_i.$$

Lower bounding $Y_{j,k}$ by $q_{j,k}\tau_k - \sqrt{3T\log T}$ we have

$$\sum_{k\in[K]}\sum_{j\in[n]} a_{ij} q_{j,k}\tau_k \leq B_i + \sum_{k\in[K]}\sum_{j\in[n]} \sqrt{3T\log T} \leq B_i + nd\sqrt{3T\log T}.$$

These suggest that conditional on $E$, any policy $\pi \in \Pi[\Lambda - 2]$ always satisfies the following constraints:

$$\begin{aligned}
\sum_{k\in[K]}\sum_{j\in[n]} a_{ij}\ q_{j,k}\ \tau_k &\leq B_i + n^2\sqrt{3T\log T} && \forall\ i \in [d]\\
\sum_{k\in[K]} \tau_k &\leq T &&\\
\tau_l &= 0 && \exists l \in [K]\\
\tau_k &\geq 0 && \forall\ k \in [K],
\end{aligned}$$

with its total revenue upper bounded by

$$\sum_{k\in[K]}\sum_{j\in[n]} p_{j,k} q_{j,k}\tau_k + nd p_{\mathsf{max}}\sqrt{3T\log T}.$$

Recall that the optimal solution to the DLP uses $\Lambda$ many price vectors. We have used $\mathcal{Z}(\boldsymbol{x}^*)$ to denote the set of price indices that are non-zero in the optimal solution to the DLP. For any $l \in \mathcal{Z}(\boldsymbol{x}^*)$, define a family of linear programs parameterized by $l$,

$$\begin{aligned}
(\textbf{DLP}_\textbf{l}) \quad \mathsf{J}_l^{\mathsf{DLP}} = \max_{(x_k)_{k\in[K]}} & \sum_{k\in[K]}\sum_{j\in[n]} p_{j,k}\ q_{j,k}\ x_k &&\\
\text{s.t.} & \sum_{k\in[K]}\sum_{j\in[n]} a_{ij}\ q_{j,k}\ x_k \leq B_i && \forall\ i \in [d]\\
& \sum_{k\in[K]} x_k \leq T &&\\
& x_l = 0 &&\\
& x_k \geq 0 && \forall\ k \in [K],
\end{aligned}$$

such that this family of linear programs use no more than $(\Lambda - 1)$ non-zero variables. Now construct the following LP's, which we denote as "perturbed LP's":

$$
\begin{aligned}
\textbf{(DLP}_\textsf{l}\textbf{ Perturbed)} \quad \mathsf{J}_l^{\mathsf{Perturbed}} = \max_{(x_k)_{k\in[K]}} & \sum_{k\in[K]} \sum_{j\in[n]} p_{j,k}\ q_{j,k}\ x_k \\
\text{s.t.} & \sum_{k\in[K]} \sum_{j\in[n]} a_{ij}\ q_{j,k}\ x_k \leq B_i + nd\sqrt{3T\log T} \quad \forall\ i \in [d] \\
& \sum_{k\in[K]} x_k \leq T \\
& x_l = 0 \\
& x_k \geq 0 \qquad \forall\ k \in [K],
\end{aligned}
$$

Since from each solution $\boldsymbol{x}^*$ of the Perturbed $\mathsf{DLP_l}$, we can find a corresponding discounted solution $\boldsymbol{x}^*/(1 + (nd\sqrt{3T\log T})/B_{\mathsf{min}})$ that is feasible to the $\mathsf{DLP_l}$. This suggests that $\mathsf{J}_l^{\mathsf{Perturbed}} \leq \mathsf{J}_l^{\mathsf{DLP}} \cdot (1 + (nd\sqrt{3T\log T})/B_{\mathsf{min}})$, because $\mathsf{DLP_l}$ is a maximization problem.

Next we define an instance-dependent gap between the maximum objective value of $\mathsf{DLP_l}$, and the objective value of DLP. Let $\Delta = (\mathsf{J}^{\mathsf{DLP}} - \max_{l\in\mathcal{Z}(\boldsymbol{x}^*)} \mathsf{J}_l^{\mathsf{DLP}})/\mathsf{J}^{\mathsf{DLP-G}}$ be such an instance-dependent gap normalized by $\mathsf{J}^{\mathsf{DLP}}$. Importantly, while $\mathsf{J}^{\mathsf{DLP}}$ scales linearly with $T$ and $\boldsymbol{B}$, $\Delta$ remain fixed as $T$ and $\boldsymbol{B}$ grow.

Putting everything together, we obtain the following result: conditional on event $E$ that happens with probability at least $1 - \frac{2}{T^3}$, for any policy $\pi \in \Pi[\Lambda - 2]$ and any possible realization of $(\tau_1, \ldots, \tau_K)$, the total revenue collected during the selling horizon is upper bounded by

$$
\begin{aligned}
& \max_{l\in\mathcal{Z}(\boldsymbol{x}^*)} \mathsf{J}_l^{\mathsf{Perturbed}} + nd\sqrt{3T\log T} p_{\mathsf{max}} \\
\leq & \max_{l\in\mathcal{Z}(\boldsymbol{x}^*)} \mathsf{J}_l^{\mathsf{DLP}} \cdot (1 + \frac{nd\sqrt{3T\log T}}{B_{\mathsf{min}}}) + nd\sqrt{3T\log T} p_{\mathsf{max}} \\
\leq & (\mathsf{J}^{\mathsf{DLP}} - \Delta \mathsf{J}^{\mathsf{DLP}}) \cdot (1 + \frac{nd\sqrt{3T\log T}}{B_{\mathsf{min}}}) + nd\sqrt{3T\log T} p_{\mathsf{max}},
\end{aligned}
$$

which suggests that

$$\mathsf{Rev}(\pi) \leq (\mathsf{J}^{\mathsf{DLP}} - \Delta\mathsf{J}^{\mathsf{DLP}}) \cdot (1 + \frac{nd\sqrt{3T\log T}}{B_{\mathsf{min}}}) + nd\sqrt{3T\log T}p_{\mathsf{max}} + 1 = \mathsf{J}^{\mathsf{DLP}} - \Omega(T).$$

$\square$

*Proof.* Proof of Theorem 3.1 under $\mathsf{SP}$ Setup. In this setup, for any problem instance $\mathcal{I} = (T, \boldsymbol{B}, K, d; \boldsymbol{C}, \boldsymbol{R})$, we consider an arbitrary policy $\pi \in \Pi[\mathsf{\Lambda} - 2]$ that only selects no more than $(\mathsf{\Lambda}-1)$ many arms. For any $k \in [K]$, let $\tau_k$ be the total number of periods that action $k$ is offered during the selling horizon, under policy $\pi$. Notice that $\tau_k$ is a random variable, i.e., it is determined by the random trajectory of reward and cost realization and action selection. Since $\pi \in \Pi[\mathsf{\Lambda} - 2]$, we know that for any realization of the random vector $(\tau_1, \ldots, \tau_K)$, it has at most $\Lambda - 1$ non-zero components.

Now denote $C^s_{i,k}$ as the random amount of resource $i$ consumed, during the $\tau_k$ periods that arm $k$ is pulled; denote $R^s_k$ as the random amount of rewards generated, during the $\tau_k$ periods that arm $k$ is pulled. Here $\tau_k$ is a random amount, so we cannot directly use Hoeffding inequality. But again we can use the "reward tape" trick demonstrated in the previous proof under the $\mathsf{NRM}$ setup. Let $C^s_{i,k}(t)$ denote the random amount of resource $i$ consumed, during the first $t$ periods that the arm $k$ is pulled; let $R^s_k(t)$ denote the random amount of rewards generated, during the first $t$ periods that arm $k$ is pulled. Now we can use Hoeffding inequality on each reward tape:

$$\forall k, \forall i, \forall t, \Pr\left(\left|C^s_{i,k}(t) - tc_{i,a_l}\right| \leq C_{\mathsf{max}}\sqrt{3t\log T}\right) \geq 1 - \frac{2}{T^6};$$
$$\forall k, \forall t, \Pr\left(\left|R^s_k(t) - tr_k\right| \leq R_{\mathsf{max}}\sqrt{3t\log T}\right) \geq 1 - \frac{2}{T^6}.$$

Denote the following event $E$:

$$\forall k, \forall i, \forall t, \left|C^s_{i,k}(t) - tc_{i,k}\right| \leq C_{\mathsf{max}}\sqrt{3t\log T};$$
$$\forall k, \forall t, \left|R^s_k(t) - tr_k\right| \leq R_{\mathsf{max}}\sqrt{3t\log T}.$$

Using a union bound we have:

$$\Pr(E) \geq 1 - \frac{4}{T^3}$$

because $K, d$ are both less than $T$, and each arm cannot be pulled longer than $T$ periods. The happening of such event implies that

$$\forall k, \forall i, \left|C^s_{i,k} - \tau_k c_{i,k}\right| \leq C_{\mathsf{max}}\sqrt{3\tau_k \log T},$$
$$\forall k, |R^s_k - \tau_k r_k| \leq R_{\mathsf{max}}\sqrt{3\tau_k \log T},$$

i.e., the realized rewards and costs are close to the expected values, suggesting that we can use LP to approximately bound the total reward collected by any policy $\pi \in \Pi[\Lambda - 2]$.

Specifically, we use the following arguments. Conditional on $E$, for any realization of $(\tau_1, \ldots, \tau_K)$, the total reward collected by policy $\pi$ during the entire horizon can be upper bounded by

$$\begin{aligned}\sum_{k\in[K]} R^s_k &\leq \sum_{k\in[K]} (r_k\tau_k + R_{\mathsf{max}}\sqrt{3T\log T}) \\ &\leq (\sum_{k\in[K]} r_k\tau_k) + dR_{\mathsf{max}}\sqrt{3T\log T},\end{aligned}$$

where the last inequality follows from $\pi \in \Pi[\Lambda - 2]$ and $\Lambda \leq d + 1$. On the other hand, the consumption of each resource $i$ must not violate the resource constraints.

$$\sum_{k\in[K]} C^s_{i,k} \leq B_i.$$

Lower bounding $C^s_{i,k}$ by $(\tau_k c_{i,k} - C_{\mathsf{max}}\sqrt{3T\log T})$ we have

$$\sum_{k\in[K]} \tau_k c_{i,k} \leq B_i + \sum_{k\in[K]} C_{\mathsf{max}}\sqrt{3T\log T} \leq B_i + dC_{\mathsf{max}}\sqrt{3T\log T}.$$

These suggest that conditional on $E$, any policy $\pi \in \Pi[\Lambda - 2]$ always satisfies the following constraints:

$$
\begin{aligned}
\sum_{k\in[K]} r_{i,k}\ \tau_k &\leq B_i + dC_{\mathsf{max}}\sqrt{3T\log T} && \forall\ i\in[d]\\
\sum_{k\in[K]} \tau_k &\leq T\\
\tau_l &= 0 && \exists l\in[K]\\
\tau_k &\geq 0 && \forall\ k\in[K],
\end{aligned}
$$

with its total collected reward is upper bounded by

$$\sum_{k\in[K]} r_{i,k}\tau_k + dR_{\mathsf{max}}\sqrt{3T\log T}.$$

Recall that the optimal solution to the DLP-G uses $\Lambda$ many prices. We have used $\mathcal{Z}(\boldsymbol{x}^*)$ to denote the set of price indices that are non-zero in the optimal solution to the DLP-G. For any $l\in\mathcal{Z}(\boldsymbol{x}^*)$, define a family of linear programs parameterized by $l$,

$$
\begin{aligned}
(\mathbf{DLP_l - G})\quad \mathsf{J}_l^{\mathsf{DLP-G}} = \max_{(x_k)_{k\in[K]}} & \sum_{k\in[K]} r_k\ x_k\\
\text{s.t.}\ & \sum_{k\in[K]} c_{i,k}\ x_k \leq B_i && \forall\ i\in[d]\\
& \sum_{k\in[K]} x_k \leq T\\
& x_l = 0\\
& x_k \geq 0 && \forall\ k\in[K].
\end{aligned}
$$

such that this family of linear programs use no more than $(\Lambda-1)$ non-zero variables. Now construct the following LP's, which we denote as "perturbed LP's":

$$(\mathbf{DLP_l - G\ Perturbed})\quad \mathsf{J}_l^{\mathsf{Perturbed-G}} = \max_{(x_k)_{k\in[K]}} \sum_{k\in[K]} r_k\ x_k$$

$$
\begin{aligned}
\text{s.t.} \sum_{k\in[K]} c_{i,k}\ x_k &\le B_i + dC_{\mathsf{max}}\sqrt{3T\log T} \quad && \forall\ i\in[d]\\
\sum_{k\in[K]} x_k &\le T \\
x_l &= 0 \\
x_k &\ge 0 && \forall\ k\in[K].
\end{aligned}
$$

Since from each solution $\boldsymbol{x}^*$ of the $\mathsf{Perturbed\ DLP_I - G}$, we can find a corresponding discounted solution $\boldsymbol{x}^*/(1+\frac{dC_{\mathsf{max}}\sqrt{3T\log T}}{B_{\mathsf{min}}})$ that is feasible to the $\mathsf{DLP_I - G}$. This suggests that $\mathsf{J}_l^{\mathsf{Perturbed}} \le \mathsf{J}_l^{\mathsf{DLP-G}} \cdot (1+\frac{dC_{\mathsf{max}}\sqrt{3T\log T}}{B_{\mathsf{min}}})$, because DLP-G is a maximization problem.

Next we define an instance-dependent gap between the maximum objective value of $\mathsf{DLP_I}$, and the objective value of DLP-G. Let $\Delta = (\mathsf{J}^{\mathsf{DLP-G}} - \max_{l\in\mathcal{Z}(\boldsymbol{x}^*)} \mathsf{J}_l^{\mathsf{DLP-G}})/\mathsf{J}^{\mathsf{DLP-G}}$ be such an instance-dependent gap normalized by $\mathsf{J}^{\mathsf{DLP-G}}$. Importantly, while $\mathsf{J}^{\mathsf{DLP}}$ scales linearly with $T$ and $\boldsymbol{B}$, $\Delta$ remain fixed as $T$ and $\boldsymbol{B}$ grow.

Putting everything together, we obtain the following result: conditional on event $E$ that happens with probability at least $1-\frac{4}{T^3}$, for any policy $\pi\in\Pi[\Lambda-2]$ and any possible realization of $(\tau_1,\ldots,\tau_K)$, the total collected reward is upper bounded by

$$
\begin{aligned}
&\max_{l\in\mathcal{Z}(\boldsymbol{x}^*)} \mathsf{J}_l^{\mathsf{Perturbed-G}} + dR_{\mathsf{max}}\sqrt{3T\log T}\\
&\le \max_{l\in\mathcal{Z}(\boldsymbol{x}^*)} \mathsf{J}_l^{\mathsf{DLP-G}} \cdot (1+\frac{dC_{\mathsf{max}}\sqrt{3T\log T}}{B_{\mathsf{min}}}) + dR_{\mathsf{max}}\sqrt{3T\log T}\\
&\le (\mathsf{J}^{\mathsf{DLP-G}} - \Delta\mathsf{J}^{\mathsf{DLP-G}})\cdot(1+\frac{dC_{\mathsf{max}}\sqrt{3T\log T}}{B_{\mathsf{min}}}) + dR_{\mathsf{max}}\sqrt{3T\log T},
\end{aligned}
$$

which suggests that

$$
\mathsf{Rev}(\pi) \le (\mathsf{J}^{\mathsf{DLP-G}} - \Delta\mathsf{J}^{\mathsf{DLP-G}})\cdot(1+\frac{dC_{\mathsf{max}}\sqrt{3T\log T}}{B_{\mathsf{min}}}) + dR_{\mathsf{max}}\sqrt{3T\log T} + 1 = \mathsf{J}^{\mathsf{DLP}} - \Omega(T).
$$

$\square$

## B.2 Proof of Theorem 3.2

In this section we prove Theorem 3.2 under two setups. For better exposition we prove it two times under the two setups.

*Proof.* Proof of Theorem 3.2 under $\mathsf{NRM}$ Setup. Let $\pi$ be any policy suggested in Algorithm 7. Let $\boldsymbol{x}^*$ be the associated optimal solution. We prove Theorem 3.2 by comparing the expected revenue earned by Algorithm 7 against a virtual policy $\pi^{\mathsf{v}}$. This virtual policy $\pi^{\mathsf{v}}$ mimics Algorithm 7 in steps 1–3. But in step 4, it sets the price vector to be $\boldsymbol{p}_{\sigma(1)}$ for the first $\gamma \cdot x^*_{\sigma(1)}$ periods, then $\boldsymbol{p}_{\sigma(2)}$ for the next $\gamma \cdot x^*_{\sigma(2)}$ periods, ..., $\boldsymbol{p}_{\sigma(\Lambda)}$ for the next $\gamma \cdot x^*_{\sigma(\Lambda)}$ periods, and finally $\boldsymbol{p}_\infty$ for the last $(T - \gamma \cdot \sum_{l=1}^{\Lambda} x^*_{\sigma(l)})$ periods. Here $\boldsymbol{p}_\infty$ is a shut-off price, under which $Q_j(\boldsymbol{p}_\infty) = 0, \forall j \in [n]$.

Policy $\pi^{\mathsf{v}}$ is virtual because it requires a shut-off price $\boldsymbol{p}_\infty$ that may or may not be available. Moreover, it requires $\Lambda$ many price changes, which is more than $(\Lambda - 1)$ many changes as suggested in Algorithm 7.

Policy $\pi^{\mathsf{v}}$ serves to bridge our analysis. It breaks our Theorem 3.2 into two inequalities that we will prove separately.

$$\mathsf{Rev}(\pi) \geq \mathsf{Rev}(\pi^{\mathsf{v}}) \geq \left(1 - 2a_{\mathsf{max}}\sqrt{\frac{nT\log T}{B^2_{\mathsf{min}}}} - \frac{d}{T^2}\right) \cdot \mathsf{J}^{\mathsf{DLP}} \tag{B.1}$$

For any policy $\pi$ as defined in Algorithm 7 and its associated virtual policy $\pi^{\mathsf{v}}$, they both solve the same DLP and have the same optimal solution. To prove the first inequality, note that both $\pi$ and $\pi^{\mathsf{v}}$ commit to the same prices in the first $\tilde{T} := \gamma \cdot \sum_{l=1}^{\Lambda} x^*_{\sigma(l)}$ time periods, and earns the same revenue following each trajectory of random demand. At the end of period $\tilde{T}$, policy $\pi$ still commits to $\boldsymbol{p}_{\sigma(\Lambda)}$, while policy $\pi^{\mathsf{v}}$ makes one change and sets $\boldsymbol{p}_\infty$. At the end of period $\tilde{T}$, if the selling horizon has ended due to inventory stock-outs, then either policy earns zero revenue, so $\pi$ and $\pi^{\mathsf{v}}$ makes no difference. If the selling horizon has not ended and there is remaining inventory for any resource, then policy $\pi$ earns non-negative revenue, while $\pi^{\mathsf{v}}$ earns zero by setup a shut-off price. Following each trajectory of random demand, policy $\pi$ earns more revenue than $\pi^{\mathsf{v}}$. As a result, $\mathsf{Rev}(\pi) \geq \mathsf{Rev}(\pi^{\mathsf{v}})$.

To prove the second inequality, we introduce the following notations. Let $\mathbb{1}_{t,k}, \forall k \in [K], t \in [T]$ be an indicator of whether or not policy $\pi^{\mathsf{v}}$ offers price $\boldsymbol{p}_k$ in period $t$.

$$\mathbb{1}_{t,k} = \begin{cases} 1, & \text{if } k = \sigma(1), t \leq \gamma x^*_{\sigma(1)}; \\ 1, & \text{if } \exists 1 < l_0 \leq \Lambda, s.t. \sigma(l_0) = k, \gamma \sum_{l=1}^{l_0-1} x^*_{\sigma(l)} < t \leq \gamma \sum_{l=1}^{l_0} x^*_{\sigma(l)}; \\ 0, & \text{otherwise} \end{cases}$$

$\mathbb{1}_{t,k}$ is deterministic once policy $\pi^{\mathsf{v}}$ is determined.

Under policy $\pi^{\mathsf{v}}$, following each trajectory of random demand, we define the length of the effective selling horizon $\tau$ as a function of a stopping time $t_0$:

$$\tau = \tilde{T} \wedge \min \left\{ t_0 - 1 \,\middle|\, \exists i, s.t. \ \sum_{t=1}^{t_0} \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} Q_{j,k} \cdot a_{ij} > B_i \right\}$$

The effective selling horizon is the minimum between (i) last period before the cumulative demand of any resource exceeds its initial inventory, and (ii) the last period before policy $\pi^{\mathsf{v}}$ switches to the shut-off price.

Let $D_{t,i}$ be the remaining inventory of resource $i$ at the end of period $t$. Under this notation, $D_{0,i} = B_i$. Note that $D_{t,i}$ are random variables, and during the effective selling horizon, inventory updates in the following fashion

$$\forall t \in [\tau], D_{t,i} = D_{t-1,i} - \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} Q_{j,k} \cdot a_{ij} \geq 0 \tag{B.2}$$

Now we calculate the expected revenue.

$$\begin{aligned} \mathsf{Rev}(\pi^{\mathsf{v}}) &\geq \mathbb{E}_{Q_{j,k}} \left[ \sum_{t=1}^{\tilde{T}} \mathbb{1}_{\{\forall i, D_{t-1,i} \geq n a_{\mathsf{max}}\}} \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} p_{j,k} Q_{j,k} \right] \\ &= \sum_{t=1}^{\tilde{T}} \Pr(\forall i, D_{t-1,i} \geq n a_{\mathsf{max}}) \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} p_{j,k} q_{j,k} \end{aligned}$$

$$\geq \Pr(\forall i, D_{\tilde{T},i} \geq na_{\mathsf{max}}) \sum_{t=1}^{\tilde{T}} \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} p_{j,k} q_{j,k} \tag{B.3}$$

We explain the inequalities. The first inequality is because we only focus on the revenue earned if event $\{\forall i, D_{t-1,i} \geq na_{\mathsf{max}}\}$ happens, while ignoring the revenue earned if event $\{\forall i, D_{t-1,i} \geq na_{\mathsf{max}}\}$ does not happen; and when event $\{\forall i, D_{t-1,i} \geq na_{\mathsf{max}}\}$ does happen, the maximum amount of any resource $i$ demanded in one single period cannot exceed $na_{\mathsf{max}}$. The first equality is expanding the expectations, where we use the fact that $\mathbb{1}_{\{\forall i, D_{t-1,i} \geq na_{\mathsf{max}}\}}$ and $\sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} p_{j,k} Q_{j,k}$ are independent, because the indicator is a random event happening up to period $t-1$, while the summation term is a random amount happening in period $t$. The third inequality is due to (B.2), $D_{t,i}$ is decreasing in $t$, so $D_{t-1,i} \geq D_{\tilde{T},i}$.

In this block of inequalities (B.3), the summation term

$$\sum_{t=1}^{\tilde{T}} \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} p_{j,k} q_{j,k} = \sum_{l=1}^{\wedge} \sum_{k \in \{\sigma(l)=k\}} \gamma x_k^* \sum_{j \in [n]} p_{j,k} q_{j,k} = \gamma \mathsf{J}^{\mathsf{DLP}},$$

since the indicators $\mathbb{1}_{t,k}$ locate which $k$ counts into this summation. So the next thing we do is to lower bound the probability.

Note that

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{\tilde{T}} \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} Q_{j,k} \cdot a_{ij}\right] &= \sum_{t=1}^{\tilde{T}} \sum_{k \in [K]} \mathbb{1}_{t,k} \sum_{j \in [n]} q_{j,k} \cdot a_{ij} \\
&= \sum_{l=1}^{\wedge} \sum_{k \in \{\sigma(l)=k\}} \gamma x_k^* \sum_{j \in [n]} q_{j,k} a_{ij} \\
&\leq \gamma B_i \\
&< B_i - na_{\mathsf{max}}
\end{aligned}$$

where the last (strict) inequality is because we plug in $\gamma = 1 - 2a_{\mathsf{max}} \sqrt{\frac{nT \log T}{B_{\mathsf{min}}^2}}$.

This above inequality suggests that for any $i \in [d]$, the expected cumulative demand generated up till period $\tilde{T}$ is strictly less than $B_i - na_{\mathsf{max}}$. So we can use concentration inequalities.

$$
\begin{aligned}
&\Pr(\forall i, D_{\tilde{T},i} \geq n a_{\mathsf{max}}) \\
&= 1 - \Pr\left(\exists i, s.t. \sum_{t=1}^{\tilde{T}} \sum_{k\in[K]} \mathbb{1}_{t,k} \sum_{j\in[n]} Q_{j,k}\cdot a_{ij} \geq B_i - n a_{\mathsf{max}}\right) \\
&\geq 1 - \sum_{i\in[d]} \Pr\left(\sum_{t=1}^{\tilde{T}} \sum_{k\in[K]} \mathbb{1}_{t,k} \sum_{j\in[n]} Q_{j,k}\cdot a_{ij} - \gamma B_i \geq (1-\gamma)B_i - n a_{\mathsf{max}}\right) \\
&\geq 1 - \sum_{i\in[d]} \exp\left(-\frac{2((1-\gamma)B_i - n a_{\mathsf{max}})^2}{a_{\mathsf{max}}^2 nT}\right) \\
&\geq 1 - d\exp\left(-\frac{2((1-\gamma)B_{\mathsf{min}} - n a_{\mathsf{max}})^2}{a_{\mathsf{max}}^2 nT}\right) \\
&\geq 1 - \frac{d}{T^2}
\end{aligned} \tag{B.4}
$$

where the first inequality is due to union bound; the second inequality is due to Hoeffding inequality, $Q_{j,k}a_{ij}$ is bounded by $a_{\mathsf{max}}$, and there are no more than $n \cdot T$ such terms; the third inequality is because we lower bound each $B_i$ by $B_{\mathsf{min}}$; the last inequality is when we plug in $1-\gamma = 2a_{\mathsf{max}}\sqrt{\frac{nT\log T}{B_{\mathsf{min}}^2}}$, and we know that $T > n$.

Putting (B.4) into (B.3) we finish the proof. Note that Theorem 3.2 follows from (B.1) and $\mathsf{J}^{\mathsf{DLP}} \leq p_{\mathsf{max}} B_{\mathsf{min}}$. □

*Proof.* Proof of Theorem 3.2 under SP Setup. Let $\pi$ be any policy suggested in Algorithm 7 under SP Setup. Let $\boldsymbol{x}^*$ be the associated optimal solution. We prove Theorem 3.2 by comparing the expected revenue earned by Algorithm 7 against a virtual policy $\pi^{\mathsf{v}}$. This virtual policy $\pi^{\mathsf{v}}$ mimics Algorithm 7 in steps 1–3. But in step 4, it pulls arm $\sigma(1)$ for the first $\gamma \cdot x^*_{\sigma(1)}$ periods, then arm $\sigma(2)$ for the next $\gamma \cdot x^*_{\sigma(2)}$ periods, ..., $\sigma(\Lambda)$ for the next $\gamma \cdot x^*_{\sigma(\Lambda)}$ periods, and finally halts for the last $(T - \gamma \cdot \sum_{l=1}^{\Lambda} x^*_{\sigma(l)})$ periods pulling no arms. Such a halting notion was introduced in Badanidiyuru et al. (2013).

Policy $\pi^{\mathsf{v}}$ serves to bridge our analysis. It breaks our Theorem 3.2 into two inequalities that we will prove separately.

$$\mathsf{Rev}(\pi) \geq \mathsf{Rev}(\pi^{\mathsf{v}}) \geq \left(1 - \frac{2C_{\mathsf{max}}}{B_{\mathsf{min}}}\sqrt{T\log T} - \frac{d}{T^2}\right) \cdot \mathsf{J}^{\mathsf{DLP-G}} \tag{B.5}$$

For any policy $\pi$ as defined in Algorithm 7 and its associated virtual policy $\pi^{\mathsf{v}}$, they both solve the same DLP-G and have the same optimal solution. To prove the first inequality, note that both $\pi$ and $\pi^{\mathsf{v}}$ pull the same arm in the first $\tilde{T} := \gamma \cdot \sum_{l=1}^{\Lambda} x^*_{\sigma(l)}$ time periods, and earns the same revenue following each trajectory of random demand. At the end of period $\tilde{T}$, policy $\pi$ still pulls arm $\sigma(\Lambda)$, while policy $\pi^{\mathsf{v}}$ halts. At the end of period $\tilde{T}$, if the selling horizon has ended due to inventory stock-outs, then both policies earn zero reward, so $\pi$ and $\pi^{\mathsf{v}}$ make no difference. If the selling horizon has not ended and there is remaining inventory for some resource, then policy $\pi$ earns non-negative reward, while $\pi^{\mathsf{v}}$ halts and earns zero. Following each trajectory of random demand, policy $\pi$ earns more revenue than $\pi^{\mathsf{v}}$. As a result, $\mathsf{Rev}(\pi) \geq \mathsf{Rev}(\pi^{\mathsf{v}})$.

To prove the second inequality, we introduce the following notations. Let $\mathbb{1}_{t,k}, \forall k \in [K], t \in [T]$ be an indicator of whether or not policy $\pi^{\mathsf{v}}$ pulls arm $k$ in period $t$.

$$\mathbb{1}_{t,k} = \begin{cases} 1, & \text{if } k = \sigma(1), t \leq \gamma x^*_{\sigma(1)}; \\ 1, & \text{if } \exists 1 < l_0 \leq \Lambda, s.t. \sigma(l_0) = k, \gamma \displaystyle\sum_{l=1}^{l_0-1} x^*_{\sigma(l)} < t \leq \gamma \sum_{l=1}^{l_0} x^*_{\sigma(l)}; \\ 0, & \text{otherwise} \end{cases}$$

$\mathbb{1}_{t,k}$ is deterministic once policy $\pi^{\mathsf{v}}$ is determined.

Under policy $\pi^{\mathsf{v}}$, following each trajectory of random demand, we define the length of the effective selling horizon $\tau$ as a function of a stopping time $t_0$:

$$\tau = \tilde{T} \wedge \min\left\{ t_0 - 1 \,\middle|\, \exists i, s.t. \sum_{t=1}^{t_0} \sum_{k\in[K]} \mathbb{1}_{t,k} C_{i,k} > B_i \right\}$$

The effective selling horizon is the minimum between (i) last period before the cumulative demand of any resource exceeds its initial inventory, and (ii) the last period before policy $\pi^{\mathsf{v}}$ halts.

Let $D_{t,i}$ be the remaining inventory of resource $i$ at the end of period $t$. Under this notation, $D_{0,i} = B_i$. Note that $D_{t,i}$ are random variables, and during the effective selling horizon, inventory updates in the following fashion

$$\forall t \in [\tau], D_{t,i} = D_{t-1,i} - \sum_{k \in [K]} \mathbb{1}_{t,k} C_{i,k} \geq 0 \tag{B.6}$$

Now we calculate the expected revenue.

$$\begin{aligned}
\mathsf{Rev}(\pi^{\mathsf{v}}) &\geq \mathbb{E}_{Q_{j,k}} \left[ \sum_{t=1}^{\tilde{T}} \mathbb{1}_{\{\forall i, D_{t-1,i} \geq C_{\mathsf{max}}\}} \sum_{k \in [K]} \mathbb{1}_{t,k} R_k \right] \\
&= \sum_{t=1}^{\tilde{T}} \Pr(\forall i, D_{t-1,i} \geq C_{\mathsf{max}}) \sum_{k \in [K]} \mathbb{1}_{t,k} R_k \\
&\geq \Pr(\forall i, D_{\tilde{T},i} \geq C_{\mathsf{max}}) \sum_{t=1}^{\tilde{T}} \sum_{k \in [K]} \mathbb{1}_{t,k} R_k
\end{aligned} \tag{B.7}$$

We explain the inequalities. The first inequality is because we only focus on the revenue earned if event $\{\forall i, D_{t-1,i} \geq C_{\mathsf{max}}\}$ happens, while ignoring the revenue earned if event $\{\forall i, D_{t-1,i} \geq C_{\mathsf{max}}\}$ does not happen; and when event $\{\forall i, D_{t-1,i} \geq C_{\mathsf{max}}\}$ does happen, the maximum amount of any resource $i$ demanded in one single period cannot exceed $C_{\mathsf{max}}$. The first equality is expanding the expectations, where we use the fact that $\mathbb{1}_{\{\forall i, D_{t-1,i} \geq C_{\mathsf{max}}\}}$ and $\sum_{k \in [K]} \mathbb{1}_{t,k} R_k$ are independent, because the indicator is a random event happening up to period $t-1$, while the summation term is a random amount happening in period $t$. The third inequality is due to (B.6), $D_{t,i}$ is decreasing in $t$, so $D_{t-1,i} \geq D_{\tilde{T},i}$.

In this block of inequalities (B.7), the summation term

$$\sum_{t=1}^{\tilde{T}} \sum_{k \in [K]} \mathbb{1}_{t,k} R_k = \sum_{l=1}^{\Lambda} \sum_{k \in \{\sigma(l)=k\}} \gamma x_k^* R_k = \gamma \mathsf{J}^{\mathsf{DLP-G}},$$

since the indicators $\mathbb{1}_{t,k}$ locate which $k$ counts into this summation. So the next thing we do is to lower bound the probability term, $\Pr(\forall i, D_{\tilde{T},i} \geq C_{\mathsf{max}})$.

Note that for any $i \in [d]$,

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{\tilde{T}} \sum_{k\in[K]} \mathbb{1}_{t,k} C_{i,k}\right] &= \sum_{t=1}^{\tilde{T}} \sum_{k\in[K]} \mathbb{1}_{t,k} c_{i,k} \\
&= \sum_{l=1}^{\Lambda} \sum_{k\in\{\sigma(l)=k\}} \gamma x_k^* c_{i,k} \\
&\leq \gamma B_i \\
&< B_i - C_{\mathsf{max}}
\end{aligned}
$$

where the last inequality is because we plug in $\gamma = 1 - \frac{2C_{\mathsf{max}}}{B_{\mathsf{min}}}\sqrt{T\log T}$, and that $2\sqrt{T\log T} > 1$.

This above inequality suggests that for any $i \in [d]$, the expected cumulative demand generated up till period $\tilde{T}$ is strictly less than $B_i - C_{\mathsf{max}}$. So we can use concentration inequalities.

$$
\begin{aligned}
\Pr(\forall i, D_{\tilde{T},i} \geq C_{\mathsf{max}}) &= 1 - \Pr\left(\exists i, s.t. \sum_{t=1}^{\tilde{T}} \sum_{k\in[K]} \mathbb{1}_{t,k} C_{i,k} \geq B_i - C_{\mathsf{max}}\right) \\
&\geq 1 - \sum_{i\in[d]} \Pr\left(\sum_{t=1}^{\tilde{T}} \sum_{k\in[K]} \mathbb{1}_{t,k} C_{i,k} - \gamma B_i \geq (1-\gamma) B_i - C_{\mathsf{max}}\right) \\
&\geq 1 - \sum_{i\in[d]} \exp\left(-\frac{2((1-\gamma)B_i - C_{\mathsf{max}})^2}{C_{\mathsf{max}}^2 nT}\right) \\
&\geq 1 - d\exp\left(-\frac{2((1-\gamma)B_{\mathsf{min}} - C_{\mathsf{max}})^2}{C_{\mathsf{max}}^2 nT}\right) \\
&\geq 1 - \frac{d}{T^2}
\end{aligned}
\tag{B.8}
$$

where the first inequality is due to union bound; the second inequality is due to Hoeffding inequality, $C_{i,k}$ is bounded by $C_{\mathsf{max}}$; the third inequality is because we lower bound each $B_i$ by $B_{\mathsf{min}}$; the last inequality is when we plug in $1-\gamma = \frac{2C_{\mathsf{max}}}{B_{\mathsf{min}}}\sqrt{T\log T}$, and that $\sqrt{T\log T} > 1$.

Putting (B.8) into (B.7) we finish the proof. Note that Theorem 3.2 follows from (B.5), $\mathsf{J}^{\mathsf{DLP}} \leq R_{\mathsf{max}} T$, and $T/B_{\mathsf{min}} = O(1)$. $\square$

# Appendix C

# Appendix to Chapter 4

## C.1 Proof of Theorem 4.4

*Proof.* Proof of Theorem 4.4. We are going to show that, for any instance of arrival sequence $S$, we have $\frac{\mathsf{ALG}_{\mathsf{N4.3}}(S)}{\mathsf{OPT}(S)} \geq c_{N4.3}$. We lower bound $\mathsf{ALG}_{\mathsf{N4.3}}(S)$ and upper bound $\mathsf{OPT}(S)$ at the same time.

First of all, Greedy always accepts something. Denote the set of items accepted by Greedy as $G$. Denote $\mathsf{size}(G) = g$. If $G = [T]$ then Greedy is optimal. In this case

$$\frac{\mathsf{ALG}_{\mathsf{N4.3}}}{\mathsf{OPT}} \geq \Pr(\tau = 0) \cdot 1 + \Pr(\tau > 0) \cdot 0 \geq F_{N4.3}(0) = 1 - c_{N4.3} \geq c_{N4.3}.$$

If $G \subsetneqq [T]$, let $M = [T] \backslash G$ denote the set of items blocked by Greedy. Since Greedy always accepts an item as long as it can fill in, any item blocked by Greedy must exceed the remaining space of the knapsack, at the moment it is blocked. We also know that $G \cup M = [T]$, $G \cap M = \phi$.

Let $m$ be the smallest size in $M$, i.e. $m = \min_{t \in M} s_t$. Define index $t_m$ for the smallest item, or the first smallest item, if there are multiple smallest items.

$$t_m = \min \{ t \in [T] \, | s_t = m \} . \tag{C.1}$$

Denote $G'$ as the set of items accepted by Greedy, at the moment $s_{t_m}$ is blocked. Let

$g' = \mathsf{size}(G')$. See Figure 4-2. A straightforward, but useful information about $m$ is:

$$g' + m > 1, \tag{C.2}$$

because $m$ is blocked by $\mathsf{Greedy}$. We wish to understand when we can admit an item of size at least $m$, by selecting a proper threshold $\tau$.

We distinguish two cases: $m > 1/2$ and $m \leq 1/2$.

**Case 1**: $m > 1/2$.
Let $S^{\mathsf{THR}}(\tau)$ be the set of items that have sizes at least $\tau$, i.e. $S^{\mathsf{THR}}(\tau) = \{t \in S \,|\, s_t \geq \tau\}$. Now define

$$\begin{aligned} q = \max \quad & \tau \\ s.t. \quad & m + \mathsf{size}(S^{\mathsf{THR}}(\tau) \cap G') > 1 \end{aligned} \tag{C.3}$$

This means that if we adopt a $\mathsf{THR}(q)$ policy, then the size $m$ item must be blocked (possibly it will also be rejected, due to $q > m$).

Now consider the items in $S^{\mathsf{THR}}(q) \cap G'$. See Figure 4-3. These items have sizes at least $q$. We count how many size $q$ items are there, and let $n$ be the number of size $q$ items. Denote the total size of the remaining items be $x$. We know that

$$\mathsf{size}(S^{\mathsf{THR}}(q) \cap G') = nq + x. \tag{C.4}$$

We make the following observations:

1. There must exist some item from $G'$ that is of size $q$, i.e.

$$\exists t_q \in G' \subseteq [T], s.t.\ s_{t_q} = q. \tag{C.5}$$

   This is because otherwise we can select the smallest item size in $G'$ that is also larger than $q$. This item size satisfies (C.3), and violates the maximum property of $q$.

2. Size $m$ items can not fit in together with items $S^{\mathsf{THR}}(q) \cap G'$, i.e.

$$nq + x + m > 1 \tag{C.6}$$

This is because $\mathsf{size}(S^{\mathsf{THR}}(q) \cap G') = nq + x$. This is implied by (C.3).

3. A size $m$ item can fit in together with items $S^{\mathsf{THR}}(\tau) \cap G', \forall \tau > q$, i.e.

$$x + m \leq 1 \tag{C.7}$$

This is because otherwise we could further increase $q$ to $\hat{q}$ so that $\mathsf{size}(S^{\mathsf{THR}}(\hat{q}) \cap G') + m > 1$, which violates the maximum property of $q$.

We further distinguish two cases: $q > m$, and $q \leq m$.

**Case 1.1**: $q > m$.

In this case, if we adopt $\mathsf{Greedy}$ then we can get as much as $g$. This is because $g$ is defined this way.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (0, q]$ then we can get no less than $q$. This is because due to (C.5) there must exist some item $t_q \in G'$ of size $q$. We either accept it, in which case we immediately earn $q$, or we have blocked it because we admitted some item $z \in [T]$ from $M$ and consumed too much space. But $\mathsf{Greedy}$ blocks item $z$ earlier than it accepts item $t_q$, which means that $s_z \geq s_{t_q} = q$. So in either case we earn $q$.

We have the following:

$$\begin{aligned}
\mathsf{ALG}_{\mathsf{N4.3}} &\geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq q) \cdot q \\
&= F_{N4.3}(0) \cdot g + (F_{N4.3}(q) - F_{N4.3}(0)) \cdot q \\
&\geq F_{N4.3}(0) \cdot (1 - 2q) + F_{N4.3}(q) \cdot q \\
&= (1 - c_{N4.3}) \cdot (1 - 2q) + \left[ 2(1 - c_{N4.3}) - \frac{1 - 2c_{N4.3}}{q} \right] \cdot q \\
&= c_{N4.3}
\end{aligned}$$

where the second inequality is because $g \geq g' > 1 - m$ (due to (C.2)) and $1 - m > 1 - q$

(Case 1.1: $q > m$); second equality is because $q > m \geq 1/2 > q_{N4.3}$, so we plug in $F_{N4.3}(\cdot)$ as defined in (4.14).

Since $\mathsf{OPT} \leq 1$, we have $\frac{\mathsf{ALG}}{\mathsf{OPT}} \geq c_{N4.3}$.

**Case 1.2**: $q \leq m$.

First we wish to upper bound $\mathsf{OPT}$. $\mathsf{OPT}$ selects some items from $[T] = G \cup M$, where $G \cap M = \phi$. Notice that $m > 1/2$ so there is at most 1 item from $M$ that $\mathsf{OPT}$ can select. If $\mathsf{OPT}$ selects no item from $M$, then $\mathsf{OPT} \leq g$. With probability $F_{N4.3}(0)$, $\mathsf{ALG_{N4.3}}$ adopts $\mathsf{Greedy}$ and earns $g$. So we have

$$\frac{\mathsf{ALG_{N4.3}}}{\mathsf{OPT}} \geq \Pr(\tau = 0) \cdot 1 + \Pr(\tau > 0) \cdot 0 \geq F_{N4.3}(0) = 1 - c_{N4.3} \geq c_{N4.3}.$$

If $\mathsf{OPT}$ selects one item from $M$, let $t_{m'} \in [T]$ be this item. So $s_{t_{m'}} = m' \geq m$. See Figure C-1.

Figure C-1: Illustration of the items accepted by $\mathsf{OPT}$


We can partition all the items in $S$ into three sets:

$$M; \qquad S^{\mathsf{THR}}(q) \cap G'; \qquad G \backslash (S^{\mathsf{THR}}(q) \cap G')$$

Let $\tilde{g} = \mathsf{size}(G \backslash (S^{\mathsf{THR}}(q) \cap G'))$. Since $S^{\mathsf{THR}}(q) \cap G'$ and $G \backslash (S^{\mathsf{THR}}(q) \cap G')$ form a partition of $G$, we have $g = (nq + x) + \tilde{g}$. From (C.6) we know that $m' + \mathsf{size}(S^{\mathsf{THR}}(q) \cap G') \geq m + \mathsf{size}(S^{\mathsf{THR}}(q) \cap G') > 1$. This means that even $\mathsf{OPT}$ cannot pack $s_{t_{m'}}$ and $S^{\mathsf{THR}}(q) \cap G'$ together. $\mathsf{OPT}$ must block at least one item from $\{t_{m'}\} \cup (S^{\mathsf{THR}}(q) \cap G')$ – and the smallest item from this union is of size $q$ (because $q \leq m \leq m'$). So we upper bound $\mathsf{OPT}$ by:

$$\begin{aligned}\mathsf{OPT} \leq & \min\left\{1, \left[m' + \mathsf{size}(S^{\mathsf{THR}}(q) \cap G')\right] - q + \mathsf{size}(G \backslash (S^{\mathsf{THR}}(q) \cap G'))\right\} \\ = & \min\left\{1, m' + (nq + x) - q + \tilde{g}\right\}\end{aligned} \tag{C.8}$$

Then we analyze $\mathsf{ALG}_{\mathsf{N4.3}}$. If we adopt $\mathsf{Greedy}$ then we can get as much as $g$. This is because $g$ is defined this way.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (0, q]$ then we get no less than $nq + x$. This is because due to (C.4) there must exist some items in $S^{\mathsf{THR}}(q) \cap G'$, which are of size $nq + x$. For any subset of items $S_0 \subseteq (S^{\mathsf{THR}}(q) \cap G')$, we either accept it, in which case we immediately earn $\mathsf{size}(S_0)$, or we have blocked it because we admitted some item $z \in [T]$ from $M$ and consumed too much space. But $\mathsf{Greedy}$ blocks item $z$ earlier than it accepts $S_0$, which means that $s_z \geq \mathsf{size}(S_0)$. So in either case we earn $\mathsf{size}(S_0)$. Since $S_0$ is chosen arbitrarily, we will always get at least $nq + x$.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (q, m]$ then we get no less than $m$. This is because due to (C.7), any item in $S^{\mathsf{THR}}(\tau) \cap G'$ will not block item $t_m$ (from expression (C.1)); and $\tau \leq m$ so we will not reject item $t_m$. We either accept item $t_m$, in which case we immediately earn $m$, or we have blocked it because we admitted some item $z \in [T]$ from M and consumed too much space. But $m$ is smallest item size in $M$, which means that $s_z \geq m$. So in either case we earn $m$.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (m, m']$ then we get no less than $\tau$. This is because $s_{t_{m'}}$ does exist, and $\mathsf{THR}(\tau)$ must accept at least one item. The least that $\mathsf{THR}(\tau)$ can get is $\tau$.

We have the following:

$$\begin{aligned}&\mathsf{ALG}_{\mathsf{N4.3}} \\ &\geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq q) \cdot (nq + x) + \Pr(q < \tau \leq m) \cdot m + \int_m^{m'} \tau \, \mathrm{d}F_{N4.3}(\tau) \\ &= F_{N4.3}(0) \cdot (nq + x + \tilde{g}) + (F_{N4.3}(q) - F_{N4.3}(0)) \cdot (nq + x) \\ &\quad + (F_{N4.3}(m) - F_{N4.3}(q)) \cdot m + \int_m^{m'} \tau \, \mathrm{d}F_{N4.3}(\tau)\end{aligned}$$

$$= F_{N4.3}(0) \cdot \tilde{g} + F_{N4.3}(q) \cdot (nq + x - m) + F_{N4.3}(m') \cdot m' - \int_{m}^{m'} F_{N4.3}(\tau)\ \mathrm{d}\tau$$

$$\geq F_{N4.3}(0) \cdot \tilde{g} + F_{N4.3}(q) \cdot (2(nq + x) - 1) + F_{N4.3}(m') \cdot m' - \int_{1-(nq+x)}^{m'} F_{N4.3}(\tau)\ \mathrm{d}\tau$$

where the second equality is due to integration by part (our definition of $F_{N4.3}(\cdot)$ in (4.14) is a continuous function); the last inequality is because $\frac{\partial \mathsf{ALG_{N4.3}}}{\partial m} = F_{N4.3}(m) - F_{N4.3}(q) \geq 0$, (because $q \leq m$, and $F_{N4.3}(\cdot)$ is a increasing function), so that $\mathsf{ALG_{N4.3}}$ is increasing in $m$. Hence, $\mathsf{ALG_{N4.3}}$ achieves its minimum when $m$ is the smallest, and $m > 1 - (nq + x)$ from (C.6).

Observe that

$$\begin{aligned}
&\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT} \\
\geq &F_{N4.3}(0) \cdot \tilde{g} + F_{N4.3}(q) \cdot (2(nq + x) - 1) + F_{N4.3}(m') \cdot m' - \int_{1-(nq+x)}^{m'} F_{N4.3}(\tau)\ \mathrm{d}\tau \\
&\quad - c_{N4.3} \cdot \min\{1, m' + (nq + x) - q + \tilde{g}\}
\end{aligned}$$

If we focus on the dependence of $\tilde{g}$, we find that

$$\frac{\partial\left(\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}\right)}{\partial \tilde{g}} \geq F_{N4.3}(0) - c_{N4.3} = 1 - 2c_{N4.3} \geq 0,$$

where the first inequality is because the subgradient of the subtracted term is either 0 or $c_{N4.3}$. Since $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}$ is a increasing function of $\tilde{g}$, it achieves its minimum when $\tilde{g} = 0$.

We have further

$$\begin{aligned}
&\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT} \\
\geq &F_{N4.3}(q) \cdot (2(nq + x) - 1) + F_{N4.3}(m') \cdot m' - \int_{1-(nq+x)}^{m'} F_{N4.3}(\tau)\ \mathrm{d}\tau \\
&\quad - c_{N4.3} \cdot \min\{1, m' + (nq + x) - q\}
\end{aligned}$$

Now let $y = (n-1)q + x$, and we plug in $F_{N4.3}(\cdot)$ as we defined in (C.6).

**Case 1.2.1**: When $q \le q_{N4.3}$, we have:

$$
\begin{aligned}
&\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT} \\
\ge& F_{N4.3}(q) \cdot (2(q+y)-1) + F_{N4.3}(m') \cdot m' - \int_{1-(q+y)}^{m'} F_{N4.3}(\tau)\ \mathrm{d}\tau - c_{N4.3} \cdot \min\{1, m'+y\} \\
=& (1-c_{N4.3})(2(q+y)-1) + \frac{(2q-1+2y)(1-2c_{N4.3})\ln(1-q)}{2q-1} + 2(1-c_{N4.3})m' - (1-2c_{N4.3}) \\
&- 2(1-c_{N4.3})m' + 2(1-c_{N4.3})[1-(q+y)] + (1-2c_{N4.3})\left[\ln m' - \ln(1-(q+y))\right] \\
&\qquad - c_{N4.3} \cdot \min\{1, m'+y\} \\
=& c_{N4.3} + \frac{(2q-1+2y)(1-2c_{N4.3})\ln(1-q)}{2q-1} + (1-2c_{N4.3})\left[\ln m' - \ln(1-(q+y))\right] \\
&\qquad - c_{N4.3} \cdot \min\{1, m'+y\}
\end{aligned}
$$

If we focus on the dependence of $m'$, we will see that $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}$ has only one local minimum: when $m' < 1-y$ we have

$$
\frac{\partial\left(\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}\right)}{\partial m'} = \frac{1-2c_{N4.3}}{m'} - c_{N4.3} \le \frac{1-2c_{N4.3}}{1/2} - c_{N4.3} = 2 - 5c_{N4.3} < 0,
$$

because $m' \ge m \ge 1/2$. So $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}$ is decreasing on $m'$ when $m' < 1-y$. When $m' > 1-y$ we have

$$
\frac{\partial\left(\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}\right)}{\partial m'} = \frac{1-2c_{N4.3}}{m'} > 0,
$$

so $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}$ is increasing on $m'$. Hence, $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}$ achieves its minimum when $m' = 1-y$.

Plugging into $m' = 1-y$, we have further

$$
\begin{aligned}
&\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT} \\
\ge& \frac{(2q-1+2y)(1-2c_{N4.3})\ln(1-q)}{2q-1} + (1-2c_{N4.3})\left[\ln(1-y) - \ln(1-(q+y))\right]
\end{aligned}
$$

If we focus on the dependence of $y$, we find that

$$\frac{\partial\left(\mathsf{ALG}_{\mathsf{N4.3}}-c_{N4.3}\mathsf{OPT}\right)}{\partial y}=(1-2c_{N4.3})\left[2\frac{\ln(1-q)}{2q-1}-\frac{1}{1-y}+\frac{1}{1-y-q}\right]>0,$$

because $ln(1-q)<0, 2q-1<2q_{N4.3}-1<0, \frac{1}{1-y-q}-\frac{1}{1-y}\geq 0$. Since $\mathsf{ALG}_{\mathsf{N4.3}}-c_{N4.3}\mathsf{OPT}$ is increasing on $y$, it achieves its minimum when $y=0$.

Finally, plugging into $y=0$, we have

$$\mathsf{ALG}_{\mathsf{N4.3}}-c_{N4.3}\mathsf{OPT}\geq\frac{(2q-1)(1-2c_{N4.3})\ln(1-q)}{2q-1}-(1-2c_{N4.3})\ln(1-q)=0$$

**Case 1.2.2**: When $q>q_{N4.3}$, we have:

$$\begin{aligned}
&\mathsf{ALG}_{\mathsf{N4.3}}-c_{N4.3}\mathsf{OPT}\\
\geq&F_{N4.3}(q)\cdot(2(q+y)-1)+F_{N4.3}(m')\cdot m'-\int_{1-(q+y)}^{m'}F_{N4.3}(\tau)\,\mathrm{d}\tau-c_{N4.3}\cdot\min\{1,m'+y\}\\
=&2(1-c_{N4.3})(2(q+y)-1)-\frac{(1-2c_{N4.3})(2(q+y)-1)}{q}+2(1-c_{N4.3})m'-(1-2c_{N4.3})\\
&-2(1-c_{N4.3})m'+2(1-c_{N4.3})[1-(q+y)]+(1-2c_{N4.3})\left[\ln m'-\ln(1-(q+y))\right]\\
&-c_{N4.3}\cdot\min\{1,m'+y\}\\
=&2(1-c_{N4.3})(y+q)-(1-2c_{N4.3})-\frac{(1-2c_{N4.3})(2(q+y)-1)}{q}\\
&+(1-2c_{N4.3})\left[\ln m'-\ln(1-(q+y))\right]-c_{N4.3}\cdot\min\{1,m'+y\}
\end{aligned}$$

Again, if we focus on the dependence of $m'$, we will see that $\mathsf{ALG}_{\mathsf{N4.3}}-c_{N4.3}\mathsf{OPT}$ has only one local minimum when $m'=1-y$.

Plugging into $m'=1-y$, we have further

$$\begin{aligned}
&\mathsf{ALG}_{\mathsf{N4.3}}-c_{N4.3}\mathsf{OPT}\\
\geq&(1-c_{N4.3})(2(y+q)-1)-\frac{(1-2c_{N4.3})(2(q+y)-1)}{q}+(1-2c_{N4.3})\left[\ln(1-y)-\ln(1-(q+y))\right]
\end{aligned}$$

Again, if we focus on the dependence of $y$, we find that

$$\frac{\partial\left(\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}\right)}{\partial y} = 2(1-c_{N4.3}-\frac{1-2c_{N4.3}}{q})+(1-2c_{N4.3})\left[-\frac{1}{1-y}+\frac{1}{1-y-q}\right] > 0,$$

because $1-c_{N4.3}-\frac{1-2c_{N4.3}}{q} \geq 1-c_{N4.3}-\frac{1-2c_{N4.3}}{q_{N4.3}} \approx 0.142 > 0$, $\frac{1}{1-y-q}-\frac{1}{1-y} \geq 0$. Since $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}$ is increasing on $y$, it achieves its minimum when $y = 0$.

Finally, plugging into $y = 0$, we have

$$\begin{aligned}
&\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT}\\
\geq&(1-c_{N4.3})(2q-1) - \frac{(1-2c_{N4.3})(2q-1)}{q} - (1-2c_{N4.3})\ln(1-q)\\
=&(2q-1)\left[(1-c_{N4.3}) - \frac{(1-2c_{N4.3})}{q}\right] - (1-2c_{N4.3})\ln(1-q)\\
\geq&(2q-1)\left[-\frac{(1-2c_{N4.3})\ln(1-q)}{1-2q}\right] - (1-2c_{N4.3})\ln(1-q)\\
=&0
\end{aligned}$$

where the second inequality is because $H(c_{N4.3}, q) = \frac{1-2c_{N4.3}}{q} - \frac{(1-2c_{N4.3})\ln(1-q)}{1-2q} - (1-c_{N4.3}) \geq 0, \forall q \in (0, 1/2)$ from (4.13), and when $q \in [1/2, 1]$, the second line expression is an increasing function of $q$ (because $2q-1$; $(1-c_{N4.3}) - \frac{(1-2c_{N4.3})}{q}$; and $-(1-2c_{N4.3})\ln(1-q)$ are all increasing in $q$), thus plugging in $q = 1/2$ we have $\mathsf{ALG_{N4.3}} - c_{N4.3}\mathsf{OPT} \geq -(1-2c_{N4.3})\ln(1-q) > 0$.

In all, $\mathsf{ALG_{N4.3}} \geq c_{N4.3}\mathsf{OPT}$.

**Case 2**: $m \leq 1/2$.

In this case, we only hope to get $m$, and a crude analysis is enough. See Figure 4-4.

If we adopt Greedy then we can get as much as $g$. This is because $g$ is defined this way.

If we adopt $\mathsf{THR}(\tau), \forall \tau \in (0, m]$ then we either get $m$, or $m$ is blocked, in which case we must have already earned at least $1-m$ to block $m$.

We have the following:

$$\mathsf{ALG} \geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq m) \cdot \min\{m, 1-m\}$$

$$
\begin{aligned}
&\geq \Pr(\tau = 0) \cdot g + \Pr(0 < \tau \leq m) \cdot m \\
&= F_{N4.3}(0) \cdot g + (F_{N4.3}(m) - F_{N4.3}(0)) \cdot m \\
&\geq F_{N4.3}(0) \cdot (1 - m) + (F_{N4.3}(m) - F_{N4.3}(0)) \cdot m \\
&= F_{N4.3}(0) \cdot (1 - 2m) + F_{N4.3}(m) \cdot m \\
&= (1 - c_{N4.3})(1 - 2m) + \left[2(1 - c_{N4.3}) - \frac{1 - 2c_{N4.3}}{m}\right] \cdot m \\
&= c_{N4.3}
\end{aligned}
$$

where the second inequality is because $m \leq 1/2$; the last inequality is because $g \geq g' > 1 - m$ (due to (C.2)); the third equality is because we plug in $F_{N4.3}(\cdot)$ as defined in (4.14).

Since $\mathsf{OPT} \leq 1$, we have $\frac{\mathsf{ALG}}{\mathsf{OPT}} \geq c_{N4.3}$.

In all, we have enumerated all the possible cases, to find $\dfrac{\mathsf{ALG}}{\mathsf{OPT}} \geq c_{N4.3}$ always holds. $\square$

# Appendix D

# Appendix to Chapter 5

## D.1 Theorems Used

We summarize here the results that we have directly used in our proofs.

*Definition* D.1 ($\phi$-Dependent Random Variables, Hoeffding and Robbins (1948)). For any sequence $\{X_1, X_2, ...\}$, if there exists $\phi$ such that for any $s - r > \phi$, the two sets

$$(X_1, X_2, ..., X_r), \quad (X_s, X_{s+1}, ..., X_n)$$

are independent, then the sequence is said to be $\phi$-dependent.

*Lemma* D.1 (Romano and Wolf (2000), Theorem 2.1). *Let $\{X_{n,i}\}$ be a triangular array of zero-mean random variables. Let $\phi \in \mathbb{N}$ be a fixed constant. For each $n = 1, 2, ...$, let $d = d_n$, and suppose that $X_{n,1}, X_{n,2}, ..., X_{n,d}$ is an $\phi$-dependent sequence of random variables. Define*

$$B^2_{n,k,a} = \mathsf{Var}\left(\sum_{i=a}^{a+k-1} X_{n,i}\right), \qquad B^2_n = B^2_{n,d,1} = \mathsf{Var}\left(\sum_{i=1}^{d} X_{n,i}\right)$$

*For some $\delta > 0$ and $-1 \leq \gamma \leq 1$, if the following conditions hold:*

1. $\mathbb{E}\left|X_{n,i}\right|^{2+\delta} \leq \Delta_n$, *for all $i$;*

2. $B^2_{n,k,a}/k^{1+\gamma} \leq K_n$, *for all $a$ and $k \geq \phi$;*

*3.* $B_n^2/(d\phi^\gamma) \geq L_n$*;*

*4.* $K_n/L_n = O(1)$*;*

*5.* $\Delta/L_n^{(2+\delta)/2} = O(1)$*,*

*then*

$$\frac{\sum_{i=1}^d X_{n,i}}{B_n} \xrightarrow{D} \mathcal{N}(0,1).$$

We explain Lemma D.1. The $\xrightarrow{D}$ notation stands for convergence in distribution. The definition of a sequence of $\phi$-dependent random variables is given in Definition D.1. To check if the conditions in Lemma D.1 hold, we will first calculate $B_{n,k,a}^2$ for any $k$ and $a$, and then construct some proper $\Delta_n, K_n$, and $L_n$.

*Lemma* D.2. *For any* $n \in \mathbb{N}$ *and* $q_1, ..., q_n \in (0,1)$*, define*

$$f(q_1, ..., q_n) = \frac{1}{\prod_{i=1}^n q_i} + \frac{1}{\prod_{i=1}^n (1-q_i)}.$$

*Then*

$$f(q_1, ..., q_n) \geq 2^{n+1},$$

*where equality holds if and only if* $q_1 = q_2 = ... = q_n = 1/2$*.*

The proof of Lemma D.2 is elegant and is of its own interests. We prove Lemma D.2 below.

*Proof.* Proof of Lemma D.2. For all $i \in [n]$ denote $\bar{q}_i = 1 - q_i$. We re-write our objective, such that we wish to find the minimum for

$$\frac{1}{\prod_{i=1}^n q_i} + \frac{1}{\prod_{i=1}^n \bar{q}_i},$$

under the constraints that $q_i + \bar{q}_i = 1$ for all $i \in [n]$. Note that $\prod_{i=1}^n (q_i + \bar{q}_i) = 1$. By expanding expand the product term and we have

$$\frac{1}{\prod_{i=1}^n q_i} = \frac{\prod_{i=1}^n (q_i + \bar{q}_i)}{\prod_{i=1}^n q_i} =$$
$$1 + \left(\frac{\bar{q}_1}{q_1} + \frac{\bar{q}_2}{q_2} + \ldots + \frac{\bar{q}_n}{q_n}\right) + \left(\frac{\bar{q}_1\bar{q}_2}{q_1q_2} + \frac{\bar{q}_1\bar{q}_3}{q_1q_3} + \ldots + \frac{\bar{q}_{n-1}\bar{q}_n}{q_{n-1}q_n}\right) + \ldots + \frac{\prod_{i=1}^n \bar{q}_i}{\prod_{i=1}^n q_i}$$

And similarly we can expand the product term for the second fractional expression. Putting them together we have:

$$\begin{aligned}
&\frac{1}{\prod_{i=1}^n q_i} + \frac{1}{\prod_{i=1}^n \bar{q}_i} \\
=&1 + \left(\frac{\bar{q}_1}{q_1} + \frac{\bar{q}_2}{q_2} + \ldots + \frac{\bar{q}_n}{q_n}\right) + \left(\frac{\bar{q}_1\bar{q}_2}{q_1q_2} + \frac{\bar{q}_1\bar{q}_3}{q_1q_3} + \ldots + \frac{\bar{q}_{n-1}\bar{q}_n}{q_{n-1}q_n}\right) + \ldots + \frac{\prod_{i=1}^n \bar{q}_i}{\prod_{i=1}^n q_i} \\
&+ 1 + \left(\frac{q_1}{\bar{q}_1} + \frac{q_2}{\bar{q}_2} + \ldots + \frac{q_n}{\bar{q}_n}\right) + \left(\frac{q_1q_2}{\bar{q}_1\bar{q}_2} + \frac{q_1q_3}{\bar{q}_1\bar{q}_3} + \ldots + \frac{q_{n-1}q_n}{\bar{q}_{n-1}\bar{q}_n}\right) + \ldots + \frac{\prod_{i=1}^n q_i}{\prod_{i=1}^n \bar{q}_i}
\end{aligned}$$

Now focus on the right hand side. There are a total of $2^{n+1}$ terms, and we match them into $2^n$ pairs. We match the first term in the first line with the first term in the second line, the second term in the first line with the second term in the second line, ..., the last term in the first line with the last term in the second line. For each pair indexed by subset $I \subseteq [T]$, we have that

$$\frac{\prod_{i \in I \subseteq [T]} \bar{q}_i}{\prod_{i \in I \subseteq [T]} q_i} + \frac{\prod_{i \in I \subseteq [T]} q_i}{\prod_{i \in I \subseteq [T]} \bar{q}_i} \geq 2,$$

where equality holds if and only if $\prod_{i \in I \subseteq [T]} q_i = \prod_{i \in I \subseteq [T]} \bar{q}_i$. Putting all the $2^n$ pairs together we finish the proof. $\square$

## D.2 Proof from Section 5.2

The only proof from Section 5.2 is the unbiasedness of the Horvitz-Thompson estimator. We prove by checking the expectations.

*Proof.* Proof of Proposition 5.2. First observe that for regular switchback experiments, both $0 < \Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}), \Pr(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1}) < 1$. So for any $t \in \{p+1 : T\}$, with probability $\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}) \neq 0$, $\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}\} = 1$, and $Y_t^{\mathsf{obs}} = Y_t(\mathbf{1}_{p+1})$.

So $\mathbb{E}\left[Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1})}\right] = Y_t(\mathbf{1}_{m+1})$. Similarly $\mathbb{E}\left[Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1})}\right] = Y_t(\mathbf{0}_{p+1})$. Sum them up for any $t \in \{p+1:T\}$ we finish the proof. $\square$

## D.3 Proofs and Discussions from Section 5.3

In Section 5.3 we focus on the case when $p = m$. Throughout this section in the appendix, we use only $m$ instead of $p$.

### D.3.1 Extra Notations Used in the Proofs from Section 5.3

Recall that any regular switchback experiment can be represented by $\mathbb{T} = \{t_0, t_1, ..., t_K\} \subseteq [T]$ and $\mathbb{Q} = (q_0, q_1, ..., q_K) \in (0,1)^{K+1}$. We first focus on the dependence on $\mathbb{T}$, the randomization points. Define $f_{\mathbb{T}} : [T] \to \mathbb{T}$ to be the "determining randomization point of period $t$", i.e.,

$$f_{\mathbb{T}}(t) = \max\{j \,|\, j \in \mathbb{T}, j \leq t\}$$

such that the coin flip in period $f_{\mathbb{T}}(t)$ uniquely determines the distribution of $W_t$, i.e., $W_t = W_{f_{\mathbb{T}}(t)}$. When $\mathbb{T}$ is clear from the context we also omit the subscript and use $f(t)$ for $f_{\mathbb{T}}(t)$.

Similarly, we define $f_{\mathbb{T}}^m(t) : [T] \to \{0,1\}^{\mathbb{T}}$, which maps a time period to a subset of $\mathbb{T}$, to be the "determining randomization points of periods $\{t-m, t-m+1, ..., t\}$", i.e.

$$f_{\mathbb{T}}^m(t) = \{j \,|\, \exists i \in \{t-m, ..., t\}, s.t.\ j = f_{\mathbb{T}}(i)\}$$

such that $f_{\mathbb{T}}^m(t) \subseteq \mathbb{T} \subseteq [T]$. And $f_{\mathbb{T}}^m(t)$ contains all the time periods whose coin flips uniquely determine the distributions of $W_{t-m}, W_{t-m+1}, ..., W_t$. Denote $|f_{\mathbb{T}}^m(t)| = J$, the cardinality of set $f_{\mathbb{T}}^m(t)$. We keep in mind that $J$ depends on $m, t$ and $\mathbb{T}$, yet they are all omitted for brevity. Since the treatment assignments $\boldsymbol{W}_{t-m:t}$ are determined by at least one randomization point $f(t-m)$, we know that $f_{\mathbb{T}}^m(t) \neq \emptyset$ is non-empty, i.e.,

$$|f_{\mathbb{T}}^m(t)| = J \geq 1. \tag{D.1}$$

Let the elements be $f^m_{\mathbb{T}}(t) = \{u_1, u_2, ..., u_J\}$, and let $u_1 < u_2 < ... < u_J$.

Finally, define "overlapping randomization points of periods $\{t-m, t-m+1, ..., t\}$ and $\{t'-m, t'-m+1, ..., t'\}$" to be

$$O_{\mathbb{T}}(t, t') = f^m_{\mathbb{T}}(t) \cap f^m_{\mathbb{T}}(t')$$

Denote $|O_{\mathbb{T}}(t, t')| = J^{\circ}$. We keep in mind that $J^{\circ}$ depends on $m, t, t'$ and $\mathbb{T}$, yet they are all omitted for brevity.

Now we introduce an important short-hand notation. Recall that for any randomization point $t_k$, the associated $q_k$ is the probability that $W_{t_k}$ receives treatment, i.e., $q_k = \Pr(W_{t_k} = 1)$. And recall that $\bar{q}_k = 1 - q_k$. Now define for any $t \in \{m+1 : T\}$,

$$\begin{aligned} \mathbb{1}_t(\mathbb{T}, \mathbb{Q}, \mathbb{Y}) = Y_t(\mathbf{1}_{m+1}) &\left[ \mathbb{1}\{\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}\} \prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1 \right] \\ &- Y_t(\mathbf{0}_{m+1}) \left[ \mathbb{1}\{\boldsymbol{W}_{t-m:t} = \mathbf{0}_{m+1}\} \prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1 \right] \end{aligned} \tag{D.2}$$

where we use $\prod_{j=1}^{J}(1/q_{u_j})$ and $\prod_{j=1}^{J}(1/\bar{q}_{u_j})$ to calculate the inverse propensity scores. When $\mathbb{T}, \mathbb{Q}$ and $\mathbb{Y}$ are clear from the context we omit them and use $\mathbb{1}_t$ for $\mathbb{1}_t(\mathbb{T}, \mathbb{Q}, \mathbb{Y})$.

Using the above notation, we could re-write

$$\hat{\tau}_m - \tau_m = \frac{1}{T-m} \sum_{t=m+1}^{T} \mathbb{1}_t$$

Similar to Proposition 5.2, we can check the expectation of $\mathbb{1}_t$ by expanding the probability governing $\boldsymbol{W}$ (the only source of randomness is our assignment path $\boldsymbol{W}$). For any $t \in \{m+1, m+2, ..., T\}$,

$$\mathbb{E}[\mathbb{1}_t] = 0. \tag{D.3}$$

### D.3.2 Preliminary Results

In this section we introduce two Lemmas for the proof of Theorem 5.4 and proof of Lemma 5.3.

*Lemma* D.3. *Under Assumptions 5.1–5.2, for any $t \in [T]$, let $|f_{\mathbb{T}}^m(t)| = J$.*

$$\mathbb{E}[\mathbb{1}_t^2] = \left(\prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1\right) Y_t(\mathbf{1}_{m+1})^2 + 2Y_t(\mathbf{1}_{m+1})Y_t(\mathbf{0}_{m+1}) + \left(\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1\right) Y_t(\mathbf{0}_{m+1})^2. \tag{D.4}$$

*Proof.* Proof of Lemma D.3.

Denote $|f_{\mathbb{T}}^m(t)| = J$. Let the elements be $f_{\mathbb{T}}^m(t) = \{u_1, u_2, ..., u_J\}$. Let $u_1 < u_2 < ... < u_J$.

Using the notations defined earlier in Section D.3.1 and, in particular, the definition of (D.2), we can directly calculate the squared terms of $\mathbb{E}[\mathbb{1}_t^2]$ by consulting the law of total expectation.

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_t^2] =& \Pr\left(\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}\right) \cdot \mathbb{E}[\mathbb{1}_t^2 \,|\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}] \\
&+ \Pr\left(\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}\right) \cdot \mathbb{E}[\mathbb{1}_t^2 \,|\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}] \\
&+ \Pr\left(\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}\right) \cdot \mathbb{E}[\mathbb{1}_t^2 \,|\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}] \\
=& \Pr\left(\boldsymbol{W}_{t-m:t} = \mathbf{1}_{m+1}\right) \cdot \left\{ Y_t(\mathbf{1}_{m+1}) \left(\prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1\right) - Y_t(\mathbf{0}_{m+1})(0-1) \right\}^2 \\
&+ \Pr\left(\boldsymbol{W}_{t-m:t} = \mathbf{0}_{m+1}\right) \cdot \left\{ Y_t(\mathbf{1}_{m+1})(0-1) - Y_t(\mathbf{0}_{m+1}) \left(\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1\right) \right\}^2 \\
&+ \Pr\left(\boldsymbol{W}_{t-m:t} \neq \mathbf{1}_{m+1} \text{ or } \mathbf{0}_{m+1}\right) \cdot \left\{ Y_t(\mathbf{1}_{m+1})(0-1) - Y_t(\mathbf{0}_{m+1})(0-1) \right\}^2 \\
=& \Pr\left((W_{u_1}, ..., W_{u_J}) = \mathbf{1}_J\right) \cdot \left\{ \left(\prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1\right) Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \right\}^2 \\
&+ \Pr\left((W_{u_1}, ..., W_{u_J}) = \mathbf{0}_J\right) \cdot \left\{ -Y_t(\mathbf{1}_{m+1}) - \left(\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1\right) Y_t(\mathbf{0}_{m+1}) \right\}^2 \\
&+ \Pr\left((W_{u_1}, ..., W_{u_J}) \neq \mathbf{1}_J \text{ or } \mathbf{0}_J\right) \cdot \left\{ -Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \right\}^2
\end{aligned}$$

$$
\begin{aligned}
=&\prod_{j=1}^{J} q_{u_j} \cdot \left\{ \prod_{j=1}^{J} \frac{1}{q_{u_j}} \cdot Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \right\}^2 \\
&+ \prod_{j=1}^{J} \bar{q}_{u_j} \cdot \left\{ - \prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} \cdot Y_t(\mathbf{0}_{m+1}) - Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \right\}^2 \\
&+ \left( 1 - \prod_{j=1}^{J} q_{u_j} - \prod_{j=1}^{J} \bar{q}_{u_j} \right) \cdot \{ -Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \}^2 \\
=&\left( \prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1 \right) Y_t(\mathbf{1}_{m+1})^2 + 2Y_t(\mathbf{1}_{m+1})Y_t(\mathbf{0}_{m+1}) + \left( \prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1 \right) Y_t(\mathbf{0}_{m+1})^2
\end{aligned}
$$

which finishes the proof. □

*Lemma* D.4. *Under Assumptions 5.1–5.2, for any $t < t' \in [T]$, when $|O_{\mathbb{T}}(t,t')| = J^{\circ} = 0$,*

$$
\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0. \tag{D.5}
$$

*When $|O_{\mathbb{T}}(t,t')| = J^{\circ} \geq 1$,*

$$
\begin{aligned}
\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] =& (\prod_{j=1}^{J^{\circ}} \frac{1}{q_{u_j^{\circ}}} - 1) Y_t(\mathbf{1}_{m+1}) Y_{t'}(\mathbf{1}_{m+1}) + Y_t(\mathbf{1}_{m+1}) Y_{t'}(\mathbf{0}_{m+1}) \\
&+ Y_t(\mathbf{0}_{m+1}) Y_{t'}(\mathbf{1}_{m+1}) + (\prod_{j=1}^{J^{\circ}} \frac{1}{\bar{q}_{u_j^{\circ}}} - 1) Y_t(\mathbf{0}_{m+1}) Y_{t'}(\mathbf{0}_{m+1}).
\end{aligned} \tag{D.6}
$$

*Proof.* Proof of Lemma D.4. Denote $|f_{\mathbb{T}}^m(t)| = J$, $|f_{\mathbb{T}}^m(t')| = J'$, and $|O_{\mathbb{T}}(t,t')| = J^{\circ}$. Let the elements be $f_{\mathbb{T}}^m(t) = \{u_1, u_2, ..., u_J\}$, $f_{\mathbb{T}}^m(t') = \{u'_1, u'_2, ..., u'_{J'}\}$, and $O_{\mathbb{T}}(t,t') = \{u_1^{\circ}, u_2^{\circ}, ..., u_{J^{\circ}}^{\circ}\}$. Let $u_1 < u_2 < ... < u_J$, $u'_1 < u'_2 < ... < u'_{J'}$, and $u_1^{\circ} < u_2^{\circ} < ... < u_{J^{\circ}}^{\circ}$.

One time period could have different numberings in $f_{\mathbb{T}}^m(t)$, $f_{\mathbb{T}}^m(t')$, and $O_{\mathbb{T}}(t,t')$. For example, $u_{J-J^{\circ}+1} = u'_1 = u_1^{\circ}$, and $u_J = u'_{J^{\circ}} = u_{J^{\circ}}^{\circ}$. See Table D.1 for an illustrator of the determining randomization points and the overlapping randomization points.

First, when $J^{\circ} = 0$, this implies that $\mathbb{1}_t$ and $\mathbb{1}_{t'}$ are independent. Then $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = \mathbb{E}[\mathbb{1}_t]\mathbb{E}[\mathbb{1}_{t'}] = 0$, where the second equality is due to (D.3).

When $J^{\circ} \geq 1$, this implies that $\mathbb{1}_t$ and $\mathbb{1}_{t'}$ are correlated. Using the notations

Table D.1: Illustrator of the determining randomization points and the overlapping randomization points

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $u_1$ | $u_2$ | ... | $u_{J-J^\circ+1}$ | ... | $u_J$ | | | |
| | | | $u_1^\circ$ | ... | $u_{J^\circ}^\circ$ | | | |
| | | | $u_1'$ | ... | $u_{J^\circ}'$ | $u_{J^\circ+1}'$ | ... | $u_{J'}'$ |

Note: Each columns stands for one time period. The first row stands for the determining randomization points of $f_{\mathbb{T}}^m(t)$; the second row for the overlapping randomization points of $O_{\mathbb{T}}(t,t')$; and the third row for the determining randomization points of $f_{\mathbb{T}}^m(t')$.

defined above,

$$\begin{aligned}\mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}] =&\mathbb{E}_{W_{u_1^\circ},...,W_{u_{J^\circ}^\circ}}\left[\mathbb{E}\left[\mathbb{1}_t\mathbb{1}_{t'}\left|W_{u_1^\circ},...,W_{u_{J^\circ}^\circ}\right.\right]\right] \\ =&\Pr\left((W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})=\mathbf{1}_{J^\circ}\right)\mathbb{E}\left[\mathbb{1}_t\mathbb{1}_{t'}\left|(W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})=\mathbf{1}_{J^\circ}\right.\right] \\ &+\Pr\left((W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})=\mathbf{0}_{J^\circ}\right)\mathbb{E}\left[\mathbb{1}_t\mathbb{1}_{t'}\left|(W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})=\mathbf{0}_{J^\circ}\right.\right] \\ &+\Pr\left((W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})\neq\mathbf{1}_{J^\circ}\text{ or }\ \mathbf{0}_{J^\circ}\right)\mathbb{E}\left[\mathbb{1}_t\mathbb{1}_{t'}\left|(W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})\neq\mathbf{1}_{J^\circ}\text{ or }\ \mathbf{0}_{J^\circ}\right.\right]\end{aligned} \tag{D.7}$$

Next we go over the three cases of $(W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})$ as decomposed above. Note that conditional on $(W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})$, $\mathbb{1}_t$ and $\mathbb{1}_{t'}$ are independent, i.e.,

$$\mathbb{E}\left[\mathbb{1}_t\mathbb{1}_{t'}\left|W_{u_1^\circ},...,W_{u_{J^\circ}^\circ}\right.\right]=\mathbb{E}\left[\mathbb{1}_t\left|W_{u_1^\circ},...,W_{u_{J^\circ}^\circ}\right.\right]\mathbb{E}\left[\mathbb{1}_{t'}\left|W_{u_1^\circ},...,W_{u_{J^\circ}^\circ}\right.\right]$$

**(1)** With probability $\prod_{j=1}^{J^\circ}q_{u_j^\circ}$, $(W_{u_1^\circ},...,W_{u_{J^\circ}^\circ})=\mathbf{1}_{J^\circ}$. In this case

$$\begin{aligned}&\mathbb{E}\left[\mathbb{1}_t\left|W_{u_1^\circ},...,W_{u_{J^\circ}^\circ}\right.\right] \\ =&\Pr\left(\boldsymbol{W}_{t-m:t}=\mathbf{1}_{m+1}\right)\cdot\left\{Y_t(\mathbf{1}_{m+1})(\prod_{j=1}^{J}\frac{1}{q_{u_j}}-1)+Y_t(\mathbf{0}_{m+1})\right\} \\ &+\Pr\left(\boldsymbol{W}_{t-m:t}\neq\mathbf{1}_{m+1}\right)\cdot\left\{Y_t(\mathbf{1}_{m+1})(0-1)+Y_t(\mathbf{0}_{m+1})\right\} \\ =&\Pr\left((W_{u_1},W_{u_2},...,W_{u_{J-J^\circ}})=\mathbf{1}_{J-J^\circ}\right)\cdot\left\{Y_t(\mathbf{1}_{m+1})(\prod_{j=1}^{J}\frac{1}{q_{u_j}}-1)+Y_t(\mathbf{0}_{m+1})\right\} \\ &+\Pr\left((W_{u_1},W_{u_2},...,W_{u_{J-J^\circ}})\neq\mathbf{1}_{J-J^\circ}\right)\cdot\left\{-Y_t(\mathbf{1}_{m+1})+Y_t(\mathbf{0}_{m+1})\right\} \\ =&\prod_{j=1}^{J-J^\circ}q_{u_j}\cdot\left\{Y_t(\mathbf{1}_{m+1})(\prod_{j=1}^{J}\frac{1}{q_{u_j}}-1)+Y_t(\mathbf{0}_{m+1})\right\}\end{aligned}$$

$$
\begin{aligned}
&+ (1 - \prod_{j=1}^{J-J^\circ} q_{u_j}) \cdot \{-Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1})\} \\
=&(\prod_{j=1}^{J^\circ} \frac{1}{q_{u_j^\circ}} - 1) Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1})
\end{aligned}
$$

where the third equality is due to (5.2). Similarly,

$$
\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{t'} \middle| W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}\right] \\
=&\Pr\left(\boldsymbol{W}_{t'-m:t'} = \mathbf{1}_{m+1}\right) \cdot \left\{Y_{t'}(\mathbf{1}_{m+1})(\prod_{j=1}^{J'} \frac{1}{q_{u'_j}} - 1) + Y_{t'}(\mathbf{0}_{m+1})\right\} \\
&+ \Pr\left(\boldsymbol{W}_{t'-m:t'} \neq \mathbf{1}_{m+1}\right) \cdot \{Y_{t'}(\mathbf{1}_{m+1})(0-1) + Y_{t'}(\mathbf{0}_{m+1})\} \\
=&\Pr\left((W_{u'_{J^\circ+1}}, W_{u'_{J^\circ+2}}, ..., W_{u'_{J'}}) = \mathbf{1}_{J'-J^\circ}\right) \cdot \left\{Y_{t'}(\mathbf{1}_{m+1})(\prod_{j=1}^{J'} \frac{1}{q_{u'_j}} - 1) + Y_{t'}(\mathbf{0}_{m+1})\right\} \\
&+ \Pr\left((W_{u'_{J^\circ+1}}, W_{u'_{J^\circ+2}}, ..., W_{u'_{J'}}) \neq \mathbf{1}_{J'-J^\circ}\right) \cdot \{-Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1})\} \\
=&\prod_{j=J^\circ+1}^{J'} q_{u'_j} \cdot \left\{Y_{t'}(\mathbf{1}_{m+1})(\prod_{j=1}^{J'} \frac{1}{q_{u'_j}} - 1) + Y_{t'}(\mathbf{0}_{m+1})\right\} \\
&+ (1 - \prod_{j=J^\circ+1}^{J'} q_{u'_j}) \cdot \{-Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1})\} \\
=&(\prod_{j=1}^{J^\circ} \frac{1}{q_{u_j^\circ}} - 1) Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1})
\end{aligned}
$$

**(2)** With probability $\prod_{j=1}^{J^\circ} \bar{q}_{u_j^\circ}$, $(W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}) = \mathbf{0}_{J^\circ}$. This case is similar to Case (1), and we can calculate the expectation similarly.

$$
\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{t} \middle| W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}\right] \\
=&\Pr\left(\boldsymbol{W}_{t-m:t} = \mathbf{0}_{m+1}\right) \cdot \left\{-Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{0}_{m+1})(\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1)\right\} \\
&+ \Pr\left(\boldsymbol{W}_{t-m:t} \neq \mathbf{0}_{m+1}\right) \cdot \{-Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{0}_{m+1})(0-1)\} \\
=&\prod_{j=1}^{J-J^\circ} \bar{q}_{u_j} \cdot \left\{-Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{0}_{m+1})(\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1)\right\}
\end{aligned}
$$

$$
\begin{aligned}
&+ (1 - \prod_{j=1}^{J-J^\circ} \bar{q}_{u_j}) \cdot \{-Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1})\} \\
&= - Y_t(\mathbf{1}_{m+1}) - (\prod_{j=1}^{J^\circ} \frac{1}{\bar{q}_{u_j^\circ}} - 1) Y_t(\mathbf{0}_{m+1})
\end{aligned}
$$

and again, similarly,

$$
\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_{t'} \,\middle|\, W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}\right] \\
&= \Pr\left(\boldsymbol{W}_{t'-m:t'} = \mathbf{0}_{m+1}\right) \cdot \left\{ -Y_{t'}(\mathbf{1}_{m+1}) - Y_{t'}(\mathbf{0}_{m+1}) (\prod_{j=1}^{J'} \frac{1}{\bar{q}_{u_j'}} - 1) \right\} \\
&\quad + \Pr\left(\boldsymbol{W}_{t'-m:t'} \neq \mathbf{0}_{m+1}\right) \cdot \{-Y_{t'}(\mathbf{1}_{m+1}) - Y_{t'}(\mathbf{0}_{m+1})(0 - 1)\} \\
&= \prod_{j=J^\circ+1}^{J'} \bar{q}_{u_j'} \cdot \left\{ -Y_{t'}(\mathbf{1}_{m+1}) - Y_{t'}(\mathbf{0}_{m+1}) (\prod_{j=1}^{J'} \frac{1}{\bar{q}_{u_j'}} - 1) \right\} \\
&\quad + (1 - \prod_{j=J^\circ+1}^{J'} \bar{q}_{u_j'}) \cdot \{-Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1})\} \\
&= - Y_{t'}(\mathbf{1}_{m+1}) - (\prod_{j=1}^{J^\circ} \frac{1}{\bar{q}_{u_j^\circ}} - 1) Y_{t'}(\mathbf{0}_{m+1})
\end{aligned}
$$

**(3)** With probability $1 - 2 \cdot (1/2^{J^\circ})$, $(W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}) \neq \mathbf{1}_{J^\circ}$ or $\mathbf{0}_{J^\circ}$. In this case

$$
\begin{aligned}
\mathbb{E}\left[\mathbb{1}_{t} \,\middle|\, W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}\right] &= - Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \\
\mathbb{E}\left[\mathbb{1}_{t'} \,\middle|\, W_{u_1^\circ}, ..., W_{u_{J^\circ}^\circ}\right] &= - Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1})
\end{aligned}
$$

Finally, putting all above together into (D.7), we have

$$
\begin{aligned}
&\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] \\
&= \prod_{j=1}^{J^\circ} q_{u_j^\circ} \cdot \left\{ (\prod_{j=1}^{J^\circ} \frac{1}{q_{u_j^\circ}} - 1) Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1}) \right\} \cdot \left\{ (\prod_{j=1}^{J^\circ} \frac{1}{q_{u_j^\circ}} - 1) Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1}) \right\} \\
&\quad + \prod_{j=1}^{J^\circ} \bar{q}_{u_j^\circ} \cdot \left\{ -Y_t(\mathbf{1}_{m+1}) - (\prod_{j=1}^{J^\circ} \frac{1}{\bar{q}_{u_j^\circ}} - 1) Y_t(\mathbf{0}_{m+1}) \right\} \cdot \left\{ -Y_{t'}(\mathbf{1}_{m+1}) - (\prod_{j=1}^{J^\circ} \frac{1}{\bar{q}_{u_j^\circ}} - 1) Y_{t'}(\mathbf{0}_{m+1}) \right\}
\end{aligned}
$$

$$+ \left\{1 - \prod_{j=1}^{J^\circ} q_{u_j^\circ} - \prod_{j=1}^{J^\circ} \bar{q}_{u_j^\circ}\right\} \cdot \{-Y_t(\mathbf{1}_{m+1}) + Y_t(\mathbf{0}_{m+1})\} \cdot \{-Y_{t'}(\mathbf{1}_{m+1}) + Y_{t'}(\mathbf{0}_{m+1})\}$$

$$=(\prod_{j=1}^{J^\circ} \frac{1}{q_{u_j^\circ}} - 1) Y_t(\mathbf{1}_{m+1}) Y_{t'}(\mathbf{1}_{m+1}) + Y_t(\mathbf{1}_{m+1}) Y_{t'}(\mathbf{0}_{m+1})$$

$$+ Y_t(\mathbf{0}_{m+1}) Y_{t'}(\mathbf{1}_{m+1}) + (\prod_{j=1}^{J^\circ} \frac{1}{\bar{q}_{u_j^\circ}} - 1) Y_t(\mathbf{0}_{m+1}) Y_{t'}(\mathbf{0}_{m+1})$$

which finishes the proof. □

### D.3.3 Lemma 5.3: Adversarial Selection of Potential Outcomes

In this section, we first prove Lemma 5.3, and then discuss the implications of Lemma 5.3.

**Proof of Lemma 5.3.**

The proof of Lemma 5.3 is through careful expansion of the risk function, the expected square loss.

*Proof.* Proof of Lemma 5.3. From Lemma D.3 and Lemma D.4, all the terms are quadratic, and all the coefficients are non-negative. After multiplying the constant $(T-m)^2$, we can expand, for any design of experiment $(\mathbb{T}, \mathbb{Q})$ and any potential outcomes $\mathbb{Y} \in \mathcal{Y}$, the following terms:

$$(T-m)^2 \cdot \mathbb{E}\left[(\hat{\tau}_m - \tau_m)^2\right]$$

$$= \sum_{t=m+1}^{T} \left\{ \left(\prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1\right) Y_t(\mathbf{1}_{m+1})^2 + 2Y_t(\mathbf{1}_{m+1})Y_t(\mathbf{0}_{m+1}) + \left(\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1\right) Y_t(\mathbf{0}_{m+1})^2 \right\}$$

$$+ \sum_{\substack{m+1 \le t < t' \le T \\ |O_{\mathbb{T}}(t,t')| \ge 1}} \left\{ \left(\prod_{j=1}^{J^\circ} \frac{1}{q_{u_j^\circ}} - 1\right) Y_t(\mathbf{1}_{m+1}) Y_{t'}(\mathbf{1}_{m+1}) + Y_t(\mathbf{1}_{m+1}) Y_{t'}(\mathbf{0}_{m+1}) \right.$$

$$\left. + Y_t(\mathbf{0}_{m+1}) Y_{t'}(\mathbf{1}_{m+1}) + \left(\prod_{j=1}^{J^\circ} \frac{1}{\bar{q}_{u_j^\circ}} - 1\right) Y_t(\mathbf{0}_{m+1}) Y_{t'}(\mathbf{0}_{m+1}) \right\}$$

where the equality is due to Lemma D.3 and Lemma D.4. Notice that in the first summation, all the coefficients in the front of $Y_t(\mathbf{1}_{m+1})^2$, $Y_t(\mathbf{1}_{m+1})Y_t(\mathbf{0}_{m+1})$, and $Y_t(\mathbf{0}_{m+1})^2$ are strictly positive, because $q_{u_j}$ are strictly between $(0,1)$. In the second summation, for those periods such that $|O_{\mathbb{T}}(t,t')| \geq 1$, all the coefficients in the front of $Y_t(\mathbf{1}_{m+1})Y_{t'}(\mathbf{1}_{m+1})$, $Y_t(\mathbf{1}_{m+1})Y_{t'}(\mathbf{0}_{m+1})$, $Y_t(\mathbf{0}_{m+1})Y_{t'}(\mathbf{1}_{m+1})$, and $Y_t(\mathbf{0}_{m+1})Y_{t'}(\mathbf{0}_{m+1})$ are strictly positive as well, because $q_{u_j}$ are strictly between $(0,1)$.

For the squared terms in the above expression, $Y_t(\mathbf{1}_{m+1})^2 \leq B^2, Y_t(\mathbf{0}_{m+1})^2 \leq B^2$ for any $t \in \{m+1 : T\}$. This is because $f(y) = y^2$ attains maximum at the end points of the interval $[-B, B]$. For the cross-product terms in the above expression, no matter if $(y_1, y_2)$ takes $(Y_t(\mathbf{1}_{m+1}), Y_t(\mathbf{0}_{m+1}))$, $(Y_t(\mathbf{1}_{m+1}), Y_{t'}(\mathbf{1}_{m+1}))$, $(Y_t(\mathbf{1}_{m+1}), Y_{t'}(\mathbf{0}_{m+1}))$, $(Y_t(\mathbf{0}_{m+1}), Y_{t'}(\mathbf{1}_{m+1}))$, or $(Y_t(\mathbf{0}_{m+1}), Y_{t'}(\mathbf{0}_{m+1}))\}$, we have that $y_1 \cdot y_2 \leq (y_1^2 + y_2^2)/2 \leq B^2$ where the first inequality is due to Cauchy-Schwarz, and the second inequality is due to convexity. Combining that fact that all coefficients are positive, $r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) \leq r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}^+) = r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}^-)$.

Moreover, for any $\mathbb{Y} \in \mathcal{Y}$ such that $\mathbb{Y} \neq \mathbb{Y}^+$ or $\mathbb{Y}^-$, if $\exists\ t \in \{m+1, ..., T\}$ such that $-B < Y_t(\mathbf{1}_{m+1}) < B$. Then from inequality (D.1), $\prod_{j=1}^{J} \frac{1}{q_{u_j}} - 1 > 0$, so the inequality is strict. Similarly, if $\exists t \in \{m+1, ..., T\}$ such that $-B < Y_t(\mathbf{0}_{m+1}) < B$, then combine $\prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} - 1 > 0$, so the inequality is strict. $\square$

**Implications of Lemma 5.3.**

Lemma 5.3 simplifies the minimax problem in (5.6). Instead of thinking it as a minimax problem, we can now replace $\mathbb{Y}$ by either $\mathbb{Y}^+$ or $\mathbb{Y}^-$, and solve only a minimization problem.

Here we state Lemma D.5 that is a direct implication of Lemma 5.3. It will be frequently used later on.

*Lemma* D.5. *When $\mathbb{Y} = \mathbb{Y}^+$ or $\mathbb{Y} = \mathbb{Y}^-$, under Assumptions 5.1–5.3, for any $t \in [T]$,*

$$\mathbb{E}[\mathbb{1}_t^2] = \left( \frac{1}{\prod_{j=1}^{J} q_{u_j}} + \frac{1}{\prod_{j=1}^{J} \bar{q}_{u_j}} \right) B^2.$$

*For any $t < t' \in [T]$, when $|O_{\mathbb{T}}(t,t')| = J^{\circ} = 0$,*

$$\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0$$

*When $|O_{\mathbb{T}}(t,t')| = J^{\circ} \geq 1$,*

$$\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = \left( \frac{1}{\prod_{j=1}^{J^{\circ}} q_{u_j^{\circ}}} + \frac{1}{\prod_{j=1}^{J^{\circ}} \bar{q}_{u_j^{\circ}}} \right) B^2$$

*Proof.* Proof of Lemma D.5. Replace $Y_t(\mathbf{1}_{m+1}) = Y_t(\mathbf{0}_{m+1})$ by $B$ or $-B$ into the expressions in Lemmas D.3 and D.4. □

### D.3.4 Theorem 5.4: Optimality of Fair Coin Flipping

In this section, we first prove Theorem 5.4, and then discuss the implications of Theorem 5.4.

**Proof of Theorem 5.4.**

The proof of Theorem 5.4 is through an elegant inequality that highlights the balance between treatment probabilities and control probabilities.

*Proof.* Proof of Theorem 5.4. Similar to the proof of Lemma 5.3, we expand the quadratic terms using Lemma D.5. After multiplying the constant $(T-m)^2$, we can expand, for any design of experiment $(\mathbb{T}, \mathbb{Q})$ and any potential outcomes $\mathbb{Y} \in \mathcal{Y}$, the following terms:

$$\begin{aligned}
&(T-m)^2 \cdot \mathbb{E}\left[(\hat{\tau}_m - \tau_m)^2\right] = \\
&\qquad \sum_{t=m+1}^{T} \left( \prod_{j=1}^{J} \frac{1}{q_{u_j}} + \prod_{j=1}^{J} \frac{1}{\bar{q}_{u_j}} \right) \cdot B^2 + \sum_{\substack{m+1 \leq t < t' \leq T \\ |O_{\mathbb{T}}(t,t')| \geq 1}} \left( \prod_{j=1}^{J^{\circ}} \frac{1}{q_{u_j^{\circ}}} + \prod_{j=1}^{J^{\circ}} \frac{1}{\bar{q}_{u_j^{\circ}}} \right) \cdot B^2
\end{aligned}$$

For each of them, due to Lemma D.2, the minimum is obtained at $q_0 = q_1 = ... = q_K = 1/2$. □

**Implications of Theorem 5.4.**

Theorem 5.4 further simplifies the minimax problem in (5.6). Now that we have identified the optimal randomization probabilities, we can directly plug in the optimal probabilities being 1/2. Here we state Lemma D.6 that is a combination of Lemma D.5 and Theorem 5.4. It will be frequently used later on.

*Lemma* D.6. *Under Assumptions 5.1–5.3, when $\mathbb{Y} = \mathbb{Y}^+$ or $\mathbb{Y} = \mathbb{Y}^-$, and when $q_0 = q_1 = ... = q_K = 1/2$, for any $t \in [T]$,*

$$\mathbb{E}[\mathbb{1}_t^2] = 2^{J+1} B^2.$$

*For any $t < t' \in [T]$, when $|O_{\mathbb{T}}(t, t')| = J^\circ = 0$,*

$$\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0$$

*When $|O_{\mathbb{T}}(t, t')| = J^\circ \geq 1$,*

$$\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 2^{J^\circ+1} B^2$$

*Proof.* Proof of Lemma D.6. Simply replace $q_0 = q_1 = ... = q_K = 1/2$ into Lemma D.5. □

### D.3.5 Structural Results of the Optimal Design

Using Lemma 5.3, we now establish two structural results that further characterize the class of optimal designs of regular switchback experiments. Lemma D.7 states the optimal starting and ending structure; Lemma D.8 states the optimal middle-case structure. The proofs to Lemma D.7 and Lemma D.8 are deferred to Sections D.3.5 and D.3.5, respectively.

*Lemma* D.7. *When $\mathbb{Y} = \mathbb{Y}^+$ or $\mathbb{Y} = \mathbb{Y}^-$, under Assumptions 5.1–5.3, any optimal* *design of experiment* $\mathbb{T}$ *must satisfy*

$$t_1 \geq m+2, \qquad \text{and} \qquad t_K \leq T-m.$$

Lemma D.7 states that the first randomization point on period 1 should be followed by at least $m$ periods that do not flip a coin, and that the last randomization point should be followed by at least $m$ periods that do not flip a coin. This guarantees that the assignments during $\{1 : m+1\}$ and during $\{T-m : T\}$ both produce observed data that can be used to estimate the lag-$m$ effect.

*Lemma* D.8. *When* $\mathbb{Y} = \mathbb{Y}^+$ *or* $\mathbb{Y} = \mathbb{Y}^-$, *under Assumptions 5.1–5.3, any optimal design of experiment* $\mathbb{T}$ *must satisfy*

$$t_{k+1} - t_{k-1} \geq m, \ \forall k \in [K].$$

Lemma D.8 suggests that in every consecutive $m+1$ periods, there could be at most 3 randomization points. Intuitively, too many randomization points in every consecutive $m+1$ periods decreases the chance of observing a useful assignment path of $\mathbf{1}_{m+1}$ or $\mathbf{0}_{m+1}$. Lemma D.8 formalizes such intuition, and suggests that as the persistence of the carryover effect increases, the optimal design randomizes less often.

Lemmas D.7 and D.8 restrict the space of possible optimal regular switchback experiment to a smaller class of switchback experiments. Under such a smaller class of switchback experiments, we can explicitly express the risk function in closed form, which we define below.

*Lemma* D.9 (Risk Function). *When* $\mathbb{Y} = \mathbb{Y}^+$ *or* $\mathbb{Y} = \mathbb{Y}^-$, *under Assumptions 5.1–5.3, as long as the following three conditions are satisfied,*

$$t_1 \geq m+2; \qquad t_K \leq T-m; \qquad t_{k+1} - t_{k-1} \geq m, \ \forall k \in [K],$$

*the risk function for any switchback experiment is given by*

$$r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) = \frac{1}{(T-m)^2}\left\{4\sum_{k=1}^{K+1}(t_k - t_{k-1})^2 + 8m(t_K - t_1) + 4m^2K - 4m^2 \right.$$
$$\left. +4\sum_{k=2}^{K}[(m - t_k + t_{k-1})^+]^2\right\} B^2 \quad \text{(D.8)}$$

Lemma D.9 explicitly describes the risk function of any optimal design of regular switchback experiments, which lies in the optimal sub-class of switchback experiments. The proof of Lemma D.9 is deferred to Section D.3.5 in the appendix.

To understand the risk function in Lemma D.9, we separately examine each term in (D.8). The first summation of the squares $\sum_{k=1}^{K+1}(t_k - t_{k-1})^2$ suggests that the gap between two consecutive randomization points should not be too large. The middle term $8m(t_K - t_1)$ formalizes Lemma D.7, suggesting that the second randomization point on period $t_1$ should not be too early and the last randomization point on period $t_K$ should not be too late. The last summation of the squares $\sum_{k=2}^{K}[(m-t_k+t_{k-1})^+]^2$ suggests that the gap should not be too small. Equation D.8 formalizes the trade-off that we have described earlier in this section. First note that when we focus on the optimal design, we treat $T$ and $m$ both as constants. So the constant of $1/(T-m)$ in the expression of the risk function does not affect the optimal design.

**Proof of Lemma D.7.**

*Proof.* Proof of Lemma D.7.

We prove the two parts separately, both by contradiction.

**(1)** Suppose there exists an optimal design $\mathbb{T} = \{t_0 = 1, t_1, t_2, ..., t_K\}$ such that $t_1 \leq m+1$. Then we try to construct another design $\tilde{\mathbb{T}}$, such that $\left|\tilde{\mathbb{T}}\right| = K = |\mathbb{T}| - 1$. And the $K$ elements are $\tilde{\mathbb{T}} = \{\tilde{t}_0 = 1, \tilde{t}_1 = t_2, \tilde{t}_2 = t_3, ..., \tilde{t}_{K-1} = t_K\}$.

Next we argue that when $\mathbb{Y} = \mathbb{Y}^+$ or $\mathbb{Y} = \mathbb{Y}^-$,

$$r(\mathbb{T}, \mathbb{Y}) > r(\tilde{\mathbb{T}}, \mathbb{Y}),$$

which suggests that $\mathbb{T}$ is not the optimal design.

First, focus on the squared terms. For any $m+1 \leq t \leq t_1+m-1$, $t_1 \in f_{\mathbb{T}}^m(t), t_1 \neq$

Table D.2: An example of two regular switchback experiments $\mathbb{T}$ and $\tilde{\mathbb{T}}$ when $m = 4$ and $t_1 = 3$

| | 1 | 2 | 3 | 4 | 5 | 6 | ... |
|---|---|---|---|---|---|---|---|
| $\mathbb{T}$ | ✓ | − | ✓ | − | − | ✓ | ... |
| $\tilde{\mathbb{T}}$ | ✓ | − | − | − | − | ✓ | ... |

Each checkmark beneath a number indicates that this number is within that set; and each dash beneath a number indicates that this number is not within that set. For example, the checkmark ✓ beneath number 3 indicates that $3 \in \mathbb{T}$; and the dash − beneath number 3 indicates that $3 \neq \tilde{\mathbb{T}}$.

$f^m_{\tilde{\mathbb{T}}}(t)$. Moreover, $t - m \leq t_1 - 1$, so that $t_0 \in f^m_{\tilde{\mathbb{T}}}(t)$. So $f^m_{\mathbb{T}}(t) - \{t_1\} = f^m_{\tilde{\mathbb{T}}}(t)$, and $\left| f^m_{\tilde{\mathbb{T}}}(t) \right| \geq 1$. As a result,

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] \geq (2^{2+1} - 2^{1+1})B^2 = 4B^2.$$

For any $t \geq t_1 + m$, either (i) $f_{\mathbb{T}}(t - m) = t_1$, in which case $f_{\tilde{\mathbb{T}}}(t - m) = t_0$. This is the only difference between $f^m_{\mathbb{T}}(t)$ and $f^m_{\tilde{\mathbb{T}}}(t)$, i.e., $f^m_{\mathbb{T}}(t) - \{t_1\} = f^m_{\tilde{\mathbb{T}}}(t) - \{t_0\}$. So $|f^m_{\mathbb{T}}(t)| = \left| f^m_{\tilde{\mathbb{T}}}(t) \right|$. The second case is (ii) $f_{\mathbb{T}}(t-m) \geq t_2$, in which case $f^m_{\mathbb{T}}(t) = f^m_{\tilde{\mathbb{T}}}(t)$. Both cases suggest that

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] = 0.$$

So we have

$$\begin{aligned}
&\sum_{t=m+1}^{T} \mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \sum_{t=m+1}^{T} \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right] \\
=&\sum_{t=m+1}^{t_1+m-1} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) + \sum_{t=t_1+m}^{T} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) \\
\geq&\sum_{t=m+1}^{t_1+m-1} (4B^2) + 0 \\
=&4(t_1 - 1)B^2 \\
>&0
\end{aligned}$$

Second, focus on the cross product terms. For any $t$ and $t'$ such that $m + 1 \leq t < t' \leq t_1 + m - 1$, $t_1 \in O_{\mathbb{T}}(t, t'), t_1 \neq O_{\tilde{\mathbb{T}}}(t, t')$. Moreover, $t - m \leq t_1 - 1$, so that

$t_0 \in O_{\mathbb{T}}(t,t')$. So $O_{\mathbb{T}}(t,t') - \{t_1\} = O_{\tilde{\mathbb{T}}}(t,t')$, and $|O_{\tilde{\mathbb{T}}}(t,t')| \geq 1$. As a result,

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] \geq (2^{2+1} - 2^{1+1})B^2 = 4B^2 > 0.$$

For any $m+1 \leq t < t' \leq T$ such that $t' \geq t_1 + m$, either (i) $f_{\mathbb{T}}(t'-m) = t_1$, in which case $f_{\tilde{\mathbb{T}}}(t'-m) = t_0$. So $O_{\mathbb{T}}(t,t') - \{t_1\} = O_{\tilde{\mathbb{T}}}(t,t') - \{t_0\}$. So $|O_{\mathbb{T}}(t,t')| = |O_{\tilde{\mathbb{T}}}(t,t')|$. The second case is (ii) $f_{\mathbb{T}}(t'-m) \geq t_2$, in which case $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t')$. Both cases suggest that

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] = 0.$$

So we have

$$\begin{aligned}
&\sum_{m+1 \leq t < t' \leq T} \mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \sum_{m+1 \leq t < t' \leq T} \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right] \\
=&\sum_{m+1 \leq t < t' \leq t_1+m-1} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right]\right) \\
&\qquad + \sum_{\substack{m+1 \leq t < t' \leq T \\ t' \geq t_1+m}} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right]\right) \\
\geq& 0
\end{aligned}$$

Combine both square terms and cross-product terms we know that

$$r(\mathbb{T}, \mathbb{Y}) > r(\tilde{\mathbb{T}}, \mathbb{Y}).$$

**(2)** Suppose there exists an optimal design $\mathbb{T} = \{t_0 = 1, t_1, t_2, ..., t_K\}$ such that $t_K \geq T - m + 1$. Then we try to construct another design $\tilde{\mathbb{T}}$, such that $\left|\tilde{\mathbb{T}}\right| = K = |\mathbb{T}| - 1$. And the $K$ elements are $\tilde{\mathbb{T}} = \{\tilde{t}_0 = 1, \tilde{t}_1 = t_1, \tilde{t}_2 = t_2, ..., \tilde{t}_{K-1} = t_{K-1}\}$.

Next we argue that when $\mathbb{Y} = \mathbb{Y}^+$ or $\mathbb{Y} = \mathbb{Y}^-$,

$$r(\mathbb{T}, \mathbb{Y}) > r(\tilde{\mathbb{T}}, \mathbb{Y}),$$

which suggests that $\mathbb{T}$ is not the optimal design.

Table D.3: An example of two regular switchback experiments $\mathbb{T}$ and $\tilde{\mathbb{T}}$ when $m = 4$ and $t_K = T - 2$

| | ... | $T-5$ | $T-4$ | $T-3$ | $T-2$ | $T-1$ | $T$ |
|---|---|---|---|---|---|---|---|
| $\mathbb{T}$ | ... | ✓ | – | ✓ | ✓ | – | – |
| $\tilde{\mathbb{T}}$ | ... | ✓ | – | ✓ | – | – | – |

Note: Each checkmark beneath a number indicates that this number is within that set; and each dash beneath a number indicates that this number is not within that set. For example, the checkmark ✓ beneath number $T-2$ indicates that $T-2 \in \mathbb{T}$; and the dash – beneath number $T-2$ indicates that $T-2 \neq \tilde{\mathbb{T}}$.

First focus on the squared terms. For any $m+1 \leq t \leq t_K - 1$, $f_{\mathbb{T}}^m(t) = f_{\tilde{\mathbb{T}}}^m(t)$ is totally unchanged.

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] = 0.$$

For any $t_K \leq t \leq T$, $t_K \notin f_{\tilde{\mathbb{T}}}^m(t), t_K \in f_{\mathbb{T}}^m(t)$. And all the other determining randomization points are unchanged. So $f_{\tilde{\mathbb{T}}}^m(t) \subset f_{\mathbb{T}}^m(t)$ and $f_{\mathbb{T}}^m(t) - \{t_K\} = f_{\tilde{\mathbb{T}}}^m(t)$ and $\left| f_{\tilde{\mathbb{T}}}^m(t) \right| \geq 1$.

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] \geq (2^{2+1} - 2^{1+1})B^2 = 4B^2.$$

So we have

$$\begin{aligned}
&\sum_{t=m+1}^{T} \mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \sum_{t=m+1}^{T} \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right] \\
=&\sum_{t=m+1}^{t_K-1} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) + \sum_{t=t_K}^{T} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) \\
\geq&\sum_{t=t_K}^{T} (4B^2) + 0 \\
=&4(T - t_K + 1)B^2 \\
>&0
\end{aligned}$$

Next we focus on the cross-product terms. For any $m+1 \leq t < t' \leq T$ such that

$t \le t_K - 1$, $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t')$ is totally unchanged.

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] = 0.$$

For any $t_K \le t < t' \le T$, since $t' - m \le T - m \le t_K - 1$, so $f_{\tilde{\mathbb{T}}}(t'-m) < t_K$ and $|O_{\tilde{\mathbb{T}}}(t,t')| \ge 1$ must contain an element. Moreover, $O_{\tilde{\mathbb{T}}}(t,t') \subset O_{\mathbb{T}}(t,t')$. So

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] \ge (2^{2+1} - 2^{1+1})B^2 \ge 4B^2 > 0.$$

So we have

$$\begin{aligned}
&\sum_{m+1\le t<t'\le T} \mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \sum_{m+1\le t<t'\le T} \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right] \\
=&\sum_{\substack{m+1\le t<t'\le T\\ t\le t_K-1}} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right]\right) + \sum_{t_K\le t<t'\le T} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right]\right) \\
\ge& 0
\end{aligned}$$

Combine both square terms and cross-product terms we know that

$$r(\mathbb{T}, \mathbb{Y}) > r(\tilde{\mathbb{T}}, \mathbb{Y}).$$

□

**Proof of Lemma D.8.**

*Proof.* Proof of Lemma D.8.

Recall that we denote $t_0 = 1$ and $t_{K+1} = T + 1$. First, from Lemma D.7, $t_1 \ge m+2, t_K \le T - m$. So $k = 1$ and $k = K$ cases both hold. Next, when $2 \le k \le K-1$, we prove by contradiction.

Suppose there exists some optimal design $\mathbb{T}$, such that $\exists 2 \le k \le K-1, s.t.\ t_{k+1} - t_{k-1} \le m - 1$. Denote

$$\mathbb{K} = \{k \in \{2 : K-1\}\,|t_{k+1} - t_{k-1} \le m-1\}.$$

Since $\mathbb{K} \neq \emptyset$, pick $j = \max \mathbb{K}$ to be the largest element in $\mathbb{K}$. Apparently $j \leq K-1$ since $j \in \{2 : K-1\}$. We also know that $t_{j+2} \geq t_j + m$, because otherwise $j+1 \in \mathbb{K}$, which contradicts the maximality of $j$.

We now construct another design $\tilde{\mathbb{T}}$ such that $\left|\tilde{\mathbb{T}}\right| = K = |\mathbb{T}| - 1$, and the $K$ elements are $\tilde{\mathbb{T}} = \{\tilde{t}_0 = 1, \tilde{t}_1 = t_1, ..., \tilde{t}_{j-1} = t_{j-1}, \tilde{t}_j = t_{j+1}, ..., \tilde{t}_{K-1} = t_K\}$.

Table D.4: An example of two regular switchback experiments $\mathbb{T}$ and $\tilde{\mathbb{T}}$ when $m = 4$ and $t_j = t_{j+1} - 1 = t_{j-1} + 2$

| | ... | $t_{j-1}$ | $t_{j-1}+1$ | $t_j$ | $t_{j+1}$ | $t_{j+1}+1$ | $t_{j+1}+2$ | $t_{j+2}$ | ... |
|---|---|---|---|---|---|---|---|---|---|
| $\mathbb{T}$ | ... | ✓ | $-$ | ✓ | ✓ | $-$ | $-$ | ✓ | ... |
| $\tilde{\mathbb{T}}$ | ... | ✓ | $-$ | $-$ | ✓ | $-$ | $-$ | ✓ | ... |

Each checkmark beneath a number indicates that this number is within that set; and each dash beneath a number indicates that this number is not within that set. For example, the checkmark ✓ beneath number $t_j$ indicates that $t_j \in \mathbb{T}$; and the dash $-$ beneath number $t_j$ indicates that $t_j \neq \tilde{\mathbb{T}}$.

Next we argue that when $\mathbb{Y} = \mathbb{Y}^+$ or $\mathbb{Y} = \mathbb{Y}^-$,

$$r(\mathbb{T}, \mathbb{Y}) > r(\tilde{\mathbb{T}}, \mathbb{Y}),$$

which suggests that $\mathbb{T}$ is not the optimal design.

First focus on the squared terms. When $t \leq t_j - 1$, $f^m_{\mathbb{T}}(t) = f^m_{\tilde{\mathbb{T}}}(t)$ is totally unchanged.

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] = 0.$$

When $t_j \leq t \leq t_j + m - 1$, this suggests that $t - m \leq t_J - 1$ so that $f_{\tilde{\mathbb{T}}} \leq t_j - 1$. So $t_j \notin f^m_{\tilde{\mathbb{T}}}(t), t_j \in f^m_{\mathbb{T}}(t)$. And all the other determining randomization points are unchanged. So $f^m_{\tilde{\mathbb{T}}}(t) \subset f^m_{\mathbb{T}}(t)$ and $f^m_{\mathbb{T}}(t) - \{t_j\} = f^m_{\tilde{\mathbb{T}}}(t)$ and $\left|f^m_{\tilde{\mathbb{T}}}(t)\right| \geq 1$.

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] \geq (2^{2+1} - 2^{1+1})B^2 = 4B^2.$$

When $t_j + m \leq t \leq T$, either (i) $f_{\mathbb{T}}(t-m) = t_j$, in which case $f_{\tilde{\mathbb{T}}}(t-m) = t_{j-1}$. This is the only difference between $f^m_{\mathbb{T}}(t)$ and $f^m_{\tilde{\mathbb{T}}}(t)$, i.e., $f^m_{\mathbb{T}}(t) - \{t_j\} = f^m_{\tilde{\mathbb{T}}}(t) - \{t_{j-1}\}$. So $|f^m_{\mathbb{T}}(t)| = \left|f^m_{\tilde{\mathbb{T}}}(t)\right|$. The second case is (ii) $f_{\mathbb{T}}(t-m) \geq t_{j+1}$, in which case $f^m_{\mathbb{T}}(t) =$

$f^m_{\tilde{\mathbb{T}}}(t)$. Both cases suggest that

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})^2] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})^2] = 0.$$

So we have

$$\begin{aligned}
&\sum_{t=m+1}^{T} \mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \sum_{t=m+1}^{T} \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right] \\
=&\sum_{t=m+1}^{t_j-1} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) + \sum_{t=t_j}^{t_j+m-1} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) \\
&\qquad + \sum_{t=t_j+m}^{T} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})^2\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})^2\right]\right) \\
\geq& 0 + \sum_{t=t_j}^{t_j+m-1} (4B^2) + 0 \\
=& 4(m-1)B^2 \\
>& 0
\end{aligned}$$

Next we focus on the cross-product terms. Let $m+1 \leq t < t' \leq T$. There are many cases which we summarize in Table D.5

Table D.5: Summary of the differences between cross-product terms under two regular switchback experiments $\mathbb{T}$ and $\tilde{\mathbb{T}}$

| | $\mathbb{T}$ | $\tilde{\mathbb{T}}$ |
|---|---|---|
| $m+1 \leq t \leq t_{j-1}, t < t' \leq T$ | unchanged | |
| $t_{j-1} \leq t \leq t_j - 1, t < t' \leq t_j + m - 1$ | unchanged | |
| $t_{j-1} \leq t \leq t_j - 1, t_j + m \leq t' \leq t_{j+1} + m - 1$ | 0 | $4B^2$ |
| $t_{j-1} \leq t \leq t_j - 1, t_{j+1} + m \leq t' \leq T$ | unchanged | |
| $t_j \leq t < t' \leq t_j + m - 1$ | $2^{\lvert O_{\mathbb{T}}(t,t')\rvert+1}B^2$ | $2^{\left\lvert O_{\tilde{\mathbb{T}}}(t,t')\right\rvert+1}B^2$ |
| $t_j \leq t \leq t_j + m - 1, t_j + m \leq t' \leq T$ | unchanged | |
| $t_j + m \leq t < t' \leq T$ | unchanged | |

We explain Table D.5. When $m+1 \leq t \leq t_{j-1}, t < t' \leq T$, all the overlapping randomization points are earlier than $t_{j-1} - 1$, i.e., $\forall a \in O_{\mathbb{T}}(t,t'), a \leq t_{j-1} - 1; \forall a \in O_{\tilde{\mathbb{T}}}(t,t'), a \leq t_{j-1} - 1$. So $t_j \notin O_{\mathbb{T}}(t,t')$, and the overlapping randomization points are unchanged, i.e., $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t')$.

When $t_{j-1} \le t \le t_j - 1, t < t' \le t_j + m - 1$, all the overlapping randomization points are earlier than $t_{j-1}$, i.e., $\forall a \in O_{\mathbb{T}}(t,t'), a \le t_{j-1}; \forall a \in O_{\tilde{\mathbb{T}}}(t,t'), a \le t_{j-1}$. So $t_j \notin O_{\mathbb{T}}(t,t')$, and the overlapping randomization points are unchanged, i.e., $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t')$.

When $t_{j-1} \le t \le t_j - 1, t_j + m \le t' \le t_{j+1} + m - 1$, changing from $\mathbb{T}$ to $\tilde{\mathbb{T}}$ increases the expected values. This is because $t' - m \ge t_j > t$. So first, $O_{\mathbb{T}}(t,t') = \emptyset$. But $f_{\tilde{\mathbb{T}}}(t'-m) = t_{j-1}$ and $t_{j-1} \in f^m_{\tilde{\mathbb{T}}}(t)$, which suggests that $t_{j-1} \in O_{\tilde{\mathbb{T}}}(t,t')$. Also, $\forall a \in f^m_{\tilde{\mathbb{T}}}(t'), a \ge t_{j-1}; \forall a \in f^m_{\mathbb{T}}(t), a \le t_{j-1}$, which suggests that $t_{j-1}$ is the only overlapping element. So, $O_{\tilde{\mathbb{T}}}(t,t') = \{t_{j-1}\}$. In this case,

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] = (0 - 2^{1+1})B^2 = -4B^2.$$

When $t_{j-1} \le t \le t_j - 1, t_{j+1} + m \le t' \le T$, since $t' - m \ge t_{j+1} > t_j > t$, $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t') = \emptyset$.

When $t_j \le t < t' \le t_j + m - 1$, $t_j \in O_{\mathbb{T}}(t,t')$ and $t_j \notin O_{\tilde{\mathbb{T}}}(t,t')$. And all the other overlapping randomization points are unchanged, so $O_{\mathbb{T}}(t,t') - \{t_j\} = O_{\tilde{\mathbb{T}}}(t,t')$ and $|O_{\tilde{\mathbb{T}}}(t,t')| \ge 1$. In this case,

$$\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] \ge (2^{2+1} - 2^{1+1})B^2 = 4B^2.$$

When $t_j \le t \le t_j + m - 1, t_j + m \le t' \le T$, either (i) $f^m_{\mathbb{T}}(t'-m) = t_j$, in which case $f_{\tilde{\mathbb{T}}}(t'-m) = t_{j-1}$. This is the only difference between $O_{\mathbb{T}}(t,t')$ and $O_{\tilde{\mathbb{T}}}(t,t')$, i.e., $O_{\mathbb{T}}(t,t') - \{t_j\} = O_{\tilde{\mathbb{T}}}(t,t') - \{t_{j-1}\}$. $|O_{\mathbb{T}}(t,t')| = |O_{\tilde{\mathbb{T}}}(t,t')|$. The second case is (ii) $f_{\mathbb{T}}(t'-m) \ge t_{j+1}$, in which case $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t')$ is unchanged. Both cases suggest that $\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] = 0$.

When $t_j + m \le t < t' \le T$, either (i) $f^m_{\mathbb{T}}(t'-m) = t_j$, in which case $f_{\tilde{\mathbb{T}}}(t'-m) = t_{j-1}$. This is the only difference between $O_{\mathbb{T}}(t,t')$ and $O_{\tilde{\mathbb{T}}}(t,t')$, i.e., $O_{\mathbb{T}}(t,t') - \{t_j\} = O_{\tilde{\mathbb{T}}}(t,t') - \{t_{j-1}\}$. $|O_{\mathbb{T}}(t,t')| = |O_{\tilde{\mathbb{T}}}(t,t')|$. The second case is (ii) $f_{\mathbb{T}}(t'-m) \ge t_{j+1}$, in which case $O_{\mathbb{T}}(t,t') = O_{\tilde{\mathbb{T}}}(t,t')$ is unchanged. Both cases suggest that $\mathbb{E}[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})] - \mathbb{E}[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})] = 0$.

So we have

$$
\begin{aligned}
&\sum_{m+1\le t<t'\le T} \mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \sum_{m+1\le t<t'\le T} \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right] \\
=& \sum_{\substack{t_{j-1}\le t\le t_j-1 \\ t_j+m\le t'\le t_{j+1}+m-1}} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right]\right) \\
&\qquad + \sum_{t_j\le t<t'\le t_j+m-1} \left(\mathbb{E}\left[\mathbb{1}_t(\mathbb{T})\mathbb{1}_{t'}(\mathbb{T})\right] - \mathbb{E}\left[\mathbb{1}_t(\tilde{\mathbb{T}})\mathbb{1}_{t'}(\tilde{\mathbb{T}})\right]\right) \\
\ge& \sum_{\substack{t_{j-1}\le t\le t_j-1 \\ t_j+m\le t'\le t_{j+1}+m-1}} \left(-4B^2\right) + \sum_{t_j\le t<t'\le t_j+m-1} \left(4B^2\right) \\
=& -(t_j-t_{j-1})(t_{j+1}-t_j)4B^2 + \frac{m(m-1)}{2}4B^2 \\
\ge& 0
\end{aligned}
$$

where the last inequality is because $j \in \mathbb{K}$, $t_{j+1}-t_{j-1} \le m-1$, so $(t_j-t_{j-1})(t_{j+1}-t_j) \le \frac{(m-1)^2}{4} \le \frac{m(m-1)}{2}$.

Combine both square terms and cross-product terms we know that

$$r(\mathbb{T}, \mathbb{Y}) > r(\tilde{\mathbb{T}}, \mathbb{Y}).$$

$\square$

**Proof of Lemma D.9.**

*Proof.* Proof of Lemma D.9.

Think of $\mathbb{E}[\mathbb{1}_t^2]$ as $\mathbb{E}[\mathbb{1}_t\mathbb{1}_t]$, so that $r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) = \sum_{t=m+1}^{T}\sum_{t'=m+1}^{T} \mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}]$. Then we can decompose the risk function to be

$$
\begin{aligned}
&(T-m)^2 \cdot r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) = \\
&\qquad \sum_{\substack{m+1\le t,t'\le T \\ \min\{t,t'\}\le t_1-1}} \mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}] + \sum_{k=1}^{K-1}\left(\sum_{\substack{t_k\le t,t'\le T \\ \min\{t,t'\}\le t_{k+1}-1}} \mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}]\right) + \sum_{t_K\le t,t'\le T} \mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}] \quad \text{(D.9)}
\end{aligned}
$$

The core of this proof is to carefully count how many values can each $\mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}], \forall t, t' \in \{m+1:T\}$ take. See Table D.6 for an illustration.

Table D.6: Illustrator of the different values of $\mathbb{E}[\mathbb{1}_t\mathbb{1}_t]$, when $T = 17, m = 4, \mathbb{T} = \{1, 6, 8, 13\}$

| | (1 2 3 4) | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (✓ – – –) | – | ✓ | – | ✓ | – | – | – | – | ✓ | – | – | – | – |
| – | | 4 | 4 | 4 | 4 | 4 | | | | | | | | |
| ✓ | | 4 | **8** | **8** | **8** | **8** | 4 | 4 | | | | | | |
| – | | 4 | **8** | **8** | **8** | **8** | 4 | 4 | | | | | | |
| ✓ | | 4 | **8** | **8** | **16** | **16** | **8** | **8** | 4 | 4 | 4 | 4 | 4 | |
| – | | 4 | **8** | **8** | **16** | **16** | **8** | **8** | 4 | 4 | 4 | 4 | 4 | |
| – | | | 4 | 4 | **8** | **8** | **8** | **8** | 4 | 4 | 4 | 4 | 4 | |
| – | | | 4 | 4 | **8** | **8** | **8** | **8** | 4 | 4 | 4 | 4 | 4 | |
| – | | | | | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | |
| ✓ | | | | | 4 | 4 | 4 | 4 | 4 | **8** | **8** | **8** | **8** | 4 |
| – | | | | | 4 | 4 | 4 | 4 | 4 | **8** | **8** | **8** | **8** | 4 |
| – | | | | | 4 | 4 | 4 | 4 | 4 | **8** | **8** | **8** | **8** | 4 |
| – | | | | | 4 | 4 | 4 | 4 | 4 | **8** | **8** | **8** | **8** | 4 |
| – | | | | | | | | | | 4 | 4 | 4 | 4 | 4 |

In the second line, each checkmark beneath number $t$ indicates that period $t \in \mathbb{T}$, i.e. there is a randomization point at period $t$. This table illustrates different values of $\mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}]$ when $t, t' \in \{m+1, T\}$, where the zero values are omitted. The $B^2$ magnitudes are also omitted.

First we calculate the first block from equation (D.9). Because $t_1 \geq m+2$, for any $t, t'$ such that $m+1 \leq \min\{t, t'\} \leq t_1 - 1$, $m+1 \leq \max\{t, t'\} \leq t_1 + m - 1$, we know that the only overlapping randomization point is $t_0$. So $\mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}] = 4B^2$. For any $t, t'$ such that $m+1 \leq \min\{t, t'\} \leq t_1 - 1$, $t_1 + m \leq \max\{t, t'\} \leq T$, there is no overlapping randomization point so $\mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}] = 0$.

$$\sum_{\substack{m+1 \leq t, t' \leq T \\ \min\{t, t'\} \leq t_1 - 1}} \mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}] = B^2 \left(4 \cdot ((t_1 - 1)^2 - m^2)\right)$$

Then we calculate the second block from equation (D.9). For any $k \in [K-1]$, consider $t_k - t_{k-1}$ and $t_{k+1} - t_k$, which jointly determine the values of $\mathbb{E}[\mathbb{1}_t\mathbb{1}_{t'}]$ for any

$t, t'$, such that $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_k \le \max\{t, t'\} \le T$. We will go over each of the four cases below.

**(1)** When $t_k - t_{k-1} \ge m, t_{k+1} - t_k \ge m$. Due to Lemma D.6, for all $t, t' \in \{t_k : t_k + m - 1\}$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 8B^2$, because both $t_{k-1} \le t - m \le t_k - 1$ and $t_{k-1} \le t' - m \le t_k - 1$, and both $t_{k-1}$ and $t_k$ are overlapping randomization points. For all $t, t'$ such that $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_k + m \le \max\{t, t'\} \le t_{k+1} + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 4B^2$, because $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_k \le \max\{t, t'\} - m \le t_{k+1} - 1$ so only $t_k$ is the overlapping randomization point. For all $t, t'$ such that $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_{k+1} + m \le \max\{t, t'\} \le T$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0$.

In this case,

$$\sum_{\substack{t_k \le t, t' \le T \\ \min\{t, t'\} \le t_{k+1} - 1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = B^2 \left(8 \cdot m^2 + 4 \cdot ((m + t_{k+1} - t_k)^2 - 2m^2)\right)$$

**(2)** When $t_k - t_{k-1} \ge m, t_{k+1} - t_k < m$. Due to Lemma D.6, for all t, t' such that $t_k \le \min\{t, t'\} \le t_{k+1} - 1$, $t_k \le \max\{t, t'\} \le t_k + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 8B^2$, because both $t, t' \le t_k + m - 1$, so $t_{k-1} \le t - m \le t_k - 1$ and $t_{k-1} \le t' - m \le t_k - 1$, and both $t_{k-1}$ and $t_k$ are overlapping randomization points. For all $t, t'$ such that $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_k + m \le \max\{t, t'\} \le t_{k+1} + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 4B^2$, because $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_k \le \max\{t, t'\} - m \le t_{k+1} - 1$ so only $t_k$ is the overlapping randomization point. For all $t, t'$ such that $t_k \le \min\{t, t'\} \le t_{k+1} - 1$ and $t_{k+1} + m \le \max\{t, t'\} \le T$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0$.

In this case,

$$\begin{aligned}\sum_{\substack{t_k \le t, t' \le T \\ \min\{t, t'\} \le t_{k+1} - 1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = B^2 \big(& 8 \cdot (m^2 - (m - t_{k+1} + t_k)^2) + \\ & 4 \cdot ((m + t_{k+1} - t_k)^2 - 2m^2 + (m - t_{k+1} - t_k)^2)\big)\end{aligned}$$

**(3)** When $t_k - t_{k-1} < m, t_{k+1} - t_k \ge m$. Due to Lemma D.6, for all $t, t' \in \{t_k : t_{k-1} + m - 1\}$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 16B^2$, because $t - m \le t_{k-1} - 1 \le t_k \le t$ and $t' - m \le t_{k-1} - 1 \le t_k \le t'$ so $t_{k-2}, t_{k-1}, t_k$ are three determining randomization points. Also $t_k - t_{k-2} \ge m$

so $t_{k-2} \leq \min\{t,t'\} - m$ and $t_{k-3}$ is not a determining randomization point. For all $t,t'$ such that $t_k \leq \min\{t,t'\} \leq t_k + m - 1, t_{k-1} + m \leq \max\{t,t'\} \leq t_k + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 8B^2$, because $\min\{t,t'\} - m \leq t_k - 1$ and $t_{k-1} \leq \max\{t,t'\} - m \leq t_k - 1$ so $t_{k-1}$ and $t_k$ are two determining randomization point. For all $t,t'$ such that $t_k \leq \min\{t,t'\} \leq t_{k+1} - 1, t_k + m \leq \max\{t,t'\} \leq t_{k+1} + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 4B^2$, because $t_k \leq \max\{t,t'\} - m$ so $t_k$ is the only determining randomization point. For all $t,t'$ such that $t_k \leq \min\{t,t'\} \leq t_{k+1} - 1, t_{k+1} + m \leq \max\{t,t'\} \leq T$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0$.

In this case,

$$\sum_{\substack{t_k \leq t,t' \leq T \\ \min\{t,t'\} \leq t_{k+1}-1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = B^2 \big(16 \cdot (m - t_k + t_{k-1})^2 + \\ 8 \cdot (m^2 - (m - t_k + t_{k-1})^2) + 4 \cdot ((m + t_{k+1} - t_k)^2 - 2m^2)\big)$$

**(4)** When $t_k - t_{k-1} < m, t_{k+1} - t_k < m$. Due to Lemma D.6, for all $t,t' \in \{t_k : t_{k-1}+m-1\}$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 16B^2$, because $t-m \leq t_{k-1}-1 \leq t_k \leq t$ and $t'-m \leq t_{k-1}-1 \leq t_k \leq t'$ so $t_{k-2}, t_{k-1}, t_k$ are three determining randomization points. Also $t_k - t_{k-2} \geq m$ so $t_{k-2} \leq \min\{t,t'\} - m$ and $t_{k-3}$ is not a determining randomization point. For all $t,t'$ such that $t_k \leq \min\{t,t'\} \leq t_{k+1} - 1, t_{k-1} + m \leq \max\{t,t'\} \leq t_k + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 8B^2$, because $\min\{t,t'\} - m < t_k - 1$ and $t_{k-1} \leq \max\{t,t'\} - m \leq t_k - 1$ so $t_{k-1}$ and $t_k$ are two determining randomization points. For all $t,t'$ such that $t_k \leq \min\{t,t'\} \leq t_{k+1} - 1, t_k + m \leq \max\{t,t'\} \leq t_{k+1} + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 4B^2$, because $t_k \leq \max\{t,t'\} - m$ so $t_k$ is the only determining randomization point. For all $t,t'$ such that $t_k \leq \min\{t,t'\} \leq t_{k+1} - 1, t_{k+1} + m \leq \max\{t,t'\} \leq T$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 0$.

In this case,

$$\sum_{\substack{t_k \leq t,t' \leq T \\ \min\{t,t'\} \leq t_{k+1}-1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = B^2 \big(16 \cdot (m - t_k + t_{k-1})^2 + \\ 8 \cdot (m^2 - (m - t_k + t_{k-1})^2 - (m - t_{k+1} + t_k)^2) + \\ 4 \cdot ((m + t_{k+1} - t_k)^2 - 2m^2 + (m - t_{k+1} + t_k)^2)\big)$$

Finally we calculate the third block from equation (D.9). Observe that $T - t_K \geq m$. **(1)** When $t_K - t_{K-1} \geq m$. Due to Lemma D.6, for all $t, t' \in \{t_K : t_K + m - 1\}$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 8B^2$, because both $t_{K-1} \leq t - m \leq t_K - 1$ and $t_{K-1} \leq t' - m \leq t_K - 1$, and both $t_{K-1}$ and $t_K$ are overlapping randomization points. For all $t, t'$ such that $t_K \leq \min\{t, t'\} \leq T, t_K + m \leq \max\{t, t'\} \leq T$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 4B^2$, because $t_K \leq \max\{t, t'\} - m$ so $t_K$ is the only determining randomization point.

In this case,

$$\sum_{t_K \leq t, t' \leq T} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = B^2 \left( 8 \cdot m^2 + 4 \cdot ((T + 1 - t_K)^2 - m^2) \right)$$

**(2)** When $t_K - t_{K-1} < m$. Due to Lemma D.6, for all $t, t' \in \{t_K : t_{K-1} + m - 1\}$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 16B^2$, because $t - m \leq t_{K-1} - 1 \leq t_K \leq t$ and $t' - m \leq t_{K-1} - 1 \leq t_K \leq t'$ so $t_{K-2}, t_{K-1}, t_K$ are three determining randomization points. Also $t_K - t_{K-2} \geq m$ so $t_{K-2} \leq \min\{t, t'\} - m$ and $t_{K-3}$ is not a determining randomization point. For all $t, t'$ such that $t_K \leq \min\{t, t'\} \leq t_K + m - 1, t_{K-1} + m \leq \max\{t, t'\} \leq t_K + m - 1$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 8B^2$, because $\min\{t, t'\} - m \leq t_K - 1$ and $t_{K-1} \leq \max\{t, t'\} - m \leq t_K - 1$ so $t_{K-1}$ and $t_K$ are two determining randomization points. For all $t, t'$ such that $t_K \leq \min\{t, t'\} \leq T, t_K + m \leq \max\{t, t'\} \leq T$, $\mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = 4B^2$, because $t_K \leq \max\{t, t'\} - m$ so $t_K$ is the only determining randomization point.

In this case,

$$\begin{aligned}\sum_{t_K \leq t, t' \leq T} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] = B^2 \big( & 16 \cdot (m - t_K + t_{K-1})^2 + \\ & 8 \cdot (m^2 - (m - t_K + t_{K-1})^2) + 4 \cdot ((T + 1 - t_K)^2 - m^2) \big)\end{aligned}$$

Now we combine all above together.

Note that whenever there exists $k \in \{2 : K\}$ such that $(t_k - t_{k-1}) < m$, this suggests that in

$$\sum_{\substack{t_k \leq t, t' \leq T \\ \min\{t, t'\} \leq t_{k+1} - 1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}]$$

there is a $16(m - t_k + t_{k-1})^2$; but in

$$\sum_{\substack{t_{k-1} \le t,t' \le T \\ \min\{t,t'\} \le t_k - 1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}]$$

there is a $8(-(m-t_k+t_{k-1})^2)$. So when we sum them up, we break $16(m-t_k+t_{k-1})^2$ into two $8(m - t_k + t_{k-1})^2$, which cancels in two sumations. By telescoping,

$$\begin{aligned}
&(T-m)^2 \cdot r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) \\
&= \sum_{\substack{m+1 \le t,t' \le T \\ \min\{t,t'\} \le t_1 - 1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] + \sum_{k=1}^{K-1} \left( \sum_{\substack{t_k \le t,t' \le T \\ \min\{t,t'\} \le t_{k+1} - 1}} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] \right) + \sum_{t_K \le t,t' \le T} \mathbb{E}[\mathbb{1}_t \mathbb{1}_{t'}] \\
&= 4B^2 \cdot \left((t_1 - 1)^2 - m^2\right) + \sum_{k=1}^{K-1} B^2 \cdot \left(8m^2 + 4\left((m + t_{k+1} - t_k)^2 - 2m^2 + \left((m - t_{k+1} + t_k)^+\right)^2\right)\right) \\
&\quad + B^2 \cdot \left(8m^2 + 4\left((T + 1 - t_K)^2 - m^2\right)\right) \\
&= B^2 \cdot \left\{ 4\sum_{k=0}^{K} (t_{k+1} - t_k)^2 + 8m(t_K - t_1) + 4m^2 K - 4m^2 + 4\sum_{k=1}^{K-1} [(m - t_{k+1} + t_k)^+]^2 \right\}
\end{aligned}$$

which finishes the proof. □

### D.3.6 Optimal Solutions to the Subset Selection Problem in Theorem 5.5

**Proof of Theorem 5.5.**

*Proof.* Proof of Theorem 5.5.

Consider the problem as we have introduced in (5.6). Due to Lemma 5.3, $\mathbb{Y}^+ = \{Y_t(\mathbf{1}_{m+1}) = Y_t(\mathbf{0}_{m+1}) = B\}_{t \in \{m+1:T\}}$ and $\mathbb{Y}^- = \{Y_t(\mathbf{1}_{m+1}) = Y_t(\mathbf{0}_{m+1}) = -B\}_{t \in \{m+1:T\}}$ are the only two dominating strategies for the adversarial selection of potential outcomes.

Then due to Lemma D.7 and Lemma D.8, the optimal design of switchback experiment must satisfy the following three conditions.

$$t_1 \geq m+2, \qquad t_K \leq T-m \qquad t_{k+1}-t_{k-1} \geq m, \ \forall k \in [K].$$

Due to Lemma D.9, the risk function of the optimal design of experiment is given by

$$r(\eta_{\mathbb{T},\mathbb{Q}}, \mathbb{Y}) = \frac{1}{(T-m)^2} \left\{ 4 \sum_{k=1}^{K+1} (t_k - t_{k-1})^2 + 8m(t_K - t_1) + \right. \\ \left. 4m^2K - 4m^2 + 4 \sum_{k=2}^{K} [(m - t_k + t_{k-1})^+]^2 \right\} B^2$$

So if we further take minimum over $\mathbb{T} \subset [T]$ in the above risk function, we find the optimal solution to the original problem introduced in (5.6). Note that $B^2$ is a constant and irrelevant to our decisions, and that $T$ and $m$ are inputs. So we solve, for any given $T$ and $m$, the following subset selection problem:

$$\min_{\mathbb{T} \subset [T]} \left\{ 4 \sum_{k=0}^{K} (t_{k+1} - t_k)^2 + 8m(t_K - t_1) + 4m^2K - 4m^2 + 4 \sum_{k=1}^{K-1} [(m - t_{k+1} + t_k)^+]^2 \right\},$$

as stated in (5.7).

In particular, if there exists some constant $n \in \mathbb{N}, n \geq 4$, such that $T = nm$, we can explicitly find the optimal design of experiment. Take the continuous relaxation of this problem, such that for any $K$, $\{1 < t_1 < t_2 < ... < t_K < T+1\} \in [1, T+1]^K$.

$$\min_{\substack{K \in \mathbb{N}, \\ \{1 < t_1 < t_2 < ... < t_K < T+1\} \in [1,T+1]^K}} \left\{ 4 \sum_{k=0}^{K} (t_{k+1} - t_k)^2 + 8m(t_K - t_1) + \right. \\ \left. 4m^2K - 4m^2 + 4 \sum_{k=1}^{K-1} [(m - t_{k+1} + t_k)^+]^2 \right\}$$

The relaxed problem provides a lower bound to the original subset selection problem as stated in (5.7). We will argue later that it is a lucky coincidence that the optimal solution to this relaxed problem is also an integer solution.

First we argue that $t_1 - t_0 = t_{K+1} - t_K$. This is because otherwise if $t_1 - t_0 \neq t_{K+1} - t_K$ then denote $a = \frac{t_1 - t_0 + t_{K+1} - t_K}{2}$. We could always pick for any $k \in \{1 : K\}$, $\tilde{t}_k = t_k + a - t_1 + 1$, such that $t_{k+1} - t_k$ is unchanged for any $k \in \{1 : K-1\}$. The only change in the objective value comes from

$$\left(2a^2\right) - \left((t_1 - t_0)^2 + (t_{K+1} - t_K)^2\right) < 0,$$

which suggests that $t_1 - t_0 \neq t_{K+1} - t_K$ is not optimal.

Second, similarly, we argue that for any $k' < k'' \in [K-1]$, $t_{k'+1} - t_{k'} = t_{k''+1} - t_{k''}$ This is because otherwise if $t_{k'+1} - t_{k'} \neq t_{k''+1} - t_{k''}$ then denote $b = \frac{t_{k'+1} - t_{k'} + t_{k''+1} - t_{k''}}{2}$. We could always pick for any $k \in \{k'+1 : k''\}$, $\tilde{t}_k = t_k + b - (t_{k'+1} - t_{k'})$, such that $t_{k+1} - t_k$ is unchanged for any $k \in \{k'+1 : k''-1\}$. The only change in the objective value comes from

$$\begin{aligned}&\left(2b^2 + 2((m-b)^+)^2\right) > \\ &\quad \left((t_{k'+1} - t_{k'})^2 + (t_{k''+1} - t_{k''})^2 + ((m - t_{k'+1} + t_{k'})^+)^2 + ((m - t_{k''+1} + t_{k''})^+)^2\right),\end{aligned}$$

where $x^2 + ((m-x)^+)^2$ is convex and the inequality holds due to Jensen's Inequality. This inequality suggests that $t_{k'+1} - t_{k'} \neq t_{k''+1} - t_{k''}$ is not optimal.

With the above two structural results, we can assume that there exists $a, b > 0$, such that $t_1 - t_0 = t_{K+1} - t_K = a$, and $t_{k+1} - t_k = b, \forall k \in [K-1]$ Also, it must be satisfied that $2a + (K-1)b = T$. Next we replace $K - 1 = \frac{T-2a}{b}$ into the relaxed problem, to have

$$\begin{aligned}&\min_{a,b>0} \left\{4(2a^2 + (K-1)b^2) + 8m(K-1)b + 4m^2(K-1) + 4(K-1)((m-b)^+)^2\right\} \\ =& \min_{a,b>0} \left\{8a^2 + 4(T-2a)b + 8m(T-2a) + 4m^2\frac{T-2a}{b} + 4\frac{T-2a}{b}((m-b)^+)^2\right\}\end{aligned}$$

Either when $b \geq m$, the above is to minimize

$$\min_{a,b>0} \left\{8a^2 + 4(T-2a)b + 8m(T-2a) + 4m^2\frac{T-2a}{b}\right\}$$

Note that

$$
\begin{aligned}
&8a^2 + 4(T-2a)b + 8m(T-2a) + 4m^2\frac{T-2a}{b} \\
=&8a^2 + 8m(T-2a) + 4(T-2a)\left(b + \frac{m^2}{b}\right) \\
\geq&8a^2 + 16m(T-2a) \\
=&8(a-2m)^2 + 16mT - 32m^2 \\
\geq&16mT - 32m^2
\end{aligned}
$$

where the first inequality takes equality if and only if $b = \frac{m^2}{b}$, which suggests $b = m$; the second inequality takes equality if and only if $a = 2m$.

Or when $b \leq m$, the above is to minimize

$$
\min_{a,b>0}\left\{8a^2 + 4(T-2a)b + 8m(T-2a) + 4m^2\frac{T-2a}{b} + 4\frac{T-2a}{b}(m-b)^2\right\}
$$

Note that

$$
\begin{aligned}
&8a^2 + 4(T-2a)b + 8m(T-2a) + 4m^2\frac{T-2a}{b} + 4\frac{T-2a}{b}(m-b)^2 \\
=&8a^2 + 8(T-2a)\left(b + \frac{m^2}{b}\right) \\
\geq&8a^2 + 16m(T-2a) \\
=&8(a-2m)^2 + 16mT - 32m^2 \\
\geq&16mT - 32m^2
\end{aligned}
$$

where the first inequality takes equality if and only if $b = \frac{m^2}{b}$, which suggests $b = m$; the second inequality takes equality if and only if $a = 2m$.

Combining both cases, the optimal solution is when $a = 2m$ and $b = m$, which happens to be an integer solution, thus optimal for the subset selection problem. Translating into $t_1, ..., t_K$ this suggests that $t_1 = 2m+1, t_2 = 3m+1, ..., t_K = (n-2)m+1$. $\square$

**Solutions in the Imperfect Cases.**

It is always worth noting that we are taking a design of experiments perspective. So when practically we have control of $T$, we can pick $T$ to be some multiples of $m$, which fits our Theorem 5.5 perfectly. If we do not have control of $T$, we can always pick a smaller $T'$ such that $T' = \lfloor T/m \rfloor \cdot m$ is some multiples of $m$.

Nonetheless, from an optimization perspective, we establish the following optimal structure for the subset selection problem as in (5.7). Recall that $t_{K+1} = T + 1$.

*Lemma* D.10. *Under Assumptions 5.1–5.3, the optimal design of regular switchback experiment must satisfy the following two conditions,*

$$|(t_1 - t_0) - (t_{K+1} - t_K)| \leq 1, \quad |(t_{j+1} - t_j) - (t_{j'+1} - t_{j'})| \leq 1, \forall 1 \leq j, j' \leq K - 1.$$

*Proof.* Proof of Lemma D.10. Prove by contradiction.

**Case 1.** Suppose there exists some optimal design $\mathbb{T}$, such that $(t_1 - t_0) - (t_{K+1} - t_K) \geq 2$. We now construct another design $\tilde{\mathbb{T}}$, such that $\left|\tilde{\mathbb{T}}\right| = K = |\mathbb{T}|$, and the $K$ elements are $\tilde{\mathbb{T}} = \{\tilde{t}_0 = 1, \tilde{t}_1 = t_1 - 1, \tilde{t}_2 = t_2 - 1, ..., \tilde{t}_K = t_K - 1\}$. Now check the expression as in (5.7). Note that $\tilde{t}_{k+1} - \tilde{t}_k = t_{k+1} - t_k$ is unchanged for any $k \in [K-1]$; $\tilde{t}_K - \tilde{t}_1 = t_K - t_1$ is unchanged; and $m - \tilde{t}_{k+1} - \tilde{t}_k = m - t_{k+1} - t_k$ in unchanged for any $k \in [K-1]$. But $(\tilde{t}_1 - \tilde{t}_0)^2 + (\tilde{t}_{K+1} - \tilde{t}_K)^2 = (t_1 - t_0 - 1)^2 + (t_{K+1} - t_K + 1)^2 \leq (t_1 - t_0)^2 + (t_{K+1} - t_K)^2$, because $(t_1 - t_0) - (t_{K+1} - t_K) \geq 2$ and due to convexity.

Similarly, if there exists some optimal design $\mathbb{T}$, such that $(t_{K+1} - t_K) - (t_1 - t_0) \geq 2$, then construct another design $\tilde{\mathbb{T}} = \{\tilde{t}_0 = 1, \tilde{t}_1 = t_1 + 1, \tilde{t}_2 = t_2 + 1, ..., \tilde{t}_K = t_K + 1\}$.

**Case 2.** Suppose there exists some optimal design $\mathbb{T}$, and there exists $1 \leq j < j' \leq K - 1$ such that $(t_{j+1} - t_j) - (t_{j'+1} - t_{j'}) \geq 2$. We now construct another design $\tilde{\mathbb{T}}$, such that $\left|\tilde{\mathbb{T}}\right| = K = |\mathbb{T}|$, and the $K$ elements are $\tilde{\mathbb{T}} = \{\tilde{t}_0 = 1, \tilde{t}_1 = t_1, ..., \tilde{t}_j = t_j, \tilde{t}_{j+1} = t_{j+1} - 1, ..., \tilde{t}_{j'} = t_{j'} - 1, \tilde{t}_{j'+1} = t_{j'+1}, ..., \tilde{t}_K = t_K\}$. Now check the expression as in (5.7). Note that $\tilde{t}_{k+1} - \tilde{t}_k = t_{k+1} - t_k$ is unchanged for any $k \in \{0 : K\}$ except $j$ and $j'$; $\tilde{t}_K - \tilde{t}_1 = t_K - t_1$ is unchanged; and $m - \tilde{t}_{k+1} - \tilde{t}_k = m - t_{k+1} - t_k$ in unchanged for any $k \in [K-1]$ except $j$ and $j'$. Now focus on $j$ and $j'$.

$$
\begin{aligned}
&(\tilde{t}_{j+1}-\tilde{t}_j)^2+(\tilde{t}_{j'+1}-\tilde{t}_{j'})^2+[(m-\tilde{t}_{j+1}+\tilde{t}_j)^+]^2+[(m-\tilde{t}_{j'+1}+\tilde{t}_{j'})^+]^2\\
=&(t_{j+1}-t_j-1)^2+(t_{j'+1}-t_{j'}+1)^2+[(m-t_{j+1}+t_j+1)^+]^2+[(m-t_{j'+1}+t_{j'}-1)^+]^2\\
\leq&(t_{j+1}-t_j)^2+(t_{j'+1}-t_{j'})^2+[(m-t_{j+1}+t_j)^+]^2+[(m-t_{j'+1}+t_{j'})^+]^2
\end{aligned}
$$

To see why this inequality holds, define $g(x)=x^2+[(m-x)^+]^2$ and note that $g(x)$ is a univariate convex function. The inequality holds due to $(t_{j+1}-t_j)-(t_{j'+1}-t_{j'})\geq 2$ and convexity.

Similarly, if there exists some optimal design $\mathbb{T}$, and there exists $1\leq j<j'\leq K-1$ such that $(t_{j'+1}-t_{j'})-(t_{j+1}-t_j)\geq 2$. Then construct another design $\tilde{\mathbb{T}}=\{\tilde{t}_0=1,\tilde{t}_1=t_1,...,\tilde{t}_j=t_j,\tilde{t}_{j+1}=t_{j+1}+1,...,\tilde{t}_{j'}=t_{j'}+1,\tilde{t}_{j'+1}=t_{j'+1},...,\tilde{t}_K=t_K\}$.

Combine both cases we finish the proof. □

## D.4 Proofs and Discussions from Section 5.4

In the first two sub-Sections of Section 5.4 we focus on the case when $p=m$. In Sections D.4.1–D.4.4 in the appendix, we also focus on the case when $p=m$, and use only $m$ instead of $p$. In Sections D.4.5–D.4.7, we will use both $p$ and $m$. Recall that $m$ is the order of the carryover effect, and $p$ is the experimenter's knowledge of $m$.

### D.4.1 Extra Notations Used in the Proofs from Section 5.4

For any $t\in\{m+1:T\}$, we use the notations of $\mathbb{1}_t$ as defined in (D.2). Denote

$$
\begin{aligned}
\bar{\mathbb{1}}_0 &= \sum_{t=m+1}^{2m}\mathbb{1}_t \\
\bar{\mathbb{1}}_k &= \sum_{t=(k+1)m+1}^{(k+2)m}\mathbb{1}_t, && \forall k\in[K] \\
\bar{\mathbb{1}}_{K+1} &= \sum_{t=(K+2)m+1}^{(K+3)m}\mathbb{1}_t
\end{aligned}
$$

It is worth noting that under the optimal design as suggested by Theorem 5.5, when $T/m = n \in \mathbb{N}$ is an integer, we have $K = n - 3$. So $(K+3)m = T$. See Example D.1 below.

*Example* D.1 (An Optimal Design and Its $\bar{\mathbb{1}}_k$ Notations). When $T = 12$, $p = m = 2$, the optimal design of regular switchback experiment is $\mathbb{T}^* = \{1, 5, 7, 9\}$, and $K = 3$. The $\bar{\mathbb{1}}_k$ notations are defined below. Each $\bar{\mathbb{1}}_k$ spans $m = 2$ periods. See Table D.7. ∎

Table D.7: An example of the optimal design $\mathbb{T}^*$ and its $\bar{\mathbb{1}}_k$ notations when $T = 12$ and $p = m = 2$

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbb{T}^*$ | ✓ | — | — | — | ✓ | — | ✓ | — | ✓ | — | — | — |
| $\{\bar{\mathbb{1}}_k\}_{k=0}^{K+1}$ | — | | $\bar{\mathbb{1}}_0$ | | $\bar{\mathbb{1}}_1$ | | $\bar{\mathbb{1}}_2$ | | $\bar{\mathbb{1}}_3$ | | $\bar{\mathbb{1}}_4$ | |

Using the above notation, we could write

$$\hat{\tau}_m - \tau_m = \frac{1}{T-m} \sum_{k=0}^{K+1} \bar{\mathbb{1}}_k,$$

and so

$$\mathsf{Var}(\hat{\tau}_m) = \frac{1}{(T-m)^2} \mathsf{Var}\left( \sum_{k=0}^{K+1} \bar{\mathbb{1}}_k \right).$$

### D.4.2 Proof of Lemma 5.6

The proof of Lemma 5.6 resembles the proof of Lemmas D.3 and D.4. The trick here is to observe that for any $k \in [K]$, the values of all the variables $\mathbb{1}_t$, where $(k+1)m + 1 \leq t \leq (k+2)m$, are all determined by the randomization at time $km + 1$ and $(k+1)m + 1$. Since they are all correlated, we can use $\bar{\mathbb{1}}_k$ to stand for $\sum_{t=(k+1)m+1}^{(k+2)m} \mathbb{1}_t$ for short.

*Proof.* Proof of Lemma 5.6. First observe that $\bar{\mathbb{1}}_k$ has zero mean for each $k \in \{0 : K+1\}$. So we can decompose the variance into squared terms and cross-product terms,

$$(T-m)^2\mathsf{Var}(\hat{\tau}_m) = \mathsf{Var}\left(\sum_{k=0}^{K+1} \bar{\mathbb{1}}_k\right) = \sum_{k=0}^{K+1} \mathbb{E}\left[\bar{\mathbb{1}}_k^2\right] + \sum_{0\leq k<k'\leq K+1} 2\mathbb{E}\left[\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k'}\right].$$

We focus on the variance of the squared terms first,

$$\mathbb{E}\left[\bar{\mathbb{1}}_k^2\right] = \begin{cases} \bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2 + 2\bar{Y}_0(\mathbf{1}_{m+1})\bar{Y}_0(\mathbf{0}_{m+1}), & \text{if} \quad k=0 \\ 3\bar{Y}_k(\mathbf{1}_{m+1})^2 + 3\bar{Y}_k(\mathbf{0}_{m+1})^2 + 2\bar{Y}_k(\mathbf{1}_{m+1})\bar{Y}_k(\mathbf{0}_{m+1}), & \text{if} \quad 1\leq k\leq K \\ \bar{Y}_{K+1}(\mathbf{1}_{m+1})^2 + \bar{Y}_{K+1}(\mathbf{0}_{m+1})^2 + 2\bar{Y}_{K+1}(\mathbf{1}_{m+1})\bar{Y}_{K+1}(\mathbf{0}_{m+1}), & \text{if} \quad k=K+1 \end{cases}$$

This is because when $k=0$ or $k=K+1$, then with probability 1/2, $\bar{\mathbb{1}}_k = \bar{Y}_0(\mathbf{1}_{m+1}) + \bar{Y}_0(\mathbf{0}_{m+1})$; with probability 1/2, $\bar{\mathbb{1}}_k = -\bar{Y}_0(\mathbf{1}_{m+1}) - \bar{Y}_0(\mathbf{0}_{m+1})$. When $k\in[K]$, with probability 1/4, $\bar{\mathbb{1}}_k = 3\bar{Y}_0(\mathbf{1}_{m+1}) + \bar{Y}_0(\mathbf{0}_{m+1})$; with probability 1/2, $\bar{\mathbb{1}}_k = -\bar{Y}_0(\mathbf{1}_{m+1}) + \bar{Y}_0(\mathbf{0}_{m+1})$; with probability 1/4, $\bar{\mathbb{1}}_k = -\bar{Y}_0(\mathbf{1}_{m+1}) - 3\bar{Y}_0(\mathbf{0}_{m+1})$.

Then for the cross-product terms, if $k'-k\geq 2$, then $\bar{\mathbb{1}}_k$ and $\bar{\mathbb{1}}_{k'}$ are independent, i.e., $\mathbb{E}\left[\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k'}\right]=0$. If $k'-k=1$, then

$$\mathbb{E}\left[\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1}\right] = (\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$$

This is because the values of $\bar{\mathbb{1}}_k$ and $\bar{\mathbb{1}}_{k+1}$ are determined by the realization at 3 randomization points, $W_{km+1}, W_{(k+1)m+1}, W_{(k+2)m+1}$. With probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (3\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(3\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (3\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(-\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (-\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(3\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (-\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(-\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (-\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(-\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (-\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1}))\cdot(-\bar{Y}_{k+1}(\mathbf{1}_{m+1}) - 3\bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (-\bar{Y}_k(\mathbf{1}_{m+1}) - 3\bar{Y}_k(\mathbf{0}_{m+1}))\cdot(-\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1}))$; with probability 1/8, $\bar{\mathbb{1}}_k\bar{\mathbb{1}}_{k+1} = (-\bar{Y}_k(\mathbf{1}_{m+1}) - 3\bar{Y}_k(\mathbf{0}_{m+1}))\cdot(-\bar{Y}_{k+1}(\mathbf{1}_{m+1}) - 3\bar{Y}_{k+1}(\mathbf{0}_{m+1}))$.

Combining the squared terms and the cross-product terms we finish the proof. $\square$

### D.4.3 Discssions and proof of Corollary 5.6.1

We first provide the details of the two variance upper bounds here.

$$\mathsf{Var}^{\mathsf{U1}}(\hat{\tau}_m) = \frac{1}{(T-m)^2}\Bigg\{3\left[\bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2\right] + \sum_{k=1}^{n-3} 6\left[\bar{Y}_k(\mathbf{1}_{m+1})^2 + \bar{Y}_k(\mathbf{0}_{m+1})^2\right]$$

$$+4\left[\bar{Y}_{n-2}(\mathbf{1}_{m+1})^2 + \bar{Y}_{n-2}(\mathbf{0}_{m+1})^2\right] + \sum_{k=0}^{n-3} 2\left[\bar{Y}_k(\mathbf{1}_{m+1})\cdot\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1})\cdot\bar{Y}_{k+1}(\mathbf{0}_{m+1})\right]\Bigg\},$$

and

$$\mathsf{Var}^{\mathsf{U2}}(\hat{\tau}_m) = \frac{1}{(T-m)^2}\Bigg\{4\left[\bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2\right] + \sum_{k=1}^{n-3} 8\left[\bar{Y}_k(\mathbf{1}_{m+1})^2 + \bar{Y}_k(\mathbf{0}_{m+1})^2\right]$$

$$+ 4\left[\bar{Y}_{n-2}(\mathbf{1}_{m+1})^2 + \bar{Y}_{n-2}(\mathbf{0}_{m+1})^2\right]\Bigg\}.$$

We prove Corollary 5.6.1 using the basic inequality that $2xy \leq x^2 + y^2$. Such an inequality is commonly used to find a conservative upper bound of the variance.

*Proof.* Proof of Corollary 5.6.1. From Lemma 5.6, the variance of the estimator is given by

$$\begin{aligned}
&(T-m)^2\mathsf{Var}(\hat{\tau}_m)\\
\leq& 2\left\{\bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2\right\} + \sum_{k=1}^{n-3} 4\left\{\bar{Y}_k(\mathbf{1}_{m+1})^2 + \bar{Y}_k(\mathbf{0}_{m+1})^2\right\} + 2\left\{\bar{Y}_{n-2}(\mathbf{1}_{m+1})^2 + \bar{Y}_{n-2}(\mathbf{0}_{m+1})^2\right\}\\
&+ \sum_{k=0}^{n-3} 2\left[\bar{Y}_k(\mathbf{1}_{m+1}) + \bar{Y}_k(\mathbf{0}_{m+1})\right]\cdot\left[\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + \bar{Y}_{k+1}(\mathbf{0}_{m+1})\right]\\
\leq& 2\left\{\bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2\right\} + \sum_{k=1}^{n-3} 4\left\{\bar{Y}_k(\mathbf{1}_{m+1})^2 + \bar{Y}_k(\mathbf{0}_{m+1})^2\right\} + 2\left\{\bar{Y}_{n-2}(\mathbf{1}_{m+1})^2 + \bar{Y}_{n-2}(\mathbf{0}_{m+1})^2\right\}\\
&+ \sum_{k=0}^{n-3} \Big\{2\bar{Y}_k(\mathbf{1}_{m+1})\bar{Y}_{k+1}(\mathbf{1}_{m+1}) + 2\bar{Y}_k(\mathbf{0}_{m+1})\bar{Y}_{k+1}(\mathbf{0}_{m+1}) +\\
&\qquad \bar{Y}_k(\mathbf{1}_{m+1})^2 + \bar{Y}_k(\mathbf{0}_{m+1})^2 + \bar{Y}_{k+1}(\mathbf{1}_{m+1})^2 + \bar{Y}_{k+1}(\mathbf{0}_{m+1})^2\Big\}\\
\leq& 3\left\{\bar{Y}_0(\mathbf{1}_{m+1})^2 + \bar{Y}_0(\mathbf{0}_{m+1})^2\right\} + \sum_{k=1}^{n-3} 6\left\{\bar{Y}_k(\mathbf{1}_{m+1})^2 + \bar{Y}_k(\mathbf{0}_{m+1})^2\right\} + 3\left\{\bar{Y}_{n-2}(\mathbf{1}_{m+1})^2 + \bar{Y}_{n-2}(\mathbf{0}_{m+1})^2\right\}
\end{aligned}$$

$$+\sum_{k=0}^{n-3}\left\{\bar{Y}_k(\mathbf{1}_{m+1})^2+\bar{Y}_k(\mathbf{0}_{m+1})^2+\bar{Y}_{k+1}(\mathbf{1}_{m+1})^2+\bar{Y}_{k+1}(\mathbf{0}_{m+1})^2\right\}$$

$$=4\left\{\bar{Y}_0(\mathbf{1}_{m+1})^2+\bar{Y}_0(\mathbf{0}_{m+1})^2\right\}+\sum_{k=1}^{n-3}8\left\{\bar{Y}_k(\mathbf{1}_{m+1})^2+\bar{Y}_k(\mathbf{0}_{m+1})^2\right\}+4\left\{\bar{Y}_{n-2}(\mathbf{1}_{m+1})^2+\bar{Y}_{n-2}(\mathbf{0}_{m+1})^2\right\}$$

where the first inequality suggests $\mathsf{Var}(\hat{\tau}_m)\leq\mathsf{Var}^{\mathsf{U1}}(\hat{\tau}_m)$, and the last inequality suggests $\mathsf{Var}^{\mathsf{U1}}(\hat{\tau}_m)\leq\mathsf{Var}^{\mathsf{U2}}(\hat{\tau}_m)$.

The unbiasedness part is due to the estimator of the variances being Horvitz-Thompson type estimators. □

### D.4.4 Proof of Theorem 5.7

We prove Theorem 5.7 by using Lemma D.1. In particular, we derive $B_{n,k,a}^2$, and then construct some proper $\Delta_n, K_n$, and $L_n$.

*Proof.* Proof of Theorem 5.7. In the $n$-replica experiment, $\hat{\tau}_m-\tau_m=\frac{1}{(n-1)m}\sum_{k=0}^{n-2}\bar{\mathbb{I}}_k$, and $\mathsf{Var}(\hat{\tau}_m)=\frac{1}{(n-1)^2m^2}\mathsf{Var}\left(\sum_{k=0}^{n-2}\bar{\mathbb{I}}_k\right)$. To use the language from Lemma D.1, denote $d=n-1$. Denote for any $i\in[n-1]$, $X_{n,i}=\frac{1}{(n-1)m}\bar{\mathbb{I}}_{i-1}$ so we know that $\phi=1$, i.e., $\{X_{n,1},X_{n,2},...\}$ is a sequence of 1-dependent random variables.

First note that $B_n^2=\mathsf{Var}(\hat{\tau}_m)$, and we calculate $B_{n,k,a}^2$ as follows.

$$\begin{aligned}
B_{n,k,a}^2&=\frac{1}{(n-1)^2m^2}\mathsf{Var}\left(\sum_{i=a}^{a+k-1}\bar{\mathbb{I}}_{i-1}\right)\\
&\leq\frac{1}{(n-1)^2m^2}\left\{\sum_{i=a}^{a+k-1}\left[3\bar{Y}_{i-1}(\mathbf{1}_{m+1})^2+3\bar{Y}_{i-1}(\mathbf{0}_{m+1})^2+2\bar{Y}_{i-1}(\mathbf{1}_{m+1})\bar{Y}_{i-1}(\mathbf{0}_{m+1})\right]\right.\\
&\qquad\left.+\sum_{i=a}^{a+k-2}2[\bar{Y}_{i-1}(\mathbf{1}_{m+1})+\bar{Y}_{i-1}(\mathbf{0}_{m+1})]\cdot[\bar{Y}_i(\mathbf{1}_{m+1})+\bar{Y}_i(\mathbf{0}_{m+1})]\right\}\\
&\leq\frac{8km^2B^2+8(k-1)m^2B^2}{(n-1)^2m^2}\\
&\leq\frac{16kB^2}{(n-1)^2}
\end{aligned}$$

Pick $\gamma=0,\delta=1$, then $\Delta_n=B^3/(n-1)^3$, $K_n=16B^2/(n-1)^2$, and $L_n=\mathsf{Var}(\hat{\tau}_m)/(n-1)$.

We check that all the five conditions from Lemma D.1 are satisfied.

1. $\mathbb{E}\left|X_{n,i}\right|^3 \leq \Delta_n = B^3/(n-1)^3$, because all the potential outcomes are bounded by $B$, so that $X_{n,i} \leq B/(n-1)$.

2. $B_{n,k,a}^2/k \leq K_n = 16B^2/(n-1)^2$.

3. $B_n^2/(n-1) \geq L_n = \mathsf{Var}(\hat{\tau}_m)/(n-1)$.

4. $K_n/L_n = 16B^2/(n-1)\mathsf{Var}(\hat{\tau}_m) = O(1)$, where the last equality is due to Assumption 5.4.

5. $\Delta_n/L_n^{3/2} = B^3/(n-1)^{3/2}\mathsf{Var}(\hat{\tau}_m)^{3/2} = O(1)$, where the last equality is due to Assumption 5.4.

Due to Lemma D.1,

$$\frac{\hat{\tau}_m - \tau_m}{\sqrt{\mathsf{Var}(\hat{\tau}_m)}} \xrightarrow{D} \mathcal{N}(0,1).$$

□

### D.4.5 Interpretation for the Horvitz-Thompson Estimator under Misspecified $m$ Case

For the remainder of this section, we discuss the cases when $m$ is misspecified. Throughout this section in the appendix, we use both $p$ and $m$. Recall that $m$ is the order of the carryover effect, and $p$ is the experimenter's knowledge of $m$.

As we have discussed in Section 5.4.3, all our estimation and inference methods will hold when $p \geq m$. When $p < m$, the Horvitz-Thompson estimator as we defined in (5.4) will no longer be unbiased in estimating the lag-$p$ causal estimand as we defined in (5.1). However, we can still interpret the Horvitz-Thompson estimator as we defined in (5.4).

When $p < m$, the lag-$p$ effect in (5.1) is not well defined. Instead, we define the $m$-misspecified lag-$p$ causal effect that pads the $p+1$ assignments with the earlier observed treatments.

$$\tau_p^{(m)}(\mathbb{Y}) = \frac{1}{T-p}\left\{\sum_{t=p+1}^{m}\left[Y_t(\boldsymbol{w}_{1:t-p-1}^{\mathsf{obs}}, \mathbf{1}_{p+1}) - Y_t(\boldsymbol{w}_{1:t-p-1}^{\mathsf{obs}}, \mathbf{0}_{p+1})\right] + \sum_{t=m+1}^{T}\left[Y_t(\boldsymbol{w}_{t-m:t-p-1}^{\mathsf{obs}}, \mathbf{1}_{p+1}) - Y_t(\boldsymbol{w}_{t-m:t-p-1}^{\mathsf{obs}}, \mathbf{0}_{p+1})\right]\right\}. \quad \text{(D.10)}$$

This is a special case of the weighted lag-$p$ causal effect introduced in Bojinov and Shephard (2019). Similarly to the average lag-$p$ causal effect, $\tau_p^{(m)}(\mathbb{Y})$ captures how administering $p+1$ consecutive treatments as opposed to $p+1$ consecutive controls impact the outcomes at time $t$, *conditional on* the observed assignment path up to time $t-p-1$.[1] See Section 5.5.4 for numerical results.

When $p > m$, Proposition 5.2 still holds, i.e., $\mathbb{E}[\hat{\tau}_p] = \tau_p(\mathbb{Y}) = \tau_m(\mathbb{Y})$. When $p < m$, sometimes we have to slightly augment the results and study the conditional expectation.

Define $f_{\mathbb{T}} : [T] \to \mathbb{T}$ to be the "determining randomization point of period $t$,"

$$f_{\mathbb{T}}(t) = \max\{j \,|\, j \in \mathbb{T}, j \le t\}$$

such that, it is the realization at time $f_{\mathbb{T}}(t)$ that uniquely determines the assignment at time $t$, *i.e.* $W_t = W_{f_{\mathbb{T}}(t)}, \forall t \in [T]$. See Example D.2 for an illustration of $f_{\mathbb{T}}(\cdot)$. When $\mathbb{T}$ is clear from the context we drop the subscript and use $f(\cdot) = f_{\mathbb{T}}(\cdot)$. Depending on if $f(t-p) \le t-m$, we establish an analogy of Proposition 5.2 for the $p < m$ case.

*Proposition* D.11 (Conditional Unbiasedness of the Estimator when $m$ is Misspecified). *Under Assumptions 5.1 and 5.2, for $p < m$, at each time $t \ge m+1$, the Horvitz-Thompson estimator is either unbiased for the lag-$m$ causal effect when $f(t-p) \le t-m$, or conditionally unbiased for the $m$-misspecified lag-$p$ causal effect when $f(t-p) > t-m$. When $p+1 \le t \le m$, the Horvitz-Thompson estimator is either unbiased for the lag-$t$ causal effect when $f(t-p) = 1$, or conditionally unbiased for the $m$-misspecified lag-$t$ causal effect when $f(t-p) > 1$.*

[1] See (Bojinov and Shephard 2019, Section 3) for an extended discussion.

To remove the conditional expectation, we can further take an outer loop of expectation averaged over the past assignment paths. Although this is somewhat different from the average lag-$p$ effect introduced earlier in (5.1), it does capture the impact of a sequence of treatment relative to a sequence of controls.

All the mathematical expressions of Proposition D.11, as well its proof, are stated in Section D.4.6 in the Appendix. See Example D.2 below for a specific illustration of Proposition D.11. For a numerical illustration of the estimand and estimator in more general setups, see Section 5.5.4.

*Example* D.2 (Misspecified $m$). Suppose $T = 4, m = 2, p = 1, \mathbb{T} = \{1, 3\}$. Then the determining randomization points are $f_{\mathbb{T}}(1) = 1, f_{\mathbb{T}}(2) = 1, f_{\mathbb{T}}(3) = 3, f_{\mathbb{T}}(4) = 3$, and

$$\begin{aligned}
&\mathbb{E}\left[Y_2^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{1:2} = (1,1)\}}{\Pr(\boldsymbol{W}_{1:2} = (1,1))} - Y_2^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{1:2} = (0,0)\}}{\Pr(\boldsymbol{W}_{1:2} = (0,0))}\right] = Y_2(1,1) - Y_2(0,0) \\
&\mathbb{E}\left[Y_3^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{2:3} = (1,1)\}}{\Pr(\boldsymbol{W}_{2:3} = (1,1))} - Y_3^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{2:3} = (0,0)\}}{\Pr(\boldsymbol{W}_{2:3} = (0,0))}\right] = Y_3(1,1,1) - Y_3(0,0,0) \\
&\mathbb{E}\left[Y_4^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{3:4} = (1,1)\}}{\Pr(\boldsymbol{W}_{3:4} = (1,1))} - Y_4^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{3:4} = (0,0)\}}{\Pr(\boldsymbol{W}_{3:4} = (0,0))}\right] = \frac{1}{2}\left[Y_4(1,1,1) + Y_4(0,1,1)\right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left. - Y_4(0,0,0) - Y_4(1,0,0)\right]
\end{aligned}$$

Note that this is the 2-misspecified lag-1 causal effect. ∎

### D.4.6 Unbiasedness of the Horvitz-Thompson Estimator when $m$ is Misspecified

We state here the omitted mathematics in Proposition D.11.

Under Assumptions 5.1 and 5.2, for $p < m$, at each time $t \geq m + 1$, the Horvitz-Thompson estimator is either unbiased for the lag-$m$ causal effect when $f(t-p) \leq t - m$, i.e.,

$$\mathbb{E}_{\boldsymbol{W}_{1:T} \sim \eta_{\mathbb{T},\mathbb{Q}}}\left[Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1})} - Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1})}\right] = Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{0}_{m+1}),$$

or conditionally unbiased for the $m$-misspecified lag-$p$ causal effect when $f(t-p) >$

$t-m$, i.e.,

$$\mathbb{E}_{\boldsymbol{W}_{1:T}\sim\eta_{\mathbb{T},\mathbb{Q}}}\left[\left\{Y_t^{\mathsf{obs}}\frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1})}-Y_t^{\mathsf{obs}}\frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1})}\right\}-\right.$$
$$\left.\left\{Y_t(\boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1},\mathbf{1}_{t-f(t-p)+1})-Y_t(\boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1},\mathbf{0}_{t-f(t-p)+1})\right\}\middle|\boldsymbol{W}_{t-m:f(t-p)-1}=\boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}\right]=0.$$

When $p+1\leq t\leq m$, the Horvitz-Thompson estimator is either unbiased for the lag-$t$ causal effect when $f(t-p)=1$, i.e.,

$$\mathbb{E}_{\boldsymbol{W}_{1:T}\sim\eta_{\mathbb{T},\mathbb{Q}}}\left[Y_t^{\mathsf{obs}}\frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1})}-Y_t^{\mathsf{obs}}\frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1})}\right]=Y_t(\mathbf{1}_t)-Y_t(\mathbf{0}_t),$$

or conditionally unbiased for the $m$-misspecified lag-$t$ causal effect when $f(t-p)>1$, i.e.,

$$\mathbb{E}_{\boldsymbol{W}_{1:T}\sim\eta_{\mathbb{T},\mathbb{Q}}}\left[\left\{Y_t^{\mathsf{obs}}\frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1})}-Y_t^{\mathsf{obs}}\frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1})}\right\}-\right.$$
$$\left.\left\{Y_t(\boldsymbol{w}^{\mathsf{obs}}_{1:f(t-p)-1},\mathbf{1}_{t-f(t-p)+1})-Y_t(\boldsymbol{w}^{\mathsf{obs}}_{1:f(t-p)-1},\mathbf{0}_{t-f(t-p)+1})\right\}\middle|\boldsymbol{W}_{1:f(t-p)-1}=\boldsymbol{w}^{\mathsf{obs}}_{1:f(t-p)-1}\right]=0.$$

To remove the conditional expectation, we can further take an outer loop of expectation averaged over the past assignment paths. So the estimator is estimating a weighted average of lag-$p$ effects. When $t\geq m+1$,

$$\sum_{\boldsymbol{w}_{t-m:f(t-p)-1}}\Pr(\boldsymbol{W}_{t-m:f(t-p)-1}=\boldsymbol{w}_{t-m:f(t-p)-1})$$
$$(Y_t(\boldsymbol{w}_{t-m:f(t-p)-1},\mathbf{1}_{t-f(t-p)+1})-Y_t(\boldsymbol{w}_{t-m:f(t-p)-1},\mathbf{0}_{t-f(t-p)+1})),$$

and when $p+1\leq t\leq m$,

$$\sum_{\boldsymbol{w}_{1:f(t-p)-1}}\Pr(\boldsymbol{W}_{1:f(t-p)-1}=\boldsymbol{w}_{1:f(t-p)-1})(Y_t(\boldsymbol{w}_{1:f(t-p)-1},\mathbf{1}_{t-f(t-p)+1})-Y_t(\boldsymbol{w}_{1:f(t-p)-1},\mathbf{0}_{t-f(t-p)+1})).$$

We prove Proposition D.11 as follows.

*Proof.* Proof of Proposition D.11. Focus on any specific $t\in\{m+1:T\}$.

When $f(t-p) \leq t-m$, both $0 < \Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}), \Pr(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1}) < 1$. With probability $\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}) \neq 0$, $\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}\} = 1$, and $Y_t^{\mathsf{obs}} = Y_t(\mathbf{1}_{m+1})$. So $\mathbb{E}\left[Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{1}_{p+1})}\right] = Y_t(\mathbf{1}_{m+1})$. Similarly $\mathbb{E}\left[Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t}=\mathbf{0}_{p+1})}\right] = Y_t(\mathbf{0}_{m+1})$. So

$$\mathbb{E}_{\boldsymbol{W}_{1:T} \sim \eta_{\mathbb{T},\mathbb{Q}}} \left[ \left\{ Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1})} - Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1})} \right\} \right] = Y_t(\mathbf{1}_{m+1}) - Y_t(\mathbf{0}_{m+1}).$$

When $f(t-p) > t-m$, both $0 < \Pr\left(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1} \,\middle|\, \boldsymbol{W}_{t-m:f(t-p)-1} = \boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}\right) < 1$ and $0 < \Pr\left(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1} \,\middle|\, \boldsymbol{W}_{t-m:f(t-p)-1} = \boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}\right) < 1$. Conditional on $\boldsymbol{W}_{t-m:f(t-p)-1} = \boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}$, it must be that with non-zero probability

$$\Pr\left(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1} \,\middle|\, \boldsymbol{W}_{t-m:f(t-p)-1} = \boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}\right) \neq 0,$$

we have $\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}\} = 1$, and $Y_t^{\mathsf{obs}} = Y_t(\boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}, \mathbf{1}_{t-f(t-p)+1})$. So

$$\mathbb{E}_{\boldsymbol{W}_{1:T}} \left[ Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{1}_{p+1})} - Y_t(\boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}, \mathbf{1}_{t-f(t-p)+1}) \,\middle|\, \boldsymbol{W}_{t-m:f(t-p)-1} = \boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1} \right] = 0.$$

Similarly, we have

$$\mathbb{E}_{\boldsymbol{W}_{1:T}} \left[ Y_t^{\mathsf{obs}} \frac{\mathbb{1}\{\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1}\}}{\Pr(\boldsymbol{W}_{t-p:t} = \mathbf{0}_{p+1})} - Y_t(\boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1}, \mathbf{0}_{t-f(t-p)+1}) \,\middle|\, \boldsymbol{W}_{t-m:f(t-p)-1} = \boldsymbol{w}^{\mathsf{obs}}_{t-m:f(t-p)-1} \right] = 0,$$

which finishes the proof. $\square$

### D.4.7 Asymptotic Normality when $m$ is Misspecified

The proof of Corollary 5.7.1 consists of two parts: $p > m$ and $p < m$. When $p > m$ we consult Theorems 5.6 and 5.7. When $p < m$ we prove Corollary 5.7.1 by using Lemma D.1. In particular, we derive $B^2_{n,k,a}$, and then construct some proper $\Delta_n, K_n$, and $L_n$.

*Proof.* Proof of Corollary 5.7.1. The proof consists of two parts: $p > m$ and $p < m$. First, when $p > m$, we know that $\hat{\tau}_p = \hat{\tau}_m, \tau_p = \tau_m, \mathsf{Var}(\hat{\tau}_p) = \mathsf{Var}(\hat{\tau}_m)$. Due to Theorems 5.6 we prove part (i) the expression in (5.10). Due to Theorem 5.7 we know that

$$\frac{\hat{\tau}_p - \tau_p}{\sqrt{\mathsf{Var}(\hat{\tau}_p)}} = \frac{\hat{\tau}_m - \tau_m}{\sqrt{\mathsf{Var}(\hat{\tau}_m)}} \xrightarrow{D} \mathcal{N}(0,1).$$

Second, when $p < m$, then we follow the same trick as in Theorem 5.7. In the $n$-replica experiment, $\hat{\tau}_p - \mathbb{E}[\tau_p^{[m]}] = \frac{1}{(n-1)p}\sum_{k=0}^{n-2} \bar{\mathbb{1}}_k$, and $\mathsf{Var}(\hat{\tau}_p) = \frac{1}{(n-1)^2p^2}\mathsf{Var}\left(\sum_{k=0}^{n-2} \bar{\mathbb{1}}_k\right)$. To use the language from Lemma D.1, denote $d = n-1$. Denote for any $i \in [n-1]$, $X_{n,i} = \frac{1}{(n-1)p}\bar{\mathbb{1}}_{i-1}$. We know that $\phi = \lceil \frac{m}{p} \rceil$, so that $\{X_{n,1}, X_{n,2}, ...\}$ is a sequence of $\phi$-dependent random variables. See Table D.8 for an illustration of $\phi$.

Table D.8: An illustration of $\phi$ when $m = 5, p = 3$

$\xrightarrow{\text{carryover effect}}$

| | $\dots$ | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | $\dots$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbb{T}^*$ | | ✓ | − | − | ✓ | − | − | ✓ | − | − | ✓ | − | − | |
| $\{\bar{\mathbb{1}}_k\}_{k=0}^{K+1}$ | | $\bar{\mathbb{1}}_3$ | | | $\bar{\mathbb{1}}_4$ | | | $\bar{\mathbb{1}}_5$ | | | $\bar{\mathbb{1}}_6$ | | | |

In this example $\phi = \lceil \frac{m}{p} \rceil = 2$. The arrow above numbers 17 through 22 means that the assignment on period 17 affects the outcome on period 22. So that $\bar{\mathbb{1}}_4$ and $\bar{\mathbb{1}}_6$ are correlated, but $\bar{\mathbb{1}}_3$ and $\bar{\mathbb{1}}_6$ are independent.

First note that $B_n^2 = \mathsf{Var}(\hat{\tau}_p)$, and we calculate $B_{n,k,a}^2$ as follows. Note that $k \geq \phi + 1$.

$$\begin{aligned}
B_{n,k,a}^2 &= \frac{1}{(n-1)^2p^2}\mathsf{Var}\left(\sum_{i=a}^{a+k-1} \bar{\mathbb{1}}_{i-1}\right) \\
&\leq \frac{1}{(n-1)^2p^2}\left(\sum_{i=a}^{a+k-1} \mathbb{E}[\bar{\mathbb{1}}_{i-1}^2] + \sum_{i=a}^{a+k-2} 2\mathbb{E}[\bar{\mathbb{1}}_{i-1}\bar{\mathbb{1}}_i] + ... + \sum_{i=a}^{a+k-1+\phi} 2\mathbb{E}[\bar{\mathbb{1}}_{i-1}\bar{\mathbb{1}}_{i-1+\phi}]\right) \\
&\leq \frac{Cp^2B^2}{(n-1)^2p^2} \cdot (k + (k-1) + ... + (k-\phi)) \\
&\leq \frac{(\phi+1)CkB^2}{(n-1)^2}
\end{aligned}$$

where $C$ is some constant bounding the number of terms in each cross-product expectation $2\mathbb{E}[\bar{\mathbb{1}}_{i-1}\bar{\mathbb{1}}_i], ..., 2\mathbb{E}[\bar{\mathbb{1}}_{i-1}\bar{\mathbb{1}}_{i-1+\phi}]$; and $\phi + 1$ is a constant as well.

Pick $\gamma = 0, \delta = 1$, then $\Delta_n = B^3/(n-1)^3$, $K_n = (\phi+1)CB^2/(n-1)^2$, and $L_n = \mathsf{Var}(\hat{\tau}_m)/(n-1)$.

We check that all the five conditions from Lemma D.1 are satisfied.

1. $\mathbb{E}|X_{n,i}|^3 \leq \Delta_n = B^3/(n-1)^3$, because all the potential outcomes are bounded by $B$, so that $X_{n,i} \leq B/(n-1)$.

2. $B_{n,k,a}^2/k \leq K_n = (\phi+1)CB^2/(n-1)^2$.

3. $B_n^2/(n-1) \geq L_n = \mathsf{Var}(\hat{\tau}_m)/(n-1)$.

4. $K_n/L_n = (\phi+1)CB^2/(n-1)\mathsf{Var}(\hat{\tau}_m) = O(1)$, where the last equality is due to Assumption 5.4.

5. $\Delta_n/L_n^{3/2} = B^3/(n-1)^{3/2}\mathsf{Var}(\hat{\tau}_m)^{3/2} = O(1)$, where the last equality is due to Assumption 5.4.

Due to Lemma D.1,

$$\frac{\hat{\tau}_p - \tau_p}{\sqrt{\mathsf{Var}(\hat{\tau}_p)}} \xrightarrow{D} \mathcal{N}(0,1).$$

□

## D.5 Additional Simulation Results

### D.5.1 Flexibility of the Outcome Models

As we will see below, it is easy to use the potential outcome framework to describe many complex relationships between assignments and outcomes.

We start with a simple model which originates from Oman and Seiden (1988):

$$Y_t(\boldsymbol{w}_{1:t}) = \mu + \alpha_t + \delta w_t + \gamma w_{t-1} + \epsilon_t \tag{D.11}$$

where $\mu$ is a fixed effect; $\alpha_t$ is a fixed effect associated to period $t$; $\delta w_t$ is the contemporaneous effect, and $\gamma w_{t-1}$ is the carryover effect from period $t-1$; $\epsilon_t$ is the random noise in period $t$. Such a model as well as a few very similar ones are widely used in the literature (Hedayat et al. 1978, Jones and Kenward 2014).

A more general variant from the above model is to consider carryover effects of any arbitrary order, which we have defined in (5.13) in the main body of the paper.

$$Y_t(\boldsymbol{w}_{1:t}) = \mu + \alpha_t + \delta^{(1)} w_t + \delta^{(2)} w_{t-1} + ... + \delta^{(t)} w_1 + \epsilon_t$$

where $\delta^{(1)}, \delta^{(2)}, ..., \delta^{(t)}$ are non-stochastic coefficients. The dotted terms are carryover effects of higher orders. And all the other parameters are as defined in (D.11). We will run simulations based on this more general model, which enables us to test the performance of our proposed optimal design under a misspecified $m$.

The autoregressive model (Arellano 2003) is even more general: $Y_1(w_1) = \delta_{1,1} w_1 + \epsilon_1$ and $\forall t > 1$

$$\begin{aligned} Y_t(\boldsymbol{w}_{1:t}) = \phi_{t,t-1} Y_{t-1}(\boldsymbol{w}_{1:t-1}) + \phi_{t,t-2} Y_{t-2}(\boldsymbol{w}_{1:t-2}) + ... + \phi_{t,1} Y_1(w_1) + \\ \delta_{t,t} w_t + \delta_{t,t-1} w_{t-1} + ... + \delta_{t,1} w_1 + \epsilon_t \end{aligned} \quad \text{(D.12)}$$

where $\phi_{t,\tilde{t}}$ and $\delta_{t,\tilde{t}}$ are non-stochastic coefficients; the dotted terms are carryover effects of higher orders; $\epsilon_t$ is the random noise in period $t$. We can iteratively replace $Y_t(w_t)$ using a linear combination of $w_t, w_{t-1}, ..., w_1$. So the autoregressive model in (D.12) can be written in a similar form of (5.13). The only difference is that the coefficients are different and dependent on $t$.

### D.5.2 Additional Simulation Results for Section 5.5.2 Asymptotic Normality

In Section 5.5.2 we have only shown simulation results for the variance distribution, when $m$ is correctly specified and under $\delta = 3$, see asymptotic normality. In this section we provide additional simulation results under $\delta = 1$ and $\delta = 2$.

See Figures D-1–D-3 for simulation results under $\delta = 1$; See Figures D-4–D-6 for simulation results under $\delta = 2$. See Figures D-7–D-8 for simulation results under $\delta = 3$.

By comparing all the results, we see that in all cases, the pink histograms approximately follow the standard normal distribution; whereas the light blue histograms, since the distributions are induced by normalizing the expectation of the conservative upper bound, are more concentrated around zero. Furthermore, as $\delta$ increases, the light blue histograms become even more concentrated around zero, i.e., the distances between the light blue histograms and the pink histograms grow larger.

Figure D-1: Approximate normality of the randomization distribution when $m = 2, p = 2, \delta = 1$.

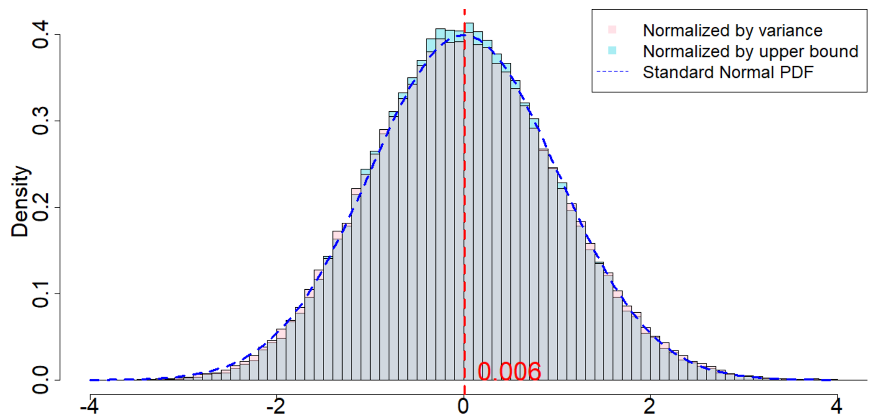
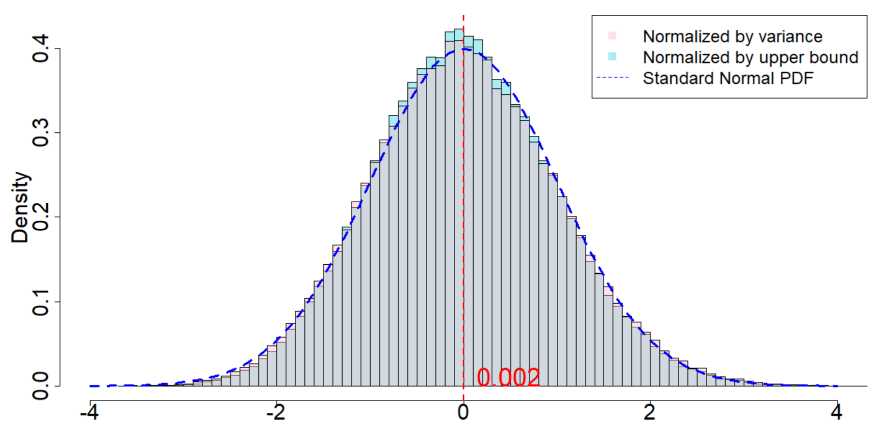



Figure D-2: Approximate normality of the randomization distribution when $m = 2, p = 3, \delta = 1$.




**Robustness Check.**

In Section 5.5.2 we have shown results when $m = 2, p = 2, \delta = 1$. In this section we provide additional simulation results under other parameters. When $T = 120$,

Figure D-3: Approximate normality of the randomization distribution when $m = 2, p = 1, \delta = 1$.

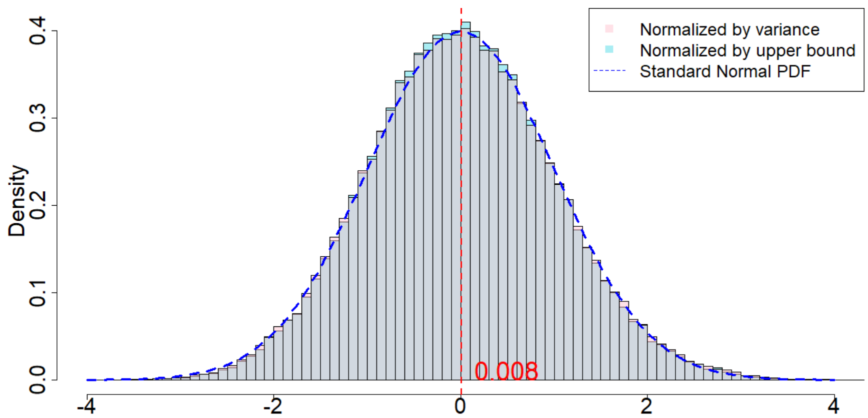



Figure D-4: Approximate normality of the randomization distribution when $m = 2, p = 2, \delta = 2$.

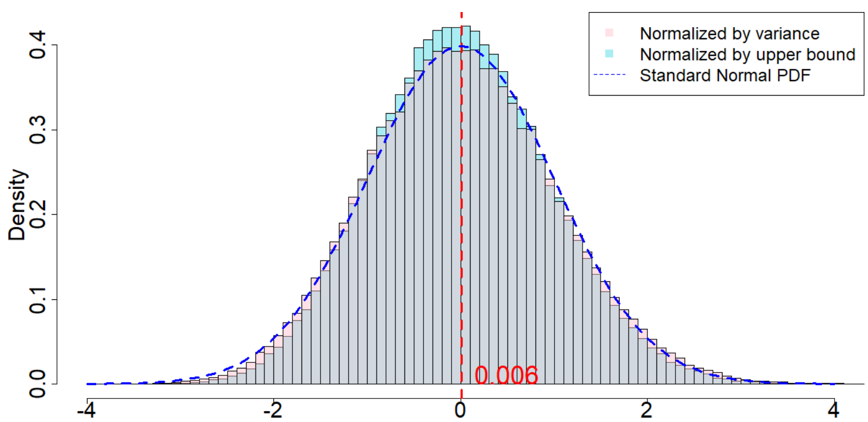



Figure D-5: Approximate normality of the randomization distribution when $m = 2, p = 3, \delta = 2$.

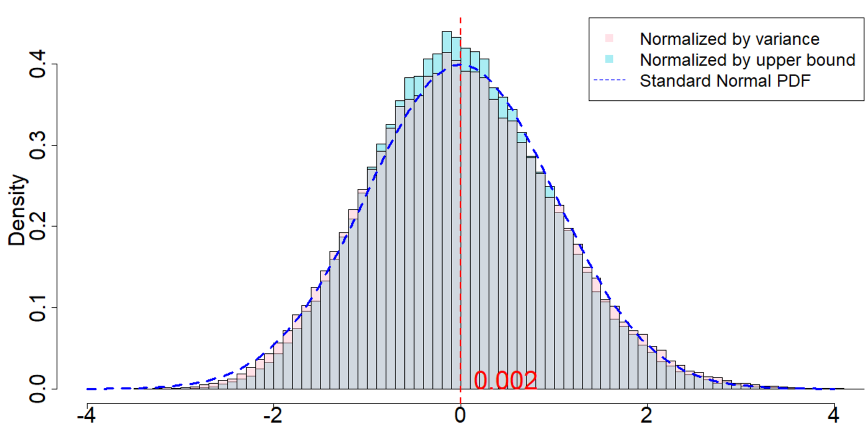


Figure D-6: Approximate normality of the randomization distribution when $m = 2, p = 1, \delta = 2$.

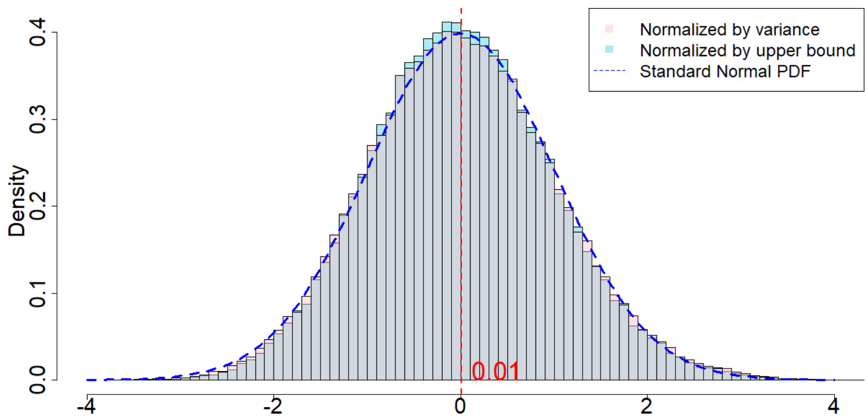



Figure D-7: Approximate normality of the randomization distribution when $m = 2, p = 3, \delta = 3$.

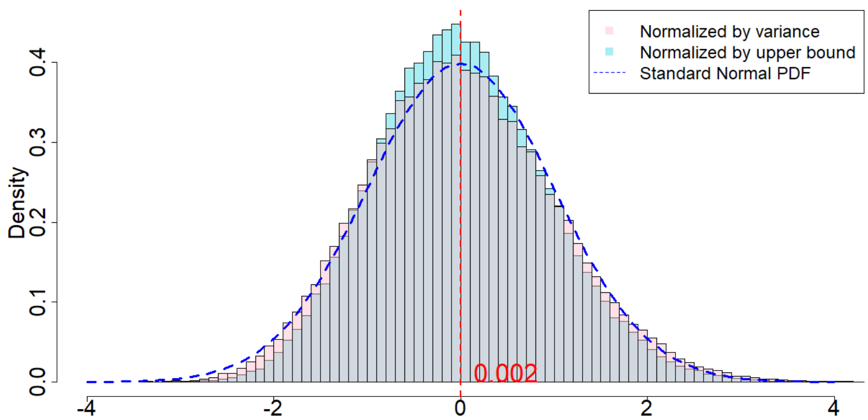
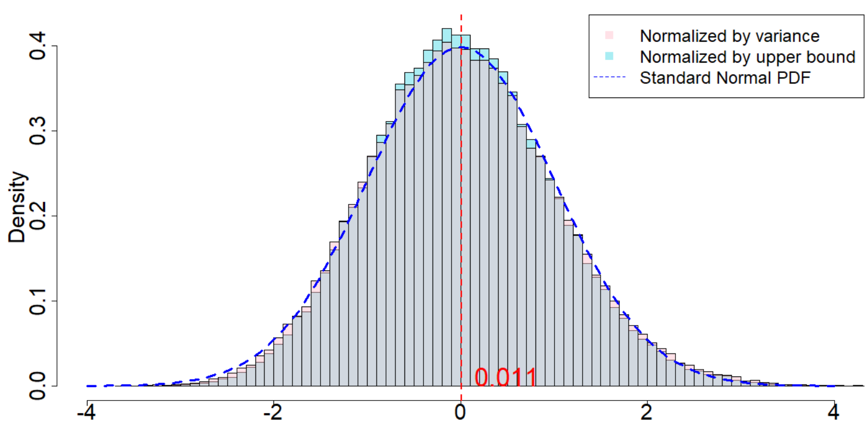


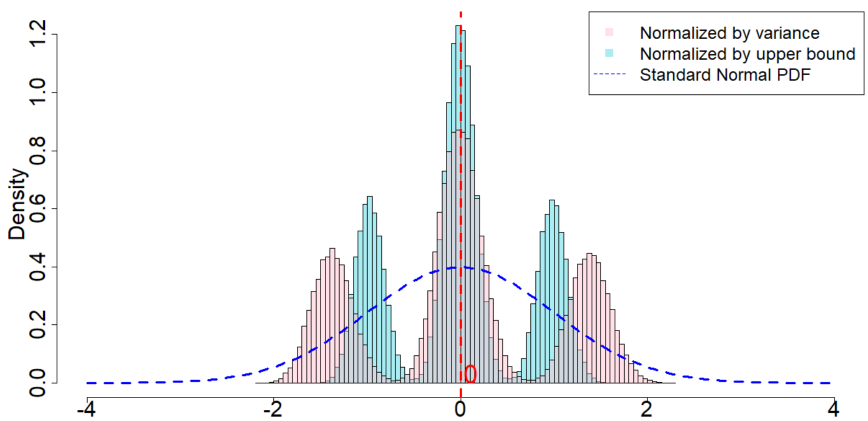

Figure D-8: Approximate normality of the randomization distribution when $m = 2, p = 1, \delta = 3$.



the empirical distributions as shown in the histograms are significantly different from normal distributions. See Figures 5-5, D-9, D-11, D-13, D-15, D-17, D-19, D-21, D-23. When $T = 1200$, the empirical distributions as shown in the histograms are much closer to normal distributions. See Figures 5-6, D-10, D-12, D-14, D-16, D-18, D-20, D-22, D-24. All the simulation results deliver the same message, that when $\epsilon_t$ noises are heavy tailed, the convergence to a standard normal distribution as we have shown in Theorem 5.7 requires longer horizon.

Interestingly, if we make the comparison between the pink histogram and the light blue histogram, we can see how much gap it incurs when we replace the true variance with the conservative upper bound. If we compare Figure D-1 and Figure 5-6, then we find that the conservative upper bound is a better approximation of the true variance when the noises $\epsilon_t$ conform normal distributions, rather than heavy-tailed distributions.

Figure D-9: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 2, \delta = 2, T = 120$.




### D.5.3 Additional Simulation Results for Section 5.5.3 Rejection Rates

In Section 5.5.3 we have provided simulation results for the rejection rates when the rejection threshold is 0.1. In this section we provide additional simulation results for the rejection rates when the rejection threshold is replaced by 0.05 and 0.01. See

Figure D-10: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 2, \delta = 2, T = 1200$.

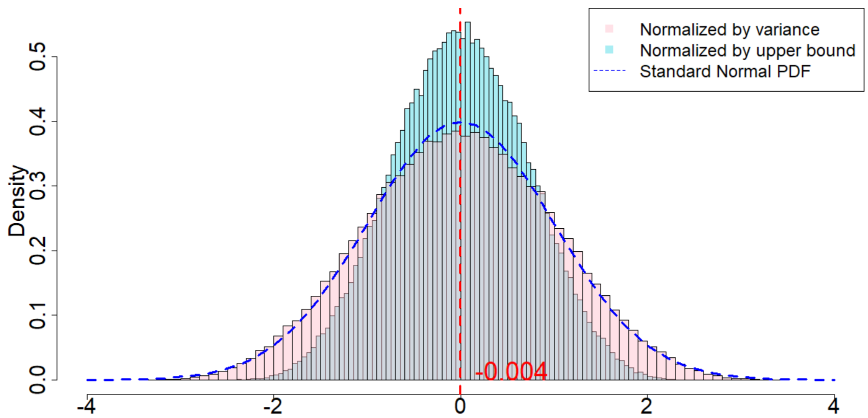



Figure D-11: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 2, \delta = 3, T = 120$.

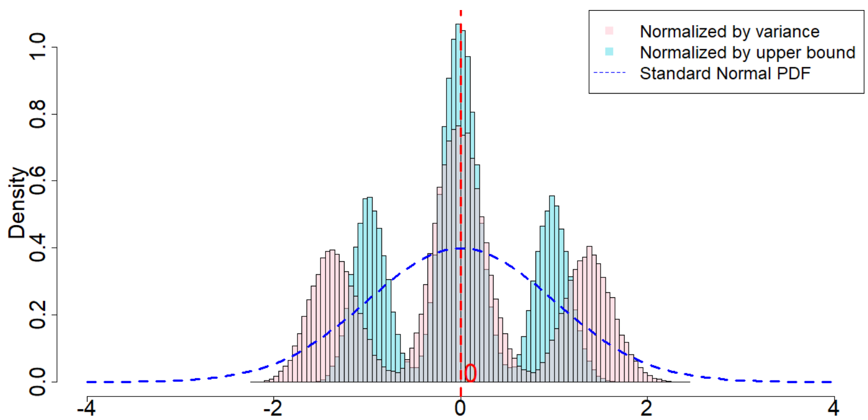



Figure D-12: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 2, \delta = 3, T = 1200$.

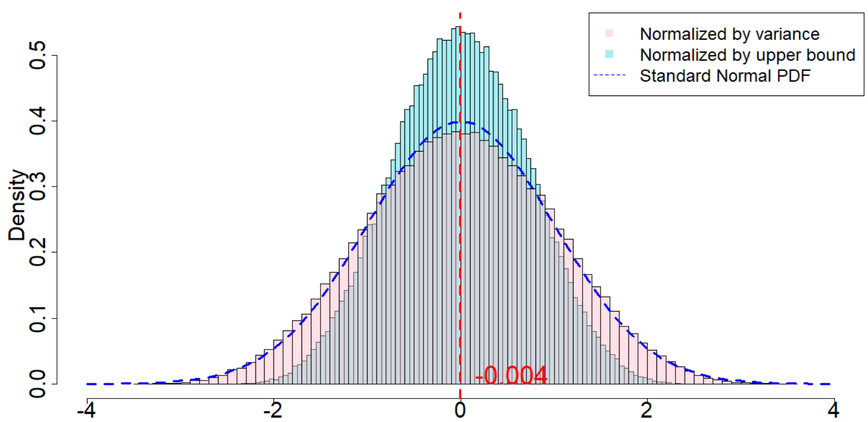


Figure D-13: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 3, \delta = 1, T = 120$.

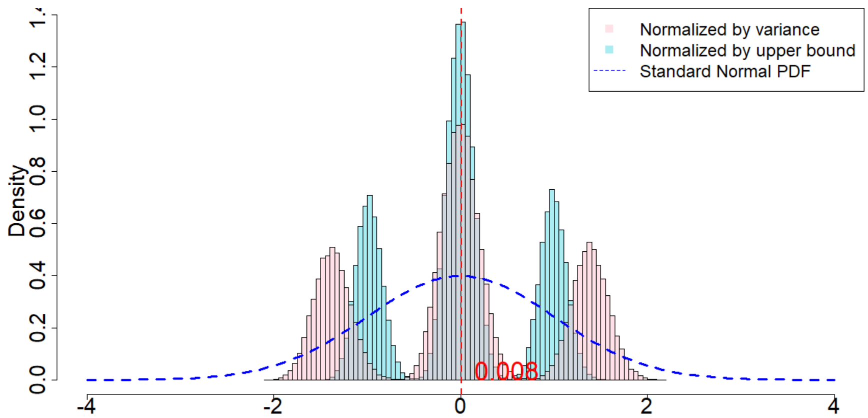



Figure D-14: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 3, \delta = 1, T = 1200$.

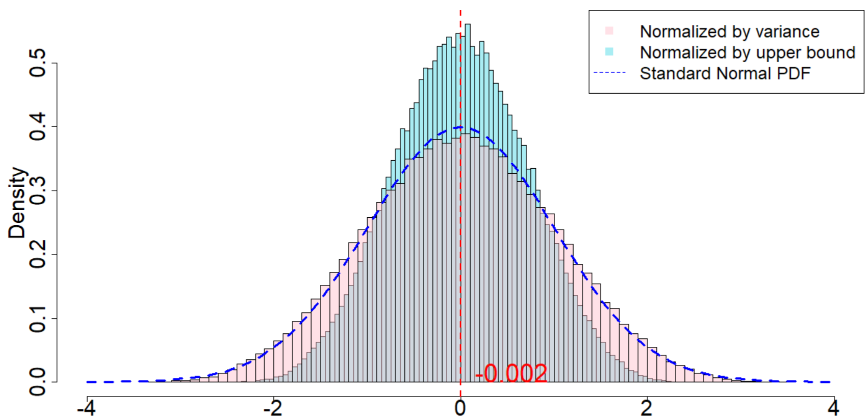



Figure D-15: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 3, \delta = 2, T = 120$.

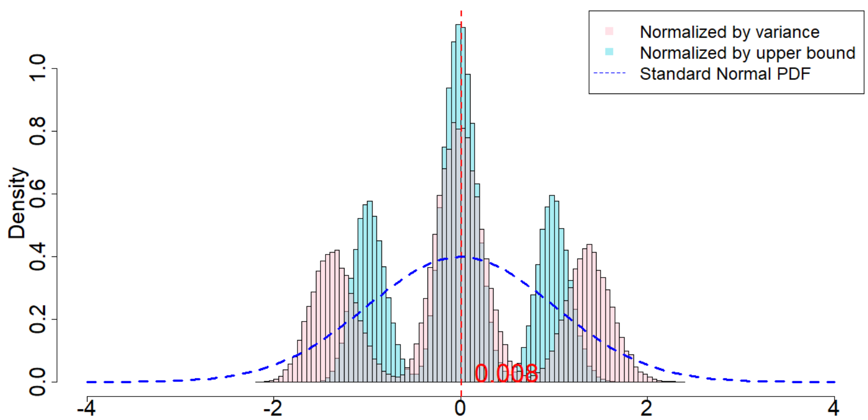


Figure D-16: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 3, \delta = 2, T = 1200$.

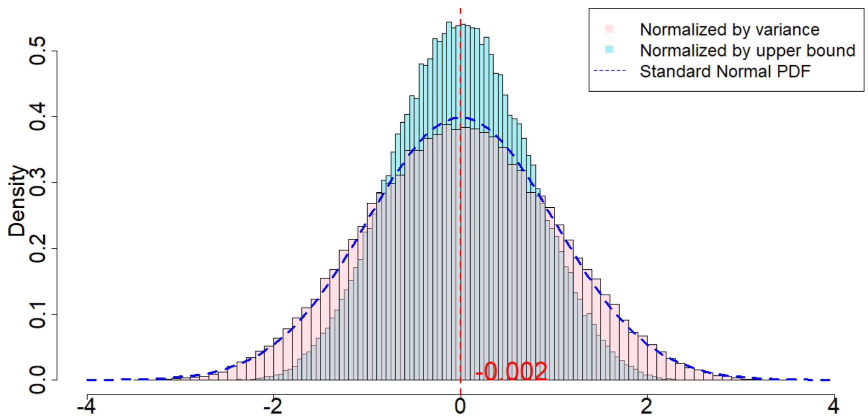



Figure D-17: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 3, \delta = 3, T = 120$.

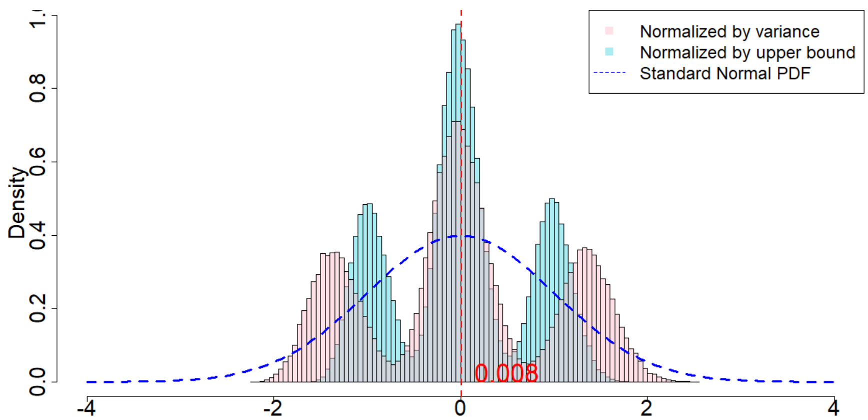



Figure D-18: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 3, \delta = 3, T = 1200$.

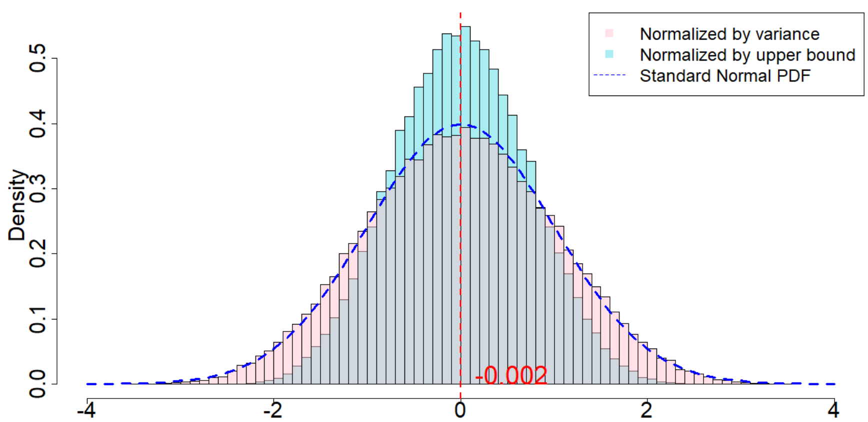


Figure D-19: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 1, \delta = 1, T = 120$.

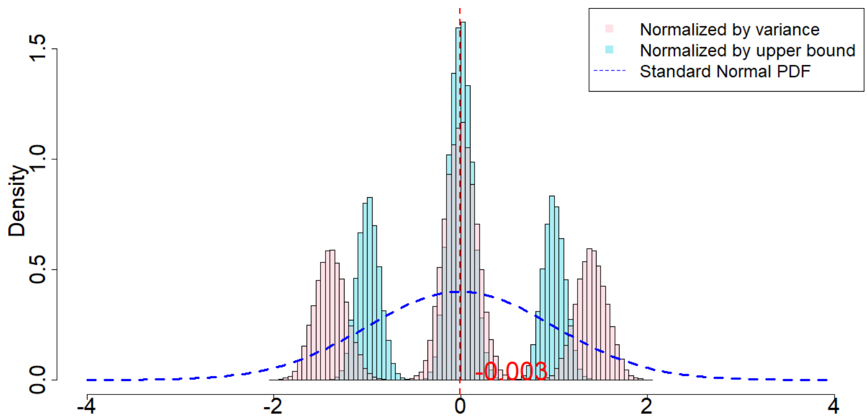



Figure D-20: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 1, \delta = 1, T = 1200$.

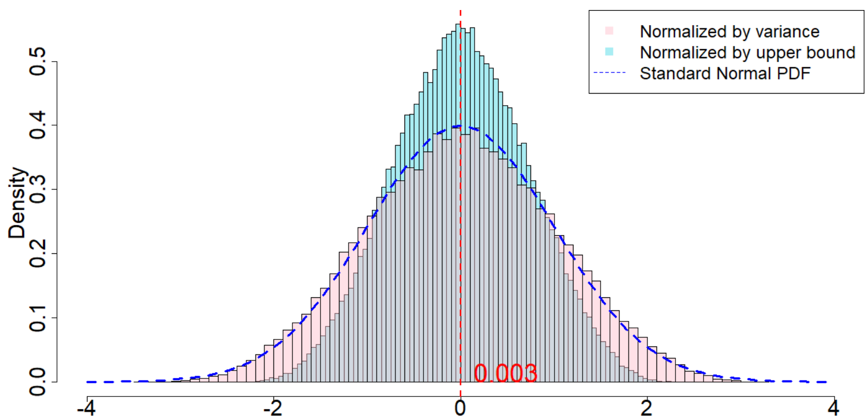



Figure D-21: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 1, \delta = 2, T = 120$.

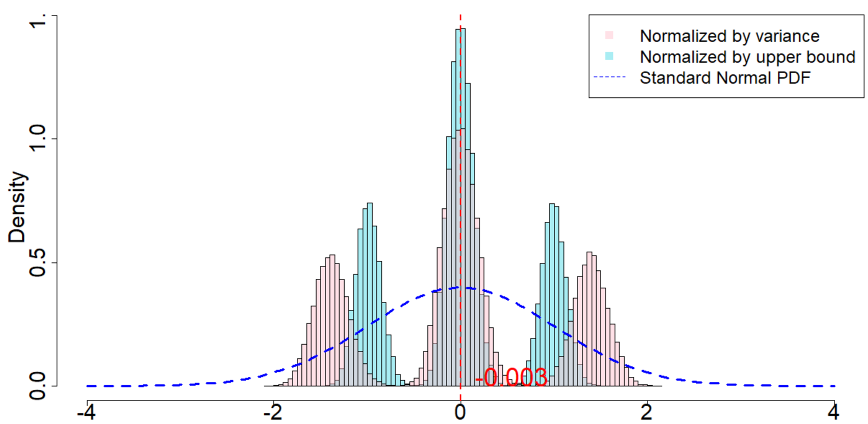


Figure D-22: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 1, \delta = 2, T = 1200$.

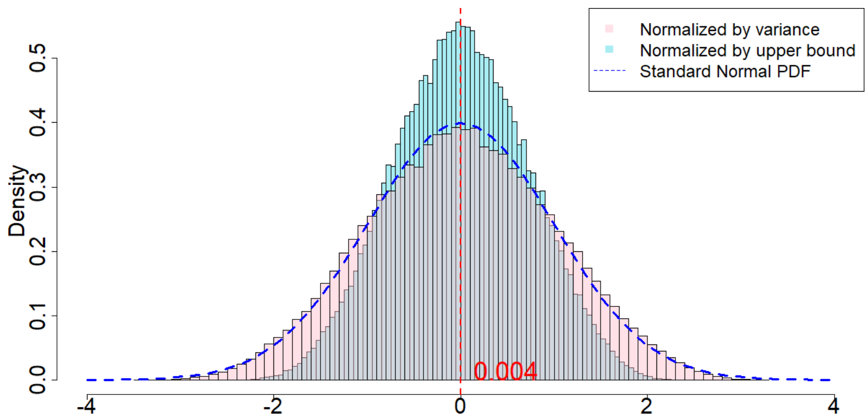



Figure D-23: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 1, \delta = 3, T = 120$.

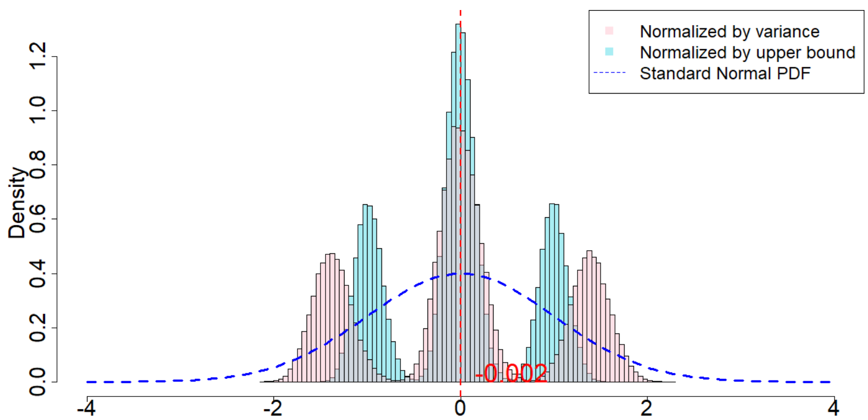
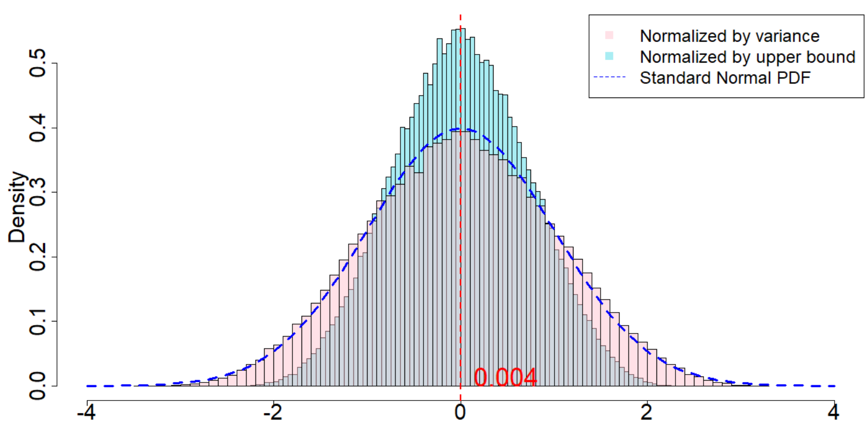



Figure D-24: Randomization distribution when random noises are Student's t-distributions, and when $m = 2, p = 1, \delta = 3, T = 1200$.



Figures D-25 and D-26.

Figure D-25: Rejection rates and their dependence on $T/m$, when the rejection threshold is 0.05. Left: $\delta = 1$; Middle: $\delta = 2$; Right: $\delta = 3$

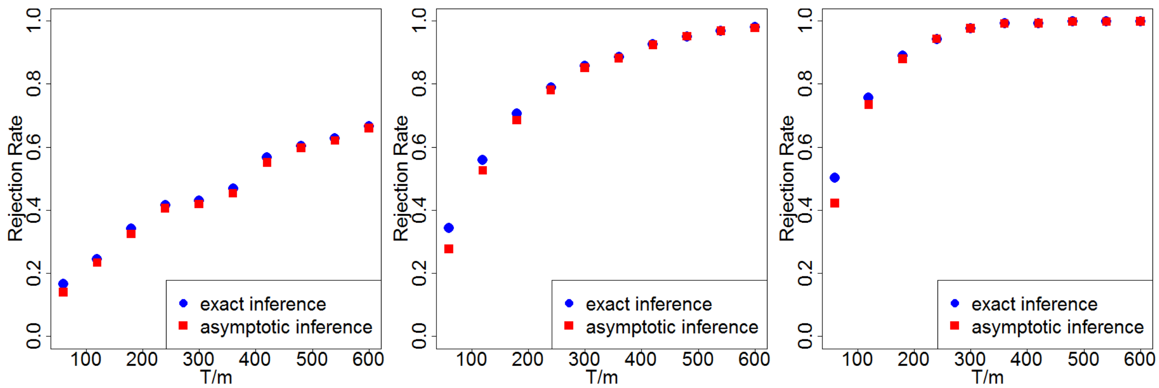
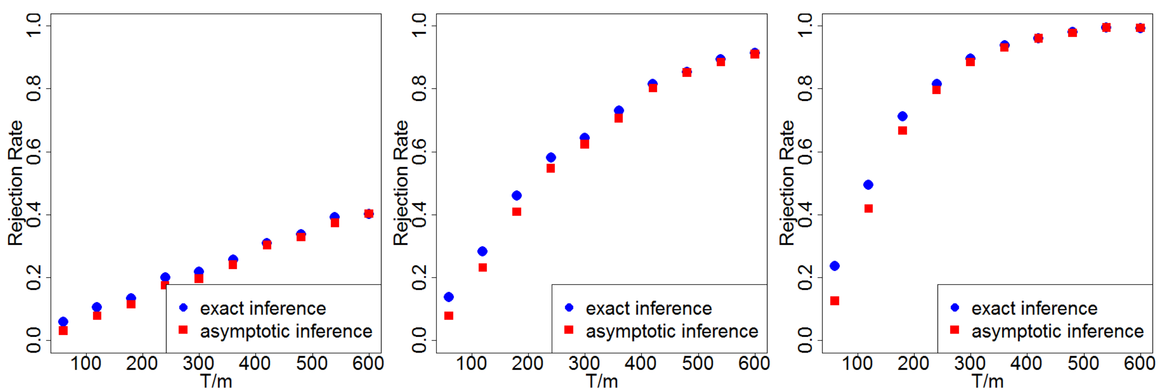



Figure D-26: Rejection rates and their dependence on $T/m$, when the rejection threshold is 0.01. Left: $\delta = 1$; Middle: $\delta = 2$; Right: $\delta = 3$




The blue dots are rejection rates under exact inference; the red dots are under asymptotic inference. Similar to the simulation results in Section 5.5.3, we would ideally wish to reject both the Fisher's null hypothesis (5.8) and the Neyman's null hypothesis (5.9). Both figures illustrate such rejection rates.

Besides the three observations we make in Section 5.5.3 (namely, dependence on $T/m$, between two inference methods, and dependence on the signal-to-noise ratio), we make an extra observation here. When we decrease the rejection threshold, we expect to reject the Neyman's null hypothesis under smaller $p$-values. As a result, as we decrease the rejection threshold, the rejection rates should be smaller, which is supported by our simulation results in Figures D-25 and D-26.

# Bibliography


Abadie A, Athey S, Imbens GW, Wooldridge JM (2020) Sampling-based versus design-based uncertainty in regression analysis. *Econometrica* 88(1):265–296.

Adelman D (2007) Dynamic bid prices in revenue management. *Operations Research* 55(4):647–661.

Agrawal S, Avadhanula V, Goyal V, Zeevi A (2017) Mnl-bandit: a dynamic learning approach to assortment selection. *arXiv preprint arXiv:1706.03880* .

Ahn HS, Ryan C, Uichanco J, Zhang M (2019) Certainty equivalent pricing under sales-dependent and inventory-dependent demand. *Available at SSRN 3502478* .

Ailawadi KL, Neslin SA (1998) The effect of promotion on consumption: Buying more and consuming it faster. *Journal of Marketing Research* 390–398.

Amine A (1998) Consumers' true brand loyalty: the central role of commitment. *Journal of strategic marketing* 6(4):305–319.

Anupindi R, Dada M, Gupta S (1998) Estimation of consumer demand with stock-out based substitution: An application to vending machine products. *Marketing Science* 17(4):406–423.

Araman VF, Caldentey R (2009) Dynamic pricing for nonperishable products with demand learning. *Operations research* 57(5):1169–1188.

Arellano M (2003) *Panel data econometrics* (Oxford university press).

Arlotto A, Gurvich I (2019) Uniformly bounded regret in the multisecretary problem. *Stochastic Systems* 9(3):231–260.

Aronow PM, Samii C, et al. (2017) Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics* 11(4):1912–1947.

Assadi S, Hsu J, Jabbari S (2015) Online assignment of heterogeneous tasks in crowdsourcing markets. *Third AAAI Conference on Human Computation and Crowdsourcing*.

Athey S, Eckles D, Imbens GW (2018) Exact p-values for network interference. *Journal of the American Statistical Association* 113(521):230–240.

Athey S, Imbens GW (2018) Design-based analysis in difference-in-differences settings with staggered adoption. Technical report, National Bureau of Economic Research.

Azevedo EM, Alex D, Montiel Olea J, Rao JM, Weyl EG (2019) A/b testing with fat tails. *Available at SSRN 3171224* .

Badanidiyuru A, Kleinberg R, Slivkins A (2013) Bandits with knapsacks. *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, 207–216 (IEEE).

Bai Y (2019) Optimality of matched-pair designs in randomized control trials. *Available at SSRN 3483834* .

Ball MO, Queyranne M (2009) Toward robust revenue management: Competitive analysis of online booking. *Operations Research* 57(4):950–963.

Balseiro SR, Brown DB, Chen C (2019) Dynamic pricing of relocating resources in large networks. *Abstracts of the 2019 SIGMETRICS/Performance Joint International Conference on Measurement and Modeling of Computer Systems*, 29–30 (ACM).

Bansak K, Ferwerda J, Hainmueller J, Dillon A, Hangartner D, Lawrence D, Weinstein J (2018) Improving refugee integration through data-driven algorithmic assignment. *Science* 359(6373):325–329.

Basse G, Ding P, Feller A, Toulis P (2019a) Randomization tests for peer effects in group formation experiments. *arXiv preprint arXiv:1904.02308* .

Basse G, Ding Y, Toulis P (2019b) Minimax crossover designs. *arXiv preprint arXiv:1908.03531* .

Bell DR, Chiang J, Padmanabhan V (1999) The decomposition of promotional response: An empirical generalization. *Marketing Science* 18(4):504–526.

Berger JO (2013) *Statistical decision theory and Bayesian analysis* (Springer Science & Business Media).

Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research* 57(6):1407–1420.

Besbes O, Zeevi A (2012) Blind network revenue management. *Operations research* 60(6):1537–1550.

Bickel PJ, Doksum KA (2015) *Mathematical statistics: basic ideas and selected topics, volume I*, volume 117 (CRC Press).

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

Bojinov I, Rambachan A, Shephard N (2020a) Panel experiments and dynamic causal effects: A finite population perspective. *arXiv preprint arXiv:2003.09915* .

Bojinov I, Saint-Jacques G, Tingley M (2020b) Avoid the pitfalls of a/b testing make sure your experiments recognize customers' varying needs. *Harvard Business Review* 98(2):48–53.

Bojinov I, Shephard N (2019) Time series experiments and causal estimands: exact randomization tests and trading. *Journal of the American Statistical Association* 114(528):1665–1682.

Bojinov I, Tu Y, Liu M, Xu Y (2019) Causal inference from observational data: Estimating the effect of contributions on visitation frequency atLinkedIn. *arXiv preprint arXiv:1903.07755* .

Boruvka A, Almirall D, Witkiewitz K, Murphy SA (2018) Assessing time-varying causal effect moderation in mobile health. *Journal of the American Statistical Association* 113(523):1112–1121.

Bumpensanti P, Wang H (2018) A re-solving heuristic for dynamic resource allocation with uniformly bounded revenue loss. *arXiv preprint arXiv:1802.06192* .

Caro F, Gallien J (2012) Clearance pricing optimization for a fast-fashion retailer. *Operations Research* 60(6):1404–1422.

Chamandy N (2016) Experimentation in a ridesharing marketplace lyft engineering. *URL: https://eng.lyft.com/experimentation-in-a-ridesharing-marketplace-b39db027a66e* .

Chamberlain G (1982) Multivariate regression models for panel data. *Journal of econometrics* 18(1):5–46.

Chawla S, Hartline JD, Malec DL, Sivan B (2010) Multi-parameter mechanism design and sequential posted pricing. *Proceedings of the forty-second ACM symposium on Theory of computing*, 311–320 (ACM).

Chen Q, Jasin S, Duenyas I (2019) Nonparametric self-adjusting control for joint learning and optimization of multiproduct pricing with finite resource capacity. *Mathematics of Operations Research* 44(2):601–631.

Chen X, Ma W, Simchi-Levi D, Xin L (2016) Assortment planning for recommendations at checkout under inventory constraints. *Available at SSRN 2853093* .

Chen Y, Farias VF, Trichakis NK (2018) On the efficacy of static prices for revenue management in the face of strategic customers. *Management Science* .

Chen Y, Shi C (2019) Network revenue management with online inverse batch gradient descent method. *Available at SSRN* .

Cheung WC, Ma W, Simchi-Levi D, Wang X (2018) Inventory balancing with online learning. *arXiv preprint arXiv:1810.05640* .

Chin A (2018) Central limit theorems via Stein's method for randomized experiments under interference. *arXiv e-prints* arXiv:1804.03105.

Ciocan DF, Farias V (2012) Model predictive control for dynamic resource allocation. *Mathematics of Operations Research* 37(3):501–525.

Cooper WL (2002) Asymptotic behavior of an allocation policy for revenue management. *Operations Research* 50(4):720–727.

Correa J, Foncea P, Pizarro D, Verdugo V (2019) From pricing to prophets, and back! *Operations Research Letters* 47(1):25–29.

Cui R, Li J, Zhang D (2017) Discrimination with incomplete information in the sharing economy: Evidence from field experiments on airbnb. *Harvard Business School* 1–35.

Cui R, Zhang DJ, Bassamboo A (2019) Learning from inventory availability information: Evidence from field experiments on amazon. *Management Science* 65(3):1216–1235.

Cygan M, Jeż Ł, Sgall J (2016) Online knapsack revisited. *Theory of Computing Systems* 58(1):153–190.

Dean BC, Goemans MX, Vondrák J (2008) Approximating the stochastic knapsack problem: The benefit of adaptivity. *Mathematics of Operations Research* 33(4):945–964.

den Boer AV (2015) Dynamic pricing and learning: historical origins, current research, and new directions. *Surveys in operations research and management science* 20(1):1–18.

Deshpande Y, Mackey L, Syrgkanis V, Taddy M (2018) Accurate inference for adaptive linear models. *International Conference on Machine Learning*, 1194–1203 (PMLR).

Dütting P, Feldman M, Kesselheim T, Lucier B (2017) Prophet inequalities made easy: Stochastic optimization by pricing non-stochastic inputs. *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, 540–551 (IEEE).

Dwork C, Hardt M, Pitassi T, Reingold O, Zemel R (2012) Fairness through awareness. *Proceedings of the 3rd innovations in theoretical computer science conference*, 214–226 (ACM).

Eckles D, Karrer B, Ugander J (2016) Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference* 5(1).

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

Eren SS, Maglaras C (2010) Monopoly pricing with limited demand information. *Journal of revenue and pricing management* 9(1-2):23–48.

Farronato C, MacCormack A, Mehta S (2018) Innovation at uber: The launch of express pool. *Harvard Business School Case* 620(062).

Feldman M, Gravin N, Lucier B (2014) Combinatorial auctions via posted prices. *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, 123–135 (SIAM).

Ferreira KJ, Lee BHA, Simchi-Levi D (2016) Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management* 18(1):69–88.

Ferreira KJ, Simchi-Levi D, Wang H (2018) Online network revenue management using thompson sampling. *Operations research* 66(6):1586–1602.

Fisher RA, et al. (1937) The design of experiments. *The design of experiments.* (2nd Ed).

Gallego G, Iyengar G, Phillips R, Dubey A (2004) Managing flexible products on a network. *Available at SSRN 3567371* .

Gallego G, Li A, Truong VA, Wang X (2016) Online personalized resource allocation with customer choice. Technical report, Working Paper. http://arxiv. org/abs/1511.01837 v1.

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

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

Garg N, Nazerzadeh H (2019) Driver surge pricing. *arXiv preprint arXiv:1905.07544* .

Glynn P, Johari R, Rasouli M (2020) Adaptive experimental design with temporal interference: A maximum likelihood approach. *arXiv preprint arXiv:2006.05591* .

Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-time optimization of personalized assortments. *Management Science* 60(6):1532–1551.

Goyal V, Levi R, Segev D (2016) Near-optimal algorithms for the assortment planning problem under dynamic substitution and stochastic demand. *Operations Research* 64(1):219–235.

Gul F, Sonnenschein H, Wilson R (1986) Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory* 39(1):155–190.

Gupta A (2009) Prophets and secretaries .

Gupta S, Kamble V (2019) Individual fairness in hindsight. *Proceedings of the 2019 ACM Conference on Economics and Computation*, 805–806 (ACM).

Gupta S, Kohavi R, Tang D, Xu Y, Andersen R, Bakshy E, Cardin N, Chandran S, Chen N, Coey D, et al. (2019) Top challenges from the first practical online controlled experiments summit. *ACM SIGKDD Explorations Newsletter* 21(1):20–35.

Hadad V, Hirshberg DA, Zhan R, Wager S, Athey S (2019) Confidence intervals for policy evaluation in adaptive experiments. *arXiv preprint arXiv:1911.02768* .

Hajiaghayi MT, Kleinberg R, Sandholm T (2007) Automated online mechanism design and prophet inequalities. *AAAI*, volume 7, 58–65.

Han KW, Bojinov I, Basse G (2021) Population interference in panel experiments. *arXiv preprint arXiv:2103.00553* .

Han X, Kawase Y, Makino K (2015) Randomized algorithms for online knapsack problems. *Theoretical Computer Science* 562:395–405.

Harshaw C, Sävje F, Spielman D, Zhang P (2019) Balancing covariates in randomized experiments using the gram-schmidt walk. *arXiv preprint arXiv:1911.03071* .

Hedayat A, Afsarinejad K, et al. (1978) Repeated measurements designs, ii. *The Annals of Statistics* 6(3):619–628.

Ho CJ, Vaughan JW (2012) Online task assignment in crowdsourcing markets. *Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Hoeffding W, Robbins H (1948) The central limit theorem for dependent random variables. *Duke Mathematical Journal* 15(3):773–780.

Holtz D, Lobel R, Liskovich I, Aral S (2020) Reducing interference bias in online marketplace pricing experiments. *arXiv preprint arXiv:2004.12489* .

Honhon D, Gaur V, Seshadri S (2010) Assortment planning and inventory decisions under stockout-based substitution. *Operations research* 58(5):1364–1379.

Hwang D, Jaillet P, Manshadi V (2018) Online resource allocation under partially predictable demand. *arXiv preprint arXiv:1810.00447* .

Hyndman RJ, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E (2020) Package 'forecast'. *Online] https://cran. r-project. org/web/packages/forecast/forecast. pdf* .

Hyndman RJ, Khandakar Y, et al. (2007) *Automatic time series for forecasting: the forecast package for R*. Number 6/07 (Monash University, Department of Econometrics and Business Statistics . . . ).

Imai K, Kim IS (2019) When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data? *American Journal of Political Science* 63(2):467–490, URL `http://dx.doi.org/10.1111/ajps.12417`.

Imbens GW, Rubin DB (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction* (Cambridge University Press), URL `http://dx.doi.org/10.1017/CBO9781139025751`.

Jacoby J, Kyner DB (1973) Brand loyalty vs. repeat purchasing behavior. *Journal of Marketing research* 10(1):1–9.

Jasin S (2014) Reoptimization and self-adjusting price control for network revenue management. *Operations Research* 62(5):1168–1178.

Johari R, Li H, Weintraub G (2020) Experimental design in two-sided platforms: An analysis of bias. *arXiv preprint arXiv:2002.05670* .

Jones B, Kenward MG (2014) *Design and analysis of cross-over trials* (CRC press).

Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, 352–358.

Kastelman D, Ramesh R (2018) Switchback tests and randomized experimentation under network effects at doordash. *URL: https://medium.com/@DoorDash/switchback-tests-and-randomized-experimentation-under-network-effects-at-doordash-f1d938ab7c2a* .

Kempthorne O (1955) The randomization theory of experimental inference. *Journal of the American Statistical Association* 50(271):946–967, ISSN 01621459, URL `http://www.jstor.org/stable/2281178`.

Kleinberg R, Weinberg SM (2012) Matroid prophet inequalities. *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, 123–136 (ACM).

Kleywegt AJ, Papastavrou JD (1998) The dynamic and stochastic knapsack problem. *Operations research* 46(1):17–35.

Kohavi R, Crook T, Longbotham R, Frasca B, Henne R, Ferres JL, Melamed T (2009) Online experimentation at microsoft. *Data Mining Case Studies* 11(2009):39.

Kohavi R, Henne RM, Sommerfield D (2007) Practical guide to controlled experiments on the web: listen to your customers not to the hippo. *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 959–967.

Kohavi R, Tang D, Xu Y (2020) *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing* (Cambridge University Press).

Kohavi R, Thomke S (2017) The surprising power of online experiments. *Harvard Business Review* 95:74–82.

Koning R, Hasan S, Chatterji A (2019) Experimentation and startup performance: Evidence from a/b testing. Technical report, National Bureau of Economic Research.

Krengel U, Sucheston L (1977) Semiamarts and finite values. *Bulletin of the American Mathematical Society* 83(4):745–747.

Laird NM, Skinner J, Kenward M (1992) An analysis of two-period crossover designs with carry-over effects. *Statistics in Medicine* 11(14-15):1967–1979.

Lattimore T, Szepesvári C (2018) Bandit algorithms. *preprint* 28.

Li H, Webster S, Mason N, Kempf K (2018) Product-line pricing under discrete mixed multinomial logit demand: Winner—2017 m&som practice-based research competition. *Manufacturing & Service Operations Management* 21(1):14–28.

Li JQ, Rusmevichientong P, Simester D, Tsitsiklis JN, Zoumpoulis SI (2015) The value of field experiments. *Management Science* 61(7):1722–1740.

Li KC (1983) Minimaxity for randomized designs: some general results. *The Annals of Statistics* 11(1):225–239.

Li X, Ding P (2017) General forms of finite population central limit theorems with applications to causal inference. *Journal of the American Statistical Association* 112(520):1759–1769.

Li X, Ding P, Rubin DB, et al. (2020) Rerandomization in $2^k$ factorial experiments. *The Annals of Statistics* 48(1):43–63.

Liaw A, Wiener M, et al. (2002) Classification and regression by randomforest. *R news* 2(3):18–22.

Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ (2011) The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Personalized medicine* 8(2):161–173.

Liu Q, Van Ryzin G (2008) On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management* 10(2):288–310.

Ma W, Simchi-Levi D (2017) Algorithms for online matching, assortment, and pricing with tight weight-dependent competitive ratios. *to appear in Operations Research* .

Ma W, Simchi-Levi D, Teo CP (2018) On policies for single-leg revenue management with limited demand information. *arXiv preprint arXiv:1810.10900* .

Ma W, Simchi-Levi D, Zhao J (2020a) Dynamic pricing (and assortment) under a static calendar. *Management Science* .

Ma Y, Rusmevichientong P, Sumida M, Topaloglu H (2020b) An approximation algorithm for network revenue management under nonstationary arrivals. *Operations Research* 68(3):834–855.

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

Mahajan S, Van Ryzin G (2001) Stocking retail assortments under dynamic consumer substitution. *Operations Research* 49(3):334–351.

Mahdian M, Nazerzadeh H, Saberi A (2007) Allocating online advertisement space with unreliable estimates. *Proceedings of the 8th ACM conference on Electronic commerce*, 288–294 (ACM).

Mankiw NG (1985) Small menu costs and large business cycles: A macroeconomic model of monopoly. *The Quarterly Journal of Economics* 100(2):529–537.

March JG (1991) Exploration and exploitation in organizational learning. *Organization science* 2(1):71–87.

Marchetti-Spaccamela A, Vercellis C (1995) Stochastic on-line knapsack problems. *Mathematical Programming* 68(1-3):73–104.

McFadden D, Train K (2000) Mixed mnl models for discrete response. *Journal of applied Econometrics* 15(5):447–470.

Mehta A, Saberi A, Vazirani U, Vazirani V (2005) Adwords and generalized on-line matching. *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, 264–273 (IEEE).

Mehta A, et al. (2013) Online matching and ad allocation. *Foundations and Trends® in Theoretical Computer Science* 8(4):265–368.

Modaresi S, Saure D, Vielma JP (2019) Learning in combinatorial optimization: What and how to explore. *Available at SSRN 3041893* .

Motwani R, Raghavan P (1995) *Randomized algorithms* (Cambridge university press).

Neyman J (1923) On the application of probability theory to agricultural experiments. essay on principles. section 9. masters thesis. portion translated by d. dabrowska and t. speed (1990). *Statistical Science* 465–472.

Nie X, Tian X, Taylor J, Zou J (2018) Why adaptively collected data have negative bias and how to correct for it. *International Conference on Artificial Intelligence and Statistics*, 1261–1269.

Oman SD, Seiden E (1988) Switch-back designs. *Biometrika* 75(1):81–89.

Owen Z, Simchi-Levi D (2017) Price and assortment optimization for reusable resources. *Available at SSRN 3070625* .

Papastavrou JD, Rajagopalan S, Kleywegt AJ (1996) The dynamic and stochastic knapsack problem with deadlines. *Management Science* 42(12):1706–1718.

Puelz D, Basse G, Feller A, Toulis P (2019) A graph-theoretic approach to randomization tests of causal effects under general interference. *arXiv preprint arXiv:1910.10862* .

Rambachan A, Shephard N (2019) Econometric analysis of potential outcomes time series: instruments, shocks, linearity and the causal response function. *arXiv preprint arXiv:1903.01637* .

Robins J (1986) A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling* 7(9-12):1393–1512.

Roehm ML, Pullins EB, Roehm Jr HA (2002) Designing loyalty-building programs for packaged goods brands. *Journal of Marketing Research* 39(2):202–213.

Romano JP, Wolf M (2000) A more general central limit theorem for m-dependent random variables with unbounded m. *Statistics & probability letters* 47(2):115–124.

Rubin DB (1980) Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association* 75(371):591–593.

Rusmevichientong P, Sumida M, Topaloglu H (2020) Dynamic assortment optimization for reusable products with random usage durations. *Management Science* .

Samuel-Cahn E, et al. (1984) Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability* 12(4):1213–1216.

Sarasvathy SD (2001) Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of management Review* 26(2):243–263.

Senn S, Lambrou D (1998) Robust and realistic approaches to carry-over. *Statistics in Medicine* 17(24):2849–2864.

Sitkin SB (1992) Learning through failure: The strategy of small losses. *Research in organizational behavior* 14:231–266.

Slivkins A (2019) Introduction to multi-armed bandits. *arXiv preprint arXiv:1904.07272* .

Sobel ME (2012) Does marriage boost men's wages?: Identification of treatment effects in fixed effects regression models for panel data. *Journal of the American Statistical Association* 107(498):521–529.

Stamatopoulos I, Bassamboo A, Moreno A (2017) The effects of menu costs on retail performance: Evidence from adoption of the electronic shelf label technology. *Available at SSRN 3007095* .

Stein C, Truong VA, Wang X (2018) Advance service reservations with heterogeneous customers. *arXiv preprint arXiv:1805.05554* .

Sun T, Viswanathan S, Huang N, Zheleva E (2018) Designing promotional incentive to embrace social sharing: Evidence from field and lab experiments. *Available at SSRN 3095094* .

Sussman DL, Airoldi EM (2017) Elements of estimation theory for causal effects in the presence of network interference. *arXiv preprint arXiv:1702.03578* .

Talluri K, Van Ryzin G (1998) An analysis of bid-price controls for network revenue management. *Management science* 44(11-part-1):1577–1593.

Talluri KT, Van Ryzin GJ (2006) *The theory and practice of revenue management*, volume 68 (Springer Science & Business Media).

Thomke S (2001) Enlightened experimentation. the new imperative for innovation. *Harvard Business Review* 79(2):66–75.

Thomke SH (2020) *Experimentation Works: The Surprising Power of Business Experiments* (Harvard Business Press).

Topaloglu H (2009) Using lagrangian relaxation to compute capacity-dependent bid prices in network revenue management. *Operations Research* 57(3):637–649.

Veblen T (2017) *The theory of the leisure class* (Routledge).

Wager S, Xu K (2019) Experimenting in equilibrium. *arXiv preprint arXiv:1903.02124* .

Wang X, Truong VA, Bank D (2015) Online advance admission scheduling for services with customer preferences. *arXiv preprint arXiv:1805.10412* .

Wu CF (1981) On the robustness and efficiency of some randomized designs. *The Annals of Statistics* 1168–1177.

Xiong R, Athey S, Bayati M, Imbens GW (2019) Optimal experimental design for staggered rollouts. *Available at SSRN* .

Yao ACC (1977) Probabilistic computations: Toward a unified measure of complexity. *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, 222–227 (IEEE).

Zhang D, Cooper WL (2005) Revenue management for parallel flights with customer-choice behavior. *Operations Research* 53(3):415–431.

Zhou Y, Chakrabarty D, Lukose R (2008) Budget constrained bidding in keyword auctions and online knapsack problems. *International Workshop on Internet and Network Economics*, 566–576 (Springer).