Essays in Consumer Choice Driven Assortment Planning

Denis R. Saure

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the Executive Committee of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2011
ABSTRACT

Essays in Consumer Choice Driven Assortment Planning

Denis R. Saure

Product assortment selection is among the most critical decisions facing retailers: product variety and relevance is a fundamental driver of consumers’ purchase decisions and ultimately of a retailer’s profitability. In the last couple of decades an increasing number of firms have gained the ability to frequently revisit their assortment decisions during a selling season. In addition, the development and consolidation of online retailing have introduced new levels of operational flexibility, and cheap access to detailed transactional information.

These new operational features present the retailer with both benefits and challenges. The ability to revisit the assortment decision frequently over time allows the retailer to introduce and test new products during the selling season, and adjust on the fly to unexpected changes in consumer preferences, and use customer profile information to customize (in real time) online shopping experience.

Our main objective in this thesis is to formulate and solve assortment optimization models addressing the challenges present in modern retail environments. We begin by analyzing the role of the assortment decision in balancing information collection and revenue maximization, when consumer preferences are initially unknown. By considering utility maximizing consumers, we establish fundamental limits on the performance of any assortment policy whose aim is to maximize long run revenues. In addition, we propose adaptive assortment policies that attain such performance limits. Our results highlight salient features of this dynamic assortment problem that distinguish it from similar problems of sequential decision making under model uncertainty.

Next, we extend the analysis to the case when additional consumer profile information is available; our primary motivation here is the emerging area of online advertisement. As
in the previous setup, we identify fundamental performance limits and propose adaptive policies attaining these limits.

Finally, we focus on the effects of competition and consumers’ access to information on assortment strategies. In particular, we study competition among retailers when they have access to *common* products, i.e., products that are available to the competition, and where consumers have full information about the retailers’ offerings. Our results shed light on equilibrium properties in such settings and the effect common products have on this behavior.
Table of Contents

1 Introduction .................................................. 1
  1.1 Motivation ....................................................... 1
  1.2 Challenges in Assortment Planning ......................... 2
  1.3 Consumer Choice Driven Assortment Planning ............... 4
    1.3.1 Dynamic Learning of Consumer Preferences ............. 4
    1.3.2 Availability and use of Profile Information ............. 5
    1.3.3 Strategic Role of the Assortment Decision ............... 6
  1.4 Organization of the Thesis .................................. 7

2 Optimal Dynamic Assortment Planning .......................... 9
  2.1 Introduction .................................................... 9
  2.2 Literature Review .............................................. 13
  2.3 Problem Formulation ........................................... 17
  2.4 Fundamental Limits on Achievable Performance .............. 21
    2.4.1 A lower bound on the performance of any admissible policy
    2.4.2 Proof outline and intuition behind Theorem 1 .......... 22
  2.5 Dynamic Assortment Planning Policies ....................... 23
    2.5.1 Intuition and a simple “separation-based” policy ......... 24
    2.5.2 A refined dynamic assortment policy .................... 28
    2.5.3 A policy customized to the multinomial Logit choice model
  2.6 Concluding Remarks ............................................ 36
Bibliography 100

A Proof of Results in Chapter 2 108
A.1 Proof of Main Results .............................................. 108
A.2 Proof of Auxiliary Results ...................................... 123

B Proof of Results in Chapter 3 126
B.1 Proof of Main Results ............................................. 126
B.2 Proof of Auxiliary Results ...................................... 134

C Proof of Results in Chapter 4 141
C.1 Proofs for Section 4.4 .............................................. 141
C.2 Proofs for Section 4.5 .............................................. 146
C.3 Minimum Margins Setting ................................. 152
List of Figures

2.1 Performance of the separation-based policy $\pi_1$. ........................ 28
2.2 Performance of the refined policy $\pi_2$. ........................................ 32
2.3 Performance of the MNL-customized policy $\pi_3$. .............................. 37
2.4 Comparison with benchmark performance. ......................................... 39
3.1 Performance of the proposed policy $\pi(w, \kappa)$. .............................. 65
3.2 Display of suboptimal ad-mixes. .................................................... 66
List of Tables

3.1 Solution to static optimization problem for the numerical example. . . . . . 64
3.2 Minimum suboptimal experimentation for each ad. . . . . . . . . . . . . . 64
4.1 Illustration of non-existence of equilibrium in the setup of Example 3] . . . 88
Acknowledgments

Being a man of few words I usually let facts speak for themselves. This is, I am afraid, one of those situations in which I am obliged to speak for them. Luckily for you, I’ll go straight to the point, and with a little bit of luck, the outcome will be both concise and accurate.

This dissertation is largely the result of a fruitful collaboration with my advisor Professor Assaf Zeevi. I am deeply indebted to him for his support and guidance. I plan to pay him back by eating as many foxes as possible.

This thesis has benefited from a rather infamous collaborative effort with Professor/friend Omar Besbes. I sincerely appreciate all the feedback and experience you share with me. I hope that we continue to collaborate in the future.

I also want to thank the other members of my Thesis Committee. The alluded group, it is my honor to divulge, consist of Professors Garrett van Ryzin, Guillermo Gallego and Garud Iyengar. One can only pause for several minutes while trying to come up with a more distinguished and qualified committee.

I might thank as well all faculty at Columbia with whom I had the opportunity to interact with and learn from. In particular, I thank Professors Gabriel Weintraub and Nelson Fraiman for their constant support and unconditional friendship.

My original interest for academia developed in a sequence of rather spaced installments, due to the interaction with my former undergraduate mentors, Professors Andres Weintraub, Rafael Epstein and Rene Caldentey (chronologically ordered).

The life of the graduate student is not easy. I must thank my fellow DRO classmates for making most of this journey fun. I feel obliged, I am afraid, to highlight the contribution of two individuals. The first, a languid and loud Chilean woman, provided invaluable support during our first years in NYC. The second, a quite unconventional Turkish man, served as an excellent comrade from day one. Natalia and Deniz, thank you.
I must acknowledge, to finish up, the unconditional support provided by my family. I cannot find words, regrettably, to express my gratitude to them, thus I have decided to step away, rather momentarily, from the premise on which this essay is based. I cannot be excused, everyone will agree, from thanking the leading character in my life. Pepa, you are the reason I am who I am, and this accomplishment is without any doubt more yours than mine.
to Dante
Chapter 1

Introduction

1.1 Motivation

A retailer’s product assortment has been defined as “the set of products carried in each store at each point in time” (see Kok et al., 2008). Product assortment selection is among the most critical decisions facing retailers: product variety and relevance is a fundamental driver of consumers’ purchase decisions, and ultimately of a retailer’s profitability. Assortment planning has received significant attention from both retailers and academia. See Kok et al., 2008 for a review of the theory and practice of assortment planning.

Traditionally, retailers have been constrained to make assortment decisions well in advance of the selling season, when limited or no information on future demand is available. Moreover, these assortment decisions are also subject to several operational constraints such as limited procurement budget, limited shelf space for displaying products, to name a few. In this traditional setup the assortment planning literature has focused on optimizing procurement decisions while considering product substitution effects on demand realization (customers might decide to purchase a different product when their preferred one is unavailable). In this context, retailers face the challenge of mapping such procurement decisions to consumer behavior, as this can be used to deconstruct the complexity of the purchase decision.

In the last couple of decades we have witnessed substantial changes on the retail industry. An increasing number of firms have implemented fundamental changes in their supply chain
architectures that allow them to reduce procurement lead-times significantly. Such firms are now able to revisit their assortments decisions more frequently during a selling season. This ability, together with the advances in information technologies, have reduced the impact of initial procurement decisions on the retailer's revenues and shifted the focus towards the role of assortment planning on managing demand throughout the selling horizon. In that regard, early sales figures might be used to revisit initial assessments of consumer preferences, which – through assortment optimization – might help in tracking a potentially changing demand environment with a more flexible and dynamic product offering.

In addition to the above, development and consolidation of online retailing has reduced retailer operational restrictions that pertain to the assortment planning task. Products are no longer required to be stocked at each store, shelf space is now limited by the consumer's own interest, and the retailer might benefit from centralizing inventory management and pooling demand. It is within realm of possibility, that in the future, online retailers might leverage this operational flexibility to enable decisions on a consumer-by-consumer basis.

### 1.2 Challenges in Assortment Planning

**Dynamic Learning of Users Preferences.** The ability to revisit the assortment decision frequently over time allows the retailer to introduce and test new products during the selling season, and to better track changes in consumer preferences. Such a level of sophistication in the assortment decision making process increases its complexity substantially; and calls for the development of more advanced information and decision support systems.

When there is limited or no demand information available prior to the selling season, retailers need to learn user preferences by dynamically adjusting their product offerings and observing consumer behavior. Consider the case of fashion retail; each season there is a need to learn the current fashion trend by exploring with different styles and colors, and to exploit such knowledge before the season is over. Customers visiting one of these retailers' stores will only see a fraction of the potential array of products the retailer has to offer, and their purchase decisions will effectively depend on the specific assortment presented at the store.
A fundamental problem arises in this setting. On the one hand, the longer a retailer spends learning consumer preferences, the less time remains to exploit that knowledge and optimize profits. On the other hand, less time spent studying consumer behavior translates into more residual uncertainty in the assortment selection process, which could hamper the revenue maximization objective. This trade-off is known as exploration versus exploitation.

**Availability and use of Profile Information.** The use of world-wide-web presents additional opportunities and challenges. Unlike traditional commerce channels, where customers remain largely anonymous, the internet allows collection of profile information at the consumer level. Such profile information might be provided by the user (e.g., by filling a subscription or survey), or collected without the consumer’s explicit consent (e.g., through internet cookies). This information is invaluable in customizing (in real time) the consumers’ online shopping experience.

The online advertisement industry has pioneered the use of profile information for targeting purposes. There, the most prevalent business models builds on the cost-per-click statistic: upon each visit to a publisher’s web-site, users are presented with a customized assortment of advertisements (henceforth, ads); and advertisers will pay a fee to the publisher each time their ads are clicked upon. To take advantage of user profiles for real time customization, one requires sophisticated and automated decision support systems.

**Strategic Role of the Assortment Decision.** The online commerce channel offers new opportunities to consumers as well: consumers have access to accurate information about retailers’ assortments and prices; they can compare retailers’ offerings in real time; and make purchase decisions, while considering product offerings from all retailers simultaneously.

The way consumers access and use information about the retailers’ offerings plays a critical role in the way consumer purchase decisions are made, and hence impacts the manner by which retailers interact and compete. Previous academic work on assortment competition has focused on the case where access to information is costly. There, consumers might make purchase decisions based only on partial knowledge about retailers’ offerings, usually engaging in hierarchical decision processes (for example, they might first decide on the retailer and then on the product to purchase, or viceversa).

When information about the retailers’ offering is readily available, consumers can form
preferences over the full set of retailer-product pairs, and make purchasing decisions accordingly. For example, when several retailers offer the same product, consumers effectively recognize such a product as being identical and make a choice based on retailer-specific attributes.

The challenge here is to understand the strategic role of the assortment decision and the effect of common products, offered by various retailers simultaneously, on competition.

### 1.3 Consumer Choice Driven Assortment Planning

Our main objective in this thesis is to formulate and solve assortment optimization models capturing the main features of the challenges identified in Section 1.2. Our first step in that direction is to specify an underlying model of consumer preferences. In this work we consider random utility models for consumer choice (see Anderson et al. [1992]). Utility based choice models assume each consumer assigns a (random) utility to each product; when offered an assortment, every consumer chooses the option providing him/her the highest utility, with an outside no-purchase option being available. Discrete choice models are extensively discussed in Ben-Akiva and Lerman (1985) and Train (2002) and the estimation and design of specialized models has been an active area in Marketing (see, e.g., Guadagni and Little (1983) for an early reference). Kok et al. (2008) provides a detailed review on the use of consumer choice models in the context of assortment planning.

Next, we describe the different formulations considered in this thesis, each addressing a different challenge in assortment planning.

#### 1.3.1 Dynamic Learning of Consumer Preferences

In Chapter 2 we focus on the impact of learning consumer behavior via suitable assortment experimentation, and on doing this in a manner that guarantees revenue maximization over the selling horizon. For that purpose we consider a population of homogeneous utility maximizing customers. Given limited display capacity and no prior information on consumers’ utility, the retailer needs to devise an assortment policy to maximize revenues over the relevant time horizon. By offering different assortments and observing the resulting purchase
behavior, the retailer learns about consumer preferences, but this experimentation should be balanced with the goal of maximizing revenues. Our stylized formulation of the dynamic assortment problem will ignore inventory considerations, additional costs (such as assortment switching costs), operational constraints (e.g. restrictions on the sequence of offered assortments) and changing demand patterns.

Chapter 2 makes several contributions:

i.) We establish fundamental bounds on the performance of any assortment policy. Specifically, we identify the magnitude of the revenue loss that any policy must incur relative to that of a retailer with prior knowledge of consumer preferences, and its dependence on the length of the selling horizon, the number of products, and the capacity constraint that limit the number of simultaneous products that can be presented.

ii.) We propose a family of adaptive policies that achieve the fundamental bounds mentioned above. These policies identify the optimal assortment of products (the one that maximizes the expected single sale profit) with high probability while successfully limiting the extent of exploration.

Chapter 2 also highlights salient features of the dynamic assortment problem that distinguish it from similar problems of sequential decision making under model uncertainty, and shows how exploiting these features helps to dramatically decrease the complexity of the assortment problem, relative to using existing non-customized strategies, e.g., from the multi-armed bandit literature.

1.3.2 Availability and use of Profile Information

Chapter 3 extends the analysis in Chapter 2 to the case when additional profile information about consumers is available. Here we are motivated by the emerging area of online advertisement. The analysis focuses on the impact of learning advertisement effectiveness via suitable assortment experimentation across different consumer/user profiles, and doing this in a manner that guarantees revenue maximization over the duration of the underlying advertisement contracts.
For that purpose we consider a population of heterogeneous (with a priori known class membership) utility maximizing visitors to a web page. Given no prior information on their utility, the publisher needs to determine, based on each user’s profile information, the ad assortment to be displayed. In our setup, users’ click decisions depend, among other factors, on each of the ads’ mean utility, which we assume is a function of the user’s profile information. By offering different assortments of ads to different types of users and observing click decisions, the publisher learns the relationship between mean utilities and user profiles. Such profile/ad experimentation should be balanced with the goal of revenue maximization.

Chapter 3 makes several contributions:

i.) We identify the magnitude of the revenue loss that any ad-mix policy must incur relative to that of a publisher with prior knowledge of consumer preferences. We show that such gain in revenue is proportional to the (minimum) amount of information needed to reconstruct the policy used by a publisher with prior knowledge of consumer preferences.

ii.) We propose a family of adaptive assortment policies that achieve the fundamental bound mentioned above. These policies focus on identifying, good user segments that can be best used for purposes of parameter estimation; these are selected so as to minimize the cost of information gathering. Exploration efforts are then focused primarily on these segments, at a frequency guided by the fundamental performance limit mentioned previously.

The results in Chapter 3 shed light on the potential benefits of using information arising from interactions with a subset of user segments, to better explain consumer preferences across all user types.

1.3.3 Strategic Role of the Assortment Decision

Chapter 4 steps away from real-time learning of consumer preferences and focuses on studying the effects of competition and consumers' access to information on assortment strategies. In particular, we aim to understanding how competition that results from the availability
CHAPTER 1. INTRODUCTION

of full information about the retailers’ offerings impacts equilibrium behavior. To that end, we analyze a model of assortment and pricing competition in a duopolistic setting, when assortment decisions are constrained by limited display capacities and retailers have access to both common products, i.e., products that are available to both competitors, and exclusive products, i.e., products that are unavailable to competition. We anchor the analysis around the well-studied Multinomial Logit choice model for consumer demand.

Chapter 4 makes several contributions:

i.) We provide a framework to analyze competition in the presence of informed consumers when demand is driven by individual utility maximization. This framework enables, among other things, the possibility of having retailers offer the same products, a possibility ignored in most of the existing literature. Our results shed light on equilibrium properties in general and on the implications of the availability of common products on such properties.

ii.) From a qualitative perspective, we establish a clear connection, in any equilibrium, between the attractiveness of the offered assortment (this concept is to be defined formally, but can be thought of as the breadth of an assortment), the profit made by a retailer, and the attractiveness of the competitor’s assortment. In addition we establish sharp uniqueness and existence results for two competitive settings: i.) the case where retailers select only their assortments given exogenously fixed prices; and ii.) the case where retailers can select both assortments and prices.

Chapter 4 shows that the introduction of common products can lead to fundamentally different properties for the equilibrium set. In addition, the unified framework we propose is fairly flexible and enables one to analyze different models of competition under a common approach.

1.4 Organization of the Thesis

The remainder of this document is organized as follows: Chapter 2 develops dynamic assortment policies for learning consumer preferences for the case of homogenous consumers.
CHAPTER 1. INTRODUCTION

Focusing on the online advertisement application area, Chapter 3 extends the results in Chapter 2 to the case of heterogeneous consumers, when additional consumer profile information is available. Finally, Chapter 4 studies assortment and price competition when information about the retailers' offerings is readily available. Each chapter is designed to be self-contained. Proofs of the results in Chapters 2, 3 and 4 are deferred to appendices A, B and C respectively.
Chapter 2

Optimal Dynamic Assortment Planning

2.1 Introduction

**Motivation and main objectives.** Product assortment selection is among the most critical decisions facing retailers. Inferring customer preferences and responding accordingly with updated product offerings plays a central role in a growing number of industries, especially for companies that are capable of revisiting product assortment decisions *during* the selling season, as demand information becomes available. From an operations perspective, a retailer is often not capable of simultaneously displaying every possible product to prospective customers due to limited shelf space, stocking restrictions and other capacity related considerations. One of the central decisions is therefore which products to include in the retailer’s product assortment. That is the essence of the assortment planning problem; see [Kok et al. (2008)] for an overview. Our interest lies in *dynamic* instances of the problem where assortment planning decisions can be revisited frequently throughout the selling season (these could correspond to periodic review schedules, for example). These instances will be referred to as *dynamic assortment planning* problems. Here are two motivating examples that arise in very different application domains.

*Example 1: Fast fashion.* In recent years “fast” fashion companies, such as Zara, Mango or World co, have implemented highly flexible and responsive supply chains that allow them
to make and revisit most product design and assortment decisions during the selling season. Customers visiting one of these retailers’ stores will only see a fraction of the potential products that the retailer has to offer, and their purchase decisions will effectively depend on the specific assortment presented at the store. The essence of fashion retail entails offering new products for which no demand information is available, and hence the ability to revisit these decisions at a high frequency is key to the “fast fashion” business model; each season there is a need to learn the current fashion trend by exploring with different styles and colors, and to exploit such knowledge before the season is over.

Example 2: Online advertising. This emerging area of business is the single most important source of revenues for thousands of web sites. Giants such as Yahoo and Google, depend almost completely on online advertisement to subsist. One of the most prevalent business models here builds on the cost-per-click statistic: advertisers pay the web site (a “publisher”) only when a user clicks on their advertisements (henceforth, ads). Upon each visit, users are presented with a finite set of ads, on which they may or may not click depending on what is being presented. Roughly speaking, the publisher’s objective is to learn ad click-through-rates (and their dependence on the set of ads being displayed) and present the set of ads that maximizes revenue within the life span of the contract with the advertiser.

The above motivating applications share common features: i.) a priori information on consumer purchase/click behavior is scarce or non-existent; ii.) products/ads are potentially substitutes, hence individual product/ad demand is affected by the assortment decision, which is relevant due to display constraints; iii.) assortment decisions can be taken in a dynamic fashion during a limited time framework.

When there is limited or no demand information available a priori, retailers must learn new products desirability/effectiveness by dynamically adjusting their product offering and observing consumer behavior. A fundamental problem arises in this setting: information collection must balance the cost associated to it and the benefits coming from having a better picture of the demand. This is the classical exploration versus exploitation trade-off: on the one hand, the longer a retailer spends learning consumer preferences, the less time remains to exploit that knowledge and optimize profits. On the other hand, less time spent on studying consumer behavior translates into more residual uncertainty, which could hamper
the revenue maximization objective. Moreover, demand information must be gathered carefully as product profitability depends on the assortments offered: the retailer/publisher may learn consumer preferences more effectively by experimenting with a particular set of assortments.

A comprehensive dynamic assortment policy will balance the aforementioned exploration versus exploitation trade-off while facing important additional operational considerations: fast fashion retailers must also consider pricing decisions, inventory replenishment, display constraints, etc; online publishers must also consider availability of users’ profile information, ads’ minimum display requirements, etc.

This chapter aims to isolate the role of assortment planning in balancing information collection and revenue maximization. We consider a family of stylized dynamic assortment problems in settings where display capacity is limited. Fisher and Vaidyanathan (2009) elaborates on the importance of considering displays constraints in assortment planning. Our formulation of the dynamic assortment problem will ignore inventory considerations, additional costs (such as assortment switching costs), operational constraints (e.g. restrictions on the sequence of assortments offered) and changing demand patterns. It is worth noting that although most of the operational constraints mentioned above are absent almost altogether from the online advertisement problem, they play an important role in shaping the overall operational strategy of fast fashion retailers. The work of Caro and Gallien (2007) presents a first attempt to operationalize assortment policies coming from the analysis of an unconstrained model, like ours.

We assume that product prices are fixed throughout the selling season. Such an assumption is common in the assortment planning literature and facilitates the analysis of our formulation. While pricing has been studied in the context of choice-based demand with limited prior information (see, e.g., Rusmevichientong and Broder (2010)), incorporating a pricing dimension into our formulation would prevent us from gaining insight regarding the role of assortment experimentation in demand inference.

Our main focus is on the impact of learning consumer behavior via suitable assortment experimentation, and doing this in a manner that guarantees revenue maximization over the selling horizon. For that purpose we consider a population of utility maximizing cus-
CHAPTER 2. OPTIMAL DYNAMIC ASSORTMENT PLANNING

tomers: each customer assigns a (random) utility to each offered product, and purchases the product that maximizes his/her utility. The retailer needs to devise an assortment policy to maximize revenues over the relevant time horizon by properly adapting the assortment offered based on observed customer purchase decisions and subject to capacity constraints that limit the size of the assortment.

Key insights and qualitative results. We consider assortment policies that can only use observed purchase decisions to adjust assortment choices at each point in time (this will be defined more formally later as a class of non-anticipating policies). Performance of such policies will be measured in terms of the expected revenue loss relative to an oracle that knows the product utility distributions in advance, i.e., the loss due to the absence of a priori knowledge of consumer behavior. Our objective is to characterize the minimum loss attainable by any non-anticipating assortment policy.

The main findings of this chapter are summarized below.

i.) We establish fundamental bounds on the performance of any policy. Specifically, we identify the magnitude of the loss, relative to the oracle performance, that any policy must incur in terms of its dependence on: the length of the selling horizon; the number of products; and the capacity constraint (see Theorem 1 for a precise statement).

ii.) We propose a family of adaptive policies that achieve the fundamental bound mentioned above. These policies quickly identify the optimal assortment of products (the one that maximizes the expected single sale profit) with high probability while successfully limiting the extent of exploration. Our performance analysis, in Section 2.5.2, makes these terms rigorous; see Theorem 3.

iii.) We prove that not all products available to the retailer need to be extensively tested: under mild assumptions, some of them can be easily and quickly identified as suboptimal. In particular, a specific subset of said products can be detected after a small number of experiments (independent of the length of the selling horizon); see Theorems 1 and 3. Moreover, we show that our proposed policy successfully limits the extent of such an exploration (see Corollary 1 for a precise statement).

iv.) We highlight salient features of the dynamic assortment problem that distinguish it
from similar problems of sequential decision making under model uncertainty, and we show how exploiting these features helps to dramatically decrease the complexity of the assortment problem, relative to using existing non-customized strategies, e.g., from the multi-armed bandit literature.

On a more practical side, our results establish that an oracle with advance knowledge of customer behavior only gains additional revenue on the order of the logarithm of the total number of customers visiting the retailer during the selling season. Moreover, we establish that this is a fundamental price that any feasible assortment policy must pay. Regarding the exploration versus exploitation trade-off, we establish the precise frequency and extent of assortment experimentation that guarantee this best achievable performance. While in general it is necessary to experiment with inferior products at a precise and critical frequency that is increasing with the time horizon, for a certain subset of these products experimentation can be kept to a minimum (a bounded number of trials independent of the time horizon). This result differs markedly from most of the literature on similar sequential decision making problems.

The remainder of the chapter. The next section reviews related work. Section 2.3 formulates the dynamic assortment problem. Section 2.4 provides a fundamental limit on the performance of any assortment policy, and analyzes its implications for policy design. Section 2.5 proposes a dynamic assortment algorithm that achieves this performance bound, and Subsection 2.5.3 customizes our proposed algorithm for the most widely used customer choice model, namely the Logit. Finally, Section 2.6 presents our concluding remarks. Proofs are relegated to two appendices, A.1 and A.2.

2.2 Literature Review

Static assortment planning. The static planning literature focuses on finding an optimal assortment that is held unchanged throughout the entire selling season. Customer behavior is assumed to be known a priori, but inventory decisions are considered; see Kok et al. (2008) for a review of the state-of-the-art in static assortment optimization. Within this area, van Ryzin and Mahajan (1999) formulate the assortment planning problem using a
CHAPTER 2. OPTIMAL DYNAMIC ASSORTMENT PLANNING

Multinomial Logit model (hereafter, MNL) of consumer choice. Assuming that customers
do not look for a substitute if their choice is stocked out (known as static substitution), they
prove that the optimal assortment is always in the “popular assortment set” and establish
structural properties of the optimal assortment and ordering quantities. In the same setting,
[108x673]Gaur and Honhon (2006) use the locational choice model and characterize properties of the
optimal solution under static substitution. In a recent paper [Goyal et al. (2009)] prove that
the assortment problem is NP-hard, in the static setting when stock-out based substitution
is allowed, and propose a near-optimal heuristic solution for a particular choice model; see
also [Mahajan and van Ryzin, 2001], [Honhon et al., 2009] and [Hopp and Xu, 2008].

Our formulation assumes perfect replenishment (and hence eliminates stock-out based
substitution considerations) while considering limited display capacity. Fisher and Vaidyanathan
(2009) studies assortment planning under display constraints and highlights how these arise
in practice. It is important to note that even in this setting the static one-period profit
maximization problem remains NP-hard in general; see [Goyal et al. (2009)]. The work
of Rusmevichientong et al. (2010) identifies a polynomial-time algorithm for the static
optimization problem when consumer preferences are represented using particular choice
models; hence at least in certain instances the problem can be solved efficiently.

**Dynamic assortment planning.** This problem setting allows revisiting assortment de-
cisions at each point in time as more information is collected about initially unknown
demand/consumer preferences. To the best of our knowledge [Caro and Gallien (2007)] were
the first to study this type of problem, motivated by an application in fast fashion. In their
formulation, customer demand for a product is exogenous, and independent of demand and
availability for other products. The rate of demand is constant throughout the selling sea-
son, and their formulation ignores inventory considerations. Taking a Bayesian approach
to demand learning, they study the problem using dynamic programming. They derive
bounds on the value function and propose an index-based policy that is shown to be near
optimal when there is certain prior information on demand. Closer to our formulation is
the work by [Rusmevichientong et al. (2010)]. There, utility maximizing customers make
purchase decisions according to the MNL choice model (a special case of the more general
setting treated in the present document), and an adaptive algorithm for joint parameter
estimation and assortment optimization is developed, see further discussion below.

**Related work in dynamic optimization with limited demand information.** Uncertainty at demand-model level has been previously studied in revenue management settings. Araman and Caldentey (2009) and Farias and van Roy (2010) for example, present dynamic programming formulations with Bayesian updating of initially unknown parameters. Closer to our approach to model uncertainty, the work by Besbes and Zeevi (2009) considers a single product firm that, given an initial inventory, needs to dynamically adjust prices so as to maximize the cumulative expected revenue over a finite horizon, when demand function is initially unknown and no prior information is available. In a similar setting, Rusmevichientong and Broder (2010) analyze the case when demand is driven by a general parametric choice model.

All studies above focus on the pricing decision problem. It is worth mentioning that arguments used to establish fundamental limits in performance for the pricing problem do not apply to our setting, due to the discrete and combinatorial nature of the assortment decision. Also, the typical solution approach in these studies, which involves evaluating a finite set of alternatives in order to reconstruct demand function, would be highly inefficient if applied to the dynamic assortment problem: see the discussion following Theorem 1. Similar formulations of revenue management problems can be found, for example, in Rusmevichientong et al. (2006) and Lim and Shanthikumar (2007).

**Connection to the multi-armed bandit literature.** In the canonical multi-armed bandit problem the decision maker can select in each period to pull a single arm out of a set of $K$ possible arms, where each arm delivers a random reward whose distribution is not known a priori, and the objective is to maximize the revenue over a finite horizon. See Lai and Robbins (1985) and Auer et al. (2002) for a classical formulation and solution approach to the problem, respectively.

The model of Caro and Gallien (2007) is in fact equivalent to a multi-armed bandit problem with multiple simultaneous plays. The dynamic programming formulation and the Bayesian learning approach aims to solve the exploration versus exploitation trade-off optimally. See also Farias and Madan (2009) for a similar bandit-formulation with multiple simultaneous plays under more restricted type of policies. In the work of Rusmevichientong
et al. (2010) the connection is less straightforward. Their proposed algorithm works in cycles. Each cycle mixes parameter estimation (exploration) and assortment optimization (exploitation). In the exploration phase order $N^2$ assortments are tested, where $N$ is the number of products. Parameter estimates based on the exploration phase are fed into the static optimization problem, which returns order $N$ assortments among which the optimal one is found with high probability. From there, a standard multi-armed bandit algorithm is prescribed to find the optimal assortment, and an upper bound on the regret of order $N^2 \log^2 T$ is established, where $T$ is the length of the planning horizon.

There is a thematic connection between multi-armed bandits and assortment planning problems, in the sense that both look to balance exploration and exploitation. However, the fact that product utility does not map directly to retailer revenues in the dynamic assortment problem is essentially what distinguishes these problems. In the bandit setting all products are ex-ante identical, and only individual product exploration allows the decision maker to differentiate them. Nevertheless, there is always the possibility that a poorly explored arm is in fact optimal. This last fact prevents limiting exploration on arms that have been observed to be empirically inferior. (In their seminal work, Lai and Robbins (1985) showed that good policies should explore each arm at least order $\log T$ times.) In the assortment planning setting, products are not ex-ante identical, and product revenue is capped by its profit margin. In Section 2.4 we show how this observation can be exploited to limit exploration on certain suboptimal products (a precise definition will be advanced in what follows). Moreover, the possibility to test several products simultaneously has the potential to further reduce the complexity of the assortment planning problem. Our work builds on some of the ideas present in the multi-armed bandit literature, most notably the lower bound technique developed by Lai and Robbins (1985) but also exploits salient features of the assortment problem in constructing optimal algorithms and highlighting key differences from traditional bandit results.
CHAPTER 2. OPTIMAL DYNAMIC ASSORTMENT PLANNING

2.3 Problem Formulation

Model primitives and basic assumptions. We consider a price-taking retailer that has \( N \) different products to sell. For each product \( i \in \mathcal{N} := \{1, \ldots, N\} \), let \( r_i \) and \( c_i \) denote the price and the marginal cost of product \( i \), respectively. As mentioned in Section 2.1 we assume both prices and marginal cost are fixed and constant throughout the selling horizon. For \( i \in \mathcal{N} \), let \( w_i := r_i - c_i > 0 \) denote the marginal profit resulting from selling one unit of the product, and let \( w := (w_1, \ldots, w_N) \) denote the vector of profit margins. Due to display space constraints, the retailer can offer at most \( C \) products simultaneously.

Let \( T \) to denote the total number of customers that arrive during the selling season after which sales are discontinued. (The value of \( T \) is in general not known to the retailer a priori.) We use \( t \) to index customers according to their arrival times, so \( t = 1 \) corresponds to the first arrival, and \( t = T \) the last. We assume the retailer has both a perfect replenishment policy, and the flexibility to offer a different assortment to every customer without incurring any switching cost. (While these assumptions do not typically hold in practice, they provide sufficient tractability and allow us to extract structural insights.)

With regard to demand, we will adopt a random utility approach to model customer preferences over products: customer \( t \) assigns a utility \( U_{i}^{t} \) to product \( i \), for \( i \in \mathcal{N} \cup \{0\} \), with

\[
U_{i}^{t} := \mu_{i} + \zeta_{i}^{t},
\]

were \( \mu_{i} \in \mathbb{R} \) denotes the mean utility assigned to product \( i \), \( \zeta_{i}^{1}, \ldots, \zeta_{i}^{T} \) are independent random variables drawn from a distribution \( F \) common to all customers, and product 0 represents a no-purchase alternative.

A more general model would consider \( \mu_{i} := \beta^T x_{i} - \alpha r_{i} \), where \( x_{i} \) denotes the profile vector of attributes associated to product \( i \), \( \beta \) denotes the vector of part-worths measuring the relative importance consumers assign to each product attribute, and \( \alpha \) measures price elasticity. In such a setup the retailer might evaluate offering any feasible product design, and not only the ones initially considered in \( \mathcal{N} \). While such a model of consumer preferences could be analyzed using the techniques presented in this chapter, it would not add to understanding the complexity of the assortment problem, hence is not pursued in this
work.

Let $\mu := (\mu_1, \ldots, \mu_N)$ denote the vector of mean utilities. We assume all customers assign $\mu_0$ to a no-purchase alternative; when offered an assortment, customers select the product with the highest utility if that utility is greater than the one provided by the no-purchase alternative. For convenience, and without loss of generality, we set $\mu_0 := 0$.

### The static assortment optimization problem

Let $S$ denote the set of possible assortments, i.e., $S := \{S \subseteq N : |S| \leq C\}$, where $|S|$ denotes the cardinality of the set $S \subset N$. For a given assortment $S \in S$ and a given vector of mean utilities $\mu$, the probability $p_i(S, \mu)$ that a customer chooses product $i \in S$ is

$$p_i(S, \mu) = \int_{-\infty}^{\infty} \prod_{j \in S \cup \{0\} \setminus \{i\}} F(x - \mu_j) \ dF(x - \mu_i),$$

and $p_i(S, \mu) = 0$ for $i \notin S$. The expected profit $r(S, \mu)$ associated with an assortment $S$ and mean utility vector $\mu$ is given by

$$r(S, \mu) = \sum_{i \in S} w_i p_i(S, \mu).$$

If the retailer knows the value of the vector $\mu$, then it is optimal to offer $S^*(\mu)$, the solution to the static optimization problem, to every customer:

$$S^*(\mu) \in \arg\max_{S \in S} r(S, \mu).$$

In what follows we will assume that the solution to the static problem is unique (this assumption simplifies our exposition of fundamental performance bounds, and can be relaxed by redefining $S^*(\mu)$ accordingly). Efficiently solving problem (2.2) is beyond the scope of our work here: we will assume that the retailer can compute $S^*(\mu)$ for any vector $\mu$.

### Remark 1 (Complexity of the static problem)

We note that for specific utility distributions there exist efficient algorithms for solving the static problem. For example, Rusmevichientong et al. (2010) present an order $N^2$ algorithm to solve the static problem when an MNL choice model is assumed, i.e., when $F$ is assumed to be a standard Gumbel distribution (location parameter 0 and scale parameter 1) for all $i \in N$. This is an important contribution given that the MNL is by far the most commonly used choice model. The
algorithm, based on a more general solution concept developed by [Megiddo (1979)] can in fact be used to solve the static problem efficiently for any attraction-based choice model.

**The dynamic optimization problem.** We assume that the retailer knows $F$, the distribution that generates the idiosyncracies of customer utilities, but *does not know* the mean vector $\mu$.

The retailer is able to observe purchase/no-purchase decisions made by each customer. S/he needs to decide what assortment to offer to each customer, taking into account all information gathered up to that point in time, in order to maximize expected cumulative profits. More formally, let $(S_t \in \mathcal{S} : 1 \leq t \leq T)$ denote an *assortment process*, with $S_t \in \mathcal{S}$ for all $t \leq T$. Let

$$Z_t^i := {1 \{ i \in S_t, U_{ti} > U_{tj}, j \in S_t \setminus \{i\} \cup \{0\} \}$$

denote the purchase decision of customer $t$ regarding product $i \in S_t$, where $Z_t^i = 1$ indicates that customer $t$ decided to purchase product $i$, and $Z_t^i = 0$ otherwise. Also, let $Z_0^i := 1 \{ U_0 > U_j, j \in S_t \}$ denote the overall purchase decision of customer $t$, where $Z_0^i = 1$ if customer $t$ opted not to purchase any product, and $Z_0^i = 0$ otherwise. Here, and in what follows, $1 \{ A \}$ denotes the indicator function of a set $A$. We denote by $Z_t := (Z_0^t, Z_1^t, \ldots, Z_N^t)$ the vector of purchase decisions of customer $t$. Let $\mathcal{F}_t = \sigma((S_u, Z_u), 1 \leq u \leq t) \ t = 1, \ldots, T$, denote the filtration or history associated with the assortment process and purchase decisions up to (including) time $t$, with $\mathcal{F}_0 = \emptyset$. An admissible assortment policy $\pi$ is a mapping from past history to assortment decisions such that the associated assortment process $(S_t \in \mathcal{S} : 1 \leq t \leq T)$ is non-anticipating (i.e., $S_t$ is $\mathcal{F}_{t-1}$-measurable, for all $t$). We will restrict attention to the set of such policies and denote it by $\mathcal{P}$. We will use $\mathbb{E}_\pi$ and $\mathbb{P}_\pi$ to denote expectations and probabilities of random variables, when the assortment policy $\pi \in \mathcal{P}$ is used.

The retailer’s objective is to choose a policy $\pi \in \mathcal{P}$ to maximize the expected cumulative revenues over the selling season

$$J^\pi(T, \mu) := \mathbb{E}_\pi \left[ \sum_{t=1}^{T} \sum_{i \in \mathcal{N}} w_i Z_t^i \right].$$

[1] These are choice models for which $p_i(S) = v_i/\left(\sum_{j \in S} v_j\right)$ for a vector $v \in \mathbb{R}^N_+$, and any $S \subseteq \mathcal{N}$. (see, for example, Anderson et al. (1992).)
If the mean utility vector $\mu$ is known at the start of the selling season, the retailer would find the assortment that maximizes the one-sale expected value, namely $S^*(\mu)$, and would offer it to every customer. The corresponding performance, denoted by $J^*(T, \mu)$, is given by

$$J^*(T, \mu) := \text{Tr}(S^*(\mu), \mu). \quad (2.3)$$

This quantity provides an upper bound on expected revenues generated by any admissible policy, i.e., $J^*(T, \mu) \geq J^\pi(T, \mu)$ for all $\pi \in \mathcal{P}$. With this in mind we define the regret $R^\pi(T, \mu)$ associated with a policy $\pi$, to be

$$R^\pi(T, \mu) := T - \frac{J^\pi(T, \mu)}{r(S^*(\mu), \mu)}.$$

The regret measures to the number of customers to whom non-optimal assortments are offered by $\pi$ over $\{1, \ldots, T\}$. One may also view this as a normalized measure of revenue loss due to the lack of a priori knowledge of consumer behavior.

Maximizing expected cumulative revenues is equivalent to minimizing the regret over the selling season, and to this end, the retailer must balance suboptimal demand exploration (which adds directly to the regret) with exploitation of the gathered information. On the one hand the retailer has incentives to explore demand extensively in order to guess the optimal assortment $S^*(\mu)$ with high probability. On the other hand the longer the retailer explores the less consumers will be offered a supposedly optimal assortment, and therefore the retailer has incentives to shorten the length of the exploration phase in favor to the exploitation phase.

**Remark 2 (Relationship to bandit problems).** One can try to interpret the assortment problem as a bandit problem in, at least, two ways: envisioning each assortment as an arm (making rewards dependent across arms); and envisioning each product as an arm (resulting in a bandit problem with multiple simultaneous plays where rewards are not directly observable, as only consumer purchase is observed). In that sense, this work shows the inefficiency of the former approach and shows how to overcome efficiently the obstacles present in the later approach. It worth noting that regardless the definition of an arm, these
are a priori distinguishable (due to differences in profit margins), a feature absent in the classical bandit formulation.)

2.4 Fundamental Limits on Achievable Performance

2.4.1 A lower bound on the performance of any admissible policy

Let us begin by narrowing down the set of “interesting” policies worthy of consideration. We say that an admissible policy is consistent if for all \( \mu \in \mathbb{R}^N \)

\[
\frac{\mathcal{R}^\pi(T,\mu)}{T^a} \to 0,
\]

as \( T \to \infty \), for every \( a > 0 \). In other words, the per-consumer normalized revenue of consistent policies converges to 1 for all possible mean utility vectors. Equation (2.4) restricts the rate of such convergence in \( T \). Let \( \mathcal{P}' \subseteq \mathcal{P} \) denote the set of non-anticipating, consistent assortment policies.

Suppose the retailer is given a-priori knowledge of the mean utility values only for the products in the optimal assortment (without revealing the optimality condition of such assortment). We call a product potentially optimal if it cannot be discarded as suboptimal solely on the base of such a priori information. That is,

\[
\mathcal{N} := \{ i \in \mathcal{N} : i \in S^*(\nu), \nu := (\mu_1, \ldots, \mu_{i-1}, v, \mu_{i+1}, \ldots, \mu_N) \text{ for some } v \in \mathbb{R} \}.
\]

(Note that the definition above does not considers changes in \( w \), the vector of profit margins.) Similarly, we let \( \overline{\mathcal{N}} := \mathcal{N} \setminus \mathcal{N} \) denote the set strictly suboptimal products.

We will assume the common distribution function \( F \) is absolutely continuous with respect to Lebesgue measure on \( \mathbb{R} \), and that its density function is positive everywhere. This assumption is quite standard and satisfied by many commonly used distributions. The result below establishes a fundamental limit on what can be achieved by any consistent assortment policy.

**Theorem 1.** For any \( \pi \in \mathcal{P}' \), and any \( \mu \in \mathbb{R}^N \),

\[
\mathcal{R}^\pi(T,\mu) \geq K \left| \mathcal{N} \setminus S^*(\mu) \right| \frac{\log T}{C},
\]

for a finite positive constant \( K \) and all \( T \).
A explicit expression for the constant $K$ is given in the proof. When all non-optimal products are strictly-suboptimal, the result suggests that a finite regret may be attainable. This last observation highlights the importance of strictly-suboptimal product detection, and hence the inefficiency of a naive multi-armed bandit approach to assortment planning: treating each possible assortment as a different arm in the bandit setting will result in the regret scaling linearly with the combinatorial term $\binom{N}{C}$, instead of the much smaller constant $(|\mathcal{N} \setminus S^*(\mu)|/C)$.

**Remark 3 (Implications for design of “good” policies).** The proof of Theorem 1, which is outlined below, suggests certain desirable properties for “optimal” policies: (i.) non-optimal products that can be made to be part of the optimal assortment are to be tested on order $\log T$ customers; (ii) this type of non-optimal product experimentation is to be conducted in batches of size $C$; and (iii.) strictly-suboptimal products (the ones that cannot be made to be part of the optimal assortment) need only be tested on a finite number of customers (in expectation), independent of $T$.

### 2.4.2 Proof outline and intuition behind Theorem 1

Proof of Theorem 1 exploits the connection between the regret and testing of suboptimal assortments. In particular, we will bound the regret by computing lower bounds on the expected number of tests involving potentially optimal suboptimal products (those in $\mathcal{N} \setminus S^*(\mu)$): each time such a product is offered, the corresponding assortment must be suboptimal, contributing directly to the policy’s regret.

To bound the number of tests involving non-optimal products we will use a change-of-measure argument introduced by Lai and Robbins (1985) for proving an analogous result for a multi-armed bandit problem, hence our proof establishes a direct connection between the two areas. To adapt this idea we consider the fact that underlying realizations of the random variables (product utilities) are non-observable in the assortment setting, which differs from the multi-armed bandit setting where reward realizations are observed directly. The argument can be roughly described as follows. Any non-optimal product $i \in \mathcal{N}$ is in the optimal assortment for at least one suitably chosen mean utility vector $\hat{\mu}$. When such a configuration is considered, any consistent policy $\pi$ must offer this non-optimal product to
all but a sub-polynomial (in $T$) number of customers. If this configuration does not differ in a significant manner from the original (made precise in Appendix A.1), then one would expect such a product to be offered to a “large” number of customers when the original mean utility vector $\mu$ is considered. In particular, we prove that for any policy $\pi$

$$\mathbb{P}_\pi \{ T_i(T) \leq \log T/K_i \} \to 0$$

(2.5)

as $T \to \infty$, where $T_i(t)$ is the number of customers product $i$ has been offered to up until customer $t - 1$, and $K_i$ is a finite positive constant. Note that this asymptotic minimum-testing requirement is inversely proportional to $K_i$, which turns out to be a measure of “closeness” of a product to “optimality” (how close the vector $\mu$ is to a configuration that makes product $i$ be part of the optimal assortment). This also has immediate consequences on the expected number of times a non-optimal product is tested: using Markov’s inequality we have that for any $i \in \mathcal{N} \setminus S^*(\mu)$,

$$\liminf_{T \to \infty} \frac{\mathbb{E}_\pi \{ T_i(T) \}}{\log T} \geq \frac{1}{K_i}.$$  

The result in Theorem 1 follows directly from the equation above and the connection between the regret and testing of suboptimal assortments mentioned at the beginning of this section.

### 2.5 Dynamic Assortment Planning Policies

This section introduces an assortment policy whose structure is guided by the key ideas gleaned from Theorem 1. Our policy is based on the idea that product experimentation on a given assortment should provide information on the performance of the same product on a different assortment. That is, one should be able to extract product-specific information from consumer decisions on any assortment including such a product. More formally, our policy is based on the assumption that one can recover the model parameters for products on a given assortment by observing purchase probabilities associated to such an assortment. With this in mind, we introduce the following assumption, which we assume holds true throughout the rest of this Chapter.
Assumption 1 (Identifiability). For any assortment $S \in \mathcal{S}$, and any vector $\rho \in \mathbb{R}_+^N$ such that $\sum_{i \in S} p_i < 1$, the system of equations $\{p_i(S, \eta) = \rho_i, \ i \in S\}$ has a unique solution $T(S, \rho)$ in $\eta \in \mathbb{R}^N$ such that $\eta_i = 0$ for all $i \notin S$. In addition $p(S, \cdot)$ is Lipschitz continuous, and $T(S, \cdot)$ is locally Lipschitz continuous in the neighborhood of $\rho$, for all $S \in \mathcal{S}$.

Under this assumption one can compute mean utilities for products in a given assortment based solely on the associated purchase probabilities. This characteristic enables our approach to parameter estimation: we will estimate purchase probabilities by observing consumer decisions during an exploration phase and we will use those probabilities to reconstruct a mean utility vector that rationalizes such observed behavior. Note that the Logit model, for which $F$ is a standard Gumbel, satisfies this assumption.

2.5.1 Intuition and a simple “separation-based” policy

To build some intuition towards the construction of our ultimate dynamic assortment policy (given in §2.5.2) it is helpful to first consider a policy that separates exploration from exploitation. The idea is to isolate the effect of imposing the right order of exploration (suggested by Theorem 2) on the regret. Assuming prior knowledge of $T$, such a policy first engages in an exploration phase over $\lceil N/C \rceil$ assortments encompassing all products, each offered sequentially to order $\log T$ customers. Then, in light of Assumption 1 an estimator for $\mu$ is computed. Based on this estimator a proxy for the optimal assortment is computed, and offered to the remaining customers. For this purpose consider the set of test-assortments $\mathcal{A} := \{A_1, \ldots, A_{\lceil N/C \rceil}\}$, where

$$A_j = \{(j - 1)C + 1, \ldots, \min\{jC, N\}\},$$

Fix $j \leq |\mathcal{A}|$. Suppose $t - 1$ customers have arrived to that point. We will use $\hat{p}_{i,t}(A_j)$ to estimate $p_i(A_j, \mu)$ when customer $t$ arrives, where

$$\hat{p}_{i,t}(A_j) := \frac{\sum_{u=1}^{t-1} Z_u^i 1\{S_u = A_j\}}{\sum_{u=1}^{t-1} 1\{S_u = A_j\}},$$

for $i \in A_j$ and $\hat{p}_{i,t}(A_j) = 0$ otherwise. Define $\hat{p}_t(A_j) := (\hat{p}_{1,t}(A_j), \ldots, \hat{p}_{N,t}(A_j))$ to be the vector of product selection probabilities. For any $i \in N$ we will use $\hat{p}_{i,t}$ to estimate $\mu_i$ when

2 In the numerical experiments we use the modified estimator $\hat{p}_{i,t}(A_j) := \frac{1 + \sum_{u=1}^{t-1} Z_u^i 1\{S_u = A_j\}}{|A_j| + \sum_{u=1}^{t-1} 1\{S_u \neq A_j\}}$ for $i \in A_j$. Such probability estimates help in overcoming the short term bias associated to MLE estimators.
customer \( t \) arrives, where 
\[
\hat{\mu}_{i,t} := (T(A_j, \hat{p}_t(A_j)))_i,
\]
\( A_j \) is the unique test assortment in \( \mathcal{A} \) such that \( i \in A_j \), \( (a)_i \) denotes the \( i \)-th component of vector \( a \). Let \( \hat{\mu}_t := (\hat{\mu}_{1,t}, \ldots, \hat{\mu}_{N,t}) \) denote the vector of mean utilities estimates.

The underlying idea is the following: when an assortment \( A_j \in \mathcal{A} \) has been offered to a large number of customers one expects \( \hat{p}_{t,i}(A_j) \) to be close to \( p_i(A_j, \mu) \) for all \( i \in A_j \). If this is the case for all assortments in \( \mathcal{A} \), by Assumption 1 we also expect \( \hat{\mu}_t \) to be close to \( \mu \). With this in mind, we propose a separation-based policy defined through a positive constant \( \kappa_1 \) that serves as a tuning parameter. The policy is summarized for convenience in Algorithm 1.

\begin{algorithm}
\caption{Algorithm 1: \( \pi_1 = \pi(\kappa_1, T, w) \)}
\begin{enumerate}
\item \textbf{STEP 1.} Exploration:
\begin{enumerate}
\item \textbf{for} \( j =1 \) to \(|\mathcal{A}|\) \textbf{do}
\begin{enumerate}
\item Offer \( A_j \) to \( \lceil \kappa_1 \log T \rceil \) customers (if possible). \hfill \text{[Exploration]}
\end{enumerate}
\end{enumerate}
\end{enumerate}
\item \textbf{STEP 2.} Exploitation:
\begin{enumerate}
\item \textbf{for} \( t =\lceil \kappa_1 (\log T) |\mathcal{A}| \rceil +1 \) to \( T \) \textbf{do}
\begin{enumerate}
\item \textbf{for} \( j =1 \) to \(|\mathcal{A}|\) \textbf{do}
\begin{enumerate}
\item Set \( \hat{p}_{i,t}(A_j) := \begin{cases}
\sum_{u=1}^{t-1} z_u \mathbf{1}\{\mathbf{S}_u=A_j\} / \sum_{u=1}^{t-1} \mathbf{1}\{\mathbf{S}_u=A_j\} & i \in A_j, \\
0 & i \notin A_j.
\end{cases} \hfill \text{[Probability estimates]}
\end{enumerate}
\end{enumerate}
\end{enumerate}
\item Set \( \hat{p}_t := \{\hat{p}_{1,t}(A_j), \ldots, \hat{p}_{N,t}(A_j)\} \).
\item Set \( \hat{\mu}_{i,t} := \eta_i \) for \( i \in A_j \), where \( \eta := T(A_j, \hat{p}_t(A_j)) \). \hfill \text{[Mean utility estimates]}
\end{enumerate}
\end{algorithm}

\textbf{Performance analysis.} This policy is constructed to guarantee that the probability of not choosing the optimal assortment after the initial exploration effort, (i.e., the magnitude of revenue loss during exploitation) balances revenue loss coming from exploration efforts,
which is order \( \log T \). This, in turn, translates into an order \( ([N/C] \log T) \) regret. The next result formalizes this.

**Theorem 2.** Let \( \pi_1 := \pi(\kappa_1, T, w) \) be defined by Algorithm 1 and let Assumption 1 hold. There exist finite constants \( \overline{K}_1 \) and \( \pi_1 \) (independent of \( T \)), such that the regret associated with \( \pi_1 \) is
\[
R^\pi(T, \mu) \leq \kappa_1 [N/C] \log T + \overline{K}_1, 
\]
for all \( T \), provided that \( \kappa_1 > \overline{\kappa}_1 \).

Constants \( \overline{K}_1 \) and \( \overline{\kappa}_1 \) depend on instance specific quantities (e.g., the minimum optimality gap), but not on the number of products, \( N \), nor on the length of the selling horizon, \( T \). The proof of Theorem 2 elucidates that \( \overline{K}_1 \) is the expected cumulative loss during the exploitation phase for an infinite horizon setting, while \( \overline{\kappa}_1 \) represents the minimum length of the exploration phase that makes \( \overline{K}_1 \) finite. This tradeoff is balanced by the construction of the policy \( \pi_1 \). The bound presented in Theorem 2 is essentially the one in Theorem 1 with \( N \) replacing \( [N \setminus S^*(\mu)] \). This indicates that: (i.) imposing the right order (in \( T \)) of exploration is enough to get the right dependence (in \( T \)) of the regret; and (ii.) to reach the fundamental limit one needs to limit exploration on strictly-suboptimal products.

**Remark 4 (Selection of the tuning parameter \( \kappa_1 \)).** We have established that the lower bound in Theorem 1 can be achieved in terms of its dependence on \( T \), for proper choice of \( \kappa_1 \). However, Theorem 2 requires \( \kappa_1 \) to be greater than \( \overline{\kappa}_1 \), whose value is not known a priori. In particular, setting \( \kappa_1 \) below the specified threshold may compromise the validity of the result. To avoid the risk of miss-specifying \( \kappa_1 \), one can increase the order of the minimum amount of exploration to, say, \( \kappa_1 \log^{1+\alpha} t \), for any \( \alpha > 0 \). With this, the upper bound above would read
\[
R^\pi(T, \mu) \leq \kappa_1 [N/C] \log^{1+\alpha} T + \overline{K}_1, 
\]
and the policy becomes optimal up to a \( \log^{\alpha} T \)-term.

**Example 1.a: performance of the Separation-based policy \( \pi_1 \) for MNL choice model.** Consider \( N = 10 \) and \( C = 4 \), with
$$w = (0.98, 0.88, 0.82, 0.77, 0.71, 0.60, 0.57, 0.16, 0.04, 0.02),$$

$$\mu = (0.36, 0.84, 0.62, 0.64, 0.80, 0.31, 0.84, 0.78, 0.38, 0.34),$$

and assume \( \{\zeta_t^i\} \) have a standard Gumbel distribution, for all \( i \in \mathcal{N} \) and all \( t \geq 1 \), i.e., we consider the MNL choice model. One can verify that \( S^*(\mu) = \{1, 2, 3, 4\} \) and \( r(S^*(\mu), \mu) = 0.76 \). Also, \( \mathcal{N} = \{5, 6, 7, 8, 9, 10\} \). One can use the test assortments \( A_1 = \{1, 2, 3, 4\} \), \( A_2 = \{5, 6, 7, 8\} \) and \( A_3 = \{9, 10\} \) to conduct the exploration phase in the algorithm described above.

**Example 1.b: performance of the Separation-based policy \( \pi_1 \) for multinomial Probit choice model.** Consider \( N = 6 \) and \( C = 2 \), with

$$w = (2.0, 1.8, 1.5, 1.4, 1.2, 1.0),$$

$$\mu = (0.2, 0.3, 0.35, 0.45, 0.50, 0.55),$$

and assume \( \{\zeta_t^i\} \) have a standard normal distribution, for all \( i \in \mathcal{N} \) and all \( t \geq 1 \), i.e., we consider the multinomial Probit choice model. One can verify that \( S^*(\mu) = \{1, 2\} \) and \( r(S^*(\mu), \mu) = 1.39 \). Also, \( \mathcal{N} = \{5, 6\} \). One can use the test assortments \( A_1 = \{1, 2\} \), \( A_2 = \{3, 4\} \) and \( A_3 = \{5, 6\} \) to conduct the exploration phase in the algorithm described above.

Panels (a) and (b) in Figure 2.1 depict the average performance of policy \( \pi_1 \) for instances in examples 1.a and 1.b. Simulation results were conducted over 500 replications, using \( \kappa_1 = 20 \), and considering selling horizons ranging from \( T = 500 \) to \( T = 10000 \). There, graphs (a) and (b) illustrates the dependence of the regret on \( T \) for instances in examples 1.a and 1.b, respectively. The dotted lines represent 95% confidence intervals for the simulation results. Two important points are worth noting: the regret is indeed of order \( \log T \), as predicted by Theorem 2; policy \( \pi_1 \) makes suboptimal decisions on a very small fraction of customers, e.g. in panel (a) it ranges from around 10% when the horizon is 2000 sales attempts, and diminishes to around 2.5% for a horizon of 10,000. (Recall that the regret is measuring the number of suboptimal sales.)

From the setting of these examples we observe that \( A_2 \) and \( A_3 \) for example 1.a, and \( A_3 \) for example 1.b, are tested on order \( \log T \) customers, despite being composed exclusively of strictly-suboptimal products. That is, the separation algorithm does not attempt to limit
testing efforts over suboptimal products. Moreover, it assumes a priori knowledge of the total number of customers, $T$. The next section proposes a policy that addresses these two issues.

![Performance of the separation-based policy $\pi_1$.](image)

Figure 2.1: Performance of the separation-based policy $\pi_1$.

### 2.5.2 A refined dynamic assortment policy

To account for strictly-suboptimal product detection it is necessary to be able to “identify” them a priori, even under partial knowledge of the mean utility vector. For that purpose we introduce the following assumption

**Assumption 2 (Optimal assortment composition).** For any vectors $\mu \in \mathbb{R}^N$ one has that

$$r(S^*(\mu), \mu) \leq w_i, \quad \text{for all } i \in S^*(\mu).$$

Assumption 2 states that no product in the optimal assortment has profit margin lower than the optimal single sale profit. The intuition behind it is that, within an assortment $S$, if a product provides a margin lower than the assortment single sale profit, then one is better off by removing a such product from the assortment. When Assumption 2 holds, the optimal single sale profit acts as a threshold that separates strictly-suboptimal products from the rest: any product with margin less than the optimal single sale profit is strictly-suboptimal.
and vice versa. That is, 
\[ N = \{ i \in \mathcal{N} : w_i < r(S^*(\mu), \mu) \} . \]

Designing test assortments based on product margins translates this into a threshold over assortments. Consider the set of valid assortments \( \mathcal{A} := \{ A_1, \ldots, A_{\lceil N/C \rceil} \} \), where
\[ A_j = \{ i_{((j-1)C+1)}, \ldots, i_{\min(jC, N)} \} , \]
and \( i_{(k)} \) corresponds to the product with the \( k \)-th highest margin in \( w \). Suppose one has a proxy for \( r(S^*(\mu), \mu) \). One can then use this value to identify assortments containing at least one potentially optimal product and to force the right order of exploration on such assortments. If successful, such a scheme will limit exploration on assortment containing only strictly-suboptimal products.

**Remark 5 (Validity of Assumption 2).** While Assumption 2 does hold for Luce-type choice models (the MNL being a special case), it does not hold for general choice models and its validity have to be study on each case. For example, one can show that if \( w_i \geq 2w_{i+1} \) for all \( i \in \mathcal{N} \), Assumption 2 holds true for any noise distribution with increasing (decreasing) failure rate, provided that \( \mu \in \mathbb{R}^N_- (\mu \in \mathbb{R}^N_+). \)

We propose an assortment policy that, for each customer executes the following logic: using the current estimate of \( \mu \) at time \( t \), the static problem is solved and \( S_t \), the estimate-based optimal assortment, and \( r_t \), the estimate of the optimal value, are obtained. If all assortments in \( \mathcal{A} \) containing products with margins greater than or equal to \( r_t \) have been tested on a minimum number of customers, then assortment \( S_t \) is offered to the \( t \)-th customer. Otherwise, we select, arbitrarily, an under-tested assortment in \( \mathcal{A} \) containing at least one product with margin greater than or equal to \( r_t \), and offer it to the current customer. The term under-tested means tested on less than order \( \log t \) customers prior to the arrival of customer \( t \). Note that this logic will enforce the correct order of exploration for any value of \( T \).

This policy, denoted \( \pi_2 \) and summarized for convenience in Algorithm 2, monitors the quality of the estimates for potentially optimal products by imposing minimum exploration on assortments containing such products. The specific structure of \( \mathcal{A} \) ensures that test
Algorithm 2: \( \pi_2 = \pi(\kappa_2, w) \)

**STEP 1. Initialization:**

for \( t = 1 \) to \(|A|\) do

Offer \( A_t \in A \) to customer \( t \) and set \( n_t = 1 \). \([\text{Initial test}]\)

end for

**STEP 2. Joint exploration and assortment optimization:**

for \( t = |A| + 1 \) to \( T \) do

for \( j = 1 \) to \(|A|\) do

Set \( \hat{p}_{i,t}(A_j) := \frac{\sum_{u=1}^{t-1} Z_{u=1}^* 1\{S_u=A_j\}}{\sum_{u=1}^{t-1} 1\{S_u=A_j\}} \) for \( i \in A_j \). \([\text{Probability estimates}]\)

Set \( \hat{\mu}_{i,t} := \eta_i \) for \( i \in A_j \), where \( \eta = \mathcal{T}(A_j, \hat{p}_t(A_j)) \). \([\text{Mean utility estimates}]\)

end for

Set \( S_t = S^*(\hat{\mu}_t) \) and \( r_t = r(S^*(\hat{\mu}_t), \hat{\mu}_t) \). \([\text{Static optimization}]\)

Set \( \widehat{N}_t = \{i \in N : w_i \geq r_t\} \). \([\text{Candidate optimal products}]\)

if \( (n_j \geq \kappa_2 \log t) \) for all \( j = 1 \) to \(|A|\) such that \( A_j \cap \widehat{N}_t \neq \emptyset \) then

Offer \( S_t \) to customer \( t \). \([\text{Exploitation}]\)

else

Select \( j \) such that \( A_j \cap \widehat{N}_t \neq \emptyset \) and \( n_j < \kappa_2 \log t \).

Offer \( A_j \) to customer \( t \). \([\text{Exploration}]\)

\( n_j \leftarrow n_j + 1 \)

end if

end for

Assortments do not “mix” high-margin products with low-margin products, thus successfully limiting exploration on strictly-suboptimal products. The policy uses a tuning parameter \( \kappa_2 \) to balance non-optimal assortment testing (which contributes directly to the regret), and the probability of choosing the optimal assortment in the exploitation phase.

**Performance analysis.** The next result characterizes the performance of the proposed assortment policy.

**Theorem 3.** Let \( \pi_2 = \pi(\kappa_2, w) \) be defined by Algorithm 2 and let Assumptions 1 and 2 hold. There exist finite constants \( \overline{K}_2 \) and \( \overline{\kappa}_2 \) (independent of \( T \)), such that the regret associated
with $\pi_2$ is accepts the bound

$$R^\pi(T, \mu) \leq \kappa_2 \left( \frac{|N|}{C} \right) \log T + \kappa_2,$$

for all $T$, provided that $\kappa_2 > \kappa_2$.

Theorem 3 implicitly states that assortments containing only strictly-suboptimal products will be tested on a finite number of customers (in expectation); see Corollary 1 below. Note that this policy attains the correct dependence on both $T$ and $N$, as prescribed in Theorem 1 (up to the size of the optimal assortment), so it is essentially optimal. Unlike Theorem 2 we see that the proposed policy successfully limits exploration on strictly-suboptimal products. The following corollary formalizes this statement. Recall $T_i(t)$ denotes the number of customers product $i$ has been offered to up to arrival of customer $t$.

**Corollary 1.** Let Assumptions 1 and 2 hold. Then, for any assortment $A_j \in A$ such that $A_j \subseteq N$, and for any selling horizon $T$

$$\mathbb{E}_\pi[T_i(T)] \leq K_2,$$

for all $i \in A_j$, where $K_2$ is a finite positive constant independent of $T$.

Regarding selection of parameter $\kappa_2$, note that the argument in Remark 4 remains valid.

**Example 1.a continued: performance of the policy $\pi_2$ for MNL choice mode.** Consider the setting of Example 1.a in Section 2.5.1, for which $N = A_2 \cup A_3$ and $S^*(\mu) = A_1$. Given that the set of test assortments separates products in $N$ from the rest, one would expect Algorithm 2 to effectively limit exploration on all strictly-suboptimal products.

**Example 1.b continued: performance of the policy $\pi_2$ for multinomial Probit choice mode.** Consider the setting of Example 1.b in Section 2.5.1, for which $N = A_3$ and $S^*(\mu) = A_1$. Given that the set of test assortments separates products in $N$ from the rest, one would expect Algorithm 2 to effectively limit exploration on all strictly-suboptimal products.
Panels (a) and (b) in Figure 2.2 depict the average performance of policies $\pi_1$ and $\pi_2$ for instances in examples 1.a and 1.b, respectively. Simulation results were conducted over 500 replications, using $\kappa_1 = \kappa_2 = 20$, and considering selling horizons ranging from $T = 500$ to $T = 10000$. There, graphs (a) and (b) compares the separation-based policy $\pi_1$, given by Algorithm 1, and the proposed policy $\pi_2$, in terms of regret dependence on $T$ for instances in examples 1.a and 1.b, respectively. The dotted lines represent 95\% confidence intervals for the simulation result. The main point to note is that policy $\pi_2$ outperforms substantially the separation-based policy $\pi_1$. In particular, for the instance in Example 1.a, the operation of $\pi_1$ results in lost sales in the range of 2.5-10\% (200-260 customers are offered non-optimal choices), depending on the length of selling horizon, while for $\pi_2$ we observe sub-optimal decisions being made only about 10-20 times (!) independent of the horizon. This constitutes more than a 10-fold improvement over the performance of $\pi_1$. In essence, $\pi_2$ adaptively identifies both $A_2$ and $A_3$ as strictly-suboptimal assortment with increasing probability as $t$ grows large. As a result, $\pi_2$ exploration efforts are eventually directed exclusively to the optimal assortment. Since incorrect choices in the exploitation phase are also controlled by $\pi_2$, we expect the regret to be finite. This is supported by the numerical results displayed in Figure 2.2.
Remark 6 (Relationship to bandit problems continued). The result in Corollary [1] stands in contrast to typical multi-armed bandit results, where all suboptimal arms/actions need to be tried at least order log \( t \) times (in expectation). In the assortment problem, product rewards are random variables bounded above by their corresponding margins, therefore, under Assumption [2] the contribution of a product to the overall profit is bounded, independent of its mean utility. More importantly, this feature makes some products a priori better than others. Such characteristic is not present in the typical bandit problem.

Remark 7 (Performance of \( \pi_2 \) in absence of Assumption [2]). In absence of Assumption [2], it seems impossible to identify strictly-suboptimal products a priori. Instead, one can modify Algorithm 2 to simply ignore strictly-suboptimal product detection. It can be then seen from the proof of Theorem [3] that the upper bound remains valid, with \( N \) replacing \(|\mathcal{N}|\).

2.5.3 A policy customized to the multinomial Logit choice model

For general utility distributions, choice probabilities depend on the offered assortment in a non-trivial way, and hence it is unclear how to combine information originating from different assortments and allow for more efficient use of data gathered on the exploitation phase. We illustrate how to modify parameter estimation to include exploitation-based product information in the case of an MNL choice model (we note that all results in this section extend directly to Luce-type choice models). As indicated earlier, for this model an efficient algorithm for solving the static optimization problem has been developed by Rusmevichientong et al. (2010).

MNL choice model properties. Taking \( F \) to have a standard Gumbel distribution, then (see, for example, Anderson et al. (1992))

\[
p_i(S, \nu) = \frac{\nu_i}{1 + \sum_{j \in S} \nu_j} \quad i \in S, \text{ for any } S \in \mathcal{S},
\]

where \( \nu_i := \exp(\mu_i), i \in \mathcal{N}, \) and \( \nu := (v_1, \ldots, v_N) \). In what follows, we will use both \( \nu \) and \( \mu \) interchangeably. Given an assortment \( S \in \mathcal{S} \) and a vector \( \rho \in \mathbb{R}^N_+ \) such that \( \sum_{i \in S} \rho_i \leq 1 \), we have that \( T(S, \rho) \), the unique solution to \( \{\rho_i = p_i(S, \nu) \text{ for } i \in S, \nu_i = 0 \text{ for } i \in \mathcal{N} \setminus S\} \)
is given by
\[ T_i(S, \rho) = \frac{\rho_i}{1 - \sum_{j=1}^{N} \rho_j} \quad \text{if} \quad i \in S. \] (2.7)

From (2.6) one can see that solving the static optimization problem is equivalent to finding the largest value of \( \lambda \) such that
\[ \sum_{i \in S} v_i (w_i - \lambda) \geq \lambda, \] (2.8)
for some \( S \in S \). One can check that (2.7) and (2.8) implies that Assumptions 1 and 2 holds, respectively.

A product-exploration-based assortment policy. We propose a customized version of the policy given by Algorithm 2, which we refer to as \( \pi_3 \), defined through a positive constant \( \kappa_3 \) that serves as a tuning parameter. The policy, which is summarized below in algorithmic form, maintains the general structure of Algorithm 2, however parameter estimation, product testing and suboptimal product detection are conducted at the product-level. In what follows, the following estimators are used. Suppose \( t-1 \) customers have shown up so far. We will use \( \hat{\nu}_{i,t} \) to estimate \( \nu_i \) when customer \( t \) arrives, where
\[ \hat{\nu}_{i,t} := \frac{\sum_{u=1}^{t-1} Z_i^u \mathbf{1}\{i \in S_u\}}{\sum_{u=1}^{t-1} Z_i^u \mathbf{1}\{i \in S_u\}} \quad i \in N. \]
The estimate above exploits the independence of irrelevant alternatives (IIA) property of the Logit model, which says that the ratio between purchase probabilities of any two products is independent of the assortment in which they are offered. This is
\[ \frac{p_i(S, \nu)}{p_j(S, \nu)} = \frac{\nu_i}{\nu_j}, \quad \text{for all products} \quad i, j \in N \cup \{0\}, \quad \text{for all} \quad S \in S. \]
As a result, all information collected (both from exploration and exploitation phases) is used to construct the parameter estimates.

Performance analysis. The tuning parameter \( \kappa_3 \) plays the same role as \( \kappa_2 \) plays in Algorithm 2. The next result characterizes the performance of the proposed assortment policy.

**Theorem 4.** Let \( \pi_3 = \pi(\kappa_3, \nu) \) be defined by Algorithm 3. There exist finite constants \( K_3 \) and \( \kappa_3 \) (independent of \( T \)), such that the regret associated with \( \pi_3 \) is accepts the bound
\[ \mathcal{R}^\pi(T, \nu) \leq \kappa_3 |N \setminus S^*(\nu)| \log T + K_3, \]
Algorithm 3: $\pi_3 = \pi(\kappa_3, w)$

**STEP 1. Initialization:**

Set $n_j = 0$ for all $j \in \mathcal{N}$.

Offer $S_1 = \text{argmax} \{ w_j : j \in \mathcal{N} \}$ to customer $t = 1$. Set $n_i = 1$.

**STEP 2. Joint exploration and assortment optimization:**

for $t = 2$ to $T$
do

for $i = 1$ to $N$
do

Set $\hat{\nu}_{i,t} := \left(\sum_{u=1}^{t-1} Z_i^u 1 \{ i \in S_u \} / \left(\sum_{u=1}^{t-1} Z_i^u 1 \{ i \in S_u \} \right)\right)$. [Estimation]

end for

Set $S_t = S^*(\hat{\nu}_t)$ and $r_t = r(S^*(\hat{\nu}_t), \hat{\nu}_t)$. [Static optimization]

Set $O_t = \{ i \in \mathcal{N} : w_i \geq r_t, n_i < \kappa_3 \log t \}$. [Candidate optimal products]

if $O_t = \emptyset$ then

Offer $S_t$ to customer $t$. [Exploitation]

else

Offer $S_t \in \{ S \in \mathcal{S} : S \subseteq O_t \}$. [Exploration]

end if

$n_i \leftarrow n_i + 1$ for all $i \in S_t$.

end for

for all $T$, provided that $\kappa_3 > \pi_3$.

Theorem 4 is essentially the equivalent of Theorem 3 for the Logit case, with the exception of the dependence on the assortment capacity $C$ (as here exploration is conducted on a product basis), and on the dependence on the set $\mathcal{N}$. The latter matches exactly the order of the result in Theorem 1 unlike policy $\pi_2$, the customized policy $\pi_3$ prevents optimal products from being offered in suboptimal assortments. Since estimation is conducting using information arising from both exploration and exploitation phases, one would expect a better empirical performance from the Logit customized policy. Note that the result implicitly states that strictly-suboptimal products will be tested on a finite number of customer, in expectation. The following corollary is the MNL-customized version of Corollary 1.
Corollary 2. For any strictly-suboptimal product \( i \in \mathcal{N} \) and for any selling horizon \( T \)

\[ \mathbb{E}_\pi[T_i(T)] \leq K_3, \]

for a positive finite constant \( K_3 \), independent of \( T \).

Regarding selection of parameter \( \kappa_2 \), note that the argument in Remark 4 remains valid.

Example 1.a continued: performance of the MNL-customized policy \( \pi_3 \). Consider the setup of Example 1.a in Section 2.5.1. Note that \( S^*(\nu) = A_2 \), i.e., the optimal assortment matches one of the test assortments. Moreover, one has that that \( \mathcal{N} = S^*(\nu) \). As a result, strictly suboptimal detection is conducted in finite time for both policies \( \pi_2 \) and \( \pi_3 \), and hence any gain in performance for policy \( \pi_3 \) over \( \pi_2 \) is tied in to the ability of the former to incorporate information gathered during both exploitation and exploration phases.

Figure 2.3 depicts the average performance of policies \( \pi_2 \) and \( \pi_3 \) over 500 replications, using \( \kappa_2 = \kappa_3 = 20 \), and considering selling horizons ranging from \( T = 1000 \) to \( T = 10000 \). There, the graph compares the more general policy \( \pi_2 \) to its Logit-customized version \( \pi_3 \), in terms of regret dependence on \( T \). The dotted lines represent 95% confidence intervals for the simulation result. Customization to a Logit nets significant, roughly 10-fold, improvement in performance of \( \pi_3 \) relative to \( \pi_2 \). Overall, the Logit-customized policy \( \pi_3 \) only offers suboptimal assortments to less than a handful of customers, regardless of the horizon of the problem. This provides picture proof that the regret (number of suboptimal sales) is finite. In particular, since \( \mathcal{N} = S^*(\nu) \) Theorem 4 predicts a finite regret for any \( T \). This suggests that difference in performance is mainly due to errors made in the exploitation phase. This elucidates the reason why the Logit customized policy \( \pi_3 \) outperforms the more general policy \( \pi_2 \): the probability of error decays much faster in the Logit customized version. If all previous exploitation efforts were successful, and assuming correct strictly-suboptimal product detection, the probability of error decays exponentially for the customized policy (\( \pi_3 \)) and polynomially for the more general policy (\( \pi_2 \)); see further details in Appendix A.1.

2.6 Concluding Remarks

Complexity of the dynamic assortment problem. Theorem 1 provides a lower bound on the regret of an optimal policy for the dynamic assortment problem. We have shown
that this lower bound can be achieved, up to constant terms, when the noise distribution on the utility of each customer is known. In particular, we proposed an assortment-exploration based algorithm whose regret scales optimally in the selling horizon $T$ and exhibits the right dependence on the number of possible optimal products $|\mathcal{N}|$. (In addition our proposed policies do not require a priori knowledge of the length the selling horizon.)

**Comparison of our policy with benchmark results.** Our results significantly improve on and generalize the policy proposed by Rusmevichientong *et al.* (2010) where an order $(N^2 (\log T)^2)$ upper bound is presented for the case of an MNL choice model. Recall the regret of our policy exhibits order $|\mathcal{N} \setminus S^*(\mu)| \log T$ performance, and we show that this cannot be improved upon. We note that the policy of Rusmevichientong *et al.* (2010) is a more direct adaptation of multi armed bandit ideas an hence does not detect strict-suboptimal products and does not limit exploration on them. We illustrate this with a simple numerical example.

Consider again Example 1.a in Section 2.5.1. Figure 2.4 compares the average performance of our proposed policies with that of Rusmevichientong *et al.* (2010) denoted RSS for short, over 500 replications, using $\kappa_1 = \kappa_2 = \kappa_3 = 20$, and considering selling horizons ranging from $T = 1000$ to $T = 10000$. There, the graph in (a) compares the separation-
based policy $\pi_1$ to the benchmark policy RSS, in terms of regret dependence on $T$. The graph in (b) compares the separation-based policy $\pi_1$, the proposed policy $\pi_2$ and its Logit-customized version $\pi_3$ in terms of regret dependence on $T$. A further analysis behind the results depicted in Figure 2.4 indicates that the performance of the benchmark behaves quadratically with $\log T$, while the performance of our proposed policies grow linearly.

Several factors explain the difference in performance. First, we consider a set of roughly $N$ test assortments while in RSS this set contains roughly $N^2$ items. As a consequence, order $N^2$ are tested, resulting in an overall performance proportional to $N^2$. This explains why even the naive separation-based policy $\pi_1$ outperforms RSS. Panel (a) in Figure 2.4 shows that the RSS policy loses sales on about 20–25% of the customers, while policy $\pi_1$ never loses more than 10%, the loss diminishes as the horizon increases to around 2.5%. Since policies $\pi_2$ and $\pi_3$ limit exploration on strictly-suboptimal products, a feature absent in both RSS and in the naive separation-based policy $\pi_1$, they exhibit far superior performance compared to either one of those benchmarks as illustrated in panel (b) of Figure 2.4. Finally, our MNL-customized policy $\pi_3$ uses all information gathered for computing parameter estimates, while the policy in RSS only uses the information collected during the exploration phase. The improvement in performance due to this feature is also illustrated in panel (b) of Figure 2.4. The overall effect is that policy $\pi_3$ improves performance by a factor of 200-1000 compared to RSS, and is able to zero in on the optimal assortment much faster than the benchmark, with a regret that is bounded independent of the horizon $T$.

---

$^3$The dynamic assortment algorithm presented in Rusmevichientong et al. (2010) is based on their clever solution to the static assortment problem. There, they propose an algorithm (for the static problem) that identifies rather efficiently (in order $N^2$ operations) a small set of assortments that contains the optimal one. For that purpose they are required to compute order $N^2$ intersection points for the set of functions $\{g_i(\cdot) := w_i(v_i \cdot) : i \in N\}$. Their dynamic implementation of the algorithm requires to estimate each intersection points by testing order $N^2$ assortments (so that each pair of products tested together on at least one assortment).
Figure 2.4: Comparison with benchmark performance.
Chapter 3

Dynamic Learning and Customization in Display-based Online Advertisement

3.1 Introduction

Motivation. Internet advertising revenues in the United States reached $22.7 billion in 2009. Display-related advertisement – where an advertiser pays an online publisher (web page) to display a static banner or logo to upcoming visitors – accounted for 35% of such revenues, up from 33% in 2008 [IAB, 2009]. This emerging area of business is an important source of revenues for thousands of online publishers.

The most prevalent business model in online advertisement builds on the cost-per-click (CPC) statistic: advertisers pay the online publisher only when a user clicks on their advertisements (henceforth, ads). The CPC model, which accounted for 59% of industry revenues in 2009, provides real-time feedback on the performance of online advertisement campaigns and improves focalization of the marketing efforts (relative to traditional media channels). Display advertisement contracts are usually the result of a direct negotiation process between the publisher and the advertiser. As a result, the publisher is entitled to display a set of ads (related to the marketing campaign) to upcoming visitors, for a certain period of
time and considering a number of requirements (e.g., budgetary constraints).

Unlike traditional media, where audiences remain largely anonymous, the online setting allows the publisher to make display decisions on a user-by-user basis. In that regard, publishers may use all information available at the moment (e.g. profile information) to customize (in real time) the set of ads to be displayed during a user’s visit. Such profile information is available to the publisher whether is consciously provided by the user (e.g., by filling a subscription or survey) or collected without the user’s explicit consent (e.g., by the use of internet cookies). This key feature of the online channel presents both potential benefits and challenges to the publisher. On one hand, customization has the potential to make the displayed content more relevant to each user, which should result on higher click-through-rates and hence on higher revenues. On the other hand, real time customization and profile information collection require the use of more advanced and automated decision tools.

The publisher’s display decision is further complicated by the high level of uncertainty associated to operating on a dynamic environment: every time a contract is signed, new ads become available for display; however, to determine how attractive an ad is to a particular user, the publisher must display the ad to similar users and observe the associated click-through-rate (and its dependence on the ad mix being displayed). While this ad experimentation allows one to recover user preferences, it also involves incurring on an opportunity cost coming from not displaying ads that are thought to be more profitable. Hence, a publisher must balance the opportunity cost associated to ad experimentation with the benefit associated to recovering user preferences. This is the classical exploration versus exploitation trade-off: on one hand, the longer one explores user preferences the less time one has to exploit such a knowledge; on the other hand, less time spent on studying user preferences translates into higher uncertainty on observed click-through-rates, which might hamper the revenue maximization objective.

**Main objectives.** This chapter studies a family of stylized problems where a publisher –facing a stream of heterogeneous utility-maximizing visitors– needs to decide, based on each user’s profile information, on the bundle of ads to be displayed, so as to maximize the cumulative revenue derived from the observed click-through-rates. In our setup, user
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

42

click decision depends, among other factors, on each ad’s mean utility, which we assume is a function of the user’s profile information. In that regard, the key feature of our model is that we assume user preferences are driven by a set of \textit{initially unknown} parameters.

Our main objective is to study the impact of learning user preferences on the publisher’s cumulative revenue; in particular, our focus is on efficient parameter estimation through adaptive ad-mix experimentation, and its impact on cumulative revenue maximization. To shed light on this issue, we ignore the effect of ad location on click-through-rates, advertisers’ budgetary constraints and click fraud, among other factors. In addition, we assume user profile information is collected and classified prior to the user arrival.

\textbf{Key insights and main contribution.} We focus on the set of policies that consider the visitors’ observable characteristics when selecting the ad bundle to display during his/her visit. Furthermore, we restrict attention to policies that make such decisions based solely on additional information collected through the interaction with previous visitors (this will be defined in Section 3.3 as the class of non-anticipating policies). The publisher would like to select the policy that maximizes the (expected) cumulative revenue. One can measure policy performance in terms of the revenue loss experienced relative to the performance of a clairvoyant publisher with prior knowledge of the model parameters.

The main result in this chapter establishes that a publisher with prior knowledge on the model parameters gains additional revenue of at most order logarithm of the total number of visitors during the relevant horizon. In addition, our result establishes that this additional gain in revenue is proportional to the \textit{minimum} amount of information needed to reconstruct the policy used by the clairvoyant publisher. Regarding information collection, the result sheds light on the benefit of using information coming from displaying an ad to given set of users segments to explain the ad’s performance on different user types. In particular, the result suggests an optimal policy should be able to use information collected through optimal ad-mix display. Finally, our result establishes the precise frequency and extent of experimentation that guarantees best achievable performance.

On the practical side, we propose a family of adaptive ad-mix policies that achieve the fundamental bound mentioned above. These policies focus on identifying, for each ad, a set of user segments to base parameter estimation on; these sets are selected so as to minimize
the cost of information gathering. Once found, exploration efforts are focalized on such sets at the frequency suggested by the fundamental performance limit mentioned above. As a result, these policies quickly identify ad/profile pairs whose performance either do not affect the optimal display decision, or can be estimated by other means rather than from direct exploration: in particular, we show that such pairs can be detected after a small number of visits (independent of the total number of visits during the relevant horizon), hence one can limit the extent of experimentation on such pairs.

Regarding the exploration versus exploitation trade-off, we quantify the benefit coming from pooling information across different user profiles. In particular, we provide a sharp characterization of these benefits in terms of the structure of the full information problem (this is the problem a clairvoyant publisher will solve to decide on the ad-mix to display to a given user).

**Organization of the chapter.** The next section reviews related work. Section 3.3 formulates the publisher’s decision problem. Section 3.4 analyzes informational structure of the full information benchmark and establishes a fundamental limit on the performance of any ad-mix policy. Section 3.5 proposes an adaptive ad-mix policy for a specific instance of the problem and establishes a performance guarantee that matches the performance limit in Section 3.4. Finally, Section 3.6 presents our concluding remarks and additional implementation challenges. Proofs are relegated to two appendices, B.1 and B.2.

### 3.2 Literature Review

**Revenue management in online advertisement.** Revenue maximization in the context of online advertising has been studied from several perspectives in the last few years. Araman and Fridgeirsdottir (2010) study the dynamic pricing problem of an online publisher facing an upcoming stream of advertisers, with whom s/he might sign cost-per-impression (CPM) contracts (here the advertiser pays the publisher every time an ad is displayed). Here, the focus is on managing an uncertain demand for display slots while considering an uncertainty supply (users’ visits). Fridgeirsdottir and Najafi (2010) studies a similar problem under the CPC pricing model when advertisers’ requests are originated by an
advertising network. Similarly, Roels and Fridgeirsdottir (2009) study the publisher’s problem when prices are exogenously given, and the publisher may accept or reject upcoming contracts.

The work of Kumar and Sethi (2009) studies the publisher’s dynamic pricing problem when revenues might be collected from both advertisement and subscription fees. There, the authors study how to dynamically adjust subscription fees as well as the size of the advertisement space so as to maximize total revenues. Radovanovic and Zeevi (2011) study dynamic allocation of advertisement inventory for the case of reservation-based online advertisement. Here, advertisers’ requests take the form of target performances (in terms of, for example, observed click-through-rates) subject to periodic budgetary constraints, and the publisher needs to decide on the bundle of online advertisement inventory to assign to each advertiser (assuming fixed prices).

Using a mechanism design approach, Chen (2010) studies the optimal dynamic allocation of display impressions between guaranteed advertisers (who sign CPM contracts with the publisher) and advertisers competing on a spot market where impressions are allocated through periodic share auctions.

It is worth noting that, to best of our knowledge, all the studies above assume click-through-rates are initially known for any given ad-mix. Instead, we focus on the case where user preferences are initially unknown, and the publisher must reconstruct a preference model while maximizing cumulative revenues.

Learning approach to interactive media. Gooley and Lattin (2000) study the problem of a marketer with limited prior knowledge on user preferences that needs to decide on the customized set of messages to present to an upcoming stream of heterogeneous consumers. Their formulation is close to ours in that consumers are assumed to make decisions according to a discrete choice model whose parameters are initially unknown: like in our model, the publisher has access to profile information before deciding on the ad-mix to be presented; however, the authors do not provide neither a fundamental performance limit nor theoretical performance guarantees for the proposed policies. On a similar context, the work of Merserau and Bertsimas (2007) presents a dynamic programming formulation of the learning problem of consumer preferences for the case of a single segment of users, when
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

Display decisions must be taken for batches of consumers.

Regarding the online advertisement application area, Rusmevichientong and Williamson (2006) develop adaptive policies for keyword selection (in the context of search-based online advertisement) when click-through-rates are initially unknown. Finally, in a related subject Radovanovic and Heavlin (2010) study the design of price-experimentation mechanisms in the context of reservation-based online advertisement. This study discusses the practical and theoretical challenges associated to reconstructing advertisers’ demand for display inventory as a function of price.

**Connection to assortment planning literature.** In some practical settings, the publisher might have limited or no access to user profile information. A plausible approach in such instances is to assume that users are homogenous in their preferences. Under such assumption the publisher’s problem can be casted as a *dynamic assortment planning* problem, where a retailer with limited initial knowledge regarding user preferences and limited display capacity must decide, in a dynamic fashion, on the set of products to display to upcoming consumers so as to maximize cumulative revenues over a finite selling horizon. In such a setting users are a-priori identical, so customization of the offerings is not considered.

To the best of our knowledge, Caro and Gallien (2007) were the first to study this problem, motivated by an application in fast fashion. Rusmevichientong *et al.* (2010) formulate the dynamic assortment problem under the hypothesis of utility maximizing consumers and considering a limit on the assortment size. Using the Multinomial Logit (henceforth, MNL) framework to model purchase decisions, the authors propose an adaptive algorithm for joint parameter estimation and assortment optimization. Considering a more general random utility model, Chapter 2 of this document establishes a fundamental limit on performance that any policy must respect, and provides an adaptive algorithm achieving such a limit.

A plausible approach to solve the publisher’s problem is to solve a separate dynamic assortment problem for each user type. Hence, one can use the aforementioned work to provide a detailed characterization of the performance of an optimal *separable* policy. We will use such an approach as a benchmark. Thus, by providing similar results for the publisher’s problem we are able to identify the benefit of pooling exploration efforts across
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

46

different user profiles.

**Connection to Multi-armed bandit literature.** Multi-armed bandit problems provide a suitable framework to study the exploration versus exploitation trade-off in sequential decision making with incomplete model specification; see Lai and Robbins (1985) and Auer et al. (2002) for a classical formulation and solution approach to the problem, respectively. In fact, the solution approaches in both Caro and Gallien (2007) and Rusmevichientong et al. (2010) map the dynamic assortment problem to equivalent bandit formulations. Similarly, Abe and Nakamura (1999) study a bandit formulation for scheduling display-based online advertisement. The work of Pandey et al. (2007) studies a bandit-based approach to match ads to display slots (across several web pages) from the perspective of a search engine. More recently, Agarwal et al. (2009) studies a multi-armed bandit formulation for finding the ad that maximizes the click-through-rate across all users, where performance is evaluated under dynamic environmental conditions. To best of our knowledge, all studies above disregard user-level information to customize the display decision.

Leaving the online advertisement literature aside, our work relates to Rusmevichientong and Tsitsiklis (2010) in the treatment of the linear dependence of the reward distribution on specific arm attributes. Similarly, the work of Woodroofe (1979) focus on bandit problems where additional contextual information on reward realizations is available; this resembles the availability of profile information in our setup. (See Zeevi and Goldenshluger (2011), Wang et al. (2005) and Rigollet and Zeevi (2010) for different approaches to such contextual bandit problems.) Back in the realm of online advertisement, Lu et al. (2010) and Kleinberg et al. (2008) use contextual multi-armed bandit formulations to study ad selection when strategy sets are metric spaces and payoff functions satisfy Lipschitz conditions with respect to the respective metrics. Our formulation differs significantly from these bandit formulations, from how arms and pulls are defined to how we model reward distributions.

### 3.3 Problem Formulation

**Model primitives.** We consider an online publisher endowed with a set of $N$ ads which can be displayed through a unique web-page to upcoming visitors. Limited by the web-page
layout, the retailer can display at most \( C \) ads simultaneously to any given visitor. Each
time a visitor clicks on ad \( i \), the publisher collects a marginal revenue \( w_i > 0, i \in \mathcal{N} \). We let
\( w := \{w_1, \ldots, w_N\} \) denote the vector of revenue margins. We assume each ad is available
for display during a finite horizon, common to all ads.

Let \( T \) to denote the total number of visitors during the relevant horizon. (The value of
\( T \) is in general not known to the publisher a priori, therefore it will be treated as random.)
We use \( t \) to index visitors according to their arrival times, so \( t = 1 \) corresponds to the first
arrival, and \( t = T \) to the last. We assume that upon each visitor’s arrival, and before making
a display decision, the publisher observes information on the visitor’s profile. Specifically,
we assume that profile information on visitor \( t \), summarized on a \( d \)-dimensional vector \( X_t \),
is available instantaneously upon arrival, and hence can be used to customize the displayed
ad selection.

While we make no assumptions on the nature of the information contained on \( X_t \), we do
assume it belongs to a finite set \( \mathcal{X} \) of possible visitor profiles. (Without loss of generality, we
assume \( \mathcal{X} \) spans \( \mathbb{R}^d \).) In that regard, we assume that the publisher observes the segment to
which a visitor belongs to, rather than a possibly unique set of user’s features. Consequently,
two visitors sharing the same profile are seen as equivalent a-priori. Similarly, a single user
is perceived as a new entity upon each visit, independent of the evolution of his/her profile
information. With this in mind, we will use the terms visitor and user interchangeably.

We assume user profiles \( X_1, \ldots, X_T \) form a sequence of i.i.d. random variables drawn
from a distribution \( G \) whose domain is \( \mathcal{X} \). In that regard, one can see \( G(x) \) as the fraction
of the user population sharing the profile \( x \in \mathcal{X} \), which we assume remains constant during
the relevant horizon.

**User click decision.** We assume user click decision is driven by utility maximization:
user \( t \) assigns a utility \( U_t^i \) to ad \( i \), for \( i \in \mathcal{N} \cup \{0\} \), with
\[
U_t^i(X_t) := \mu_i(X_t) + \xi_t^i,
\]
where \( \mu_i(x) \in \mathbb{R} \) denotes the mean utility assigned to ad \( i \) by all users sharing the profile
\( x \in \mathcal{X} \), \( \{\xi_t^i : i \in \mathcal{N}, t = 1, \ldots, T\} \) are i.i.d. random variables drawn from a common dis-
tribution \( F \), and ad 0 represents an always-available no-click alternative. We assume \( F \) to
be absolutely continuous with respect to Lebesgue measure in \( \mathbb{R} \). During his/her visit, user
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

$t$ clicks on the ad with the highest utility (among those displayed) if that utility is greater than the one provided by the no-click alternative, which we assume is profile-independent. For convenience, and without loss of generality, we set $\mu_0(x) := 0$ for all $x \in X$.

Our discrete-choice approach to click-behavior implicitly assumes that each user clicks on at most a single ad, and that the location of an ad within a given ad-mix does not impact its click-through-rate. While the first of these assumptions is a reasonable one (considering click-through-rates observed in practice), the second one does not typically hold (at least not for premium locations within a web-page). These assumptions provide sufficient tractability for analysis purposes, and allow us to extract structural insights.

We complete the specification of the choice model by assuming that $\mu_i(x)$ can be cast as a parameterized function of the user profile $x \in X$, for $i \in \mathcal{N}$. In particular, we assume the publisher (linearly) regresses mean utilities on the user’s profile attributes. That is

$$\mu_i(x) := \beta_i^\top x,$$

$x \in X$, where $\beta_i \in \mathbb{R}^d$ is a vector of factors associated to ad $i \in \mathcal{N}$. Our assumption is quite general: more complex functions of the user profile can be considered by redefining the profile vector components (e.g., quadratic functions can be considered by augmenting the profile vector to include any term of the form $x_j x_i$). We let $B := \{\beta_{ji}\} \in \mathbb{R}^{d \times \mathcal{N}}$ denote the matrix of factors.

The static optimization problem. Let $\mathcal{S} := \{S \subseteq \mathcal{N} : |S| \leq C\}$ denote the set of all possible ad-mixes, where here, and in what follows, $|A|$ denotes the cardinality of a set $A$. For a given set $S \in \mathcal{S}$ and a given matrix $M \in \mathbb{R}^{d \times \mathcal{N}}$, the probability $p_i(S, M)$ that a user with profile $x \in X$ clicks on ad $i$ is given by

$$p_i(S, M, x) = \int_{-\infty}^{\infty} \prod_{j \in S \cup \{0\} \setminus \{i\}} F(z - M_j^\top x) \ dF(z - M_i^\top x),$$

for $i \in \mathcal{S} \cup \{0\}$, where $M_i$ denote the $i$-th column of matrix $M$, and $p_i(S, M, x) = 0$ otherwise. The expected revenue $r(S, M, x)$ associated to displaying ad-mix $S \in \mathcal{S}$ to a user with profile $x \in X$ when the matrix of factors is $M$ is given by

$$r(S, M, x) := \sum_{i \in S} w_i \ p_i(S, M, x).$$
Were the value of \( B \) known to the publisher, it would be optimal for him/her to display \( S^*(B, X_t) \) to user \( t \), where

\[
S^*(B, x) \in \arg\max_{S \in \mathcal{S}} \{ r(S, B, x) \},
\]

for \( x \in \mathcal{X} \). We will assume throughout this chapter that the solution to the (3.2) is unique (i.e., \(|S^*(B, x)| = 1\) for all \( x \in \mathcal{X} \). (This assumption simplifies our construction of fundamental performance bounds, and can be relaxed without compromising such result: see proof of Theorem 5). Solving (3.2) for a given profile amounts to find the best ad-mix over a combinatorial number of possibilities. Such a combinatorial problem in general hard to solve: we assume the publisher has access to \( S^*(M, x) \) for any matrix \( M \) and profile \( x \in \mathcal{X} \). (We will see \( S^*(M, x) \) can be computed efficiently for the case of the model specification we analyze in Section 3.5.)

The dynamic optimization problem. We assume that the publisher knows the distributions \( F \) and \( G \), but does not know the matrix \( B \) driving user decisions. That is, the publisher knows how likely is that an upcoming visitor belongs to a given user segment, and understands how users with same profile vary in their valuation for the same ad, but has no prior information on how such a valuation varies across users with different profiles and across different ads.

Upon a user’s arrival, the retailer observes the user’s profile and, taking into account all information available at the time, decides on the ad-mix to display during the user’s visit. More formally, consider the ad-mix process \( \{S_t \in \mathcal{S} : 1 \leq t \leq T\} \) and let

\[
Z^t_i := 1 \{ i \in S_t, U^t_i > U^t_j, j \in S_t \setminus \{i\} \cup \{0\} \}
\]
denote the click decision of user \( t \) regarding ad \( i \in S_t \), where \( Z^t_i = 1 \) indicates that user \( t \) clicks on ad \( i \), and \( Z^t_i = 0 \) otherwise. Similarly, we let \( Z^t_0 := 1 \{ U_0 > U_j, j \in S_t \} \) denote the overall click decision of user \( t \), where \( Z^t_0 = 1 \) if user \( t \) chooses not to click on any ad, and \( Z^t_0 = 0 \) otherwise. Throughout this chapter, we let \( 1 \{ A \} \) denote the indicator function of a set \( A \). We let \( Z^t := (Z^t_0, Z^t_1, \ldots, Z^t_N) \) denote the decision made by user \( t \) and define \( \mathcal{F}_t \) as the history associated with the ad-mix process \( \{S_u : 1 \leq u < t\} \), click decisions \( \{Z^u : 1 \leq u < t\} \), user profiles \( \{X_u : 1 \leq u < t\} \), and the profile \( X_t \), for \( t \in \{1, \ldots, T\} \) (with \( \mathcal{F}_0 = \emptyset \)).
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

We say an ad mix process is non-anticipating if each of its elements is determined solely on the basis of information collected through interaction with previous visitors (and the user profile), i.e., $S_t$ is $F_t$-measurable, for all $t \in \{1, \ldots, T\}$. An admissible ad-mix policy $\pi$ provides a mapping from past history and user profiles to $S$ such that the induced ad-mix process is non-anticipating. We restrict attention to the set of such policies, and denote it by $\mathcal{P}$. We use $\mathbb{E}_\pi$ and $\mathbb{P}_\pi$ to denote expectations and probabilities of random variables when the ad-mix policy $\pi \in \mathcal{P}$ is used.

The publisher’s objective is to choose a policy $\pi \in \mathcal{P}$ to maximize $J^\pi(T, B)$, the expected revenue cumulated throughout the interaction with all $T$ users. That is,

$$J^\pi(T, B) := \sum_{t=1}^T \mathbb{E}_\pi \left\{ \sum_{i \in S_t} w_i Z^t_i \right\}.$$ 

Consider the case of publisher that knows upfront the value of the matrix $B$: such a publisher would display $S^*(B, X_t)$ to user $t$, for all $t \leq T$. Let $J^*(T, B)$ denote the expected cumulative revenue achieved by such a publisher, throughout the interaction with all $T$ users, i.e.,

$$J^*(T, B) := T \mathbb{E}_x \{ r(S^*(B, x), B, x) \}.$$ 

This quantity provides an upper bound on the expected revenue generated by any admissible policy, i.e., $J^*(T, B) \geq J^\pi(T, B)$ for all $\pi \in \mathcal{P}$. For any $T > 0$, we define $\mathcal{R}^\pi(T, B)$ to be the cumulative expected revenue loss associated to following policy $\pi$ instead of the one used by the clairvoyant publisher. That is,

$$\mathcal{R}^\pi(T, B) := J^*(T, B) - J^\pi(T, B).$$

The regret associated to policy $\pi$ can be viewed as a measure of revenue loss due to the lack of prior knowledge on the parameters governing user preferences. We see that the publisher’s maximization objective is equivalent to that of regret minimization. In that regard, the publisher objective is to gather enough information to reconstruct the clairvoyant solution (with increasing probability), and doing it so while losing the smallest possible revenue in the process.

The next section focuses on characterizing the minimum amount of information required to solve (3.2) exactly, and on establishing fundamental limits on the revenue loss associated
to policies that balance the cost of information collection and immediate revenue maximization in an optimal fashion.

### 3.4 Fundamental Limit on Achievable Performance

We begin by characterizing the minimum information necessary to recover the solution to the static optimization problem. After that, we focus on how to collect such information while minimizing the revenue loss.

**Information requirements.** Fix $x \in \mathcal{X}$ and let $\mathcal{N}(x)$ include any ad $i$ that can be made to enter the optimal ad-mix $S^*(B, x)$ by unilaterally modifying its mean utility $\mu_i(x)$. That is,

$$\mathcal{N}(x) := \left\{ i \in \mathcal{N} : \exists M \in \mathbb{R}^{d \times N} \text{ s.t. } i \in S^*(M, x), M_{-i} = B_{-i} \right\},$$

where $M_{-i}$ denotes the matrix $M$ with column $i$ removed. Suppose $\mu_i(x)$ is known for all $i \in S^*(B, x)$: in order to verify the optimality of $S^*(B, x)$ one must know the value of $\mu_i(x)$ for all $i \in \mathcal{N}(x)$. In contrast, mean utilities of ads outside $\mathcal{N}(x)$ do not affect the optimality of $S^*(B, x)$. One concludes that, in order to solve (3.2) correctly, one needs to reconstruct the mean utilities only for ads in $\mathcal{N}(x)$.

In addition, we argue that the solution to (3.2) can be recovered even if there is not complete certainty about the underlying parameters: by the absolute continuity of $F$ with respect to Lebesgue measure on $\mathbb{R}$ one has that the click probabilities given in (3.1) are continuous functions of the model parameters. This, together with our assumption on the uniqueness of the solution to the static problem, implies that

$$\Lambda := \left\{ M \in \mathbb{R}^{d \times N} : S^*(x, M) = S^*(x, B), x \in \mathcal{X} \right\},$$

has a non-empty interior, hence one can recover the clairvoyant’s solution by displaying $S^*(M, x)$ to users with profile $x \in \mathcal{X}$, as long as $M$ is not too far ($M \in \Lambda$) from $B$.

Fix $i \in \mathcal{N}$ and define $\mathcal{X}(i)$ as the set of profiles for which ad $i$ is potentially optimal. That is,

$$\mathcal{X}(i) := \left\{ x \in \mathcal{X} : i \in \mathcal{N}(x) \right\},$$

for $i \in \mathcal{N}$. In order to recover the clairvoyant’s solution, one might as well solve (3.2) using a matrix $M$ such that $M_i^\top x \approx \beta_i^\top x$ for all $x \in \mathcal{X}(i)$, for all $i \in \mathcal{N}$ (more precisely, $M \in \Lambda$).
Efficient information gathering. Any good ad-mix policy must collect enough information on ad $i$ to (approx.) reconstruct $\mu_i(x)$ for all $x \in \mathcal{X}(i)$, $i \in \mathcal{N}$. Consider a set $E \subseteq \mathcal{X}$: for any $x \in \text{span} \{E\}$ one has that

$$
\mu_i(x) = \sum_{y \in E} \alpha_y \mu_i(y),
$$

for all $i \in \mathcal{N}$, for some finite constants $\{\alpha_y : y \in E\}$, where $\text{span} \{A\}$ refers to the linear span (linear hull) of the set $A$. Hence, a good policy might focus on approximating $\mu_i(x)$ only for $x \in E$, provided that $\mathcal{X}(i) \subseteq \text{span} \{E\}$, $i \in \mathcal{N}$.

Consider now the cost of information collection. For ad $i \in \mathcal{N}$, define $O(i)$ as the set of profiles in $\mathcal{X}$ for which ad $i$ is optimal. That is

$$
O(i) := \{x \in \mathcal{X} : i \in S^*(B, x)\},
$$

for $i \in \mathcal{N}$. It is easy to see that $O(i) \subseteq \mathcal{X}(i)$, for all $i \in \mathcal{N}$. Fix $i \in \mathcal{N}$ and consider a profile $x \in O(i)$: information on ad $i$ might be collected without incurring in revenue loss (in expectation) if one displays ad-mix $S^*(B, x)$ to users with profile $x$. Hence, a good policy might collect sufficient information to approximate $\mu_i(x)$ for all $x \in O(i)$ while keeping revenue loss to a minimum.

Suppose ad $i \in \mathcal{N}$ is such that $\mathcal{X}(i) \subseteq \text{span} \{O(i)\}$; information collected on profiles in $O(i)$ is enough to reconstruct the solution to the static problem when $B_{-i}$ is known. Moreover, such information might be collected without incurring in revenue loss. On the other hand, when ad $i \in \mathcal{N}$ is such that $\text{span} \{O(i)\} \subset \mathcal{X}(i)$, information collection is likely to incur in revenue loss, as information collected on profiles in $O(i)$ does not suffice to reconstruct the solution to (3.2) even when $B_{-i}$ is known. Here, one would expect a good policy to minimize the revenue loss associated to information collection outside $O(i)$.

Next, we present a fundamental limit on what can be achieved by any good ad-mix policy and connect such a result to the intuition on minimum information requirements and efficient information gathering developed here.

---

1. This follows from assuming that mean utilities are affine functions of the profile vector, and the fact that any vector spanned by $E$ can be written as a linear combination of profiles in $E$. Our approach can extend to more general parametric functions by augmenting the profile vector definition.

2. The linear span of a set $A = \{v_1, \ldots, v_a\}$ is defined as $\text{span} \{A\} := \{v_1 \lambda_1 + \ldots + v_a \lambda_a : \lambda \in \mathbb{R}^a\}$. 
3.4.1 A lower bound on the performance of any admissible policy

We say that an admissible policy is consistent if for all $M \in \mathbb{R}^{d \times N}$

$$\frac{\mathcal{R}^\pi(T, M)}{T^a} \to 0,$$

as $T \to \infty$, for every $a > 0$. In other words, the frequency at which a consistent policy offers the optimal assortment to upcoming visitors converges to 1 for all possible values for $M$. The restriction in (3.3) imposes a uniform bound on the grow rate of the regret across all possible factor matrices. Let $\mathcal{P}' \subseteq \mathcal{P}$ denote the set of non-anticipating, consistent ad-mix policies. We restrict our attention to consistent policies; this prevents us from considering policies that perform well for specific values of $B$ by mere chance. The result below establishes a fundamental limit on what can be achieved by any consistent ad-mix policy.

**Theorem 5.** For any $\pi \in \mathcal{P}'$, and any $B \in \mathbb{R}^{d \times N}$,

$$\mathcal{R}^\pi(T, B) \geq \left( \sum_{i \in \mathcal{N}} K_i \right) \log T \text{ a.s.,}$$

for a finite positive constants $K_i$, $i \in \mathcal{N}$.

The regret of a policy is directly related to its ability to reconstruct the solution to the static optimization problem. In that regard, this result establishes an asymptotic lower bound on the number of suboptimal display decisions that any consistent policy must take (in expectation). Moreover, such a bound is derived by establishing a series of lower bounds on the number of times each ad is displayed in a suboptimal display. In that regard, our result asserts that ad $i \in \mathcal{N}$ must be displayed in a suboptimal display to at least order $K_i \log T$ users. Moreover, from proof of Theorem 5 one has that $K_i$ will be proportional (we will be precise in the proof outline below) to rank $\{X(i) \setminus \text{span} \{O(i)\}\}$, were rank $\{A\}$ denotes the maximal number of linearly independent elements in the set $A$.

The above is aligned with the discussion in the previous section: rank $\{X(i) \setminus \text{span} \{O(i)\}\}$ is precisely the minimum number of profiles that, in addition to the ones in $O(i)$, allow to reconstruct the mean utility of ad $i$ for all profiles in $X(i)$. In that regard, the result above asserts that the regret associated to a good policy follows from revenue losses incurred while collection information on those profiles.
Connection to previous results in Dynamic Assortment Planning. The publisher’s problem for the case when $\mathcal{X} = \{x_1\}$ corresponds to an instance of the dynamic assortment planning problem. There, a single optimal assortment (ad-mix) must be recovered by collecting information from the interaction with homogenous consumers (visitors). In Chapter 2, we show that the regret in this setting is at least of order $(|\mathcal{N}(x_1) - S^*(B, x)|) \log T$. Such a lower bound coincides with the one in Theorem 5, however our result suggests the inefficiency of solving the publisher problem by decomposing it into $\mathcal{X}$ assortment problems: such an approach will result in a performance bound of order $\sum_{x \in \mathcal{X}} (|\mathcal{N}(x) - S^*(B, x)|) \log T$.

In particular one has that

$$\sum_{i \in \mathcal{N}} \text{rank} \{\mathcal{X}(i) \setminus \text{span} \{O(i)\}\} \leq \sum_{i \in \mathcal{N}} (\mathcal{X}(i) - O(i)) = \sum_{x \in \mathcal{X}} (|\mathcal{N}(x) - S^*(B, x)|).$$

The comparison above quantifies the value of combining information collected through interaction with some type of users to explain the clicking behavior of a larger set of visitors. For example, suppose $B$ is such that $S^*(B, x) \subset \mathcal{N}(x)$ for all $x \in \mathcal{X}$, and that $\text{rank} \{O(i)\} = d$ for all $i \in \mathcal{N}$, then one has that

$$\sum_{i \in \mathcal{N}} \text{rank} \{\mathcal{X}(i) \setminus \text{span} \{O(i)\}\} = 0 < |\mathcal{X}| \leq \sum_{x \in \mathcal{X}} |\mathcal{N}(x) \setminus S^*(B, x)|,$$

and hence, while Theorem 5 suggests a finite revenue loss is attainable, revenue loss under the (myopic) separable approach will be at least of order $|\mathcal{X}| \log T$.

Remark 8 (Implications for design of “good” policies). The early analysis in this section, together with the result in Theorem 5 suggest a number of properties that good policy should have. First, one might reconstruct the solution to the static optimization problem by collecting information on ad $i$ when displayed to users with profiles spanning $\mathcal{X}(i)$; moreover, these profiles should include those on which product $i$ is optimal. Second, information collected on profiles for which ad $i$ is not optimal must be collected through interaction with at most order $(\log T)$ users; in addition, the number of profiles on which information is to be gathered must be kept at the minimum possible. We use these properties in Section 3.5 to design an efficient ad-mix policy.
3.4.2 Proof outline for Theorem 5

We bound the regret of any consistent policy by computing a lower bound on the expected number of suboptimal displays for each ad $i \in \mathcal{N}$, thus the lower bound in Theorem 5 is presented as a sum of individual contributions made by each ad in $\mathcal{N}$. For computing the lower bound associated to a given ad we use a change-of-measure argument introduced by Lai and Robbins (1985) for proving an analogous result for a multi-armed bandit problem.

The argument can be roughly described as follows. If $\mathcal{X}(i) \subseteq \text{span}\{O(i)\}$ for ad $i \in \mathcal{N}$, then optimal display of the ad is sufficient to collect all information (related to the ad) necessary for recovering the solution to (3.2). If this is not the case, then one can find an alternative factor-matrix such that: i) it only differs from the original matrix $B$ on the factors associated to ad $i$; ii) mean utilities for profiles in $O(i)$ remain unchanged; and iii) ad $i$ is optimal for at least one profile outside $O(i)$. When such alternative factor-matrix is considered, any consistent policy must display ad $i$ to order $T$ users with profiles outside $O(i)$. These are users for which ad $i$ is optimal under the alternative factor-matrix but it is not under the original configuration. We show that this later fact implies that, under the original configuration, ad $i$ must be displayed to at least order $\log T$ (in expectation) users with profiles outside $O(i)$, and that such display is proportional to the distance between the alternative and original factor-matrices (we formalize this notion of proximity in the proof of Theorem 5). In particular, we prove that for any consistent policy $\pi$

$$
\mathbb{P}_{\pi}\{T_i \leq \log T/H_i\} \to 0
$$

as $T \to \infty$, where $T_i$ denotes the total number of users with profiles outside $O(i)$ to which ad $i$ is displayed to, and $H_i$ is a finite positive constant proportional to the distance between the alternative and original configurations. Each constant $K_i$ in Theorem 5 is derived (up to a scale factor) by minimizing $H_i$ across alternative configurations that increases the set of profiles for which ad $i$ is optimal. Using Markov’s inequality we have that for any $i \in \mathcal{N}$,

$$
\liminf_{t \to \infty} \frac{\mathbb{P}_{\pi}\{T_i\}}{\log T} \geq C K_i.
$$

The result in Theorem 5 follows directly from the equation above, the maximum size of any ad-mix, and the connection between revenue loss and suboptimal display of ads in
3.5 The Proposed Dynamic Ad-mix Policy

This section builds on the insight derived from the proof of Theorem 5 to design an ad-mix policy attaining (essentially) such a performance. To keep things concrete, we will anchor our analysis around the well known MNL model which arises when $F$ is a Gumbel distribution. Later, in Section 3.6 we explain how our analysis (proposed policy and its performance guarantee) extends to the case of more general distribution functions (under some mild regularity assumptions).

Next, we first identify the main challenges in ad-mix policy design. Then, we explain our approach to face such challenges. Finally, we propose an ad-mix policy and show it is optimal in a precise mathematical sense.

Multinomial Logit choice model. Taking $F$ to have a standard Gumbel distribution is equivalent to say that each user makes his/her clicking decision according to a MNL choice model. While the MNL model admits some deficiencies that are well documented (see, e.g., Kok et al. (2008) for a discussion of those in the context of assortment planning), it is central to the existing literature in fields such as economics, marketing and operations research.

We initially consider the MNL specification for several reasons. First, one can derive a closed form expression for (3.1). In particular, for $S \subseteq N$ and $x \in \mathcal{X}$ one has that (see Anderson et al. (1992))

$$p_i(S, B, x) = \frac{\exp(\beta_i^\top x)}{1 + \sum_{j \in S} \exp(\beta_j^\top x)},$$

for $i \in S$, and $p_i(S, B, x) = 0$ otherwise. Second, one can use (3.5) to describe both $N(x)$ and the solution to (3.2) in a precise manner, for all $x \in \mathcal{X}$. Third, there is an efficient algorithm to solve the static optimization problem for the MNL specification; see
Finally, we make use of the *independence from irrelevant alternatives* (IIA) property of the MNL model to estimate the model parameters efficiently by combining information coming from displaying different ad-mixes to different users. All these elements facilitate the performance analysis of the proposed algorithm.

### 3.5.1 Main implementation challenges

Fix $i \in \mathcal{N}$. Any consistent policy should collect information on ad $i$ over a set of profiles $E(i)$ allowing the reconstruction of its mean utility for all profiles for which ad $i$ is potentially optimal. Moreover, such set $E(i)$ should be as small as possible while still including profiles for which ad $i$ is optimal. Unfortunately, such a set can only be identified on the basis of knowledge of $B$, which is initially unknown.

One can see that it is possible for parameter estimation to affect information collection efforts, which at the same time will condition parameter estimation. Hence, one risks entering into a vicious circle where failing to approximate the model parameters will result on failing to recover the problem structure which might deepen the error in parameter estimation. The challenge here is to simultaneously estimate model parameters and guide exploration efforts in such a way that both $B$ and $\mathcal{X}(i)$ are correctly approximated.

Regarding the extent of the exploration efforts, Theorem 5 hints on the amount of information any consistent policy should base parameter estimation on: suboptimal exploration must be conducted on at most order $\log T$ users. Unfortunately, the total number of visitors $T$ is, in most instances, initially unknown. This feature prevents the use of ad-mix policies that separate exploration and exploitation a priori. (For that purpose one would need to know upfront the amount of information to collect.) The challenge here is to devise a policy that will perform well (in the sense of Theorem 5) for any realization of $T$.

### 3.5.2 Preliminaries

**Parameter estimation.** Fix $i \in \mathcal{N}$. Suppose we want to estimate $\beta_i$ based on the information collected through interacting with the first $t-1$ visitors. In particular, we look for a vector $\hat{\beta}_i$ explaining ad $i$’s click-through-rates for users with profiles on a set $E(i) \subseteq \mathcal{X}$. 

[Rusmevichientong *et al., 2010]
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

(we explain how to select \(E(i)\) later on this section). From (3.5) we have that

\[
p_i(S, B, x) = \exp(\beta_i^T x) \cdot p_0(S, B, x),
\]

for all \(S \subseteq \mathcal{S}\) and \(x \in \mathcal{X}\). We use this property to establish the following lemma.

**Lemma 1.** For \(x \in \mathcal{X}\) and \(i \in \mathcal{N}\) one has that

\[
\mathbb{E}_\pi \left\{ \sum_{s=1}^{t-1} \mathbf{1}_{X_s = x, Z_s^i = 1} \right\} = \exp(\beta_i^T x).
\]

According to Lemma 1, for \(i \in \mathcal{N}\) one has that

\[
\beta_i \in \Lambda_i := \left\{ \beta \in \mathbb{R}^d : \log \left( \frac{\mathbb{E}_\pi \left\{ \sum_{s=1}^{t-1} \mathbf{1}_{X_s = x, Z_s^i = 1} \right\}}{\mathbb{E}_\pi \left\{ \sum_{s=1}^{t-1} \mathbf{1}_{X_s = x, i \in S_s, Z_s^0 = 1} \right\}} \right) = \beta^T x, \quad x \in E(i) \right\},
\]

for any \(E(i) \subseteq \mathcal{X}\) such that rank \(\{E(i)\} \leq d\). In addition, suppose \(E(i)\) is such that \(\mathcal{X}(i) \subseteq \text{span} \{E(i)\}\), for all \(i \in \mathcal{N}\), and take \(\hat{\beta}_i \in \Lambda_i\) arbitrarily, \(i \in \mathcal{N}\): one has that \(\hat{B} := (\hat{\beta}_1, \ldots, \hat{\beta}_N) \in \Lambda\), thus one can recover the solution to the static problem by solving (3.2) using \(M = \hat{B}\) (i.e., \(S^*(B, x) = S^*(\hat{B}, x)\) for all \(x \in \mathcal{X}\)).

With this in mind, and assuming \(E(i)\) is properly chosen, we will estimate \(B\) by constructing a proxy for \(\Lambda_i\), from which we will choose \(\hat{\beta}_i\), for all \(i \in \mathcal{N}\). More precisely, we use the estimate \(\hat{B} := \{\hat{\beta}_{j,i}\}\), where

\[
\hat{\beta}_i \in \left\{ \beta \in \mathbb{R}^d : \log \left( \frac{\sum_{s=1}^{t-1} \mathbf{1}_{X_s = x, Z_s^i = 1}}{\sum_{s=1}^{t-1} \mathbf{1}_{X_s = x, i \in S_s, Z_s^0 = 1}} \right) = \beta^T x, \quad x \in E(i) \right\}.
\]

**Estimation set selection.** We aim to select a set of profiles \(E(i)\) that: \((i)\) is contained in \(\mathcal{X}(i)\); and \((ii)\) explains the mean utility of ad \(i\) for all profiles in \(\mathcal{X}(i)\), for all \(i \in \mathcal{N}\). The next lemma characterizes \(\mathcal{N}(x)\) for all \(x \in \mathcal{X}\), for the MNL model.

**Lemma 2.** For each \(i \in \mathcal{N}\) one has that

\[
\mathcal{N}(x) = \{i \in \mathcal{N} : w_i \geq r(S^*(B, x), B, x)\}.
\]

Estimating \(\mathcal{X}(i)\) requires not only to guess the optimal ad-mix \(S^*(B, x)\) but also to approximate its (expected) revenue. While the former might be recovered by solving (3.2) using any matrix in \(\Lambda\), the later can only be approximated. In that regard, a policy might
be able recover $\mathcal{X}(i)$ only if there is a minimum separation between profit margins and optimal expected revenues. We introduce the following assumption, which we assume holds throughout the rest of this chapter.

**Assumption 3.** The matrix $B$ is such that $|w_i - r(S^*(B, x), B, x)| > 0$ for all $i \in \mathcal{N}$ and $x \in \mathcal{X}$.

Suppose $\hat{B}$ is computed according to (3.7). We use Lemma 2 to approximate $\hat{X}(i)$. That is, we use

$$\hat{\mathcal{X}}(i) := \left\{ x \in \mathcal{X} : w_i \geq r(S^*(\hat{B}, x), \hat{B}, x) \right\}. \quad (3.8)$$

We select $E(i)$ (used for estimating $\hat{B}$) so that it spans $\hat{\mathcal{X}}(i)$. The idea here is the following: if enough information is collected for estimation one would expect $\hat{\beta}_i^\top x$ to be close to its real value, for all $x \in \text{span}\{E(i)\}$, in particular one would have that $\hat{\beta}_i^\top x$ is close to its real value for all $i \in S^*(\hat{B}, x)$, for all $x \in \mathcal{X}$; hence $r(S^*(\hat{B}, x)$ would be close to its real value, and from the optimality of $S^*(B, x)$ one will have that $\mathcal{X}(i) \subseteq \hat{\mathcal{X}}(i)$. This implies that continuous improvement on parameter estimation will translate into improvement on the selection of the estimation sets.

Given $\hat{\mathcal{X}}(i)$ there are many possible choices for $E(i)$: ours will prioritize inclusion of profiles where ad $i$ has been displayed the most. The idea here is that, if the overall scheme works, one should identify the solution to (3.2) with increasing probability, thus ad $i$ should be displayed to users with profiles in $O(i)$ on a linear fashion, and as a consequence some of such profiles will become part of $E(i)$.

**Length of exploration.** Theorem 5 suggests that parameter estimation should be based on interaction with order $\log T$ users. Since $T$ is initially unknown, we aim to perform parameter estimation for user $t$ based on the interaction with order $\log t$ users. Our policy considers a variation of the **doubling** trick (see Cesa-Bianchi and Lugosi (2006)), where time (visitors) is partitioned into intervals of exponentially increasing length, and were the suggested amount of information is collected within each interval. For $n \geq 0$ we define

$$\tau(n) = \inf \left\{ t \geq 1 : \kappa \log t \geq n \right\},$$

where $\kappa$ is an arbitrarily chosen (for now) positive and finite constant. During the $n$-th interval, encompassing users $\tau(n)$ to $\tau(n + 1) - 1$, we look to collect information on ad $i$
from the interaction with at least \( n + 1 \) users with profile \( x \), for each profile \( x \in E(i) \), and for each ad \( i \in \mathcal{N} \).

### 3.5.3 The Proposed Ad-mix Policy

Our proposed ad mix policy attempts to reconstruct the problem’s structure at specific moments in time, using all information available at those times. Specifically, the policy updates parameter estimates and estimation sets at the beginning of each user interval. Parameter estimates are used to (probably) maximize revenue collection throughout the correspondent user interval. In addition, our policy collects a minimum amount of information on each ad for profiles on the newly computed estimations sets, preparing the ground for updating parameter estimates at the beginning of the next user interval. Next, we elaborate on how our policy implements the ideas above.

At the beginning of the \( n \)-th user interval, model parameters are re-estimated and used to update a proxy for each \( \mathcal{X}(i) \), which will drive information collection during the \( n \)-th interval. Starting from arbitrarily chosen matrix \( \hat{B}^0 \) and initial estimation sets \( \{E^0(i) : i \in \mathcal{N}\} \), each one spanning \( \mathcal{X} \), our policy executes the following steps at the beginning of the \( n \)-th interval, \( n \geq 1 \):

(a) Parameter estimate \( \hat{B}^n \) is computed according to (3.7), using estimation sets \( E^n(i) \), \( i \in \mathcal{N} \). In addition, estimates are chosen so that \( x \in \text{span} \{E^{n-1}(i)\} \) for each ad \( i \) in \( S^*(\hat{B}^n, x) \), for all \( x \in \mathcal{X} \).

(b) Using \( \hat{B}^n \), we select each \( E^n(i) \) to span \( \hat{X}^n(i) \) (computed as in (3.8)), prioritizing most displayed profiles.

Step (a) above updates model parameters at the beginning of the \( n \)-th interval: such an update will be computed with information on \( n \) user interactions (for each add/profile pair) with high probability; in addition, such an update is selected so that ads though to be optimal for a given profile during the \( n \)-th interval have the correspondent mean utility estimates computed using information on \( n \) user interactions. Step (b) computes the new estimation set for each ad \( i \in \mathcal{N} \); exploration efforts during the \( n \)-th interval will be directed
to collect enough information on such profiles; specific details about the selection of $E^n(i)$ are shown in Algorithm 4.

Within the $n$-th interval, we select the ad-mix to display during the visit of user $t$ as follows:

(i) We select any ad $i$ that has been displayed to less than $n + 1$ users with profile $X_t$, provided that $X_t \in E^n(i)$, $i \in \mathcal{N}$.

(ii) If no ad is found in the step above, then we display $S^*(\hat{B}^n, X_t)$.

Step (i) above aims to collect enough information to update model parameters at the beginning of the next $(n + 1)$-th user interval. Step (ii) exploits the information collected during the first $n - 1$ user intervals by offering a probably optimal ad mix.

The proposed policy is summarized for convenience in Algorithm 4. There, $T_i(t, x)$ denotes the number of users with profile $x \in \mathcal{X}$ to whom ad $i \in \mathcal{N}$ has been displayed to before user $t$. That is

$$T_i(t, x) := \sum_{u=1}^{t-1} 1\{i \in S_t, X_t = x\},$$

for $i \in \mathcal{N}$, $x \in \mathcal{X}$ and $t \geq 1$. The tuning parameter $\kappa$ controls the extent of information collection on estimation sets.

### 3.5.4 Performance Analysis

The next result characterizes the performance of the proposed ad-mix policy.

**Theorem 6.** Let $\pi = \pi(w, \kappa)$ be defined by Algorithm 4. Then, the regret associated to $\pi$ is bounded for all $T$ as follows

$$R^\pi(T, B) \leq \kappa \left( \sum_{i \in \mathcal{N}} \text{rank}\{\mathcal{X}(i) \setminus \text{span}\{O(i)\}\} \right) \log T + K,$$

provided that $\kappa \geq \kappa$, where $\kappa$ and $K$ are finite positive constants.

The constants $K$ and $\kappa$ depend on instance specific quantities, such as minimum optimality gaps for the solution to (3.2), but not on the size of product set $\mathcal{N}$, or total number of users $T$. Proof of Theorem 6 portraits $K$ as an upper bound on the expected cumulative
Algorithm 4: $\pi = \pi(w, \kappa)$

**STEP 1. Initialization:**

Set $E_0(i) := E \subseteq X$, for all $i \in \mathcal{N}$, with $E$ such that $X \subseteq \text{span}\{E\}$ and $\text{rank}\{E\} \leq d$.

Set $B^0 \in \mathbb{R}^{d \times N}$ with $\beta_{j,i} = 0$ for all $1 \leq j \leq d$ and $i \in \mathcal{N}$.

Set $n = 1$.

**STEP 2. Joint Exploration and Optimization:**

for $t = 1$ to $T$ do

Set $F := \{i \in \mathcal{N} : X_t \in E^{n-1}(i), T_i(t, X_t) < n\}$ [Candidates for exploration]

if $F \neq \emptyset$ then

Display $S_t \subseteq F$, with $S_t \in \mathcal{S}$.[Exploration]

else

Display $S_t \in \mathcal{S}(\hat{B}^{n-1}, X_t)$.[Exploitation]

end if

if $t = \tau(n)$ then

Set $\hat{B}^n \in \mathbb{R}^{d \times N}$ to be any solution to

$$
\log \left( \frac{\sum_{s=1}^{t-1} 1 \{X_s = x, Z_i^s = 1\}}{\sum_{s=1}^{t-1} 1 \{X_s = x, i \in S_s, Z_0^s = 1\}} \right) = \hat{\beta}_i^\top x, \ x \in E^{n-1}(i), \ i \in \mathcal{N},
$$

such that $i \notin S^*(\hat{B}^n, x)$ for all $x \notin \text{span}\{E^{n-1}(i)\}, i \in \mathcal{N}$.[Estimation]

for $i \in \mathcal{N}$ do

Set $\hat{X}_i^n := \left\{ x \in X : w_i \geq r(S^*(\hat{B}^n, x), \hat{B}^n, x) \right\}$.[Interesting profiles]

Set $E^n(i) := \{x_1, \ldots\}$ such that

i) $x_l \in \hat{X}_i^n$, for all $l$,

ii) $x_l \notin \text{span}\{x_1, \ldots, x_{l-1}\}$, for all $l$, and

iii) $T_i(t, x_l) \geq T_i(t, x)$ for all $x \in \hat{X}_i^n \setminus \text{span}\{x_1, \ldots, x_{l-1}\}$, for all $l$.

end for

$n \leftarrow n + 1$

end if

end for
revenue loss due to errors while reconstructing the solution to (3.2), while \( \pi \) relates to the minimum exploration intensity that makes \( K \) finite when \( T \to \infty \).

The performance guarantee in Theorem 6 matches the performance limit provided in Theorem 5 up to an scaling factor: the regret associated to the proposed policy can be expressed as the sum of individual contributions made by every ad \( i \in \mathcal{N} \), each of them being order \( \log T \) and proportional to the minimum number of profiles (in addition to those on which the ad is optimal) from which one needs to collect information to recover the solution to the static optimization problem.

The analysis performed in the proof of Theorem 6 reveals an important feature of the proposed policy: it bases parameter estimation on a right set of profiles with increasing probability. That is, our policy identifies both \( \mathcal{X}(i) \) and \( \mathcal{O}(i) \) with increasing probability, for all \( i \in \mathcal{N} \). As a consequence, revenue loss due to suboptimal exploration is kept at the minimum possible (in expectation). This means that our policy successfully limits exploration of ad \( i \) on user profiles for which the ad is not potentially optimal (i.e., profiles outside \( \mathcal{X}(i) \)), for all \( i \). In addition, the analysis implies that information coming from optimal display of each ad is used for parameter estimation with increasing probability.

**Remark 9 (Selection of the tuning parameter \( \kappa \)).** Theorem 6 requires one to select \( \kappa > \pi \), however our characterization of \( \pi \) (see proof of Theorem 6) depends on the factor-matrix \( B \), which is initially unknown. Thus one faces the risk of selecting \( \kappa < \pi \), compromising the validity of the performance guarantee. One can avoid such a risk by redefining

\[
\tau(n) = \inf \{ t \geq 1 : \kappa (\log t)^\alpha \geq n \},
\]

for \( \alpha > 1 \). With this, the upper bound in Theorem 6 would read

\[
\mathcal{R}^\pi(T, B) \leq \kappa \left( \sum_{i \in \mathcal{N}} \text{rank} \{ \mathcal{X}(i) \setminus \text{span} \{ \mathcal{O}(i) \} \} \right) (\log T)^\alpha + K,
\]

independent of the value of \( \kappa \), and \( \pi \) becomes optimal up to an order \( (\log T)^{\alpha-1} \)-term.

### 3.5.5 Numerical Illustration

In this section we illustrate the results established so far by means of a simple numerical example. We consider the case of \( N = 4 \) and set \( w = \{0.63, 0.59, 0.56, 0.60\} \). User profiles
Profile & $S^*(B, x)$ & \{1, 2\} & \{2, 3\} & \{2, 4\} \\
\hline
$r(S^*(B, x), B, x)$ & 0.587 & 0.546 & 0.578 \\
\hline

Table 3.1: Solution to static optimization problem for the numerical example.

<table>
<thead>
<tr>
<th>Ad</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O(i)$</td>
<td>{x_1}</td>
<td>{x_1, x_2, x_3}</td>
<td>{x_2}</td>
<td>{x_3}</td>
</tr>
<tr>
<td>$X(i)$</td>
<td>{x_1, x_2, x_3}</td>
<td>{x_1, x_2, x_3}</td>
<td>{x_2}</td>
<td>{x_1, x_2, x_3}</td>
</tr>
<tr>
<td>rank ${X(i) \setminus \text{span}{O(i)}}$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3.2: Minimum suboptimal experimentation for each ad.

are drawn from $\mathcal{X} = \{x_1, x_2, x_3\}$ with equal probability. We consider the case of $d = 2$ and set

\[
x_1 = \begin{pmatrix} 0.1 \\ 0.9 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 0.1 \\ 0.9 \end{pmatrix}.
\]

We set the display capacity $C = 2$, and set the model parameter $B$ to be

\[
B = \begin{pmatrix} -1.30 & 2.00 & 2.75 & 3.00 \\ 3.00 & 2.00 & 2.75 & -1.30 \end{pmatrix},
\]

i.e., $\beta_1^T = (-1.30 \ 3.00)$.

**Static optimization.** Table 3.1 depicts the solution to the static optimization problem and the optimal expected revenue, for each profile in $\mathcal{X}$. We use Lemma 2 to quantify the theoretical contribution each ad makes to the regret. According to this information, summarized in Table 3.2, our results predict that the regret associated to the proposed policy will be of order $\log T$, due to suboptimal display of ads $i = 1, 4$. (There, rank $\{X(i) \setminus \text{span}\{O(i)\}\}$ quantifies the amount of suboptimal experimentation required to recover the solution to (3.2), for ad $i \in \mathcal{N}$.)

**Proposed policy performance.** Figure 3.1 depicts the average performance of policy $\pi(w, \kappa)$ over 500 replications, using $\kappa = 25$, and considering $T = 1$ to $T = 100000$. There, graphs (a) and (b) illustrate the regret of the proposed policy as a function of $T$ and $\log T$, respectively.
respectively. The dotted lines represent 95% confidence intervals for the simulation results. From panels (a) and (b) we observe that the associated regret is indeed of order $\log T$, as predicted.

![Graph](image)

**Figure 3.1**: Performance of the proposed policy $\pi(w, \kappa)$.

In this example, the proposed policy makes suboptimal display decisions on a small fraction of users, ranging from around 10% when the total number of visitors is $T = 1000$, to around 0.5% when the total number of visitors is $T = 100000$. (Recall that the regret relates to the number of users to whom a suboptimal ad-mix is displayed to.)

Figure 3.2 provides further information on the performance of the proposed policy on this example. There, graph (a) depicts display of ad 3 outside $\mathcal{X}(3)$ as a function of $T$. The graph in (b) shows total number of visitors to whom suboptimal ad-mixes are displayed to, as a function of $T$. The dotted lines represent 95% confidence intervals for the simulation results. In panel (a) we observe that $\pi(w, \kappa)$ successfully limits display of ad 3 to users with profiles $x_1$ and $x_3$ (i.e., with profiles outside $\mathcal{X}(3)$). Panel (b) shows the number of users to whom a suboptimal ad-mix is displayed to, when in exploitation phase. Thus, one can argue that the regret in Figure 3.1 scales with $T$ exclusively due to suboptimal information collection.
CHAPTER 3. DYNAMIC LEARNING AND CUSTOMIZATION IN DISPLAY-BASED ONLINE ADVERTISEMENT

Figure 3.2: Display of suboptimal ad-mixes.

Quality of parameter estimation. Policy \( \pi(w, \kappa) \) focuses exploration efforts into identifying a matrix \( M \in \Lambda \), and not necessarily on approximating \( B \). For example, in our numerical illustration one only needs to estimate \( \beta_3^\top x_2 \) rather than \( \beta_3 \) itself \( (X(3) = \{x_2\}) \).

This fact can be observed, for example, in the output of the algorithm for an arbitrary replication: after \( T = 100000 \) visitors one has that \( \hat{\beta}_3^\top x_2 = (5.465 \ 0.000) \); while \( \hat{\beta}_3 \) is not close to \( \beta_3 \) componentwise, one has that \( \beta_3^\top x_2 = 2.75 \approx 2.73 = \hat{\beta}_3^\top x_2 \).

3.6 Concluding Remarks

This chapter has studied the optimization problem faced by a publisher deciding on the (customized) ad-mixes to display to a heterogeneous stream of upcoming visitors, when click decisions are driven by utility maximization and model parameters are initially unknown.

We have established fundamental limits of performance for any consistent policy, and we have proposed a family of ad-mix policies attaining such a performance limit for a particular instance of the problem, namely when users make decisions according to a MNL model.

Complexity of the publisher problem. Theorem 5 provides a lower bound on the regret every consistent policy must incur on. Such result is derived under mild assumptions and
is valid for any absolutely continuous (w.r.t. Lebesgue measure in \( \mathbb{R} \)) distribution function \( F \). Our analysis of the lower bound links revenue loss to suboptimal information collection necessary to reconstruct the full information solution, given by (3.2). In particular, our bound predicts that finite revenue loss might be attainable when displaying the optimal ad-mix to each visitor provides enough information to reconstruct the solution to the static optimization problem.

In the practical side, we have shown that the theoretical lower bound in Theorem 5 can be achieved, up to constant terms, when idiosyncratic shocks to user utility are Gumbel distributed. In particular, we proposed an ad-mix policy whose regret scales optimally with the number of visitors, and exhibits the right dependence on the problem structure. We did this by exploiting properties of the MNL model that allowed us to estimate the model parameters by combining information coming from the display of different ad-mixes.

A more general ad-mix policy. We see several challenges in designing a policy for an arbitrary distribution function \( F \). Most discrete choice models do not allow for closed form expression for the click probabilities in (3.1). This does not only affect parameter estimation but also prevents us from characterizing \( N(x) \) for any profile \( x \in \mathcal{X} \). An ad-mix policy must be able to identify the set of profiles for which an ad is potentially optimal, so information collection efforts can be correctly focalized.

A more general ad-mix policy, for which the performance guarantee in Theorem 6 still applies, can be easily devise by imposing the following identifiability restrictions on the distribution function \( F \).

**Assumption 4 (Identifiability).** For any ad mix \( S \in \mathcal{S} \), profile \( x \in \mathcal{X} \) and vector \( \rho \in \mathbb{R}_+^N \) such that \( \sum_{i \in S} \rho_i < 1 \) one has that;

(i) the system of equations \( \{ p_i(S, M, x) = \rho_i, \ i \in S, \ M^T x = \eta \} \) has a unique solution \( T(S, \rho) \) in \( \eta \in \mathbb{R}^N \) such that \( \eta_i = 0 \) for all \( i \notin S \);

(ii) \( p(S, \cdot, x) \) is Lipschitz continuous, and \( T(S, \cdot) \) is locally Lipschitz continuous in the neighborhood of \( \rho \), for all \( S \in \mathcal{S} \);

(iii) for any matrix \( M \in \mathbb{R}^{d \times N} \) and profile \( x \in \mathcal{X} \) one has that

\[
r(S^*(M, x), M, x) \leq w_i, \quad \text{for all } i \in S^*(M, x).
\]
Assumptions \( (i) \) and \( (ii) \) ensure that parameter estimation is possible and that the interior of \( \Lambda \) is not empty. Assumption \( (iii) \) allows one to characterize \( \mathcal{N}(x) \) through Lemma 2 for all \( x \in \mathcal{X} \). When Assumption \( (i) \) hold true, one can modify our proposed policy so parameter estimates are based on the display of fixed ad-mixes (covering \( E(i) \) rather than on individual ad display. This will limit the amount of information used for estimation purposes (in particular, it will require to separate parameter estimation for profiles in \( O(i) \) from that for profiles in \( \mathcal{X}(i) \setminus O(i) \)), nevertheless one will have that the same arguments in proof of Theorem 6 will apply, despite such modifications.

**Comparison of our policy with benchmark results.** As advanced in Section 3.4, our policy improves on a naive dynamic assortment approach to the publisher’s problem. By exploiting the relationship between mean utilities across different user profiles, our policy keeps overall exploration efforts to a bare minimum; this results on significant gains (in terms of revenue) when compared to an approach that separates the publisher problem across user segments. In that regard, the performance guarantee in Theorem 6 indicates that the relative revenue gap between the performance of our policy and the coming from an optimal dynamic assortment approach to the problem is of order \( \left( \sum_{i \in \mathcal{N}} |\mathcal{X}(i) \setminus O(i)| - \text{rank} \{\mathcal{X}(i) \setminus \text{span} \{O(i)\}\} \right) \log T \).
Chapter 4

Product Assortment and Price

Competition with Informed Consumers

4.1 Introduction

Motivation and main objectives. Assortment planning and pricing decisions are fundamental drivers of consumers’ purchase decisions and ultimately of a retailer’s profitability. Retailers face significant challenges to understand the mapping from such decisions to consumer behavior as this mapping should synthesize complex aspects of purchase decisions such as, for example, substitution behavior, consumers’ collection and aggregation of information, consumer heterogeneity, and the effect of competition. In particular, consumers’ access and use of information about the retailers’ offerings play a critical role with respect to the way consumers purchase decisions are made. If access to information is costly, consumers might make purchase decisions based only on partial knowledge about retailers’ offerings. In cases in which obtaining such information does not require significant effort and cost, which is often true for online retailers, consumers can form preferences over the full set of retailer-product pairs, and make purchasing decisions accordingly. For example, when retailers offer the same product, consumers will select the retailer to buy from based
on retailer-specific attributes, such as price.

Once the mapping between offerings and purchase decisions is at hand, retailers face the additional challenge of making appropriate assortment and pricing decisions and the latter is often further complicated by the presence of operational constraints such as limited shelf space, stocking restrictions and other capacity related considerations.

The objective of this chapter is to study equilibrium behavior in assortment and pricing competition in the presence of informed consumers, i.e., when consumers have full information regarding retailers’ offerings. In particular, we aim at understanding how the competition that results from the availability of such information impacts equilibrium behavior. To that end, we analyze a model of assortment and pricing competition in a duopolistic setting, when assortment decisions are constrained by limited display capacities and retailers have access to both common products, i.e., products that are available to both competitors, and exclusive products, i.e., products that are unavailable to competition. We anchor the analysis around the well studied Multinomial Logit choice model for consumer demand. While the Logit model admits some deficiencies that are well documented (see, e.g., Kok et al. [2008] for a discussion of those in the context of assortment planning), it is central to both the existing assortment literature and practice, and as a result serves as a good starting point to study competition in the presence of informed consumers.

Summary of main contributions. We consider two competitive settings: i.) the case where retailers select only their assortments given exogenously fixed prices; and ii.) the case where retailers can select both assortments and prices. The main findings of this chapter can be summarized as follows.

i.) When prices are exogenously given and retailers have access only to exclusive products, we show that an equilibrium always exists (see Theorem 7) and provide a bound on the number of equilibria (see Theorem 8 and the discussion that follows). In addition, we establish that, when multiple equilibria exist, retailers will always prefer the same equilibrium, as laid out in Proposition 2.

---

1We note here that in cases where assortment information is costly, a model where customers would first decide on a retailer and then on the product to purchase (such as a nested Logit) might be more appropriate.
ii.) When prices are exogenously fixed and retailers have access to both common and exclusive products, we provide sufficient conditions for existence of an equilibrium. In addition, we show that in the general case, it is now possible that an equilibrium fails to exist or that the number of equilibria grows exponentially with the retailers’ display capacity.

iii.) When retailers compete in both assortment and price decisions, we prove that at most one equilibrium exists. We show that the existence of an equilibrium is driven by the number of exclusive products (in comparison to the display capacity) that retailers have. In particular, we identify the only possible equilibrium candidate and provide a simple procedure to confirm or invalidate its equilibrium status. For the case where only exclusive products are available, an equilibrium is always guaranteed to exist. These results are summarized in Theorem 9.

Many of the results for the setting with exclusive products can be generalized to the case of an arbitrary number competing retailers and we comment on this point throughout this chapter.

The current work contributes to the existing assortment and pricing literature on various fronts. From a modeling point of view, we provide a framework to analyze competition in the presence of informed consumers when demand is driven by individual utility maximization. This framework enables among other things the possibility of having retailers offer the same products, a possibility ignored in most of the existing literature. The current work sheds light on equilibrium properties in general and on the implications of the availability of common products on such properties. From a qualitative perspective, it shows, for example, that, starting from an equilibrium, the introduction of a new product to the set of available products of both retailers may lead to a situation where no equilibrium exists. For the case of assortment-only competition with exclusive products, the chapter establishes a clear connection, in any equilibrium, between the attractiveness of the offered assortment (this concept will be defined more formally in Section 4.4 but can be thought of as the breadth of an assortment), the profit made by a retailer, and the attractiveness of the competitor’s assortment. More precisely, the results show that the higher the attractiveness of the
offered assortment, the lower the profit achieved, and the higher the attractiveness of the competitor’s assortment, in equilibrium. In particular, the results indicate that in general, competition will lead to a broader assortment offering, i.e., a retailer that faces competition will offer a broader set of products than if s/he were operating as a monopolist with the same display capacity. At the same time, when multiple equilibria exist, the above implies that both retailers will prefer, among all equilibria, the one that minimizes the overall attractiveness of the product assortments.

From a methodological viewpoint, the analysis builds on the idea of computing best responses via a simple problem transformation. Such an approach has been previously used in various settings when faced with a combinatorial optimization problem with a rational objective function. It was, for example, used by Dantzig et al. (1966) for finding the minimal cost to time ratio cycle in a network, by Megiddo (1979) for computational complexity results on the optimization of rational objective functions, and more recently by Rusmevichientong et al. (2010) in the closely related context of monopolistic assortment optimization with Logit demand. The current work leverages this transformation to present a unified framework for the analysis of both the cases of assortment-only and assortment and price competition. This framework enables one to derive a crisp characterization of equilibrium properties and can be used to analyze extensions to the original model, as illustrated in Section 4.6.

The remainder of the chapter. The next section reviews related work. Section 4.3 formulates the model of assortment and price competition. Section 4.4 studies assortment-only competition, where prices are fixed exogenously, while Section 4.5 focuses on joint assortment and price competition. Finally, Section 4.6 presents possible extensions and some associated challenges. Proofs are relegated to Appendix C.

4.2 Literature review

A key building block in any assortment optimization problem lies in the definition of the process through which customers choose among a selection of products. In this work we focus on the multinomial Logit (MNL) choice model. Discrete choice models are extensively
CHAPTER 4. PRODUCT ASSORTMENT AND PRICE COMPETITION WITH INFORMED CONSUMERS

discussed in Ben-Akiva and Lerman (1985) and Train (2002), and the estimation and design of specialized models has been an active area in Marketing (see, e.g., Guadagni and Little (1983) for an early reference).

In the current work, both assortment and pricing decisions are considered in a competitive environment. We begin by reviewing the relevant literature on assortment planning before focusing on the existing literature on competitive models.

**Assortment Optimization.** The key challenges associated with assortment optimization often stem from i.) the need to account for substitution effects when introducing products in an assortment; and ii.) the presence of operational constraints captured through the costs associated with different assortments or display constraints.

The problem of assortment planning has often been studied in conjunction with inventory decisions, starting with the work of van Ryzin and Mahajan (1999). In the latter, considering a Logit demand model and assuming customers do not look for a substitute if their choice is stocked out (known as static substitution), the authors show that one can significantly simplify the assortment problem as there are only a limited number of candidates to consider for the optimal assortment. Maddah and Bish (2007) study a similar model, where in addition, the retailer could select prices; see also Aydin and Ryan (2000) for a study in the absence of inventory considerations. In a similar setting as the one in van Ryzin and Mahajan (1999) Gaur and Honhon (2006) analyze the implications of the alternative modeling of demand through a location choice model, building on the early work of Hotelling (1929) on the broad conclusions obtained under Logit demand.

The case of customers looking for substitutes if their choice is stocked out, known as stock-out based substitution, was studied in conjunction with inventory decisions by Smith and Agrawal (2000) and Mahajan and van Ryzin (2001) In this setting, Goyal et al. (2009) showed that the assortment problem is in general NP-Hard and proposed near-optimal heuristics for a particular choice model.

In the present work, we do not consider inventory decisions and assume that products that are included in a retailer’s assortment are always available when requested; hence stock-out based substitution does not arise. In particular, we focus on the case where the retailers have display constraints. Such a setting with Logit demand and fixed prices in
a monopolistic context has been studied in Chen and Hausman (2000), where the authors analyze mathematical properties of the problem and in Rusmevichientong et al. (2010), where the authors provide an efficient algorithm for finding an optimal assortment. Fisher and Vaidyanathan (2009) also study assortment optimization under display constraints and highlight how such constraints arise in practice. In the absence of display constraints, the assortment problem arises in the network revenue management literature; see Gallego et al. (2004) and Liu and van Ryzin (2008) for results under the MNL model. When demand is generated by a mixture of Logit (also referred to as a latent class model), Miranda et al. (2009) show that the assortment optimization problem is NP-Hard (see also Rusmevichientong et al. (2009)).

Alternative consumer choice models have been considered in the assortment planning literature; for example, Cachon et al. (2005) study assortment decisions when consumers might search across different stores for additional products and illustrate the consequences of failing to incorporate consumer search in the assortment optimization. In a similar modeling spirit but with a computational focus in the context of network revenue management, Gallego (2010) incorporates the possibility that consumer choice may depend on the products that are not offered. A detailed review of the literature on monopolistic assortment optimization and of industry practices can be found in Kok et al. (2008).

**Competitive environment.** Our work builds on the monopolistic studies above to derive conclusions in competitive settings, where the literature is less extensive. As in the case of monopolistic studies, locational choice models have also been considered (see, e.g., Alptekinoglu and Corbett (2008)) but most of the studies are anchored around the Logit model and some of its extensions, as is the present one.

Price competition under choice models has been studied and is still an active area of research. Anderson et al. (1992) study oligopoly pricing for single-product firms under Logit demand and study pricing and assortment depth for multi-product firms in a duopoly with a nested Logit demand, restricting attention to symmetric equilibria. When firms offer a single product and customer choice is described by an attraction model, Bernstein and Federgruen (2004) establish existence and uniqueness of an equilibrium for profit maximizing firms and Gallego et al. (2006) generalize this result for different cost structures. For the Logit
model, Konovalov and Sándor (2009) provide guarantees for the existence and uniqueness of an equilibrium for affine cost functions when firms may have multiple products. Allon et al. (2010) provide conditions that ensure existence and uniqueness of an equilibrium under MNL demand with latent classes.

Misra (2008) studies joint assortment and price competition of retailers offering exclusive products with MNL demand and in the presence of display constraints, and conducts an empirical study to analyze the impact of competition on assortment size and prices. The analytical results obtained focus on best response analysis but do not provide equilibrium existence or uniqueness results. We also refer the reader to Draganska et al. (2009) for an empirical investigation of assortment and pricing strategies in oligopolistic markets.

Additional dimensions of competition as well as alternative consumer choice models have also been analyzed. Hopp and Xu (2008) study joint pricing, service rates and assortment competition under MNL demand with random market size and operational costs associated with the assortment size. They show existence of an equilibrium and provide sufficient conditions for uniqueness. Cachon et al. (2008) analyze how consumer search (across retailers) may influence equilibrium assortments and prices, focusing on symmetric equilibria. Cachon and Kök (2007) study price and assortment competition for retailers offering two categories when basket shopping consumers are present, with an emphasis on the impact of centralized category management. Recently, Kök and Xu (2010) investigate assortment competition under a hierarchical customer choice model, a nested MNL, focusing on the differences in the properties of best responses for the cases where customers choose a product by first selecting a brand or by first selecting a category.

It is worth noting that, to the best of our knowledge, all the competitive studies above focus on the case where firms offer exclusive products. We will show that the introduction of common products can lead to fundamentally different properties for the equilibrium set. In addition, the unified framework we propose is fairly flexible and enables one to analyze different models of competition under a common approach.
4.3 Model

We next describe the operational setting in which retailers compete and the demand model considered, and then present two competitive settings: one where retailers compete on assortments when prices are predetermined and one where retailers compete on both assortments and prices.

4.3.1 Operational setting

We consider duopolistic retailers that compete in product assortment and pricing decisions. We will index retailers by 1 and 2, and whenever we use \( n \) to denote a retailer’s index, we use \( m \) to denote her/his competitor’s (e.g., if \( n = 1 \), then \( m = 2 \)).

We assume retailer \( n \) has access to a subset \( \mathcal{N}_n \) of products, from which s/he must select her/his product assortment. In addition, we assume that, due to display space constraints, retailer \( n \) can offer at most \( C_n \geq 1 \) products. Such display constraints have been used and motivated in previous studies (see, e.g., Rusmevichientong et al. (2010), Misra (2008), and Fisher and Vaidyanathan (2009)). Without loss of generality, we assume that \( C_n \leq |\mathcal{N}_n| \), where here and in the remainder of the chapter, \(|A|\) denotes the cardinality of a set \( A \). We will let \( \mathcal{N} \) denote the set of all products, i.e., \( \mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 \), and denote its elements by \( \{1, \ldots, S\} \). For each product \( i \) in \( \mathcal{N} \), we let \( c_i \geq 0 \) denote the marginal cost resulting from acquiring a unit of the product, which will be assumed constant and common to both retailers.

We will say that product \( i \) is exclusive to retailer \( n \) if it belongs to \( \mathcal{N}_n \) but not to \( \mathcal{N}_m \); we denote the set of exclusive products for retailer \( n \) by \( \mathcal{N}_n \setminus \mathcal{N}_m \), where here and in all that follows, \( A \setminus B := A \cap B^c \) stands for the set difference between sets \( A \) and \( B \), and the complement of a set is taken relative to \( \mathcal{N} \). Similarly, we say that product \( i \) is common if it is available to both retailers, i.e., if \( i \) belongs to \( \mathcal{N}_1 \cap \mathcal{N}_2 \).

For \( n \) in \( \{1, 2\} \), we define \( \mathcal{S}_n \) to be the set of feasible assortment selections for retailer \( n \), i.e.,

\[
\mathcal{S}_n := \{ A \subseteq \mathcal{N}_n : |A| \leq C_n \}.
\]

We also let \( A_n \) and \( p_n := (p_{n,1}, \ldots, p_{n,S}) \) denote the assortment selection and vector of prices.
offered by retailer \( n \), respectively. Note that we include all products in \( \mathcal{N} \) for notational convenience but it should be clear that the only prices that will matter in this vector are those that correspond to the assortment selection of the retailer. Additionally, we say that retailer \( n \) offers a full-capacity assortment if \( |A_n| = C_n \), and that an assortment is exclusive if it consists only of exclusive products.

### 4.3.2 Demand model and retailers’ objective

We assume that customers have perfect information about product assortments and prices from both retailers. For a given assortment decision \( A_n \) in \( \mathcal{S}_n \) made by retailer \( n \), we assume that customer \( t \) assigns a utility \( U_{n,i}(t) \) to buying product \( i \) in \( A_n \) from retailer \( n \), and utility \( U_{n,0}(t) \) to not purchasing any product, where

\[
U_{n,i}(t) := \mu_i - \alpha p_{n,i} + \xi^t_i, \quad n = 1, 2 \\
U_{n,0}(t) := \mu_0 + \xi^t_0.
\]

In the above, the \( \mu_i \)'s are finite and deterministic parameters representing an adjusted mean utility assigned to product \( i \), and \( \alpha > 0 \) is a parameter of price sensitivity. We assume \( \{\xi^t_i : i \in \mathcal{N} \cup \{0\}\} \) are independent Gumbel distributed random variables with location parameter 0 and scale parameter 1. Note that these random variables, which represent idiosyncratic shocks to utility, are independent of the retailer \( n \) and hence we are implicitly assuming that consumers do identify the fact that products are identical. In the current setting, since the mean utilities (\( \mu_i \)'s) and costs (\( c_i \)'s) are assumed not to depend on the retailers, the same product will always generate the same utility for a consumer if offered at the same price. However, it is worth noting that the case of retailer-dependent mean utilities or costs can be analyzed using the same techniques we use throughout this work. We do not pursue this analysis to keep the notation as simple as possible. For convenience, and without loss of generality, we set \( \mu_0 := 0 \).

Customers are utility maximizers; customer \( t \) computes the best option from each retailer, \( i_n \) in \( \text{argmax} \{U_{n,i}(t) : i \in A_n \cup \{0\}\} \), for \( n \in \{1, 2\} \), and then selects option \( i \) that belongs to \( \text{argmax} \{U_{1,i}(t), U_{2,i}(t)\} \). As noted above, utility maximization may be attained simultaneously at a single product offered by both retailers at the same price. In such a
case, we assume customers select any of the retailers, with equal probability.

For \(i = 1, \ldots, S\), define the attraction factor of product \(i\) when offered at price \(p\) as follows

\[
\nu_i(p) := e^{\mu_i - \alpha p}.
\]

We will refer to the sum of attraction factors over products in a given assortment as the \textit{attractiveness} of such an assortment.

The above setup leads to MNL demand where the customers’ consideration set is obtained after eliminating options that are strictly dominated, i.e., products that are offered at a lower price at another retailer. In particular, one can show that for given assortment and price decisions \(\{ (A_n, p_n) : n = 1, 2 \}\), the probability that a customer elects to purchase product \(i\) in \(A_n\) from retailer \(n\), \(q_{n,i}\), is given by

\[
q_{n,i}(A_n, p_n, A_m, p_m) := \frac{\nu_i(p_{n,i}) \left( 1\{i \notin A_m\} + \delta_{n,i}(p_n, p_m) 1\{i \in A_m\} \right)}{1 + \sum_{i \in A_n \setminus A_m} \nu_i(p_{n,i}) + \sum_{i \in A_n \cap A_m} \nu_i(\min(p_{n,i}, p_{m,i})) + \sum_{i \in A_m \setminus A_n} \nu_i(p_{m,i})},
\]

where \(1\{\cdot\}\) denotes the indicator function and

\[
\delta_{n,i}(p_n, p_m) := \begin{cases} 
1 & \text{if } p_n < p_m, \\
\frac{1}{2} & \text{if } p_n = p_m, \\
0 & \text{if } p_n > p_m,
\end{cases}
\]

defines the split of product \(i\)'s market share between the retailers (when offered by both) as a function of the prices. The expected profit per customer for retailer \(n\), is then written as \(\pi_n(A_n, p_n, A_m, p_m) := \sum_{i \in A_n} (p_{n,i} - c_i) q_{n,i}(A_n, p_n, A_m, p_m)\), or, equivalently

\[
\pi_n(A_n, p_n, A_m, p_m) = \frac{\sum_{i \in A_n \setminus A_m} (p_{n,i} - c_i) \nu_i(p_{n,i}) + \sum_{i \in A_n \cap A_m} \delta_{n,i}(p_n, p_m)(p_{n,i} - c_i) \nu_i(p_{n,i})}{1 + \sum_{i \in A_n \setminus A_m} \nu_i(p_{n,i}) + \sum_{i \in A_n \cap A_m} \nu_i(\min(p_{n,i}, p_{m,i})) + \sum_{i \in A_m \setminus A_n} \nu_i(p_{m,i})}.
\]

Each retailer’s objective is to maximize her/his expected profit per customer, given the competitors’s decisions.

### 4.3.3 Competition in product assortment and pricing decisions

**Assortment-only competition.** In this setup, prices are predetermined and not under the control of retailers, and these compete through their assortment selection. This setting
is relevant when, e.g., prices are determined by the manufacturers/service providers and not the retailers. Given retailer \( m \)'s assortment decision, retailer \( n \) selects an assortment so as to maximize her/his expected profit per customer subject to the display constraint on the number of products that can be offered. Mathematically, the problem that retailer \( n \) solves can be written as follows

\[
\max_{A_n \in S_n} \{ \pi_n(A_n, p_n, A_m, p_m) \}.
\] (4.1)

Problem (4.1) is a combinatorial problem where the retailer attempts to find the best set of products to offer among all the possible assortments in \( S_n \). We say that a feasible assortment \( A_n \) is a best response to \( A_m \) if \( A_n \) maximizes the profit per customer for retailer \( n \), i.e., if \( A_n \) solves problem (4.1). Given that there is a finite number of feasible assortments, there always exists at least one best response for every assortment \( A_m \) in \( S_m \). For \( n = 1, 2 \), let \( B_n(A_m) \) denote the set of best responses to \( A_m \), i.e., the set of assortments that achieve the maximum value for problem (4.1). We say that an assortment pair \( (A_1, A_2) \) is an equilibrium if \( A_n \) belongs to \( B_n(A_m) \) for \( n = 1, 2 \). Formally, this corresponds to the concept of a pure strategy Nash equilibrium.

**Joint assortment and price competition.** In this setting, retailers can decide on both the products to include in their assortments and the prices at which to offer them in a simultaneous fashion. In such a case, given an assortment selection \( A_m \) offered by retailer \( m \) with a corresponding vector of prices \( p_m \), retailer \( n \)'s objective is to maximize its expected profit per customer, which can be formalized as follows.

\[
\sup_{A_n \in S_n, p_n \in \mathbb{R}^S} \{ \pi_n(A_n, p_n, A_m, p_m) \}.
\] (4.2)

Problem (4.2) now combines continuous decision variables (prices) and discrete ones (assortments). As in the previous case, we will say that an assortment-price pair \( (A_n, p_n) \) is a best response to \( (A_m, p_m) \) if \( (A_n, p_n) \) maximizes the profit per customer for retailer \( n \), i.e., if \( (A_n, p_n) \) solves problem (4.2). With some abuse of notation, for \( n = 1, 2 \), we also let \( B_n(A_m, p_m) \) denote the set of all such solutions, i.e., the best response correspondence. (When problem (4.2) admits no solution, the latter set is empty.) We say that a 4-tuple \( (A_1, p_1, A_2, p_2) \) is an equilibrium in assortment and price decisions if \( (A_n, p_n) \) belongs to \( B_n(A_m, p_m) \) for \( n = 1, 2 \).
It is worth noting that while we pursue simultaneous assortment and price competition, there are settings where pricing decisions might only be made after assortment decisions are taken by both firms. The framework we use enables one to also analyze sequential competition and we comment on the type of results this leads to in Section 4.6.

4.4 Assortment-only Competition: Main Results

We consider retailers competing in assortment decisions when product prices are exogenously given. We start the analysis by studying the problem of computing the best response correspondence $B_n(\cdot)$, for $n = 1, 2$. We assume throughout this section that all products have a non-negative profit margin. Note also that while we do not impose any particular relationship between the $\mu_i$‘s, the product prices $p_{n,i}$’s and the costs $c_i$‘s, the analysis will of course apply to cases where such a relationship exists.

**Best response correspondence.** The analysis of the best response correspondence will rely on a simple equivalent formulation of the profit maximization problem. This idea of focusing on such an equivalent formulation when faced with a combinatorial problem with an objective function in the form of a ratio has previously been used in various settings as mentioned in the introduction. The reader is in particular referred to Gallego et al. (2004) and Rusmevichientong et al. (2010) for applications to monopolistic assortment optimization with Logit demand. The treatment that follows accounts for the appropriate modifications to the competitive setting we study. Consider the following problem

$$\max \lambda \quad \text{s.t.} \quad \max_{A \in S_n} \left\{ \sum_{i \in A \setminus A_m} (p_{n,i} - c_i - \lambda) \nu_i(p_{n,i}) + \sum_{i \in A \cap A_m} \left( \delta_{n,i}(p_n, p_m)(p_{n,i} - c_i) - \lambda \right) \nu_i(\min\{p_{n,i}, p_{m,i}\}) - \lambda \sum_{i \in A_m \setminus A} \nu_i(p_{m,i}) \right\} \geq \lambda. \quad (4.4)$$

It is easy to see that Problems (4.1) and (4.3) are equivalent in the following sense: the optimal values for both problems are equal and an assortment is optimal for problem (4.1) if and only if it maximizes the left-hand-side of (4.4) when $\lambda$ is equal to the maximal
value. Hence the value of problem (4.3) is the maximal expected profit that retailer \(n\) can achieve given its competitor’s assortment, \(A_m\). Fix \(\lambda \in \mathbb{R}\) and let \(S^\lambda(A_m)\) denote the set of optimal solutions to the maximization on the left-hand-side of (4.4). One could solve (4.3) (and hence (4.1)) by computing \(S^\lambda(A_m)\) for all possible values of \(\lambda\), and then selecting any assortment in \(S^\lambda(A_m)\), where \(\lambda^*\) corresponds to the highest value of \(\lambda\) for which \(\pi_n(a, p_n, A_m, p_m) \geq \lambda\) holds, for \(a\) in \(S^\lambda(A_m)\). This observation will prove useful in the equilibrium analysis we conduct for this setting, as well as throughout the rest of the chapter. We now outline how to compute \(S^\lambda(A_m)\). To that end, for \(i \in \mathcal{N}_n\), we define

\[
\theta_i(\lambda) := \begin{cases} 
(p_{n,i} - c_i - \lambda) \nu_i(p_{n,i}) & \text{if } i \not\in \mathcal{N}_n \setminus A_m, \\
(\delta_{n,i}(p_n, p_m)(p_{n,i} - c_i) - \lambda) \nu_i(\min\{p_{n,i}, p_{m,i}\}) + \lambda \nu_i(p_{m,i}) & \text{otherwise}. 
\end{cases}
\]

(4.5)

Given this, the maximization problem in (4.4) can be rewritten as

\[
\max_{A \in \mathcal{A}_n} \left\{ \sum_{i \in A} \theta_i(\lambda) \right\} \geq \lambda \left(1 + \sum_{i \in A_m} \nu_i(p_{m,i})\right).
\]

Now, one can solve for \(S^\lambda(A_m)\) by ordering the products in \(\mathcal{N}_n\) according to the corresponding values of \(\theta_i(\lambda)\), from highest to lowest, and selecting the maximum number of products in the assortment (up to \(C_n\)) with positive values of \(\theta_i(\lambda)\).

**Product ranking.** Note that the product ranking according to the \(\theta_i\)'s (and hence the selected assortment) will vary depending on the value of \(\lambda\) and on which products are included in the competitor’s assortment \(A_m\). This last observation implies that, for a fixed value of \(\lambda\), a product that is not “appealing” (i.e., a product that is not included in a best response) if not offered by the competitor might become appealing when the latter offers it. This can be seen from (4.5); indeed, if, for example, retailer \(n\) offers the lowest price for product \(i\), \(\theta_i(\lambda)\) increases by a factor of \(\lambda \nu_i(p_{m,i})\) when product \(i\) is offered by retailer \(m\). This gain can be interpreted as the value of profiting from product \(i\) without having to expand the consideration set of customers. In addition, as already highlighted for a monopoly in Rusmevichientong et al. (2010) such a ranking for an optimal value of \(\lambda\) need not coincide with the ranking of the profit margins.
4.4.1 The case of exclusive products

This section studies the case of retailers having only exclusive products, i.e., $N_1 \cap N_2 = \emptyset$. We begin by specializing the best response computation to this setting, and then study equilibrium behavior.

**Best response properties.** In this setting, since retailers cannot offer common products, (4.3) can be written as

$$\max \left\{ \lambda \in \mathbb{R} : \max_{A \in S_n} \left\{ \sum_{i \in A} (p_{n,i} - c_i - \lambda) \nu_i(p_{n,i}) \right\} \geq \lambda \left( 1 + \sum_{i \in A_m} \nu_i(p_m) \right) \right\}.$$  \hfill (4.6)

Note that the inner maximization in (4.6) does not depend on the competitor’s assortment, $A_m$, and hence neither does the collection $\{ S^\lambda(A_m) : \lambda \in \mathbb{R} \}$. In addition, the solution to (4.6) (and hence retailer $n$’s profit) depends on $A_m$ only through $\sum_{i \in A_m} \nu_i(p_{m,i})$. For $n = 1, 2$, and $A$ in $S_n$, we can express the attractiveness of assortment $A$ as follows

$$E_n(A) := \sum_{i \in A} \nu_i(p_{n,i}).$$

For given assortment offerings, $A_1$ and $A_2$, this quantity is related to the market share of retailer $n$ as the latter is given by $E_n(A_n)/(1 + E_1(A_1) + E_2(A_2))$. Define $\lambda_n(e)$ to be the retailer $n$’s expected profit per customer when retailer $m$ offers assortment $A_m$ with attractiveness $e$. That is

$$\lambda_n(e) := \max \left\{ \lambda \in \mathbb{R} : \max_{A \in S_n} \left\{ \sum_{i \in A} (p_{n,i} - c_i - \lambda) \nu_i(p_{n,i}) \right\} \geq \lambda(1 + e) \right\}.$$  \hfill (4.7)

Similarly, let $a_n(e)$ denote retailer $n$’s set of best responses to any assortment with attractiveness $e$, i.e.,

$$a_n(e) := \arg\max_{A \in S_n} \left\{ \sum_{i \in A} (p_{n,i} - c_i - \lambda_n(e)) \nu_i(p_{n,i}) \right\}.$$  \hfill (4.7)

The next result establishes monotonicity properties of the best response correspondence in terms of attractiveness and profit level.

**Proposition 1** (best response properties). Suppose that $N_1 \cap N_2 = \emptyset$.

i.) Retailer $n$’s best response profit is decreasing in the attractiveness of its competitor’s assortment, i.e., $\lambda_n(e)$ is decreasing in $e$. 

ii.) The attractiveness of retailer $n$’s best response assortments is non-decreasing in the attractiveness of the competitor’s assortment, $e$, in the following sense: for any $e > e' \geq 0$,

$$\max \{ E_n(a) : a \in a_n(e') \} \leq \min \{ E_n(a) : a \in a_n(e) \}.$$ 

Proposition 1 i.) states that a retailer’s (optimized) profits will decrease if the competitor increases the attractiveness/breadth of its offerings, which is in line with intuition. Proposition 1 ii.) provides an important qualitative insight: if one retailer increases the attractiveness of the products it is offering, then so will the other one. In particular, it implies that if one compares a retailer with capacity $C_1$ operating in a duopoly with a monopolist with the same capacity, then the one operating in a duopoly will offer a set of products with higher attractiveness than the monopolist (whose assortment attractiveness corresponds to the attractiveness of the best response of the retailer operating in a duopoly when the competitor does not offer any product). As we will see next, this result plays a key role in characterizing the set of equilibria in this scenario.

The conclusions of Proposition 1 are usually obtained in the context of supermodular games. However, it is worth noting that, in general, it is not clear whether one could obtain a supermodular representation of the assortment game with the exception of the case where margins are equal across products. (In the latter case, the game can be seen to be log-supermodular on the discrete lattice of possible attractiveness levels induced by all feasible assortments.) However, once properties outlined in Proposition 1 are at hand, one can establish existence and ordering of equilibria in a similar fashion as is usually performed for supermodular games, which we do next.

**Equilibrium behavior.** The following result guarantees that an equilibrium exists.

**Theorem 7** (equilibrium existence). Suppose that $N_1 \cap N_2 = \emptyset$. Then there always exists an equilibrium in assortment decisions.

Theorem 7 establishes the existence of an equilibrium, but leaves open the possibility of having multiple equilibria. While there might indeed exist multiple equilibria, we show next that if such a case occurs, both retailers will prefer the same equilibrium.
Proposition 2 (best equilibrium). Suppose that $\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$ and that multiple equilibria exist. Then, among these equilibria, retailer $n$ receives maximal profits in the one that minimizes the attractiveness $E_n(A_n)$, $n = 1, 2$. Moreover, both retailers prefer the same equilibrium.

In other words, when multiple equilibria exist, retailers would prefer to select the one with the least breadth of offerings. The result is a direct consequence of the relationship between profit level and attractiveness of the offering, established in Proposition 1. We note that Proposition 1, Theorem 7 and Proposition 2 can be generalized to an arbitrary number of retailers and we briefly indicate how one might do so in the proofs of those results.

We next focus on the question of uniqueness of an equilibrium and more generally the number of equilibria that may exist. As a prelude to that analysis, we first consider the following example.

Example 1. Consider a setting where $\mathcal{N}_1 = \{1, 2, 3\}$, $\mathcal{N}_2 = \{4, 5, 6\}$, and $\alpha = 1$, $p_{n,i} = 1$ for $n = 1, 2$, $\mu_i = 1$, for $i = 1, \ldots, 6$. Setting $C_1 = C_2 = 1$ and $c_i = 0$ for $i = 1, \ldots, 6$, one gets that any pair $(i_1, i_2)$ in $\mathcal{N}_1 \times \mathcal{N}_2$ constitutes an equilibrium in assortment decisions. However, in each equilibrium $(A_1, A_2)$, one has $E_n(A_n) = 1$ and $\pi_n(A_n, p_n, A_m, p_m) = 1/3$, for $n = 1, 2$. This situation stems from the fact that retailer $n$ is indifferent between the competitor offering $A_m$ or $A'_m$, provided that $E_m(A_m) = E_m(A'_m)$.

Hence, one can have situations in which while there may be many equilibria, some of these share very important properties, which motivates the following definition.

Definition 1 (equivalent equilibria for assortment competition). We say that two equilibria, $(A_1, A_2)$ and $(A'_1, A'_2)$ are equivalent if $E_n(A_n) = E_n(A'_n)$, for $n = 1, 2$. Otherwise, we say that two equilibria are fundamentally different.

Note that two equilibria are equivalent if each retailer offers assortments with the same attractiveness level and hence, each retailer collects the same profits in both of them.

A bound on the number of equilibria. We will restrict attention to fundamentally different equilibria. These are generated by best response assortments with different attractiveness levels. Let $\mathcal{E}_n$ denote the set of attractiveness levels of all possible best response assortments.
assortments offered by retailer \( n \), i.e.,

\[
E_n := \{ E_n(a) \in \mathbb{R}_+ : a \in a_n(e), e \geq 0 \}, \quad \text{for } n = 1, 2.
\]

The next proposition introduces a bound on the number of fundamentally different equilibria, based on the cardinality of these sets.

**Theorem 8** (bound on the number of equilibria). Suppose that \( N_1 \cap N_2 = \emptyset \). Then there are at most \(|E_1| + |E_2| - 1\) fundamentally different equilibria.

A priori, a trivial bound on the number of fundamentally different equilibria is the number of combinations of best response attractiveness levels, \(|E_1||E_2|\). Theorem 8 provides a significantly sharper bound. The proof of Theorem 8 relies on the strong monotonicity property established in Proposition II.), which enables one to eliminate a large number of candidates in the set of possible equilibria.

**Corollary 3** (Sufficient condition for a unique equilibrium). Suppose that all products offered by a given retailer have the same margin, i.e., \( p_{n,i} - c_i = r_n \) for all \( i \in N_n \), \( n = 1, 2 \), where \( r_1 \) and \( r_2 \) are given positive constants. Then, all equilibria are equivalent.

This result follows from the fact that when margins are equal across products, \(|E_n| = 1\) for \( n = 1, 2 \). In particular, \( \{a_n(e) : e \in \mathbb{R}_+\} = \mathcal{P}_n(N_n, C_n) \), where for \( n = 1, 2 \) and a set of products \( Z \subseteq N_n \) one defines

\[
\mathcal{P}_n(Z, C) := \arg\max_{A \subseteq Z, |A| \leq C} \left\{ \sum_{i \in A} \nu_i(p_{n,i}) \right\}. \tag{4.8}
\]

\( \mathcal{P}_n(Z, C) \) will be referred to as the \( C \)-popular sets of products and a set in \( \mathcal{P}_n(Z, C) \) is formed by \( \min\{C, |Z|\} \) products in \( Z \) with the highest attraction factors. The above, in conjunction with Theorem 8 yields that all equilibria are equivalent (i.e., there is a unique equilibrium in the fundamentally different sense).

We illustrate below that the bound of Theorem 8 can be applied in other important cases and also show that the bound cannot be improved upon in general.

**Case of monotonic margins.** Suppose that a higher attraction factor is synonymous with a higher profit margin, i.e., that \( \nu_i(p_{n,i}) \geq \nu_j(p_{n,j}) \) if and only if \( p_{n,i} - c_i \geq p_{n,j} - c_j \), for all \( i, j \in N_n \), \( n = 1, 2 \). Recalling the definition of the \( \theta_i \)'s in (4.5) and of \( a_n(e) \) in (4.7), for a
given $e \in \mathbb{R}$, one has that assortments in $a_n(e)$ consist of some subset of the products with highest $\theta_i(\lambda(e))$, which implies that

$$\{a_n(e) : e \in \mathbb{R}_+\} \subseteq \{P_n(N_n, C) : 1 \leq C \leq C_n\}. $$

The latter set of assortments, is commonly referred to as the set of popular assortments; see, for example, van Ryzin and Mahajan (1999). Hence, one has that $|E_n| \leq C_n$ for $n = 1, 2$ and, by Theorem 8, there exist at most $C_1 + C_2 - 1$ fundamentally different equilibria.

There are other cases of interest that can be analyzed. One of these was recently studied in Rusmevichientong et al. (2010) in the context of assortment planning for a monopolist. There, a number of assumptions (detailed below) are introduced with the purpose of limiting the cardinality of $E_n$ by limiting the number of different product rankings (based on the $\theta_i(\lambda)$’s) one could obtain for all values of $\lambda$.

Case of distinct customer preferences and intersection points. For $n = 1, 2$, for $i < j$ in $N_n$, define

$$\mathcal{I}_n(i, j) := \frac{(p_{n,i} - c_i)\nu_i(p_{n,i}) - (p_{n,j} - c_j)\nu_j(p_{n,j})}{\nu_i(p_{n,i}) - \nu_j(p_{n,j})}. $$

Suppose products are such that $\nu_i(p_{n,i}) \neq \nu_j(p_{n,j})$ for $i \neq j$, and that $\mathcal{I}_n(i, j) \neq \mathcal{I}_n(i', j')$ for any $(i, j) \neq (i', j')$, for $n = 1, 2$. Under this assumption, when one varies $\lambda$, the $\theta_i(\lambda)$-based ranking of products changes by swapping at most two consecutively ranked products. This limits the number of possible assortments in $a_n(e)$ for any $e$, which limits the cardinality of $E_n$. Under such assumptions, Rusmevichientong et al. (2010) Theorem 2.5 show that $|E_n| \leq C_n(|N_n| - C_n + 1)$, hence the number of fundamentally different equilibria is bounded by $C_1(|N_1| - C_1 + 1) + C_2(|N_2| - C_2 + 1) - 1$.

Theorem 8 provides an upper bound on the number of equilibria one may sustain. The next example shows that such a bound is tight for the case of monotonic margins.

**Example 2** (matching bound). Consider the case where $\nu_i(p_{n,i}) \neq \nu_j(p_{n,j})$ for all $i, j$ in $P_n(N_n, C_n)$, $n = 1, 2$. Let $i_n = P_n(N_n, 1)$, for $n = 1, 2$. Suppose that for all $j$ in $N_n \setminus \{i_n\}$, $p_{n,j} - c_j = r_n$, where

$$r_1 := \frac{(p_{1,i_1} - c_{i_1})\nu_{1,i_1}(p_{1,i_1})}{1 + \nu_{1,i_1}(p_{1,i_1}) + E_2(P_2(N_2, C_2))} \quad \text{and} \quad r_2 := \frac{(p_{2,i_1} - c_{i_2})\nu_{2,i_2}(p_{2,i_2})}{1 + \nu_{2,i_2}(p_{2,i_2}) + E_1(P_1(N_1, 1))}.$$
CHAPTER 4. PRODUCT ASSORTMENT AND PRICE COMPETITION WITH INFORMED CONSUMERS

In other words, for each retailer, all products except the one with the highest attraction factor have the same profit margin, and the latter is strictly lower than that of the former. This is an instance of a problem with monotonic margins. In such a case, as shown above, the set of all possible best response attractiveness levels for retailer \( n \), \( \mathcal{E}_n \), is exactly the set of popular assortments.

Given the construction of profit margins above, it is possible to show that when retailer 1 selects the assortment \( \mathcal{P}_1(\mathcal{N}_1, 1) \), retailer 2 is indifferent between all the popular assortments, i.e., \( \mathcal{B}_2(\mathcal{P}_1(\mathcal{N}_1, 1)) = \{\mathcal{P}_2(\mathcal{N}_2, C) : 1 \leq C \leq C_2\} \). Similarly, when retailer 2 selects the largest popular assortment \( \mathcal{P}_2(\mathcal{N}_2, C_2) \), retailer 1 is indifferent between all popular assortments, i.e., \( \mathcal{B}_1(\mathcal{P}_2(\mathcal{N}_2, C_2)) = \{\mathcal{P}_1(\mathcal{N}_1, C) : 1 \leq C \leq C_1\} \). This, in conjunction with the increasing property of the best response attractiveness in the competitor’s attractiveness (see Proposition 1\( \Box \)) implies that

\[(\mathcal{P}_1(\mathcal{N}_1, 1), A), A \in \{\mathcal{P}_2(\mathcal{N}_2, C) : 1 \leq C \leq C_2\} \text{ and } (A, \mathcal{P}_2(\mathcal{N}_2, C_2)), A \in \{\mathcal{P}_1(\mathcal{N}_1, C) : 1 \leq C \leq C_1\},\]

are all fundamentally different equilibria. Hence, the number of fundamentally different equilibria is given by \(|\mathcal{E}_1| + |\mathcal{E}_2| - 1 = C_1 + C_2 - 1\), which matches the bound in Theorem 8.

\[\Box\]

Computing the set of fundamentally different equilibria. We have not addressed the issue of computing the equilibria so far. It can be shown that, in general, one can build on efficient procedures available to compute optimal assortments of a monopoly (see, e.g., Rusmevichientong \textit{et al.} (2010)) in order to compute all fundamentally different equilibria in an efficient manner. In particular, it is possible to construct a superset of all possible best responses \( \{\mathcal{S}^\lambda(A_m) : A_m \in \mathcal{A}_m, \lambda > 0\} \) which would yield a set of candidate equilibria. In turn, one may conduct an efficient search across all candidate equilibria by using the monotonicity properties established in Proposition 1.

4.4.2 The case of both exclusive and common products

We now turn to the case when retailers may offer the same products in their respective assortments, i.e., when \( \mathcal{N}_1 \cap \mathcal{N}_2 \) is not empty. We provide sufficient conditions for existence of an equilibrium but show that in general the key structural results derived in Section
4.4.1 will fail to hold. In particular, we show that i.) an equilibrium may fail to exist; and
ii.) the set of fundamentally different equilibria might have a number of elements growing
exponentially with $C_1$ and $C_2$, even in the special case of equal margins for all products.

**Example 3** (non-existence of equilibrium). Consider a setting with two retailers, each
having access to the three same products $\mathcal{N}_1 = \mathcal{N}_2 = \{1, 2, 3\}$, and with display capacities
$C_1 = 2$, $C_2 = 1$. Suppose that prices and costs are uniform across products and retailers
and given by $p_{n,1} = p_{n,2} = p_{n,3} = p > 1$ and $c_1 = c_2 = c_3 = p - 1$ for $n = 1, 2$, and that
the mean utilities are such that $\nu_1(p) = 1.1, \nu_2(p) = 1.01, \nu_3(p) = 1$. Table 4.1 depicts the
rewards for each retailer for feasible pairs of assortment decisions $(A_1, A_2)$. There, each
entry corresponds the profit of retailer 1 and retailer 2 (in that order).

<table>
<thead>
<tr>
<th>$A_1$</th>
<th>{1}</th>
<th>{2}</th>
<th>{3}</th>
<th>{1,2}</th>
<th>{1,3}</th>
<th>{2,3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>{1}</td>
<td>(0.262, 0.262)</td>
<td>(0.354, 0.325)</td>
<td>(0.355, 0.323)</td>
<td>(0.502, 0.177)</td>
<td>(0.500, 0.177)</td>
<td>(0.489, 0.268)</td>
</tr>
<tr>
<td>{2}</td>
<td>(0.325, 0.354)</td>
<td>(0.251, 0.251)</td>
<td>(0.336, 0.332)</td>
<td>(0.516, 0.162)</td>
<td>(0.511, 0.246)</td>
<td>(0.500, 0.168)</td>
</tr>
<tr>
<td>{3}</td>
<td>(0.323, 0.355)</td>
<td>(0.332, 0.336)</td>
<td>(0.250, 0.250)</td>
<td>(0.513, 0.243)</td>
<td>(0.516, 0.161)</td>
<td>(0.502, 0.166)</td>
</tr>
</tbody>
</table>

Table 4.1: Illustration of non-existence of equilibrium in the setup of Example 3.

One can verify that no equilibrium exists. Intuitively, the latter stems from the fact that
retailer 1, with a capacity of 2, will always prefer to incorporate in its assortment the product
that retailer 2 is offering, while retailer 2 prefers to offer an exclusive product. Recalling
the discussion following the definition of $\theta_i(\lambda)$ in (4.5), the current example illustrates how
a product gains in appeal (measured by $\theta_i$) when offered by the competitor. In this setting,
this prevents the possibility of an equilibrium.

While the previous example shows that one may not ensure in general the existence
of an equilibrium in assortment decisions, the next result provides sufficient conditions for
existence of an equilibrium, exists in this scenario.

**Proposition 3** (sufficient conditions for equilibrium existence). Suppose that one of the
following two conditions is satisfied:
CHAPTER 4. PRODUCT ASSORTMENT AND PRICE COMPETITION WITH INFORMED CONSUMERS

i.) $N_1 = N_2 = N$, $C_1 = C_2 = C$ and $p_i - c_i = r$ for all $i \in N$.

ii.) $C_n \geq |S_n|$, $n = 1, 2$, and $p_{1,i} = p_{2,i}$ for all $i \in N$.

Then, an equilibrium in assortment decisions is guaranteed to exist.

Condition i.) essentially says that as long firms have only common products, and those have the same profit margin, an equilibrium is guaranteed to exist. So the combination of exclusive and common products, in conjunction with different product margins seems to be a driver of the potential failure for an equilibrium to exist. Condition ii.) ensures that both retailers operate without any capacity constraints. In such a case, it is always possible to construct an equilibrium in which both retailers offer all common products. Indeed, when one retailer does so, it is always optimal for the competing retailer to also do so and given this, the equilibrium analysis boils down to the selection of exclusive products, for which an equilibrium is guaranteed to exist by the analysis of Section 4.4.1.

In the cases where an equilibrium is guaranteed to exist, the implications of the introduction of common products on the number of equilibria still remains. The example below, for which condition i.) in Proposition 3 holds, shows that the number of equilibria may be exponential in the capacities of the retailers.

**Example 4** (exponential number of equilibria). Consider the following setup. Suppose that the set of available products is common to both retailers ($N_1 = N_2$) and has $S = 2C$ elements, where $C_1 = C_2 = C$, that all products are priced at the same uniform price $p$, that their marginal cost is zero, and that the $\mu_i$’s are ordered so that

$$\nu_1(p) > \nu_2(p) > \ldots > \nu_S(p), \quad \text{and} \quad \nu_1(p) < \frac{3}{2}\nu_S(p).$$

In addition, suppose that

$$\frac{\sum_{i=1}^{C} \nu_i(p)}{1 + \sum_{i=1}^{C} \nu_i(p)} \leq \frac{1}{4}.$$ 

This condition corresponds to assuming that the maximum share any retailer can achieve (under any scenario) is below 25%. Under the setup above, we show that if retailer 2 offers an arbitrary selection of products $A_2$, then the best response of retailer 1 is to offer the
set of $C$ products with the highest $\nu_i$’s in $\mathcal{N}_1 \setminus A_2$, i.e., to offer the remaining $C$ products. Recalling (4.3), retailer 1 solves for the maximal $\lambda$ such that

$$
\max_{A_1 \subseteq \mathcal{N}} \left\{ \sum_{j \in A_1 \setminus A_2} (p - \lambda)\nu_j(p) + \sum_{j \in A_1 \cap A_2} \left( \frac{p}{2} - \lambda \right)\nu_j(p) - \lambda \sum_{j \in A_2 \setminus A_1} \nu_j(p) \right\} \geq \lambda.
$$

Given any assortment offered by retailer 2, $A_2$, the revenues of retailer 1, $\lambda$, are bounded by the revenues of a monopolist (with display capacity $C$), i.e., $\lambda \leq p(\sum_{i=1}^{C} \nu_i(p))(1 + \sum_{i=1}^{C} \nu_i(p))^{-1}$. This, in conjunction with the market share condition above, implies that $\lambda \leq p/4$.

For any products $j$ in $\mathcal{N}_1 \setminus A_2$ and $j'$ in $A_2$, given that $\nu_{j'}(p) < (3/2)\nu_j(p)$, it will always be the case that $\nu_j(p)(p - \lambda) \geq (3/4)\nu_j(p)p > \nu_{j'}(p) p/2$. Hence $\theta_j(\lambda) > \theta_{j'}(\lambda)$ and the best response of retailer 1 to $A_2$ will never include any product in $A_2$. In addition, since $\lambda \leq p/4$, retailer 1 will always include the remaining $C$ products in $\mathcal{N}_1$.

Given the above, it is possible to verify that any pair of assortments $(A_1, A_2)$ that belongs to the set

$$
\mathcal{EQ} := \{ A_1 \subseteq \mathcal{N}, A_2 = \mathcal{N} \setminus A_1 : |A_1| = C \}
$$

is an equilibrium. It is also possible to see that one can choose the $\nu_i(p)$’s so that all equilibria are fundamentally different. In particular, the cardinality of $\mathcal{EQ}$ is given by $\binom{2C}{C}$. This illustrates that in general, even when prices are uniform, the number of fundamentally different equilibria may be exponential in the capacities of the retailers in contrast with what was observed in the case of exclusive products (see Corollary 3).

It is also worth noting that each firm will prefer a different equilibrium in this example (each retailer prefers the equilibrium where s/he offers the products with the highest attraction factors), in stark contrast with the case of exclusive products, where both firms always prefer the same equilibrium (see Proposition 2).

$\square$

### 4.5 Joint Assortment and Price Competition: Main Results

We now turn attention to the case where in addition to assortment decisions retailers also set prices for the products they offer. We follow a parallel exposition to that of Section 4.4.
CHAPTER 4. PRODUCT ASSORTMENT AND PRICE COMPETITION WITH INFORMED CONSUMERS

by first studying the best response correspondence before separating the analysis for the case of exclusive products and that of both exclusive and common products. Without loss of generality, we assume throughout this section that prices are restricted to be greater or equal than $c_i$ for any product $i$.

**Best response correspondence.** Fix an assortment $A_m$ in $S_m$ and corresponding price vector $p_m$. Retailer $n$’s best response to $(A_m, p_m)$ consists of the assortment-price pairs solving Problem (4.2). We will establish that problem (4.2) is equivalent, in some sense, to a problem where the relevant decision is reduced to the retailer’s assortment. For that purpose, define for any assortment $A$ in $S_n$ and any $\lambda > 0$,

$$p_{n,i}^*(A, \lambda, A_m, p_m) := \begin{cases} \frac{1}{\alpha} + c_i + \lambda & \text{if } i \in A \setminus A_m, \\ \min \{\frac{1}{\alpha} + c_i + \lambda, p_{m,i}\} & \text{if } i \in A \cap A_m. \end{cases} \quad (4.9)$$

Loosely speaking, $p_{n,i}^*$ is the optimal price for product $i$ when assortment $A$ is offered and retailer $n$’s profit is $\lambda^2$. Throughout this section we let

$$\tilde{\nu}_i := \frac{1}{\alpha} e^{\mu_i - (c_i + \alpha + 1)}, \text{ for } i \in N.$$ 

As in the previous section, for $n = 1, 2$ and a set of products $Z \subseteq N_n$, we define sets of $C$-popular products as follows

$$\tilde{P}_n(Z, C) := \arg\max_{A \subseteq Z : |A| \leq C} \left\{ \sum_{i \in A} \tilde{\nu}_i \right\},$$

i.e., a set in $\tilde{P}_n(Z, C)$ is formed by $C$ products in $Z$ with the highest value of $\tilde{\nu}_i$. Note that the definition of the $C$-popular sets matches the one provided in (4.8) when prices are set as in (4.9) and no product is offered by both retailers simultaneously.

More formally, any sequence $\{p_{n,i}^k : k \geq 1\}$ such that $p_{n,i}^k < p_{n,i}^*$ for all $i$ in $A$ and $k \geq 1$ and $p_{n,i}^k \to p_{n,i}^*$ yields profits that converge to the supremum in (4.2) as $k \to \infty$ and $A_n$ is fixed to $A$. (4.9) can be seen to be an expression of “equal margins” across offered products, with the modifications to account for the possibility of common products. Such a property has previously appeared in various related settings; see, e.g., Anderson et al. (1992).
Consider the following problem

\[
\max \quad \lambda \\
\text{s.t.} \quad \max_{A \in S_n} \left\{ \sum_{i \in A \setminus A_m} \tilde{\nu}_i e^{-\lambda \alpha} + \sum_{i \in A \cap A_m} \left( p^*_n,i - c_i - \lambda \right) \nu_i(p^*_n,i) - \lambda \sum_{i \in A \setminus A_m} \nu_i(p_{m,i}) \right\} \geq \lambda,
\]

(4.11)

**Lemma 3.** Problems (4.10) and (4.11) are equivalent in the following sense: the optimal values for both problems are equal and an assortment is optimal for problem (4.10) if and only if it maximizes the left-hand-side of (4.11) when \( \lambda \) is equal to the supremum value, \( \lambda^* \). Moreover, for any maximizing assortment \( a \) in \( S_n \), for any sequence \( \{p^k : k \geq 1\} \) of prices such that \( p^k_{n,i} < p^*_n(a, \lambda^*, A_m, p_m) \) for all \( i \) in \( a \) and \( p^k_n \to p^*_n(a, \lambda^*, A_m, p_m) \),

\[
\lim_{k \to \infty} \pi_n(a, p^k, A_m, p_m) = \lambda^*.
\]

Hence, one has that the optimal solution to (4.10) and the corresponding assortments that maximize the left-hand-side of (4.11) represent the optimal profit and the optimal assortments, respectively.\(^3\) When offering any maximizing assortment \( a \) in \( S_n \) one can get arbitrarily close to achieve the profit level of \( \lambda^* \) by pricing products in \( a \) just below \( p^*_n(\lambda^*, a, A_m, p_m) \). Fix \( \lambda \in \mathbb{R} \) and let \( S^\lambda(A_m, p_m) \) be the solution set to the inner maximization in (4.11). For \( i \in N_n \) define

\[
\theta_i(\lambda) := \begin{cases} 
\tilde{\nu}_i e^{-\lambda \alpha} & \text{if } i \in N_n \setminus A_m, \\
\nu_i(p^*_n,i)(p^*_n,i - c_i - \lambda) + \lambda \nu_i(p_{m,i}) & \text{if } i \in N_n \cap A_m.
\end{cases}
\]

(4.12)

As in the case of assortment competition (see the discussion immediately following (4.3)), one can solve for \( S^\lambda(A_m, p_m) \) by selecting the products with highest positive values of \( \theta_i(\lambda) \). From (4.12) one observes that \( \theta_i(\lambda) \) is always non-negative, for \( i \in N_n \). As a result, one can show that elements of \( S^\lambda(A_m, p_m) \) are necessarily full-capacity assortments, for any \( \lambda > 0 \). This is formalized in the next result.

**Lemma 4.** Both retailers will always offer full-capacity assortments in an equilibrium.

---

\(^3\)One notes that when retailers offer a product \( i \) in common, optimal profits are not achieved unless \( p_{n,i} = p_{m,i} = c_i \) due to the discontinuity of profits when \( p_{n,i} = p_{m,i} > c_i \).
In other words, a retailer will always benefit from including an additional product into its assortment. This contrasts with the case of assortment-only competition where retailers may not want to offer the maximal number of products. Hence, the ability to modify prices enables retailer \( n \) to always have \( C_n \) products that can be included in an optimal assortment.

4.5.1 The case of exclusive products

This section studies the case of retailers having only exclusive products, i.e., \( \mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset \).

**Best response correspondence.** In this setting, problem (4.10) can be specialized and rewritten as

\[
\max \left\{ \lambda \in \mathbb{R} : \max_{A \in \mathcal{S}_n} \left\{ \sum_{i \in A} \tilde{\nu}_i e^{-\lambda \alpha} \right\} \geq \lambda \left( 1 + \sum_{i \in A_m} \nu_i(p_{m,i}) \right) \right\}. \tag{4.13}
\]

Note that the inner maximization in (4.13) does not depend on \((A_m, p_m)\), hence neither does the collection of all best response candidates \( \{S^\lambda : \lambda \in \mathbb{R}\} \) (hence we suppress its dependence on \( A_m \) and \( p_m \)). From (4.12) one has that \( \{S^\lambda : \lambda \in \mathbb{R}\} = \mathcal{P}_n(\mathcal{N}_n, C_n) \), i.e., the assortment best response is always formed by the \( C_n \) most attractive products (in terms of the \( \tilde{\nu}_i \)'s), independent of the competitor’s decisions, for \( n = 1, 2 \). Note that, for a given \( \lambda > 0 \), all \( C_n \)-popular sets of \( \mathcal{N}_n \) will provide the same attractiveness. This justifies the following definition for \( n = 1, 2 \).

\[
V_n := \sum_{i \in a} \tilde{\nu}_i,
\]

where \( a \) is any assortment in \( \mathcal{P}_n(\mathcal{N}_n, C_n) \). With this, (4.13) reduces to

\[
\max \left\{ \lambda \in \mathbb{R} : V_n e^{-\lambda \alpha} \geq \lambda \left( 1 + \sum_{i \in A_m} \nu_i(p_{m,i}) \right) \right\}. \tag{4.14}
\]

With the optimal value to the problem above, \( \lambda^* \), one can compute optimal prices for any equilibrium assortment \( a \) in \( \mathcal{P}_n(\mathcal{N}_n, C_n) \) through \( p_{n,i}^* = 1/\alpha + c_i + \lambda^* \).

**Equilibrium behavior.** From (4.14) above, one observes that all assortments in \( \mathcal{P}_n(\mathcal{N}_n, C_n) \) are essentially equivalent for purposes of equilibrium computation. The next definition is the equivalent of Definition 1 adapted to the current setting.

**Definition 2** (equivalent equilibria for joint assortment and price competition). *We say that two equilibria, \((A_1, p_1, A_2, p_2)\) and \((A'_1, p'_1, A'_2, p'_2)\) are equivalent if for \( n = 1, 2 \), \( \sum_{i \in A_n} \tilde{\nu}_i = \sum_{i \in A'_n} \tilde{\nu}_i \), \( V_n = V'_n \), and \( p_{n,i}^* = p'_{n,i}^* \) for all \( i \in A_n \).*
\[ \sum_{i \in A_n} \tilde{\nu}_i. \]

Otherwise, we say that two equilibria are fundamentally different.

As in section 4.4, retailers offer assortments with the same attractiveness level and collect the same profits in equivalent equilibria. The next proposition establishes existence and uniqueness of an equilibrium in this setting.

**Proposition 4** (equilibrium existence and uniqueness). Suppose that \( N_1 \cap N_2 = \emptyset \). Then, there always exists an equilibrium. Moreover, all equilibria are equivalent, and characterized by retailer \( n \) offering any assortment \( a \) in \( \tilde{P}_n(N_n, C_n) \), and charging \( p_{n,i} = \frac{1}{\alpha} + c_i + \lambda_n \), for \( i \in a \), where \( (\lambda_1, \lambda_2) \) is the unique solution to the system of equations

\[
V_n e^{-\lambda_n \alpha} = \lambda_n (1 + \alpha V_n e^{-\lambda_m \alpha}), \quad n = 1, 2. \tag{4.15}
\]

For the existence result, one can use similar monotonicity properties as those established in Proposition 1. Uniqueness (in the fundamentally different sense) follows from the fact all assortments in \( \tilde{P}_n(N_n, C_n) \) have the same attractiveness, in equilibrium. Recall from Theorem 7 and Corollary 3 that there is a unique equilibrium (in the fundamentally different sense) when margins are equal across products offered by a given retailer and exogenously fixed. In the current setting, a similar conclusion is obtained; however, it is worth noting that while margins are necessarily equal across products offered by a given retailer, the margins are endogenously determined. We note that Proposition 4 can be generalized to an arbitrary number of retailers and we indicate how one may do so in the proof. We also observe that \( (\lambda_1, \lambda_2) \), the solution to (4.15), is such that \( \lambda_1 \geq \lambda_2 \) if and only if \( V_1 \geq V_2 \), i.e., the retailer with the broader assortment (better products in terms of \( V_n \)) will achieve the highest profit, in equilibrium.

The analysis of best responses in this setting has already appeared in the literature (see, e.g., Misra (2008)). The result above complements such an analysis by establishing existence and uniqueness of an equilibrium, but also illustrates along the way the general applicability of the framework we use.

---

Note that this is also true when retailers offer overlapping product sets. This is a consequence of Lemma 5 that appears in Section 4.5.2.
4.5.2 The case of both exclusive and common products

We now consider the case where retailers have overlapping product sets, i.e., \( \mathcal{N}_1 \cap \mathcal{N}_2 \neq \emptyset \). We first establish properties that any equilibrium must satisfy, which will limit the set of candidates when searching for equilibria.

**Lemma 5.** Suppose \( |\mathcal{N}_n| > C_n \) for some \( n = 1, 2 \), and that \((A_1, p_1, A_2, p_2)\) is an equilibrium. Then \( A_1 \cap A_2 = \emptyset \). Moreover, it is necessarily the case that at least one retailer offers an exclusive assortment, i.e., \( A_n \cap \mathcal{N}_m = \emptyset \) for some \( n = 1, 2 \).

Lemma 5 states that, as long as a retailer has more products to select from than what s/he can offer, no product will be offered by both retailers simultaneously in equilibrium. This stems from the fact that if a product is offered by both retailers in equilibrium, then it is necessarily priced at cost and the above retailer will always benefit by substituting it for a different product. At an intuitive level, the second part of the Lemma is a consequence of the fact that if both retailers want to offer common products, then both of them will want to offer the same common products, which is not possible in equilibrium. Hence, the only possibility is that a single retailer selects products from the pool of common products. Note that is a consequence of the intense price competition the retailers engage in for common products. As soon as such competition is softened, then retailers may offer the same products in equilibrium. The latter may occur if prices are fixed as in Section 4.4 or when a minimum price is imposed by manufacturers as discussed in Section 4.6.

The result above motivates the following definition. For \((n, k)\) in \( \{1, 2\} \times \{1, 2\} \), we let

\[
\mathcal{N}^k_n := \begin{cases} 
\mathcal{N}_n \setminus \mathcal{N}_k & \text{if } n \neq k, \\
\mathcal{N}_n & \text{otherwise}.
\end{cases}
\]

Equilibria in which only retailer \( n \) offers products in \( \mathcal{N}_1 \cap \mathcal{N}_2 \) must also be equilibria when retailer \( n \) is restricted to offer products in \( \mathcal{N}^n_n \) and retailer \( m \) is restricted to offer products in \( \mathcal{N}^m_m \). Since these scenarios correspond to cases of exclusive products, Proposition 4 ensures that all equilibria are equivalent in the latter case. Also, we know that such equilibria can be computed solely based on the knowledge of the following quantities,

\[
V^k_n := \sum_{i \in \tilde{\nu}} \tilde{\nu}_i, \quad n = 1, 2, \ k = 1, 2,
\]
where \( a \) corresponds to any assortment in \( \tilde{P}_n(\lambda_n^k, C_n) \). For \( k = 1, 2 \), let \((\lambda_1^k, \lambda_2^k)\) be the unique solution to the system of equations

\[
V_n^k e^{-\lambda_n^k \alpha} = \lambda_n^k \left( 1 + \alpha V_m^k e^{-\lambda_m^k \alpha} \right), \quad n = 1, 2,
\]

and define the set of 4-tuples of type \( k \) as

\[
CE_k := \left\{ (A_1, p_1, A_2, p_2) : A_n \in \tilde{P}_n(N_n^k, C_n), p_{n,i} = c_i + \lambda_n^k + \frac{1}{\alpha}, \ i \in A_n, \ n = 1, 2 \right\}.
\]

\( CE_k \) contains all the equivalent equilibria associated with the case where retailer 1 and 2 are restricted to select products in \( S_1^k \) and \( S_2^k \), respectively. Lemma 5 in conjunction with the above discussion implies that in the absence of such restrictions, the set of equilibria is contained in \( CE_1 \cup CE_2 \) as long as \(|N_n| > C_n\) for some \( n = 1, 2 \). The next result provides a characterization of equilibrium behavior.

**Theorem 9** (equilibrium existence and uniqueness). All equilibria are necessarily equivalent. In addition, we have the following characterization.

i.) Suppose \( C_n \leq |S_n \setminus S_m| \) for at least one \( n \) in \( \{1, 2\} \). Then, all candidates in \( CE_1 \cap CE_2 \) are equilibria and only one of \( CE_1 \setminus CE_2 \) or \( CE_2 \setminus CE_1 \) can contain an equilibrium.

ii.) Suppose \( C_n > |S_n \setminus S_m| \) for \( n = 1, 2 \).

a) If \( C_n = |S_n| \) for \( n = 1, 2 \), then there exists a unique equilibrium.

b) If \( C_n < |S_n| \) for some \( n \) in \( \{1, 2\} \), then there is no equilibrium.

The first part of Theorem 9 describes a setting in which at least one retailer has an ample set of exclusive products to select from (i.e., \( C_n \leq |S_n \setminus S_m| \)). When both retailers prefer to offer exclusive assortments rather than offer common products, then \( CE_1 \cap CE_2 \) is necessarily non-empty and an equilibrium always exists. When common products have the potential to increase the profit of a retailer, it is not possible to guarantee existence of an equilibrium, however, one can show that the set of candidate equilibria can be restricted to at most one of \( CE_1 \setminus CE_2 \) or \( CE_2 \setminus CE_1 \). We provide in Proposition 5 and the discussion that follows a procedure to first identify which set may contain an equilibrium and second to verify whether an equilibrium exists.
The second part of Theorem 9 establishes existence and uniqueness results in setting where no retailer has an ample set of exclusive products to select from, i.e., no retailer can offer a full-capacity and exclusive assortment. In particular, an equilibrium will only exist if retailers do not have any option with regard to what products to offer (case ii.b) when selecting a full-capacity assortment. (Recall that retailers necessarily offer full-capacity assortments in equilibrium (see Lemma 4).) In such a case, it is established in the proof of Theorem 9 that the unique equilibrium is such that each retailer offers all products and prices product \( i \in N_1 \cap N_2 \) at \( p_{n,i} = c_i \) and product \( i \in N_n \setminus N_m \) at \( p_{n,i} = c_i + 1/\alpha + \lambda_n \), where \((\lambda_1, \lambda_2)\) is the unique solution to

\[
\sum_{i \in N_n \setminus N_m} \bar{\nu}_i e^{-\lambda_n \alpha} = \lambda_n \left( 1 + \sum_{i \in N_n \cap N_m} \alpha \bar{\nu}_i e + \alpha \sum_{i \in N_m \setminus N_n} \bar{\nu}_i e^{-\lambda_m \alpha} \right), \quad n = 1, 2.
\]

Now, if retailers have some flexibility when selecting a full-capacity assortment (case ii.b)), then no equilibrium can be sustained. Indeed, in any equilibrium, at least one product would need to be offered by both retailers (since \( C_n > |S_n \setminus S_m| \) for \( n = 1, 2 \)). However, each retailer would want to deviate from such a scenario by replacing a product offered by both retailers (which is necessarily offered at its marginal cost in equilibrium) by either another common product that s/he is not currently offering (and charging a slightly lower price than the competitor), or by an exclusive product that is not currently being offered.

It is worth noting that the absence of equilibrium is driven by the intense price competition the retailers engage in for common products, and when such competition is softened then some candidate equilibria may become sustainable (see, e.g., Proposition 3 for an illustration when prices are fixed).

**Checking equilibrium existence.** While Theorem 9 solves the issue of equilibrium uniqueness, existence is guaranteed only if \( \mathcal{CE}_1 \cap \mathcal{CE}_2 \neq \emptyset \). When the latter condition fails to hold one is left with the task of checking which of \( \mathcal{CE}_1 \setminus \mathcal{CE}_2 \) or \( \mathcal{CE}_2 \setminus \mathcal{CE}_1 \) contains equilibria, if any. The next proposition provides a procedure to eliminate one of the candidate sets. Before stating the result, we note that it will always be the case that \( V_n^m \geq V_m^m \) for some \( n \) in \( \{1, 2\} \).

**Proposition 5 (elimination of equilibrium candidates).** Take \( n \) in \( \{1, 2\} \) such that \( V_n^m \geq V_m^m \). If \( a \cap N_n \neq \emptyset \) for some \( a \) in \( \bar{\mathcal{P}}_m(N_m, C_m) \), then no candidate in \( \mathcal{CE}_n \setminus \mathcal{CE}_m \) is an
equilibrium. Otherwise, no candidate in $C \varepsilon_m \setminus C \varepsilon_n$ is an equilibrium.

Note that the procedure above does not require the computation of actual equilibria in the candidate sets.

Suppose that $C \varepsilon_n \setminus C \varepsilon_m$ is a valid candidate for containing equilibria, after checking the condition above. Note that it must be the case that retailer $n$ offers common products in any element of the candidate set. To check whether all elements of $C \varepsilon_n \setminus C \varepsilon_m$ are equilibria or not, it is sufficient to evaluate the impact of a single deviation starting from a single candidate equilibrium. One needs to evaluate the impact of retailer $m$ offering the most attractive common product instead of the least attractive exclusive product currently offered. More formally, if $C \varepsilon_n \setminus C \varepsilon_m$ remains a valid candidate for containing equilibria, any of its elements is an equilibrium if and only if

$$
\tilde{\nu}_j \left( \alpha \min \{\lambda^m_n - \lambda^m_m, 0\} + 1 \right) e^{-\alpha \min\{\lambda^m_n, \lambda^m_m\}} + \alpha \lambda^m_n \tilde{\nu}_j e^{-\alpha \lambda^m_n} \leq \tilde{\nu}_{\hat{j}} e^{-\alpha \lambda^m_n}
$$

where $\hat{j}$ is the index of the most attractive product in $N_1 \cap N_2$ (i.e., $\hat{j} = \arg\max\{\tilde{\nu}_i : i \in N_1 \cap N_2\}$) and $\hat{j}_m$ is the index of the least attractive product in any assortment $A_m$ offered by retailer $m$ in any candidate equilibrium (i.e., $\hat{j}_m = \arg\min\{\tilde{\nu}_i : i \in A_m\}$). The condition in (4.16) follows from comparing $\theta_j(\lambda^m_n)$ and $\theta_{\hat{j}}(\lambda^m_m)$ and noting that $\theta_j(\lambda^m_n) \geq \theta_i(\lambda^m_n)$ for all $i \in N_m \setminus A_m$ and $\theta_{\hat{j}}(\lambda^m_m) \leq \theta_i(\lambda^m_m)$ for all $i \in A_m$. Such a comparison determines whether $A^m_n$ maximizes the left-hand-side of (4.11).

4.6 Extensions and Concluding Remarks

This chapter has studied equilibrium behavior for retailers competing in assortment and pricing. For both the cases of assortment-only and joint assortment and pricing, a crisp characterization of equilibrium properties has been presented, highlighting the role that the presence of common products may have on such properties. The framework outlined to analyze equilibrium properties can be used to study other extensions of the setting presented, as we illustrate next.

Minimum profit margins. In the assortment and price competition setting that preceded, no constraints have been imposed on the prices charged by the retailers. This introduced
the possibility of intense Bertrand-like price competition and was a key driver of the result that ensured that at most one equilibrium existed (in the fundamentally different sense) and that firms may not offer the same products in equilibrium.

Here, we discuss an important extension where retailers must collect a minimum margin from the sale of some products. It is often the case that minimal prices are imposed directly or indirectly by manufacturers through, e.g., a Manufacturer’s Suggested Retail Price (MSRP). We highlight some of the implications of such constraints on the results of Section 4.5. A summary of the main steps that could be taken to adapt part of the earlier analysis to the current setting is presented in Appendix C.3.

At a qualitative level, the two extreme cases where minimum margins are relatively small or high are very informative with regard to the type of modifications to expect. Recall from (4.9) that whenever a product is offered and its price is selected in an unconstrained fashion, its profit margin is always greater than $1/\alpha$. Hence, if the minimum margins are small (lower than $1/\alpha$), then the results for joint price and assortment competition of Section 4.5 would not be affected. In contrast, when the minimum margins are very high, it will always be optimal for the retailers to price at the minimum margin, and hence the retailers would essentially be facing fixed prices. In such a case, the retailers’ joint price and assortment problem would reduce to an assortment-only problem, and the analysis and results of Section 4.4 would hold. Hence, in general, the analysis of the minimum profit margins case will build on both Sections 4.4 and 4.5.

Overall, the above discussion, in conjunction with that appearing in Appendix C.3, illustrates that, in general, multiple price and assortment equilibria can be sustained in the presence of minimal profit margins and those may enable retailers to operate at equilibria where some products are present in both assortments.

**Extensions to an arbitrary number of retailers.** The analysis has focused on a duopoly of retailers. It is worth noting that many of the results can be generalized to the case of an arbitrary number of retailers. In particular, for the case of assortment-only competition with exclusive products, Proposition 1, Theorem 7 and Proposition 2 admit parallel results with an arbitrary number of retailers. For the case of assortment and price competition with exclusive products, the conditions that an equilibrium must satisfy laid out in Proposition
4 can also be generalized to the case of an arbitrary number of retailers and existence and uniqueness can also be established. For the case of both exclusive and common products, the number of possible configurations becomes difficult to summarize, however, one observes that it is possible to exhibit sufficient conditions in particular configurations of interest such as in Proposition 3 ii.) to ensure existence of an equilibrium.

The case of sequential competition. While Section 4.5 has focused on the case of simultaneous assortment and price competition, the approach taken in the current work can also be used to analyze the case where firms compete first on assortments and then on prices. In such a setting, one can solve for equilibria in a backward fashion and establish that there always exists an equilibrium (for both the cases of exclusive and common products) and one can provide a sufficient and necessary condition for the uniqueness of such an equilibrium, which essentially says that the set of $C_n$-popular products for each retailer should not intersect. In particular, this condition is satisfied for the case of exclusive products and the unique equilibrium coincides with the unique equilibrium when retailers select actions simultaneously.

Additional generalizations and challenges. In the case of assortment-only competition, the analysis only relied on the attraction form of the demand model. However, the particular Logit assumption played a key role in the joint assortment and pricing analysis. It is worth noting the challenges one faces under more general models. For example, under mixed Logit demand, while under some conditions, for a given assortment, existence and uniqueness of an equilibrium in prices can be guaranteed (see Allon et al. (2010)), the assortment-only problem becomes intractable, as highlighted in Rusmevichientong et al. (2009) where it is shown that the monopolist’s problem is in general NP-Hard, even in the absence of display capacities.
Bibliography


Appendix A

Proof of Results in Chapter 2

A.1 Proof of Main Results

Proof of Theorem 1. The lower bound is trivial when $\mathcal{N} = S^*(\mu)$, so assume $S^*(\mu) \subset \mathcal{N}$. For $i \in \mathcal{N}$ define $T_i(t)$ as the number of customers product $i$ has been offered to, before customer $t$’s arrival,

$$T_i(t) := \sum_{u=1}^{t-1} 1 \{ i \in S_u \} , \ t \geq 1.$$ 

Similarly, for $n \geq 1$ define $t_i(n)$ as the customer to whom product $i$ is offered for the $n$-th time,

$$t_i(n) := \inf \{ t \geq 1 : T_i(t + 1) = n \} , \ n \geq 1.$$ 

For $i \in \mathcal{N} \setminus S^*(\mu)$, define $\Theta_i$ as the set of mean utility vectors for which product $i$ is in the optimal assortment, but that differs from $\mu$ only on its $i$-th coordinate. That is,

$$\Theta_i := \{ \nu \in \mathbb{R}^N : \nu_i \neq \mu_i , \nu_j = \mu_j \ \forall \ j \in \mathcal{N} \setminus \{ i \} , \ i \in S^*(\nu) \} .$$

We will use $E_\nu^\pi$ and $P_\nu^\pi$ to denote expectations and probabilities of random variables, when the assortment policy $\pi \in \mathcal{P}$ is used, and the mean utilities are given by the vector $\nu$. Let $I_i(\mu \parallel \nu)$ denote the Kullback-Leibler divergence between $F(\cdot - \mu_i)$ and $F(\cdot - \nu_i)$,

$$I_i(\mu \parallel \nu) := \int_{-\infty}^{\infty} \left[ \log \left( \frac{dF(x - \mu_i)}{dF(x - \nu_i)} \right) \right] dF(x - \mu_i).$$

This quantity measures the “distance” between $P_\mu^\pi$ and $P_\nu^\pi$. We have that $0 < I_i(\mu \parallel \nu) < \infty$ for all $\nu \neq \mu, i \in \mathcal{N} \setminus S^*(\mu)$. Fix $i \in \mathcal{N}$ and consider a configuration $\nu \in \Theta_i$. For $n \geq 1$
define the log-likelihood function
\[ L_i(n) := \sum_{u=1}^{n} \left[ \log(dF(U_{it}^{t_i(u)} - \mu_i)/dF(U_{it}^{t_i(u)} - \nu_i)) \right]. \]

Note that \( L_i(\cdot) \) is defined in terms of utility realizations that are unobservable to the retailer. Define \( \delta(\eta) \) as the minimum (relative) optimality gap when the mean utility vector is given by \( \eta \in \mathbb{R}^N \),
\[ \delta(\eta) := \inf \left\{ 1 - r(S, \eta)/r(S^*(\eta), \eta) > 0 : S \in \mathcal{S} \right\}. \] (A.1)

Fix \( \alpha \in (0, 1) \). For any consistent policy \( \pi \) one has that for any \( \epsilon > 0, \)
\[ R^\pi(T, \nu) \geq \delta(\nu) \mathbb{E}^\nu_{\pi} \{ T - T_i(T) \} \geq \delta(\nu) \left( T - \frac{(1 - \epsilon)}{I_i(\mu||\nu)} \log T \right) \mathbb{P}^\nu_{\pi} \{ T_i(T) < (1 - \epsilon) \log T / I_i(\mu||\nu) \}, \]
and by assumption on \( \pi \) \( R^\pi(T, \nu) = o(T^\alpha) \). From the above, we have that
\[ \mathbb{P}^\nu_{\pi} \{ T_i(T) < (1 - \epsilon) \log T / I_i(\mu||\nu) \} = o(T^{\alpha - 1}). \] (A.2)

Define the event
\[ \beta_i := \left\{ T_i(T) \leq \frac{(1 - \epsilon)}{I_i(\mu||\nu)} \log T, \ L_i(T_i(T)) \leq (1 - \alpha) \log T \right\}. \]
From the independence of utilities across products and the definition of \( \beta_i \), we have that
\[ \mathbb{P}^\nu_{\pi} \{ \beta_i \} = \int_{\omega \in \beta_i} d\mathbb{P}^\nu_{\pi} \]
\[ = \int_{\omega \in \beta_i} \prod_{u=1}^{T-1} \prod_{i \in S_u} dF(U_{it}^u - \nu_i) \]
\[ = \int_{\omega \in \beta_i} \prod_{u=1}^{T-1} \prod_{i \in S_u} dF(U_{it}^u - \nu_i) d\mathbb{P}^\mu_{\pi} \]
\[ = \int_{\omega \in \beta_i} \exp(-L_i(T_i(T))) d\mathbb{P}^\mu_{\pi} \]
\[ \geq \exp(-(1 - \alpha) \log T) \mathbb{P}^\mu_{\pi} \{ \beta_i \}. \]
From (A.2) one has that \( \mathbb{P}^\nu_{\pi} \{ \beta_i \} = o(T^{\alpha - 1}) \). It follows by (A.2) that as \( T \to \infty \)
\[ \mathbb{P}^\mu_{\pi} \{ \beta_i \} \leq \mathbb{P}^\nu_{\pi} \{ \beta_i \} / T^{\alpha - 1} \to 0. \] (A.3)
Indexed by \( n \), \( \mathcal{L}_i(n) \) is the sum of finite mean identically distributed independent random variables, therefore, by the strong law of large numbers (SLLN).

\[
\limsup_{n \to \infty} \frac{\max \{ \mathcal{L}_i(l) : l \leq n \}}{n} \leq \mathcal{I}_i(\mu \parallel \nu) \frac{1}{1 - \alpha} \, \text{a.s.,}
\]

i.e., the log-likelihood function grows no faster than linearly with slope \( \mathcal{I}_i(\mu \parallel \nu) . \) This implies that

\[
\limsup_{n \to \infty} \mathbb{P}_\pi^n \{ \exists l \leq n, \mathcal{L}_i(l) > n \mathcal{I}_i(\mu \parallel \nu) / (1 - \epsilon) \} = 0.
\]

In particular,

\[
\lim_{T \to \infty} \mathbb{P}_\pi^n \left\{ T_i(T) < \frac{(1 - \epsilon)}{\mathcal{I}_i(\mu \parallel \nu)} \log T , \mathcal{L}_i(T_i(T)) > \frac{(1 - \epsilon)}{1 - \alpha} \log T \right\} = 0.
\]

Taking \( \alpha < \epsilon \) small enough, and combining with (A.3) one has that

\[
\lim_{T \to \infty} \mathbb{P}_\pi^n \left\{ T_i(T) > (1 - \epsilon) \log T / H_i^\mu \right\} = 1.
\]

By Markov’s inequality, and letting \( \epsilon \) shrink to zero we get

\[
\liminf_{T \to \infty} \mathbb{E}_\pi^n \left\{ T_i(T) \right\} \log T \geq \frac{1}{H_i^\mu}.
\]  

By the definition of the regret, we have that for any consistent policy \( \pi \in \mathcal{P}' \),

\[
\mathcal{R}_\pi^n (T, \mu) \geq (a) \delta(\mu) \mathbb{E}_\pi^n \left[ \sum_{t=1}^{T} \mathbb{P}_\pi^n 1 \{ S_t \neq S^*(\mu) \} \right] \geq (b) \delta(\mu) \frac{1}{C} \sum_{i \in \mathcal{N} \setminus S^*(\mu)} \mathbb{E}_\pi^n [T_i(T)] .
\]

where \((a)\) follows from the non-optimal assortments contributing at least \( \delta(\mu) \) to the regret, and \((b)\) follows by assuming non-optimal products are always tested in batches of size \( C \), discarding products in \( \mathcal{N} \). Thus

\[
\sum_{u=1}^{T} 1 \{ S_u \neq S^*(\mu) \} \geq \sum_{u=1}^{T} 1 \{ S_u \cap \mathcal{N} \setminus S^*(\mu) \neq \emptyset \} \geq \frac{1}{C} \sum_{i \in \mathcal{N} \setminus S^*(\mu)} \sum_{u=1}^{T} 1 \{ i \in S_u \} = \frac{1}{C} \sum_{i \in \mathcal{N} \setminus S^*(\mu)} T_i(T).
\]
Combining the above with (A.4) we have that, asymptotically,
\[ R^\pi(T, \mu) \geq \frac{\delta(\mu)}{C} \left( \sum_{i \in \mathcal{N} \setminus S^*(\mu)} \frac{1}{H_i^\mu} \right) \log T. \]

Taking \( K := \delta(\mu) \min_{i \in \mathcal{N} \setminus S^*(\mu)} \{ (H_i^\mu)^{-1} \} \) gives the desired result.

**Proof of Theorem 2**. We prove the result in 3 steps. First, we compute an upper bound on the probability of the estimates deviating from the true mean utilities. Second, we address the quality of the solution to the static problem, when using estimated mean utilities. Finally, we combine the above and analyze the regret. For purposes of this proof, let \( P \) denote probability of random variables when the assortment policy \( \pi_1 \) is used, and the mean utilities are given by the vector \( \mu \). With a slight abuse of notation define \( p_i := \{ p_i(A_j, \mu) : A_j \in \mathcal{A} \text{ s.t. } i \in A_j \} \), for \( i \in \mathcal{N} \), and \( p := (p_1, \ldots, p_N) \).

**Step 1.** Define \( T^j(t) \) to be the number of customers \( A_j \) has been offered to, up to customer \( t - 1 \), for \( A_j \in \mathcal{A} \). That is,
\[ T^j(t) = \sum_{u=1}^{t-1} 1 \{ S_u = A_j \}, \quad j = 1, \ldots, |\mathcal{A}|. \]

We will need the following side lemma, whose proof is deferred to Appendix A.2.

**Lemma 6.** Fix \( j \leq |\mathcal{A}| \) and \( i \in A_j \). Then, for any \( n \geq 1 \) and \( \epsilon > 0 \)
\[ \mathbb{P} \left\{ \left| \sum_{u=1}^{t-1} (Z_i^u - p_i(A_j, \mu)) 1 \{ S_u = A_j \} \right| \geq \epsilon T^j(t), \quad T^j(t) \geq n \right\} \leq 2 \exp(-c(\epsilon)n), \]
for a positive constant \( c(\epsilon) < \infty \).

For any vector \( \nu \in \mathbb{R}^N \) and set \( A \subseteq \mathcal{N} \) define \( \|\nu\|_A = \max \{ \nu_i : i \in A \} \). Consider \( \epsilon > 0 \) and fix \( t \geq 1 \). By Assumption 1 we have that for any assortment \( A_j \subseteq \mathcal{A} \)
\[ \|\mu - \hat{\mu}_t\|_{A_j} \leq \kappa(\epsilon)\|p - \hat{p}_t\|_{A_j}, \quad (A.5) \]
for some constant $1 < \kappa(\epsilon) < \infty$, whenever $\|p - \hat{p}_t\|_{A_j} < \epsilon$. We have that, for $n \geq 1,$

$$P \left\{ \|\mu - \hat{\mu}_t\|_{A_j} > \epsilon, T^j(t) \geq n \right\} = P \left\{ \|\mu - \hat{\mu}_t\|_{A_j} > \epsilon, \|p - \hat{p}_t\|_{A_j} \geq \epsilon, T^j(t) \geq n \right\} + P \left\{ \|\mu - \hat{\mu}_t\|_{A_j} > \epsilon, \|p - \hat{p}_t\|_{A_j} < \epsilon, T^j(t) \geq n \right\} \leq P \left\{ \|p - \hat{p}_t\|_{A_j} \geq \epsilon, T^j(t) \geq n \right\} + P \left\{ \|\mu - \hat{\mu}_t\|_{A_j} > \epsilon, \|p - \hat{p}_t\|_{A_j} < \epsilon, T^j(t) \geq n \right\} \leq 2P \left\{ \|p - \hat{p}_t\|_{A_j} \geq \epsilon/\kappa(\epsilon), T^j(t) \geq n \right\} \leq 2 \sum_{i \in A_j} P \left\{ |p_i(A_j, \mu) - \hat{p}_{i,t}| \geq \epsilon/\kappa(\epsilon), T^j(t) \geq n \right\} \leq 2 |A_j| \exp(-c(\epsilon/\kappa(\epsilon))n), \tag{A.8}$$

where (a) follows from [A.5], (b) follows from the definition of $\hat{p}_{i,t}$, and (c) follows from Lemma 6.

**Step 2.** Fix an assortment $S \in \mathcal{S}$. By the Lipschitz-continuity of $p(S, \cdot)$ we have that, for $t \geq 1, \max \{|p_i(S, \mu) - p_i(S, \hat{\mu}_t)| : i \in S\} \leq K\|\mu - \hat{\mu}_t\|_S,$ for a positive constant $K < \infty$, and therefore

$$|r(S, \mu) - r(S, \hat{\mu}_t)| \leq \|w\|_{\infty} K C \|\mu - \hat{\mu}_t\|_S. \tag{A.9}$$

From here, we conclude that

$$r(S^*(\hat{\mu}_t), \mu) \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t) - \|w\|_{\infty} K C \|\mu - \hat{\mu}_t\|_{S^*(\hat{\mu}_t)} \geq r(S^*(\mu), \hat{\mu}_t) - \|w\|_{\infty} K C \|\mu - \hat{\mu}_t\|_{S^*(\hat{\mu}_t)} \geq r(S^*(\mu), \mu) - 2\|w\|_{\infty} K C \|\mu - \hat{\mu}_t\|_{(S^*(\mu) \cup S^*(\hat{\mu}_t))}.$$

As a consequence, if

$$\|\mu - \hat{\mu}_t\|_{(S^*(\mu) \cup S^*(\hat{\mu}_t))} < (2\|w\|_{\infty} K C)^{-1} \delta(\mu) r(S^*(\mu), \mu)$$
then $S^*(\mu) = S^*(\hat{\mu}_t)$, where $\delta(\mu)$ is the minimum (relative) optimality gap (see (A.1) in proof of Theorem 1). This means that if the mean utility estimates are uniformly close to the underlying mean utility values, then solving the static problem using estimates returns the same optimal assortment as when solving the static problem with the true parameters.

In particular we will use the following relation:

$$\{S^*(\mu) \neq S^*(\hat{\mu}_t)\} \subseteq \{\|\mu - \hat{\mu}_t\|_{(S^*(\mu) \cup S^*(\hat{\mu}_t))} \geq (2\|w\|_{\infty}KC)^{-1}\delta(\mu)r(S^*(\mu), \mu)\}. \quad (A.10)$$

**Step 3.** Let $NO(t)$ denote the event that a non-optimal assortment is offered to customer $t$. That is

$$NO(t) := \{S_t \neq S^*(\mu)\},$$

Define $\xi := (2\|w\|_{\infty}KC)^{-1}\delta(\mu)r(S^*(\mu), \mu)$. For $t \geq |A|\lceil \kappa_1 \log T \rceil$ one has that

$$\mathbb{P}\{NO(t)\} \leq a \sum_{A_j \in A} \mathbb{P}\{\|\mu - \hat{\mu}_t\|_{A_j} \geq \xi\} \leq b \sum_{A_j \in A} \mathbb{P}\{\|\mu - \hat{\mu}_t\|_{A_j} \geq \xi, T^j(t) \geq \kappa_1 \log T\} \leq (a) \sum_{A_j \in A} 2|A_j|T^{-\kappa_1 c(\xi/\kappa\xi)}, \quad (A.11)$$

were (a) follows from (A.10) and (b) follows from (A.8). Considering $\kappa_1 > c(\xi/\kappa\xi)^{-1}$ results in the following bound for the regret:

$$R^\pi(T, \mu) \leq \sum_{t=1}^T \mathbb{P}\{NO(t)\} \leq |A|\lceil \kappa_1 \log T \rceil + \sum_{t>|A|\lceil \kappa_1 \log T \rceil}^{\infty} 2|A_j|T^{-\kappa_1 c(\xi/\kappa\xi)} \leq |A|\kappa_1 \log T + K_1 = [N/C] \kappa_1 \log T + K_1,$$

for a finite constant $K_1$. Setting $K_1 = c(\xi/\kappa\xi)^{-1}$ gives the desired result. 

**Proof of Theorem 3.** The proof follows the arguments of the proof of Theorem 2. Steps 1 and 2 are identical.
Step 3. Let $NO(t)$ denote the event that a non-optimal assortment is offered to customer $t$, and $G(t)$ the event that there is no forced testing for customer $t$. That is,

$$
NO(t) := \{ S_t \neq S^*(\mu) \},
$$

$$
G(t) := \{ T^j(t) \geq \kappa_2 \log t, \ j \leq |A| \text{ such that } \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t) \} . \tag{A.12}
$$

Define $\xi := (2\|w\|_{\infty} KC)^{-1}\delta(\mu)r(S^*(\mu), \mu)$. We have that

$$
\mathbb{P}\{ NO(t), G(t) \} \overset{(a)}{=} \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{(S^*(\mu) \cup S^*(\hat{\mu}_t))} > \xi, G(t) \} \leq \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{S^*(\mu)} > \xi, G(t) \} + \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{S^*(\hat{\mu}_t)} > \xi, G(t) \} \overset{(b)}{=} \sum_{j : A_j \cap S^*(\hat{\mu}_t) \neq \emptyset} \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{A_j} > \xi, T^j(t) > \kappa_2 \log t \} + \sum_{j : A_j \cap S^*(\mu) \neq \emptyset} \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{A_j} > \xi, G(t) \} \overset{(c)}{=} \sum_{j : A_j \cap S^*(\mu) \neq \emptyset} 2|A_j| t^{-c(\xi/\kappa(\xi))\kappa_2} + \sum_{j : A_j \cap S^*(\hat{\mu}_t) \neq \emptyset} \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{A_j} > \xi, G(t) \},
$$

where: (a) follows from (A.10); (b) follows from the fact that Assumption 2 guarantees $w_i \geq r(S^*(\nu), \nu)$ for all $i \in S^*(\nu)$ for any vector $\nu \in \mathbb{R}^N$; and (c) follows from (A.8).

Fix $j$ such that $A_j \cap \bar{S}^*(\mu) \neq \emptyset$. For such an assortment we have that

$$
\mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{A_j} > \xi, G(t) \} \leq \mathbb{P}\{ \|\mu - \hat{\mu}_t\|_{A_j} > \xi, T^j(t) \geq \kappa_2 \log t, G(t) \} + \mathbb{P}\{ T^j(t) < \kappa_2 \log t, G(t) \} .
$$

The first term on the right-hand-side above can be bounded using (A.8). For the second one, note that $\{ T^j(t) < \kappa_2 \log t, G(t) \} \subseteq \{ \|w\|_{A_j} < r(S^*(\hat{\mu}_t), \hat{\mu}_t) \}$, and that

$$
\|w\|_{A_j} - r(S^*(\mu), \mu)\delta(\mu)/2 \overset{(a)}{=} r(S^*(\mu), \mu)(1 - \delta(\mu)/2) \overset{(b)}{=} r(S^*(\hat{\mu}_t), \mu) \overset{(c)}{=} r(S^*(\hat{\mu}_t), \hat{\mu}_t) - \|w\|_{\infty} KC \|\mu - \hat{\mu}_t\|_{S^*(\hat{\mu}_t)},
$$

where: (a) follows from Assumption 2; (b) follows from the definition of $\delta(\mu)$; and (c) follows...
The above implies that \( \{ \| w \|_{A_j} < r(S^*(\bar{\mu}_t), \bar{\mu}_t) \} \subseteq \{ \| \mu - \bar{\mu}_t \|_{S^*(\bar{\mu}_t)} > \xi \} \), i.e.,

\[
\mathbb{P} \{ T^j(t) < \kappa_2 \log t, G(t) \} \leq \mathbb{P} \{ \| w \|_{A_j} < r(S^*(\bar{\mu}_t), \bar{\mu}_t) \} \\
\leq \mathbb{P} \{ \| \mu - \bar{\mu}_t \|_{S^*(\bar{\mu}_t)} > \xi, G(t) \} \\
\leq \sum_{k: A_k \cap S^*(\bar{\mu}_t) \neq \emptyset} \mathbb{P} \{ \| \mu - \bar{\mu}_t \|_{A_k} > \xi, G(t) \} \\
= \sum_{j: A_j \cap \bar{S}^*(\bar{\mu}_t) \neq \emptyset} \left( 2 |A_j| t^{-c(\xi/\kappa)\kappa_2} + \sum_{k: A_k \cap \bar{S}^*(\bar{\mu}_t) \neq \emptyset} 2 |A_k| t^{-c(\xi/\kappa)\kappa_2} \right)
\]

where the last step follows from (A.8). Using the above we have that

\[
\mathbb{P} \{ NO(t) \wedge G(t) \} \leq \sum_{j: A_j \cap \bar{S}^*(\bar{\mu}_t) \neq \emptyset} 2 |A_j| t^{-c(\xi/\kappa)\kappa_2} + \sum_{j: A_j \cap S^*(\mu) \neq \emptyset} \left( 2 |A_j| t^{-c(\xi/\kappa)\kappa_2} + \sum_{k: A_k \cap S^*(\bar{\mu}_t) \neq \emptyset} 2 |A_k| t^{-c(\xi/\kappa)\kappa_2} \right)
\]

Consequently, we have that

\[
\mathbb{P} \{ NO(t) \wedge G(t)^c \} \leq \sum_{A_j \in \mathcal{A}} \mathbb{P} \{ S_t = A_j, G(t)^c \}
\]

For the first term above, we have from the policy specification that

\[
\sum_{u=1}^{T} \sum_{j: \| w \|_{A_j} \geq r(S^*(\mu), \mu)} \mathbb{P} \{ S_u = A_j, G(u)^c \} \leq \frac{|\mathcal{A}|}{C} (\kappa_2 \log T + 1).
\]

To analyze the second term, fix \( j \) such that \( \| w \|_{A_j} < r(S^*(\mu), \mu) \), and define \( L(t) \) as the last customer (previous to customer \( t \)) to whom the empirical optimal assortment (according to estimated mean utilities) was offered. That is

\[
L(t) := \sup \{ u \leq t - 1 : G(u) \},
\]

with \( G(u) \) given in (A.12). Note that \( L(t) \in \{ t - \lfloor |\mathcal{A}| \kappa_2 \log t \rfloor, \ldots, t - 1 \} \) for \( t \geq \tau \), where \( \tau \) is given by

\[
\tau := \inf \{ u \geq 1 : \log(u - \lfloor |\mathcal{A}| \kappa_2 \log u \rfloor) + \kappa_2^{-1} > \log u \}.
\]
Consider \( t \geq \tau \) and \( u \in \{ t - \lfloor |A| \kappa_2 \log t \rfloor, \ldots, t - 1 \} \). Then

\[
\mathbb{P} \{ S_t = A_j, G(t)^c, L(t) = u \} \leq \mathbb{P} \{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t), G(t)^c, L(t) = u \}
\]

\[
\leq \mathbb{P} \{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t), G(t)^c, G(u) \}
\]

\[
= \mathbb{P} \{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t), G(t)^c, G(u), NO(u) \} + \mathbb{P} \{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t), G(t)^c, G(u), NO(u)^c \}
\]

\[
\leq \mathbb{P} \{ NO(u), G(u) \} + \mathbb{P} \{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t), T^k(t) \geq \kappa_2 \log t, \forall k \text{ s.t. } A_k \cap S^*(\mu) \neq \emptyset \},
\]

where the last step follows from the fact that offering \( S^*(\mu) \) to customer \( u \) implies (from \( G(u) \)) that \( T^j(u) \geq \kappa_2 \log u \), and therefore (for \( t \geq \tau \)) that \( T^j(t) \geq \kappa_2 \log t \), for all \( j \) such that \( A_j \cap S^*(\mu) \neq \emptyset \). The first term in the last inequality can be bounded using (A.13). For the second, observe that

\[
r(S^*(\mu), \hat{\mu}_t) - \|w\|_{A_j} \geq r(S^*(\mu), \mu) - \|w\|_{\infty} KC \|\mu - \hat{\mu}_t\|_{S^*(\mu)} - \|w\|_{A_j},
\]

which follows from (A.9). Define \( \delta := \inf \{ (\|w\|_{\infty} KC)^{-1} (1 - \|w\|_{A_j}/r(S^*(\mu), \mu)) > 0 : A_j \in \mathcal{A} \} \). From the above,

\[
\{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t) \} \subseteq \{ \|w\|_{A_j} \geq r(S^*(\mu), \hat{\mu}_t) \} \subseteq \{ \|\mu - \hat{\mu}_t\|_{S^*(\mu)} > \delta r(S^*(\mu), \mu) \}.
\]

Define the event \( \Xi = \{ \|w\|_{A_j} \geq r(S^*(\hat{\mu}_t), \hat{\mu}_t), T^k(t) \geq \kappa_2 \log t, \forall k \text{ s.t. } A_k \cap S^*(\mu) \neq \emptyset \} \) and \( \delta := \delta r(S^*(\mu), \mu) \). One has that

\[
\mathbb{P} \{ \Xi \} \leq \mathbb{P} \{ \|\mu - \hat{\mu}_t\|_{S^*(\mu)} > \delta, T^k(t) \geq \kappa_2 \log t, \forall k \text{ s.t. } A_k \cap S^*(\mu) \neq \emptyset \}
\]

\[
\leq \sum_{k: A_k \cap S^*(\mu) \neq \emptyset} \mathbb{P} \{ \|\mu - \hat{\mu}_t\|_{A_k} > \delta, T^k(t) \geq \kappa_2 \log t \}
\]

\[
\leq \sum_{k: A_k \cap S^*(\mu) \neq \emptyset} 2|A_k| t^{-c(\delta/\kappa(\delta))\kappa_2}.
\]

Using Lemma[2], we have that, when \( \kappa_2 > c(\delta/\kappa(\delta)) \),

\[
\mathbb{P} \{ S_t = A_j, G(t)^c, L(t) = u \} \leq C^2(2 + C) u^{-c(\xi/\kappa(\xi))\kappa_2} + \sum_{k: A_k \cap S^*(\mu) \neq \emptyset} 2|A_k| t^{-c(\delta/\kappa(\delta))\kappa_2}
\]

\[
\leq C^2(2 + C) \left( t - \lfloor |A| \kappa_2 \log t \rfloor \right)^{-c(\xi/\kappa(\xi))\kappa_2} + C^2 t^{-c(\delta/\kappa(\delta))\kappa_2}.
\]
Since the right hand side above is independent of \( u \), one has that

\[
\mathbb{P}\{S_t = A_j, G(t)^c\} \leq C^2(2 + C)\left(t - |A|\kappa_2 \log t\right)^{-c(\xi/\kappa(\xi))\kappa_2} + C^2t^{-c(\delta/\kappa(\delta))\kappa_2},
\]  

(A.15)

for \( j \) such that \( \|w\|_{A_j} < r(S^*(\mu), \mu) \), and \( t \geq \tau \). Considering \( \kappa_2 > \max\{c(\xi/\kappa(\xi))^{-1}, c(\delta/\kappa(\delta))^{-1}\} \) results in the following bound for the regret:

\[
\mathcal{R}^\pi(T, \mu) \leq \sum_{t=1}^{T} \mathbb{P}\{NO(t), G(t)\} + \sum_{t=1}^{T} \mathbb{P}\{NO(t), G(t)^c\}
\]

\[
\leq \sum_{t=1}^{T} \mathbb{P}\{NO(t), G(t)\} + \sum_{t=1}^{T} \sum_{j: \|w\|_{A_j} \geq r(S^*(\mu), \mu)} \mathbb{P}\{S_t = A_j, G(t)^c\} + \sum_{t=1}^{T} \sum_{j: \|w\|_{A_j} < r(S^*(\mu), \mu)} \mathbb{P}\{S_t = A_j, G(t)^c\}
\]

\[
\leq \sum_{t=1}^{\infty} C^2(2 + C)u^{-c(\xi/\kappa(\xi))\kappa_2} + \mathbb{N}/C(\kappa_2 \log(T) + 1) + \sum_{t=1}^{\infty} \sum_{j: \|w\|_{A_j} < r(S^*(\mu), \mu)} C^2(2 + C)\left(t - |A|\kappa_2 \log t\right)^{-c(\xi/\kappa(\xi))\kappa_2} + C^2t^{-c(\delta/\kappa(\delta))\kappa_2}
\]

\[
\leq \mathbb{N}/C\kappa_2 \log T + K_2,
\]

for a finite constant \( K_2 < \infty \), where: (a) follows from (A.13), (A.14) and (A.15); and (b) uses the summability of the series, implied by the terms in (A.13) and (A.15). Taking \( \kappa_2 > \max\{c(\xi/\kappa(\xi))^{-1}, c(\delta/\kappa(\delta))^{-1}\} \) provides the desired result.

Proof of Corollary 1. Fix \( i \in \mathcal{N} \), and fix \( j = \{k \leq |A| : i \in A_k\} \). We have that

\[
\mathbb{E}_\pi[T_i(T)] \leq \tau + \sum_{t=\tau+1}^{T} \mathbb{P}\{NO(t), G(t)\} + \mathbb{P}\{S_t = A_j, G(t)^c\}
\]

\[
\leq K_2,
\]

for a finite constant \( K_2 \), where we have used the summability of the terms in (A.13) and (A.15). This completes the proof.
Proof of Theorem 4. The proof is an adaptation of the one for Theorem 3 customized for the MNL choice model. However, we provide an explanation version of each step with the objective of highlighting how the structure of the MNL model is exploited.

Step 1. We will need the following side lemma, whose proof is deferred to Appendix A.2.

Lemma 7. Fix $i \in \mathcal{N}$. For any $n \geq 1$ and $\epsilon > 0$ one has

$$
\mathbb{P} \left\{ \frac{1}{t-1} \sum_{u=1}^{t-1} \left( Z_j^u - \mathbb{E} \{ Z_j^u \} \right) \mathbf{1} \{ i \in S_u \} \geq \epsilon T_i(t), \; T_i(t) \geq n \right\} \leq 2\exp(-c(\epsilon)n),
$$

for $j \in \{i, 0\}$ and a positive constant $c(\epsilon) < \infty$.

Consider $\epsilon > 0$ and fix $t \geq 1$ and $i \in \mathcal{N}$. Define $\varrho = 1/2 (1 + C||w||_{\infty})^{-1}$. From Assumption 2 we have that $p_0(S, \mu) \geq 2\varrho$, for all $S \in \mathcal{S}$. For $n \geq 1$ define the event $\Xi := \{ |\nu_i - \hat{\nu}_{i,t}| > \epsilon, \; T_i(t) \geq n \}$. We have that

$$
\mathbb{P} \{ \Xi \} = \mathbb{P} \left\{ \frac{1}{t-1} \sum_{u=1}^{t-1} Z_j^u \mathbf{1} \{ i \in S_u \} - \nu_i \geq \epsilon, \; T_i(t) \geq n \right\} \leq \mathbb{P} \left\{ \frac{1}{t-1} \sum_{u=1}^{t-1} Z_j^u \mathbf{1} \{ i \in S_u \} - \nu_i > \epsilon, \right\}

\sum_{u=1}^{t-1} \left( Z_j^u - \mathbb{E} \{ Z_j^u \} \right) \mathbf{1} \{ i \in S_u \} \geq \varrho T_i(t), \; T_i(t) \geq n \}

\mathbb{P} \left\{ \sum_{u=1}^{t-1} \left( Z_j^u - \mathbb{E} \{ Z_j^u \} \right) \mathbf{1} \{ i \in S_u \} \geq \varrho T_i(t), \; T_i(t) \geq n \right\}

\mathbb{P} \left\{ \sum_{u=1}^{t-1} (Z_j^u - \mathbb{E} \{ Z_j^u \}) \mathbf{1} \{ i \in S_u \} > \epsilon \varrho T_i(t), \; T_i(t) \geq n \right\} + 2\exp(-c(\varrho)n)

\mathbb{P} \left\{ \sum_{u=1}^{t-1} (Z_j^u - \mathbb{E} \{ Z_j^u \}) \mathbf{1} \{ i \in S_u \} > \epsilon \varrho/2 T_i(t), \; T_i(t) \geq n \right\} + 2\exp(-c(\varrho)n)

\mathbb{P} \left\{ \sum_{u=1}^{t-1} (Z_j^u - \mathbb{E} \{ Z_j^u \}) \mathbf{1} \{ i \in S_u \} > \epsilon \varrho/(2\nu_i) T_i(t), \; T_i(t) \geq n \right\} + 2\exp(-c(\varrho)n)

\leq 2\exp(-c(\epsilon/2)n) + 2\exp(-c(\epsilon/(2\nu_i))n) + 2\exp(-c(\varrho)n).

where: (a) follows from Lemma 7 and from the fact that

$$
\left| \sum_{u=1}^{t-1} Z_j^u \mathbf{1} \{ i \in S_u \} \right| \geq \left| \sum_{u=1}^{t-1} \mathbb{E} \{ Z_j^u \} \mathbf{1} \{ i \in S_u \} \right| - \left| \sum_{u=1}^{t-1} (Z_j^u - \mathbb{E} \{ Z_j^u \}) \mathbf{1} \{ i \in S_u \} \right| \geq \varrho T_i(t),
$$
when \( \left| \sum_{u=1}^{t-1} (Z_u^u - \mathbb{E} \{ Z_0^u \}) 1 \{ i \in S_u \} \right| < \rho T_i(t) \); and (b) follows from the fact that \( \mathbb{E} Z_i^u = \nu_i \mathbb{E} Z_0^u \), for all \( u \geq 1 \) such that \( i \in S_u, i \in \mathcal{N} \). For \( \epsilon > 0 \) define

\[
\tilde{c}(\epsilon) := \min \{ c(\epsilon \rho/2), c(\epsilon \rho/(2\|\nu\|_\mathcal{N})), c(\rho) \}.
\]

From above we have that for \( \epsilon > 0 \)

\[
P \{ |\nu_i - \hat{\nu}_{i,t}| > \epsilon, T_i(t) \geq n \} \leq 6 \exp(\tilde{c}(\epsilon)n), \tag{A.16}
\]

for all \( i \in \mathcal{N} \).

**Step 2.** Consider two vectors \( \nu, \eta \in \mathbb{R}_+^\mathcal{N} \). From (2.8), for any \( S \in \mathcal{S} \) one has

\[
\sum_{i \in S} \nu_i (w_i - r(S, \nu)) = r(S, \nu)
\]

\[
\sum_{i \in S} \eta_i (w_i - r(S, \nu)) \geq r(S, \nu) - C\|w\|_\infty \|\nu - \eta\|_S
\]

\[
\sum_{i \in S} \eta_i (w_i - r(S, \nu)) - C\|w\|_\infty \|\nu - \eta\|_S \geq r(S, \nu) - C\|w\|_\infty \|\nu - \eta\|_S
\]

This implies that

\[
r(S, \eta) \geq r(S, \nu) - C\|w\|_\infty \|\eta - \nu\|_S. \tag{A.17}
\]

From the above we conclude that

\[
\{ S^*(\hat{\nu}_t) \neq S^*(\hat{\nu}_t) \} \subseteq \left\{ \|\nu - \hat{\nu}_t\|_{S^*(\nu) \cup S^*(\nu)} \geq (2\|w\|_\infty C)^{-1} \delta(\nu) r(S^*(\mu), \mu) \right\}, \tag{A.18}
\]

were \( \delta(\nu) \) refers to the minimum optimality gap, in terms of the adjusted terms \( \exp(\mu) \).

**Step 3.** Let \( NO(t) \) denote the event that a non-optimal assortment is offered to customer \( t \), and \( G(t) \) the event that there is no “forced testing” on customer \( t \). That is

\[
NO(t) := \{ S_t \neq S^*(\nu) \},
\]

\[
G(t) := \{ T_i(t) \geq \kappa_3 \log t, \forall i \in \mathcal{N} \text{ such that } w_i \geq r(S^*(\hat{\nu}_t), \hat{\nu}_t) \}.
\]
Define \( \xi := (2\|w\|_{\infty}C)^{-1}\delta(\nu)r(S^*(\mu), \mu) \). We have that

\[
\mathbb{P}\{NO(t) \mid G(t)\} \leq (a) \mathbb{P}\{\|\nu - \hat{\nu}_t\|_{(S^*(\nu) \cup S^*(\hat{\nu}_t))} > \xi, G(t)\} \\
\leq \mathbb{P}\{\|\nu - \hat{\nu}_t\|_{S^*(\nu)} > \xi, G(t)\} + \mathbb{P}\{\|\nu - \hat{\nu}_t\|_{S^*(\nu)} > \xi, G(t)\} \\
\leq (b) \sum_{i \in S^*(\hat{\nu}_t)} \mathbb{P}\{\|\nu - \hat{\nu}_i,t\| > \xi, T_i(t) \geq \kappa_3 \log t\} + \sum_{i \in S^*(\nu)} \mathbb{P}\{\|\nu - \hat{\nu}_i,t\| > \xi, G(t)\} \\
\leq (c) 6Ct^{-\kappa_3\bar{c}(\xi)} + \sum_{i \in S^*(\nu)} \mathbb{P}\{\|\nu - \hat{\nu}_i,t\| > \xi, G(t)\}
\]

where: (a) follows from \([A.18]\); (b) follows from the fact that Assumption 2 guarantees \( w_i \geq r(S^*(\eta), \eta) \) for all \( i \in S^*(\eta) \) and for any vector \( \eta \in \mathbb{R}^N \); and (c) follows from \([A.16]\).

Fix \( i \in S^*(\nu) \). We have that

\[
\mathbb{P}\{\|\nu - \hat{\nu}_i,t\| > \xi, G(t)\} \leq \mathbb{P}\{\|\nu - \hat{\nu}_i,t\| > \xi, T_i(t) \geq \kappa_3 \log t\} + \mathbb{P}\{G(t), T_i(t) < \kappa_3 \log t\}.
\]

The first term above can be bounded using \([A.16]\). Regarding the second one, note that \( \{G(t), T_i(t) < \kappa_3 \log t\} \subseteq \{w_i < r(S^*(\hat{\nu}_i), \hat{\nu}_t)\} \), and that

\[
w_i - r(S^*(\nu), \nu)\delta(\nu)/2 \geq (a) r(S^*(\nu), \nu)(1 - \delta(\nu)/2) \\
\geq (b) r(S^*(\hat{\nu}_i), \nu) \\
\geq (c) r(S^*(\hat{\nu}_i), \hat{\nu}_t) - \|w\|_{\infty}C\|\nu - \hat{\nu}_t\|_{S^*(\hat{\nu}_i)},
\]

where (a) follows from Assumption 2, (b) follows from the definition of \( \delta(\nu) \), and (c) follows from \([A.17]\). The above implies that \( \{w_i < r(S^*(\hat{\nu}_i), \hat{\nu}_t)\} \subseteq \{\|\nu - \hat{\nu}_i\|_{S^*(\hat{\nu}_i)} > \xi\} \), i.e.,

\[
\mathbb{P}\{T_1(t) < \kappa_3 \log t, G(t)\} \leq (a) \mathbb{P}\{w_i < r(S^*(\hat{\nu}_i), \hat{\nu}_t), G(t)\} \\
\leq (b) \mathbb{P}\{\|\nu - \hat{\nu}_i\|_{S^*(\hat{\nu}_i)} > \xi, G(t)\} \\
\leq (c) \sum_{j \in S^*(\hat{\nu}_i)} \mathbb{P}\{|\nu_j - \hat{\nu}_j,t| > \xi, G(t)\} \\
\leq (d) 6Ct^{-\kappa_3\bar{c}(\xi)},
\]

where the last step follows from \([A.16]\). Using the above we have that

\[
\mathbb{P}\{NO(t) \mid G(t)\} \leq 6C(1 + C)t^{-\kappa_3\bar{c}(\xi)}.
\]
From here, we have that
\[
\mathbb{P}\{NO(t), G(t)^c\} \leq \sum_{i : w_i < f^*(S^*(\nu), \nu)} \mathbb{P}\{i \in S_t, G(t)^c\} + \sum_{i : w_i \geq f^*(S^*(\nu), \nu)} \mathbb{P}\{i \in S_t, G(t)^c\}
\]
\[
= \sum_{i : w_i < f^*(S^*(\nu), \nu)} \mathbb{P}\{i \in S_t, G(t)^c\} + \mathbb{P}\{i \in S_t, G(t)^c\} + |\mathcal{N}| (\kappa_3 \log T + 1).
\]

where (a) follows from the specification of the policy. Fix \(i\) such that \(w_i < r(S^*(\nu), \nu)\), and define \(L(t)\) as the last customer (previous to customer \(t\)) to whom the empirical optimal assortment, according to estimated mean utilities, was offered. That is
\[
L(t) := \sup \{u \leq t - 1 : G(u)\}.
\]

Note that \(L(t) \in \{t - \lfloor N \kappa_3 \log t \rfloor, \ldots, t - 1\}\) for \(t \geq \tau\), where \(\tau\) is given by
\[
\tau := \inf \{u \geq 1 : \log(u - \lfloor N \kappa_3 \log u \rfloor) + \kappa_3^{-1} > \log u\}.
\]

Consider \(t \geq \tau\) and \(u \in \{t - \lfloor N \kappa_3 \log t \rfloor, \ldots, t - 1\}\). Then
\[
\mathbb{P}\{i \in S_t, G(t)^c, L(t) = u\} \leq \mathbb{P}\{w_i \geq r(S^*(\hat{\nu}_t), \hat{\nu}_t), G(t)^c, L(t) = u\}
\]
\[
\leq \mathbb{P}\{w_i \geq r(S^*(\hat{\nu}_t), \hat{\nu}_t), G(t)^c, G(u)\}
\]
\[
\leq \mathbb{P}\{G(u), NO(u)\} + \mathbb{P}\{w_i \geq r(S^*(\hat{\nu}_t), \hat{\nu}_t), G(t)^c, G(u), NO(u)^c\}
\]
\[
\leq 6C(1 + C)u^{-\kappa_3 \hat{\tau}(t)} + \mathbb{P}\{w_i \geq r(S^*(\hat{\nu}_t), \hat{\nu}_t), T_j(t) \geq \kappa_3 \log t \forall j \in S^*(\nu)\},
\]

where (a) follows from \(\{L(t) = u\} \subseteq \{G(u)\}\), and (b) from \(\text{[A.19]}\) and the fact that offering \(S^*(\nu)\) to customer \(u\) implies (from \(G(u)\)) that \(T_j(u) \geq \kappa_3 \log u\) and therefore (from \(t \geq \tau\)) that \(T_j(t) \geq \kappa_3 \log t\), for all \(j \in S^*(\nu)\). From \(\text{[A.17]}\) we have that
\[
r(S^*(\nu), \hat{\nu}_t) - w_i \geq r(S^*(\nu), \nu) - \|w\|_\infty C\|\nu - \hat{\nu}_t\|_{S^*(\nu)} - w_i.
\]

Define \(\delta := \inf \{(\|w\|_\infty C)^{-1} (1 - w_i / r(S^*(\nu), \nu)) > 0 : i \in \mathcal{N}\}\). From the above, we have that
\[
\{w_i \geq r(S^*(\hat{\nu}_t), \hat{\nu}_t)\} \subseteq \{w_i \geq r(S^*(\nu), \hat{\nu}_t)\} \subseteq \{\|\nu - \hat{\nu}_t\|_{S^*(\nu)} > \delta r(S^*(\nu), \nu)\}.
\]
Define $\tilde{\delta} := \delta r(S^*(\nu), \nu)$, and the event $\Xi = w_i \geq r(S^*(\nu_t), \nu_t), T_j(t) \geq \kappa_3 \log t \forall j \in S^*(\nu).$

It follows that

$$
P \{ \Xi \} \leq P \{ \| \nu - \nu_t \|_{S^*(\nu)} > \delta, T_j(t) \geq \kappa_3 \log t \forall j \in S^*(\nu) \}$$

$$\leq \sum_{i \in S^*(\nu)} P \{ |\nu_i - \nu_{t,i}| > \delta, T_i(t) \geq \kappa_3 \log t \}$$

$$\leq 6Ct^{-\kappa_3 \tilde{\delta}(\delta)}.$$ 

Using the above one gets that, when $\kappa_3 > \tilde{c}(\xi)^{-1}$

$$P \{ i \in S_t, G(t)^c, L(t) = u \} \leq 6C(1 + C)u^{-\kappa_3 \tilde{c}(\xi)} + 6Ct^{-\kappa_3 \tilde{\delta}(\delta)}$$

$$\leq 6C(1 + C)(t - \lfloor N \kappa_3 \log t \rfloor)^{-\kappa_3 \tilde{c}(\xi)} + 6Ct^{-\kappa_3 \tilde{\delta}(\delta)}.$$

Since the right hand side above is independent of $u$, one has that

$$P \{ i \in S_t, G(t)^c \} \leq 6C(1 + C)(t - \lfloor N \kappa_3 \log t \rfloor)^{-\kappa_3 \tilde{c}(\xi)} + 6Ct^{-\kappa_3 \tilde{\delta}(\delta)}, \quad (A.20)$$

for all $i \in N$ such that $w_i < r(S^*(\nu), \nu)$, and $t \geq \tau$.

Now fix $i \in S^*(\mu)$, and consider $t \geq \tau, u \in \{ t - \lfloor N \kappa_3 \log t \rfloor, \ldots, t - 1 \}$ and $\kappa_3 > \tilde{c}(\xi)^{-1}$. Then

$$P \{ i \in S_t, G(t)^c, L(t) = u \} \leq P \{ T_i(t) < \kappa_3 \log t, G(t)^c, L(t) = u \}$$

$$\overset{(a)}{\leq} P \{ T_i(t) < \kappa_3 \log t, G(u) \}$$

$$\leq P \{ G(u), NO(u) \} + P \{ T_i(t) < \kappa_3 \log t, G(u), NO^c(u) \}$$

$$\overset{(b)}{\leq} 6C(1 + C)u^{-\kappa_3 \tilde{c}(\xi)}$$

$$\leq 6C(1 + C)(t - \lfloor N \kappa_3 \log t \rfloor)^{-\kappa_3 \tilde{c}(\xi)},$$

where $(a)$ follows from $\{ L(t) = u \} \subseteq \{ G(u) \}$, and $(b)$ from $\overset{[A.19]}{}$ and the fact that offering $S^*(\nu)$ to customer $u$ implies (from $G(u)$) that $T_i(u) \geq \kappa_3 \log u$ and therefore (from $t \geq \tau$) that $T_i(t) \geq \kappa_3 \log t$. Since the right hand side above is independent of $u$, one has that

$$P \{ i \in S_t, G(t)^c \} \leq 6C(1 + C)(t - \lfloor N \kappa_3 \log t \rfloor)^{-\kappa_3 \tilde{c}(\xi)}, \quad (A.21)$$

for all $i \in S^*(\mu)$ and $t \geq \tau$. 

Considering $\kappa_3 > \max \{ \tilde{c}(\xi)^{-1}, \tilde{c}(\delta)^{-1} \}$ results in the following bound for the regret

$$R^\pi(T, \nu) \leq \sum_{t=1}^{T} P\{NO(t), G(t)\} + \sum_{t=1}^{T} P\{NO(t), G(t)^c\}$$

\[(a)\]

$$\leq 6C(1 + C) \sum_{t=1}^{\infty} t^{-\kappa_3 \tilde{c}(\xi)} + |\mathcal{N} \setminus S^*(\mu)| \kappa_3 \log T + \tau + 6C |\mathcal{N} \cup S^*(\mu)| \sum_{t=\tau}^{\infty} (1 + C)(t^{-\kappa_3 \tilde{c}(\xi)} + (t - [N \kappa_3 \log t])^{-\kappa_3 \tilde{c}(\xi)} + t^{-\kappa_3 \tilde{c}(\delta)})$$

\[(b)\]

$$\leq |\mathcal{N} \setminus S^*(\mu)| \kappa_3 \log T + \mathcal{K}_3,$$

for a finite constant $\mathcal{K}_3 < \infty$, where (a) follows from (A.19), (A.20) and (A.21), and (b) uses the summability of the series, implied by the terms in (A.19), (A.20) and (A.21). Taking $\kappa_3 > \max \{ \tilde{c}(\xi)^{-1}, \tilde{c}(\delta)^{-1} \}$ provides the desired result. \[\blacksquare\]

**Proof of Corollary 2.** Fix $i \in \mathcal{N}$. We have that

$$E_\pi[T_i(T)] \leq \tau + \sum_{t=\tau+1}^{T} P\{NO(t), G(t)\} + P[i \in S_t, G(t)^c]$$

$$\leq K_3 < \infty,$$

for a finite constant $K_3$, where we have used the summability of the terms in (A.19) and (A.20). This concludes the proof. \[\square\]

### A.2 Proof of Auxiliary Results

**Proof of Lemma 6.** Fix $i \in \mathcal{N}$. For $\theta > 0$ consider the process \( \{M_t(\theta) : t \geq 1\} \), defined as

$$M_t(\theta) := \exp \left( \sum_{u=1}^{t} 1\{S_u = A_j\} [\theta (Z_i^u - p_i(A_j, \mu)) - \phi(\theta)] \right),$$

where

$$\phi(\theta) := \log \mathbb{E} \{\exp(\theta (Z_i^u - p_i(A_j, \mu)))\} = -\theta p_i(A_j, \mu) + \log(p_i(A_j, \mu) \exp(\theta) + 1 - p_i(A_j, \mu)),$$

and $A_j \in \mathcal{A}$ such that $i \in A_j$. One can check that $M_t(\theta)$ is an $\mathcal{F}_t$-martingale, for any $\theta > 0$ (see §2.3 for the definition of $\mathcal{F}_t$). Note that

$$\exp \left( \theta \sum_{u=1}^{t} 1\{S_u = A_j\} ((Z_i^u - p_i(A_j, \mu)) - \epsilon) \right) = \sqrt{M_t(2\theta)} \exp \left( \sum_{u=1}^{t} 1\{S_u = A_j\} (\phi(2\theta)/2 - \theta \epsilon) \right).$$

(A.22)
Let $\chi_i$ denote the event we are interested in. That is
\[
\chi_i := \left\{ \sum_{u=1}^{t-1} (Z_i^u - p_i(A_j, \mu)) \mathbb{1}\{S_u = A_j\} \geq T^j(t)\epsilon, T^j(t) \geq n \right\}.
\]

Let $\psi(t)$ denote the choice made by the $t$-th user. Using the above one has that
\[
P\{\chi_i\} \overset{(a)}{=} \mathbb{E} \left\{ \exp \left( \theta \sum_{u=1}^{t-1} \mathbb{1}\{S_u = A_j\} (Z_i^u - p_i(A_j, \mu) - \epsilon) \right) ; T_i(t) \geq n \right\}
\overset{(b)}{=} \left( \mathbb{E} \{M_{t-1}(2\theta)\} \mathbb{E} \left\{ \exp \left( \sum_{u=1}^{t-1} \mathbb{1}\{\psi(u) = i\} (\phi(2\theta) - 2\theta \epsilon) \right) ; T_i(t) \geq n \right\} \right)^{1/2}
\overset{(c)}{=} \left( \mathbb{E} \left\{ \exp \left( \sum_{u=1}^{t-1} \mathbb{1}\{\psi(u) = i\} (\phi(2\theta) - 2\theta \epsilon) \right) ; T_i(t) \geq n \right\} \right)^{1/2},
\]
where: (a) follows from Chernoff’s inequality; (b) follows from the Cauchy-Schwartz inequality and (A.22); and (c) follows from the properties of $M_t(\theta)$. Note that when $\epsilon < (1 - p_i(A_j, \mu))$ minimizing $\phi(\theta) - \theta \epsilon$ over $\theta > 0$ results on
\[
\theta^* := \log \left( 1 + \frac{\epsilon}{p_i(A_j, \mu)(1 - p_i(A_j, \mu) - \epsilon)} \right) > 0,
\]
with
\[
c(\epsilon) := \phi(2\theta^*)/2 - \theta^* \epsilon < 0.
\]

Using this we have
\[
P\left\{ \sum_{u=1}^{t-1} (Z_i^u - p_i) \mathbb{1}\{S_u = A_j\} \geq T^j(t)\epsilon, T^j(t) \geq n \right\} \leq \sqrt{\mathbb{E}\{\exp(-2c(\epsilon)T_i(t)); T_i(t) \geq n\}} \leq \exp(-c(\epsilon)n).
\]

Using the same arguments one has that
\[
P\left\{ \sum_{u=1}^{t-1} (Z_i^u - p_i) \mathbb{1}\{S_u = A_j\} \leq -T^j(t)\epsilon, T^j(t) \geq n \right\} \leq \exp(-c(\epsilon)n).
\]

The result follows from the union bound.

**Proof of Lemma 7.** The proof follows almost verbatim the steps in the proof of Lemma 6. Fix $i \in N$. For $\theta > 0$ consider the process \( \{M^j_t(\theta) : t \geq 1\} \), defined as
\[
M^j_t(\theta) := \exp \left( \sum_{u=1}^{t} \mathbb{1}\{i \in S_u\} \left[ \theta(Z_j^u - p_j(S_u, \mu)) - \phi^j_u(\theta) \right] \right) \quad j \in \{i, 0\},
\]
where

\[
\phi^j_s(\theta) := \log \mathbb{E} \left\{ \exp(\theta(Z^u_j - p_j(S_u, \mu))) \right\} = \log \mathbb{E} \left\{ \exp(-\theta p_j(S_u, \mu)) (\exp(\theta)p_j(S_u, \mu) + 1 - p_j(S_u, \mu)) \right\}.
\]

One can verify that \( M^j_t(\theta) \) is an \( \mathcal{F}_t \)-martingale, for any \( \theta > 0 \) and \( j \in \{i, 0\} \) (see [2.3 for the definition of \( \mathcal{F}_t \)). Fix \( j \in \{i, 0\} \) and note that

\[
\exp \left( \theta \sum_{u=1}^{t} \mathbf{1} \{i \in S_u\} (Z^u_j - p_j(S_u, \mu)) - \epsilon \right) = \sqrt{M^j_t(2\theta)} \exp \left( \sum_{u=1}^{t} \mathbf{1} \{i \in S_u\} (\phi^j_s(2\theta)/2 - \theta \epsilon) \right).
\]

Put

\[
\chi_j := \left\{ \sum_{u=1}^{t-1} (Z^u_j - p_j(S_u, \mu)) \mathbf{1} \{i \in S_u\} \geq T_i(t)\epsilon, T_i(t) \geq n \right\}.
\]

Let \( \psi(t) \) denote the choice made by the \( t \)-th customer. Using the above one has that

\[
P \{ \chi_j \} \overset{(a)}{\leq} \mathbb{E} \left\{ \exp \left( \theta \sum_{u=1}^{t-1} \mathbf{1} \{i \in S_u\} (Z^u_j - p_j(S_u, \mu) - \epsilon) \right) ; T_i(t) \geq n \right\}
\]

\[
\overset{(b)}{\leq} \left( \mathbb{E} \left\{ M^j_{t-1}(2\theta) \right\} \right)^{1/2}
\]

\[
\overset{(c)}{\leq} \left( \mathbb{E} \left\{ \exp \left( \sum_{u=1}^{t-1} \mathbf{1} \{\psi(u) = j, i \in S_u\} (\phi^j_s(2\theta) - 2\theta \epsilon) \right) ; T_i(t) \geq n \right\} \right)^{1/2},
\]

where; (a) follows from Chernoff’s inequality; (b) follows from the Cauchy-Schwartz inequality and (A.22); and (c) follows from the properties of \( M^j_t(\theta) \). Note that \( \phi^j_s(\cdot) \) is continuous, \( \phi^j_s(0) = 0, (\phi^j_s)'(0) = 0 \), and \( \phi^j_s(\theta) \to \infty \text{ when } \theta \to \infty, \text{ for all } s \geq 1 \). This implies that there exists a positive constant \( c(\epsilon) < \infty \) (independent of \( n \)), and a \( \theta^* > 0 \), such that \( \phi^j_s(2\theta^*) - 2\theta^* \epsilon < -2c(\epsilon) \) for all \( s \geq 1 \). Using this we have that

\[
P \left\{ \sum_{u=1}^{t-1} (Z^u_j - p_j(S_u, \mu)) \mathbf{1} \{i \in S_u\} \geq T_i(t)\epsilon, T_i(t) \geq n \right\} \leq \sqrt{\mathbb{E} \left\{ \exp(-2c(\epsilon)T_i(t)); T_i(t) \geq n \right\}} \leq \exp(-c(\epsilon)n).
\]

Using the same arguments one has that

\[
P \left\{ \sum_{u=1}^{t-1} (Z^u_j - p_j(S_u, \mu)) \mathbf{1} \{i \in S_u\} \leq -T_i(t)\epsilon, T_i(t) \geq n \right\} \leq \exp(-c(\epsilon)n).
\]

The result follows from the union bound.
Appendix B

Proof of Results in Chapter 3

B.1 Proof of Main Results

Proof of Theorem 5. For $X \subset X$ and $t \leq T$ define $T_i(t, X)$ as the number of times ad $i$ has been displayed to users with profiles in the set $X$, prior to arrival of user $t$, i.e.,

$$T_i(t, X) := \sum_{u=1}^{t-1} 1 \{i \in S_u, X_u \in X\},$$

and $T(t, X)$ as the number of users with profile in set $X$ arrived before user $t$, i.e.,

$$T(t, X) := \sum_{u=1}^{t-1} 1 \{X_t \in X\}.$$

Similarly, define $t_i(l, X)$ as the index of the user for which ad $i$ is displayed by the $l$-th time on the set $X$,

$$t_i(l, X) := \inf \{t \geq 0 : T_i(t + 1, X) = l\}, l \geq 1.$$

For $i \in \mathcal{N}$, define $O_i(B)$ as the set of user profiles for which ad $i$ belongs to the optimal ad-mix. That is

$$O_i(B) := \{x \in X : i \in S^*(B, x)\}.$$

Similarly, define $\Theta_i$ as the set ad-factor matrices for which ad $i$ has the same mean utility values for profiles in $O_i(B)$, but it is optimal for a broader set of profiles, leaving ad-factors of every other ad intact. That is,

$$\Theta_i := \left\{M \in \mathbb{R}^{d \times N} : O_i(B) \subset O_i(M), M_{-i} = B_{-i}, M_i^\top x = \beta_i^\top x, \forall x \in O_i(B)\right\}.$$
APPENDIX B. PROOF OF RESULTS IN CHAPTER 3

We will use $\mathbb{E}_\pi^M$ and $\mathbb{P}_\pi^M$ to denote expectations and probabilities of random variables, when the ad-mix policy $\pi \in \mathcal{P}'$ is used, and ad factors are given by the matrix $M \in \mathbb{R}^{d \times N}$. For $x \in \mathcal{X}$, let $\mathcal{I}_{i,x}(B\|M)$ denote the Kullback-Leibler divergence between $F(\cdot - \beta_i^T x)$ and $F(\cdot - M_i^T x)$,

$$\mathcal{I}_{i,x}(B\|M) := \int_{-\infty}^{\infty} [\log(dF(u - \beta_i^T x)/dF(u - M_i^T x))] dF(u - \beta_i^T x).$$

This quantity measures the “distance” between $\mathbb{P}_\pi^B$ and $\mathbb{P}_\pi^M$. We assume that $0 < \mathcal{I}_{i,x}(B\|M) < \infty$ for all $x \in O_i(B)^c$ and $M \in \Theta_i$ (this is the case for most commonly used distribution functions). Fix an ad $i \in \mathcal{N}$ for which $\Theta_i \neq \emptyset$ and consider a matrix $M \in \Theta_i$. For $n \geq 1$ and a set $X \subseteq \mathcal{X}$ define the log-likelihood function

$$\mathcal{L}_i(X, n) := \sum_{l=1}^{n} \log(dF(U_i^{t(l,X)} - \beta_i^T X_{t(l,X)})/dF(U_i^{t(l,X)} - M_i^T X_{t(l,X)})).$$

Note that $\mathcal{L}_i(\cdot)$ is defined in terms of utility realizations that are unobservable to the retailer. Define $\delta(M,X)$ as the minimum (relative) optimality gap when ad factors are given by the matrix $M$ and the user profile belongs to $X \subseteq \mathcal{X}$.

$$\delta(M,X) := \inf \left\{1 - r(S, M, x)/r(S^*(M, x), M, x) : S \in \mathcal{S}, x \in X\right\}.$$  

Fix $\alpha \in (0, 1)$ and $\rho \in (0, 1)$. For any consistent policy $\pi$ one has that, for any $\epsilon > 0$,

$$R_\pi(T, M) \geq \delta(M, \Delta_i) \mathbb{E}_\pi^M \{T(T, \Delta_i) - T_i(T, \Delta_i)\} \geq \delta(M, \Delta_i) \left((1 - \rho)q_i T - \frac{(1 - \epsilon)}{K_i(B\|M)} \log T\right).$$

$$\mathbb{P}_\pi^M \left\{T_i(T, \Delta_i) < \frac{(1 - \epsilon) \log T}{K_i(B\|M)}, T(T, \Delta_i) > (1 - \rho)q_i T\right\} \geq \delta(M, \Delta_i) \left((1 - \rho)q_i T - \frac{(1 - \epsilon)}{K_i(B\|M)} \log T\right).$$

$$\mathbb{P}_\pi^M \left\{T_i(T, O_i(B)^c) < \frac{(1 - \epsilon) \log T}{K_i(B\|M)}, T(T, \Delta_i) > (1 - \rho)q_i T\right\},$$

where $\Delta_i := O_i(M) \setminus O_i(B)$, $q_i := \mathbb{P}\{X_t \in \Delta_i\}$ and $K_i(B\|M) := \max \{\mathcal{I}_{i,x}(B\|M) : x \in O_i(\beta)^c\}$. From the above and [3.3], we have that

$$\mathbb{P}_\pi^M \left\{T_i(T, O_i(B)^c) < \frac{(1 - \epsilon) \log T}{K_i(B\|M)}, T(T, \Delta_i) > (1 - \rho)q_i T\right\} = o(T^{\alpha - 1}). \quad (B.1)$$
APPENDIX B. PROOF OF RESULTS IN CHAPTER 3

Define the event
\[ \gamma_i := \left\{ T_i(T, O_i(B)^c) < \frac{(1 - \epsilon) \log T}{\mathcal{K}_i(B\|M)}, \mathcal{L}_i(O_i(B)^c, T_i(T, O_i(B)^c)) \leq (1 - \alpha) \log T, \right. \\
\left. T(T, \Delta_i) > (1 - \rho)q_i T \right\}. \]

From the independence of utilities across ads and the definition of \( \gamma_i \), we have that
\[
\mathbb{P}_\pi^M \{ \gamma_i \} = \int_{\omega \in \gamma_i} d\mathbb{P}_\pi^M \\
= \int_{\omega \in \gamma_i} \prod_{u=1}^T \prod_{j \in S_u} dF(U_u^j - M_j^T X_u) \\
= \int_{\omega \in \gamma_i} \prod_{u=1}^T \prod_{j \in S_u} dF(U_u^j - M_j^T X_u) d\mathbb{P}_\pi^B \\
= \int_{\omega \in \gamma_i} \prod_{u=1}^T \prod_{j \in S_u} \frac{dF(U_u^j - M_j^T X_u)}{dF(U_u^j - \beta_j^T X_u)} d\mathbb{P}_\pi^B \\
= \int_{\omega \in \gamma_i} \exp(-\mathcal{L}_i(O_i(B)^c, T_i(T, O_i(B)^c))) d\mathbb{P}_\pi^B \\
\geq \exp(-(1 - \alpha) \log T) \mathbb{P}_\pi^B \{ \gamma_i \}.
\]

From (B.1) one has that \( \mathbb{P}_\pi^B \{ \gamma_i \} = o(T^{\alpha - 1}) \). It follows by (B.1) that as \( T \to \infty \)
\[
\mathbb{P}_\pi^B \{ \gamma_i \} \leq \mathbb{P}_\pi^M \{ \gamma_i \} / T^{\alpha - 1} \to 0.
\] (B.2)

Indexed by \( n \), \( \mathcal{L}_i(O_i(B)^c, n) \) is the sum of finite-mean independent random variables, which is in turn the sum of \( |O_i(B)^c| \) sums of at most \( n \) identically distributed random variables.

By the strong law of large numbers (SLLN).
\[
\limsup_{n \to \infty} \frac{\max \{ \mathcal{L}_i(O_i(B)^c, l) : l \leq n \}}{n} \leq \mathcal{K}_i(B\|M) \quad \mathbb{P}_\pi^B - a.s..
\]

(Otherwise one can find at least one \( x \in O_i(\beta)^c \) such that its asymptotic average contribution to \( \mathcal{L}_i(O_i(B)^c, n) \) is greater than \( \mathcal{K}_i(B\|M) \).) This result says that the log-likelihood function grows no faster than linearly with slope \( \mathcal{K}_i(B\|M) \). This implies that
\[
\limsup_{n \to \infty} \mathbb{P}_\pi^B \{ \exists l \leq n, \mathcal{L}_i(O_i(B)^c, l) > n\mathcal{K}_i(B\|M) \} = 0.
\]

In particular,
\[
\lim_{T \to \infty} \mathbb{P}_\pi^B \left\{ T_i(T, O_i(B)^c) < \frac{(1 - \epsilon)}{\mathcal{K}_i(B\|M)} \log T, \mathcal{L}_i(O_i(B)^c, T_i(T, O_i(B)^c)) > (1 - \epsilon) \log T, T(t, \Delta_i) > (1 - \rho)q_i T \right\} = 0.
\]
Taking $\alpha < \epsilon$ and combining with (B.2) one has that
\[
\lim_{T \to \infty} \mathbb{P}_\pi^B \left\{ T_i(T, O_i(B)^c) < \frac{(1 - \epsilon)}{K_i(B\|M)} \log T, \ T(T, \Delta_i) > (1 - \rho)q_iT \right\} = 0.
\]
Define the positive finite constant $H_i(B) := \inf \{ K_i(B\|M) : M \in \Theta_i \}$. It follows that that
\[
\lim_{T \to \infty} \mathbb{P}_\pi^B \left\{ T_i(T, O_i(B)^c) \geq \frac{(1 - \epsilon)}{H_i(B)} \log T \right\} + \lim_{T \to \infty} \mathbb{P}_\pi^B \left\{ T(T, \Delta_i) \leq (1 - \rho)q_iT \right\} \geq 1.
\]
By Hoeffding’s inequality one has that
\[
\lim_{T \to \infty} \mathbb{P}_\pi^B \left\{ T(T, \Delta_i) \leq (1 - \rho)q_iT \right\} \leq \lim_{T \to \infty} \exp\left(-\frac{2(\rho q_i)^2}{T}\right) = 0,
\]
thus we conclude that
\[
\lim_{T \to \infty} \mathbb{P}_\pi^B \left\{ T_i(T, O_i(B)^c) \geq \frac{(1 - \epsilon)}{H_i(B)} \log T \right\} = 1.
\]
By Markov’s inequality, and letting $\epsilon$ shrink to zero we get
\[
\liminf_{T \to \infty} \frac{\mathbb{P}_\pi^B \left\{ T_i(T, O_i(B)^c) \right\}}{\log T} \geq \frac{1}{H_i(B)}. \tag{B.3}
\]
By the definition of the regret, we have that for any consistent policy $\pi \in \mathcal{P}'$,
\[
\mathcal{R}^\pi(T, B) \geq \mathbb{E}_\pi^B \left\{ \sum_{t=1}^T \delta(B, X_t)1 \{ S_t \neq S^*(B, X_t) \} \right\} \geq \frac{1}{C} \sum_{i \in \mathcal{N}} \delta(B, O_i(B)^c) \mathbb{E}_\pi^B \left\{ T_i(T, O_i(B)^c) \right\},
\]
where (a) follows from the non-optimal ad-mixes contributing at least $\delta(\mu, X_t)$ to the regret on period $t$, and (b) follows by assuming non-optimal ads for any given profile are always tested in batches of size $C$. Thus
\[
\sum_{u=1}^T 1 \{ S_u \neq S^*(B, X_u) \} \geq \sum_{u=1}^T 1 \{ S_u \setminus S^*(B, X_u) \neq \emptyset \} \geq \frac{1}{C} \sum_{i \in \mathcal{N}} \sum_{u=1}^T 1 \{ i \in S_u \cap X_u \in O_i(B)^c \} = \frac{1}{C} \sum_{i \in \mathcal{N}} T_i(T, O_i(B)^c).
\]
Combining the above with (B.3) we have that, asymptotically,
\[
\mathcal{R}^\pi(T, B) \geq \frac{1}{C} \left( \sum_{i \in \mathcal{N}} \frac{\delta(B, O_i(B)^c)}{H_i(B)} \right) \log T.
\]
Taking $K_i := \delta(B, O_i(B)^c) / (C H_i(B))$ gives the desired result. \qed
APPENDIX B. PROOF OF RESULTS IN CHAPTER 3

130

Proof of Lemma 1. From (3.6) we have that

\[
\mathbb{P}_\pi \{X_t = x, Z^t_i = 1\} = \mathbb{P}_\pi \{Z^t_i = 1 | X_t = x, i \in S_t\} \mathbb{P}_\pi \{i \in S_t | X_t = x\} g(x)
\]

\[
= \exp(\beta_i^T x) \mathbb{E}_\pi \left\{ \left( 1 + \sum_{j \in S_t} \exp(\beta_j^T x) \right)^{-1} \mathbbm{1}_{\{i \in S_t\}} | X_t = x \right\} g(x),
\]

for all \(i \in \mathcal{N}\), where \(g(x) := \mathbb{P}\{X_t = x\}\). The same way one can show that

\[
\mathbb{P}_\pi \{X_t = x, i \in S_t, Z^t_0 = 1\} = \mathbb{E}_\pi \left\{ \left( 1 + \sum_{j \in S_t} \exp(\beta_j^T x) \right)^{-1} \mathbbm{1}_{\{i \in S_t\}} | X_t = x \right\} g(x).
\]

With this we have that

\[
\frac{\mathbb{E}_\pi \left\{ \sum_{s=1}^{t-1} \mathbbm{1}_{\{X_s = x, Z^s_i = 1\}} \right\}}{\mathbb{E}_\pi \left\{ \sum_{s=1}^{t-1} \mathbbm{1}_{\{X_s = x, i \in S_s, Z^s_0 = 1\}} \right\}} = \frac{\sum_{s=1}^{t-1} \mathbb{P}_\pi \{X_s = x, Z^s_i = 1\}}{\sum_{s=1}^{t-1} \mathbb{P}_\pi \{X_s = x, i \in S_s, Z^s_0 = 1\}} = \exp(\beta_i^T x).
\]

\[\square\]

Proof of Lemma 2. Fix \(x \in \mathcal{X}\). Solving for \(S^*(B, x)\) amounts to find the highest value of \(\lambda\) for which

\[
\max_{S \subseteq \mathcal{S}} \sum_{i \in S} \exp(\mu_i(x))(w_i - \lambda) \geq \lambda
\]

holds. It is easy to see that such value is \(\lambda^* = r(S^*(B, x), B, x)\). With this value at hand one can find \(S^*(B, x)\) by selecting the \(C\) ads with the highest (positive) values for \(\exp(\mu_i(x))(w_i - \lambda^*)\).

For each ad \(i \in \mathcal{N}\) with \(w_i > \lambda^*\), one can select \(\mu_i(x)\) big enough so \(i\) becomes one of the \(C\) most attractive ads, i.e., \(i\) is potentially optimal (by definition). This implies that \(\mathcal{N}(x) = \{i \in \mathcal{N} : w_i \geq r(S^*(B, x), B, x)\}\). The Lemma follows directly from this last observation.

\[\square\]

Proof of Theorem 6. We begin by stating a series of side lemmas, whose proofs are deferred to Appendix B.2. For purposes of this proof, let \(\mathbb{P}\) denote probability of random variables when the ad-mix policy \(\pi = \pi(w, \kappa)\) is used, and the matrix of ad features is given by \(B\).

Our first side lemma provides an upper bound on the estimation error as a function of the number of samples used for computing such estimates.
Lemma 8. For \( i \in \mathcal{N} \) and \( n \geq 1 \) we have that for any \( \epsilon > 0 \) and \( k \geq 0 \)

\[
\mathbb{P}\left\{ \left| \exp(\beta_i^\top x) - \exp((\hat{\beta}^n_i)^\top x) \right| > \epsilon, T_i(\tau(n), y) \geq k, y \in E^{n-1}(i) \right\} \leq c_1(\epsilon) \exp(c_2(\epsilon) k),
\]

for finite positive constants \( c_1(\epsilon) \) and \( c_2(\epsilon) \), for all \( x \in \text{span} \{E^{n-1}(i)\} \cap \mathcal{X} \).

The next lemma establishes that in order to recover the optimal display \( S^*(B, x) \) for a given profile \( x \in \mathcal{X} \), it is sufficient to solve the static optimization problem (3.2) using an estimate matrix \( \hat{B} \), as long as that estimate is sufficiently close to \( B \).

Lemma 9. For any \( x \in \mathcal{X} \) and \( M \in \mathbb{R}^{N \times d} \) one has that

\[
\left\{ \sum_{i \in S^*(B, x) \cup S^*(M, x)} 2w_i \left| \exp(M_i^\top x) - \exp(B_i^\top x) \right| < \xi \right\} \subseteq \{ S^*(B, x) = S^*(M, x) \}, \tag{B.4}
\]

where \( \xi := \min \{ \delta(B, x) r(S^*(B, x), B, x) : x \in \mathcal{X} \} \).

For \( n \geq 1 \) define the event \( \Theta^n(S, x, \epsilon) := \left\{ \sum_{i \in S} 2w_i \left| \exp(\beta_i^\top x) - \exp((\hat{\beta}^n_i)^\top x) \right| > \epsilon \right\} \), for \( x \in \mathcal{X} \), \( S \in \mathcal{S} \) and \( \epsilon > 0 \). The next lemma bounds the probability of \( \Theta^n(S, x, \epsilon) \) for two specific \( S \).

Lemma 10. Consider \( S \in \left\{ S^*(\hat{B}^n, x), S^*(B, x) \right\} \). For any \( \epsilon > 0 \) and \( x \in \mathcal{X} \),

\[
\mathbb{P}\{ \Theta^n(S, x, \epsilon) \} \leq K_X(\epsilon) \tau(n)^{-c_X(\epsilon) \kappa},
\]

for \( n \geq \pi := \inf \{ n \geq 0 : \kappa \log \tau(n) < n + 1 \} \), and finite positive constants \( c_X(\epsilon) \) and \( K_X(\epsilon) \).

The proof is divided in four steps. Step 1 provides auxiliary results. Step two establishes a bound on the probability of making errors when trusting the estimate parameters for solving the static problem. Step 3 bounds the amount of exploration conducted throughout the execution of the algorithm. Finally, step 4 uses the results of step 2 and 3 to bound the total regret of the algorithm.

**Step 1.** We begin by establishing a bound on the probability of changing the set of profiles generating the estimation sets. Let \( \xi_w \) be the minimum distance between ad margins and the optimal expected revenue, uniformly across profiles. That is,

\[
\xi_w := \inf \{ |w_i - r(S^*(B, x), B, x)| : x \in \mathcal{X}, i \in \mathcal{N} \}.
\]
By Assumption A, we have that $\xi_w > 0$. For $n \geq \overline{n}$ and $i \in \mathcal{N}$, one has that for $\epsilon > 0$,

$$\mathbb{P}\{\mathcal{X}_i^n \neq \mathcal{X}(i)\} \leq \sum_{x \in \mathcal{X}} \mathbb{P}\left\{\left|r(S^*(\hat{B}^n, x), \hat{B}^n, x) - r(S^*(B, x), B, x)\right| > \xi_w\right\}$$

which (a) follows from the definition of $\xi_w$, (b) follows from the proof of Lemma 9, and (c) follows from Lemma 10. Using this result, one obtains that for $n > \overline{n}$,

$$\mathbb{P}\{\mathcal{X}_i^n \neq \mathcal{X}_i^{n-1}, \text{ for some } i \in \mathcal{N}\} \leq \sum_{i \in \mathcal{N}} \mathbb{P}\{\mathcal{X}_i^{n-1} \neq \mathcal{X}(i)\} + \sum_{i \in \mathcal{N}} \mathbb{P}\{\mathcal{X}_i^n \neq \mathcal{X}(i)\}$$

which is a bound on the probability of making an error when solving the static optimization problem using the parameter estimates. Fix $n > \overline{n}$; from Lemma 9 and (B.5) one has that, for any $x \in \mathcal{X}$,

$$\mathbb{P}\left\{S^*(\hat{B}^n, x) \neq S^*(B, x)\right\} \leq \mathbb{P}\left\{\sum_{i \in S^*(\hat{B}^n, x) \cup S^*(B, x)} 2w_i \left|\exp(\beta_i^T X_t) - \exp((\hat{\beta}_i^n)^T X_t)\right| > \xi\right\} + \mathbb{P}\{\Theta^n(S^*(\hat{B}^n, x), x, \xi/2)\} + \mathbb{P}\{\Theta^n(S^*(B, x), x, \xi/2)\}$$

which is a bound on the amount of exploration performed during the execution of the algorithm. Let $I(n)$ the amount of exploration performed between periods $\tau(n)$ and $\tau(n+1)$, and define the event $\Xi := \{\mathcal{X}_i^n = \mathcal{X}_i^{n-1} = \mathcal{X}(i), i \in \mathcal{N}\}$. We have that

$$\mathbb{E}\{I(n)\} = \mathbb{E}\{I(n)1\{\Xi\}\} + \mathbb{E}\{I(n)1\{\mathcal{X}_i^n \neq \mathcal{X}(i) \text{ or } \mathcal{X}_i^{n-1} \neq \mathcal{X}(i), \text{ for some } i \in \mathcal{N}\}\}$$

We now establish upper bounds for both terms on the right-hand-side of the above. We have that

$$\mathbb{E}\{I(n)(1 - 1\{\Xi\})\} \leq (|\mathcal{N}| d(n+1)) \sum_{i \in \mathcal{N}} \sum_{k=n-1}^n \mathbb{P}\{\mathcal{X}_i^k \neq \mathcal{X}(i)\}$$

which is a bound on the amount of exploration performed between periods $\tau(n)$ and $\tau(n+1)$, and define the event $\Xi := \{\mathcal{X}_i^n = \mathcal{X}_i^{n-1} = \mathcal{X}(i), i \in \mathcal{N}\}$. We have that

$$\mathbb{E}\{I(n)(1 - 1\{\Xi\})\} \leq 4 \left(|\mathcal{X}| \sum_{i \in \mathcal{N}} \mathbb{E}\{1\{\mathcal{X}_i^n \neq \mathcal{X}(i) \text{ or } \mathcal{X}_i^{n-1} \neq \mathcal{X}(i), \text{ for some } i \in \mathcal{N}\}\}\right)$$
where (a) follows from (B.5). We now focus on bounding the first term on the right-hand-side of (B.7). For that, define \( I_i(n) \) as the amount of exploration on performed between \( \tau(n) \) and \( \tau(n + 1) \), associated to \( i \in \mathcal{N} \). Note that \( \tau(n) - \tau(n - 1) \geq \tau(n)(1 - \exp(-2/\kappa)) \) for \( n \geq \pi \). We have that
\[
\mathbb{E} \{ I(n) 1 \{ \Xi \} \} \leq \sum_{i \in \mathcal{N}} \mathbb{E} \{ I_i(n) 1 \{ \Xi \} \}
\]
\[
\leq \sum_{i \in \mathcal{N}} \text{rank} \{ \mathcal{X}(i) \} \mathbb{P} \{ \Xi, T_i(\tau(n), x) \leq n \text{ for some } x \in O(i) \} + \sum_{i \in \mathcal{N}} \text{rank} \{ \mathcal{X}(i) \setminus \text{span} \{ O(i) \} \} \mathbb{P} \{ \Xi, T_i(\tau(n) > n, i \in O(i) \} + d |\mathcal{N}| (n + 1) O(\exp(-\tau(n)),
\]
for \( n \geq \pi \), where the last term above, derived from applying Hoeffdings inequality, subsumes the possibility of not collecting the required amount of experimentation between \( \tau(n - 1) \) and \( \tau(n) \). Fix \( i \in \mathcal{N} \). For each \( x \in O(i) \) one has that
\[
\mathbb{P} \{ \Xi, T_i(\tau(n), x) \leq n \} \leq \mathbb{P} \left\{ \Xi, T_i(\tau(n), x) \leq n, i \in S^*(\hat{B}^{n-1}, x) \right\} + \mathbb{P} \left\{ \Xi, i \notin S^*(\hat{B}^{n-1}, x) \right\}
\]
(a) \[
\leq \mathbb{P} \left\{ \sum_{t = \tau(n-1)}^{\tau(n)-1} 1 \{ X_t = x \} \leq n |\mathcal{N}| \right\} + \mathbb{P} \left\{ S^*(\hat{B}^{n-1}, x) \neq S^*(B, x) \right\}
\]
(b) \[
\leq O(-\exp(\tau(n)) + 2K_\mathcal{X}(\xi/2)(\tau(n - 1))^{-c_\mathcal{X}(\xi/2)\kappa},
\]
where (b) follows from (B.6) and a direct application of Hoeffding’s inequality. With this, one has that
\[
\mathbb{E} \{ I(n) \} \leq \sum_{i \in \mathcal{N}} \text{rank} \{ \mathcal{X}(i) \setminus \text{span} \{ O(i) \} \} + 4 \left( |\mathcal{X}| |\mathcal{N}|^2 d(n + 1) \right) K_\mathcal{X}(\xi_w/2) (\tau(n - 1))^{-c_\mathcal{X}(\xi_w/2)\kappa}
\]
\[
+ \sum_{i \in \mathcal{N}} \text{rank} \{ \mathcal{X}(i) \} |O(i)| \left( O(\exp(-\tau(n)) + 2K_\mathcal{X}(\xi/2)(\tau(n - 1))^{-c_\mathcal{X}(\xi/2)\kappa} \right). \quad \text{(B.8)}
\]

**Step 4.** Now, we combine the results from the previous steps to bound the regret associated to the algorithm. In particular, using the results in **Step 2** and **Step 3** we have the following.

\[
\mathcal{R}^*(T, B) \leq \sum_{n=1}^{n(T)} \|w\|_\infty (\tau(n) - \tau(n - 1)) \sum_{x \in \mathcal{X}} \mathbb{P} \left\{ S^*(\hat{B}^{n-1}, x) \neq S^*(B, x) \right\} + \sum_{n=1}^{n(T)} \mathbb{E} \{ I(n) \}.
\]
The first term above represents the regret coming from offering the approximate solution to the static optimization problem, assuming one is to present such solution to every customer. The second term bounds the amount of exploration during the execution of the algorithm. Note that all elements in the right hand side of (B.8) define convergent series in \( n \) (but the first), provide that \( \kappa > 1 / c_X(\xi_w/2) \)\(^1\) hence one obtains that

\[
\sum_{n=1}^{n(T)} \mathbb{E}\{I(n)\} \leq \left( \sum_{i \in \mathcal{N}} \text{rank}\{\mathcal{X}(i) \setminus \text{span}\{O(i)\}\} \right) \kappa \log T + \bar{K},
\]

for some finite and positive constant \( \bar{K} \), independent of \( T \). Similarly, all elements in (B.6) are \( o(\tau(n)^2) \), provided that \( \kappa > 2 / c_X(\xi_w/2) \). Using this one obtains that

\[
\sum_{n=1}^{n(T)} (\tau(n) - \tau(n-1)) \mathbb{P}\left\{ S^*(\bar{B}^n, x) \neq S^*(B, x) \right\} \leq \bar{K},
\]

for some finite and positive constant \( \bar{K} \), independent of \( T \), for all \( x \in \mathcal{X} \). Combining the results above one obtains

\[
R^x(T, B) \leq \left( \sum_{i \in \mathcal{N}} \text{rank}\{\mathcal{X}(i) \setminus \text{span}\{O(i)\}\} \right) \kappa \log T + \bar{K},
\]

for a suitable chosen finite and positive constant \( \bar{K} \), independent of \( T \). The result follows from defining \( \bar{K} := \bar{K} + \bar{K} \), \( \bar{K} := 2 / c_X(\xi_w/2) \) and \( \bar{K}_i := \text{rank}\{\mathcal{X}(i) \setminus \text{span}\{O(i)\}\} \), \( i \in \mathcal{N} \).

**B.2 Proof of Auxiliary Results**

We begin by stating and proving a couple of intermediate results.

**Lemma 11.** Fix \( i \in \mathcal{N} \) and \( x \in \mathcal{X} \). For any \( n \geq 1 \), \( t > 1 \) and \( \epsilon > 0 \) one has that

\[
\mathbb{P}\left\{ \left| \sum_{u=1}^{t-1} (Z^u_j - \mathbb{E}\{Z^u_j\}) \mathbf{1}\{i \in S_u, X_u = x\} \right| \geq \epsilon \right\} \leq 2 \exp\left(-c(\epsilon)n\right),
\]

for \( j \in \{i, 0\} \) and a positive constant \( c(\epsilon) < \infty \).

\(^1\)From proof of Lemma 10 one has that \( c_X(\xi_w/2) \leq c_X(\xi/2) \) by construction.
Proof of Lemma \ref{lemma:proof}. The proof follows almost verbatim the steps in the proof of Lemma 2 in \cite{SaureandZeevi2011}. Fix $\theta > 0$ consider the process $\{M_t^j(\theta) : t \geq 1\}$, defined as

$$M_t^j(\theta) := \exp \left( \sum_{u=1}^{t} 1 \{ i \in S_u, X_u = x \} \left[ \theta(Z_j^u - p_j(S_u, B)) - \phi_u^j(\theta) \right] \right) \quad j \in \{i, 0\},$$

where

$$\phi_u^j(\theta) := \log \mathbb{E} \left\{ \exp(\theta(Z_j^u - p_j(S_u, B))) \right\}$$

$$:= \log \mathbb{E} \left\{ \exp(-\theta p_j(S_u, B)) \right\} \exp(\theta)p_j(S_u, B, x) + 1 - p_j(S_u, B, x) \right\}.$$ 

One can verify that $M_t^j(\theta)$ is an $\mathcal{F}_t$-martingale, for any $\theta > 0$ and $j \in \{i, 0\}$ (see Section 3.3 for the definition of $\mathcal{F}_t$). Fix $j \in \{i, 0\}$ and note that

$$\frac{\exp \left( \theta \sum_{u=1}^{t} 1 \{ i \in S_u, X_u = x \} \left( (Z_j^u - p_j(S_u, B, x)) - \epsilon \right) \right)}{\sqrt{M_t^j(2\theta)}} = \sqrt{M_t^j(2\theta)}. \quad (B.1)$$

Define

$$\chi_j := \left\{ \sum_{u=1}^{t-1} (Z_j^u - p_j(S_u, B, x)) \right\} 1 \{ i \in S_u, X_u = x \} \geq T_i(t, x) \epsilon, T_i(t, x) \geq n \right\}.$$ 

Using the above one has that

$$\mathbb{P} \{ \chi_j \} \leq \mathbb{E} \left\{ \exp \left( \theta \sum_{u=1}^{t} 1 \{ i \in S_u, X_u = x \} (Z_j^u - p_j(S_u, B, x)) - \epsilon \right) \right\} 1 \{ T_i(t, x) \geq n \} \leq \mathbb{E} \left\{ M_{t-1}^j(2\theta) \right\} \mathbb{E} \left\{ \exp \left( \sum_{u=1}^{t-1} 1 \{ i \in S_u, X_u = x \} (\phi_u^j(2\theta) - 2\theta\epsilon) \right) 1 \{ T_i(t, x) \geq n \} \right\}^{1/2} \leq \mathbb{E} \left\{ \exp \left( \sum_{u=1}^{t-1} 1 \{ i \in S_u, X_u = x \} (\phi_u^j(2\theta) - 2\theta\epsilon) \right) 1 \{ T_i(t, x) \geq n \} \right\}^{1/2},$$

where; (a) follows from Chernoff’s inequality; (b) follows from the Cauchy-Schwartz inequality and (B.1); and (c) follows from the properties of $M_t^j(\theta)$. Note that $\phi_u^j(\cdot)$ is continuous, $\phi_u^j(0) = 0$, $(\phi_u^j)'(0) = 0$, and $\phi_u^j(\theta) \to \infty$ when $\theta \to \infty$, for all $s \geq 1$. This implies that there exists a positive constant $c(\epsilon) < \infty$ (independent of $n$), and a $\theta^* > 0$, such that $\phi_u^j(\theta) \to -2\theta^* \epsilon$ when $\theta \to \infty$, for all $s \geq 1$. Using this we have that

$$\mathbb{P} \{ \xi_j \} \leq \mathbb{E} \{ \exp(-c(\epsilon)T_i(t, x))1 \{ T_i(t, x) \geq n \} \} \leq \exp(-c(\epsilon)n).$$
Using the same arguments one has that
\[
\Pr \left\{ \sum_{u=1}^{t-1} (Z_{u}^n - p_j(S_u, B, x)) \mathbf{1} \{ i \in S_u, X_u = x \} \leq -T_i(t, x) \epsilon, T_i(t, x) \geq n \right\} \leq \exp(-c(\epsilon)n).
\]

The result follows from the union bound.

**Lemma 12.** For \( i \in \mathcal{N} \) and \( n \geq 1 \) we have that for any \( \epsilon > 0, k \in \mathbb{N}^+ \)
\[
\Pr \left\{ \left| \exp(\beta_i^\top x) - \exp((\hat{\beta}_i^n)^\top x) \right| > \epsilon, T_i(\tau(n), x) \geq k \right\} \leq c_1(\epsilon) \exp(c_2(\epsilon)n),
\]
for finite positive constants \( c_1(\epsilon) \) and \( c_2(\epsilon) \), for any \( x \in E^{n-1}(i) \).

**Proof of Lemma 12.** Define \( \rho \) as half the minimum probability of no-click over all possible ad-mixes and user profiles. That is
\[
\rho := \frac{1}{2} \min \{ p_0(S, B, x) : S \in \mathcal{S}, x \in \mathcal{X} \}.
\]

Define the event \( \Xi := \left\{ \left| \exp(\beta_i^\top x) - \exp((\hat{\beta}_i^n)^\top x) \right| > \epsilon, T_i(\tau(n), x) \geq k \right\} \). Fix \( i \in \mathcal{N} \) and set \( t = \tau(n) \). For \( x \in E^{n-1}(i) \) one has that
\[
\exp(\hat{\beta}_i^\top x) = \frac{\sum_{u=1}^{t-1} Z_u^i \mathbf{1} \{ i \in S_u, X_u = x \}}{\sum_{u=1}^{t-1} Z_0^i \mathbf{1} \{ i \in S_u, X_u = x \}}.
\]
hence one can write
\[ P \{ \Xi \} \leq P \left\{ \sum_{i=1}^{t-1} (Z_i^u - \mathbb{E} \{ Z_i^u \}) 1 \{ i \in S_u, X_u = x \} < \varrho \, T_i(t, x) \right\} + \]
\[ P \left\{ \sum_{i=1}^{t-1} (Z_i^u - \mathbb{E} \{ Z_i^u \}) 1 \{ i \in S_u, X_u = x \} \geq \varrho \, T_i(t, x) \right\} \]
\[ \leq P \left\{ \sum_{i=1}^{t-1} Z_i^u 1 \{ i \in S_u, X_u = x \} - \exp(\beta^\top_i x) \right\} > \varepsilon, T_i(t, x) \geq k, \]
\[ \quad \sum_{i=1}^{t-1} (Z_i^u - \mathbb{E} \{ Z_i^u \}) 1 \{ i \in S_u, X_u = x \} \right\} < \varrho \right\} + 2 \exp(-c(\varrho)k) \]
\[ \leq 2 \exp(-c(\varepsilon)k) + 2 \exp\left(-c\left(\frac{\varepsilon \varrho}{2}\right)k\right) + 2 \exp\left(-c\left(\frac{\varepsilon \varrho}{2 \exp(\beta^\top_i x)}\right)k\right), \]

where: (a) follows from Lemma 11 and from the fact that
\[ \sum_{i=1}^{t-1} Z_i^u 1 \{ i \in S_u, X_u = x \} \geq \sum_{i=1}^{t-1} E[Z_i^u] 1 \{ i \in S_u, X_u = x \} - \]
\[ \sum_{i=1}^{t-1} (Z_i^u - E[Z_i^u]) 1 \{ i \in S_u, X_u = x \} \geq \varrho \, T_i(t, x), \]
when \[ \sum_{i=1}^{t-1} (Z_i^u - E[Z_i^u]) 1 \{ i \in S_u, X_u = x \} \leq \varrho \, T_i(t, x); \] and (b) follows from the fact that \[ E \{ Z_i^u \} = \exp(\beta^\top_i x) \mathbb{E} \{ Z_i^u \}, \] for all \( u \geq 1 \) such that \( i \in S_u \) and \( X_u = x \). For \( \varepsilon > 0 \) define
\[ \bar{c}(\varepsilon) := \min \left\{ c(\varepsilon), c\left(\frac{\varepsilon \varrho}{2}\right), c\left(\frac{\varepsilon \varrho}{2 \exp(\beta^\top_i x)}\right) \right\}. \]

From above one have that
\[ P \left\{ \exp(\beta^\top_i x) - \exp((\hat{\beta}_i^u)^\top x) > \varepsilon, T_i(t, x) \geq k \right\} \leq 6 \exp(-\bar{c}(\varepsilon)k). \]
APPENDIX B. PROOF OF RESULTS IN CHAPTER 3

138

Proof of Lemma 8. Since \( x \in \text{span} \{ E^{n-1}(i) \} \), one has that \( x = \sum_{y \in E^{n-1}(i)} \alpha_y y \), for some set \( \alpha := \{ \alpha_y : y \in E^{n-1}(i) \} \). Suppose \( |\exp(\beta_i^\top x) - \exp((\beta_i^n)^\top y)| < \epsilon/d \) for all \( y \in E^{n-1}(i) \), then

\[
|\exp(\beta_i^\top x) - \exp((\beta_i^n)^\top x)| \leq \left( \prod_{y \in E^{n-1}(i)} \left( \exp(\beta_i^\top y) \right)^{\alpha_y} - \prod_{y \in E^{n-1}(i)} \left( \exp((\beta_i^n)^\top y) \right)^{\alpha_y} \right)^{(a)} \\
\leq c_3(\epsilon, x, E^{n-1}(i)) \sum_{y \in E^{n-1}(i)} |\exp(\beta_i^\top y) - \exp((\beta_i^n)^\top y)| \leq \epsilon,
\]

for a finite positive constant \( c_3(\epsilon, x, E^{n-1}(i)) \), where (a) comes from the fact that the function \( f(z) = \prod_{i=1}^{d} (z_i^{\alpha_i}) \) is locally Lipschitz for \( z \in \mathbb{R}_+^d \), for any \( \alpha \). For \( \epsilon > 0 \) define

\[
\tilde{c}(\epsilon) = \max \{ c_3(\epsilon, x, S) : x \in \mathcal{X}, S \subseteq \mathcal{X}, 1 \leq |S| \leq C \}.
\]

It is easy to see that \( \tilde{c}(\epsilon) \) is finite and positive for all \( \epsilon > 0 \).

Define \( \Xi = \left\{ \left| \exp(\beta_i^\top x) - \exp((\beta_i^n)^\top x) \right| > \epsilon, T_i(t, y) \geq k, y \in E^{n-1}(i) \right\} \). From above one has that

\[
\mathbb{P} \{ \Xi \} \leq \sum_{y \in E^{n-1}(i)} \mathbb{P} \left\{ \left| \exp(\beta_i^\top y) - \exp((\beta_i^n)^\top y) \right| > \epsilon, T_i(\tau(n), y) \geq k \right\} \overset{(a)}{\leq} 6 \exp \left( -\tilde{c}(\epsilon) k \right),
\]

where \( \epsilon := \epsilon/(d \cdot \tilde{c}(\epsilon)) \), and (a) follows from Lemma 12.

Proof of Lemma 9. From proof of Lemma 2 one sees that, for any \( M \in \mathbb{R}^{N \times d} \),

\[
\sum_{i \in S} \exp(M_i^\top x)(w_i - r) \geq r, \tag{B.2}
\]

for all \( S \subseteq \mathcal{S} \) and \( x \in \mathcal{X} \), for \( r \leq r(S^*(M, x), M, x) \). In particular, the above holds as an equality for \( S = S^*(M, x) \). For \( M \) and \( L \) in \( \mathbb{R}^{N \times d} \) define

\[
\Delta(M, L) := \sum_{i \in S^*(M, x)} w_i \left| \exp(M_i^\top x) - \exp(L_i^\top x) \right|.
\]

Fix \( M \in \mathbb{R}^{N \times d} \). From (B.2) one has that

\[
\begin{align*}
& r(S^*(M, x), M, x) - \Delta(M, B) \\
& = \sum_{i \in S^*(M, x)} \exp(M_i^\top x)(w_i - r(S^*(M, x), M, x)) - \Delta(M, B) \\
& \leq \sum_{i \in S^*(M, x)} \exp(B_i^\top x)(w_i - r(S^*(M, x), M, x)) \\
& \leq \sum_{i \in S^*(M, x)} \exp(B_i^\top x) \left( w_i - (r(S^*(M, x), M, x) - \Delta(M, B)) \right).
\end{align*}
\]
APPENDIX B. PROOF OF RESULTS IN CHAPTER 3

139

Since $\Delta(M, B) < r(S^*(M, x), M, x)$, the above and (B.2) imply that

$$r(S^*(M, x), B, x) \geq r(S^*(M, x), M, x) - \Delta(M, B).$$  \hfill (B.3)

Inverting the roles of $M$ and $B$ above one gets

$$r(S^*(B, x), M, x) \geq r(S^*(B, x), B, x) - \Delta(B, M).$$

Combining these two equations one obtains

$$r(S^*(M, x), B, x) \geq r(S^*(B, x), B, x) - \Delta(B, M) - \Delta(M, B).$$

From this one concludes that, if $\Delta(M, B) + \Delta(B, M) < \delta(B, x)$, then $S^*(M, x) = S^*(B, x)$. The above implies that

$$\{S^*(B, x) \neq S^*(M, x)\} \subseteq \{\Delta(M, B) + \Delta(B, M) \geq \xi\}.$$  \hfill (B.4)

The result follows from noting that

$$\Delta(B, M) + \Delta(B, M) \leq \sum_{i \in S^*(B, x) \cup S^*(M, x)} 2w_i \left| \exp(M_i^\top x) - \exp(B_i^\top x) \right|. \hfill \square$$

**Proof of Lemma 10** Let us begin proving the result for $S = S^*(\hat{B}^n, x)$. Define $\varepsilon := \varepsilon/(2\|w\|_{\infty})$.

$$P\left\{ \Theta^n \left( S^*(\hat{B}^n, x), x, \varepsilon \right) \right\} \leq \sum_{i \in S^*(\hat{B}^n, x)} P\left\{ T_i(\tau(n), y) < n \text{ for some } y \in E^{n-1}(i) \right\} + \sum_{i \in S^*(\hat{B}^n, x)} P\left\{ \left| \exp(\beta_i^\top x) - \exp(\hat{\beta}_i^\top x) \right| > \varepsilon, T_i(\tau(n), y) \geq n, y \in E^{n-1}(i) \right\} \hfill (B.4)$$

$$\leq (a) C \sum_{y \in \mathcal{X}} \sum_{\tau(n-1)} \{ \sum_{\tau(n-1)} 1 \{ X_t = y \} < dn \} + Cc_1(\varepsilon) \exp(-c_2(\varepsilon)n)$$

$$\leq (b) \sum_{y \in \mathcal{X}} \exp \left( -2 \frac{(G(y)(\tau(n) - \tau(n-1)) - dn)^2}{\tau(n) - \tau(n-1)} \right) + \tilde{K}_1(\varepsilon) \tau(n)^{-c_2(\varepsilon)}, \hfill (B.5)$$

where (a) follows from Lemma 8 and the fact that $x \in \text{span} \left\{ E_{x^k(t)-1}(i) \right\}$ for all $i \in S^*(\hat{B}_t, x)$, for all $x \in \mathcal{X}$, and (b) follows from noting that $\kappa \log \tau(n) < n + 1$, for $n > \bar{n}$, and
from a direct application of Hoeffding’s inequality to the first term on the right-hand-side.

By construction one has that
\[ \tau(n) - \tau(n-1) \geq \tau(n)(1 - \exp(-2/\kappa)), \]
for \( n > \bar{n} \). With this, one can group terms above to obtain
\[ P \{ \Theta^n \left( S^*(\hat{B}^n, x), x, \epsilon \right) \} \leq \tilde{K}_2(\epsilon)\tau(n)^{-c_2(\epsilon)\kappa}, \]
for all \( n \geq \tilde{n} \), for some finite constant \( \tilde{n} \).

For \( \Theta^n(S^*(B, x), x, \epsilon) \), the argument above holds only if \( x \in \text{span} \left\{ E^{n-1}(i) \right\} \) for all \( i \in S^*(B, x) \), which is required to apply Lemma 8. Next we prove that is indeed the case, with high probability. For \( i \in S^*(B, x) \) we have
\[ P \{ x / \notin \text{span} \left\{ E^{n-1}(i) \right\} \} \leq P \{ w_i \leq r(S^*(\hat{B}^{n-1}, x), \hat{B}^{n-1}, x) \} \leq P \{ \Theta^{n-1}(S^*(\hat{B}^{n-1}, x), x, \xi/2) \} \leq \tilde{K}_3(\xi/2)\tau(n)^{-c_3(\xi/2)\kappa}, \] (B.6)
were (a) follows from (B.3) and the definition of \( \xi^2 \) and (b) follows from the fact that \( \tau(n-1) \geq \tau(n)\exp(-2/\kappa) \) for \( n \geq \bar{n} \). Using this result, we have that for \( x \in \mathcal{X} \),
\[ P \{ \Theta^n \left( S^*(B, x) \right), x, \epsilon \} \leq \sum_{i \in S^*(B, x)} P \left\{ \left| \exp(\beta_i^T x) - \exp((\hat{\beta}^n)^T x) \right| > \epsilon, x \in \text{span} \left\{ E^{n-1}(i) \right\} \right\} + \sum_{i \in S^*(B, x)} P \left\{ x / \notin \text{span} \left\{ E^{n-1}(i) \right\} \right\} \leq \tilde{K}_2(\epsilon)\tau(n)^{-c_2(\epsilon)\kappa} + C\tilde{K}_3(\xi/2)\tau(n)^{-c_3(\xi/2)\kappa}, \]
where the last inequality follows from (B.6) and the arguments leading to (B.5). The results follows from defining \( K_\mathcal{X}(\epsilon) := \tilde{K}_2(\epsilon) + C\tilde{K}_3(\xi/2) \) and \( c_\mathcal{X}(\epsilon) := \min \{ c_2(\epsilon), c_3(\xi/2) \} \). \( \square \)

\footnote{Lemma 9 defines \( \xi \) as the minimum optimality gap across profiles.}
Appendix C

Proof of Results in Chapter 4

C.1 Proofs for Section 4.4

Proof of Proposition 1. For any \( e \geq 0 \), by the definition of \( a_n(e) \) and \( \lambda_n(e) \), one has that for any assortment \( a \) in \( a_n(e) \),

\[
\sum_{i \in a} \nu_i(p_n,i)(p_i - c_i - \lambda_n(e)) = \lambda_n(e)(1 + e).
\]

Proof of part i.) Consider now any \( e, e' \) such that \( 0 \leq e' < e \). It is necessarily the case that for all \( \lambda \) in a neighborhood of \( \lambda_n(e) \),

\[
\max_{A_n \in S_n} \left\{ \sum_{j \in A_n} \nu_j(p_j - c_j - \lambda_n(e)) \right\} > \lambda_n(e)(1 + e').
\]

Since both \( \lambda \mapsto \max_{A_n \in S_n} \left\{ \sum_{j \in A_n} \nu_j(p_j - c_j - \lambda) \right\} \) and \( \lambda \mapsto \lambda(1 + e') \) are continuous in \( \lambda \), one has that

\[
\max_{A_n \in S_n} \left\{ \sum_{j \in A_n} \nu_j(p_j - c_j - \lambda) \right\} > \lambda(1 + e')
\]

for all \( \lambda \) in a neighborhood of \( \lambda_n(e) \). Noting that the left-hand-side above is decreasing in \( \lambda \) and the right-hand-side is increasing in \( \lambda \), it is necessarily the case that \( \lambda_n(e') > \lambda_n(e) \). This completes the proof of part i.).

Proof of part ii.) Fix \( e, e' \) such that \( 0 \leq e' < e \) and define \( \Delta \lambda := \lambda_n(e') - \lambda_n(e) \), which is positive by part i.). Let \( a' \) be any best response to an assortment with attractiveness \( e' \), i.e., \( a' \in a_n(e') \). Then, recalling the definition of the \( \theta_i \)'s in (4.5) and the discussion that
followed, it is necessarily the case that
\[
\theta_i(\lambda_n(e')) \geq \theta_j(\lambda_n(e')) \text{ for any } i \in a', j \in \mathcal{N}_n \setminus a'. \tag{C.1}
\]
Also, since \(\theta_i(\lambda) = \nu_i(p_n,i)(p_{n,i} - c_i - \lambda)\) in the case of exclusive products, one has that for all \(a \in a_n(e)\), for any \(i \in a\), \(j \in \mathcal{N}_n \setminus a\),
\[
\theta_i(\lambda_n(e')) + \nu_i(p_{n,i}) \Delta \lambda = \theta_i(\lambda_n(e)) \geq \theta_j(\lambda_n(e)) = \theta_j(\lambda_n(e')) + \nu_j(p_{n,j}) \Delta \lambda. \tag{C.2}
\]
Combining \(\text{(C.1)}\) and \(\text{(C.2)}\), we conclude that for any pair \((a, a') \in a_n(e) \times a_n(e')\), for all \(i\) in \(a \setminus a'\) and \(j\) in \(a' \setminus a\), \(\nu_i(p_{n,i}) \geq \nu_j(p_{n,j})\).

For any given \(\lambda \in \mathbb{R}_+\), all assortments \(a\) in \(\mathcal{S}^\lambda\) consist only of products with non-negative values of \(\theta_i(\lambda)\). Since \(\theta_i(\lambda)\) is strictly decreasing in \(\lambda\) for all products, higher values of \(\lambda\) translate into fewer products with non-negative \(\theta_i(\lambda)\). Thus, the cardinality of the assortments in \(\mathcal{S}^\lambda\) is non-increasing in \(\lambda\), i.e., \(|a| \leq |a'|\) for any \(a \in \mathcal{S}^\lambda\) and any \(a' \in \mathcal{S}^{\lambda'}\), for \(\lambda > \lambda'\). Since \(\lambda_n(e') > \lambda_n(e)\), one has that \(|a'| \leq |a|\), for any \(a'\) in \(a_n(e')\) and \(a\) in \(a_n(e)\). This in turns implies that
\[
|a' \setminus a| = |a'| - |a \cap a'| \leq |a| - |a \cap a'| = |a \setminus a'|.
\]
Observing that \(\nu_i(p_{n,i}) > 0\) for all \(i \in \mathcal{N}_n\), one has that
\[
\sum_{i \in a \setminus a'} \nu_i(p_{n,i}) \geq |a \setminus a'| \min \{\nu_i(p_{n,i}) : i \in a \setminus a'\} \geq |a' \setminus a| \max \{\nu_i(p_{n,i}) : i \in a' \setminus a\} \geq \sum_{i \in a' \setminus a} \nu_i(p_{n,i}).
\]
This, in turn, implies that
\[
E_n(a) = \sum_{i \in a \cap a'} \nu_i(p_{n,i}) + \sum_{i \in a \setminus a'} \nu_i(p_{n,i}) \geq \sum_{i \in a' \cap a} \nu_i(p_{n,i}) + \sum_{i \in a' \setminus a} \nu_i(p_{n,i}) = E_n(a'),
\]
for all \(a\) in \(a_n(e)\) and \(a'\) in \(a_n(e')\). This concludes the proof.

**Proof of Theorem 7** For any retailer \(n\), consider the set \(\{E_n(A) : A \in \mathcal{A}_n\}\) of all possible attraction levels corresponding to assortments, and denote those levels by \(e_1 < e_2 < \ldots < \)
$e_{kn}$. Let $Z_n = \{e_1, \ldots, e_{kn}\}$ denote the ordered set of those levels. In addition, for any attraction level $e^m$ offered by firm $m$, let $Y_n(e^m) = \{E_n(a) : a \in a_n(e^m)\}$ denote the set attractiveness levels corresponding to best responses to $e^m$. Finally, let $Y(e^1, e^2) = (Y_1(e^2), Y_2(e^1))$ denote the correspondence from $Z_1 \times Z_2$ into $Z_1 \times Z_2$. Now, note that $Z_1 \times Z_2$ is a non-empty complete lattice and that Proposition 1(ii) implies that $Y(e^1, e^2)$ is a non-decreasing correspondence. These two facts, in conjunction with the fixed point result of Topkis (1998), Theorem 2.5.1, imply that $Y(e^1, e^2)$ admits a fixed point in $Z_1 \times Z_2$. Selecting the assortments that correspond to the attractiveness levels associated with this fixed points yields an equilibrium in assortment decisions and the proof is complete.

We now comment on the fact that the reasoning above extends to the case of an arbitrary number of retailers. Indeed, when there are $N \geq 2$ retailers, one can define $Y(e^1, \ldots, e^N) = (Y_1(e^N - e^1), \ldots, Y_N(e - e^N))$, with $e := \sum_{i=1}^N e_i$. By using Proposition 1 one can prove that $Y(e^1, \ldots, e^N)$ is a non-decreasing correspondence from $Z_1 \times \ldots \times Z_N$ into itself, where $Z_n$ is defined as in the two-retailer case. Since $Z_1 \times \ldots \times Z_N$ is a non-empty complete lattice one can again use the fixed point result of Topkis (1998), Theorem 2.5.1 to establish existence of an equilibrium in assortment decisions.

Proof of Proposition 2. Following the argument in the proof of Theorem 7, Topkis (1998) Theorem 2.5.1 also yields that the set of fixed points of $Y(e^1, e^2)$ is a nonempty complete lattice relative to $\leq$ (component-wise). Hence there exist a fixed point $(e_1, e_2)$ such that $e_n \leq e'_n$, $n = 1, 2$, for all fixed points $(e'_1, e'_2)$ of $Y(\cdot)$. From proposition 1(i) and noticing that fixed points of $Y(\cdot)$ map to assortment equilibria, retailer $n$ prefers an equilibrium involving the attractiveness pair $(e_1, e_2)$ as it minimizes $e_m$. This applies for both $n$, thus both retailers prefer the same equilibrium.

We now comment on the fact that the reasoning above extends to the case of an arbitrary number of retailers. Indeed, following the reasoning on the proof of Theorem 7, Topkis (1998) Theorem 2.5.1 establishes the existence of a fixed point $(e_1, \ldots, e_N)$ such that $e_n \leq e'_n$, $n = \{1 \ldots, N\}$, for all fixed points $(e'_1, \ldots, e'_N)$ of $Y(\cdot)$, where $N$ is the number of retailers. The result would then follow from the fact that $\sum_{k \neq n} e_k \leq \sum_{k \neq n} e'_k$ for all fixed points $(e'_1, \ldots, e'_N)$ of $Y(\cdot)$, hence retailer $n$ prefers $(e_1, \ldots, e_N)$ and this applies to all $n$. □
Proof of Theorem 8. For \( n = 1, 2 \) and \( e \geq 0 \), let \( R_n(e) \) denote the set of attractiveness levels corresponding to best responses of retailer \( n \) when retailer \( m \) offers an assortment with an attractiveness of \( e \). That is, \[ R_n(e) := \{ E_n(a) : a \in a_n(e) \} \]

We observe that the number of fundamentally different equilibria is bounded above by \( \sum_{e \in \mathcal{E}_2} |R_1(e)| \). We next provide a bound on this sum. Define \( \mathcal{E}_2 := \{ e \in \mathcal{E}_2 : |R_1(e)| > 1 \} \). Let \( k \) denote the cardinality of \( \mathcal{E}_2 \) and let us denote the elements of \( \mathcal{E}_2 \) by \( e_1 < \ldots < e_k \).

One has that
\[
\sum_{e \in \mathcal{E}_2} |R_1(e)| = \sum_{e \in \mathcal{E}_2} |R_1(e)| + \sum_{e \in \mathcal{E}_2 \setminus \mathcal{E}_2} |R_1(e)| = \sum_{j=1}^k |R_1(e_j)| + |\mathcal{E}_2 \setminus \mathcal{E}_2|.
\]

Note that for any pair \((e_i, e_j)\) with \( i \neq j \), part ii.) of Lemma 1 implies that
\[
|R_n(e_i) \cap R_n(e_j)| \leq 1.
\]

In addition, the latter result, in conjunction with the fact that \( |R_1(e_j)| > 1 \) for all \( j \in \{1, \ldots, k\} \) implies that \( |R_1(e_j) \cap R_1(e_i)| = 0 \) for any \( j < i + 1 \) and \( i, j \) in \( \{1, \ldots, k\} \). Hence \( \sum_{j=1}^k |R_1(e_j)| \leq |\mathcal{E}_1| + |\mathcal{E}_2| - 1 \). We deduce that
\[
\sum_{e \in \mathcal{E}_2} |R_1(e)| \leq |\mathcal{E}_1| + |\mathcal{E}_2| - 1 + |\mathcal{E}_2 \setminus \mathcal{E}_2| = |\mathcal{E}_1| + |\mathcal{E}_2| - 1.
\]

This completes the proof. \( \Box \)

Proof of Proposition 3. We prove that each condition ensures existence of an equilibrium separately.

i.) We assume without loss of generality that \( C \leq |S| \), that \( r = 1 \) and that the products are indexed so that \( \nu_1(p_1) \geq \nu_2(p_2) \geq \ldots \geq \nu_{|N|}(p_{|N|}) \).

Let \( A_1 = \{1, \ldots, C\} \) consist of the set of \( C \) best products. Let \( A_2 \) denote the best response of retailer 2 to \( A_1 \). We establish that \((A_1, A_2)\) is always an equilibrium.
Step 1. We first establish that in the configuration \((A_1, A_2)\), the profit generated by retailer 1 satisfies \(\lambda_1 \leq 1/2\). We argue by contradiction. Suppose for a moment that \(\lambda_1 > 1/2\). Let \(E(A_n)\) denote the attractiveness of assortment \(A_n\), i.e., \(E(A_n) = \sum_{i \in A_n} \nu_i (1 - (1/2)\mathbf{1}\{i \in A_m\})\). The condition \(\lambda_1 > 1/2\) can be rewritten as

\[
\frac{E(A_1)}{1 + E(A_1) + E(A_2)} > 1/2.
\]

This implies that \(E(A_1) > 1 + E(A_2)\). Using the latter, we would have that

\[
\lambda_2 = \frac{E(A_2)}{1 + E(A_1) + E(A_2)} < \frac{1}{2} \frac{E(A_2)}{1 + E(A_2)} \overset{(a)}{\leq} \frac{1}{2} \frac{\sum_{i=1}^C \nu_i}{\sum_{i=1}^C \nu_i},
\]

where \((a)\) follows from the facts that \(E(A_2) \leq \sum_{i=1}^C \nu_i\) and \(x \mapsto x/(1 + x)\) is increasing on \([0, +\infty)\). However, this is a contradiction with the fact that \(A_2\) is a best response to \(A_1\). Indeed, given that retailer 1 offers \(A_1\), retailer 2 can generate at least

\[
\frac{1}{2} \frac{\sum_{i=1}^C \nu_i}{\sum_{i=1}^C \nu_i}
\]

by simply offering products \(\{1, ..., C\}\). We conclude that it must be the case that \(\lambda_1 \leq 1/2\).

Step 2. We now establish that no retailer has any incentive to deviate from \((A_1, A_2)\).

Case 1: \(A_2 \cap A_1 = \emptyset\). In such a case, we have that for retailer 1, \(\theta_i(\lambda_1) = (r - \lambda_1)\nu_i \geq r\nu_i/2 \geq r\nu_j/2\) for all \(i \in A_1, j \in A_2\) and hence retailer 1 has no incentive to deviate from \(A_1\) given that retailer 2 offers \(A_2\). Hence \((A_1, A_2)\) is an equilibrium.

Case 2: \(A_2 \cap A_1 \neq \emptyset\). For all products \(i \in A_1 \setminus A_2\), note that \(\theta_i(\lambda_1) = (r - \lambda_1)\nu_i \geq \theta_j(\lambda_1)\) for all \(j \in S \setminus A_1\).

Consider now \(i \in A_1 \cap A_2\) and suppose that there is a product \(j \notin A_1\) such that, \(\theta_j(\lambda_1) > \theta_i(\lambda_1) = r\nu_i/2\). If \(j \notin A_2\), then this would imply that \(\theta_i(\lambda_2) = r\nu_i/2 \geq \theta_j(\lambda_2)\). However, \(\lambda_1 \geq \lambda_2\) and hence \(\theta_j(\lambda_1) \leq \theta_j(\lambda_2) \leq r\nu_i/2 = \theta_i(\lambda_1)\), which is a contradiction.

If \(j \in A_2\), then we would have \(\theta_j(\lambda_1) = r\nu_j/2 > \theta_i(\lambda_1) = r\nu_i/2\), i.e., \(\nu_j > \nu_i\) or \(j < i\). However, since \(i \in A_1\), one must have \(i \leq C\) and this would imply that \(j < C\), which contradicts the fact that \(A_1 = \{1, ..., C\}\) by construction.

We conclude that in case 2, retailer 1 has no incentive to deviate from \(A_1\) and \((A_1, A_2)\) is an equilibrium.

The above establishes the existence of an equilibrium when condition \(i.)\) is satisfied.
APPENDIX C. PROOF OF RESULTS IN CHAPTER 4

ii.) Suppose now that $C_n \geq |S|, n = 1, 2$ and that $p_{1,i} = p_{2,i}$ for all $i \in \mathcal{N}$. We exhibit an equilibrium by restricting attention to cases where both retailers offer all common products. Noting that any best response to $A_m$ will include all common products, retailer $n$ solves for an attraction level of $e_m$ of $A_m$

$$\max_{A \in \mathcal{N}_n \setminus \mathcal{N}_m} \left\{ \sum_{i \in A \setminus A_m} \nu_i(p_i - c_i - \lambda_n) \right\} + \sum_{i \in \mathcal{N}_1 \cap \mathcal{N}_2} \nu_i(p_i - c_i)/2 \geq \lambda_n(1 + e_m).$$

Let $a'_n(e_m)$ denote the set of optimal assortments. One needs to find a pair of assortments $(A_1, A_2)$ such that $A_1 \in a'_1(E_2(A_2))$ and $A_2 \in a'_2(E_1(A_1))$, which reduces to a problem of assortment selection with exclusive products. Existence of such a pair can be proven as in proof of Theorem 1. □

C.2 Proofs for Section 4.5

Proof of Lemma 3. One can establish that problem (4.2) is equivalent to the following problem

$$\max \lambda$$

s.t. $\sup_{A \in \mathcal{S}_n, p_n} \left\{ \sum_{i \in A \setminus A_m} (p_{n,i} - c_i - \lambda) \nu_i(p_{n,i}) \right\}$

$$+ \sum_{i \in A \cap A_m} \left( \delta_{n,i}(p_n,p_m)(p_{n,i} - c_i) - \lambda \right) \nu_i(\min\{p_{n,i}, p_{m,i}\}) - \lambda \sum_{i \in A_m \setminus A} \nu_i(p_{m,i}) \right\} \geq \lambda,$$

in the sense that optimal values for both problems are equal, and a sequence of decisions achieves profits that converge to the supremum in (4.2) if and only if such a sequence achieves profits that converge to the supremum of the left-hand-side of (C.3) when $\lambda$ is equal to the supremum value. Now, the left-hand-side of (C.3) can be written as

$$\max_{A \in \mathcal{S}_n} \left\{ \sup_{p_n} \left\{ \sum_{i \in A \setminus A_m} (p_{n,i} - c_i - \lambda) \nu_i(p_{n,i}) \right\} + \sum_{i \in A \cap A_m} \left( \delta_{n,i}(p_n,p_m)(p_{n,i} - c_i) - \lambda \right) \nu_i(\min\{p_{n,i}, p_{m,i}\}) - \lambda \sum_{i \in A_m \setminus A} \nu_i(p_{m,i}) \right\}. $$
It is possible to verify that the inner-supremum above is attained by any sequence converging to \( p^*_n,i \) (strictly) from below, as defined in (4.9). Thus, the expression above is equal to

\[
\max_{A \in S_n} \left\{ \sum_{i \in A \setminus A_m} \tilde{\nu}_i e^{-\lambda n} + \sum_{i \in A \cap A_m} (p^*_n,i - c_i - \lambda)\nu_i(p^*_n,i) - \lambda \sum_{i \in A \setminus A} \nu_i(p_{m,i}) \right\}.
\]

The result follows by substituting the above in (C.3).

**Proof of Lemma 4.** We establish the result by contradiction. Suppose that \((A_1, p_1, A_2, p_2)\) is an equilibrium such that \(|A_1| < C_1\).

Let \( \lambda_1 \) denote the profits achieved by retailer 1. Select \( i \) in \( N_1 \setminus A_1 \) (which is non-empty since \( C_1 \leq |N_1| \) by assumption). Note that, through (4.12), it is necessarily the case that \( \theta_i(\lambda_1) \geq 0 \). The latter follows from the fact that \( p^*_{1,i} \) is chosen to maximize \( \nu_i(p-c_i-\lambda_1) \) for \( p \in [c_i, p_{2,i}] \) (note that it is always the case that \( p_{2,i} \geq c_i \) in any equilibrium).

**Case 1:** \( \theta_i(\lambda_1) > 0 \). In such a case, retailer 1 could strictly increase its profits by adding product \( i \) to its assortment, contradicting that \((A_1, p_1, A_2, p_2)\) is an equilibrium.

**Case 2:** \( \theta_i(\lambda_1) = 0 \). In such a case, analyzing (4.12), it must be that \( i \in A_2 \setminus A_1 \). In addition, the fact that \( \theta_i(\lambda_1) = \max_{c_i \leq p \leq p_{m,i}} \nu_i(p-c_i-\lambda_1) + \lambda_1 \nu_i(p_{2,i}) \geq \nu_i(p_{2,i})\nu_i(p_{2,i}-c_i) \) implies that \( p_{2,i} = c_i \). Retailer 2 could strictly increase its profits by removing product \( i \) from its assortment. This, again, contradicts that \((A_1, p_1, A_2, p_2)\) is an equilibrium.

We conclude that in any equilibrium, both retailers necessarily offer full-capacity assortments and the proof is complete.

**Proof of Proposition 4.** Lemma 4 implies that retailers offer full-capacity assortments and these need to be in the set of \( C_n \)-popular assortments for retailer \( n = 1,2 \). Hence, the attractiveness of assortments is uniquely defined for any equilibrium, implying that all equilibria are equivalent. Suppose that \((\lambda_1, \lambda_2)\) is a solution to (4.15), then \( \lambda_n \) solves (4.14) when retailer \( m \) sets \( p_{m,i} := c_i + \frac{1}{\alpha} + \lambda_m \), and selects \( A_m \in \tilde{P}_m(N_m, C_m) \), hence \( A_n \in \tilde{P}_n(N_n, C_n) \) with \( p_{n,i} := c_i + \frac{1}{\alpha} + \lambda_n \) is a best response to retailer \( m \)'s action. We conclude \((A_1, p_1, A_2, p_2)\) is an equilibrium. We complete the proof by showing that (4.15) always has a unique solution.
Let \( \Lambda_n \) denote the unique solution to \( V_n e^{-\alpha \lambda_n} = \lambda_n \) and note that the profit achieved by firm \( n \) always lies in \([0, \Lambda_n]\). Let \( Y_n(\lambda_m) \in [0, \Lambda_n] \) denote the unique solution to \( V_n e^{-\alpha \lambda_n} = \lambda_n + \alpha V_m e^{-\alpha \lambda_m} \) and define \( Y(\lambda_1, \lambda_2) := (Y_1(\lambda_2), Y_2(\lambda_1)) \). By the implicit function theorem, \( Y(\lambda_1, \lambda_2) \) is differentiable and
\[
\frac{\partial Y(\lambda_1, \lambda_2)}{\partial \lambda_m} = \frac{Y_n(\lambda_m) \alpha^2 V_m e^{-\lambda_m} \alpha}{(1 + \alpha V_m e^{-\lambda_m} \alpha)(1 + \alpha Y_n(\lambda_m))}.
\]
From the above we observe that \( 0 \leq J_Y(\lambda_1, \lambda_2) < 1 \) for all \((\lambda_1, \lambda_2) \in [0, \Lambda_1] \times [0, \Lambda_2] \) where \( J_Y(\cdot) \) denotes the Jacobian of \( Y(\cdot) \). We deduce that \( Y(\cdot) \) is a contraction from \([0, \Lambda_1] \times [0, \Lambda_2] \) into itself and hence it admits a unique fixed point (see Judd (1998), Theorem 5.4.1). The proof is completed by noting that \((\lambda_1, \lambda_2)\) solves (4.15) if and only if it is also a fixed point of \( Y(\cdot) \).

We now comment on the fact that the reasoning above extends to the case of an arbitrary number of retailers. Indeed, when there are say, \( N \geq 2 \) retailers, firms will still offer full-capacity assortments in equilibrium and again one can reduce the search for equilibria to settings where each firm selects a constant margin across products, or equivalently a value \( \lambda_n \). Now, the best response of firm \( n \) would be to select the unique value of \( \lambda_n \) satisfying
\[
V_n e^{-\alpha \lambda_n} = \lambda_n \left( 1 + \alpha \sum_{j=1, j\neq n}^{N} V_j e^{-\lambda_j} \alpha \right).
\]
One can now define \( Y_n(\lambda_{-n}) \) as the above solution and \( Y(\lambda_1, \ldots, \lambda_N) = (Y_1(\lambda_{-1}), \ldots, Y_N(\lambda_{-N})) \), where \( \lambda_{-n} = (\lambda_1, \ldots, \lambda_{n-1}, \lambda_{n+1}, \ldots, \lambda_N) \). \( Y \) is differentiable on \([0, \Lambda_1] \times \ldots \times [0, \Lambda_N] \) and such that
\[
\|J_Y(\lambda_1, \ldots, \lambda_N)\|_\infty = \max_{n=1, \ldots, N} \sum_{\ell=1}^{N} \left| \frac{\partial Y_n(\lambda_{-n})}{\partial \lambda_\ell} \right| = \max_{n=1, \ldots, N} \frac{\alpha^2 \sum_{\ell=1, \ell\neq n}^{N} V_\ell e^{-\lambda_\ell} \alpha}{(1 + \alpha \sum_{j=1, j\neq n}^{N} V_j e^{-\lambda_j} \alpha)(1 + \alpha Y_n(\lambda_{-n}))} < 1.
\]
On obtains that \( Y \) is a contraction and hence admits a unique fixed point. \( \square \)

**Proof of Lemma 5.** Assume, without loss of generality, that \(|\mathcal{N}_1| > C_1\) and suppose \((A_1, p_1, A_2, p_2)\) is an equilibrium.
Suppose that $A_1 \cap A_2 \neq \emptyset$. Necessarily, $p_{1,i} = p_{2,i} = c_i$ for all $i \in A_1 \cap A_2$ (since otherwise, one of the retailers could benefit by an infinitesimal decrease in price for product $i$). Let $\lambda_n$ be retailer $n$’s profit level in this equilibrium. Problem (4.11) is equivalent to
\[
\max_{A \in S_1} \left\{ \sum_{i \in A} \theta_i(\lambda_1) \right\} \geq \lambda_1 \left( 1 + \sum_{i \in A_2} \nu_i(p_{2,i}) \right).
\]
From (4.12), one has that $\theta_i(\lambda_1) > 0$ for all $i \in N_1 \setminus A_2$, and since $p_{1,i} = c_i$ for $i \in A_1 \cap A_2$, one concludes that $\theta_i(\lambda_1) = 0$ for such products. Therefore the expression above cannot be maximized at $A = A_1$, since retailer 1 could substitute $i$ with a product in $N_1 \setminus A_2$ that s/he does not currently offer. This contradicts that $(A_1, p_1, A_2, p_2)$ is an equilibrium. We deduce that one must have $A_1 \cap A_2 = \emptyset$

Now, suppose $A_n \cap N_m \neq \emptyset$ for both $n = 1, 2$. Let $j_n$ denote the product in $N_1 \cap N_2$ that provides the highest profit contribution to retailer $n$. That is
\[
j_n = \arg\max \left\{ \nu_j(p_{n,j}^*)(p_{n,j}^* - c_j) : j \in A_n \cap N_m \right\}.
\]
Note that $j_1 \neq j_2$ since $A_1 \cap A_2 = \emptyset$. Fix $n$ in $\{1, 2\}$ and assume, without loss of generality, that $\nu_{j_m}(p_{m,j_m}^*)(p_{m,j_m}^* - c_{j_m}) \geq \nu_{j_n}(p_{n,j_n}^*)(p_{n,j_n}^* - c_{j_n})$. This corresponds to a case where product $j_m$ contributes a higher amount to the profits of retailer $m$ than what product $j_n$ contributes to the profit of retailer $n$. We next show that retailer $n$ offering $j_m$ instead $j_n$ constitutes a profitable deviation. Indeed, we have
\[
\theta_{j_m}(\lambda_n) = \nu_{j_m}(p_{n,j_m}^*)(p_{n,j_m}^* - c_{j_m}) + (\nu_{j_m}(p_{n,j_m}^*) - \nu_{j_m}(p_{n,j_m}^*)) \lambda_n
\]
\[
\geq \nu_{j_m}(p_{m,j_m}^*)(p_{m,j_m}^* - c_{j_m})
\]
\[
\geq \nu_{j_n}(p_{n,j_n}^*)(p_{n,j_n}^* - c_{j_n}) > \theta_{j_n}(\lambda_n),
\]
where $(a)$ follows from the fact that $p_{n,i}^*$ is selected to maximize $\nu_i(p)(p - c_i - \lambda_n)$ for $p \in [c_i, p_{n,i}^*]$. The above contradicts the optimality of $A_n$ in maximizing the left-hand-side of (4.11) (since retailer $n$ could increase its profits by offering product $j_m$ instead of $j_n$ and pricing it just below $p_{n,j_m}^*$). This contradicts that $(A_1, p_1, A_2, p_2)$ is an equilibrium. Hence, necessarily $A_n \cap N_m = \emptyset$ for some $n = 1, 2$. This concludes the proof. \hfill \Box
APPENDIX C. PROOF OF RESULTS IN CHAPTER 4

Proof of Theorem 9. From Lemma 5 one has that equilibria must be contained in $\mathcal{CE}_1 \cup \mathcal{CE}_2$ when $|N_n| > C_n$ for some $n \in \{1, 2\}$. Equivalence of the equilibria in this setting follows directly from the statement in part i.), that we prove next.

Part i.) Consider a candidate $(A_1, p_1, A_2, p_2)$ in $\mathcal{CE}_1 \cap \mathcal{CE}_2$. It is necessarily the case that $A_n \subset N_n^m$, for $n = 1, 2$. Then, from (4.12) one has that $\theta_i(\lambda) = \tilde{\nu}_i e^{-\lambda \alpha}$ for all $i$. Therefore the popular set $A_n$ maximizes the left-hand-side of (4.11) when retailer $m$ offers $A_m$, for $n = 1, 2$. Hence $(A_1, p_1, A_2, p_2)$ is a equilibrium.

Now, suppose $\mathcal{CE}_n \setminus \mathcal{CE}_m \neq \emptyset$ for $n = 1, 2$. We will show that one cannot find equilibria in both sets. We argue by contradiction. Suppose that there are two elements $(A_1^n, p_1^n, A_2^n, p_2^n) \in \mathcal{CE}_n \setminus \mathcal{CE}_m$ for $n = 1, 2$ that constitute equilibria. For each element, let $(\lambda_n^1, \lambda_n^2)$ denote the corresponding equilibrium profits.

Let $\jmath$ be the index of the most attractive product in $N_1 \cap N_2$ (i.e., $\jmath = \arg\max\{\tilde{\nu}_i : i \in N_1 \cap N_2\}$, and $\jmath_n$ be the index of the least attractive product in $A_n^m$ (i.e., $\jmath_n = \arg\min\{\tilde{\nu}_i : i \in A_n^m\}$).

Case I: $A_n^1 \cap N_m \neq \emptyset$ for $n = 1, 2$. In such a case, it must be that $\tilde{\nu}_{\jmath} \geq \tilde{\nu}_{\jmath_n}$, for $n = 1, 2$.

Suppose $\lambda_2^1 \leq \lambda_1^1$. Let us analyze the $\theta_i$ values associated with $\jmath$ and $\jmath_n$ for retailer 2. From (4.12) one has that

$$\theta_{\jmath} (\lambda_2^1) = \tilde{\nu}_{\jmath} e^{-\lambda_2^1 \alpha} + \lambda_2^1 \tilde{\nu}_{\jmath} e^{-\lambda_2^1 \alpha} > \tilde{\nu}_{\jmath_n} e^{-\lambda_2^1 \alpha} = \theta_{\jmath_n} (\lambda_2^1).$$

This implies that retailer 2 would benefit from substituting product $\jmath_n$ by product $\jmath$ in its assortment (this would strictly increase the quantity being maximized on the left-hand-side of (4.11)). However, this contradicts the fact that $(A_1^1, p_1^1, A_2^1, p_2^1)$ is an equilibrium.

Suppose $\lambda_2^1 > \lambda_1^1$. Then, recalling the discussion that followed Proposition 4 it is necessarily the case $V_2^1 > V_1^1$. This, in conjunction, with the inequalities $V_2^2 \geq V_2^1$ and $V_1^1 \geq V_1^2$ implies, again through the discussion that followed Proposition 4, that $\lambda_2^2 > \lambda_1^2$. One can then use the argument above to obtain a contradiction with the fact that $(A_1^2, p_1^2, A_2^2, p_2^2)$ is an equilibrium.

Case II: $A_n^1 \cap N_m = \emptyset$ for some $n = 1, 2$. Suppose that it is the case for retailer 1. Since we assume that $(A_1^1, p_1^1, A_2^1, p_2^1)$ is an equilibrium, then it must be the case that $A_2^1$ is contained in $\overline{P}_2(N_2, C_2)$. In addition, given that $(A_1^1, p_1^1, A_2^1, p_2^1)$ belongs to $\mathcal{CE}_1 \setminus \mathcal{CE}_2$, one must have
A^1_2 \cap N_1 \neq \emptyset. However, this contradicts the fact that A^1_2 \subset N^1_2, which holds by construction. We deduce that (A^1_1,p^1_1,A^1_2,p^1_2) cannot be an equilibrium in this case.

**Part ii.) a.** By Lemma 4 one has that the only possible equilibrium involves A_n = N_n for n = 1, 2. Also, one must have that p_{n,j} = c_j for all j in N_1 \cap N_2 and n = 1, 2. From (4.10), in equilibrium one must have that

\[ \sum_{i \in N_n \setminus N_m} \tilde{\nu}_i e^{-\lambda_n \alpha} = \lambda_n \left( 1 + \sum_{i \in N_n \cap N_m} \alpha \tilde{\nu}_i e + \sum_{i \in N_m \setminus N_n} \tilde{\nu}_i e^{-\lambda_m \alpha} \right), \]  

(C.4)

for n = 1, 2. The arguments presented in the proof of Proposition 4 are valid to prove that (N_1,p_1,N_2,p_2) constitutes the unique possible equilibrium, where

\[ p_{n,i} = \begin{cases} \lambda_n + c_i + \frac{1}{\alpha} & i \in N_n \setminus N_m, \\ c_i & \text{otherwise}, \end{cases} \]

for n = 1, 2, and (\lambda_1, \lambda_2) is the unique solution to (C.4). In addition, it is clear that (N_1,p_1,N_2,p_2) is indeed an equilibrium.

**Part ii.) b.** Suppose (A_1,p_1,A_2,p_2) is an equilibrium. Lemma 4 implies that \(|A_n| = C_n\) for n = 1, 2. Hence A_n \cap (N_1 \cap N_2) \neq \emptyset for n = 1, 2. However this contradicts the second result of Lemma 5. We deduce that no equilibrium can exist.

The proof is complete. \(\square\)

**Proof of Proposition 5** Suppose first that a \cap N_n \neq \emptyset for some a in \(\widetilde{P}_m(N_m,C_m)\). Consider a candidate \((A^n_1,p^n_1,A^n_2,p^n_2)\) in \(\mathcal{C}E_n \setminus \mathcal{C}E_m\). By similar arguments as in the proof of Theorem 9 part i. (Case I), one has that retailer m will find a profitable deviation (by substituting the least attractive product in A^n_m (in terms of the \(\tilde{\nu}_i\) by product \(\overline{j}\)). Hence, no element of \(\mathcal{C}E_n \setminus \mathcal{C}E_m\) can be an equilibrium.

Suppose now that no popular assortment a in \(\widetilde{P}_m(N_m,C_m)\) overlaps with N_n. This implies that A m \cap N_n = \emptyset for each candidate \((A^m_1,p^m_1,A^m_2,p^m_2)\) in \(\mathcal{C}E_m \setminus \mathcal{C}E_n\). For such a candidate to be an equilibrium, it must be the case that A m belongs to \(\widetilde{P}_n(N_n,C_n)\). In addition, the fact that A m \cap N_m = \emptyset implies that the candidate belongs to \(\mathcal{C}E_m \cap \mathcal{C}E_n\). This contradicts the fact that the candidate belongs to \(\mathcal{C}E_m \setminus \mathcal{C}E_n\). This completes the proof. \(\square\)
C.3 Minimum Margins Setting

In this section, we consider the setting where the retailers solve problem (4.2) subject to the additional constraints that \( p_{n,i} \geq c_i + r_i, \ i = 1, \ldots, S \), i.e., there is a minimum profit margin \( r_i > 0 \) for product \( i \). In particular, retailer \( n \) solves

\[
\sup_{A_n \in \mathcal{S}_n, \ p_{n,i} \geq c_i + r_i, \ i \in A} \{ \pi_n(A_n, p_n, A_m, p_m) \}.
\]

We next outline the key steps that can be taken to analyze equilibrium behavior.

**Best response correspondence and equivalent problem formulation.** Proceeding in a similar fashion as in Section 4.5, one can reduce Problem (C.5) to the following assortment-only problem

\[
\begin{align*}
\sup & \quad \lambda \\
\text{s.t.} & \quad \max_{A \in \mathcal{S}_n} \left\{ \sum_{i \in A \setminus A_m, \ i \in A \cap A_m : p_{m,i} \neq r_i + c_i} \left( p_{n,i}^* - c_i - \lambda \right) \nu_i(p_{n,i}^*) \ight. \\
& \quad \left. + \sum_{i \in A \cap A_m : p_{m,i} = r_i + c_i} 1/2 (r_i - \lambda) \nu_i(r_i + c_i) - \lambda \sum_{i \in A \setminus A} \nu_i(p_{m,i}) \right\} \geq \lambda \tag{C.7}
\end{align*}
\]

where

\[
p_{n,i}^* := \max \left\{ r_i + c_i, \min \left\{ \frac{1}{\alpha} + \lambda + c_i, p_{m,i} \right\} \right\}, \quad i \in A,
\]

and where it is assumed that \( p_{m,i} := \infty \) if \( i \) is not in \( A_m \). Note that maximizing prices are set in a similar form as in the original setting, with a correction to account for the minimum margin constraint.

In particular, any assortment that solves (C.6) will enable retailer \( n \) to achieve the supremum in (C.5) when prices converge to \( p_{n,i}^* \) (strictly) from below.

As in Sections 4.4 and 4.5, one can analyze the set of best response assortments for a given \( \lambda, \mathcal{S}^\lambda \) (i.e., the solution set to the inner maximization in (C.7)). For \( i \) in \( \mathcal{N}_n \) define

\[
\theta_i(\lambda) := \begin{cases} 
\nu_i(p_{n,i}^*) (p_{n,i}^* - c_i - \lambda) & \text{if } i \in \mathcal{N}_n \setminus A_m, \\
\nu_i(p_{n,i}^*) (p_{n,i}^* - c_i - \lambda) + \lambda \nu_i(p_{m,i}) & \text{if } i \in \mathcal{N}_n \cap A_m, \ p_{m,i} \neq r_i + c_i \\
\nu_i(r_i + c_i) r_i/2 & \text{if } i \in \mathcal{N}_n \cap A_m, \ p_{m,i} = r_i + c_i.
\end{cases}
\]

One can solve for \( \mathcal{S}^\lambda \) by selecting the products with highest positive values for \( \theta_i(\lambda) \) (with a maximum of \( C_n \) products). Note that the expression is similar to (4.12), except that now, one needs to account for the fact that one cannot price below \( c_i + r_i \).
**APPENDIX C. PROOF OF RESULTS IN CHAPTER 4**

**Discussion of the case of exclusive products.** When $N_1 \cap N_2 = \emptyset$, problem (C.6) can be further simplified (similarly to what was done in Section 4.5.1), which enables one to establish a parallel result to Proposition 1 on the monotonicity of best response assortment attractiveness. In turn, this would allow one to prove existence and ordering of equilibria. It is worth noting that even though the $w$-based ranking does not depend directly on $A_m$, as in Section 4.5.1 it does now depend on the specific value of $\lambda$. This means that $\{S^\lambda : \lambda \in \mathbb{R}\}$ may contain assortments with different attractiveness, opening the possibility of having multiple equilibria that yield different profit levels. This behavior is more in line with the case of exogenous prices, analyzed in Section 4.4.1.

More specific results can be obtained for example, when the minimum margins are equal across products for both retailers (i.e., a minimum profit margin $r_n$ is imposed on all products $i \in N_n$, $n = 1, 2$). In such a case, it is possible to establish that there exists a unique equilibrium (in the fundamentally different sense), by building on the analyses in Sections 4.4.1 and 4.5.1.

When retailers have both exclusive and common products, while one can still prove that both retailers will necessarily offer full-capacity assortments in equilibrium, following the lines of the proof of Lemma 4 it is not possible anymore to ensure that their offerings will not overlap in equilibrium. In addition, in sharp contrast with the results of Section 4.5.2 there will be cases where the number of equilibria grows exponentially with the display capacities. As an illustration, one can verify that the setup of Example 4 in Section 4.4.2 (where one allows retailers to select prices) falls into that category when there is a uniform minimum margin on all products that is greater than $(4/3)\alpha$. 