Assortment and Price Optimisation under non-conventional customer choice models

Alvaro Flores

A thesis submitted for the degree of
Doctor of Philosophy at
The Australian National University

February 2020

Except where otherwise indicated, this thesis is my own original work.

Alvaro Flores
25 February 2020

to my wife, family and friends.

Acknowledgements

First and foremost, I would like to thank my whole supervisory panel. Thanks to Pascal Van Hentenryck and Gerardo Berbeglia for providing their time, guidance and insightful discussions that made my Ph.D. journey extremely enjoyable, challenging and exciting. Without their encouragement, support and patience, I would not be in this privileged position. To Lexing Xie, who warmly accepted me as one of her students, and was one of the kindest professors that I had the luck to engage with. Her understanding on how to conduct research and give the right tip at the right moment, even when not closely related with my research topic itself, helped me greatly to finish my candidature. To Richard Weber for being a good mentor through these years. I would also like to extend my gratitude to Guillermo Gallego, who inspired me quite a lot in my visit to HKUST. His approach to research and the way he looked at problems was inspiring, and pushed me to embrace knowledge in a different light, with more responsibility and focus. To Data 61 (Phil Kilby et al.), thanks for providing the funding and an amazing place to work full of enriching interactions. I learnt a lot from every one of you. To all of you, thank you so much.

I have deep gratitude to my friends, who for things that I'm yet to fully understand, bare with me through this whole process. Including many of them living here in Australia and making this adventure way more enjoyable. Their support, and willingness to hear my research ramblings is something that helped a lot during my candidature. Thanks are also due to my friends at Hivery, who supported me in the latest stage of my candidature, by asking interesting questions and providing a platform where this research can be taken into reality. Many thanks as well to **CONICYT**[1] for funding this research too. Thanks are also due to the *Proofed* team, who kindly helped proofread the Introduction, Conclusion and Future Work sections.

I am very grateful also to all the members in my family. I carry every single one of them in the deepest place of my heart. Their motivation, unconditional support and even visit to this wonderful country, made my journey possible and enjoyable. It would not be possible to be in this position without the early teachings and experiences that I had the luck to live with them, and everything that I learnt being part of that awesome family.

Finally, I would like to express my deepest gratitude to my wife, Daniela, for being the most amazing human being that I've known so far. You allocated so much confidence in this endeavour, that it was not an option to fail. Your smile and joyful attitude towards everything is the engine that keeps me going every day. Without you, I would not be the person that I'm now. Thank you so much.

[1]BECAS CHILE, Doctorado Becas Chile/2016 - 72160300

Declaration of Authorship

This thesis is an account of research undertaken between September 2015 and August 2019 at the College of Engineering & Computer Science, Research School of Computer Science, The Australian National University, Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Alvaro Alexis Flores Rios
February, 2020

Abstract

Nowadays, extensive research is being done in the area of revenue management, with applications across industries. At the heart of this area lie two fundamental problems: (1) the assortment problem, which aims to find the subset of products a company should offer to maximise revenue, provided customers follow a certain model of choice, and (2) the pricing problem, which aims to determine the prices a company should offer to best meet its objectives (profit maximisation, revenue maximisation, market share maximisation, etc.), based on how customers might respond to different prices and the interaction between price and intrinsic features of each product.

Most models studied satisfy the following property: when the offered set is enlarged, the probability of selecting a specific product decreases. This property is called *regularity* in the literature. However, customer behaviour often shows violations of this property, such as the decoy effect, where adding extra options sometimes leads to a positive effect for some products, whose probabilities of being selected increase in relative terms compared to other products (for example, including a medium size popcorn slightly cheaper than the large one, with the purpose of making the latter more attractive by comparison). We study two models of customer choice where regularity violations can be accommodated, and show that the assortment optimisation problem can still be solved in polynomial time.

First we analyse the Sequential Multinomial Logit (SML) model, where products are partitioned into two levels to capture differences in attractiveness, brand awareness, and/or visibility in the market. When a consumer is presented with an assortment of products, she first considers products in the first level and, if none of them are purchased, products in the second level are considered. This model is a special case of the Perception-adjusted Luce model (PALM) recently proposed by Echenique et al. [2018]. It can explain many behavioural phenomena such as attraction, compromise, similarity effects, and choice overload, which cannot be explained by the Multinomial Logit (MNL) model or any discrete choice model based on random utility. We show that the seminal concept of revenue-ordered assortment sets, which contain the optimal assortment under the MNL model, can be generalised to the SML model. More precisely, we show that all optimal assortments under the SML are revenue-ordered by level, a natural generalisation of revenue-ordered assortments that contains, at most, a quadratic number of assortments. As a corollary, the assortment optimisation problem under the SML model is solvable in polynomial time.

Secondly, we study the two-stage Luce model (2SLM), which is a discrete choice model introduced by Echenique and Saito [2018], that generalises the standard MNL model. The 2SLM does not satisfy the Independence of Irrelevant Alternatives (IIA)

property or regularity, and to model customer behaviour, each product has an intrinsic utility and uses a dominance relation between products. Given a proposed assortment S, consumers first discard all dominated products in S before using the MNL model on the remaining products. As a result, the model can capture behavior that cannot be replicated by any choice model that belongs to the RUM class. We show that the assortment problem under the 2SLM is solvable in polynomial time. Moreover, we prove that the capacitated assortment optimisation problem is NP-hard and presents polynomial-time algorithms for the cases where (1) the dominance relation is attractiveness correlated and (2) its transitive reduction is a forest. The proofs exploit a strong connection between assortments under the 2SLM and independent sets in comparability graphs.

The third and final contribution is an in-depth study of the pricing problem under the 2SLM. We first note that changes in prices should be reflected in the dominance relation if the differences between the resulting attractiveness are large enough. This is formalised by solving the joint assortment and pricing problem under the Threshold Luce model, where one product dominates another if the ratio between their attractiveness is greater than a fixed threshold. In this setting, we show that this problem can be solved in polynomial time.

Preface

The contributions made in this thesis are listed in the following peer-reviewed publications, conferences or in preparation papers. Each chapter is associated to one or more publications

Chapter 2:

- **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment optimization under the Sequential Multinomial Logit Model. *European Journal of Operational Research*, Volume 273, Issue 3, 16 March 2019, Pages 1052-1064

Chapters 3 and 4:

- **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment and Pricing Optimization Under the Two-Stage Luce model. *Submitted to Operations Research (13th of April of 2019).* Under Review.
- **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment and Pricing Optimization Under the Two-Stage Luce model. *Presented at the Informs Revenue Management & Pricing Conference, Stanford (Informs RM&P), 7th of June 2019.*

Appendix A:

- Van Hentenryck, P.; **Flores, A.**; Berbeglia, G., 2017. Trial-Offer Markets with Continuation. *Presented in the 21st International Federation of Operational Research Societies Conference, Quebec (IFORS 2017).*

Alvaro Alexis Flores Rios
Research School of Computer Science
The Australian National University
25 February 2020

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Revenue management (RM) refers to the managerial practice of modifying the availability and the prices of products to maximise revenue or profit, usually using IT-supported means. The origin of this discipline extends back to the 1970s, following the deregulation of the US airline market. A large volume of research has been devoted to this area over the last 45 years, with successful results across many industries such as airlines, hospitality, retail, online marketplaces, etc. [McGill and van Ryzin, 1999; Kök et al., 2005; Vulcano et al., 2010; Strauss et al., 2018].

The term revenue management was originally coined in the airline industry as low variable costs and high fixed costs led to a point where maximising revenue was almost equivalent to maximising profit. However, this does not currently limit the purpose of RM to the sole objective of profit maximisation. It is also used to drive other business decisions with other potential objectives, such as market share maximisation, customer engagement, churn minimisation, etc.

In the last decades, there has been an increase in the use of technology to inform managerial decisions, due to an increase in competition, availability of better prediction models, and more computing power. This has created an environment where it is paramount for companies to understand customer behaviour and leverage their understanding to maximise profit, or attract more customers through careful and planned modification of their sales strategies. This trend will not end any time soon, since the interaction between competition and the use of technology goes both ways: competition forces the use of technology to find ways to maintain profitability, and technology itself increases competition severity, by pushing practitioners to adopt said practices to stay in the market.

It is important to note that even when a firm can perfectly model customer behaviour, a lot of work still needs to be done to take advantage of this knowledge, since trying all potential alternatives of products and their prices is impractical and computationally prohibitive. Thus, to maximise its revenue, the firm needs a strategy on how to handle at least two key problems: product variety and product pricing. These two problems lie at the heart of RM theory and practice, and are known as the assortment problem and the pricing problem respectively.

The assortment problem consists of selecting the subset of products a company should offer to maximise revenue. For example, consider a retailer with limited space allocated to mobile phones. If more than 500 mobile phones can be acquired through distributors (in various combinations of brands and sizes) and the mobile phone aisle can only fit 50 phones on the shelves, the store manager has to decide the subset of products a company should offer based on product costs and customer preferences. The potential number of assortments grows exponentially with the number of products available. Therefore, developing algorithms to find optimal or near-optimal solutions that exploit customer choice structure is an important topic of research in RM.

The pricing problem aims to determine the prices a company should offer to best meet its objectives (profit maximisation, revenue maximisation, market share maximisation, etc.), while considering how customers might respond to different prices and the interaction between price and intrinsic features of each product. The complexity of this problem will vary greatly depending on the customer choice model, and so will the corresponding techniques needed to solve it.

When customers face multiple products, with their corresponding prices, they choose their preference based on many factors such as product features, brand awareness, prices, product placement, social influence, among others. Utility maximisation theory poses that a customer will always select the product that give her the most value. Since there might be some uncertainty in the valuation function of each product, probabilistic models are better suited to explain customer choice. The customer selects each product with some probability distribution that varies according to the assumptions we use for the utility and random uncertainty functions. The family of choice models that arises from this framework of utility maximisation are called *Random Utility Models* (RUMs) [Block and Marschak, 1960]. This class of discrete choice models has been extensively studied and used in several industries, and its relevance in terms of describing customer choice and aiding better decision making is well documented in the RM literature [Bodea and Ferguson, 2014; Chiang et al., 2007; Strauss et al., 2018]. One of the most studied models of to this category is the *Multinomial Logit* (MNL) model, which is extremely attractive to use, mainly due to the following reasons: (1) it is easy to understand because the formula that describes customer choice probabilities is intuitive, (2) it is easy to estimate given the concavity of the log-likelihood function [McFadden, 1978], and (3) the associated optimisation problems under this model of choice tend to have tractable solutions. For example, to find the optimal assortment under the MNL model, it is enough to consider the family of *revenue-ordered assortments*, which is basically setting a threshold price and including all available products whose price is more than said threshold. Thus, is easy to show that only a linear number potential assortments needs to be considered to find a revenue maximising assortment [Talluri and Van Ryzin, 2004]. The specific details of all these benefits will be discussed in the next sections within this chapter.

Despite the apparent generality of the RUM class and associated efficient estimation methods, certain phenomena require considering models that do not belong to this customer class to be able to capture them [Tversky and Kahneman, 1974; Huber et al., 1982]. One of the reasons behind these limitations is a shared property that all RUMs satisfy: the regularity property. Regularity states that the probability of choosing an alternative cannot increase if the offer set is enlarged. Although this property seems intuitive, psychologists and marketing researchers have carried out several controlled experiments [Simonson and Tversky, 1992; Tversky and Kahneman, 1974; Huber et al., 1982] where customers in certain contexts violate it. One prominent example is the *decoy effect* [Herne, 1997], where the impact of adding a similar but inferior alternative into the current assortment ends up increasing the probability of products that are now easily seen as *better* in comparison to this *decoy* alternative (a well-known example is the inclusion of medium-size fries with a price tag close to the large-size alternative). It is easy to show that every RUM must satisfy this property, and thus, in contexts where regularity does not hold, it is impossible to fit a RUM regardless of how sophisticated the estimation procedure is.

In Section 1.2.1, we present some recent models that attempt to bridge this gap, and go beyond random utility. Among them, we describe two variants of the MNL model that allow regularity violations, providing more flexibility to accommodate substitution patterns that cannot be captured by any RUM, and will be studied in detail in this thesis. First, we consider the *Sequential Multinomial Logit* (SML) [Echenique et al., 2018] model, where products are partitioned into two levels, to capture differences in attractiveness, brand awareness, and/or visibility in the market. When a consumer is presented with an assortment of products, she first considers products on the first level and, if none of them are purchased, products in the second level are considered. In Chapter 2, we show that the assortment optimisation problem under the SML is polynomial-time solvable, by showing that all optimal assortments under the SML model are revenue-ordered by level, a natural generalisation of revenue-ordered assortments, which are optimal for the MNL model.

The second variant of the MNL model considered in this thesis is the *two-stage Luce model* (2SLM) [Echenique and Saito, 2018]. The 2SLM does not satisfy the Independence of Irrelevant Alternatives (IIA) property or regularity, and to model customer behaviour, each product has an intrinsic utility and uses a dominance relation between products. Given a proposed assortment S, consumers first discard all dominated products in S before using the MNL model on the remaining products. As a result, the model can capture behaviour that cannot be replicated by any discrete choice model based on random utility. In Chapter 3 we study the assortment problem under this model of customer choice, providing a polynomial time solution for it, and proving that when the number of products is bounded, the problem becomes NP-hard. We study particular cases for the latter where an optimal assortment can still be found in polynomial time, exploiting the structure of the dominance graph.

Chapter 4 provides a polynomial time solution to the pricing problem for a particular case of the 2SLM, called the *Threshold Luce model* [Echenique and Saito, 2018] where one product dominates another if the ratio between their attractiveness is greater than a fixed threshold. The solution is similar to the one for the MNL model (fixed mark-up, [Anderson et al., 1992; Hopp and Xu, 2005; Gallego and Stefanescu, 2009; Besbes and Sauré, 2016]) but with some price adjustments in both extremes of the attractiveness spectrum, to avoid low attractiveness products being dominated by higher attractiveness ones.

In the reminder of this chapter, we provide a review of the literature on the main topics that will help understanding the rest of this thesis. In Section 1.2, we provide a review of customer choice models, with focus on the distinction of models belonging to RUM or not. In Section 1.3), we take a look at the results of assortment optimisation for many of the models discussed in Section 1.2. We finish this chapter with a review of pricing optimisation results in Section 1.4. Across all these sections, we focus primarily on results closely related to this thesis, and in particular, with emphasis over the MNL model [Luce, 1959; McFadden, 1978] and its known variants, since the models we study share some structural similarities by the fact of being extensions of it. The more interested reader can be referred to for several in-depth reviews of the literature concerning RM [McGill and van Ryzin, 1999; Chiang et al., 2007; Strauss et al., 2018], discrete choice models with their corresponding estimation methods [Berbeglia et al., 2018], assortment optimisation [Kök et al., 2005], and pricing optimisation [Bodea and Ferguson, 2014].

1.2 Customer Choice models

Discrete choice models, which have been studied for more than 50 years, are essential to understand and make predictions about choices made by individuals in different settings. For example, choice models are used to estimate customer purchases in several markets, such as retail, air travel, and accommodation. These sales estimates obtained with choice models are key components of RM, where the availability of products as well as their prices are optimised to maximise expected profits.

The first choice model used in RM was the independent demand model. Under this model, the probability of purchasing a product is considered to be independent of other products on offer. This assumption holds when the products are not substitutes of each other, for example, when they participate in different markets. This assumption clearly fails when a customer selects among competing alternatives, and substitution patterns need to be taken into consideration. This assumption makes the models easy to study but precludes the ability to include simple substitution patterns, such as the fact that when a product is removed from an assortment, its demand can potentially be captured by other products (spilled demand). On the other hand, the addition of a new alternative to an assortment can cause cannibalisation of the demand of other products in the assortment, or cause demand to flow from pre-existing products to this new alternative.

Independent demand models were prevalent until the early 2000s, mostly in quasi-monopolistic settings where different alternatives were fenced off, e.g. for the airline industry, where different fare types were tied with specific restrictions (such as cancellation policy, the requirement of early purchase, seat type, etc.). This was designed to appeal to different customer segments. In recent years, with technological advances that provide information in real-time through the internet and online agencies, as well as the rise of low-cost airlines, the independent demand model is no longer an accurate depiction of customer choice due to higher visibility of alternatives and increasing competition.

Since then, the RM community has studied more sophisticated discrete choice models that incorporate substitution behaviour. The benefits of incorporating substitution patterns have shown an improvement in demand predictions [Talluri and Van Ryzin, 2004; Newman et al., 2014; van Ryzin and Vulcano, 2015; Ratliff et al., 2008; Vulcano et al., 2010].

An important class of customer choice models is the RUMs, originally proposed by Thurstone [1927], where there is a deterministic component to the utility of each product, and a random one. Customers draw a sample of the joint distribution of utilities and select the one with the highest utility. In Block and Marschak [1960], the authors showed that the RUM class is equivalent to distributions over strict preference rankings.

Arguably the most prominent model belonging to RUM is the MNL model [Luce, 1959], also known as the Luce model, which is widely used in discrete choice theory. Since the model was introduced by Luce [1959], it was applied to a wide variety of demand estimation problems arising in transportation [McFadden, 1978; Catalano et al., 2008], marketing [Guadagni and Little, 1983; Gensch, 1985; Rusmevichientong et al., 2010], and revenue management [Talluri and Van Ryzin, 2004; Rusmevichientong et al., 2010]. One of the reasons for its success stems from its small number of parameters (one for each product). This allows for simple estimation procedures that generally avoid overfitting problems even when there is limited historical data [McFadden, 1974]. However, as mentioned in Section 1.1, one of the flaws of the MNL is the property known as the IIA, which states that the ratio between the probabilities of choosing elements x and y is constant regardless of the offered subset. This property does not hold when products cannibalise each other or are perfect substitutes [Ben-Akiva and Lerman, 1985; Debreu, 1960; Anderson et al., 1992].

Despite the IIA property, the MNL model is widely used. Indeed, for many applications the mean utility of a product can be modelled as a linear combination of its features. If the features capture the mean utility associated with each product, then the error between the utilities and their means may be considered as independent noise, and the MNL model emerges as a natural candidate for modelling customer choice. In addition, the MNL model parameters can be estimated from customer choice data, even if limited data is available [Ford, 1957; Negahban et al., 2012], because the associated estimation problem has a concave log-likelihood function [McFadden, 1974] and it is possible to measure how good the fitted MNL model approximates the data [Hausman and McFadden, 1984]. Moreover, it is possible to improve model estimation when the IIA property is likely to be satisfied [Train, 2003].

To overcome the IIA limitation in settings where it is likely to not be satisfied, more complex choice models have been proposed in the literature such as the Nested Logit (NL) model [Williams, 1977]. Under the NL model, alternatives are organised into nests, where products in the same nests are closer substitutes of each other compared to products in different nests. Under the NL model, costumer selection occurs in two stages. First, she selects one of the nests or decides to not make a purchase based on the alternatives, and if a nest was chosen, then she selects one of the products within that nest.

It is sensible to think that the customer base might have heterogeneous preferences, and being able to model these differences will lead to a better exploitation of this knowledge to maximise profit. Hence, a natural extension of the MNL model is to consider multiple customer classes, each assumed to follow the MNL model. This model is known as *Finite-mixture Logit* model or *Latent Class* model [Greene and Hensher, 2003]. The mixed MNL (MMNL) model represents an MNL model where utilities are drawn from a cumulative distribution, where the Finite-mixture Logit is a special case when the distribution has finite support. Interestingly enough, under mild regularity conditions, McFadden and Train [2000] shows that any choice model based on random utility maximisation can be arbitrarily closely approximated by a Mixed Multinomial Logit model.

Gallego et al. [2015] proposed the General Attraction model (GAM), where the probabilities of choosing a product depend on all products (not only the offered subset as in the MNL model). This involves a shadow attraction value associated with each product that influences the choice probabilities when the product is not offered.

To model customer behaviour in a more graphical way, [Blanchet et al., 2016] proposed the *Markov Chain* model, where the process of selecting an alternative can be described as a Markov chain, having one state per alternative, and one state representing the outside option. In this model, customers have an initial distribution over alternatives and whenever faced with an assortment, they arrive according to this distribution. If the product is available, they purchase it, if not, they transition to another alternative, with a probability that approximates the substitution rate between those products, when the first one is not available. This process is repeated until she selects a product from the assortment, or she falls in the outside option state, in which case she leaves without purchasing anything from the assortment. This model is fully characterised by the initial distribution and the transition probabilities. Moreover, Berbeglia [2016] proves that the Markov Chain model belongs to RUMs. Recently, Şimşek and Topaloglu [2018] proposed an Expectation-Maximization algorithm to estimate model parameters.

The *Exponomial* model [Daganzo, 1979; Alptekinoğlu and Semple, 2016] incorporates negatively skewed distributions of customer utilities. In this model, choice probabilities are written as a linear combination of exponential terms, thus obtaining this name. This model is well suited to describe markets where customers tend to not overpay if they are well informed about alternatives and their corresponding prices. In Alptekinoğlu and Semple [2016], the authors show that the log-likelihood function is concave, and the parameter estimation can be computed using maximum likelihood estimates.

All the models previously presented share two common features. First, they impose an a priori specific structure for choice probabilities. In the last decade, studies have been trying to estimate a general RUM without pre-imposing any structure on the choice probabilities. In Farias et al. [2013], the authors used constraint sampling to find a distribution over strict preference rankings, which achieves the purpose since distributions over strict preference rankings was shown to be equivalent to RUM in Block and Marschak [1960]. Later, in van Ryzin and Vulcano [2015], a column-generation procedure was proposed to compute the maximum likelihood estimates. The cost of going away from models that have a particular structure is that model estimation gets more complicated due to increasing time complexity, overfitting issues due to all the flexibility, and in some cases non-identifiability (since a particular RUM may have more than one description). However, these costs come with an associated reward: not imposing a specific structure allows modelling more nuanced substitution behaviour that cannot be described by parametric RUMs. To quantify this benefit, Farias et al. [2013] tracked the sales data of 14 products (a range of small SUVs) from a major US automaker over 16 months. The authors compared the performance of an MNL model, a mixed MNL model, and a non-parametric RUM by training these models and comparing their predictions against never-seen-before data. Compared to the MNL and the mixed MNL, their non-parametric model improved prediction accuracy by roughly 20%. Which in turn can be translated to up to 10% increase in revenues. More recently, Farias et al. [2017] developed a non-parametric approach to model customer choice and applied it to a large US fashion retailer to increase purchases by customers. Their solution, that includes an assortment optimisation procedure, has increased the retailer's revenue by around 7%.

Secondly, all these models satisfy *regularity*, and although this notion seems natural and intuitive, it is well known that it is sometimes violated by individuals [Debreu, 1960; Tversky, 1972a,b; Tversky and Simonson, 1993; Herne, 1997]. Recently, there have been efforts to develop discrete choice models that go beyond random utility and therefore can explain complex choice behaviours such as the violation of regularity. The following section offers a review of why this is relevant and presents the models that we will cover in this thesis.

1.2.1 Beyond Random Utility

In the last few years, there has been an increasing effort to propose and study models that go beyond random utility. The reason behind this increasing interest lies in the flexibility to capture customer behaviour, and the ability to accommodate substitution patterns and effects that cannot be explained by any model within the RUM class.

Examples of such effects include attraction, [Doyle et al., 1999], the compromise effect [Simonson and Tversky, 1992], the similarity effect [Debreu, 1960; Tversky, 1972b], and the paradox of choice (also known as choice overload) [Iyengar and Lepper, 2000; Schwartz, 2004; Haynes, 2009; Chernev et al., 2015].

The attraction effect stipulates that, under certain conditions, adding a product to an existing assortment can increase the probability of choosing a product in the original assortment. We briefly describe two experiments of this effect. Simonson and Tversky [1992] considered a choice among three microwaves x, y and z. Microwave y is a Panasonic oven, perceived as a good quality product. z is a more expensive version of y. Product x is an Emerson microwave oven, perceived as a lower quality product. The authors asked a set of 60 individuals ($N = 60$) to choose between x and y; they also asked another set of 60 participants ($N = 60$) to choose among x, y and z. They found out that the probability of choosing y increases when product z is shown. This is a direct violation of regularity, which states that the probability of choosing a product does not increase when the choice set is enlarged, as described by McCausland and Marley [2013]. Another demonstration of the attraction effect was carried by Doyle et al. [1999], who analysed the choice behaviour of two sets of participants ($N = 70$ and $N = 82$) in a grocery store in the UK, varying the choice set of baked beans. To the first group, they showed two types of baked beans: Heinz baked beans and a local (and cheaper) brand called Spar. In this setting, the Spar beans option was chosen 19% of the time. To the second group, the authors introduced a third option: a more expensive version of the local brand Spar. After adding this new option, the cheap Spar baked beans option was chosen 33% of the time. It is worth highlighting that the choice behaviour in these two experiments cannot be explained by an MNL model or by any choice model based on random utility. This effect, also known as the decoy effect, was even observed in animals. In Lea and Ryan [2015], the authors showed that female túngara frogs violate regularity in the following sense: When selecting mating partners, selecting between two alternatives exhibited a clear preference of one over the other. However, the addition of a third alternative consistently reversed that preference order. Building upon this interesting result, Natenzon [2019] proposed the Bayesian probit. This model explains context-dependent choices as the optimal response of an agent facing imperfect information, to the ease of option comparison. This model can accommodate both the attraction and the compromise effect.

The compromise effect [Simonson and Tversky, 1992] captures the fact that individuals are averse to extremes, which helps products that represent a "compromise" over more extreme options (either in price, familiarity, quality, etc.). As a result, adding extreme options sometimes leads to a positive effect on compromise products, whose probabilities of being selected compared to other products increase in relative terms. This phenomenon again violates the IIA axiom of the Luce model, and the regularity axiom satisfied by all random utility models [Berbeglia and Joret, 2017].

The similarity effect is discussed in Tversky [1972b], elaborating on an example presented in Debreu [1960]: Consider x and z to represent two recordings of the same Beethoven symphony and y to be a suite by Debussy. The intuition behind the effect is that x and z jointly compete against y, rather than being separate individual alternatives. As a result, the ratio between the probability of choosing x and the probability of choosing y when the customer is shown the set $\{x, y\}$ is larger than the ratio when the customer is shown the complete set $\{x, y, z\}$. Intuitively, z takes the market share of product x, rather than the market share of product y.

Finally, choice overload effect occurs when the probability of making a purchase decreases when the assortment of available products is enlarged. To our knowledge, the first paper that shows the empirical existence of choice overload is Iyengar and

Lepper [2000]. In their experimental setup, customers are offered jams from a tasting booth displaying either 6 (limited selection) or 24 (extensive selection) different flavours. All customers were given a discount coupon for making a purchase of one of the offered jams. Surprisingly, 30% of the customers offered the *limited selection* used the coupon, while only 3% of customers offered the *extensive selection* used the coupon. Other studies of choice overload have been done on 401(k) plans [Sethi-Iyengar et al., 2004], chocolates [Chernev, 2003b], consumer electronics [Chernev, 2003a], and pens [Shah and Wolford, 2007]. For a more in depth discussion of this effect, the reader is referred to Schwartz [2004]. Readers are also referred to Chernev et al. [2015] for a review and meta-analysis of this topic.

Recently, Echenique et al. [2018] proposed the *Perception-Adjusted Luce Model* (PALM), which aims specifically to accommodate the effects just described. In Chapter 2, we study the *Sequential Multinomial Logit* (SML), which is a special case of the PALM. In the SML model, products are partitioned *a priori* into two sets, which we call levels. This product segmentation into two levels can capture different degrees of attractiveness. For example, it can model customers who check promotions/special offers first before considering the purchase of regular-priced products. It can also model consumer brand awareness, where customers first check products of specifics brands before considering the rest. Finally, the SML model can model product visibilities in a market, where products are placed in specific positions (aisles, shelves, web-pages, etc.) that induce a sequential analysis, even when all products are in sight.

While the PALM and the NL model can both be conceived as sequential choice processes, they have important differences. Probably the most important difference is that the NL model belongs to the family of RUMs[1], and therefore cannot accommodate regularity violations. On the other hand, the PALM does not belong to the RUM class and allows regularity violations as well as choice overload. In terms of the choice process, in the NL model customers first select a nest, and then a product within the nest. In the PALM, products are separated by preference levels, so when a customer is offered a set of products, she first chooses from the offered products belonging to the lowest available level, and if none of them are chosen then she selects from the next available level, and keeps repeating this process until no more levels are available or until a purchase is made.

Another phenomena that makes customer behaviour difficult to model is that customers tend to use rules to simplify decisions that are not always easy to infer. Before making a purchase decision, they often narrow down the set of alternatives to choose from, using different heuristics to make the decision process simpler. Several *consider-then-choose* models have been proposed in the literature, related with attention filters, search costs, feature filters, among others. Another reasonable way to discard alternatives, is when the difference between attractiveness is so evident, that the less attractive alternative, even when it is offered, is never picked (as in the Threshold Luce model, Echenique and Saito [2018]). Any of the heuristics mentioned

[1]This is unless nest specific parameters are greater than one, a case rarely studied in the literature.

earlier allow the consumer to restrict her attention to a smaller set usually referred in the literature as *consideration set*. This effect also proposes that an offered product might result in zero-probability choice, since it is not being considered.

Several extensions to the MNL model have been introduced to overcome the IIA property and some of its other weaknesses. They include the nested MNL model and the latent class MNL model. These models however do not handle zero-probability choices well. Consider two products a and b: The MNL model states that the probability of selecting a over b depends on the relative attractiveness of a compared to the attractiveness of b. Consider the case in which b is never selected when a is offered. Under the MNL model, this means that b must have zero attractiveness. But this would prevent b from being selected even when a is not offered in an assortment.

To overcome zero-probability choices, Masatlioglu et al. [2012] proposed a theoretical foundation for maximising a single preference under limited attention, i.e., when customers select among the alternatives that they pay attention to. To incorporate the role of attention into stochastic choice, Manzini and Mariotti [2014] proposed a model in which customers consider each offered alternative with a probability and choose the alternative maximising a preference relation within the considered alternatives. This was axiomatised and generalised in Brady and Rehbeck [2016], by introducing the concept of *random conditional choice set rule* (RCCSR), which captures correlations in the availability of alternatives. This concept also provides a natural way to model substitutability and complementarity. The paper also states that RCCSR is not a RUM.

Payne [1976] showed that a considerable portion of the subjects in his experimental setting use a decision process involving a consideration set. Numerous studies in marketing also validated a consider-then-choose decision process. In his seminal work, Hauser [1978] observed that most of the heterogeneity in consumer choice can be explained by consideration sets. He shows that nearly 80% of the heterogeneity in choice is captured by a richer model based in the combination of consideration sets and logit-based rankings. The rationale behind this observation is that first stage filters eliminate a large fraction of alternatives, thus the resulting consideration sets are composed of a few products in most of the studied categories [Belonax Jr and Mittelstaedt, 1978; Hauser and Wernerfelt, 1990]. Pras and Summers [1975] and Gilbride and Allenby [2004] empirically showed that consumers form their consideration sets by a *conjunction* of elimination rules. Furthermore, there are empirical results showing that a two-stage model including consideration sets better fits consumer search patterns than sequential models [De los Santos et al., 2012].

From a customer standpoint, the use of consider-then-choose models alleviates the cognitive burden of deciding when facing too many alternatives Tversky [1972a,b]; Tversky and Kahneman [1974]; Payne et al. [1996]. When dealing with a decision under limited time and knowledge, customers often return to screening heuristics as shown in Gigerenzer and Goldstein [1996]. Psychologically speaking, customers as decision makers need to carefully balance search efforts and opportunity costs with potential gains, and consideration sets help to achieve that goal [Roberts and Lattin, 1991; Hauser and Wernerfelt, 1990; Payne et al., 1996]. Recently Jagabathula and Rusmevichientong [2017] proposed a two-stage model where customers consider only the products that are contained within a certain range of their willingness to pay. Aouad et al. [2015] explored consider-then-choose models where each costumer has a consideration set, and a ranking of the products within it. The customer then selects the higher ranked product offered. Dai et al. [2014] considered a revenue management model where an upcoming customer might discard one offered itinerary alternative due to individual restrictions, such as time of departure. Wang and Sahin [2018] studied a choice model that incorporates product search costs, so the set that a customer considers might differ from what is being offered, depending on how *expensive* is to acquire extra information on products.

Another recently proposed extension to the MNL model is the *General Luce Model* (GLM), independently developed by Echenique and Saito [2018] and Ahumada and Ülkü [2018]. The GLM generalises the MNL model and falls outside the RUM class. In the GLM, each product has an intrinsic utility and the choice probability depends upon a dominance relationship between the products. Given an assortment S, consumers first discard all dominated products in S and then select a product from the remaining ones using the standard MNL model. Chapters 3 and 4 study this model, and provide solutions to the assortment and pricing problems.

It is worth to point out that after we posted all our results from Chapter 3 and Chapter 4 on ArXiv and submitted them for publication (first version: June 2017, second version: April 2019), the following paper was released: *Consumer Choice with Consideration Set: Threshold Luce model*, by Ruxian Wang, posted on June 1^{st} 2019. This paper studies a particular case of the General Luce model, called the Threshold Luce model, which we also studied in detail. He focused solely on the Threshold Luce model, and was able to obtain interesting results. Wang arrived to the same pricing results. He also studied price competition and show that depending on the value of the threshold, the market can have a varying number of Nash equilibria. He also showed that the assortment optimisation problem under the Threshold Luce model with heterogeneous customers is NP-hard, which we originally have as an open question, and provided estimation methods to calibrate the threshold Luce model. He showed that this model can improve the goodness of fit and prediction accuracy significantly in comparison with the MNL, which suggests the threshold effect should be taken into consideration when studying customer choice.

Recently, Berbeglia [2018] proposed the *Generalised Stochastic Preference* (GSP) model, which is a non-parametric model capable to explain easily (and exactly) well known examples that fall outside RUM. The model also has a nice intuitive interpretation: it is based on the representation of RUMs as stochastic preference, but allows some consumer types to be non-rational.

In Cattaneo et al. [2017], the authors introduce the *Random Attention Model* (RAM), which consists of a ranking of the available alternatives, and a probability function that is a distribution over consideration sets that are subsets of the offered set. Given its non-parametric nature, the RAM model, is very general: allows regularity violations, contains the RUM class and *Random Conditional Choice Set Rule* (RCCSR) [Brady and Rehbeck, 2016].

In an effort to understand the connection between some of the existing classes of choice models, Feng et al. [2017] studied the relation between RUMs, the *representative agent model* and the *semi-parametric choice model* (SCM) [Natarajan et al., 2009]. In the representative agent model, a *representative agent* makes a choice among alternatives on behalf of the entire population. The agent may choose any fractional amount of the alternatives, and to make his choice, she maximises the expected utility while rewarding diversification to some extent. On the other hand, unlike RUM where a distribution of the random utility is specified, the SCM considers a set of distributions for it. Thus, RUM can be viewed as a particular case of SCM (where the distributions set is a singleton). Despite the fact that these two models are based on different approaches, the authors made a key observation about their underlying structure: they are essentially functions from the vector of utilities, to a vector of choice probabilities, mediated by a welfare function that captures the expected utility that a customer can obtain from the choice model. Furthermore, the choice probability can be understood as the gradient of this welfare function with respect to the utilities. Building on this observation, they proposed the *Welfare-Based Choice model*. The authors showed that this model is equivalent to the representative agent model and the semi-parametric model, showing the equivalence between the two. Additionally, they showed that the welfare-based model strictly subsumes RUM when there are three or more alternatives, while they are equivalent when there are only two.

Another recently proposed model that can accommodate attraction, compromise and similarity effects (and hence, violate regularity) is the *Gradual Pairwise Comparison Rule* (GPCR) [Dutta, 2018]. This model is built on the premise that the choice procedure is primarily made through a sequence pairwise comparisons, as the evidence of eye-tracking studies of choice suggests [Russo and Rosen, 1975]. Dutta proposed that the customer performs a series of pairwise comparisons against her underlying strict rational preference, in a menu-independent manner to remove inferior alternatives from an offered set. While she can make all possible comparisons with positive probability, it is not certain. Once the sequential pairwise comparisons stops (which occurs with a possibly menu-dependent probability distribution), she selects among the non-removed alternatives with equal probability. The author showed (by construction) that this model can rationalise any Luce rule. In terms of the data that this model can rationalise, this model neither contains nor is contained by any of the models proposed in Manzini and Mariotti [2014]; Brady and Rehbeck [2016]; Cattaneo et al. [2017].

Chen and Mišić [2019], proposed a non-parametric model where each customer type is associated with a binary decision tree, representing the decision process of making a purchase based on checking the existence of specific products from the assortment. Together with a probability distribution over customer types, the authors show that this model is able to represent any customer choice model with arbitrary precision, including models that are inconsistent regularity. A recent paper [Chen et al., 2019] proposed a similar tree-based discrete choice model, and although the results are similar, in the latter case the estimation step is based on random forest, whereas in Chen and Mišić [2019] the estimation procedure is carried over by an optimisation approach based on column generation. Additionally, Chen et al. [2019] shows that the estimation procedure allows the inclusion of pricing information, and even aggregate choice data.

It is natural to wonder how much is lost in prediction accuracy by assuming that consumers follow a RUM when attempting to model choice behaviour. In Jagabathula and Rusmevichientong [2018], the authors created an efficient procedure to quantify the loss of rationality (LoR), which is defined as the cost of approximating the observed choice fractions from data, with those from the best-fitting probability distribution over rankings (which coincides with RUM). Using this methodology, they analysed real transactions of consumer packaged goods, and show that for some specific categories of products, the LoR is high. This fact showed that irrespective of the RUM that we select to approximate choice behaviour in this context, we will not be able to provide a good approximation.

Alternatively, there are other factors that can drive customers to violate regularity, and exhibit choice patterns that apparently are non-rational given only product features and customer valuation of them. Among the common explanations proposed in the literature, two of them are pervasive in online markets, and slowly making progress in retail settings: social influence and position bias. The next section will briefly explain how these two effects modify customer behaviour, and some of the attempts at characterise their impact.

1.2.2 Social Markets, Social Influence and Position Bias

To inform purchase decisions, customers often rely in what other previous customers have done before. This is especially relevant when there is no prior information about the alternatives. Websites take advantage of this fact and provide information on past consumption of their offered alternatives, or they stablish rating systems. The effect that this information has over customer behaviour is known as *wisdom of the crowd* Lorenz et al. [2011], where the opinion of the majority impacts the decision carried by individuals. This effect can cause that new users entering the market, inform their decision based on the behaviour of old users, assuming they had a strong reason or they are better informed to choose among alternatives and thus decide to copy their behaviour.

Understanding how exactly this works, and being able to predict the extent of this effect, has proven to be extremely difficult: experts constantly fail trying to predict the success of books, songs and movies Hirsch [1972]; De Vany and Walls [1999]; Caves [2000]. In De Vany and Walls [1999], the authors showed that in the movie industry, there are only a few rare and unpredictable movies that concentrate most of the ticketing revenue. This trend is practically the norm in online markets, where web pages [Broder et al., 2000], books [Newman, 2005], youtube videos [Cha et al., 2007] receive most of the attention while the majority of the available alternatives are barely noticed.

Customers are also influenced by how alternatives are presented to them. Experimental analysis using eye-tracking inspired the so-called *Cascade Models*, introduced first by Craswell et al. [2008]; Kempe and Mahdian [2008a], and showed fit experimental data better than separable models in presence of position bias. The intuition behind this model, is that users consider products in a top-to-bottom fashion, as they were presented in an ordered list, and they only look to the next product if the current one was not selected. This phenomena has been observed in online stores such as `Amazon`, `Itunes`, flight search engines, and retail stores [Lim et al., 2004].

Examples of these two phenomena can be observed in trial-offer markets, where customers can sample a product before deciding whether to buy it or not, and might be influenced by social signals, or alternative placement. The MusicLab experiment Salganik and Watts [2008] showed that social influence steer the markets towards unpredictability and inequality, where few products dominate the market, and predicting which ones is extremely complex. However, one characteristic of the MusicLab experiment, was that the songs were displayed in decreasing order of popularity. This intensifies the social signal with the addition of position bias Payne [1951], since customers tend to pay more attention to the items at the top, amplifying the *rich get richer* effect that is already in place given they are already ranked by popularity. Is then the interaction of these two factors that can lead to more predictable outcomes.

Following this line of thought, Watts [2011] suggested that when social influence affects the market, popularity prediction is nearly impossible to achieve. However, not all is lost: we can attempt an approach that he calls *measure and react*, where instead of focusing on predicting how users will behave or create rules to induce specific customer behaviour, we should observe directly the market response and react accordingly.

As a result of Watt's advice, researchers focused on designing appropriate ranking systems that can control how information is being transferred to users, and with that, control the negative effects created by social influence, showing that the unpredictability is not an inherent property of markets where customers are influenced by what other customers are doing, but rather a consequence of market design.

These results have been obtained using extensions of the MNL model [Krumme et al., 2012; Lerman and Hogg, 2014; Abeliuk et al., 2015; Van Hentenryck et al., 2016; Abeliuk et al., 2017; Maldonado et al., 2018]. One alternative is to use a myopic approach and find a ranking (the performance ranking) that maximises the expected downloads at each time step, given the current state of the market. Abeliuk et al. [2016] show that there is a polynomial time solution for this problem, with position bias and social influence. The authors also show that by using this approach, the result always profits from both position bias and social influence, as compared with a market without these effects.

Abeliuk et al. [2017] shows that using the *quality ranking* (which instead of ranking products by popularity, ranks them by decreasing order of quality) reduces significantly the unpredictability attributed to social influence. They reproduced the results in an experimental study, where participants participate in a market where in each interaction they view a list of ten science related stories, ordered in one column and were asked to pick one, and decide to recommend it if they deemed worthy of it. Participants were randomly assigned into one of four different experimental settings, that differ in the way stories are ordered, and whether social signals were displayed in the screen or not. In the setting where no social signal was available for participants, they only saw the titles and abstracts. On the other hand, when social signals were enabled, each participant had the additional information of how many recommendations each story received so far.

Trial-offer markets occupy a special place in my candidature, since they introduced me to the RM literature. One question related with this market that was often asked is: *what if customers can try more than one product before making a purchase decision?*. Following this line of inquiry, we extended some results on predictability and efficiency in Abeliuk et al. [2016, 2017] to a setting where consumers can try more than one alternative before making a purchase decision, while being influenced by social signal and position bias. In Appendix A we show that a MNL model with continuation can be reduced to a standard trial-offer market under the MNL model with different appeal and product qualities. We examine the consequences of this reduction on the performance and predictability of the market, the role of social influence, and the ranking policies (popularity ranking, performance ranking and quality ranking). For a more in depth review of trial-offer markets, the reader is referred to [Abeliuk et al., 2017; Maldonado et al., 2018].

1.3 The Assortment Problem

The assortment optimisation problem is a central problem in revenue management, where a firm wishes to offer a set of products with the goal of maximising the expected revenue. This problem has many relevant applications in brick-and-mortar retail, fashion industry, online markets, and RM in general [Kök et al., 2005]. For example, a publisher must decide which set of advertisements to show, an airline must decide which fare classes to offer on each flight, and a retailer must decide which products to show in a limited shelf space.

One of the first positive results of the assortment problem under the MNL model was obtained by Talluri and Van Ryzin [2004], where the authors showed that the optimal assortment can be found by greedily adding products to the offered assortment in the order of decreasing revenues, thus evaluating at most a linear number of subsets. Rusmevichientong et al. [2010] studied the assortment problem under the MNL but with a *capacity constraint* limiting the products that can be offered. Under these conditions, the optimal solution is not necessarily a revenue-ordered assortment but it can still be found in polynomial time.

Davis et al. [2013] showed that solutions to the assortment problem when the choice probabilities follow the GAM (which extends the MNL) can be found in polynomial time when the assortments are constrained by a set of totally unimodular constraints. Some examples of such constraints are: capacity constraints (limiting the number of products), assortment problems where display allocations must be chosen, pricing with finite menu, quality consistent pricing (where prices must follow the same order as product qualities). The author showed that all those problems can be solved as linear programs, and constraints can be combined as long as total unimodularity is preserved. Abeliuk et al. [2016] also proposed polynomial time algorithms to solve the assortment problem under the MNL model with capacity constraint and position bias, where position bias means that customer choices are affected by the positioning of the products from the assortment. The approach is different from Davis et al. [2013], since rather than solve a linear program, it shows that the number of potential assortments that needs to be considered as candidates is at most quadratic in the number of products, and offer a procedure to obtain them.

Under the Mixed Multinomial Logit model, the problem becomes NP-hard [Bront et al., 2009] and it remains NP-hard even for two customer types [Rusmevichientong et al., 2014]. A branch-and-cut algorithm was proposed by Méndez-Díaz et al. [2014]. Feldman and Topaloglu [2015] proposed methods to obtain good upper bounds on the optimal revenue. Rusmevichientong and Topaloglu [2012] considered a model where customers follow a MNL model and the parameters belong to a compact uncertainty set. The firm wants to hedge against the worst-case scenario and the problem amounts to finding the optimal assortment under these uncertainty conditions. Surprisingly, despite the uncertainty, when there is no capacity constraint the revenue-ordered strategy is still optimal in this setting.

There are also studies on how to solve the assortment problem when customers follow a Mixed Multinomial Logit model. Bront et al. [2009] showed that this problem is NP-hard in the strong sense using a reduction from the *minimum vertex cover problem* [Garey and Johnson, 1979]. Méndez-Díaz et al. [2014] proposed a branch-and-cut algorithm to solve the optimal assortment under the Mixed Logit model. An algorithm to obtain an upper bound of the revenue of the optimal assortment solution under this choice model was proposed by Feldman and Topaloglu [2015]. Rusmevichientong et al. [2014] showed that the problem remains NP-hard even when there are only two customers classes.

Another model that attracted researchers' attention is the *NL model* [Williams, 1977]. Under the NL model, products are partitioned into nests, and the selection process for a customer goes by first selecting a nest, and then a product within the selected nest. It also has a dissimilarity parameter associated with each nest that serves the purpose of magnifying or dampening the total preference weight of the nest. For the two-level NL model, Davis et al. [2014] studied the assortment problem and showed that when the dissimilarity parameters are bounded by 1 and the no-purchase option is contained in a nest of its own, the optimal assortment can be found in polynomial time. If either of these two conditions is relaxed, the resulting problem becomes NP-hard, using a reduction from the *partition problem* [Garey and Johnson, 1979]. The polynomial-time solution was further extended by Gallego and Topaloglu [2014], who showed that even if there is a capacity constraint per nest, the problem remains solvable in polynomial time. Li et al. [2015] extended this result to a d-level NL model (both results under the same assumptions over the dissimilarity parameters and the no-purchase option).

Jagabathula [2014] proposed a local-search heuristic for the assortment problem under an arbitrary discrete choice model. This heuristic is optimal in the case of the MNL model, even with a capacity constraint. Aouad et al. [2015] studied the assortment problem under several consideration sets and ranking structures, and provide a dynamic programming approach capable of returning the optimal assortment in polynomial time for families of consideration set functions originated by screening rules Hauser et al. [2009]. Wang and Sahin [2018] has studied the assortment optimisation in a context in which consumer search costs are non-negligible. The authors showed that the strategy of revenue-ordered assortments is not optimal. Another interesting model sharing similar choice probabilities to those of the PALM, is the one proposed in Manzini and Mariotti [2014] which is based on consider first and choose second process. Echenique et al. [2018] showed that the PALM and the model by Manzini and Mariotti are in fact disjoint. The assortment optimisation problem under the Manzini and Mariotti model was recently studied by Gallego and Li [2017], where they show that revenue-ordered assortments strategy is optimal.

The assortment problem was also studied under the negative exponential distribution (NED) model [Daganzo, 1979], also known as the exponomial model [Alptekinoğlu and Semple, 2016] in which customer utilities follow negatively skewed distribution. Alptekinoğlu and Semple [2016] proved that when prices are exogenous, the optimal assortment might not be revenue-ordered assortment, because a product can be skipped in favour of a lower-priced one depending on the utilities. This last result differs from what happens under the MNL and the NL model (within each nest).

Recently, Berbeglia and Joret [2017] studied how well revenue-ordered assortments approximate the optimal revenue for a large class of choice models, namely all choice models that satisfy regularity. They provide three types of revenue guarantees that are exactly tight for the RUM family. In the last few years, there has been also progress in studying the assortment problem in choice models that incorporate visibility or position biases. In these models, the likelihood of selecting an alternative depends not only on the offer set but also on the specific positions in which each product is displayed [Abeliuk et al., 2016; Aouad and Segev, 2015; Davis et al., 2013; Gallego et al., 2016].

1.4 The Pricing Problem

Multi-product price optimisation under the MNL model and the NL model has been studied since the models were introduced in the literature. One of the first results on the structure of the problem is due to Hanson and Martin [1996], where they show that the profit function for a company selling substitutable products when customers follow the MNL model is not jointly concave in price. To overcome this issue, in Song and Xue [2007] and later in Dong et al. [2009], the authors show that even when the profit function is not concave in prices, it is concave in the market share and there is a one-to-one correspondence between price and market share. Multiple studies shown that under the MNL where all products share the same price sensitivity parameter, the *mark-up* which is simply the difference between price and cost, remains constant for all products at optimality [Anderson et al., 1992; Hopp and Xu, 2005; Gallego and Stefanescu, 2009; Besbes and Sauré, 2016]. Furthermore, the profit function is also uni-modal on this constant quantity and it has a unique optimal solution, which can be determined by studying the first order conditions.

Li and Huh [2011] showed the same result for the NL model. Up to that point, all previous results assumed an identical price sensitivity parameter for all products. Under the MNL, there is empirical evidence that shows the importance of allowing different price sensitivity parameters for each product [Berry et al., 1995; Erdem et al., 2002]. There is also evidence in Börsch-Supan [1990] that restricting the nest specific parameters to the unit interval results in rejection of the NL model when fitting the data, thus recommending relaxing this assumption. The problem when relaxing this condition, is that the profit function is no longer concave on the market share, which complicates the optimisation task. In Gallego and Wang [2014], the authors considered a NL model with differentiated price sensitivities, and found that the *adjusted mark-up*, defined as price minus cost minus the reciprocal of the price sensitivity is constant for all products within a nest at optimality. Furthermore, each nest also has an *adjusted next-level mark-up* which is also invariant across nests, which reduces the original problem to a one variable optimisation problem. Additional theoretical development can be found in Rayfield et al. [2015]; Kouvelis et al. [2015] but these are restricted to the two-stage NL model. In Huh and Li [2015] some of the results were extended to a multi-stage NL model for specific settings, but also show that the equal mark-up property fails to hold in general for products that do not share the same immediate *parent* node in the nested choice structure, even when considering identical price sensitivity parameters. Li and Huh [2011] and Gallego and Wang [2014] extend to the multi-stage NL model and show that the optimal pricing solution can still be found by means of maximising a scalar function.

There are some interesting results for other models that share similarities with the MNL model, and therefore are closely related with the model that we are studying. In Wang and Sahin [2018], the authors incorporate search cost into consumer choice model. The joint assortment and pricing results are similar to the ones that we study in Section 4.1, in that many structural results that holds at optimality for their model, are also satisfied in our studied case. They show that the *quasi-same price* policy (that charges the same price for all products but one, the least attractive one) was optimal for this model. Interestingly, the joint assortment and pricing results under the Threshold Luce model has a slightly different result: The optimal pricing is a fixed price for all products, except for the most attractive and least attractive ones. This led to a situation where there are many possible prices, not just two. We note that this result was also recently and independently obtained by Wang [2019].

For the exponomial model, Alptekinoğlu and Semple [2016] shows that the optimal pricing policy allows for variable mark-ups in optimal prices that increase with expected utilities. This is apparently a consequence of the skewed distribution of consumer utilities.

1.5 Thesis Outline

The rest of this thesis is organised as follows. Chapter 2, studies the assortment optimisation problem under the SML model, where products are partitioned into two levels, to capture differences in attractiveness, brand awareness, and/or visibility in the market. When a consumer is presented with an assortment of products, she first considers products on the first level and, if none of them are purchased, products in the second level are considered. We show that the concept of revenue-ordered assortment sets, which contain the optimal assortment under the MNL model, can be generalised to the SML model. More precisely, we show that all optimal assortments under the SML are revenue-ordered by level, a natural generalisation of revenue-ordered assortments that contains, at most, a quadratic number of assortments. As a corollary, the assortment optimisation problem under the SML is solvable in polynomial time.

In Chapter 3 we study the assortment problem under the two-stage Luce model. We show that the assortment problem under the 2SLM is solvable in polynomial time. Moreover, we prove that the capacitated assortment optimisation problem is NP-hard and presents polynomial-time algorithms for the cases where (1) the dominance relation is attractiveness correlated and (2) its transitive reduction is a forest.

Chapter 4 studies the pricing problem under the two-stage Luce model. We first note that changes in prices should be reflected in the dominance relation if the differences between the resulting attractiveness are large enough. This is formalised by solving the joint assortment and pricing problems under the Threshold Luce model, where one product dominates another if the ratio between their attractiveness is greater than a fixed threshold. In this setting, we show that this problem can be solved in polynomial time.

Finally, Chapter 5 summarises the main contributions of this thesis, and explores potential avenues for future work.

Chapter 2

Sequential Multinomial Logit

This chapter is reproduced with minor changes from:

Flores, A.; Berbeglia, G; Van Hentenryck, P. 2019. Assortment optimisation under the Sequential Multinomial Logit Model. *European Journal of Operational Research*, Volume 273, Issue 3, 16 March 2019, Pages 1052-1064

In this chapter, we study the assortment optimisation problem under the *Sequential Multinomial Logit* (SML), a discrete choice model that generalizes the multinomial logit (MNL). Under the SML model, products are partitioned into two levels to capture differences in attractiveness, brand awareness and/or visibility of the products in the market. When a consumer is presented with an assortment of products, she first considers products in the first level and, if none of them is purchased, products in the second level are considered. This model is a special case of the Perception-Adjusted Luce Model (PALM) recently proposed by Echenique et al. [2018]. It can explain many behavioral phenomena such as the attraction, compromise, similarity effects and choice overload which cannot be explained by the MNL model or any discrete choice model based on random utility. In particular, the SML model allows violations to *regularity* which states that the probability of choosing a product cannot increase if the offer set is enlarged.

The chapter shows that the seminal concept of revenue-ordered assortment sets, which contain an optimal assortment under the MNL model, can be generalized to the SML model. More precisely, we prove that all optimal assortments under the SML are revenue-ordered by level, a natural generalization of revenue-ordered assortments that contains, at most, a quadratic number of assortments. As a corollary, assortment optimisation under the SML is polynomial-time solvable. This result is particularly interesting given that the SML model does not satisfy the regularity condition and, therefore, it can explain choice behaviours that cannot be explained by any choice model based on random utility.

2.1 Motivation

The *Sequential Multinomial Logit* (SML) for brevity, is a special case of the recently proposed model known as the perception-adjusted Luce model (PALM) [Echenique et al., 2018]. In the SML model, products are partitioned *a priori* into two sets, which we call levels. This product segmentation into two levels can capture different degrees of attractiveness. For example, it can model customers who check promotions/special offers first before considering the purchase of regular-priced products. It can also model consumer brand awareness, where customers first check products of specifics brands before considering the rest. Finally, the SML can model product visibilities in a market, where products are placed in specific positions (aisles, shelves, web-pages, etc.) that induce a sequential analysis, even when all the products are at sight. Our main contribution is to provide a polynomial-time algorithm for the assortment problem under the SML and to give a complete characterization of the resulting optimal assortments.

A key feature of the PALM and the SML, is their ability to capture several effects that cannot be explained by any choice model based in random utility (such as for example the MNL, the mixed MNL, the markov chain model, and the stochastic preference models). Examples of such effects include attraction [Doyle et al., 1999], the compromise effect [Simonson and Tversky, 1992], the similarity effect [Debreu, 1960; Tversky, 1972b], and the paradox of choice (also known as choice overload) [Iyengar and Lepper, 2000; Schwartz, 2004; Haynes, 2009; Chernev et al., 2015]. These effects are discussed in the next section. In particular, the SML allows for violations of regularity. There are very few analyses of assortment problems under a choice model outside the RUM class.

The decision on the first level is just as a conventional MNL, and when nothing is selected on this level, the selection process in the second level is again an MNL model that inflate the outside alternative with the attractiveness of the products already rejected on the first level. We might argue that this can be attributed to the consumer being either *overwhelmed* by the selection process so far, or that she/he is still comparing with products on the first level, and this burden carries to the second level, decreasing the overall purchasing probability. Another way to understand the reasoning behind the functional form of this model, is in settings where all products are visible but the customer still decides sequentially among a partition of the products, and the cognitive efforts of making this partition provokes the exacerbation of the outside option. One might argue that the extent of this effect should be weighted, but this variation is not explored in this Chapter. A similar model is currently being studied in Liu et al. [2018]. Here, the functional form is slightly different than the PALM, but it is also characterised by iterative applications of the MNL.

Our algorithm is based on an in-depth analysis of the structure of the SML. It exploits the concept of revenue-ordered assortments that underlies the optimal algorithm for the assortment problem under the MNL. The key idea in our algorithm is to consider an assortment built from the union of two sets of products: a revenue-ordered assortment from the first level and another revenue-ordered assortment from the second level. Several structural properties of optimal assortments under the SML are also presented.

On Section 2.3, we establish bounds on the following: any the product offered on any optimal solution, and also for an assortment considered on an optimal solution on each level. Then, we hypothesize that any optimal solution must by revenue ordered assortment by level. To demonstrate this result, in Section 2.4 we use the structural insights developed on Section 2.3, and show by contradiction that if an optimal solution is not revenue ordered assortment by level, then we can slightly modify that assortment to find another one yielding more revenue, contradicting optimality.

2.2 Problem Formulation

This section presents the sequential multinomial logit model considered in this chapter and its associated assortment optimisation problem. For the purpose of completeness, we describe the perception-adjusted Luce model (PALM) proposed by in Echenique et al. [2018]. The authors recover the role of perception among alternatives using a weak order $\succsim$. The idea is that if $x \succ y$ then x tends to be perceived before y, and whenever $x \sim y$ then x and y are perceived at the same time.

A *Perception-Adjusted Luce Model* (PALM) is described by two parameters: a weak order $\succsim$, and a utility function u. She perceives elements of an offered set $S \subseteq X$ sequentially according to the equivalence classes induced by $\succsim$ (in our representation, we called them *levels*). Each alternative is selected with probability defined by μ, a function depending on u and closely related to Luce's formula.

Definition 1. A *Perception-Adjusted Luce Model* (**PALM**) is a pair $(\succsim, u)$, where the probability of choosing a product x when offering the set S is:

$$\rho(x,S) = \mu(x,S) \cdot \prod_{\alpha \in S/\succsim : \alpha > x} (1 - \mu(\alpha, S)), \tag{2.1}$$

where:

$$\mu(x,S) = \frac{u(x)}{\sum_{y \in S} u(y) + u_0}. \tag{2.2}$$

$S/\succsim$ corresponds to the set of equivalence classes in which $\succsim$ partitions S.

Thus, the probability of choosing a product is the probability of not choosing any products belonging to the previous levels and then selecting the product according a Luce's Model considering all the offered products. We can also write the probability of not choosing any product of the assortment:

$$\rho(x_0, S) = 1 - \sum_{y \in S} \rho(y, S), \tag{2.3}$$

or equivalently, as the probability of not choosing on each one of the equivalence classes.

$$\rho(x_0, S) = \prod_{\alpha \in S/\succsim} (1 - \mu(\alpha, S)) \tag{2.4}$$

In the Perception-Adjusted Luce Model, the customer makes her selection following a sequential procedure. She considers the alternatives in sequence, following a predefined perception priority order. Choosing an alternative is conditioned to not choosing any other alternative perceived before. If none of the offered alternatives is selected, then the outside option is chosen.

Is interesting to note that setting the outside option utility to zero does not result in zero choice probability for the outside option. As explained in Echenique et al. [2018], there are two sources behind choosing the outside option in the PALM. One is the utility of the outside option, which is to the same extent as in Luce's model with an outside option. The second, and to our appreciation the one that differentiate this model from other models of choice, is due the sequential nature of choice. When a customer chooses sequentially following a perception priority order, it can happen that she checks all products in the offered set without making a choice. When this occurs, this seems to increase or bias the value of the outside option probability.

Another consequence of the functional form of the outside option probability (Equation (2.4)), is the ability to model choice overload. In addition, this model allows violations to regularity and violations to stochastic transitivity. The interested reader is referred to Echenique et al. [2018] for more details.

The *Sequential Multinomial Logit Model* (**SML** which is a special sub-case of the Perception-Adjusted Luce Model) is a discrete choice model where the probability $\rho(x, S)$ of choosing a product x in an assortment S is given by:

$$\rho(x, S) = \begin{cases} \frac{u(x)}{\sum_{y \in S} u(y) + u_0} & \text{if } x \in S_1, \\ \left[1 - \frac{\sum_{z \in S_1} u(z)}{\sum_{y \in S} u(y) + u_0}\right] \cdot \frac{u(x)}{\sum_{y \in S} u(y) + u_0} & \text{if } x \in S_2. \end{cases}$$

where u_0 denotes the intrinsic utility of the no-choice option, which has a probability

$$\rho(x_0, S) = 1 - \sum_{i \in S} \rho(i, S)$$

of being chosen.

Observe that the probability of choosing a product $x \in S_1$ (which implies that $l(x) = 1$ and $x \in S$) is given by the standard MNL formula, whereas the probability of choosing a product y that belongs to the second level is given by the probability of not choosing any product belonging to the level 1 multiplied by the probability of selecting product y according the MNL again. Note that, if all the offered products belong to the same level, this model is equivalent to the classical MNL model. The SML corresponds to PALM restricted to two levels.

Let $r : X \cup \{x_0\} \rightarrow \mathbb{R}^+$ be the revenue function which assigns a per-unit revenue to each product and let $r(x_0) = 0$. We use $R(S)$ to denote the *expected revenue* of an assortment S, i.e.,

$$R(S) = \sum_{x \in S} \rho(x, S) \cdot r(x). \tag{2.5}$$

The assortment optimisation problem under the SML consists in finding an assortment S^* that maximises R, i.e.,

$$S^* = \underset{S \subseteq X}{\operatorname{argmax}} R(S). \tag{2.6}$$

We use R^* to denote the maximum expected revenue, i.e.,

$$R^* = \max_{S \subseteq X} R(S). \tag{2.7}$$

Without loss of generality, we assume that $u(i) > 0$ in the rest of this chapter. We use x_{ij} to denote the j^{th} product of the i^{th} level ($i = 1, 2$), and m_i to denote the number of products in level i. Also, we assume that the products in each level are indexed in a decreasing order by revenue (breaking ties arbitrarily), i.e.,

$$\forall i \in \{1, 2\}, r(x_{i1}) \geq r(x_{i2}) \geq \ldots \geq r(x_{im_i}).$$

It is useful to illustrate how the SML allows for violations of the regularity condition, a property first observed by Echenique et al. [2018]. Our first example captures the attraction effect presented earlier.

Example 1 (Attraction Effect in the SML)**.** Consider a retail store that offers different brands of chocolate. Suppose that there is a well-known brand A and the brand B owned by the retail store. There is one chocolate bar a_1 from brand A and there are two chocolate bars b_1 and b_2 from Brand B, with b_2 being a more expensive version of b_1. When shown the assortment $\{a_1, b_1\}$, 71% of the clients purchase a_1 and 8.2% buy b_1. When shown the assortment $\{a_1, b_1, b_2\}$, customers select a_1 49.8% of the time and, surprisingly, bar b_1 increases its market share to about 10%, while bar b_2 accounts for 15% of the market. The introduction of b_2 to the assortment increases the purchasing probability of b_1, violating regularity. The numerical example can be explained with the SML as follows: Consider $A = \{a_1\}$, $B = \{b_1, b_2\}$ and $X = A \uplus B$. With $u(a_1) = 100, u(b_1) = 40, u(b_2) = 60$, and $u_0 = 1$ as the utility of the outside option, we have:

$$\rho(b_1, \{a_1, b_1\}) = \frac{40}{141} \cdot \left[1 - \frac{100}{141}\right] \approx 8.2\%.$$

and

$$\rho(b_1, \{a_1, b_1, b_2\}) = \frac{40}{201} \cdot \left[1 - \frac{100}{201}\right] \approx 10\%.$$

Hence $\rho(b_1, \{a_1, b_1\}) < \rho(b_1, \{a_1, b_1, b_2\})$ which contradicts regularity.

Our second example shows that the SML can capture the so-called *paradox of choice* or *choice overload effect* (e.g., Schwartz [2004]; Chernev et al. [2015]): The overall purchasing probability may decrease when the assortment is enlarged. Once again, this effect cannot be explained by any random utility model and it is sometimes called the effect of "too much choice".

Example 2 (Paradox of Choice in the SML)**.** Let $X_1 = \{x_{11}\}$, $X_2 = \{x_{21}, x_{22}\}$, $X = X_1 \uplus X_2$, $u(x_{11}) = 10$,$u(x_{21}) = 1$,$u(x_{22}) = 10$, and $u_0 = 1$. We have

$$\begin{aligned}\rho(x_0, \{x_{11}, x_{21}\}) &= 1 - \rho(x_{11}, \{x_{11}, x_{21}\}) - \rho(x_{21}, \{x_{11}, x_{21}\}) \\ &= 1 - \frac{10}{12} - \left(1 - \frac{10}{12}\right) \cdot \frac{1}{12} = 0.152\overline{7},\end{aligned}$$

and

$$\begin{aligned}\rho(x_0, \{x_{11}, x_{21}, x_{22}\}) &= 1 - \rho(x_{11}, \{x_{11}, x_{21}, x_{22}\}) - \rho(x_{21}, \{x_{11}, x_{21}, x_{22}\}) - \rho(x_{22}, \{x_{11}, x_{21}, x_{22}\}) \\ &= 1 - \frac{10}{22} - \left(1 - \frac{10}{22}\right) \cdot \frac{1}{22} - \left(1 - \frac{10}{22}\right) \cdot \frac{10}{22} = 0.\overline{27}.\end{aligned}$$

Hence $\rho(x_0, \{x_{11}, x_{21}\}) < \rho(x_0, \{x_{11}, x_{21}, x_{22}\})$.

2.3 Properties of Optimal Assortments

In this section we derive properties of the optimal solutions to the assortment problem under the SML. These properties are extensively used in the proof of our main result (Theorem 1) in Section 2.4. We establish bounds on the following: any product offered on any optimal solution, and also the assortments considered on an optimal solution on each level. We assume a set of products $X = X_1 \uplus X_2$ and use the following notations

$$U(S) = \sum_{x \in S} u(x), \quad \alpha(S) = \frac{\sum_{x \in S} u(x) r(x)}{\sum_{x \in S} u(x)} = \frac{\sum_{x \in S} u(x) r(x)}{U(S)} \quad \text{and} \quad \lambda(Z, S) = \frac{U(Z)}{U(S) + u_0} \tag{2.8}$$

where $Z \subseteq S$ and $Z, S \subseteq X$. Note that $\alpha(S)$ is the usual MNL formula for the revenue and, when $S = \{x\}$ for some $x \in X$, $\alpha(S) = r(x)$. With these notations, the revenue of an assortment $S = S_1 \uplus S_2$ is

$$\begin{aligned}R(S) &= \frac{\alpha(S_1)U(S_1)}{U(S_1) + U(S_2) + u_0} + \frac{\alpha(S_2)U(S_2)}{U(S_1) + U(S_2) + u_0} \cdot \left(1 - \frac{U(S_1)}{U(S_1) + U(S_2) + u_0}\right) \\ &= \frac{\alpha(S_1)U(S_1)}{U(S) + u_0} + \frac{\alpha(S_2)U(S_2)}{U(S) + u_0} \cdot \left(1 - \frac{U(S_1)}{U(S) + u_0}\right).\end{aligned} \tag{2.9}$$

The following proposition is useful to divide a set into disjoint sets, which can then be analyzed separately.

Proposition 1. *Let $S \subseteq X$ and $S = H \cup T$ with $H \cap T = \emptyset$. We have*

$$\alpha(S) = \frac{\alpha(H)U(H) + \alpha(T)U(T)}{U(S)}. \tag{2.10}$$

Proof.

$$\begin{aligned}
\frac{\alpha(H)U(H) + \alpha(T)U(T)}{U(S)} &= \frac{\frac{\sum_{x\in H} r(x)u(x)}{U(H)} \cdot U(H) + \frac{\sum_{x\in T} r(x)u(x)}{U(T)} \cdot U(T)}{U(S)} && \text{/using definition of } \alpha(\cdot) \\
&= \frac{\sum_{x\in H} r(x)u(x) + \sum_{x\in T} r(x)u(x)}{U(S)} && \text{/cancelling } U(H) \text{ and } U(T) \\
&= \frac{\sum_{x\in S} r(x)u(x)}{U(S)} && \text{/using that } H \cup T = S \\
&= \alpha(S). && \text{/definition of } \alpha(S)
\end{aligned}$$

□

The next proposition is useful to bound expected revenues.

Proposition 2. *Let $S_1, S_2 \subseteq X$. If $\forall x \in S_1, \forall y \in S_2, r(x) \geq r(y)$, then $\alpha(S_1) \geq \alpha(S_2)$.*

Proof. If $\forall x \in S_1, \forall y \in S_2 : r(x) \geq r(y)$, then $\min_{x\in S_1} r(x) \geq \max_{y\in S_2} r(y)$. We have

$$\alpha(S_1) = \frac{\sum_{x\in S_1} u(x)r(x)}{\sum_{x\in S_1} u(x)} \geq \min_{x\in S_1} r(x) \cdot \underbrace{\frac{\sum_{x\in S_1} u(x)}{\sum_{x\in S_1} u(x)}}_{1} \geq \max_{y\in S_2} r(y) \underbrace{\frac{\sum_{y\in S_2} u(y)}{\sum_{y\in S_2} u(y)}}_{1} \geq \alpha(S_2).$$

□

The following proposition bounds the MNL revenue of the products in the first level. We use $S_i^* = S^* \cap X_i$ to denote the products in level i in the optimal assortment, i.e., $S^* = S_1^* \uplus S_2^*$.

Proposition 3 (Bounding Level 1)**.** *$\alpha(S_1^*) \geq R^*$.*

Proof. The proof shows that the optimal revenue is a convex combination of $\alpha(S_1^*)$ and another term by using Equation (2.9) and multiplying/dividing the revenue associated with the second level by $\frac{U(S_2^*)+u_0}{U(S_2^*)+u_0}$. We have

$$\begin{aligned}
R^* =& \frac{\alpha(S_1^*)U(S_1^*)}{U(S^*) + u_0} + \frac{\alpha(S_2^*)U(S_2^*)}{U(S^*) + u_0} \cdot \left(1 - \frac{U(S_1^*)}{U(S^*) + u_0}\right) \\
=& \frac{\alpha(S_1^*)U(S_1^*)}{U(S^*) + u_0} + \frac{\alpha(S_2^*)U(S_2^*)}{U(S_2^*) + u_0} \cdot \underbrace{\frac{U(S_2^*) + u_0}{U(S^*) + u_0}}_{\left(1-\lambda(S_1^*, S^*)\right) \in (0,1)} \cdot \left(1 - \frac{U(S_1^*)}{U(S^*) + u_0}\right) \\
=& \alpha(S_1^*)\lambda(S_1^*, S^*) + R(S_2^*)\left(1 - \lambda(S_1^*, S^*)\right)^2.
\end{aligned}$$

R^* is a convex combination of $\alpha(S_1^*)$ and $R(S_2^*)(1 - \lambda(S_1^*, S^*))$. By optimality of R^*, $R(S_2^*) \leq R^*$ and hence $\alpha(S_1^*) \geq R^*$. □

We now prove a stronger lower bound for the value $\alpha(S_2^*)$ of the second level.

Proposition 4. *(Bounding Level 2)* $\alpha(S_2^*) \geq \frac{R^*}{1-\lambda(S_1^*,S^*)}$.

Proof. The proof is similar to the one in Proposition 3.

$$\begin{aligned}
R^* =& \frac{\alpha(S_1^*)U(S_1^*)}{U(S^*)+u_0} + \frac{\alpha(S_2^*)U(S_2^*)}{U(S^*)+u_0} \cdot \left(1 - \frac{U(S_1^*)}{U(S^*)+u_0}\right) \\
=& \frac{\alpha(S_1^*)U(S_1^*)}{U(S_1^*)+u_0} \cdot \underbrace{\frac{U(S_1^*)+u_0}{U(S^*)+u_0}}_{1-\lambda(S_2^*,S^*)} + \alpha(S_2^*) \cdot \underbrace{\frac{U(S_2^*)}{U(S^*)+u_0}}_{\lambda(S_2^*,S^*)} \cdot \underbrace{\left(1 - \frac{U(S_1^*)}{U(S^*)+u_0}\right)}_{1-\lambda(S_1^*,S^*)} \\
=& R(S_1^*) \cdot (1-\lambda(S_2^*,S^*)) + (\alpha(S_2^*)(1-\lambda(S_1^*,S^*))) \cdot \lambda(S_2^*,S^*).
\end{aligned}$$

R^* is a convex combination of and $R(S_1^*)$ and $\alpha(S_2^*)(1-\lambda(S_1^*,S^*))$. By optimality of R^*, $R(S_1^*) \leq R^*$ and $\alpha(S_2^*) \geq \frac{R^*}{(1-\lambda(S_1^*,S^*))}$. □

The following example shows that it not always the case that the inequality proved above holds if one considers the products in S_2^* separately. That is, the inequality $r(x) \geq \frac{R^*}{1-\lambda(S_1^*,S^*)}$ for all $x \in S_2^*$ is not always true.

Example 3. Let $X_1 = \{x_{11}\}$, $X_2 = \{x_{21}, x_{22}\}$, and $X = X_1 \uplus X_2$. Let the revenues be $r(x_{11}) = 10, r(x_{21}) = 9$, and $r(x_{22}) = 6$ and the utilities be $u(x_{11}) = u(x_{21}) = 1, u(x_{22}) = 3$, and $u_0 = 1$. The expected revenue for all possible subsets are given by

S	$R(S)$
$\{x_{11}\}$	5
$\{x_{21}\}$	4.5
$\{x_{22}\}$	4.5
$\{x_{11}, x_{21}\}$	$5.\overline{3}$
$\{x_{11}, x_{22}\}$	4.88
$\{x_{21}, x_{22}\}$	5.4
$\{x_{11}, x_{21}, x_{22}\}$	$5.41\overline{6}$

Table 2.1: Revenue for all potential assortments in X.

The optimal assortment is $S^* = \{x_{11}, x_{21}, x_{22}\}$ with an expected revenue of $R^* = 5.41\overline{6}$. By definition of $\lambda(\cdot)$, we have

$$\lambda(S_1^*, S^*) = \frac{U(S_1^*)}{U(S^*)+u_0} = \frac{1}{6} = 0.1\overline{6}.$$

It follows that $r(x_{22}) = 6 < \frac{R^*}{1-\lambda(S_1^*,S^*)} = \frac{5.41\overline{6}}{1-0.1\overline{6}} = 6.488$, showing that the bound does not hold for product x_{22}.

However, the weaker bound holds for every product, and more generally, we have the following proposition.

Proposition 5. *In every optimal assortment S^*, if $Z \subseteq S_i^*$ ($i = 1, 2$), then $\alpha(Z) \geq R^*$.*

Proof. The proof of this proposition relies on the following technical lemma.

Lemma 1. *Consider an assortment $S = S_1 \uplus S_2$ and $Z \subseteq S_i$ for some $i = 1, 2$. $R(S)$ can be expressed in terms of the following convex combinations:*

- *If $Z \subseteq S_1$,*

$$\begin{aligned} R(S) =& R(S \setminus Z) \cdot (1 - \lambda(Z, S)) \\ &+ \left[\alpha(Z) - \frac{\alpha(S_2)U(S_2)(U(S_2) + u_0)(1 - \lambda(Z, S)}{(U(S) - U(Z) + u_0)^2}\right] \cdot \lambda(Z, S). \end{aligned} \tag{2.11}$$

- *if $Z \subseteq S_2$,*

$$\begin{aligned} R(S) =& R(S \setminus Z) \cdot (1 - \lambda(Z, S)) \\ &+ \left[\frac{\alpha(Z)(U(S_2) + u_0)}{U(S) + u_0} + \frac{\alpha(S_2 \setminus Z)(U(S_2) - U(Z))}{U(S) - U(Z) + u_0} \cdot \frac{U(S_1)}{U(S) + u_0}\right] \cdot \lambda(Z, S). \end{aligned} \tag{2.12}$$

Proof. For $Z \subseteq X$ and $Z_0 \subseteq Z$, we have:

$$\begin{aligned} \alpha(Z \setminus Z_0) &= \frac{\sum_{x \in Z \setminus Z_0} r(x)u(x)}{U(Z \setminus Z_0)} \\ &= \frac{\sum_{x \in Z} r(x)u(x)}{U(Z \setminus Z_0)} - \frac{\alpha(Z_0)U(Z_0)}{U(Z \setminus Z_0)} \\ &= \frac{\sum_{x \in Z} r(x)u(x)}{U(Z \setminus Z_0)} \cdot \frac{U(Z)}{U(Z)} - \frac{\alpha(Z_0)U(Z_0)}{U(Z \setminus Z_0)} \\ &= \underbrace{\frac{\sum_{x \in Z} r(x)u(x)}{U(Z)}}_{\alpha(Z)} \cdot \frac{U(Z)}{U(Z) - U(Z_0)} - \frac{\alpha(Z_0)U(Z_0)}{U(Z) - U(Z_0)} \\ &= \frac{\alpha(Z)U(Z) - \alpha(Z_0)U(Z_0)}{U(Z) - U(Z_0)} \\ &= \frac{\alpha(Z)U(Z) - \alpha(Z_0)U(Z_0)}{U(Z) - U(Z_0)}, \end{aligned} \tag{2.13}$$

which can be rewritten as

$$\alpha(Z \setminus Z_0)(U(Z) - U(Z_0)) = \alpha(Z)U(Z) - \alpha(Z_0)U(Z_0). \tag{2.14}$$

Note also that, when $Z_0 = Z$, $\alpha(Z \setminus Z_0) = \alpha(\emptyset) = 0$.

The rest of the proof is by case analysis on the level. If $Z \subseteq S_1$, $\lambda(Z,S) = \frac{U(Z)}{U(S)+u_0}$. We have:

$$\begin{aligned}
R(S) =& \frac{\alpha(S_1)U(S_1)}{U(S)+u_0} + \frac{\alpha(S_2)U(S_2)}{U(S)+u_0} \cdot \left(1 - \frac{U(S_1)}{U(S)+u_0}\right) \\
=& \frac{\alpha(S_1)U(S_1) - \alpha(Z)U(Z)}{U(S)+u_0} + \frac{\alpha(Z)U(Z)}{U(S)+u_0} + \frac{\alpha(S_2)U(S_2)}{U(S)+u_0} \cdot \left(1 - \frac{U(S_1)}{U(S)+u_0}\right) \\
=& \frac{\alpha(S_1)U(S_1) - \alpha(Z)U(Z)}{U(S)+u_0} \cdot \frac{U(S)-U(Z)+u_0}{U(S)-U(Z)+u_0} + \frac{\alpha(Z)U(Z)}{U(S)+u_0} + \frac{\alpha(S_2)U(S_2)}{U(S)+u_0} \cdot \left(1 - \frac{U(S_1)}{U(S)+u_0}\right) \\
=& \frac{\alpha(S_1)U(S_1) - \alpha(Z)U(Z)}{U(S)-U(Z)+u_0} \cdot (1-\lambda(Z,S)) + \alpha(Z)\lambda(Z,S) \\
&+ \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)}{(U(S)-U(Z)+u_0)^2} \cdot \left(\frac{U(S)-U(Z)+u_0}{U(S)+u_0}\right)^2,
\end{aligned}$$

where we first add and subtract $\frac{\alpha(Z)U(Z)}{U(S)+u_0}$, and multiply and divide the first term by $(U(S) - U(Z) + u_0)$. The last step uses the definition of $\lambda(Z,S)$ and multiplies and divides the last term by $(U(S) - U(Z) + u_0)^2$. Now applying Equation (2.14) to S_1 and Z in the last equation, we obtain

$$\begin{aligned}
R(S) =& \frac{\alpha(S_1 \setminus Z)(U(S_1) - U(Z))}{U(S)-U(Z)+u_0} \cdot (1-\lambda(Z,S)) + \alpha(Z)\lambda(Z,S) \\
&+ \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)}{(U(S)-U(Z)+u_0)^2} \cdot (1-\lambda(Z,S))^2 \\
=& \underbrace{\left[\frac{\alpha(S_1 \setminus Z)(U(S_1) - U(Z))}{U(S)-U(Z)+u_0} + \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)}{(U(S)-U(Z)+u_0)^2}\right]}_{R(S_1 \setminus Z \cup S_2)} \cdot (1-\lambda(Z,S)) \\
&+ \left[\alpha(Z) - \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)\,(1-\lambda(Z,S))}{(U(S)-U(Z)+u_0)^2}\right] \cdot \lambda(Z,S) \\
=& R(S_1 \setminus Z \cup S_2)\,(1-\lambda(Z,S)) + \left[\alpha(Z) - \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)\,(1-\lambda(Z,S))}{(U(S)-U(Z)+u_0)^2}\right] \cdot \lambda(Z,S) \\
=& R(S \setminus Z)\,(1-\lambda(Z,S)) + \left[\alpha(Z) - \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)\,(1-\lambda(Z,S))}{(U(S)-U(Z)+u_0)^2}\right] \cdot \lambda(Z,S).
\end{aligned}$$

If $Z \subseteq S_2$, the proof is essentially similar. It also uses $\lambda(Z,S) = \frac{U(Z)}{U(S)+u_0}$ and apply

Equation (2.14) to S_2 and Z to obtain

$$\begin{aligned}
R(S) =& \frac{\alpha(S_1)U(S_1)}{U(S)+u_0} + \frac{\alpha(S_2)U(S_2)}{U(S)+u_0} \cdot \left(1 - \frac{U(S_1)}{U(S)+u_0}\right) \\
=& \frac{\alpha(S_1)U(S_1)}{U(S)-U(Z)+u_0} \cdot (1-\lambda(Z,S)) + \frac{\alpha(Z)U(Z)(U(S_2)+u_0)}{(U(S)+u_0)^2} \\
&+ \frac{(\alpha(S_2)U(S_2)-\alpha(Z)U(Z))(U(S_2)+u_0)}{(U(S)+u_0)^2} \\
=& \frac{\alpha(S_1)U(S_1)}{U(S)-U(Z)+u_0} \cdot (1-\lambda(Z,S)) + \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))(U(S_2)-U(Z)+u_0)}{(U(S)+u_0)^2} \\
&+ \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))U(Z)}{(U(S)+u_0)^2} + \frac{\alpha(Z)U(Z)(U(S_2)+u_0)}{(U(S)+u_0)^2},
\end{aligned}$$

where we multiply and divide the first term by $(U(S)-U(Z)+u_0)$ and use the definition of $\lambda(Z,S)$ and then add and subtract $\frac{\alpha(Z)U(Z)(U(S_2)+u_0)}{(U(S)+u_0)^2}$. The last step uses Equation (2.14) and adds and subtracts $\frac{\alpha(S_2\setminus Z)(U(S_2)-U(Z))U(Z)}{(U(S)+u_0)^2}$. The goal of this manipulation is to form $R(S \setminus Z)\,(1-\lambda(Z,S))$. We then obtain

$$\begin{aligned}
=& \frac{\alpha(S_1)U(S_1)}{U(S)-U(Z)+u_0} \cdot (1-\lambda(Z,S)) + \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))(U(S_2)-U(Z)+u_0)}{(U(S)+u_0)^2} \\
&+ \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))U(Z)}{(U(S)+u_0)^2} + \frac{\alpha(Z)U(Z)(U(S_2)+u_0)}{(U(S)+u_0)^2} \\
=& \frac{\alpha(S_1)U(S_1)}{U(S)-U(Z)+u_0} \cdot (1-\lambda(Z,S)) \\
&+ \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))(U(S_2)-U(Z)+u_0)}{(U(S)-U(Z)+u_0)^2}\,(1-\lambda(Z,S))^2 \\
&+ \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))U(Z)}{(U(S)+u_0)^2} + \frac{\alpha(Z)U(Z)(U(S_2)+u_0)}{(U(S)+u_0)^2} \\
=& \underbrace{\left[\frac{\alpha(S_1)U(S_1)}{U(S)-U(Z)+u_0} + \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))(U(S_2)-U(Z)+u_0)}{(U(S)-U(Z)+u_0)^2}\right]}_{R(S_1 \cup S_2 \setminus Z)} \cdot (1-\lambda(Z,S)) \\
&+ \frac{\alpha(Z)U(Z)(U(S_2)+u_0)}{(U(S)+u_0)^2} + \frac{\alpha(S_2 \setminus Z)(U(S_2)-U(Z))U(Z)}{(U(S)+u_0)^2} \\
&- \frac{\lambda(Z,S)\,(1-\lambda(Z,S))\,\alpha(S_2 \setminus Z)(U(S_2)-U(Z))(U(S_2)-U(Z)+u_0)}{(U(S)-U(Z)+u_0)^2},
\end{aligned}$$

where the second step multiplies and divides the second term by $(U(S)-U(Z)+u_0)$ and uses the definition of $\lambda(Z,S)$. The third step splits the second term by expressing $(1-\lambda(Z,S))^2$ as $(1-\lambda(Z,S)) - \lambda(Z,S)(1-\lambda(Z,S))$ in order to identify the term $R(S_1 \cup S_2 \setminus Z)$. Now, putting together the remaining terms, we obtain:

$$
\begin{aligned}
R(S) =& R(S_1 \cup S_2 \setminus Z)\,(1-\lambda(Z,S)) \\
&+\lambda(Z,S)\left[\frac{\alpha(Z)(U(S_2)+u_0)}{U(S)+u_0}+\frac{\alpha(S_2\setminus Z)(U(S_2)-U(Z))}{U(S)+u_0}\right. \\
&\left.-\frac{(1-\lambda(Z,S))\,\alpha(S_2\setminus Z)(U(S_2)-U(Z))(U(S_2)-U(Z)+u_0)}{(U(S)-U(Z)+u_0)^2}\right] \\
=& R(S_1 \cup S_2 \setminus Z)\,(1-\lambda(Z,S)) \\
&+\lambda(Z,S)\left[\frac{\alpha(Z)(U(S_2)+u_0)}{U(S)+u_0}+\frac{\alpha(S_2\setminus Z)(U(S_2)-U(Z))}{U(S)+u_0}\left[1-\left(1-\frac{U(S_1)}{U(S)-U(Z)+u_0}\right)\right]\right] \\
=& R(S\setminus Z)\,(1-\lambda(Z,S))+\left[\frac{\alpha(Z)(U(S_2)+u_0)}{U(S)+u_0}+\frac{\alpha(S_2\setminus Z)(U(S_2)-U(Z))}{U(S)-U(Z)+u_0}\cdot\frac{U(S_1)}{U(S)+u_0}\right]\cdot\lambda(Z,S).
\end{aligned}
$$

where the second line uses the definition of $\lambda(Z,S)$ to factorize the expression and the last step just simplifies the resulting expressions. □

Now we can continue with the proof of Proposition 5. Let $Z \subseteq S_i^*$ for some $i = 1,2$. Consider first the case in which the optimal solution S^* contains only products from level i ($i \in \{1,2\}$). Then,

$$
\begin{aligned}
R^* &= \frac{\sum_{y\in S^*} r(y)u(y)}{\sum_{y\in S^*} u(y)+u_0} \\
&= \frac{\sum_{y\in S^*\setminus Z} r(y)u(y)}{\sum_{y\in S^*\setminus Z} u(y)+u_0}\cdot\frac{\sum_{y\in S^*} u(y)-U(Z)+u_0}{\sum_{y\in S^*} u(y)+u_0}+\frac{\alpha(Z)U(Z)}{\sum_{y\in S^*} u(y)+u_0} \\
&= \underbrace{\frac{\sum_{y\in S^*\setminus Z} r(y)u(y)}{\sum_{y\in S^*\setminus Z} u(y)+u_0}}_{R(S^*\setminus Z)}\cdot(1-\lambda(Z,S^*))+\alpha(Z)\lambda(Z,S^*) \\
&= R(S^*\setminus Z)\cdot(1-\lambda(Z,S^*))+\alpha(Z)\lambda(Z,S^*).
\end{aligned}
$$

The optimal solution is a convex combination of $R(S^* \setminus Z)$ and $\alpha(Z)$. By optimality of R^*, $R(S^* \setminus Z) \leq R^*$ and hence $\alpha(Z) \geq R^*$.

Consider the case in which the solution is non-empty in both levels, and suppose that $\alpha(Z) < R^*$. We now show that this is not possible. The proof considers two independent cases, depending on the level that contains Z.

If $Z \subseteq S_1^*$, by Lemma 1, the revenue of S^* can be expressed as

$$
R(S^*) = R(S^*\setminus Z)\cdot(1-\lambda(Z,S^*))+\underbrace{\left[\alpha(Z)-\frac{\alpha(S_2^*)U(S_2)(U(S_2^*)+u_0)(1-\lambda(Z,S^*))}{(U(S^*)-U(Z)+u_0)^2}\right]}_{\Gamma_Z}\cdot\lambda(Z,S^*). \tag{2.15}
$$

R^* is a convex combination of $R(S^* \setminus Z)$ and Γ_Z. We show that $\Gamma_Z < R^*$.

$$\begin{aligned}
\Gamma_Z &= \alpha(Z) - \frac{\alpha(S_2^*)U(S_2^*)(U(S_2^*)+u_0)\,(1-\lambda(Z,S^*))}{(U(S^*)-U(Z)+u_0)^2} \\
&= \alpha(Z) - \frac{\alpha(S_2^*)U(S_2^*)(U(S_2^*)+u_0)}{(U(S^*)-U(Z)+u_0)(U(S^*)+u_0)} && \text{/using definition of } \lambda(Z,S^*) \\
&= \alpha(Z) - \frac{\alpha(S_2^*)U(S_2^*)(1-\lambda(S_1^*,S^*))}{(U(S^*)-U(Z)+u_0)} && \text{/by definition of } \lambda(S_1^*,S^*) \\
&\leq \alpha(Z) - \frac{R^*U(S_2^*)}{U(S^*)-U(Z)+u_0} && \text{/Using proposition 4} \\
&< R^*\left(1-\frac{U(S_2^*)}{U(S^*)-U(Z)+u_0}\right) && \text{/using the assumption } \alpha(Z) < R^* \\
&< R^*.
\end{aligned}$$

Since $R(S^*/Z) \leq R^*$, we have that $R^* < R^*$ and hence it must be the case that $\alpha(Z) \geq R^*$.
Now, if $Z \subseteq S_2^*$, by Lemma 1, the revenue of S^* can be expressed as

$$\begin{aligned}
R(S^*) &= R(S^* \setminus Z)\,(1-\lambda(Z,S^*)) \\
&\quad + \underbrace{\left[\frac{\alpha(Z)(U(S_2^*)+u_0)}{U(S^*)+u_0} + \frac{\alpha(S_2^*\setminus Z)(U(S_2^*)-U(Z))}{U(S^*)-U(Z)+u_0}\cdot\frac{U(S_1^*)}{U(S^*)+u_0}\right]}_{\Gamma_Z}\cdot\lambda(Z,S^*). \qquad (2.16)
\end{aligned}$$

R^* is thus a convex combination of $R(S^* \setminus Z)$ and Γ_Z. Again, we show that $\Gamma_Z < R^*$:

$$\begin{aligned}
\Gamma_Z &= \frac{\alpha(Z)(U(S_2^*)+u_0)}{U(S^*)+u_0} + \frac{\alpha(S_2^*\setminus Z)(U(S_2^*)-U(Z))}{U(S^*)-U(Z)+u_0}\cdot\frac{U(S_1^*)}{U(S^*)+u_0} \\
&= \alpha(Z)(1-\lambda(S_1^*,S^*)) + \frac{\alpha(S_2^*\setminus Z)(U(S_2^*)-U(Z))}{U(S^*)-U(Z)+u_0}\cdot\lambda(S_1^*,S^*) && \text{/by definition of } \lambda(S_1^*,S^*) \\
&< \alpha(Z)(1-\lambda(S_1^*,S^*)) + \underbrace{\frac{\alpha(S_2^*\setminus Z)(U(S_2^*)-U(Z))}{U(S_2^*)-U(Z)+u_0}}_{R(S_2^*\setminus Z)}\cdot\lambda(S_1^*,S^*) && \text{/replacing } U(S^*) \text{ by } U(S_2^*) \\
&< R^*(1-\lambda(S_1^*,S^*)) + R(S_2^*\setminus Z)\cdot\lambda(S_1^*,S^*) && \text{/using that } \alpha(Z) < R^* \\
&< R^*(1-\lambda(S_1^*,S^*)) + R(S_2^*\setminus Z)\lambda(S_1^*,S^*) && \text{/Using the optimality of } R^* \\
&< R^*.
\end{aligned}$$

Hence, it must be the case that $\alpha(Z) \geq R^*$, completing the proof.

□

The converse of Proposition 5 does not hold: Example 4 presents an instance where the optimal solution does not contain all the products with a revenue higher than R^*.

Example 4. We show that some products with revenue greater or equal than R^* may not be included in an optimal assortment. Let $X_1 = \{x_{11}\}$, $X_2 = \{x_{21}\}$, and $X = X_1 \uplus X_2$. Let the revenues be $r(x_{11}) = r(x_{21}) = 1$ and the utilities be $u(x_{11}) = 10, u(x_{21}) = 1$, and $u_0 = 1$. Consider the possible assortments and their expected revenues:

$$
\begin{aligned}
R(\{x_{11}\}) &= \tfrac{u(x_{11})r(x_{11})}{u(x_{11})+u_0} = \tfrac{10\cdot 1}{10+1} = 0.9\overline{09} \\
R(\{x_{21}\}) &= \tfrac{u(x_{21})r(x_{21})}{u(x_{21})+u_0} = \tfrac{1\cdot 1}{1+1} = 0.5 \\
R(\{x_{11}, x_{21}\}) &= \tfrac{u(x_{11})r(x_{11})}{u(x_{11})+u(x_{21})+u_0} + \left(1 - \tfrac{u(x_{11})}{u(x_{11})+u(x_{21})+u_0}\right) \cdot \tfrac{u(x_{21})r(x_{21})}{u(x_{11})+u(x_{21})+u_0} \\
&= \tfrac{10\cdot 1}{10+1+1} + \left(1 - \tfrac{10}{10+1+1}\right) \cdot \tfrac{1\cdot 1}{10+1+1} \\
&= \tfrac{10}{12} + \left(1 - \tfrac{10}{12}\right) \cdot \tfrac{1}{10+1} \\
&= 0.847\overline{2}
\end{aligned}
$$

The optimal assortment is $S^* = \{x_{11}\}$. However, we have that $r(x_{21}) = 1 > R^*$, but x_{21} is not part of the optimal assortment.

The following corollary is a direct consequence of Proposition 5.

Corollary 1. *For any non-empty subset $S_0 \subseteq S^*$, where S^* is an optimal solution, we have $\alpha(S_0) \geq R^*$.*

Proof. By Proposition 5, we have $\alpha(\{x\}) = r(x) \geq R^*$ for all $x \in S^*$. For each set $S_0 \subseteq S^*$, we have:

$$
\begin{aligned}
\alpha(S_0) =& \frac{\sum_{x \in S_0} u(x) r(x)}{\sum_{x \in S_0} u(x)} \\
\geq& R^* \frac{\sum_{x \in S_0} u(x)}{\sum_{x \in S_0} u(x)} \quad \text{/ by Proposition 5} \\
=& R^*.
\end{aligned}
$$

□

The corollary above implies that $\alpha(\{x\}) = r(x) \geq R^*$ for all $x \in S^*$. Thus, every product in an optimal assortment has a revenue greater than or equal to R^*.

In the following example we show that the well known revenue-ordered assortment strategy for the assortment problem does not always lead to optimality.

Example 5 (Revenue-Ordered assortments are not optimal)**.** Let $X_1 = \{x_{11}\}$, $X_2 = \{x_{21}\}$, and $X = X_1 \uplus X_2$. Let $r(x_{11}) = 10$ and $r(x_{21}) = 12$. Let the utilities be $u(x_{11}) = 10, u(x_{21}) = 2$, and $u_0 = 1$. A direct calculation shows that the revenues for all possible assortments under this setting are:

S	$R(S)$
$\{x_{11}\}$	$9.\overline{09}$
$\{x_{21}\}$	8
$\{x_{11}, x_{21}\}$	8.12

Table 2.2: Revenue for all potential assortments in X.

The optimal assortment is $S^* = \{x_{11}\}$, yielding a revenue of $R^* = 9.\overline{09}$. However, the best revenue ordered assortment is $S' = \{x_{11}, x_{21}\}$, obtaining a revenue of $R' = 8.12$, this means that the revenue-ordered assortment strategy provides an approximation ratio of $\frac{R'}{R^*} \approx 89.3\%$ for this particular instance.

2.4 Optimality of Revenue-Ordered Assortments by level

This section proves that optimal assortments under the SML are revenue-ordered by level, generalizing the traditional results for the MNL [Talluri and Van Ryzin, 2004]. As a corollary, the optimal assortment problem under the SML is polynomial-time solvable.

Definition 2 (Revenue-Ordered Assortment by Level)**.** Denote by N_{ij} the set of all products in level i with a revenue of at least r_{ij} ($1 \leq j \leq m_i$) and fix $N_{i0} = \emptyset$ by convention. A revenue-ordered assortment by level is a set $S = N_{1j_1} \uplus N_{2j_2} \subseteq X$ for some $0 \leq j_1 \leq m_1$ and $0 \leq j_2 \leq m_2$. We use $\mathcal{A}$ to denote the set of revenue-ordered assortments by level.

When an assortment $S = S_1 \uplus S_2$ is not revenue-ordered by level, it follows that

$$\exists k \in \{1, 2\}, z \in X_k \setminus S_k, y \in S_k : \quad r(z) \geq r(y).$$

We say that S has a gap, the gap is at level k, and z belongs to the gap. We now define the concept of *first gap*, which is heavily used in the proof.

Definition 3 (First Gap of an Assortment)**.** Let $S = S_1 \uplus S_2$ be an assortment with a gap and let k be the smallest level with a gap. Let $\hat{r} = \max_{y \in X_k \setminus S_k} r(y)$ be the maximum revenue of a product in level k not contained in S_k. The *first gap* of S is a set of products $G \subseteq X_k \setminus S$ defined as follows:

- If $\max_{x \in S_k} r(x) < \hat{r}$, then the gap G consists of all products with higher revenues than the products in assortment S, i.e.,

$$G = \{y \in X_k \setminus S_k \mid r(y) \geq \max_{x \in S_k} r(x)\}.$$

- Otherwise, when $\max_{x \in S_k} r(x) \geq \hat{r}$, define the following quantities:

$$r_M = \min_{\substack{x \in S_k \\ r(x) \geq \hat{r}}} r(x) \quad \text{and} \quad r_m = \max_{\substack{x \in S_k \\ r(x) \leq \hat{r}}} r(x). \tag{2.17}$$

  The set G contains products with revenues in $[r_m, r_M]$, i.e.,

$$G = \{y \in X_k \setminus S_k \mid r_m \leq r(y) \leq r_M\}.$$

We are now in a position to state the main theorem of this chapter.

Theorem 1. *Under the SML, any optimal assortment is revenue-ordered by level.*

Proof. The intuition behind the proof is the following: Assume that S is an optimal solution with at least one gap as in Definition 3. Let G be the first gap of S, and that G occurs at level k. Define $S_k = H \cup T$ with $H, T \subseteq X_k$ and

$$H = \{x \in S_k \mid r(x) \geq \max_{g \in G} r(g)\}$$

and

$$T = \{x \in S_k \mid r(x) \leq \min_{g \in G} r(g)\}.$$

We call the set H as the *head* and the set T is called the *tail*. Figure 2.1 illustrates these concepts visually.

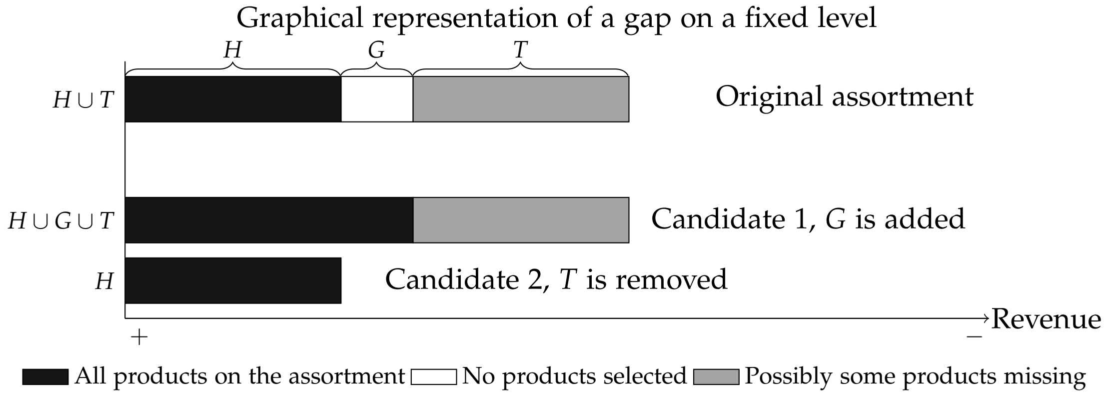


Figure 2.1: Representation of a level containing a gap G at the top, and the two proposed candidates fixing the gap by either adding G, or removing T.

We will prove that is always possible to select an assortment that is revenue-ordered by level and has revenue greater than $R(S)$ (contradicting optimality). The proof shows that such an assortment can be obtained either by including the gap G in S or by eliminating T from S. For the purpose of contradiction, assume that S is an optimal solution with at least one gap, G is the first gap of S, and G occurs at level k. Define $S_k = H \cup T$ with $H, T \subseteq X_k$ and

$$H = \{x \in S_k \mid r(x) \geq \max_{g \in G} r(g)\}$$

and

$$T = \{x \in S_k \mid r(x) \leq \min_{g \in G} r(g)\}.$$

In the following, the set H is called the *head* and the set T is called the *tail*. We prove that is always possible to select an assortment that is revenue-ordered by level and has revenue greater than $R(S)$. The proof shows that such an assortment can be obtained either by including the gap G in S or by eliminating T from S. Figure 2.1

illustrates these concepts visually. The proof is by case analysis on the level of G.

Consider first the case where G is in the first level. We can define $S = S_1 \uplus S_2$, with $S_1 = H \cup T$ as defined above and $S_2 \subseteq X_2$. The revenue for S is

$$\begin{aligned} R(S_1 \cup S_2) &= \frac{\alpha(S_1)U(S_1)}{U(S)+u_0} + \left(1 - \frac{U(S_1)}{(U(S)+u_0)}\right) \cdot \frac{\alpha(S_2)U(S_2)}{U(S)+u_0} \\ &= \frac{\alpha(H)U(H)}{U(S)+u_0} + \frac{\alpha(T)U(T)}{U(S)+u_0} + \frac{U(S_2)+u_0}{(U(S)+u_0)^2} \cdot \alpha(S_2)U(S_2) \end{aligned}$$

where we used Proposition 1 on S_1 for deriving the second equality. We show that assortment $H \cup S_2$ or assortment $H \cup G \cup T \cup S_2$ provides a revenue greater than $R(S = H \cup T \cup S_2)$, contradicting our optimality assumption for S. The proof characterizes the differences between the revenues of S and the two considered assortments, adds those two differences, and shows that this value is strictly less than zero, implying that at least one of the differences is strictly negative and hence that one of these assortments has a revenue larger than $R(S)$. $R(H \cup S_2)$ can be expressed as

$$\frac{\alpha(H)U(H)}{U(H)+U(S_2)+u_0} + \frac{U(S_2)+u_0}{(U(H)+U(S_2)+u_0)^2} \cdot \alpha(S_2)U(S_2). \tag{2.18}$$

Let $\theta = U(H) + U(T) + U(S_2) + u_0$ (or, equivalently, $\theta = U(S) + u_0$). The difference $R(H \cup T \cup S_2) - R(H \cup S_2)$ is

$$\frac{U(T)}{\theta(\theta - U(T))}\left[-\alpha(H)U(H) + \alpha(T)(\theta - U(T)) - \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)(2(\theta - U(T)) + U(T))}{\theta(\theta - U(T))}\right]. \tag{2.19}$$

$R(H \cup G \cup T \cup S_2)$ can be expressed as

$$\frac{1}{U(S)+U(G)+u_0} \cdot [\alpha(H)U(H) + \alpha(G)U(G) + \alpha(T)U(T)] + \frac{U(S_2)+u_0}{(U(S)+U(G)+u_0)^2} \cdot \alpha(S_2)U(S_2).$$

The difference $R(H \cup T \cup S_2) - R(H \cup G \cup T \cup S_2)$ is given by

$$\frac{U(G)}{\theta(\theta + U(G))} \cdot \left[\alpha(H)U(H) + \alpha(T)U(T) - \alpha(G)\theta + \frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)(2\theta + U(G))}{\theta(\theta + U(G))}\right]. \tag{2.20}$$

By optimality of S, these two differences must be positive. However, their sum, dropping the positive multiplying term on each difference, which must also be positive, is given by

$$\begin{aligned}
(2.19)+(2.20) &= \alpha(T)\theta-\alpha(G)\theta+\frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)}{\theta}\cdot\left[\frac{2\theta+U(G)}{\theta+U(G)}-\frac{2(\theta-U(T))+U(T)}{(\theta-U(T))}\right]\\
&=(\alpha(T)-\alpha(G))\theta+\frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)}{\theta}\cdot\left[1+\frac{\theta}{\theta+U(G)}-2-\frac{U(T)}{(\theta-U(T))}\right]\\
&=\underbrace{\underbrace{(\alpha(T)-\alpha(G))\theta}_{\leq 0,\text{ by Proposition 2 and }\theta\geq 0}+\frac{\alpha(S_2)U(S_2)(U(S_2)+u_0)}{\theta}\cdot\left[\underbrace{\left(\frac{\theta}{\theta+U(G)}-1\right)}_{<0}\underbrace{-\frac{U(T)}{(\theta-U(T))}}_{<0}\right]}_{<0}
\end{aligned}$$

which contradicts the optimality of S.

Consider now the case where the gap is in the second level. Using the definition of the head and the tail discussed above, S can be written as $S=S_1\uplus H\cup T$. The revenue $R(S)$ is given by

$$R(S_1\cup H\cup T)=\frac{\alpha(S_1)U(S_1)}{\theta}+\frac{\alpha(H)U(H)}{\theta}+\frac{\alpha(T)U(T)}{\theta}-\frac{U(S_1)\alpha(H)U(H)}{\theta^2}-\frac{U(S_1)\alpha(T)U(T)}{\theta^2} \tag{2.21}$$

and the proof follows the same strategy as for the case of the first level. The revenue $R(S_1\cup H)$ is given by

$$R(S_1\cup H)=\frac{\alpha(S_1)U(S_1)}{\theta-U(T)}+\frac{\alpha(H)U(H)}{\theta-U(T)}-\frac{U(S_1)\alpha(H)U(H)}{(\theta-U(T))^2} \tag{2.22}$$

and the difference $R(S_1\cup H\cup T)-R(S_1\cup H)$ by

$$\begin{aligned}
&\frac{U(T)}{\theta(\theta-U(T))}\cdot\Bigg[-\alpha(S_1)U(S_1)-\alpha(H)U(H)+\alpha(T)(\theta-U(T))\\
&+\frac{\alpha(H)U(H)U(S_1)(2\theta-U(T))}{\theta(\theta-U(T))}-\frac{U(S_1)\alpha(T)(\theta-U(T))}{\theta}\Bigg]
\end{aligned} \tag{2.23}$$

The revenue $R(S_1\cup H\cup G\cup T)$ is given by

$$\begin{aligned}
&\frac{\alpha(S_1)U(S_1)}{\theta+U(G)}+\frac{\alpha(H)U(H)}{\theta+U(G)}+\frac{\alpha(G)U(G)}{\theta+U(G)}+\frac{\alpha(T)U(T)}{\theta+U(G)}\\
&-\frac{U(S_1)}{(\theta+U(G))^2}\cdot[\alpha(H)U(H)+\alpha(G)U(G)+\alpha(T)U(T)]
\end{aligned} \tag{2.24}$$

and the difference $R(S_1 \cup H \cup T) - R(S_1 \cup H \cup G \cup T)$ by

$$\frac{U(G)}{\theta(\theta+U(G))} \cdot \Bigg[\alpha(S_1)U(S_1) + \alpha(H)U(H) - \alpha(G)\theta + \alpha(T)U(T) - \frac{\alpha(H)U(H)U(S_1)(2\theta+U(G))}{\theta(\theta+U(G))} - \frac{\alpha(T)U(T)U(S_1)(2\theta+U(G))}{\theta(\theta+U(G))} + \frac{U(S_1)\alpha(G)\theta}{\theta+U(G)}\Bigg] \tag{2.25}$$

Adding (2.23) and (2.25) and dropping the positive multiplying terms on each difference gives

$$\begin{aligned}
&\theta\left(\alpha(T)-\alpha(G)\right) + U(S_1)\left(\alpha(G)-\alpha(T)\right) + U(S_1)\left[\frac{\alpha(T)U(T)}{\theta} - \frac{\alpha(G)U(G)}{\theta+U(G)}\right] \\
&+ \frac{U(S_1)}{\theta}\left[\alpha(H)U(H)\left(1+\frac{\theta}{\theta-U(T)}\right) - \alpha(H)U(H)\left(1+\frac{\theta}{\theta+U(G)}\right) - \alpha(T)U(T)\left(1+\frac{\theta}{\theta+U(G)}\right)\right] \\
&= \left(\theta-U(S_1)\right)\left(\alpha(T)-\alpha(G)\right) + \frac{U(S_1)\alpha(H)U(H)}{\theta-U(T)} - \frac{U(S_1)}{\theta+U(G)} \cdot \left[\alpha(H)U(H)+\alpha(G)U(G)+\alpha(T)U(T)\right] \\
&= \underbrace{\left(\theta-U(S_1)\right)\left(\alpha(T)-\alpha(G)\right)}_{\leq 0,\text{ by Proposition 2 and }\theta\geq U(S_1)} + \underbrace{\frac{U(S_1)\alpha(H)U(H)(U(G)+U(T))}{(\theta-U(T))(\theta+U(G))} - \frac{U(S_1)}{\theta+U(G)} \cdot \left[\alpha(G)U(G)+\alpha(T)U(T)\right]}_{\Gamma}
\end{aligned} \tag{2.26}$$

Γ cannot be greater or equal than zero, since otherwise

$$\begin{aligned}
&\frac{U(S_1)\alpha(H)U(H)(U(G)+U(T))}{(\theta-U(T))(\theta+U(G))} - \frac{U(S_1)}{\theta+U(G)} \cdot \left[\alpha(G)U(G)+\alpha(T)U(T)\right] \geq 0 \\
&\frac{U(S_1)}{(\theta+U(G))} \cdot \left[\frac{\alpha(H)U(H)(U(G)+U(T))}{(\theta-U(T))} - \left(\alpha(G)U(G)+\alpha(T)U(T)\right)\right] \geq 0.
\end{aligned} \tag{2.27}$$

The factor on the left is always positive, so Inequality (2.27) implies that the term between brackets is greater than zero. We now show that this contradicts the optimality of S. We do this by manipulating Inequality (2.27) and showing that, if this inequality holds, then $R(H) > R(S)$.

$$
\begin{aligned}
&\frac{\alpha(H)U(H)(U(G)+U(T))}{(\theta-U(T))}-(\alpha(G)U(G)+\alpha(T)U(T))\geq 0\\
&\frac{\alpha(H)U(H)}{(\theta-U(T))}\geq\frac{\alpha(G)U(G)+\alpha(T)U(T)}{(U(G)+U(T))}\\
&\frac{\alpha(H)U(H)}{(U(S_1)+U(H)+u_0)}\geq\frac{\alpha(G)U(G)+\alpha(T)U(T)}{(U(G)+U(T))}\\
&\underbrace{\frac{\alpha(H)U(H)}{(U(H)+u_0)}}_{R(H)}\cdot\left[1-\frac{U(S_1)}{U(S_1)+U(H)+u_0}\right]\geq\frac{\alpha(G)U(G)+\alpha(T)U(T)}{(U(G)+U(T))}\\
&R(H)\geq\underbrace{R(H)\cdot\frac{U(S_1)}{U(S_1)+U(H)+u_0}}_{>0}+\frac{\alpha(G)U(G)+\alpha(T)U(T)}{(U(G)+U(T))}>R(S)\cdot\frac{U(G)+U(T)}{U(G)+U(T)}>R(S).
\end{aligned}
\tag{2.28}
$$

Inequality (2.28) follows from Proposition 5 applied to $T\subset S_2$, which implies $\alpha(T)\geq R(S)$, and from Proposition 2, which implies $\alpha(G)\geq\alpha(T)$ and hence $\alpha(G)\geq R(S)$. □

The following corollary follows directly from the fact that there are at most $\mathcal{O}(|X|^2)$ revenue-ordered assortments by level and the fact that the revenue obtained from a given assortment can be computed in polynomial time.

Corollary 2. *The assortment problem under the sequential multinomial logit is polynomial-time solvable.*

2.5 Numerical Experiments

In this section, we analyse numerically the performance of revenue-ordered assortments (RO) against our proposed strategy (ROL) by varying the number of products, and the utility of the outside option. In our experiments with up to 100 products, we found that the optimality gap can be as large as 26.319%.

Each family or class of instances we tested is defined by three numbers: the number of products in the first and second level (n_1, n_2), and the utility of the outside option u_0. In total, we tested 20 classes or family of instances, each containing 100 instances. In each specific instance, revenues and product utilities are drawn from an uniform distribution between 0 and 10. We ran both strategies (RO and ROL) and we report the average and the worst optimality gap for the RO strategy, and the time taken for both strategies. These numerical experiments were conducted in Python 3.6, at a computer with 4 processors (each with 3.6 GHz CPU) and 16 GB of RAM. The computing time is reported in seconds is the average among the 100 instances in each class (or family).

Based on the results on Table 2.3 we can observe the following:

(n_1, n_2)	u_0	RO			ROL
		Avg. Gap (%)	Worst Gap (%)	Avg. Time RO (s)	Avg. Time ROL (s)
(5,5)	0	0	0	0	0
(5,5)	1	1.811	18.098	0	0.001
(5,5)	2.5	3.43	17.416	0	0
(5,5)	5	3.413	13.183	0	0
(5,5)	10	1.923	12.406	0	0.001
(10,10)	0	0	0	0	0.004
(10,10)	1	3.23	20.669	0.001	0.004
(10,10)	2.5	5.613	20.359	0.001	0.004
(10,10)	5	5.975	15.521	0.001	0.004
(10,10)	10	5.347	15.694	0.001	0.004
(20,20)	0	0	0	0.003	0.04
(20,20)	1	4.331	19.427	0.003	0.04
(20,20)	2.5	8.523	21.873	0.003	0.039
(20,20)	5	9.776	19.719	0.003	0.038
(20,20)	10	9.682	18.771	0.003	0.039
(50,50)	0	0	0	0.016	0.564
(50,50)	1	6.315	26.319	0.016	0.549
(50,50)	2.5	11.662	24.117	0.016	0.55
(50,50)	5	14.94	24.445	0.016	0.551
(50,50)	10	15.543	22.232	0.016	0.547

Table 2.3: Numerical experiments comparing the revenue ordered assortment strategy (RO) and the revenue-ordered assortments by level (ROL, which is optimal). For each class of instances, we display the average optimality gap and the worst-case gap, as well as the computing time.

1. As expected, although ROL takes more time than RO, it takes a small amount of time to solve the instances.
2. When the utility of the outside option is $u_0 = 0$, the average and worst gap are identically zero. This is because in those cases, the optimal solution is simply selecting the highest revenue product and therefore both strategies coincide.
3. The average gap is generally increasing as the outside option utility increases. With a high outside option, we typically expect to select more products to counterbalance the effect of the no choice alternative. This can amplify the difference between ROL and RO as the likelihood that the optimal solution of the revenue-ordered by level is indeed a revenue ordered assortment decreases.

Chapter 3

Assortment Optimisation under the Two-Stage Luce Model

This chapter is reproduced with minor changes from:

1. **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment and Pricing optimisation Under the Two-Stage Luce model. *Submitted to Operations Research (13th of April of 2019)*. Under Review.

2. **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment and Pricing optimisation Under the Two-Stage Luce model. *Presented at the Informs Revenue Management & Pricing Conference, Stanford (Informs RM&P), 7th of June 2019.*

This chapter studies the assortment problem under the Two-Stage Luce model (2SLM), a discrete choice model introduced by Echenique and Saito [2018] that generalizes the multinomial logit model (MNL). The model employs an utility function as in the MNL, and a dominance relation between products. When consumers are offered an assortment S, they first discard all dominated products in S and then select one of the remaining products using the standard MNL. This model may violate the regularity condition, which states that the probability of choosing a product cannot increase if the offer set is enlarged. Therefore, the 2SLM falls outside the large family of discrete choice models based on random utility which contains almost all choice models studied in revenue management.

The first key contribution is to show that the assortment problem can be solved in polynomial time under the 2SLM. The proof is built upon two unrelated results in optimisation: the polynomial-time solvability of the maximum-independent set in a comparability graph [Möhring, 1985] and a seminal result by Megiddo [1979] that provides an algorithm to solve a class of combinatorial optimisation problems with rational objective functions in polynomial time. This is particularly appealing since the 2SLM is one of the very few choice models that goes beyond the random utility model and it allows violations the property known as *regularity*: the probability of choosing an alternative cannot increase if the offer set is enlarged. Since many decades ago, there are well-documented lab experiments where the regularity property is violated [Huber et al., 1982; Tversky and Simonson, 1993; Herne, 1997].

The second key contribution is to show that the capacitated assortment problem under the 2SLM is NP-hard, which contrasts with results on the MNL. We then propose polynomial algorithms for two interesting subcases of the capacitated assortment problem: (1) When the dominance relation is attractiveness-correlated and (2) when the transitive reduction of the dominance relation can be represented as a forest. The proofs use a strong connection between assortments under the 2SLM and independent sets.

3.1 The Two-Stage Luce model

The 2SLM [Echenique and Saito, 2018] overcomes a key limitation of the MNL: The fact that a product must have zero attractiveness if it has zero probability to be chosen in a particular assortment. This limitation means that the product cannot be chosen with positive probability in any other assortment. The 2SLM eliminates this pathological situation through the concept of *consideration function* which, given a set of products S, returns a subset of S where each product has a positive probability of being selected. Let X denotes the set of all products and let $a(x) > 0$ be the attractiveness of product $x \in X$. For notational convenience, we use a_x to denote the attractiveness of product x, i.e., $a_x = a(x)$. We extend the attractiveness function to consider the outside option, with index 0 and $a_0 = a(0) \geq 0$, to model the fact that customers may not select any product. As a result, the attractiveness function has signature $a : X \cup \{0\} \to \mathbb{R}^+$. Given an assortment $A \subseteq X$, a stochastic choice function ρ returns a probability distribution over A, i.e., $\rho(x, A)$ is the probability of picking x in the assortment A. The 2SLM is a sub-case of the general Luce model presented in Echenique and Saito [2018], and independently discovered in Ahumada and Ülkü [2018], which is defined below.

Definition 4 (General Luce Function[1], Echenique and Saito [2018])**.** A stochastic choice function ρ is called a general Luce function if there exists an attractiveness function $a \cup \{0\} : X \to \mathbb{R}^+$ and a function $c : 2^X \setminus \emptyset \to 2^X \setminus \emptyset$ with $c(A) \subseteq A$ for all $A \subseteq X$ such that

$$\rho(x, A) = \begin{cases} \frac{a_x}{\sum_{y \in c(A)} a_y + a_0} & \text{if } x \in c(A), \\ 0 & \text{if } x \notin A. \end{cases} \tag{3.1}$$

for all $A \subseteq X$. We call the pair (a, c) a general Luce model.

The function c (which is arbitrary) provides a way to capture the support of the stochastic choice function ρ. As observed in Echenique and Saito [2018], there are two interesting cases worthy of being mentioned:

1. If $c(S)$ is a singleton for all $S \subseteq X$, then $\rho(x, S)$ is a deterministic choice.
2. If $c(S) = S$ for all $S \subseteq X$, then the 2SLM coincides with the MNL.

[1]The definition is slightly different: It makes the outside option effect a_0 explicit in the denominator.

Two special cases of this model were provided in Echenique and Saito [2018]. The first is the two-stage Luce model. This model restricts c, such that the $c(A)$ represents the set of all *undominated alternatives* in A.

Definition 5 (two-Stage Luce model (2SLM), Echenique and Saito [2018])**.** A general Luce model (a, c) is called a 2SLM if there exists a strict partial order (i.e. transitive, antisymmetric and irreflexive binary relation) $\succ$ such that:

$$c(A) = \{x \in A \mid \nexists y \in A : y \succ x\}. \tag{3.2}$$

We call $\succ$ *dominance relation*.

As a result, any 2SLM can be described by an irreflexive, transitive, and antisymmetric relation $\succ$ that fully captures the relation between products. This observation allow us to describe the Two-Stage Luce Model as a Directed Acyclic Graph (DAG). The construction of this graph can be made using the following steps: Take as input $\succ$ which allow us to explicitly know whether two products $s, t \in X$ are related, i.e $s \succ t$, $t \succ s$ or they are not related at all. To construct the graph, we simply create one node for each product, and create a directed edge between s and t if $s \succ t$. An example of the graph representation is given below:

Example 6. Let $X = \{1, 2, 3, 4, 5, 6, 7\}$ and let the dominance relation $\succ$ consists of the following pairs $\succ = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (5, 6)\}$. The resulting graph can be seen below:

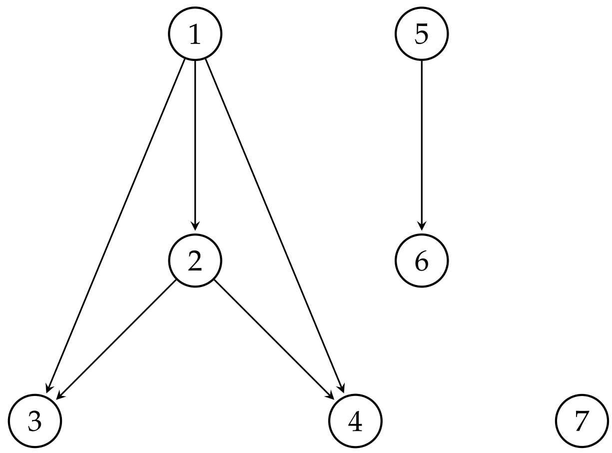


Figure 3.1: DAG representation of $\succ$ over X

Is interesting to note that the directed acyclic graph can have more than one connected component, as shown in Figure 3.1, which mean that it can be that clusters of products can be unrelated in terms of dominance. The intuition and practical use behind a graph generated in this way (as a graphical representation of a 2SLM), is that selecting a particular product to be into an assortment, implies that all his *children* or descendants are dominated and have probability zero to be chosen. In the graph, sets where $c(S) = S$ are called *anti-chains*, which means that no distinct products are connected by an edge. These particular sets are going to play a crucial role later on.

An important application of the 2SLM can be found in assortment problems where there exists a direct way to compare the products over a set of features. For illustration, consider a telecommunication company offering phone plans to consumers. A plan is characterized by a set of features such as price per month, free minutes in peak hours, free minutes in weekends, free data, price for additional data, and price per minute to foreign countries. Given two plans x and y, we say that plan x *dominates* plan y, if the price per month of x is less than that of y, and x is *at least* as good as y in every single feature. In the past, the company offered consumers a certain set of plans S_t each month t such that no plan in S_t is dominated by another plan (in S_t). The offered plans however were different each month. Using historical data and assuming that consumers preferences can be approximated using a multinomial logit, it is possible to perform a robust estimation procedure to obtain the parameters of such MNL model. Once the parameters are obtained, the assortment problem consists in finding the best assortment of phones plans S^* to maximise the expected revenue. A natural constraint in this problem consisting in enforcing that every phone plan offered in S^* cannot be dominated by any other. Section 3.1 shows that the problem discussed here can be modelled using the 2SLM and thus solving this problem is reduced to solving an assortment problem under the 2SLM.

The second model presented in Echenique and Saito [2018], which is a particular case of the 2SLM, is the *Threshold Luce Model* (TLM), where they explain dominance in terms of how big the attractiveness are when compared with each other, so c is strongly tied to a. More specifically, for a given threshold $t > 0$, the consideration set $c(S)$ for a set $S \subseteq X$ is defined as:

$$c(S) = \{y \in S \mid \nexists x \in S : a_x > (1+t)a_y\}. \tag{3.3}$$

In other words, $x \succ y$ if and only if $\frac{a_x}{a_y} > (1+t)$. Intuitively, an attractiveness ratio of more than $(1+t)$ means that the less-preferred alternative is dominated by the more-preferred alternative. Observe that the relation $\succ$ is clearly irreflexive, transitive, and antisymmetric.

The dominance relation $\succ$ can thus be represented as a Directed Acyclic Graph (DAG), where nodes represent the products and there is a directed edge (x, y) if and only if $x \succ y$. Sets satisfying $c(S) = S$ are *anti-chains* in the DAG, meaning that there are no arcs connecting them. For instance, consider the *Threshold Luce model* defined over $X = \{1, 2, 3, 4, 5\}$ with attractiveness values $a_1 = 12, a_2 = 8, a_3 = 6, a_4 = 3$ and $a_5 = 2$, and threshold $t = 0.4$. We have that $i \succ j$ iff $a_i > 1.4\, a_j$.

The DAG representing this dominance relation is depicted in Figure 3.2.

In the following example, we show that the 2SLM admits regularity violations, meaning that it is possible that the probability of choosing a product can increase when we enlarge the offered set. Since regularity is satisfied by any choice model based on random utility (RUM), this shows that the 2SLM is not contained in the

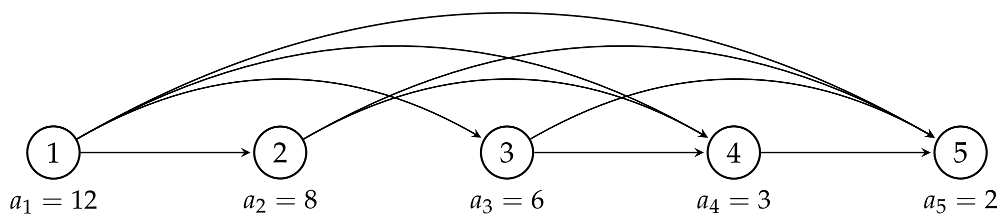


Figure 3.2: Example of a DAG for the Threshold Luce model

RUM class [2].

Example 7. Consider the following instance of the Threshold Luce model (which is a special case of the 2SLM). Let $X = \{1, 2, 3, 4\}$ with attractiveness $a_1 = 5, a_2 = 4, a_3 = 3$ and $a_4 = 3$. Consider $t = 0.4$ and the attractiveness of the outside option $a_0 = 1$. For the offer set $\{2, 3, 4\}$, the probability of selecting product 2 is 4/11 since no product dominates each other. However, if we add product 1 to the offer set, i.e. if we offer all four products, then the probability of selecting product 2 increases to 4/10, because products 3 and 4 are now dominated by product 1.

The Two-Stage Luce Model allows to accommodate different decision heuristics and market scenarios by specifying the dominance relation responding to a specific set of rules. Two examples where this can be observed are provided below.

Example 8. Feature Difference Threhsold: Assume that each product has a set of features $\mathcal{F} = \{1, \ldots, m\}$. A product x can then be represented by a m-dimensional vector $x \in \mathbb{R}^m$. Assume that the perceived relevance of each feature k is measured by a weight ν_k, so that the utility perceived by the customers can be expressed as a weighted combination of their features $u(x) = \sum_{k=1}^{m} \nu_k \cdot x_k$. The dominance relation can be defined as $x \succ y \iff u(x) - u(y) = \sum_{k=1}^{m} \nu_k (x_k - y_k) \geq T$, where $T > 0$ is a tolerance parameter that represents how much difference a customer allows before considering that an alternative dominates another. The dominance relation is irreflexive, transitive, and antisymmetric and hence it can be used to define an instance of the 2SLM. One can easily show that this model is a special case of the TLM.

Example 9. Price Levels Suppose we have N products, each product i has k_i price levels. Let x_{il} be product i with price p_{il} attached and it corresponding attractiveness a_{il}, we assume that for each product i prices p_{ik} satisfy $p_{i1} < p_{i2} < \ldots, p_{ik_i}$. Naturally, $x_{i1} \succ x_{i2} \succ \ldots \succ x_{ik_i}$, because for the same product the customer is going to select the one with the lowest price available. Each price level for each product can still dominate or be dominated by other products as well, as long as the dominance relation is irreflexive, transitive and antisymmetric. This setting can be modelled by the *Two-Stage Luce model* in a natural way.

[2] Observe that this implies that the 2SLM is not contained by the Markov chain model proposed by [Blanchet et al., 2016] since this last one belongs to the RUM class [Berbeglia, 2016].

3.2 Assortment Problems Under the Two-Stage Luce model

This section studies the assortment problem for the 2SLM using the definitions and notations presented earlier. Let $r : X \cup \{0\} \to \mathbb{R}^+$ be a revenue function associated with each product and satisfying $r(0) = 0$. The expected revenue of a set $S \subseteq X$ is given by

$$R(S) = \sum_{i \in c(S)} \rho(i, S) r(i). \tag{3.4}$$

The assortment problem amounts to finding a set

$$S^* \in \operatorname*{argmax}_{S \subseteq X} R(S)$$

yielding an optimal revenue of

$$R^* = \max_{S \subseteq X} R(S).$$

Observe that every subset $S \subseteq X$ can be uniquely represented by a binary vector $x \in \{0,1\}^n$ such that $i \in S$ if and only if $x_i = 1$. Using this bijection, the search space for S^* can be restricted to

$$\mathcal{D} = \{x \in \{0,1\}^n \mid \forall s \succ t : x_s + x_t \leq 1\}$$

where $\mathcal{D}$ represents all the subsets satisfying $S = c(S)$, which means that no product on S dominates another product in S. There is always an optimal solution S^* that belongs to $\mathcal{D}$ because $R(S) = R(c(S))$ and $c(S) \in D$ for all sets S in X. As a result, the Assortment Problem under the 2SLM (`AP-2SLM`) can be formulated as

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \frac{\sum_{i=1}^n r_i a_i x_i}{\sum_{i=1}^n a_i x_i + a_0} \\ \text{subject to} \quad & x \in \mathcal{D} \end{aligned} \tag{AP-2SLM}$$

where r_i and a_i represent $r(i)$ and $a(i)$ for simplicity.

An effective strategy for solving many assortment problems consists in considering revenue-ordered assortments, which are obtained by choosing a threshold ρ and selecting all the products with revenue at least ρ. This strategy leads to an optimal algorithm for the assortment problem under the MNL. Unfortunately, it fails under the 2SLM because adding a highly attractive product may remove many dominated products whose revenues and utilities would lead to a higher revenue.

Example 10 (Sub-Optimality of Revenue-Ordered Assortments)**.** Consider a Threshold Luce model with $X = \{1, 2, 3\}$, revenues $r_1 = 88, r_2 = 47, r_3 = 46$, attractiveness $a_0 = 55, a_1 = 13, a_2 = 26, a_3 = 15$ and $t = 0.6$. Then $x \succ y$ iff $a_x > 1.6\ a_y$ which gives $2 \succ 1$ and $2 \succ 3$. Consider the sets $S \subseteq X$ satisfying $S = c(S)$:

S	$R(S)$
$\{1\}$	16.824
$\{2\}$	15.086
$\{3\}$	9.857
$\{1,3\}$	22.096

Table 3.1: List of subsets S and their associated revenues $R(S)$

The optimal revenue is given by assortment $\{1,3\}$, while the best revenue-ordered assortment under the 2SLM is $S = \{1\}$, yielding almost 24% less revenue.

Our next example indicates that the 2SLM may also violate the regularity condition.

Example 11. [Violation of the Regularity Condition] Consider $X = \{1,2,3\}$, the utilities are $a(1) = 1, a(2) = 1, a(3) = 2$, the attractiveness of the outside option is $a(0) = 1$ and the only dominance relation is $2 \succ 3$. Whenever product 2 is offered, product 3 is never selected. Consider the following assortments $S = \{1,3\}$ and $S' = \{1,2,3\}$. According the 2SLM, the probability of selecting product 1 on each of these assortments is:

$$\rho(1,S) = \frac{a(1)}{a(1) + a(3) + a(0)} = \frac{1}{4} = 0.25$$

$$\rho(1,S') = \frac{a(1)}{a(1) + a(2) + a(0)} = \frac{1}{3} = 0.\overline{3}$$

We have that $S \subset S'$ but $\rho(1,S) < \rho(1,S')$, which violates regularity..

To solve problem AP-2SLM, consider first the `MaxAtt` problem defined over the same set of constraints. Given weights $c_i \in \mathbb{R}$ $(1 \leq i \leq n)$, the `MaxAtt` problem is defined as follows:

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} \quad & x \in \mathcal{D} \end{aligned} \tag{MaxAtt}$$

We now show that (MaxAtt) can be reduced to the maximum weighted independent set problem in a directed acyclic graph with positive vertex weights. An independent set is a set of vertices I such that there is no edge connecting any two vertices in I. The maximum weighted independent set problem (`MWIS`) can be stated as follows:

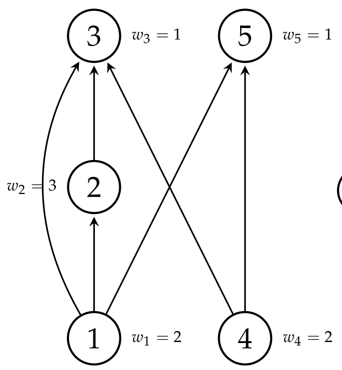
Definition 6. *Maximum Weighted Independent Set Problem:* Given a graph $G = (V, E)$ with a weight function $w : V \to \mathbb{R}$, find an independent set $I^* \in \text{argmax}_{I \in \mathcal{I}} \sum_{i \in I} w(i)$, where $\mathcal{I}$ is the set of all independent sets.

Recall that the dominance relation can be represented as a DAG G which includes an arc (u, v) whenever $u \succ v$. As a result, the condition $x \in \mathcal{D}$ implies that any feasible solution to (MaxAtt) represents an independent set in G and maximising $\sum_{i=1}^{n} c_i x_i$ amounts to finding the independent set maximising the sum of the weights. Since the dominance relation is a partial order, the DAG representing the dominance relation is a comparability graph. The following result is particularly useful.

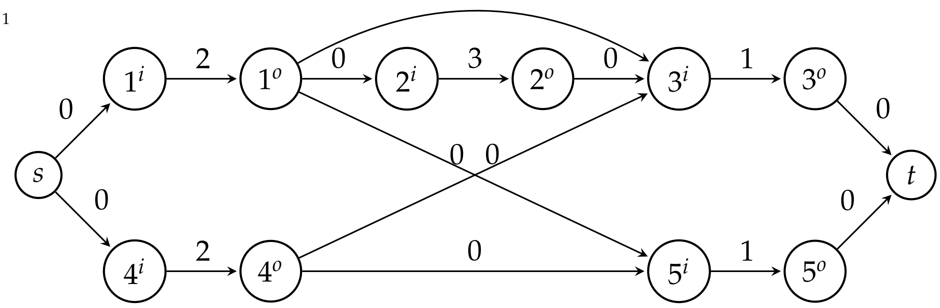
Theorem 2 (Möhring [1985])**.** *The maximum weighted independent set is polynomially-solvable for comparability graphs with positive weights.*

To give intuition about how this works, the reduction proposed in Möhring [1985] is as follows: starting with the Directed Acyclic Graph G, the vertex set V, $\mathcal{R}$ as the edge set, and $w : V \to \mathbb{R}^+$ a positive weight function over the vertex set, we construct a network N_G through the following steps: first, add two new vertices s, t. Add edges (s, a) and (b, t) for each minimal vertex a and maximal vertex b in G, where a vertex is *minimal* (resp. *maximal*) if v does not have any ingoing (resp. outgoing) edges. These edges and all the edges $(u, v) \in \mathcal{R}$ have lower capacities 0 and upper capacities $+\infty$. Then, each vertex v in V is split into two vertices v^i, v^o_i (in and out vertices), and all incoming edges of v now point at v^i and all the outgoing edges of v have v^o as starting point. Also the vertices v^i, v^o are joined by a new edge (v^i, v^o) with lower capacity $w(v)$ (the weight of v) and upper capacity $+\infty$. An example of the reduction from G to N_G can be seen in the following figure:

Example 12. [[Möhring, 1985, p.66]] Let $X = \{1, 2, 3, 4, 5\}$ and let $\succ$ defined by $\mathcal{R} = \{(1, 2), (2, 3), (1, 3), (1, 5), (4, 3), (4, 5)\}$. Then G and the network N_G are given by:



The original graph G



The Network N_G

By the Min-Flow Max-Cut Theorem (see Ford and Fulkerson [2010]), the minimal value of a feasible s,t-flow in N_G is the same as the maximum capacity of an s,t-cut in N_G, this is:

$$v = \max_{S,T} \left[\sum_{\substack{(uv) \in E' \\ u \in S, v \in T}} l((u, v)) - \sum_{\substack{(uv) \in E' \\ u \in T, v \in S}} c((u, v)) \right], \tag{3.5}$$

where S,T is an s,t-cut of $N_G = (V', E')$, i.e. $s \in S$, $t \in T$, $S \cup T = V'$ and $S \cap T = \emptyset$. Also, $l((u,v))$ and $c((u,v))$ denotes the lower and upper capacity of an edge $(u,v) \in E'$.

Möhring [1985] showed that the set $M = \{v \in V, (v^i, u) \in E', v^i \in S, u \in T\}$ induced by an optimal cut, is the maximum weighted anti-chain in G (Theorem 1.25). In the previous example, the optimal cut is $S = \{s, 1^i, 1^o, 2^i, 4^o\}$, $T = V' \setminus S$ and $M = \{2, 4\}$, with a total weight of 5. We can solve this problem using any min-flow algorithm [Even, 1979; Ford and Fulkerson, 2010].

Is also worth to notice that another method to solve the *maximum weighted independent set* problem MaxAtt, is provided in Grötschel et al. [1981], since the ellipsoid method can be used to solve this problem over perfect graphs, and comparability graphs are perfect by Mirki's Theorem [Mirsky, 1971]. We are now in position to present our first result.

Lemma 2. (MaxAtt) *is polynomial-time solvable.*

Proof. We first show that we can ignore those products with a negative weight. Let $\hat{X} = \{i \in X \mid c_i > 0\}$ and $\hat{\mathcal{D}} = \{x \in \{0,1\}^n \mid \forall s, t \in \hat{X}, s \succ t : \ x_s + x_t \leq 1\}$. Solving (MaxAtt) is equivalent to solving:

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \sum_{i \in \hat{X}} c_i x_i \\ \text{subject to} \quad & x \in \hat{\mathcal{D}} \end{aligned} \tag{Reduced MaxAtt}$$

Indeed, consider an optimal solution x^* to Problem MaxAtt and assume that there exists $i \in X$ such that $c_i < 0$ and $x_i^* = 1$. Define $\hat{x}$ like x^* but with $\hat{x}_i = 0$. $\hat{x}$ has a strictly greater value for the objective function in Reduced MaxAtt than x^* has, and is feasible since setting a component to zero cannot violate any constraint (i.e., $\hat{x} \in \mathcal{D}$). This contradicts the optimality of x^*. Now Problem `Reduced MaxAtt` can be reduced to solving an instance of Problem `MWIS` in a DAG with positive weights that corresponds to the dominance relation. This DAG is a comparability graph and the result follows from Theorem 2. □

The next step in solving the assortment problem under the 2SLM relies on a result by Megiddo Megiddo [1979]. Let D be a domain defined by some set of constraints and consider Problem A

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} \quad & x \in D \end{aligned} \tag{A}$$

and its associated Problem B:

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \frac{a_0 + \sum_{i=1}^{n} a_i x_i}{b_0 + \sum_{i=1}^{n} b_i x_i} \\ \text{subject to} \quad & x \in D. \end{aligned} \tag{B}$$

Using this notation, Megiddo's theorem can be stated as follows.

Theorem 3 (Megiddo [1979]). *If Problem A is solvable within $O(p(n))$ comparisons and $O(q(n))$ additions, then Problem B is solvable in $O(p(n)(q(n)+p(n)))$ time.*

We are now in position to state our main theorem of this section.

Theorem 4. *The assortment problem under the Two-Stage Luce model is polynomial-time solvable.*

Proof. Recall that the assortment problem under the 2SLM (`AP-2SLM`) can be formulated as

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \frac{\sum_{i=1}^{n} r_i a_i x_i}{\sum_{i=1}^{n} a_i x_i + a_0} \\ \text{subject to} \quad & x \in \mathcal{D} \end{aligned} \tag{3.6}$$

where $\mathcal{D} = \{x \in \{0,1\}^n \mid \forall s \succ t : x_s + x_t \leq 1\}$.

The problem of maximising the numerator in (3.6) is exactly the `MaxAtt` problem. By Lemma 2, this is polynomial-time solvable. Now observe that (3.6) (i.e., problem AP-2SLM) can be seen as a Problem B. Therefore, by Theorem 3, the assortment problem under the 2SLM is solvable in polynomial time. □

In addition to solving the assortment problem under the 2SLM, Theorem 4 is interesting in that it solves the assortment problem under a Multinomial Logit with a specific class of constraints. It can be contrasted with the results by Davis et al. [2013], where feasible assortments satisfy a set of totally unimodular constraints. They show that the resulting problem can be solved as a linear program. However, the 2SLM introduces constraints that are not necessarily totally unimodular as we now show.

Example 13. Consider $X = \{1,2,3,4\}$ and $1 \succ 3, 1 \succ 4, 2 \succ 3, 2 \succ 4$, and $3 \succ 4$. The constraint matrix that defines the feasible space ($\mathcal{D}$) for this instance is:

$$M = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

where each row represents a constraint $x_u + x_v \leq 1$. meaning that just one end of the edge can be selected at the time. Camion [1965] proved that M is totally unimodular if and only if, for every (square) Eulerian submatrix A of M, $\sum_{i,j} a_{ij} \equiv 0 \pmod 4$. Consider the sub-matrix corresponding to the first, second, and fifth rows and the first, third, and fourth columns

$$N = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Matrix N is eulerian (The sums of every element on each row or on each column is a multiple of 2). But the sum of all elements of N is $6 \not\equiv 0 \pmod 4$ and hence M is not totally unimodular.

We close this section by explaining how our results can be extended to a more general setting. Gallego et al. [2015] proposed the general attraction model (GAM) to describe customer behaviour, that alleviates some deficiencies of the MNL. More specifically, the intuition behind this choice model is that whenever a product is not offered, then its absence can potentially increase the probability of the no-purchase alternative, as consumers can potentially look for the product elsewhere, or at a later time. To achieve this effect, for each product j the model considers two different weights: v_j and w_j, usually with $0 \leq w_j \leq v_j$. If product j is offered, then its preference weight is v_j. But if j is not offered, then the preference weight of the outside option is increased by w_j. For all $j \in X$, let $\tilde{v}_j = v_j - w_j$ and $\tilde{v}_0 = v_0 + \sum_{k \in X} w_k$. Using this notation, the probabilities associated with the GAM model can be recovered by means of the following equation:

$$\rho(j, S) = \begin{cases} \frac{v_j}{\sum_{i \in S} \tilde{v}_j + \tilde{v}_0} & \text{if } j \in S, \\ 0 & \text{if } j \notin S. \end{cases} \tag{3.7}$$

Observe that the resulting assortment problem will has the same functional form than problem AP-2SLM, with a slight modification on the coefficients in the denominator. Thus, we can apply the same solution technique described in Theorem 4 to find the optimal assortment for the GAM.

3.3 The Capacitated Assortment Problem

In many applications, the number of products in an assortment is limited, giving rise to capacitated assortment problems. Let C $(1 \leq C \leq n)$ be the maximum number of products allowed in an assortment. The Capacitated Assortment Problem under the Two-Stage Luce Model (`C2SLMAP`) is given by

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \frac{\sum_{i=1}^{n} r_i a_i x_i}{\sum_{i=1}^{n} a_i x_i + a_0} \\ \text{subject to} \quad & x \in \mathcal{D}_C \end{aligned} \tag{C2SLMAP}$$

where $\mathcal{D}_C = \{x \in \{0,1\}^n \mid \forall (s,t) \in \mathcal{R} \quad x_s + x_t \leq 1 \wedge \sum_{i=1}^{n} x_i \leq C\}$. As before, it is useful to define its capacitated maximum-attractiveness counterpart (`C-MaxAtt`), i.e.,

$$\begin{aligned} \underset{x}{\text{maximise}} \quad & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} \quad & x \in \mathcal{D}_C \end{aligned} \tag{C-MaxAtt}$$

This section first proves that the capacitated assortment problem under the 2SLM is NP-hard. The reduction uses the Maximum Weighted Budgeted Independent Set (`MWBIS`) problem proposed by Bandyapadhyay [2014] which amounts to finding a maximum weighted independent set of size not greater than C. Kalra et al. [2017] showed that Problem (`MWBIS`) is NP-hard for bipartite graphs.

Theorem 5. *Problem* (C2SLMAP) *is NP-hard (under Turing reductions).*

Proof. The proof considers four problems:

1. Problem (MWBISBP): Maximum weighted independent set of size at most C for bipartite graphs.

2. Problem (MWEBISBP): Maximum weighted independent set of size equal to C for bipartite graphs.

3. Problem (EC2SLMAP): Optimal assortment under the General Luce model of size C.

4. Problem (C2SLMAP): Optimal capacitated assortment under the Two-Stage Luce model of size at most C.

The proof shows that Problems (MWEBISBP), (EC2SLMAP), and (C2SLMAP) are NP-hard, using the NP-hardness of Problem (MWBISBP) [Kalra et al., 2017] as a starting point.

First observe that Problem (MWEBISBP) is NP-hard under Turing reductions. Indeed, Problem (MWBISBP) can be reduced to solving C instances of Problem (MWEBISBP) with budget c ($1 \leq c \leq C$).

We now show that Problem (EC2SLMAP) is NP-hard. Consider Problem (MWEBISBP) over a bipartite graph $G = (V = V_1 \cup V_2, E)$, where $V_1 \cap V_2 = \emptyset$, every edge $(v_1, v_2) \in E$ satisfies $v_1 \in V_1$ and $v_2 \in V_2$, w_v is the weight of vertex v, and C is the budget. We show that Problem (MWEBISBP) over this bipartite graph can be polynomially reduced to Problem (EC2SLMAP). The reduction assigns each vertex v to a product with $a(v) = 1$ and $r_v = w_v$, sets $a_0 = 0$, and has a capacity C. Moreover, the reduction uses the following dominance relation: $v_1 \succ v_2$ iff $(v_1, v_2) \in E$. This dominance relation is irreflexive, anti-symmetric, and transitive, since the graph is bipartite. A solution to Problem (MWEBISBP) is a feasible solution to Problem (EC2SLMAP), since the independent set cannot contain two vertices v_1, v_2 with $v_1 \succ v_2$ by construction. Similarly, a feasible assortment is an independent set, since the assortment cannot select two vertices $v_1 \in V_1$ and $v_2 \in V_2$ with $(v_1, v_2) \in E$, since $v_1 \succ v_2$. The objective function of Problem (EC2SLMAP) reduces to maximising

$$\frac{1}{C} \sum_{v \in V} r_v x_v$$

which is equivalent to maximising $\sum_{v \in V} r_v x_v$ since exactly C products will be selected by every feasible assortment. The result follows by the NP-hardness of Problem (MWEBISBP).

Finally, Problem (C2SLMAP) is NP-hard under Turing reductions. Indeed, Problem (C2SLMAP) can be reduced to solving C instances of Problem (EC2SLMAP) with capacity c ($1 \leq c \leq C$). $\square$

It is interesting to mention that Problem (C-MaxAtt) is equivalent to finding an antichain of maximum weight among those of cardinality at most C. This problem (MWLA) was proposed by Shum and Trotter [1996] and its complexity was left open, but the above results show that it is also NP-hard. Bandyapadhyay [2014] studied Problem (MWBIS) for various types of graphs (e.g., trees and forests), but the dominance relation of the 2SLM can never be a tree since it is transitive (unless we consider a graph with a single vertex).

In light of this NP-hardness result, the rest of this section presents polynomial-time algorithms for two special cases of the dominance relation.

3.3.1 The Two-Stage Luce model over Tree-Induced Dominance Relations

Let $\mathcal{R}_\succ$ be the transitive reduction of the irreflexive, antisymmetric, and transitive relation $\succ$. This section considers the capacitated assortment problem when the relation $\mathcal{R}_\succ$ can be represented as a tree. Without loss of generality, we can assume that the tree contains all products. Otherwise, we can add another product with zero weight that dominates all original products. This new product will be the root of the tree and the products not in the original tree will be the children of the root. Similarly, the same transformation applies to the case when $\mathcal{R}_\succ$ is a forest. Here all the trees in the forest will be children of the new product.

We show how to solve Problem (C-MaxAtt). The result follows again by applying Megiddo's theorem. The first step of the algorithm simply removes all products with negative weight: Their children can be added to the parent of the deleted vertex. The main step then solves (C-MaxAtt) bottom-up using dynamic programming from the leaves. For simplicity, we present the recurrence relations to compute the weight of the optimal assortment. It is easy to recover the optimal assortment itself. The recurrence relations compute two functions:

1. $\mathcal{A}(k,c)$ which returns the weight of an optimal assortment using product k and its descendants in the tree representation of $\mathcal{R}_\succ$ for a capacity c;

2. $\mathcal{A}^+(S,c)$ which, given a set S of vertices that are children of a vertex k, returns the weight of an optimal assortment using the products in S and their descendants for a capacity c.

The key intuition behind the recurrence is as follows. If v is a vertex and v_1 and v_2 are two of its children, v_1 does not dominate v_2 or any of its descendants. Hence, it suffices to compute the best assortments producing $\mathcal{A}(v_1,0),\ldots,\mathcal{A}(v_1,C)$ and $\mathcal{A}(v_2,0),\ldots,\mathcal{A}(v_2,C)$ and to combine them optimally. The recurrence relations are defined as follows ($v \in X$ and $1 \le c \le C$):

$$\begin{aligned}
&\mathcal{A}(v,0) = 0;\\
&\mathcal{A}(v,c) = \max(c_v, \mathcal{A}^+(\mathit{children}(v),c));
\end{aligned}$$

and

$$\mathcal{A}^+(\emptyset, 0) = 0;$$
$$\mathcal{A}^+(S, c) = \max_{\substack{n_1, n_2 \geq 0 \\ n_1 + n_2 = c}} \mathcal{A}^+(S \setminus \{e\}, n_1) + \mathcal{A}(e, n_2) \text{ where } e = \operatorname*{argmax}_{i \in S} c_i.$$

where *children(p)* denotes the children of product p in the tree. Note that $\mathcal{A}^+(S, c)$ is computed recursively to obtain the best assortment from the products in S and their descendants. Using these recurrence relation, the following Theorem follows:

Theorem 6. *Let $\succ$ a dominance relation whose relation $\mathcal{R}_\succ$ is a tree containing all products. The capacitated assortment problem under the 2SLM and $\succ$ is polynomial-time solvable.*

Proof. By Theorem 3, it suffices to show that Problem (C-MaxAtt) is solved by the recurrences in polynomial time. The correctness of recurrence $\mathcal{A}(v, c)$ comes from the fact that vertex v dominates all its descendants and cannot be present in any assortment featuring any of them. The correctness of recurrence $\mathcal{A}^+(S, c)$ follows from the fact that e is not dominated by, and does not dominate, any element in S, since they are all children of the same node. This also holds for the descendants of e and the descendants of the elements in S. Hence, the optimal assortment is obtained by splitting the capacity c into n_1 and n_2 and merging the best assortment for $\mathcal{A}^+(S, n_1)$ and $\mathcal{A}(e, n_2)$ for some $n_1, n_2 \geq 0$ summing to c. The recurrences can be solved in polynomial time since the computation for each vertex v and capacity c takes $O(n\ C)$ time, giving an overall time complexity of $O(n^2\ C^2)$. □

3.3.2 The Attractiveness-Correlated Two-Stage Luce model

The second special case considers a dominance relation that is correlated with attractiveness.

Definition 7 (Attractiveness-Correlated Two-Stage Luce model)**.** A Two-Stage Luce model is attractiveness-correlated if the dominance relation satisfies the following two conditions:

1. If $x \succ y$, then $a_x > a_y$.
2. If $x \succ y$ and $a_z > a_x$, then $z \succ y$.

The first condition simply expresses that product x can only dominate product y if the attractiveness of x is greater than the attractiveness of y. The second condition ensures that, if x dominates y, then any product whose attractiveness is greater than x also dominates y. The induced dominance relation is irreflexive, anti-symmetric, and transitive. A particular case of this model, is the Threshold Luce model.

When customers follow the Threshold Luce model, they form their consideration sets based on the attractiveness of products. Without loss of generality, we can assume $a_1 \geq a_2 \geq \ldots \geq a_n$, unless stated otherwise. For a set S, the associated consideration set $c(S)$ may be a proper subset of S, but for the purpose of assortment optimisation, we don't have incentives to offer sets including products that are not even consider by customers, so we can restrict our search for optimal solutions to sets where $c(S) = S$. A necessary and sufficient condition for this to happen is $\frac{\max_{i \in S} a_i}{\min_{i \in S} a_i} \leq 1 + t$. Meaning that largest ratio between attractiveness is not greater than $1 + t$, so no dominance relation appears.

The firm now needs to carefully balance the inclusion of high-attractiveness products and their prices to maximise the revenue. In the following example we show that *revenue ordered assortments* are not optimal under the Threshold Luce Model. In fact, this strategy can be arbitrarily bad.

Example 14 (Revenue ordered assortments are not optimal)**.** Consider the following product configuration. Let $N + 1$ products, with prices p_1 for the first product, and αp_1 for the rest of them, with $\alpha < 1$. The attractiveness for all products is a_1 for the first product and γa_1 for all the rest, such as in the presence of product 1, all the rest of the products are ignored. To complete the set up, let a_0 the attractiveness of the outside option. The best revenue ordered assortment is to consider product 1, given a revenue of:

$$R' = R(\{1\}) = \frac{p_1 a_1}{a_1 + a_0}$$

But, if N is big enough (at least bigger than $\frac{1}{\alpha\gamma}$), is more profitable to show $S_N = X \setminus \{1\}$, resulting in a revenue of:

$$R^* = R(S_N) = \frac{N \cdot \alpha p_1 \gamma a_1}{N \cdot \alpha \gamma a_1 + a_0}$$

Now, if we calculate the ratio if this two values, R' and R^* and let N tend to infinity we have:

$$\begin{aligned}
\frac{R'}{R^*} &= \lim_{N\to\infty} \frac{\frac{p_1 a_1}{a_1 + a_0}}{\frac{N \cdot \alpha p_1 \gamma a_1}{N \cdot \alpha\gamma a_1 + a_0}} \\
\frac{R'}{R^*} &= \lim_{N\to\infty} \frac{p_1 a_1}{a_1 + a_0} \cdot \frac{N \cdot \alpha\gamma a_1 + a_0}{N \cdot \alpha p_1 \gamma a_1} \\
\frac{R'}{R^*} &= \frac{a_1}{a_1 + a_0}
\end{aligned} \tag{3.8}$$

Observe that this last expression is the market share of offering just product 1, which can be arbitrarily bad by either making a_1 as small as desired, or making the outside option more attractive.

The capacitated assortment optimisation can be solved in polynomial time under the Attractiveness-Correlated Two-Stage Luce model. Consider an assortment whose product with the largest attractiveness is k. This assortment cannot contain any product dominated by k. Moreover, if k_1 and k_2 are two other products in this assortment, then k_1 cannot dominate k_2 since k would also dominate k_2. As a result, consider the set

$$X_k = \{i \in X \mid a_i \leq a_k \ \& \ k \not\succ i\}.$$

No product in X_k dominates any other product in X_k and hence the `C2SLMAP` reduces to a traditional assortment problem under the MNL. This idea is formalized in Algorithm 1, where `CMLMAP` is a traditional algorithm for the MNL. The algorithm considers each product in turn and the products that it does not dominate and applies a traditional capacitated assortment optimisation under the MNL. The best such assortment is the solution to the capacitated assortment under the attractiveness-correlated 2SLM.

Algorithm 1: Capacitated Assortment optimisation under the Attractiveness-Correlated 2SLM.

```
Data: X, ≻, r, a
Result: Optimal Assortment S*
R(S*) = 0 for k = 1, ..., n do
    X_k = {i ∈ X | a_i ≤ a_k & k ⊁ i}
    S_k = CMLMAP(X_k, r, a)
    if R(S_k) > R(S*) then
        S* = S_k
    end
end
return S*
```

Theorem 7. *`C2SLMAP` can be solved in polynomial time for Attractiveness-Correlated instances.*

Proof. To show correctness, it suffices to show that the optimal assortment must be a subset of one of the X_k $(1 \leq k \leq n)$. Let A be the optimal assortment and assume that k is its product with the largest attractiveness (break ties randomly). A must be included in X_k since otherwise it would contain a product x such that $k \succ x$ (contradicting feasibility) or such that $a(x) > a(k)$ (contradicting our hypothesis). The correctness then follows since there is no dominance relationship between any two elements in each of X_k. The claim of polynomial-time solvability follows from the availability of polynomial-time algorithms for the assortment problem under the MNL and the fact that are exactly n calls to such an algorithm. □

3.4 Assortment Optimisation - Numerical Results

This section presents some numerical results on the performance of revenue-ordered assortments (RO) against our proposed strategy detailed in Section 3.2, which we call `2SLM-OPT`. In order to do this, we variate the number of products n, the attractiveness of the outside option a_0 and the density d of the graph, which we use as the probability that a dominance relation is active for each pair of products[3]. Theoretically, as shown in Example 14, the optimality gap can be as large desired. But in practice, we were able to found gaps as large as 95.40%.

Each tested family or class of instances is defined by essentially three numbers: the number of products n, the attractiveness of the outside option a_0, and the density d, that controls the probability that a dominance edge exists, and then we also compute the transitive closure over the resulting graph. It is worth noticing that we did not consider the case $a_0 = 0$ because in those cases, the optimal solution is simply selecting the highest revenue product and therefore both strategies coincide. In total, we experimented with 48 classes or families of instances, each containing 250 instances. In each specific instance, revenues and utilities are drawn from an uniform distribution between 0 and 10. We ran both strategies (RO and `2SLM-OPT`) and report the average and worst optimality gap for the RO strategy. We are not providing running times, because as expected, `2SLM-OPT` takes more time than RO, but all instances can be solved very fast in practice (less than half a second). Table 3.2 presents the results which can be summarized as follows:

1. The average gap tends to increase with the number of products, reaching about 14% for 30 products. The worst gap is more instance-dependent (as it strongly depends on the dominance structure, and how revenues are matched with attractiveness) so it can be large both in smaller and larger instances. However, it tends to increase with the density of the dominance graph, as it is more likely for RO to choose a product that dominates potential contributors whose inclusion can be more profitable than keeping the higher attractiveness one.

2. The average gap generally widens as the outside option attractiveness increases. With a high outside option, we typically expect to select more products to counterbalance the effect of the no-choice alternative. This can amplify the difference between `2SLM-OPT` and RO as the likelihood that the optimal solution turns out to be revenue-ordered decreases, given the randomness of the dominance relation.

3. With higher densities, is more likely to *make a mistake* using revenue ordered assortments and include a product that dominates many *potential contributors* that considered together, might be more profitable. Thus, both the average and worst gap widens as the density increases in general. The exception occurs at the higher end of densities where not many products can be included without provoking dominances. Here the solutions of both strategies tends to be similar and select a few higher revenue products. This is also interesting from a managerial standpoint: when customers have more clarity on what products are clearly superior in comparison, this might drift the offered assortment to be smaller, compared against when customers does not have a clear hierarchy among products.

[3] used as the probability that an edge in the dominance graph occurs

(n, a_0, d)	RO Assortments			2SLM-OPT
	Avg. Gap (%)	Worst Gap (%)	Avg. Cardinality	Avg. Cardinality
(5,1,0.2)	0.476	48.899	1.484	1.496
(5,1,0.4)	1.532	80.404	1.384	1.376
(5,1,0.8)	2.812	71.888	1.08	1.084
(5,2,0.2)	1.173	73.387	1.82	1.816
(5,2,0.4)	1.827	69.8	1.504	1.536
(5,2,0.8)	4.529	94.759	1.116	1.14
(5,4,0.2)	1.835	69.574	2.108	2.104
(5,4,0.4)	3.133	61.627	1.784	1.82
(5,4,0.8)	5.378	69.555	1.2	1.228
(5,8,0.2)	1.789	64.546	2.284	2.34
(5,8,0.4)	5.927	70.854	1.884	1.988
(5,8,0.8)	6.933	91.335	1.168	1.244
Avg. $n = 5$	3.112	72.219	1.568	1.597667
(10,1,0.2)	0.68	51.339	1.872	1.896
(10,1,0.4)	2.388	63.414	1.5	1.524
(10,1,0.8)	3.997	95.49	1.092	1.076
(10,2,0.2)	1.385	49.292	2.272	2.296
(10,2,0.4)	3.275	90.659	1.604	1.664
(10,2,0.8)	6.495	73.787	1.132	1.148
(10,4,0.2)	1.984	61.872	2.612	2.764
(10,4,0.4)	5.734	90.983	1.92	2.112
(10,4,0.8)	7.107	86.55	1.156	1.224
(10,8,0.2)	3.509	41.995	3.08	3.2
(10,8,0.4)	6.592	82.358	2.028	2.304
(10,8,0.8)	8.916	92.576	1.172	1.304
Avg. $n = 10$	4.3385	73.35958333	1.786666667	1.876
(20,1,0.2)	1.067	36.45	2.18	2.216
(20,1,0.4)	2.664	82.68	1.448	1.556
(20,1,0.8)	2.884	74.534	1.1	1.092
(20,2,0.2)	2.349	40.095	2.46	2.652
(20,2,0.4)	3.452	41.717	1.696	1.856
(20,2,0.8)	5.112	83.79	1.132	1.192
(20,4,0.2)	3.786	34.659	2.84	3.184
(20,4,0.4)	8.575	73.075	1.848	2.14
(20,4,0.8)	7.749	86.321	1.152	1.284
(20,8,0.2)	5.938	68.465	3.352	3.856
(20,8,0.4)	8.88	52.627	2.088	2.616
(20,8,0.8)	10.204	94.021	1.152	1.392
Avg. $n = 20$	5.221666667	64.03616667	1.870666667	2.086333
(30,1,0.2)	1.762	20.877	2.068	2.228
(30,1,0.4)	3.34	83.702	1.44	1.616
(30,1,0.8)	3.773	62.764	1.056	1.108
(30,2,0.2)	3.084	43.736	2.544	2.864
(30,2,0.4)	5.554	79.378	1.64	1.968
(30,2,0.8)	5.499	86.544	1.072	1.148
(30,4,0.2)	4.721	53.873	2.984	3.464
(30,4,0.4)	8.046	74.267	1.876	2.3
(30,4,0.8)	9.045	92.51	1.14	1.304
(30,8,0.2)	7.623	46.498	3.368	4.188
(30,8,0.4)	14.266	91.851	1.916	2.684
(30,8,0.8)	11.422	75.239	1.132	1.412
Avg. $n = 30$	6.51125	67.60325	1.853	2.190333

Table 3.2: Numerical experiments comparing the revenue ordered assortment strategy (RO) and our proposed strategy 2SLM-OPT. For each class of instances, we display the average optimality gap and the worst-case gap, as well as the computing time and the cardinality of the offered set.

3.5 Threshold Luce model: assortment optimisation problem extensions

3.5.1 Market Share Maximisation with heterogeneous customers

Being able to offer a set that maximises the likelihood of a purchase is a problem often studied in the litertaure. This problem is known in the literature as *market share maximisation*. Using the same notation than for the TLM in Section 3.2, the market share obtained by showing a set $S \subseteq X$ under the TLM can be written as:

$$MS(S) = \sum_{i \in c(S)} \rho(i, S). \tag{3.9}$$

We can extend this definition to the following setting. Suppose that the customer base is heterogeneous on the Threshold value, meaning that they are more or less tolerant when faced relative differences between products perceived attractiveness. Let $M > 1$ be the number of customer classes. For each one of the classes j, let $t_j > 0$ and $\alpha_j \geq 0$ be the threshold and the proportion of the population associated to class j, where $\sum_{j \in [M]} \alpha_j = 1$, capturing the fact that the fractions of the population add up to 1. We need to slightly redefine the consideration set as follows:

Definition 8. Given a threshold t, a set S , we say that $i \succ_t j$ with $i, j \in S$ if and only if:

$$\frac{a_i}{a_j} > (1 + t)$$

And the consideration set associated with this threshold t is:

$$c(S, t) = \{j \in S \mid \nexists i \in S : i \succ_t j\}, \tag{3.10}$$

Given these two definitions, we can calculate the expected market share for the heterogeneous case as follows:

$$MS(S, p) = \sum_{j \in [M]} \alpha_j \cdot \frac{\sum_{i \in c(S, t_j)} a_i}{\sum_{i \in c(S, t_j)} a_i + a_0} \tag{3.11}$$

Without loss of generality, we label the customer classes such that $t_1 \leq t_2 \leq \ldots \leq t_m$. This indexing implies that customers are progressively more tolerant to more dissimilar attractiveness. This indexing allow us to state the following proposition.

Proposition 6 (Nested consideration sets). *Let $S \subseteq X$. Then $c(S, t_1) \subseteq c(S, t_2) \subseteq \ldots \subseteq c(S, t_m)$.*

Proof. Let $i < j$. Is enough to show that if $k \in c(S, t_i) \implies k \in c(S, t_j)$. Indeed, if $k \in c(S, t_i)$ implies:

$$\begin{aligned}
\frac{a_k}{a_{k'}} \leq& (1+t_i) \quad \forall k' \in S \\
\frac{a_k}{a_{k'}} \leq& (1+t_i) \leq (1+t_j) \quad \forall k' \in S \quad \text{because } t_i \leq t_j
\end{aligned}$$

so we have $\frac{a_k}{a_{k'}} \leq (1+t_j)$, which means that product k belong to $c(S, t_j)$, concluding the proof. □

In many customer choice models, like the Multinomial Logit, and all Random Utility Models (RUM), offering more products *always* yields to a greater market share. However, this is not the case for the Threshold Luce model as we can see in the following example [4].

Example 15 (Market Share Maximisation - Showing more can decrease the market share)**.** Let us revisit example 7. Consider four products with attractiveness $a_1 = 5, a_2 = 4, a_3 = 3$ and $a_4 = 3$, the threshold $t = 0.4$ and the attractiveness of the outside option $a_0 = 1$. For the offer set $\{2, 3, 4\}$, the total market share is $10/11$, since no product dominates each other. However, if we add product 1 to the offer set, i.e. if we offer all four products, then the total market share drops to $9/11$, because products 3 and 4 are now dominated by product 1, and this effect outweighs the benefit of adding alternative 1.

In fact, if we stretch this example a little bit further, we can show that the showing all products is arbitrarily bad when compared to the optimal solution. A detailed example is provided below.

Example 16 (Market Share Maximisation - Show all strategy can be arbitrarily bad)**.** Suppose we have a product (let us call it product 1) with attractiveness a and N products with attractiveness $a - \epsilon_t$. Where t is the threshold value and $\epsilon_t > 0$ serves the purpose that in presence of the first product, all the rest of the N products are simply ignored by any upcoming customer. The Market Share of the Show-all strategy is:

$$MS(X) = \frac{a}{a + a_0}$$

Let S_N be the set of the N products having lower attractiveness. If we show just the S_N products instead, we achieve a market share of:

$$MS(S_N) = \frac{N(a - \epsilon_t)}{N(a - \epsilon_t) + a_0}$$

So, if we calculate the ratio between $MS(X)$ and $MS(S_N)$ and let N tend to infinity we have:

[4] Since the Markov chain model proposed by [Blanchet et al., 2016] belongs to the RUM class [Berbeglia, 2016], neither the Threshold Luce or the two-stage Luce models are not contained on it.

$$
\begin{aligned}
\lim_{N\to\infty} \frac{MS(X)}{MS(S_N)} &= \lim_{N\to\infty} \frac{\frac{a}{a+a_0}}{\frac{N(a-\epsilon_t)}{N(a-\epsilon_t)+a_0}} \\
\lim_{N\to\infty} \frac{MS(X)}{MS(S_N)} &= \lim_{N\to\infty} \frac{a}{a+a_0} \cdot \frac{N(a-\epsilon_t)+a_0}{N(a-\epsilon_t)} \\
\lim_{N\to\infty} \frac{MS(X)}{MS(S_N)} &= \lim_{N\to\infty} \frac{a\,[N(a-\epsilon_t)+a_0]}{(a+a_0)\cdot[N(a-\epsilon_t)]} \\
\lim_{N\to\infty} \frac{MS(X)}{MS(S_N)} &= \frac{a}{a+a_0} \qquad (3.12)
\end{aligned}
$$

We can make a as small as desired, and the ratio would be arbitrarily bad.

The intuition that justifies *more is not always better*, relies on the fact that when high attractiveness products are shown, then the consideration set might not include many low-attractiveness products, reducing the overall market share. Thus, it might be convenient for the firm to not offer high-attractiveness products that can dominate many low attractiveness alternatives. Motivated by this intuition, we characterize the optimal solution for the Market Share Optimsation problem under the Threshold Luce model. We first need to present the following definition:

Definition 9 (Attractive Windowed Assortments)**.** Subsets $S = \{k_1, \ldots, k_2\}$ are called **attractive windowed assortments**, if $k_1 \leq k_2$ and $\forall k$ such as $k_1 \leq k \leq k_2$, we have $k \in S$ as well.

Theorem 8 (Optimality of attractiveness windowed assortments)**.** *For the market share maximisation problem in equation* (3.9)*, any optimal solution is of the form* $S = [k_1, \ldots, k_2]$ *(i.e. considering all the elements indexed between* k_1 *and* k_2*) with* $1 \leq k_1 \leq k_2 \leq n$*. This result also holds for the market share maximisation problem with heterogeneous customers.*

Proof. Let S be an optimal solution to the market share optimisation problem, and suppose that there exist a product k such that $k \notin S$ and there exist $k_1, k_2 \in S$ such that $k_1 < k < k_2$. We now show that the addition of product k is always beneficial for the market share. For each customer class, the addition of product k does not induce any dominance, because k_1 is in the assortment and has a higher attractiveness. Therefore, for each customer class, we are just simply increasing the total attractiveness, or in the worst case, keeping it the same. Thus, the total market share also increases, or at least remains the same. To show this, let $x \geq 0$ represents a value for the total attractiveness, if we consider the function:

$$m(x) = \frac{x}{x+a_0} = 1 - \frac{a_0}{x+a_0} \qquad (3.13)$$

is clearly strictly increasing in x, so the proof follows.

□

This result was also independently shown in Wang [2019]. The author also showed that the assortment optimisation problem with heterogeneous customers is NP-hard, using the partition problem [Garey and Johnson, 1979], a known NP-complete problem, and provide a fully polynomial time approximation scheme (FPTAS).

3.5.2 Assortment Optimisation with Position Bias

Another well studied extension of the Assortment Optimisation Problem, is to consider position bias. In this extension, product placement of the products influence how consumers choose among them. Several models had been proposed in the literature to capture this effect [Kempe and Mahdian, 2008a; Hummel and McAfee, 2014; Abeliuk et al., 2016; Davis et al., 2013] among others. We use the same way of describing this effect as in Abeliuk et al. [2016], using a weight associated with each position that can increase or decrease the probability of selecting a product, depending on how attractive the position is.

Let $\theta_j > 0$, with $0 \leq j \leq n$ be the weight associated with position j. Without loss of generality, we can assume $\theta_1 \geq \theta_2 \geq \ldots \geq \theta_n > 0$. A position assignment is an injective function $\sigma : S \rightarrow [n]$ that maps each product in S to a position. If product i is shown in position j, means that the position bias produces a shift on the intrinsic utility of product i of a factor $\ln \theta_j$. So if we let a_i the attractiveness of product i, the new attractiveness of this product when is shown in position j is $\theta_j a_i$. We need to slightly change the concept of dominance under the Threshold Luce model to reflect this change in relative attractiveness caused by positions. We do that in the following way:

Definition 10. Given a threshold t, a set S and an assignment σ, we say that $i \succ_\sigma j$ with $i, j \in S$ if:

$$\frac{\theta_{\sigma_i} a_i}{\theta_{\sigma_j} a_j} > (1+t),$$

thus, the consideration set is defined as:

$$c(S,\sigma) = \{j \in S \mid \nexists i \in S : i \succ_\sigma j\}, \tag{3.14}$$

And consequently, the revenue given an assortment S and assignment σ and considering product revenues $r \in \mathbb{R}^n_+$ is:

$$R(S,p) = \frac{\sum_{i \in c(S,\sigma)} \theta_{\sigma_i} a_i r_i}{\sum_{i \in c(S,\sigma)} \theta_{\sigma_i} a_i + a_0} \tag{3.15}$$

We are now interested in solving the assortment optimisation problem: finding the optimal assortment and position for each product in order to maximise revenue (Equation (3.15)).

Abeliuk et al. [2016] proposed a solution to the assortment problem under the usual Multinomial Logit Model with position bias, which is based on the rearrangement inequality and it can be performed in polynomial time. Let us call this algorithm the *Multinomial Logit Model Assortment Problem with Position Bias Algorithm* and call it as $\texttt{MLMAPPBA}(X, r, a, a_0, \theta)$, which returns (S^*, σ^*, R^*), the optimal subset to offer, the optimal position assignment, and the optimal expected revenue respectively. We now propose a polynomial time algorithm to solve the assortment optimisation problem with position bias under the Threshold Luce model, based on using MLMAPPBA a polynomial number of times.

Theorem 9. *The Assortment optimisation Problem under the Threshold Luce model with Position Bias can be solved in polynomial time.*

Proof. Let $(X, r, a, a_0, \theta, t)$ be an instance of the Threshold Luce model with Position Bias. Let A be the matrix generated using a and θ, where $a_{ij} = a_i\theta_j$ and products and are indexed in decreasing order of attractiveness. Suppose that we include product i_1 in position j_1. This inclusion means that any other product i_2, allocated in position j_2 gets dominated by product i_1 in position j_1 if:

$$\frac{a_{i_1 j_1}}{a_{i_2 j_2}} > (1+t), \tag{3.16}$$

Consider for each pair (i_1, j_1), the corresponding pair $(\hat{i_1}, \hat{j_1})$ satisfying the following conditions:

- $\frac{a_{i_1 j_1}}{a_{\hat{i_1}\hat{j_1}}} \leq (1+t)$, meaning that product $\hat{i_1}$ in position $\hat{j_1}$ is not dominated by product i_1 in position j_1.
- $\frac{a_{i_1 j_1}}{a_{\hat{i_1}+1\hat{j_1}}} > (1+t)$, which means that if we allocate product $\hat{i_1}+1$ instead in position $\hat{j_1}$ it gets dominated by product i_1 in position j_1.
- $\frac{a_{i_1 j_1}}{a_{\hat{i_1}\hat{j_1}+1}} > (1+t)$, which means that if we allocate product $\hat{i_1}$ in position $\hat{j_1}+1$ instead, it gets dominated by product i_1 in position j_1.

Now, restricted to products $X_{i_1 j_1} = \{i_1, \ldots, \hat{i_1}\}$ and positions $\mathcal{P}_{i_1 j_1} = \{j_1, \ldots, \hat{j_1}\}$, no dominance relation appears for any assignment of products to positions. So restricted to these products and positions, we can solve the Assortment Problem to optimality by using MLMAPPBA. Noting that we need to perform this process at most a quadratic number of times (one for each element in matrix A) and simply keep track of the maximum revenue, the proof follows. □

Chapter 4

Joint Assortment and Pricing under the Threshold Luce model

This chapter is reproduced with minor changes from:

1. **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment and Pricing optimisation Under the Two-Stage Luce model. *Submitted to Operations Research (13th of April of 2019).* Under Review.

2. **Flores, A.**; Berbeglia, G; Van Hentenryck, P. 2019. Assortment and Pricing optimisation Under the Two-Stage Luce model. *Presented at the Informs Revenue Management & Pricing Conference, Stanford (Informs RM&P), 7th of June 2019.*

Chapter 3 provides solutions to the Assortment optimisation problem under the Two-Stage Luce model. But, what about the pricing optimisation problem? how can we attach prices to products in order to maximise the expected revenue? Recall Example 9 in the previous chapter, where each product has different price levels that can be attached to it (with their corresponding attractiveness, reflecting the price point). This setting appears naturally for several reasons: rounded-to-dollar prices, market regulations, or discrete demand information that does not allow to recover a clear relation between price and attractiveness. In this case there is no explicit known functional relationship between attractiveness and prices, and therefore it allows any price-attractiveness combination. Theorem 4 shows that this problem is actually polynomial time solvable, because essentially amounts to solve an assortment problem with extra dominances within each product and their corresponding price levels, preserving the 2SLM structure. However, this only solves a discrete version of the pricing problem for the 2SLM.

If we want to solve the pricing problem for the 2SLM having the freedom to choose any price for any product, we need a way of *update* the dominance relation if a price assignment changes the perceived attractiveness of the products too much. To overcome this issue, in this chapter, we study the Joint Assortment and Pricing Problem under the Threshold Luce model making the attractiveness of each product dependent upon the price attached to it.

4.1 Threshold Luce Model: Joint Assortment and Pricing Problem

In this section, we study the Joint Assortment and Pricing Problem under the Threshold Luce model, by making the attractiveness of each product dependent upon its price. Let $p = (p_1, \dots, p_n)$ be the price vector, where such that $p_i \in \mathbb{R}_+ \cup \{\infty\}$ represents the price of product i. Since the price will affect the attractiveness a_i of product i, the presentation makes this dependency explicit by writing $a_i(p_i)$ whose form in this chapter is specified by

$$a_i(p_i) = \exp(u_i - p_i) \tag{4.1}$$

where u_i is the *intrinsic utility* of product i and the value $v_i = u_i - p_i$ is called the *net utility* of product i. Assigning an infinite price to a product is equivalent to not offering the product, as the attractiveness, and therefore the probability of selecting the product, becomes 0. Without loss of generality, products are indexed in a decreasing order by intrinsic utility.

The following definition is an extension of the definition of a consideration set given an assortment S when each product i has a price p_i.

Definition 11. Given an assortment S, a price vector $p = (p_1, p_2, \dots, p_n)$ and a threshold t, the consideration set $c(S, p)$ for the Threshold Luce model is defined as:

$$c(S, p) = \{j \in S \mid \nexists i \in S : a_i(p_i) > (1 + t)a_j(p_j)\}. \tag{4.2}$$

The influence of the price vector over the dominance relations is given by the following example:

Example 17 (Price effect on the dominance relation)**.** Consider the *Threshold Luce model* defined over $X = \{1, 2, 3, 4\}$ with utilities $u_1 = \ln(10), u_2 = \ln(8), u_3 = \ln(6)$ and $u_4 = \ln(3)$, and consider first a scenario where all products have the same price $p_i = \ln(3) \quad \forall i = 1, \dots, 4$. Consider also a second scenario with prices equal to $p'_1 = \ln(4), p'_2 = \ln(4), p'_3 = \ln(3)$ and $p'_4 = \ln(2)$. For a threshold $t = 0.5$, we have that $i \succ j$ iff $a_i(p_i) > 1.5\ a_j(p_j)$. A table summarizing the utilities, prices, and attractiveness for both scenarios is given in Table 4.1 and the DAGs depicting the dominance relations for the two scenarios are given in Figures 4.1 and 4.2.

i	u_i	p_i	$a_i(p_i)$	p'_i	$a_i(p'_i)$
1	ln(10)	ln(3)	$3.\overline{3}$	ln(4)	2.5
2	ln(8)	ln(3)	$2.\overline{6}$	ln(4)	2
3	ln(6)	ln(3)	2	ln(3)	2
4	ln(3)	ln(3)	1	ln(2)	1.5

Table 4.1: Summary of utilities, prices and attractiveness for the two proposed scenarios.

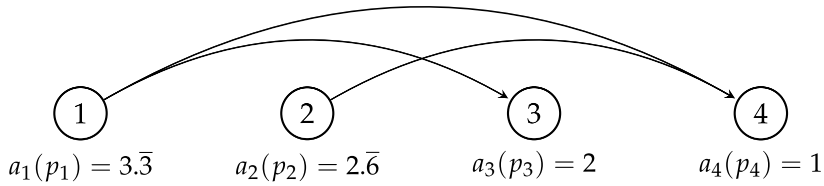


Figure 4.1: The DAG for the first scenario where all prices are fixed to $\ln(3)$ and the threshold is $t = 0.5$. Product 1 dominates products 3 and 4, and product 2 dominates product 4.

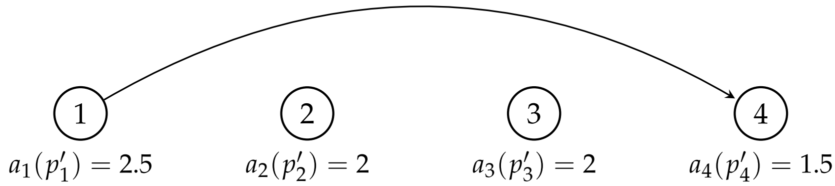


Figure 4.2: The DAG for the second scenario where all prices are fixed to $(\ln(4), \ln(4), \ln(3), \ln(2))$ and the threshold is $t = 0.5$. Only product 1 dominates product 4.

It is also necessary to update the definition of ρ in Definition 4, since it now depends on the price of all products in the assortment. The definition of $\rho : X \cup \{0\} \times 2^X \times (\mathbb{R}_+ \cup \infty)^n \to [0, 1]$ becomes:

$$\rho(i, S, p) = \begin{cases} \frac{a_i(p_i)}{\sum_{j \in c(S,p)} a_j(p_j) + a_0}, & \text{if } i \in c(S, p), \\ 0 & \text{if } i \notin c(S, p). \end{cases} \tag{4.3}$$

where a_0 is the attractiveness of the outside option.
The expected revenue (ER) of an assortment $S \subseteq X$ and a price vector $p \in \mathbb{R}^n_+$ is given by

$$R(S, p) = \sum_{i \in c(S,p)} \rho(i, S, p) p_i. \tag{ER}$$

A pair (S, p) with $S \subseteq X$ and $p \in (\mathbb{R}_+ \cup \infty)^n$ is *valid* if $S = \{i : p_i < \infty\}$ and $c(S, p) = S$. Let $\mathcal{V}$ be the set of all valid pairs (S, p). Observe that one can always restrict the search for optimal solutions to $\mathcal{V}$. Indeed, all dominated products can be given an infinite price and removing them from the original assortment yields the exact same revenue.

The Joint Assortment and Pricing problem aims at finding a set S^* and a price vector p^* satisfying

$$(S^*, p^*) \in \operatorname*{argmax}_{(S,p) \in \mathcal{V}} R(S, p)$$

and yielding an optimal revenue of

$$R^* = R(S^*, p^*).$$

First observe that the strategy used to solve this problem under the multinomial logit does not carry over to the Threshold Luce Model. Under the multinomial logit, the optimal solution for the joint assortment and pricing problem is a fixed adjusted margin policy [Wang, 2012] which, for equal price sensitivities and normalised costs, translates to a fixed price policy. As shown in Li and Huh [2011], the optimal solution for the pricing problem under the multinomial logit can be expressed in closed form using the Lambert function $W(x) : [0, \infty) \to [0, \infty)$ which is defined as the unique function satisfying:

$$x = W(x)e^{W(x)} \quad \forall x \in [0, \infty). \tag{4.4}$$

Using this function, the optimal revenue can be expressed as:

$$R^* = W\left(\frac{\sum_{i \in X} \exp(u_i - 1)}{a_0}\right) \tag{4.5}$$

The prices are all equal and satisfy: $p_i = 1 + R^* \quad \forall i \in X$. The following example shows that fixed-price policy is not optimal under the Threshold Luce Model.

Example 18 (Fixed-Price policy is not optimal)**.** Consider 11 products with product 1 having utility $u = 2$ and all remaining 10 products having utility $u' = 1$. Consider $a_0 = 1$ and $t = 1$. Observe that, for any fixed price, product 1 always dominates the other 10 products having lower utility, as $\exp(u - u') = \exp(1) = e > (1 + t) = 2$. Therefore, the optimal revenue for a a fixed price strategy is:

$$R_{fixed} = W\left(\frac{\exp(u - 1)}{a_0}\right) = W(e) = 1.$$

As a result, the 10 lower utility products are completely ignored and only product 1 contributes to the revenue.

Consider the following price scheme now: let the price for product 1 be $p = 1.8$ and let the price be $p' = 1.4$ for the remaining products. Product 1 does not dominate any other product now. Indeed, for any $1 < k \leq 11$,

$$\frac{a_1}{a_k} = \exp((u - p) - (u' - p')) = \exp((2 - 1.8) - (1 - 1.4)) \approx 1.822 < 1 + t = 2,$$

which yields a revenue of:

$$R' = \frac{p \cdot \exp(u - p) + 10 \cdot p' \exp(u' - p')}{\exp(u - p) + 10 \cdot \exp(u' - p') + a_0} = \frac{1.8 \cdot \exp(2 - 1.8) + 10 \cdot 1.4 \exp(1 - 1.4)}{\exp(2 - 1.8) + 10 \cdot \exp(1 - 1.4) + 1} \approx 1.298,$$

This pricing scheme improves upon the fixed-price policy, yielding a revenue almost %30 higher.

The intuition behind this example is as follows: For a fixed price strategy, the only factor affecting dominance is the intrinsic utilities because the prices vanish when calculating the ratio between two attractiveness. This means that the solution can potentially miss the benefits of low attractiveness products which are dominated by the most attractive product.

It is thus important to understand the structure of an optimal solution for the Joint Assortment and Pricing problem under the Threshold Luce model. The first result states that, for any optimal solution (S^*, p^*), all product prices are greater or equal than R^*, where R^* denotes the revenue achieved at optimality.

Proposition 7. *In any optimal solution (S^*, p^*), for all $i \in S^*$, $p_i^* \geq R^*$.*

Proof. The proof is by contradiction: Removing products with a price lower than R^* yields a greater revenue. Indeed, suppose $p_i^* < R^*$ for some $i \in S$, then $\hat{S} = S^* \setminus \{i\}$ has better revenue than the optimal solution if we keep the same prices and $p_i^* < R^*$. Indeed, let us calculate $R(\hat{S})$:

$$R(\hat{S}) = \frac{\sum_{j\in\hat{S}} e^{u_j - p_j^*} \cdot p_j^*}{\sum_{j\in\hat{S}} e^{u_j-p_j^*} + a_0}$$

$$R(\hat{S}) = \frac{\sum_{j\in S^*} e^{u_j - p_j^*} \cdot p_j^* - e^{u_i-p_i^*} \cdot p_i^*}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0}$$

$$R(\hat{S}) = \frac{\sum_{j\in S^*} e^{u_j - p_j^*} \cdot p_j^*}{\sum_{j\in S^*} e^{u_j-p_j^*} + a_0} \cdot \frac{\sum_{j\in S^*} e^{u_j-p_j^*} + a_0}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0} - \frac{e^{u_i-p_i^*} \cdot p_i^*}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0}$$

$$R(\hat{S}) = \frac{\sum_{j\in S^*} e^{u_j - p_j^*} \cdot p_j^*}{\sum_{j\in S^*} e^{u_j-p_j^*} + a_0} \cdot \left[1 + \frac{e^{u_i-p_i^*}}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0}\right] - \frac{e^{u_i-p_i^*} \cdot p_i^*}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0}$$

$$R(\hat{S}) = R^* \cdot \left[1 + \frac{e^{u_i-p_i^*}}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0}\right] - \frac{e^{u_i-p_i^*} \cdot p_i^*}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0}$$

$$R(\hat{S}) = R^* + \underbrace{\frac{e^{u_i-p_i^*}}{\sum_{j\in S^*} e^{u_j-p_j^*} - e^{u_i-p_i^*} + a_0} \cdot [R^* - p_i^*]}_{\Gamma}$$

Now Γ is positive because $p_i^* < R^*$, but this implies $R(\hat{S}) > R^*$, contradicting the optimality of R^*. □

The next proposition characterises the optimal assortment of products of any optimal solution to the Joint Assortment and Pricing problem. Recall that the products are indexed by decreasing utility u_i. Thus, the set of products $[k] := \{1, \ldots, k\}$, (with $0 < k \leq n$) is said to be an *intrinsic utility ordered set*. The following proposition holds:

Proposition 8. *Let (S^*, p^*) denote an optimal solution. Then $S^* = [k]$ for some $k \leq n$.*

Proof. Let (S^*, p^*) be an optimal solution. We can assume that $(S^*, p^*) \in \mathcal{V}$. We proceed by contradiction. Suppose that there is a product i not included in the optimal solution and another product j with **smaller intrinsic utility** included in S^*. We show that we can include product i, and remove j and get a greater revenue. Let $\hat{S} = (S^* \setminus \{j\}) \cup \{i\}$, be the set where we removed product j, and included product i. Let $\hat{p}_i = u_i - u_j + p_j^*$, this means that the total attractiveness remains unchanged, and no new domination relations appear, given that product j already had the same level attractiveness that product i now has. Observe that given that $u_i \geq u_j$, we have that $\hat{p}_i \geq p_j^*$. Let us calculate $R(\hat{S}, \hat{p})$, where $\hat{p}$ is the same as p^*, but with the proposed changes in price:

$$
\begin{aligned}
R(\hat{S}, \hat{p}) &= \frac{\sum_{k \in \hat{S}} e^{u_k - \hat{p}_k} \cdot \hat{p}_k}{\sum_{k \in \hat{S}} e^{u_k - \hat{p}_k} + a_0} \\
R(\hat{S}, \hat{p}) &= \frac{\sum_{k \in S^*} e^{u_k - p_k^*} \cdot p_k^* - e^{u_j - p_j^*} \cdot p_j^* + e^{u_i - \hat{p}_i} \cdot \hat{p}_i}{\sum_{k \in \hat{S}} e^{u_k - \hat{p}_k} + a_0} \\
R(\hat{S}, \hat{p}) &= \underbrace{\frac{\sum_{k \in S^*} e^{u_k - p_k^*} \cdot p_k^*}{\sum_{k \in S^*} e^{u_k - \hat{p}_k} + a_0}}_{R^*} + \frac{e^{u_i - \hat{p}_i} \cdot \hat{p}_i - e^{u_j - p_j^*} \cdot p_j^*}{\sum_{k \in \hat{S}} e^{u_k - \hat{p}_k} + a_0} \\
R(\hat{S}, \hat{p}) &= R^* + \underbrace{\frac{e^{u_j - p_j^*}}{\sum_{k \in \hat{S}} e^{u_k - \hat{p}_k} + a_0}}_{\geq 0} \cdot \underbrace{\left[\hat{p}_i - p_j^*\right]}_{>0} \\
R(\hat{S}, \hat{p}) &> R^*
\end{aligned}
$$

Where we first rewrite $R(\hat{S}, \hat{p})$ using (S^*, p^*) because we just swapped product i for product j, and the total attractiveness remain the same, so the denominator does not change. Then we identify $R(S, p)$, and we use $u_i - \hat{p}_i = u_j - p_j$ to being able to factorize the remaining terms. So we found a pair $(\hat{S}, \hat{p})$, yielding strictly more revenue than (S, p), but adding product i, which contradicts the optimality of (S^*, p^*). □

The following Lemma due to Wang and Sahin [2018] is useful to prove some of the upcoming propositions.

Lemma 3 (Lemma 1, Wang and Sahin [2018])**.** *Let $H(p_i, p_j) := p_i \cdot \exp(u_i - p_i) + p_j \cdot \exp(u_j - p_j)$, where $\exp(u_i - p_i) + \exp(u_j - p_j) = T$. Then, $H(p_i, p_j)$ is strictly unimodal with respect to p_i or p_j, and it achieves the maximum at the following point:*

$$p_i^* = p_j^* = \ln\left((\exp(u_i) + \exp(u_j))/T\right) \tag{4.6}$$

Proof. The proof (due to Wang and Sahin [2018]) is useful because it provides intuition on how the optimal price variates when constrained to a fixed additive market share among any two products. By the equality constraint, we have $p_j = u_j - \ln(T - \exp(u_i - p_i))$, so $H(p_i, p_j)$ can be rewritten purely as a function of p_i as:

$$H(p_i) = p_i \cdot \exp(u_i - p_i) + (u_j - \ln(T - \exp(u_i - p_i))) \cdot (T - \exp(u_i - p_i)). \tag{4.7}$$

Now, let us calculate the first derivative of $H(p_i)$ w.r.t. p_i:

$$\frac{\partial H(p_i)}{\partial p_i} = \left(-p_i + (u_j - \ln(T - \exp(u_i - p_i)))\right) \cdot \exp(u_i - p_i) \tag{4.8}$$

Clearly the left-hand side term on the multiplication is monotonically decreasing from positive to negative values as p_i increases from 0 to ∞. Therefore $H(p_i)$ is strictly unimodal and reaches its maximum value at:

$$p_i^* = p_j^* = \ln\left((\exp(u_i) + \exp(u_j))/T\right).$$

□

Observe that setting the price of a product to ∞ is equivalent to not showing it to consumers. By Proposition 8, one can always find an optimal solution that is *intrinsic utility ordered*. Given a price vector $p \in \mathbb{R}^n$, let $\gamma(p) : \mathbb{R}^n \to [n]$ be defined as $\gamma(p) \doteq \left\{\max_{i \in [n]} i \text{ s.t } p_i < \infty\right\}$. Intuitively, this is the last non-infinite price. Proposition 9 shows that, at optimality, the finite prices are non-increasing in i, meaning that lower prices are assigned to lower utility products.

Proposition 9. *The prices at an optimal solution (S^*, p^*) satisfy $p_i^* \geq p_{i+1}^* \quad \forall i \in [\gamma(p) - 1]$. Moreover, if $i, j \in S^*$ satisfy $u_i = u_j$, then $p_i^* = p_j^*$.*

Proof. We prove this result by contradiction. Let i be the first index where this condition does not hold, this means that $p_i^* < p_{i+1}^*$. Using Lemma 3, we found $\hat{p}$ satisfying $p_i^* < \hat{p} < p_{i+1}^*$. Does this new price alter the consideration set? We show that this is not the case. Indeed, the effect is two-fold: the price for product i increases, and the price for product $i+1$ decreases. We analyse the effect of these two consequences:

- Increase on price for product i: This means $a(i, p)$ decreases. Note that $u_i - \hat{p} \geq u_{i+1} - p_{i+1}^*$, so neither $i \succ i+1$ or $i+1 \succ i$, because their attractiveness are now even closer than before. Can i be dominated now by another product? No, because given that $u_i \geq u_{i+1}$ we have $u_i - \hat{p} \geq u_{i+1} - \hat{p} \geq u_{i+1} - p_{i+1}^*$. Therefore the new attractiveness of i is still larger than the new attractiveness of $i+1$, and the last inequality implies that the new attractiveness of i is larger than the old attractiveness of $i+1$, and $i+1$ was not previously dominated either by any other product.

- Decrease on price for product $i+1$: Previously $i+1$ was not dominated by any product. Can $i+1$ be dominated now? No, because if $i+1$ was not dominated before, now with a smaller price $\hat{p}$ its attractiveness is larger and therefore can't be dominated now either (the only other product that changed attractiveness was i, and it now has smaller attractiveness). Can $i+1$ dominate another product now with its new higher attractiveness? No, because given that $u_i \geq u_{i+1}$ we have $u_i - p_i^* \geq u_{i+1} - p_i^* \geq u_{i+1} - \hat{p}$, so the old attractiveness of product i is larger than the new attractiveness of product $i+1$, and given that i did not dominate another product before, the new price does not make $i+1$ dominate another product either.

So, letting p^{fix} exactly the same as p^*, but replacing both p_i^* and p_{i+1}^* with $\hat{p}$, means that the pair (S^*, p^{fix}) yields strictly more revenue than (S^*, p^*) (by Lemma 3), contradicting the optimality assumption. The fact that equal intrinsic utility implies equal price at optimality, can be easily demonstrated by the following: if two equal intrinsic utility products have different prices, then using Lemma 3 we obtain strictly better revenue by assigning them the same price, and no new domination occurs, because the new price is confined between the previous prices. □

Recall that the *net utility* of product i was defined as: $v_i = u_i - p_i$. The following proposition shows that at optimality, net utility follows the same order as intrinsic utility.

Proposition 10. *Let p^* be the price of an optimal solution of the Joint Assortment and Pricing Problem. The following condition holds: $u_i - p_i^* \geq u_{i+1} - p_{i+1}^* \quad \forall i \in [\gamma(p) - 1]$.*

Proof. We prove this by contradiction. Let p^* be the optimal solution and i be the first index where this condition does not hold. This means that $u_i - p_i^* < u_{i+1} - p_{i+1}^*$. We can extrapolate this inequality further and say:

$$u_{i+1} - p_i^* < u_i - p_i^* < u_{i+1} - p_{i+1}^* < u_i - p_{i+1}^*, \tag{4.9}$$

because $u_i \geq u_{i+1}$ and $p_i \geq p_{i+1}$ by Propositions 8 and 9 respectively. We now do the following: Define p_i' and p_{i+1}' such as $\exp(u_i - p_i') + \exp(u_{i+1} - p_{i+1}') = \exp(u_i - p_i^*) + \exp(u_{i+1} - p_{i+1}^*)$ and $\exp(u_i - p_i') = \exp(u_{i+1} - p_{i+1}')$. This means that:

$$p_i' = u_i - \ln\left(\frac{\exp(u_i - p_i^*) + \exp(u_{i+1} - p_{i+1}^*)}{2}\right)$$
$$p_{i+1}' = u_{i+1} - \ln\left(\frac{\exp(u_i - p_i^*) + \exp(u_{i+1} - p_{i+1}^*)}{2}\right)$$

Consider $H(p_i, p_{i+1}) = p_i \cdot \exp(u_i - p_i) + p_j \cdot \exp(u_{i+1} - p_{i+1})$, where $\exp(u_i - p_i) + \exp(u_i - p_i) = \exp(u_i - p_i^*) + \exp(u_{i+1} - p_{i+1}^*)$. By Lemma 3, $H(p_i, p_{i+1})$ is strictly increasing in p_i for $p_i \leq \hat{p}$ and strictly decreasing for $p_i \geq \hat{p}$, with $\hat{p} = \ln\left(\frac{\exp(u_i)+\exp(u_{i+1})}{\exp(u_i - p_i^*)+\exp(u_{i+1} - p_{i+1}^*)}\right)$ the solution of the corresponding maximization problem of Lemma 3. We can verify that $\hat{p} < p_i' < p_i^*$. The first inequality is straightforward. Indeed:

$$
\begin{aligned}
p'_i =& u_i - \ln\left(\frac{\exp(u_i - p_i) + \exp(u_{i+1} - p_{i+1})}{2}\right)\\
p'_i =& \ln\left[\frac{2\exp(u_i)}{\exp(u_i - p^*_i) + \exp(u_{i+1} - p^*_{i+1})}\right]\\
p'_i >& \underbrace{\ln\left[\frac{\exp(u_i) + \exp(u_{i+1})}{\exp(u_i - p^*_i) + \exp(u_{i+1} - p^*_{i+1})}\right]}_{\hat{p}}\\
p'_i >& \hat{p}
\end{aligned}
$$

proving the desired inequality. Now, for the second one:

$$
\begin{aligned}
p'_i &= u_i - \ln\left(\frac{\exp(u_i - p_i) + \exp(u_{i+1} - p_{i+1})}{2}\right)\\
p'_i &= \ln\left[\frac{2\exp(u_i)}{\exp(u_i - p^*_i) + \exp(u_{i+1} - p^*_{i+1})}\right]\\
p'_i &\leq \ln\left[\frac{2\exp(u_i)}{\exp(u_i - p^*_i) + \exp(u_i - p^*_{i+1})}\right]\\
p'_i &= \ln\left[\frac{2\exp(u_i)}{\exp(u_i)(\exp(-p^*_i) + \exp(-p^*_{i+1}))}\right]\\
p'_i &< \ln\left[\frac{2}{2\exp(-p^*_i)}\right]\\
p'_i &< p^*_i,
\end{aligned}
$$

thus we have:

$$
p'_i \cdot \exp(u_i - p'_i) + p'_{i+1} \cdot \exp(u_{i+1} - p'_{i+1}) > p^*_i \cdot \exp(u_i - p^*_i) + p^*_{i+1} \cdot \exp(u_{i+1} - p^*_{i+1}).
$$

Meaning that we have the same assortment, but with prices p'_i and p'_{i+1} generating strictly more revenue than the optimal prices, which is a contradiction. The only thing that we have left to show that with these new prices we are still on the same consideration set. It would be enough to show that the new net utilities are bounded by previous values of net utilities. Indeed, we can verify that $p^*_{i+1} \leq p'_{i+1} \leq p'_i \leq p^*_i$, by simply using the definitions. We also know, by hypothesis that $u_i - p'_i = u_{i+1} - p'_{i+1}$, then $u_i - p'_i = u_{i+1} - p'_{i+1} \leq u_{i+1} - p^*_{i+1}$. So even when the price of product i decreased, the new attractiveness is bounded above by a previously existing attractiveness, thus not changing the consideration set. By the same reasoning, $u_{i+1} - p'_{i+1} = u_i - p'_i \geq u_i - p^*_i$, meaning that the new attractiveness is bounded below by a pre-existing one, so $i+1$ is not dominated with this new prices either. So the consideration set stays the same, concluding the proof. □

The above propositions make it possible to filter out non-efficient assortments and prices by restricting the search space to intrinsic utility ordered assortments and providing insights on how the optimal solution behaves regarding prices and their relation with utilities. Based on these propositions, the joint assortment and pricing optimisation problem for the TLM can be written in a more succinct way. From Proposition 8, the solution is an intrinsic utility ordered set $S_k = [k]$ for some $k \leq n$. Suppose there exists an optimal solution in the form (S_k, p) for a fixed value k. In that case, recall that it is sufficient to restrict to valid pairs (S_k, p), meaning that $c(S_k, p) = S_k$. Consider a fixed $k \leq n$. By Proposition 10, at optimality, $u_i - p_i \geq u_j - p_j \quad \forall 1 \leq i < j \leq k$. Therefore, the condition that $c(S_k, p) = S_k$ can be written as

$$g_{ij}(p) := \exp(u_i - p_i) - (1+t) \cdot \exp(u_j - p_j) \leq 0, \quad \forall 1 \leq i < j \leq k \tag{4.10}$$

As a result, the joint k-assortment and pricing optimisation problem for the TLM (`JAPTLM-k`), which aims at finding an optimal assortment S_k of size k with $k \leq n$, can be written as:

$$\begin{aligned} \underset{p}{\text{maximise}} \quad & R^{(k)}(p) := \frac{\sum_{i \in S_k} p_i \cdot \exp(u_i - p_i)}{\sum_{i \in S_k} \exp(u_i - p_i) + a_0} \\ \text{subject to} \quad & g_{ij}(p) \leq 0, \quad \forall 1 \leq i < j \leq k \end{aligned} \tag{JAPTLM-k}$$

Note that, if $\exp(u_1 - u_k) \leq (1+t)$, then the solution is the same as the unconstrained case, because any fixed price can be assigned without creating dominances. Hence, the optimal revenue $\boldsymbol{R}^{(k)}$ can be calculated using equation (4.5), and all prices are equal to $1 + \boldsymbol{R}^{(k)}$. On the other hand, if $\exp(u_1 - u_k) > 1 + t$, as in Example 18, the prices need to be adjusted in order to avoid dominances.

The next theorem is the main result of this section.

Theorem 10. *Problem JAPTLM-k can be solved in polynomial time.*

Proof. The intuition behind the proof is based on Proposition 10 and the study of the Lagrangean relaxation of problem (JAPTLM-k). Observe that, since $u_i - p_i \geq u_j - p_j \quad (i \leq j)$ at optimality, then the largest ratio between attractiveness is obtained for products 1 and k. This ratio can also occur for more products but only if they have the same net utility as products 1 or k. Thus, it must be the case that there are non-negative integers k_1 and k_2 with $k_1 + k_2 \leq k$, such that letting $I_1 = [k_1]$ and $I_2 = \{k - k_2 + 1, k - k_2 + 2, \ldots, k\}$, the set of constraints $C(k_1, k_2) = \{g_{ij}(p) \mid i \in I_1, j \in I_2\}$ are satisfied at equality for the optimal solution. Since it is only necessary to study a polynomial number of combinations of constraints satisfied at equality and, for each one of those combinations a closed form solution is provided, the result follows.

Indeed, we first write problem (JAPTLM-k) in minimization form to directly apply the Karush-Khun-Tucker conditions (KKT)[Karush, 1939].

$$\begin{aligned} \underset{p}{\text{minimize}} \quad & -R^{(k)}(p) \\ \text{subject to} \quad & g_{ij}(p) \leq 0, \quad \forall 1 \leq i < j \leq k \end{aligned} \tag{4.11}$$

The associated Lagrangean function is:

$$\mathcal{L}_k(p, \mu) = -R^{(k)}(p) + \sum_{1 \leq i < j \leq k} \mu_{ij} \cdot g_{ij}(p), \tag{4.12}$$

where $\mu_{ij} \geq 0$ are the associated Lagrange multipliers. Recall that if $\exp(u_1 - u_k) \leq (1 + t)$, the optimal revenue $\boldsymbol{R^{(k)}}$ can be calculated using equation (4.5), and the solution corresponds to a fixed price policy as for the regular multinomial logit.

On the other hand, if $\exp(u_1 - u_k) > (1 + t)$, any fixed price causes product k to be dominated by product 1. Thus, to include product k in the assortment we need to adjust the prices. Let $p = (p_1, \ldots, p_k)$ be the optimal price vector for problem (4.11). Observe that it can't happen that $\frac{a_1(p_1)}{a_k(p_k)} < 1 + t$, since by Proposition 10, it will also means that $\frac{a_1(p_1)}{a_2(p_2)} < 1 + t$ and using Lemma 3 we can find $\hat{p}$ such that assigning $\hat{p}$ to products 1 and 2 yields a larger revenue (and no dominance relation appears, since the attractiveness of product 1 was reduced, and the attractiveness of product 2 increased, but is still less than the one of product 1), which contradicts optimality. Therefore, g_{1k} must be satisfied with equality, meaning $\frac{a_1(p_1)}{a_k(p_k)} = 1 + t$.

Furthermore, at optimality it holds $u_i - p_i \geq u_j - p_j \quad \forall i \leq j$ (by Proposition 10), and thus the biggest ratio between attractiveness is observed for products 1 and k, and is exactly equal to $1 + t$. This ratio can be replicated for other pairs of products, but only if they share the same net utility (and thus attractiveness) to the one of products 1 or k. Therefore, it must be the case that there are integers k_1 and k_2 with $k_1 + k_2 \leq k$, such that all products in $I_1 = [k_1]$ share the same attractiveness ($a_1(p_1)$) and all products in $I_2 = \{k - k_2 + 1, k - k_2 + 2, \ldots, k\}$ share the same attractiveness as well ($a_k(p_k)$). This means that the set of constraints $C(k_1, k_2) = \{g_{ij}(p) \mid i \in I_1, j \in I_2\}$ are all satisfied with equality at optimality.

We now study the derivative of equation (4.12) with respect to each price p_i to obtain the KKT conditions. We here assume that the first k_1 values share the same net utility value, meaning $u_s = u_1 - p_1 = u_i - p_i \quad \forall i \in I_1$, and for the last k_2 products, we also have the same value of net utility, that we call u_f, this is: $u_f = u_k - p_k = u_i - p_i \quad \forall i \in I_2$. Where these two quantities satisfy:

$$u_s - u_f = \ln(1 + t),$$

Let us write the derivatives of the Lagrangean depending on where the index i belongs. If $i \in I_1$, then:

$$\frac{dL_k}{dp_i} = \frac{\exp(u_i - p_i)}{\sum_{j \in S_k} \exp(u_j - p_j) + a_0} \cdot \left[p_i - 1 - R^{(k)}(p) \right] - \exp(u_i - p_i) \cdot \sum_{j \in I_2} \mu_{ij}, \tag{4.13}$$

if $i \in I_2$, we have:

$$\frac{dL_k}{dp_i} = \frac{\exp(u_i - p_i)}{\sum_{j \in S_k} \exp(u_j - p_j) + a_0} \cdot \left[p_i - 1 - R^{(k)}(p)\right] + (1+t)\exp(u_i - p_i) \cdot \sum_{j \in I_1} \mu_{ji}, \tag{4.14}$$

And finally, if $i \in \bar{I}_k = [k] \setminus (i_1 \cup I_2)$, the derivative takes the following form:

$$\frac{dL_k}{dp_i} = \frac{\exp(u_i - p_i)}{\sum_{j \in S_k} \exp(u_j - p_j) + a_0} \cdot \left[p_i - 1 - R^{(k)}(p)\right] \tag{4.15}$$

Observe that $\forall i \in \bar{I}_k$, $\frac{dL_k}{dp_i} = 0 \implies p_i = 1 + R^{(k)}(p)$, and the right hand side is not dependent on i, so all products in $\bar{I}_k$ share the same price, which we denote $\bar{p}$. We can rewrite all prices and the revenue depending on u_s and $\bar{p}$, using the following relations:

1. $\forall i \in I_1 \quad u_1 - p_1 = u_i - p_i \implies p_i = u_i - u_s$
2. $\forall i \in I_2 \quad u_1 - p_1 = u_i - p_i + \ln(1+t) \implies p_i = u_i - u_s + \ln(1+t)$

Note now that at optimality, for a fixed k, prices are determined by k_1 and k_2. Thus, the optimal revenue can be written explicitly depending on k, k_1 and k_2, taking the following form:

$$R^{(k)}(k_1,k_2) = \frac{\sum_{i \in I_1}(u_i - u_s)\exp(u_s) + \bar{p}\exp(-\bar{p})\sum_{i \in \bar{I}_k}\exp(u_i) + \sum_{i \in I_2}(u_i - u_s + \ln(1+t))\exp(u_s - \ln(1+t))}{\sum_{i \in I_1}\exp(u_s) + \exp(-\bar{p})\sum_{i \in \bar{I}_k}\exp(u_i) + \sum_{i \in I_2}\exp(u_s + \ln(1+t)) + a_0} \tag{4.16}$$

Note that $\bar{p} = 1 + R^{(k)}(k_1,k_2)$ (Equation (4.15)) and let $E(k_1,k_2) = \sum_{i \in \bar{I}_k}\exp(u_i)$. Using these two relations, we can rewrite the optimal revenue as:

$$R^{(k)}(k_1,k_2) = \frac{e^{u_s}\sum\limits_{i \in I_1}(u_i - u_s) + \frac{e^{u_s}}{1+t} \cdot \sum\limits_{i \in I_2}(u_i - u_s + \ln(1+t)) + E(k_1,k_2)(1 + R^{(k)}(k_1,k_2))e^{-(1+R^{(k)}(k_1,k_2))}}{e^{u_s}\left[k_1 + \frac{k_2}{1+t}\right] + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0} \tag{4.17}$$

Up to this point, we have an equation relating the optimal revenue $R^{(k)}(k_1,k_2)$ and u_s. From equation (4.13), after reordering terms we have:

$$\begin{aligned}
\frac{p_i - 1 - R^{(k)}(k_1,k_2)}{e^{u_s}(k_1 + k_2(1+t)) + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0} &= \sum_{j \in I_2}\mu_{ij}, \qquad \forall i \in I_1 \\
\frac{u_i - u_s - 1 - R^{(k)}(k_1,k_2)}{e^{u_s}(k_1 + k_2(1+t)) + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0} &= \sum_{j \in I_2}\mu_{ij}, \qquad \forall i \in I_1
\end{aligned} \tag{4.18}$$

Analogously, from equation (4.14), after reordering terms we have $\forall i \in I_2$:

$$\frac{p_i - 1 - R^{(k)}(k_1,k_2)}{e^{u_s}\left(k_1 + k_2(1+t)\right) + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0} = -(1+t)\sum_{j\in I_1}\mu_{ji}, \qquad \forall i \in I_2$$

$$\frac{1}{1+t}\cdot\frac{u_i - u_s + \ln(1+t) - 1 - R^{(k)}(k_1,k_2)}{e^{u_s}\left(k_1 + k_2(1+t)\right) + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0} = -\sum_{j\in I_1}\mu_{ji}, \qquad \forall i \in I_2 \tag{4.19}$$

Now, if we add equations (4.18) $\forall i \in I_1$ then take equations (4.19) and also add them $\forall i \in I_2$, and add those two results we can derive the value $R^{(k)}(k_1,k_2)$ as follows.

$$\underbrace{\sum_{\substack{i\in I_1\\ j\in I_2}}\mu_{ij} - \sum_{\substack{i\in I_1\\ j\in I_2}}\mu_{ij}}_{0} = \frac{\sum_{i\in I_1}(u_i - u_s - 1 - R^{(k)}(k_1,k_2)) + \frac{\sum_{i\in I_2}(u_i - u_s + \ln(1+t) - 1 - R^{(k)}(k_1,k_2))}{1+t}}{e^{u_s}\left(k_1 + k_2(1+t)\right) + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0}$$

$$R^{(k)}(k_1,k_2)\left(k_1 + \frac{k_2}{1+t}\right) = \sum_{i\in I_1} u_i + \frac{1}{1+t}\cdot\sum_{i\in I_2} u_i - (1+u_s)\cdot\left(k_1 + \frac{k_2}{1+t}\right) + \frac{k_2\ln(1+t)}{1+t}$$

$$R^{(k)}(k_1,k_2) = \frac{(1+t)\sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i + k_2\ln(1+t)}{k_1(1+t) + k_2} - 1 - u_s \tag{4.20}$$

We now have two equations relating $R^{(k)}(k_1,k_2)$ and u_s in (4.17) and (4.20). Using these equations we can find the values of the optimal revenues and all the pricing structure while varying k_1 and k_2. If we define the following constant:

$$C_1(k_1,k_2) = \frac{(1+t)\sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i + k_2\ln(1+t)}{k_1(1+t) + k_2} - 1, \tag{4.21}$$

Equation (4.20) becomes:

$$R^{(k)}(k_1,k_2) = C_1(k_1,k_2) - u_s, \tag{4.22}$$

and from Equation (4.22), we can deduce the following relations:

$$1 + R^{(k)}(k_1,k_2) = C_1(k_1,k_2) - u_s + 1, \qquad \text{and} \qquad e^{-(1+R^{(k)}(k_1,k_2))} = e^{u_s - C_1(k_1,k_2) - 1}. \tag{4.23}$$

We will use these relations on Equation (4.17). Let us first multiply both sides by the denominator on the right side:

$$R^{(k)}(k_1,k_2)\cdot\left(e^{u_s}\left[k_1 + \frac{k_2}{1+t}\right] + E(k_1,k_2)e^{-(1+R^{(k)}(k_1,k_2))} + a_0\right)$$
$$= e^{u_s}\sum_{i\in I_1}(u_i - u_s) + \frac{e^{u_s}}{1+t}\cdot\sum_{i\in I_2}(u_i - u_s + \ln(1+t)) + E(k_1,k_2)(1 + R^{(k)}(k_1,k_2))e^{-(1+R^{(k)}(k_1,k_2))}$$

using equations (4.23) to replace the value of $R^{(k)}(k_1, k_2)$ and write everything depending on u_s we have:

$$(C_1(k_1,k_2) - u_s)\left(e^{u_s}\left[k_1 + \frac{k_2}{1+t}\right] + E(k_1,k_2)e^{u_s - C_1(k_1,k_2)-1} + a_0\right) = e^{u_s}\sum_{i\in I_1}(u_i - u_s) + \frac{e^{u_s}}{1+t}\cdot\sum_{i\in I_2}(u_i - u_s + \ln(1+t)) + E(k_1,k_2)\cdot(C_1(k_1,k_2) - u_s - 1)\,e^{u_s - C_1(k_1,k_2)-1} \tag{4.24}$$

We focus first on the left hand side (LHS) of Equation (4.24):

$$LHS = (C_1(k_1,k_2) - u_s)\left(e^{u_s}\left[\left(k_1 + \frac{k_2}{1+t}\right) + E(k_1,k_2)e^{-C_1(k_1,k_2)-1}\right] + a_0\right)$$

For ease of notation, define $C_2(k_1,k_2)$ as:

$$C_2(k_1,k_2) = \left(k_1 + \frac{k_2}{1+t}\right) + E(k_1,k_2)e^{-C_1(k_1,k_2)-1} \tag{4.25}$$

Rewriting the LHS using the value for $C_2(k_1,k_2)$:

$$LHS = (C_1(k_1,k_2) - u_s)\,[e^{u_s}\cdot C_2(k_1,k_2) + a_0)] \tag{4.26}$$

We now focus on the right side (RHS) of equation (4.24):

$$\begin{aligned}
RHS =& e^{u_s}\sum_{i\in I_1}(u_i - u_s) + \frac{e^{u_s}}{1+t}\cdot\sum_{i\in I_2}(u_i - u_s + \ln(1+t)) \\
&+ E(k_1,k_2)\cdot(C_1(k_1,k_2) - u_s - 1)\,e^{u_s - C_1(k_1,k_2)-1} \\
RHS =& e^{u_s}\left[\sum_{i\in I_1} u_i + \frac{1}{1+t}\cdot\sum_{i\in I_2} u_i + \frac{k_2\ln(1+t)}{1+t} - u_s\left(k_1 + \frac{k_2}{1+t}\right)\right] \\
&+ e^{u_s}e^{-C_1(k_1,k_2)-1}E(k_1,k_2)\cdot(C_1(k_1,k_2) - u_s + 1) \\
RHS =& e^{u_s}\cdot\left(k_1 + \frac{k_2}{1+t}\right)[C_1(k_1,k_2) - u_s + 1] + e^{u_s}e^{-C_1(k_1,k_2)-1}E(k_1,k_2)\cdot(C_1(k_1,k_2) - u_s + 1) \\
RHS =& e^{u_s}\cdot(C_1(k_1,k_2) - u_s + 1)\cdot\underbrace{\left[\left(k_1 + \frac{k_2}{1+t}\right) + E(k_1,k_2)\cdot e^{-C_1(k_1,k_2)-1}\right]}_{C_2(k_1,k_2)} \\
RHS =& e^{u_s}\cdot(C_1(k_1,k_2) - u_s + 1)\cdot C_2(k_1,k_2)
\end{aligned} \tag{4.27}$$

Putting together equations (4.26) and (4.27), we have:

$$\begin{aligned} LHS =& RHS \\ (C_1(k_1,k_2) - u_s)\left[e^{u_s} \cdot C_2(k_1,k_2) + a_0)\right] =& e^{u_s} \cdot (C_1(k_1,k_2) - u_s + 1) \cdot C_2(k_1,k_2) \\ (C_1(k_1,k_2) - u_s)\, e^{u_s} \cdot C_2(k_1,k_2) + (C_1(k_1,k_2) - u_s) \cdot a_0 =& (C_1(k_1,k_2) - u_s)\, e^{u_s} \cdot C_2(k_1,k_2) + e^{u_s} \cdot C_2(k_1,k_2) \\ (C_1(k_1,k_2) - u_s) \cdot a_0 =& e^{u_s} \cdot C_2(k_1,k_2) \\ e^{u_s} =& -\frac{a_0}{C_2(k_1,k_2)} \cdot (u_s - C_1(k_1,k_2)) \end{aligned} \tag{4.28}$$

Equation (4.28) has a known explicit closed form solution, and can be found using the following Lemma:

Lemma 4. *Let $a, b \neq 0$ and c be real numbers and $W(\cdot)$ be the Lambert function [Corless et al., 1996]. The solution to the transcendental algebraic equation for x:*

$$e^{-ax} = b\,(x - c)\,, \tag{4.29}$$

is:

$$x = c + \frac{1}{a} \cdot W\left(\frac{ae^{-ac}}{b}\right). \tag{4.30}$$

Proof of Lemma 4. Let us start with equation (4.29) and find an explicit solution to it.

$$\begin{aligned} e^{-ax} &= b\,(x - c) \\ e^{-ax+ac-ac} &= b\,(x - c) && /\text{multiplying both sides by } \frac{a}{b} \cdot e^{a(x-c)} q \\ \frac{a \cdot e^{-ac}}{b} &= a \cdot (x - c) \cdot e^{a(x-c)} && /\text{using definition of } W(\cdot) \text{ as in Eq. (4.4)} \\ W\left(\frac{ae^{-ac}}{b}\right) &= a \cdot (x - c) && /\text{reorganising and isolating } x \\ x &= c + \frac{1}{a} \cdot W\left(\frac{ae^{-ac}}{b}\right) && , \end{aligned}$$

completing the proof. □

Identifying terms on equation (4.28), the solution for u_s is:

$$u_s = C_1(k_1,k_2) - W\left(\frac{C_2(k_1,k_2)}{a_0} \cdot e^{C_1(k_1,k_2)}\right) \tag{4.31}$$

Let us call this value $u_s(k, k_1, k_2)$, meaning that is a function of the integers k, k_1 and k_2. To get the revenue for this specific combination of parameters, we can simply use equation (4.22), giving us:

$$R^{(k)}(k_1,k_2) = C_1(k_1,k_2) - u_s(k,k_1,k_2) \tag{4.32}$$

Thus, the optimal revenue $\boldsymbol{R^{(k)}}$ given a specific integer k can be obtained by:

$$\boldsymbol{R^{(k)}} = \max_{\substack{k_1,k_2 \geq 1 \\ k_1+k_2 \leq k}} R^{(k)}(k_1,k_2) \tag{4.33}$$

Noting that there are $O(k^2)$ pairs (k_1,k_2) to evaluate, the proof follows. $\square$

For the non-trivial case with $\exp(u_1-u_k) > 1+t$, where a fixed price fails to be optimal, the prices need to be adjusted in order to avoid the dominances. Let $\boldsymbol{R^{(k)}}$ and $\boldsymbol{p^{(k)}}$ be the optimal revenue and price vector. The following Lemma characterizes the structure of the optimal solution for problem JAPTLM-k.

Lemma 5. *The optimal solution to problem* (JAPTLM-k) *is either the same as the unconstrained case (i.e. fixed price, in the case that* $\exp(u_1-u_k) \leq (1+t)$*) or the following holds at optimality:*

$$\frac{a_1(p_1)}{a_k(p_k)} = 1+t. \tag{4.34}$$

Moreover, there are non-negative integers k_1^*, k_2^**, with* $k_1^*+k_2^* \leq k$ *such that:*

$$\boldsymbol{R^{(k)}} = W\left(\frac{\left(k_1^*+\frac{k_2^*}{1+t}\right)\cdot\exp\left(\frac{(1+t)\sum_{i\in I_1}u_i+\sum_{i\in I_2}u_i+k_2^*\ln(1+t)}{k_1^*(1+t)+k_2^*}-1\right)+\sum_{i\in\bar{I}_k}\exp(u_i-1)}{a_0}\right),$$

where $I_1=[k_1^*]$*,* $I_2=\{k-k_2^*+1, k-k_2^*+2,\ldots,k\}$ *and* $\bar{I}_k=[k]\setminus(I_1\cup I_2)$*. The optimal prices can be obtained as follows:*

$$\boldsymbol{p_i^{(k)}} = \begin{cases} 1+\boldsymbol{R^{(k)}}+u_i-\frac{(1+t)\sum_{i\in I_1}u_i+\sum_{i\in I_2}u_i+k_2^*\ln(1+t)}{k_1^*(1+t)+k_2^*} & \text{if } i\in I_1, \\ 1+\boldsymbol{R^{(k)}}+u_i-\frac{(1+t)\sum_{i\in I_1}u_i+\sum_{i\in I_2}u_i+k_2^*\ln(1+t)}{k_1^*(1+t)+k_2^*}+\ln(1+t) & \text{if } i\in I_2, \\ 1+\boldsymbol{R^{(k)}} & \text{if } i\in\bar{I}_k. \end{cases} \tag{4.35}$$

Proof. The optimal revenue is already calculated in Equation (4.33). The proof follows by first obtaining $u_s^*(k)$ from Equation (4.20). Then, for products in I_1, the price can be obtained directly since their net utility is the same as $u_s^*(k)$. For products in I_2, since g_{1k} is satisfied with equality, all products share the same net utility and equal to $u_s^*(k)-\ln(1+t)$. Finally, for products in $\bar{I}_k$, we can use the relation provided in equation (4.15) to obtain the prices. More explicitly, let (k_1^*,k_2^*) be the integers satisfying $\boldsymbol{R^{(k)}} = R^{(k)}(k_1^*,k_2^*)$. To obtain the optimal prices, let $u_s^*(k)=u_s(k,k_1^*,k_2^*)$. By Equation (4.32) $u_s^*(k)$ can be written as:

$$u_s^*(k) = \frac{(1+t)\sum_{i\in I_1}u_i+\sum_{i\in I_2}u_i+k_2^*\ln(1+t)}{k_1^*(1+t)+k_2^*}-1-\boldsymbol{R^{(k)}} \tag{4.36}$$

Therefore, the optimal prices are given by:

$$\boldsymbol{p}_i^{(k)}(\boldsymbol{k}) = \begin{cases} u_i - u_s^*(k) & \text{if } i \in I_1, \\ u_i - u_s^*(k) + \ln(1+t) & \text{if } i \in I_2, \\ 1 + \boldsymbol{R}^{(k)} & \text{if } i \in \bar{I}_k \end{cases} \tag{4.37}$$

□

Let `TLM-Opt`(X, u, a_0, k) be the procedure to obtain the optimal solution for problem (JAPTLM-k). Using `TLM-Opt`(X, u, a_0, k) at most n times (once for each $k \leq n$) to obtain the assortment and prices yielding the highest $\boldsymbol{R}^{(k)}$, one can find the optimal assortment and price vector for any given instance. A description of the full algorithm to solve the Joint Pricing and Assortment Optimization under the Threshold Luce Model is provided below.

Algorithm 2: Joint Pricing and Assortment Optimization under the Threshold Luce Model

Data: X, u, a_0
Result: A set of products S and their prices p maximizing the expected revenue $R(S, p)$
Initialization:
$S^* = \emptyset$
$R^* = 0$
$p^* = 0$
Recall that products are indexed by decreasing utility.
for $k = 1, \ldots, |X|$ **do**
  $(\boldsymbol{R}^{(k)}, \boldsymbol{p}^{(k)}))$ =`TLM-Opt`(u, X, a_0, k)
  if $\boldsymbol{R}^{(k)} > R^*$ **then**
    $R^* \leftarrow \boldsymbol{R}^{(k)}$
    $p^* \leftarrow \boldsymbol{p}^{(k)}$
    $S^* \leftarrow [k]$
end
return S^*, R^*, p^*

The intuition behind `TLM-Opt`(X, u, a_0, k) is to mimic the optimal strategy for the regular MNL (Fixed-Price Policy) as much as possible. However, given that it needs to accommodate prices in order to avoid dominances, the algorithm adjusts prices for the higher intrinsic utility products (making prices larger, hence less attractive) and reduces the price of lower intrinsic utility ones, making them more attractive for customers and preventing them from being dominated. This allows the optimal strategy to have an edge over strategies ignoring the Threshold induced dominances, such as Fixed-Price Policy and, to a lesser extent, the Quasi-Same Price [Wang and Sahin, 2018]. The Quasi-Same Price policy policy only adjusts the price of the lowest attractiveness product, instead of adjusting both extremes of the attractiveness spectrum and potentially multiple products.

4.2 Numerical Experiments

This section presents some numerical results related to solve the Joint Assortment and Pricing Problem discussed in Section 4.1. We analyse the performance of algorithm `TLM-Opt`, compared against Fixed-Price strategy, which is optimal for the MNL and Quasi-Same Price strategy [Wang and Sahin, 2018], which is optimal for the MNL variant considered in their paper that takes into consideration search cost, and it basically a fixed price for all products but one, which share some similarities with our proposed pricing policy, as it is fixed price in general but the higher and lower ends of the utility spectrum.

Each tested family or class of instances is characterized by three numbers: the number of products n; the threshold t, that controls how tolerant are customers with respect to differences in attractiveness and the attractiveness of the outside option a_0, which controls how likely is that customers review all products without purchasing. In total, we experimented with 48 classes or families of instances, each containing 250 instances. In each specific instance, revenues and utilities are drawn from an uniform distribution between 0 and 10. We ran the three strategies: Fixed Price, Quasi-Same Price and `TLM-Opt`, and report the average and worst optimality gap for Fixed Price and Quasi-Same Price strategies, as well as cardinality of the offered set for both strategies. These numerical experiments were conducted in Python 3.6 on a computer with 8 processors (each with 3.6 GHz CPU) and 16 GB of RAM. Table 4.2 presents the results which can be summarized as follows:

1. As expected `TLM-Opt` outperformed the other two algorithms in terms of revenue, and being quite fast to execute (less than half of a second for all the instances simulated).

2. Fixed-Price policy performs the worst across the board, which is expected given that it has the lowest degrees of freedom, as shown in example 18. Although the average gap is quite low, it can be as high as 43.027%. In fact, fixed-price policy can be arbitrarily bad. A proof of this fact is provided in Appendix B, Lemma 6.

3. Quasi-Same price policy also performs well on average, and the worst gap obtained was 29.964%, which is significantly better than the worst gap for Fixed Price policy.

4. The cardinality of the optimal solution is always at least the same or greater than Fixed-Price policy. This can be observed empirically, or deduced analytically. The intuition behind it is that given the functional form of the revenue for Fixed-Price and the fact that the Lambert function is strictly increasing the strategy always try to show as much as possible. This, and the fact that under same price,the dominance relation only depends upon intrinsic utilities, imply that there is a limit on the number of products that the fixed price policy can

offer [1] without causing any domination for low intrinsic utility products. On the other hand, under `TLM-Opt` (or Quasi Same price) we can go further and add products in such a way that the dominance relations are not triggered, and therefore we can include more products.

5. The main difference stems from the fact that our strategy leverage both ends of the utility spectrum, and reveals the following interesting insight. Sometimes in order to avoid low attractiveness products to be dominated, we want to: increase the price of the higher utility products (to make them less attractive) and at the same time, reduce the price for lower utility products, in order to make them more attractive, and making them *visible* for the consumer.

[1]the last product 'k' where $\frac{a_1(p_1)}{a_k(p_k)} = \exp(u_1 \check{\ } u_k) \leq (1+t)$

(n,t,a_0)	Fixed Price			Quasi Same Price			`TLM-Opt`
	Avg. Gap (%)	Worst Gap (%)	Avg. Cardinality	Avg. Gap (%)	Worst Gap (%)	Avg. Cardinality	Avg. Cardinality
(5,0.5,1)	2.164728	17.514	1.212	0.442364	8.609	2.212	1.92
(5,0.5,10)	2.925244	27.779	1.26	0.54798	11.074	2.248	1.888
(5,0.5,100)	4.136064	43.027	1.24	1.16004	29.964	2.108	1.856
(5,1,1)	1.575348	13.446	1.384	0.253996	4.638	2.38	2.028
(5,1,10)	2.074672	24.984	1.472	0.362416	11.116	2.448	2.04
(5,1,100)	2.726188	33.938	1.416	0.52712	14.404	2.308	1.952
(5,2,1)	0.9865	8.881	1.58	0.723548	8.881	1.812	2.132
(5,2,10)	1.4685	10.592	1.536	1.040004	10.592	1.84	2.116
(5,2,100)	2.343244	32.77	1.624	1.520068	22.581	1.924	2.196
(5,5,1)	0.415064	5.103	2.012	0.153776	3.748	2.64	2.468
(5,5,10)	0.918528	17.044	1.844	0.460556	17.044	2.424	2.384
(5,5,100)	1.092544	11.539	1.972	0.466548	7.574	2.552	2.468
Avg. $n = 5$	1.902219	20.55142	1.546	0.638201	12.51875	2.241333	2.120667
(10,0.5,1)	3.63332	14.951	1.408	1.079656	6.774	2.408	2.896
(10,0.5,10)	4.710012	30.328	1.512	1.499744	15.575	2.512	3.132
(10,0.5,100)	6.7165	24.489	1.42	2.465028	23.517	2.364	2.912
(10,1,1)	2.69928	12.027	1.748	0.835872	7.352	2.748	3.264
(10,1,10)	3.563196	15.769	1.672	1.264116	10.296	2.672	3.18
(10,1,100)	4.822544	26.928	1.756	1.704816	15.125	2.74	3.264
(10,2,1)	1.38662	8.541	2.076	0.803492	8.541	2.576	3.236
(10,2,10)	2.37252	15.852	2.016	1.284344	15.852	2.544	3.38
(10,2,100)	3.115392	18.694	2.076	1.53734	18.694	2.612	3.34
(10,5,1)	0.611308	4.19	2.888	0.322156	2.961	3.492	3.908
(10,5,10)	0.931108	5.537	2.804	0.523108	4.975	3.432	3.84
(10,5,100)	1.323312	12.683	2.828	0.705452	12.683	3.44	3.936
Avg. $n = 10$	2.990426	15.83242	2.017	1.16876	11.86208	2.795	3.357333
(20,0.5,1)	5.227892	16.189	1.964	2.406408	10.383	2.964	5.412
(20,0.5,10)	6.505472	18.556	1.844	2.926688	11.734	2.844	4.868
(20,0.5,100)	9.65628	30.904	1.844	4.633812	20.602	2.84	5.104
(20,1,1)	3.917928	11.225	2.332	1.87528	7.241	3.332	5.476
(20,1,10)	4.640684	20.635	2.32	2.227456	14.112	3.32	5.384
(20,1,100)	6.765772	26.431	2.368	3.075284	17.929	3.368	5.48
(20,2,1)	2.197276	9.372	3.324	1.210484	9.372	4.112	6.164
(20,2,10)	2.669532	10.396	3.316	1.449576	9.11	4.168	6.252
(20,2,100)	3.708316	15.228	3.28	2.055808	14.009	4.092	5.924
(20,5,1)	0.878752	4.584	4.636	0.577612	4.584	5.236	6.976
(20,5,10)	1.244528	5.209	4.632	0.805216	5.209	5.26	7.1
(20,5,100)	1.718416	11.142	4.664	1.138056	7.856	5.312	7.192
Avg. $n = 20$	4.094237	14.98925	3.043667	2.031807	11.01175	3.904	5.944333
(30,0.5,1)	6.343964	16.145	2.24	3.642808	10.606	3.24	7.344
(30,0.5,10)	8.202948	22.238	2.224	4.572832	15.339	3.224	7.32
(30,0.5,100)	10.6693	26.659	2.228	6.018692	18.973	3.224	7.2
(30,1,1)	4.0957	14.803	3.044	2.266944	10.581	4.044	7.464
(30,1,10)	5.039272	19.686	3.24	2.906664	13.583	4.24	7.892
(30,1,100)	7.451636	23.606	3.084	4.349844	14.428	4.084	7.884
(30,2,1)	2.193896	12.267	4.344	1.32162	9.526	5.252	8.588
(30,2,10)	3.110752	13.315	4.252	1.911924	9.177	5.148	8.64
(30,2,100)	4.005252	19.963	4.476	2.473688	19.963	5.4	8.684
(30,5,1)	1.023416	4.25	6.26	0.74478	4.25	6.88	10.132
(30,5,10)	1.328936	5.534	6.328	0.93428	4.053	7.012	10.232
(30,5,100)	1.730272	6.284	6.436	1.189296	6.164	7.136	10.104
Avg. $n = 30$	4.599612	15.39583	4.013	2.694448	11.38692	4.907	8.457

Table 4.2: Numerical experiments comparing Fixed-Price and Quasi-Same price against `TLM-Opt`. For each class of instances, for non-optimal strategies we display the average optimality gap, worst-case gap and the cardinality of the offered set. We also provide the average of those metrics for each value of n considered.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, we studied models for consumer choice that fall outside RUM. Going outside of this well studied class allows practitioners to model effects that were not conventionally picked up using models belonging to RUM [Simonson and Tversky, 1992; Tversky and Simonson, 1993], where there are clear drawbacks of not being able to accommodate them [Jagabathula and Rusmevichientong, 2018]. Our efforts were mainly devoted to provide optimisation techniques to solve optimisation problems of recently proposed models that overcome some of the shortcomings of RUMs.

Chapter 2 studied the assortment optimisation problem under the *Sequential Multinomial Logit* (SML) Model, a discrete choice model that generalises the MNL Model. Under the SML Model, products are partitioned into two levels. When a consumer is presented with such an assortment, she first considers products in the first level and, if none of them are appropriate, products in the second level are considered. The SML is a special case of the PALM recently proposed by Echenique et al. [2018]. It can explain many behavioural phenomena such as the attraction, compromise, and similarity effects which cannot be explained by the MNL Model or any discrete choice model based on random utility.

Through a structural analysis of the optimal solutions, we showed that the seminal concept of revenue-ordered assortments can be generalised to the SML Model. We proved that all optimal assortments under the SML Model are revenue-ordered by level, a natural generalisation of revenue-ordered assortments. As a corollary, the assortment optimisation problem under the SML Model is solvable in polynomial time.

Chapter 3 studied the assortment optimisation problem under the 2SLM, a discrete choice model introduced by Echenique and Saito [2018] that generalises the standard MNL Model with a dominance relation and may violate regularity. We proved that the assortment problem under the 2SLM can be solved in polynomial time. We considered the capacitated assortment optimisation problem under the 2SLM and proved that the problem becomes NP-hard in this setting. We also provide polynomial-time algorithms for cases of the capacitated assortment optimisation problem where (1) the dominance relation is attractiveness correlated and when (2) its transitive reduction is a forest. We provide a set of numerical experiments to highlight the performance of the proposed algorithm against the classical revenue-ordered assortment strategy. Theoretically, as we show in Example 14, the gap can be arbitrarily bad. In our experimental setting, we found cases where the gap was as large as 95% (see Table 3.2 for the detailed information).

Chapter 4 is an in-depth study of the pricing problem under the 2SLM. We first note that changes in prices should be reflected in the dominance relation, if the differences between the resulting attractiveness are large enough. This is formalised by solving the joint assortment and pricing problems under the Threshold Luce Model, where one product dominates another if the ratio between their attractiveness is greater than a fixed threshold. In this setting, we show that this problem can be solved in polynomial time. The main difference between our proposed pricing scheme, and fixed-price policy as a solution for the classical MNL, stems from the fact that our strategy leverages both ends of the utility spectrum, and reveals the following interesting insight: sometimes to avoid low attractiveness products to be dominated, we want to: increase the price of the higher utility products (to make them less attractive) and at the same time, reduce the price for lower utility products, to make them more attractive, and making them *visible* to the consumer. We also provide numerical experiments to show the benefits of considering the algorithm developed in this chapter, against two strategies: fixed-price policy, and the quasi-same price policy [Wang and Sahin, 2018], given the similarity of having a cut-off point in the form of a consideration set induced by prices. We found that although the average gap was relatively low (compared to both alternative strategies), the worst gap observed for the fixed-price policy was as high as 43%, and for the quasi-same price policy was as high as 30%, showing that the impact of leveraging pricing on both sides of the spectrum of utility can provide substantial benefits if customer behaviour follows this model of customer choice.

It is important to note that the TLM offers better prediction power than the usual MNL model. Recently, [Wang, 2019] showed that compared to the MNL model, the TLM improved prediction accuracy as much as 11% on synthetic data, and improve market share and revenue prediction by roughly 15%. This makes the results from Chapters 3 and 4 relevant for practitioners.

5.2 Future Work

There are many interesting avenues for future research. The main open issue in Chapter 2, is to extend the results to the PALM, which has an arbitrary number of levels. Note that one can easily extend the algorithm of revenue-ordered by level to the PALM and it would take at most $\mathcal{O}(|X|^{k+1})$ time, where k is the number of levels. [1]

We tried to generalise the results to three levels, but the same strategy applied for the two-level SML model in this thesis failed for three levels. When building the foundations for the two-level SML model, we based our proofs in expressing the

[1] Note that $\mathcal{O}(|X|^k)$ is the number assortments that are revenue-ordered by level.

optimal solution as convex combinations of non-optimal subsets and Kindly check format. using this we obtain the desired bounds. Attempting this strategy for the third level on the three level SML model led to a point where there is no guarantee that the coefficients add up to less than one. So we cannot apply the same strategy developed for the two level SML model, given that this result is used to prove the main theorem (Theorem 1) of Chapter 2.

We also executed our algorithm over a series of PALM instances by varying the number of levels and the revenue-ordered assortment algorithm always returned the optimal solution. Our conjecture is that the optimality result of revenue ordered assortments by level holds for the general PALM, but the problem remains open. A second interesting research avenue is to consider a new discrete choice model that allows decision makers to change the order in which the levels are presented to consumers. In the SML, the level ordering is intrinsic to products, but one may consider settings in which decision makers can choose, not only what to show, but also the priority associated with each of the displayed products. A similar model is currently being studied in Liu et al. [2018]. The functional form is slightly different than the PALM, but it is also characterised by iterative applications of the MNL. The authors showed that the assortment optimisation problem, including the location of the items in the corresponding levels is NP-hard, and provide a fully polynomial-time approximation scheme. Another research direction is to study the assortment optimisation problem under the SML with cardinality or space constraints. Finally, it is important to develop efficient procedures to estimate the parameters of the SML model based on historical data (e.g., van Ryzin and Vulcano [2017]).

Based on Chapter 3 and Chapter 4, there are many interesting avenues for future research. First, one may wish to study how to further generalise the 2SLM, while keeping the assortment problem solvable in polynomial time. For example, one can try to check whether there exists a model that unifies the 2SLM and the elegant work in Davis et al. [2013], where the assortment problem is still solvable in polynomial time. Second, given that the capacitated version of the 2SLM is NP-hard under Turing reductions (theorem 5), it is interesting to see whether there exist good approximation algorithms for this problem. Third, one can explore different forms of dominance. For example, one may consider dominance specified by a discrete relation or a continuous functional form between products. Fourth, one can try to generalise our results for the joint assortment and pricing problems under the TLM to a more general setting, where price sensitivity is dependent on each product. Finally, one can try to mix attention models with dominance relations, meaning that a customer first perceives a subset of products, dictated by an attention filter, and then filters the products even more using dominance relations.

Another alternative for future research, given that the TLM produces a better fit than the MNL model Wang [2019], is to check if the results of trial-offer markets can be replicated using the TLM instead, where dominance relation can be induced by imbalances in position bias, social influence, or any combination of the two. An open general question that arises as a consequence of this research is to find what is the inherent condition (or set of conditions) that allows a model outside RUM to still produce polynomial-time exact solutions for the assortment problem, and if there is any way to characterise those conditions.

Finally, it would be interesting to analyse both the assortment and pricing problems, with the addition of planning in a finite time horizon. In this setting, the firm has to plan in advance considering customer strategic behaviour, potentially taking into account scarcity in the products to be offered and considering that customers follow the 2SLM and dominance relations are present among products.

Appendix A

On the robustness of the MusicLab model: continuation and further analysis

This appendix is reproduced with minor changes from:

Van Hentenryck, P.; **Flores, A.**; Berbeglia, G., 2017. Trial-Offer Markets with Continuation. *Presented in the 21st International Federation of Operational Research Societies Conference, Quebec (IFORS 2017)*

Trial-offer markets, where customers can sample a product before deciding whether to buy it, are ubiquitous in the online experience. Their static and dynamic properties are often studied by assuming that consumers follow a multinomial logit model and try exactly one product. This chapter, is an attempt at generalizing Multinomial Logit Models to account for a richer class of customer behavior. It endows the Multinomial Logit Model with a notion of continuation, which enables participants to sample multiple products before making a purchase. We studied how this generalization affects market efficiency and the role of social influence. The main contributions can be summarized as follows:

1. We show that a trial-offer market with continuation can be reduced to a traditional trial-offer market by adjusting the quality and appeal of the products and we quantify how the continuation model affects market efficiency;
2. We show that, under a natural continuation model, the quality-ranking policy, where the products are ranked by quality, is preserved by the reduction, but not the performance ranking, which optimises the market performance at each step. We also show that social influence remains beneficial in this setting under the quality ranking;
3. Finally, we show experimental results that indicate that the popularity ranking, which ranks the product by popularity, benefits more from the generalization than the quality and performance ranking, unless the continuation is strongly dependent of the product just sampled. This improvement however is not enough to bridge the gap with the performance and quality rankings.

A.1 Trial-Offer Markets

We consider trial-offer markets in which participants can try a product before deciding whether to buy it. Such settings are common in online cultural markets (e.g., books, songs, and videos). In this chapter, the trial-offer market is composed of n products and each product $i \in \{1, \ldots, n\}$ is characterized by two values:

1. Its *appeal* A_i representing the inherent preference of trying product i;

2. Its *quality* q_i representing the probability of purchasing product i given that it was tried.

Each participant, when entering the market, is presented with a product list π: She then tries a product s in π and decides whether to purchase s with a certain probability. The product list is a permutation of $\{1, \ldots, n\}$ and each position p in the list is characterized by its *visibility* $v_p > 0$ which is the inherent probability of trying a product in position p. Since the list π is a bijection from positions to products, its inverse is well-defined and is called a ranking. We denote rankings by σ in the following, π_i denotes the product in position i of the list π, and σ_i denotes the *position* of product i in the ranking σ. Therefore v_{σ_i} denotes the visibility of the position of product i.

The probability of trying product i given a list σ is

$$p_i(\sigma) = \frac{v_{\sigma_i} A_i}{\sum_{j=1}^{n} v_{\sigma_j} A_j}.$$

Given a ranking σ, the expected number of purchases is

$$\lambda(\sigma) = \sum_{i=1}^{n} p_i(\sigma)\, q_i. \tag{A.1}$$

The traditional static market optimisation problem consists of finding a ranking σ^* maximising $\lambda(\sigma)$, i.e.,

$$\sigma^* = \underset{\sigma \in S_n}{\operatorname{argmax}} \sum_{i=1}^{n} p_i(\sigma)\, q_i \tag{A.2}$$

where S_n represents the symmetry group over $\{1, \ldots, n\}$. Observe that consumer choice preferences for trying the products are essentially modeled as a discrete choice model based on a multinomial logit Luce [1959] in which product utilities are affected by their position.

Social Influence Following Krumme et al. [2012], we considers a dynamic market where the appeal of each product changes over time according to a social influence signal. Given a social signal $d = (d_1, \ldots, d_n)$, where d_i denotes the number of purchases of product i, the appeal of i becomes $A_i + d_i$ and hence the probability of trying i given a list σ becomes

$$p_i(\sigma, d) = \frac{v_{\sigma_i}(A_i + d_i)}{\sum_{j=1}^{n} v_{\sigma_j}(A_j + d_j)}.$$

Note that the probability of trying a product depends on its position in the list, its appeal, and its number of purchases ($d_{i,t}$) at time t. As the market evolves over time, the number of purchases could dominate the appeal, and the sampling probability of a product becomes its market share. Without social influence, a dynamic market reduces to solving the static optimisation problem repeatedly. This set-up is the *independent condition*.

In the following, without loss of generality, we assume that the qualities and visibilities are non-increasing, i.e., $q_1 \geq q_2 \geq \cdots \geq q_n$ and $v_1 \geq v_2 \geq \cdots \geq v_n$. We also assume that the qualities and visibilities are known. In practical situations, the product qualities are obviously not known. But, as shown by Abeliuk et al. [2015], they can be recovered accurately and quickly, either before or during the market execution. For simplicity, we use $a_{i,t} = A_i + d_{i,t}$ to denote the appeal of product i at step t. When the step t is not relevant, we omit it and use a_i instead.

Ranking policies Following Abeliuk et al. [2015], we explore several ranking policies. The *performance ranking* maximises the expected number of purchases at each iteration, exploiting all the available information globally, i.e., the appeal, the visibility, the purchases, and the quality of the products. More precisely, the performance ranking at step k produces a ranking σ_k^* defined as

$$\sigma_k^* = \operatorname*{argmax}_{\sigma \in S_n} \sum_{i=1}^{n} p_i(\sigma, d_k) \cdot q_i$$

where $d_k = (d_{1,k}, \ldots, d_{n,k})$ is the social influence signal at step k. The performance ranking uses the probability $p_i(\sigma, d_k)$ of trying products i at iteration k given ranking σ, as well as the quality q_i of product i. The performance ranking can be computed in strongly polynomial time and the resulting policy is scalable to large markets Abeliuk et al. [2015]. The *quality ranking* simply orders the products by quality, assigning the product of highest quality to the most visible position and so on. With the above assumptions, a quality ranking σ satisfies $\sigma_i = i \quad (1 \leq i \leq n)$. The *popularity ranking* was used by Salganik et al. [2006] to show the unpredictability caused by social influence in cultural markets. At iteration k, the popularity ranking orders the products by the number of purchases $d_{i,k}$, but these purchases do not necessarily reflect the inherent quality of the products, since they depend on how many times the products were tried. We also follow Abeliuk et al. [2015] and use Q-RANK, D-RANK, and P-RANK to denote the policies using the quality, popularity, and performance rankings respectively. We also use R-RANK to denote the policy that simply presents a random order at each period.

A.2 Trial-Offer Markets With Continuation

The main goal of this chapter is to study trial-offer markets with continuation, i.e., a setting where market participants can continue shopping even when they decline to purchase the product just sampled. We model such a trial-offer market by adding a continuation probability

$$c_i = f(\cdot)(1 - q_i) \tag{A.3}$$

to continue shopping after a participant has declined to purchase product i. In the above probability, the $(1 - q_i)$ term represents the fact that the participant has declined to purchase product i and the $f(\cdot)$ term represents a function that might depend on the product quality, the current position, or even on another overall measure (or a combination of all these factors). Figure A.1 shows a graphic representation of a trial-offer market with continuation. It uses $\bar{c}_i = 1 - c_i$ to denote the probability that a participant leaves the market place after sampling product i.

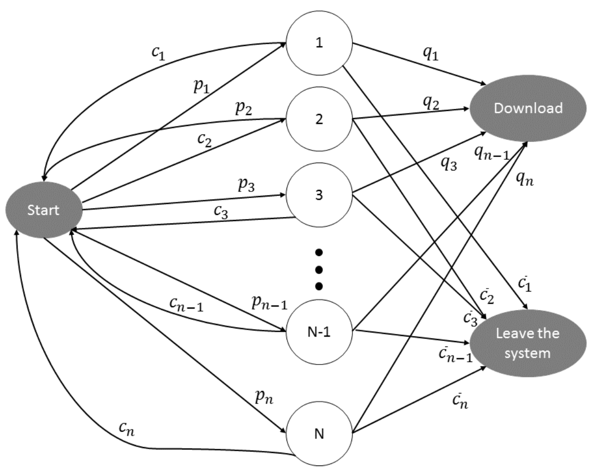


Figure A.1: Trial-Offer market with continuation.

The expected number of purchases in the static version of the trial-offer market with continuation for a ranking σ is denoted by $\overline{\lambda(\sigma)}$ and defined by

$$\overline{\lambda(\sigma)} = \sum_{i=1}^{n} p_i(\sigma)(q_i + c_i\overline{\lambda(\sigma)}) \tag{A.4}$$

Our primary objective is to maximise market efficiency, i.e., the expected purchases:

$$\sigma^* = \underset{\sigma \in S_n}{\operatorname{argmax}}\, \overline{\lambda(\sigma)}. \tag{A.5}$$

Note that the higher this objective is, the lower the probability that consumers try a product but then decide not to purchase it. Hence, if we interpret this last action as an inefficiency, maximising the expected efficiency of the market minimizes unproductive trials.

We prove a number of results when the continuation c_i depends polynomially on q_i, i.e.,

$$c_i = \rho q_i^r (1 - q_i) \tag{A.6}$$

where $\rho \leq 1$ controls the overall tendency to continuation and $r \geq 0$ represents the influence of q. This choice is justified intuitively by the fact that a market participant is more likely to continue sampling if the product she tried is of high quality, because it reflects on how good the other products potentially are. Figure A.2 depicts various choices of ρ and r.

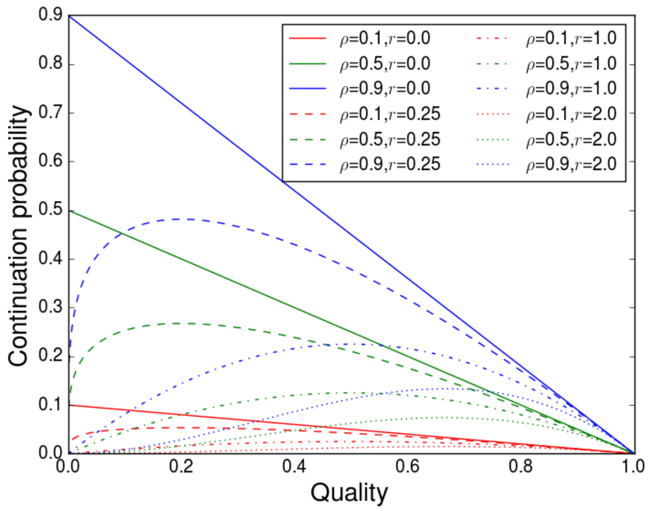


Figure A.2: Examples of the continuation probabilities for different values of ρ and r; the r parameter defines where the peak is (the maximum is always attained at $q = \frac{r}{r+1}$), and ρ modulates how strong the continuation is.

A.3 Reduction to the Trial-Offer Model

This section shows that the trial-offer market with continuation can be reduced to a trial-offer market. Indeed, rearranging the terms in Equation (A.4) leads to

$$\overline{\lambda(\sigma)} = \sum_{i=1}^{n} p_i(\sigma)(q_i + c_i \overline{\lambda(\sigma)})$$

$$\overline{\lambda(\sigma)} = \frac{\sum_{i=1}^{n} p_i(\sigma) q_i}{1 - \sum_{i=1}^{n} p_i(\sigma) c_i}$$

Defining

$$\overline{p_i(\sigma)} = \frac{p_i(\sigma)}{1 - \sum_{i=1}^{n} p_i(\sigma) c_i} \tag{A.7}$$

we obtain

$$\overline{\lambda(\sigma)} = \sum_{i=1}^{n} \overline{p_i(\sigma)}\, q_i$$

By definition of $p_i(\sigma)$, we have

$$\overline{p_i(\sigma)} = \frac{v_{\sigma_i} a_i}{\sum_{i=1}^{n} v_{\sigma_i} a_i} \cdot \frac{1}{1 - \sum_{i=1}^{n} (c_i \cdot \frac{v_{\sigma_i} a_i}{\sum_{i=1}^{n} v_{\sigma_i} a_i})}$$
$$\overline{p_i(\sigma)} = \frac{v_{\sigma_i} a_i}{\sum_{i=1}^{n} (1 - c_i) v_{\sigma_i} a_i}$$

Now, by defining $\overline{a_i} = a_i(1 - c_i)$ and $\overline{q_i} = \frac{q_i}{(1-c_i)}$, we obtain

$$\overline{\lambda(\sigma)} = \sum_{i=1}^{n} \frac{v_{\sigma_i} \overline{a_i}\, \overline{q_i}}{\sum_{i=1}^{n} v_{\sigma_i} \overline{a_i}}$$

To understand this reduction intuitively, we can rewrite Equation A.7 as:

$$\overline{p_i(\sigma)} = p_i(\sigma) \cdot \sum_{j=1}^{\infty} (\sum_{i=1}^{n} p_i(\sigma) c_i)^j$$

The value $\overline{p_i(\sigma)}$ can thus be interpreted as the probability of sampling product i in any number of steps. The rewriting uses the fact that $\sum_{i=1}^{n} p_i(\sigma) c_i < 1$ to obtain an infinite sum and the term $(\sum_{i=1}^{n} p_i(\sigma) c_i)^j$ captures all the possible ways to sampling i in j steps.

We have proven the following theorem:

Theorem 11. *A trial-offer market with continuation can be reduced to a trial-offer market by using the product qualities $\overline{q_i}$ and appeals $\overline{a_i}$ defined as follows:*

$$\overline{q_i} = \frac{q_i}{1 - c_i}$$
$$\overline{a_i} = a_i(1 - c_i).$$

In the following, $\overline{q_i}$ and $\overline{a_i}$ are called the *continuation qualities* and *continuation appeals*, and Figure A.3 depicts the continuation quality for different values of ρ and r. Observe how the continuation model typically boosts the quality of the products, sometimes substantially.

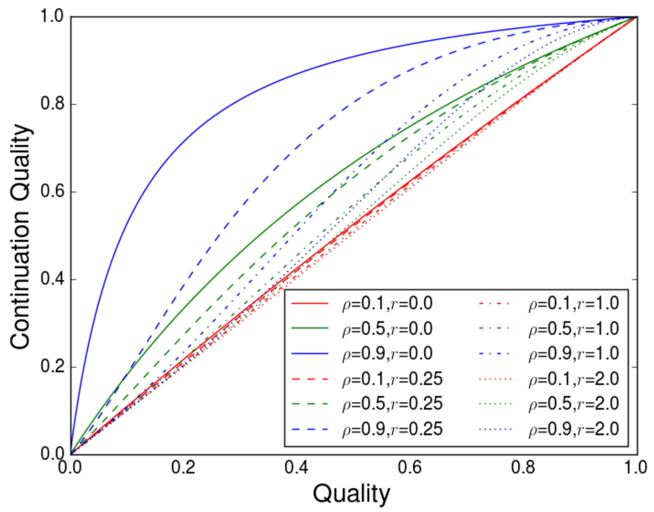


Figure A.3: Continuation qualities for different values of the ρ and r parameters; the larger ρ is, the more *concave* the continuation quality becomes.

A.4 Properties of the Market

Market Efficiency The first result links the expected number of purchases of the market with and without continuation under the performance ranking.

Theorem 12. *Let π_c^* and π^* be optimal permutations for the trial-offer markets with and without continuation. Then,*

$$\lambda(\pi^*) \leq \overline{\lambda}(\pi_c^*) \leq \frac{\lambda(\pi^*)}{1 - \max_i c_i}.$$

Proof. The lower bound can be derived as follows:

$$\begin{aligned}
\lambda(\pi^*) &= \frac{\sum_{i=1}^n v_i a_{\pi_i^*} q_{\pi_i^*}}{\sum_{i=1}^n v_i a_{\pi_i^*}} \\
\lambda(\pi^*) &= \frac{\sum_{i=1}^n v_i a_{\pi_i^*} q_{\pi_i^*}}{\sum_{i=1}^n v_i a_{\pi_i^*}} \cdot \frac{1 - \sum_{i=1}^n p_i(\pi^*) c_{\pi_i^*}}{1 - \sum_{i=1}^n p_i(\pi^*) c_{\pi_i^*}} \\
\lambda(\pi^*) &= \underbrace{\frac{\sum_{i=1}^n p_i(\pi^*) q_{\pi_i^*}}{1 - \sum_{i=1}^n p_i(\pi^*) c_{\pi_i^*}}}_{\overline{\lambda}(\pi^*)} \underbrace{(1 - \sum_{i=1}^n p_i(\pi^*) c_{\pi_i^*})}_{\leq 1} \\
\lambda(\pi^*) &\leq \overline{\lambda}(\pi^*) \leq \overline{\lambda}(\pi_c^*)
\end{aligned}$$

where the last inequality holds because of optimality of π_c^* for $\overline{\lambda}(\cdot)$. The upper bound follows from

$$\overline{\lambda}(\pi_c^*) \leq \lambda(\pi_c^*) \cdot \frac{1}{1 - \sum_{i=1}^{n} p_i c_{\pi_i^*}} \leq \lambda(\pi^*) \cdot \frac{1}{1 - \max_i c_i}.$$

□

The following corollary considers the case where continuations depend polynomially on qualities

Corollary 3. *Assume that $c_i = \rho q_i^r (1 - q_i)$. It follows that*

$$\lambda(\pi^*) \leq \overline{\lambda}(\pi_c^*) \leq \lambda(\pi^*) \frac{1}{1 - \frac{\rho r^r}{(r+1)^{r+1}}} \tag{A.8}$$

Proof. The proof follows from examining the maximum value for the c_i.

$$\max_i c_i = \max_i \rho q_i^r (1 - q_i) \leq \rho \max_{x \in [0,1]} x^r (1 - x) = \frac{\rho r^r}{(r+1)^{r+1}}$$

where the last equality holds because the maximum value of $x^r(1 - x)$ is reached when $x = \frac{r}{r+1}$. □

When $\rho = 1$ and $r = 1$, $\lambda(\pi_c^*) \leq \frac{4}{3} \cdot \lambda(\pi^*)$ indicating a market that is at most 33% more efficient.

Prior work on trial-offer markets with the social influence signal considered here has shown that the quality ranking is asymptotically optimal Van Hentenryck et al. [2016]: The market converges towards a monopoly for the product of highest quality. We now show that, when the continuations are polynomial in product qualities, the quality ranking is preserved by the reduction and hence the two markets, with and without continuation, converge to the same equilibrium in market shares.

Proposition 11. *Let $c_i = \rho q_i^r (1 - q_i)$ with $\rho \in (0, 1)$ and $r \geq 0$. Then $q_i \leq q_j \Leftrightarrow \overline{q_i} \leq \overline{q_j}$.*

Proof. It is sufficient to show that $\overline{q_i}$, when viewed as a function of q_i, is increasing in $(0, 1)$. Consider such function

$$h(x) = \frac{x}{1 - \rho x^r (1 - x)}$$

and its derivative

$$\frac{dh(x)}{dx} = \frac{\rho x^r \left[(r-1) - rx\right] + 1}{((1 - \rho x^r (1 - x))^2}.$$

The denominator is greater than zero, so it remains to show that the numerator also is. The term ρx^r is increasing in x and the term $[(r - 1) - rx]$ is a line decreasing in x. The product is minimized when $x = 1$, in which case the product has a value of $-\rho$. Since $\rho \in (0, 1)$, the minimum value of the numerator is $1 - \rho \geq 0$, which concludes the proof. □

More importantly, it is also possible to show that, under the quality ranking, the probability that the next purchase is product i is the same in the markets with and without continuation. Hence, from a product standpoint, the markets behave very similarly.

Proposition 12. *The probability p_i that the next purchase (after any number of steps) is product i is*

$$p_i = \frac{v_i a_i q_i}{\sum\limits_{j=1}^{n} v_j a_j q_j}.$$

Proof. The probability that product i is purchased in the first step is

$$p_i^{1st} = \frac{v_i a_i}{\sum\limits_{j=1}^{n} v_j \overline{a_j}} q_i,$$

More generally, the probability that product i is purchased in step m while no product was purchased in earlier steps is:

$$p_i^{mth} = \left(\frac{\sum\limits_{j=1}^{n} v_j \overline{a_j}(1 - \overline{q_j})}{\sum\limits_{j=1}^{n} v_j \overline{a_j}} \right)^{m-1} \frac{v_i a_i}{\sum\limits_{j=1}^{n} v_j \overline{a_j}} q_i.$$

Defining $\beta = (\sum\limits_{j=1}^{n} v_j a_j q_j)/(\sum\limits_{j=1}^{n} v_j \overline{a_j})$, Equation A.4 becomes

$$p_i^{mth} = \left(1 - \beta\right)^{m-1} \frac{v_i a_i}{\sum\limits_{j=1}^{n} v_j \overline{a_j}} q_i.$$

Hence the probability that the next purchased product is product i is given by

$$p_i = \sum_{m=0}^{\infty} \left(1 - \beta\right)^{m} \frac{v_i a_i}{\sum\limits_{j=1}^{n} v_j \overline{a_j}} q_i.$$

Given the fact that $\beta < 1$ we have:

$$\sum_{m=0}^{\infty} \left(1 - \beta\right)^{m} = \frac{1}{\beta},$$

the probability that the next purchase is product i is given by

$$p_i = \frac{\sum_{j=1}^{n} v_j \overline{a_j}}{\sum_{j=1}^{n} v_j a_j q_j} \cdot \frac{v_i a_i}{\sum_{j=1}^{n} v_j \overline{a_j}} q_i.$$

$$p_i = \frac{v_i a_i q_i}{\sum_{j=1}^{n} v_j a_j q_j}$$

□

In contrast, the same results do not hold for the performance ranking, which may change when a continuation is used, as shown by the following example.

Example 19. Consider the following instance with 3 songs:

- Visibilities: $v_1 = 0.8$, $v_2 = 0.5$ and $v_3 = 0.1$
- Qualities: $q_1 = 0.9$, $q_2 = 0.2$ and $q_3 = 0.6$
- Appeals: $a_1 = 0.9$, $a_2 = 0.1$ and $a_3 = 0.3$
- Continuation parameters: $\rho = 0.8$ and $r = 0.7$

In this case, the performance ranking for the market without continuation is $\sigma^* = [1, 2, 3]$; It is $\sigma_c^* = [1, 3, 2]$ for the continuation model.

Position Bias This result generalizes the result shown in Van Hentenryck et al. [2016], to the continuation setting, and it means that we can always benefit from position bias. The formalization of this claim can be seen below

Theorem 13. *Position bias increases the expected number of purchases under the quality-ranking policy, i.e., for all visibilities v_i, appeals a_i, qualities q_i $(1 \leq i \leq n)$ and continuation probabilities c_i.This is, after we make the reduction to the Associated Multinomial Logit, we have:*

$$\frac{\sum_i v_i \overline{a_i}\ \overline{q_i}}{\sum_j v_j \overline{a_j}} \geq \frac{\sum_i \overline{a_i}\ \overline{q_i}}{\sum_j \overline{a_j}}.$$

Proof. Let $\overline{\lambda} = \frac{\sum_i v_i \overline{a_i} \overline{q_i}}{\sum_j v_j \overline{a_j}}$ be the expected number of purchases for the quality ranking. We have

$$\sum_i v_i \overline{a_i} \left(\overline{q_i} - \overline{\lambda}\right) = 0.$$

Consider the index k such that $\left(\overline{q_k} - \overline{\lambda}\right) \geq 0$ and $\left(\overline{q_{k+1}} - \lambda\right) < 0$. Since $v_1 \geq \ldots \geq v_n$, we have

$$\sum_{i=1}^{k} v_k \overline{a_i} \left(\overline{q_i} - \overline{\lambda}\right) + \sum_{i=k+1}^{n} v_k \overline{a_i} \left(\overline{q_i} - \overline{\lambda}\right) \leq \sum_i v_i \overline{a_i} \left(\overline{q_i} - \overline{\lambda}\right) = 0$$

and, given the fact that $v_k \geq 0$,

$$\sum_{i=1}^{n} \overline{a_i} \left(\overline{q_i} - \overline{\lambda} \right) \leq 0.$$

We had the desired result: $\lambda \geq \frac{\sum_{i=1}^{n} \overline{a_i q_i}}{\sum_{i=1}^{n} \overline{a_i}}$. □

Social Influence The last result in this section shows that the social influence signals always benefit trial-offer markets with continuation. The result is independent of the structure of the continuation probabilities. The proof is a generalization of the result in Van Hentenryck et al. [2016].

Theorem 14. *The expected marginal rate of purchases is non-decreasing over time for the quality ranking under social influence in trial-offer markets with continuation.*

Proof. Let

$$\mathbb{E}[D_t] = \frac{\sum_i v_i \overline{a_i q_i}}{\sum_i v_i \overline{a_i}} = \overline{\lambda} \tag{A.9}$$

be the expected number of purchases at time t. The expected number of purchases at time $t+1$ conditional to time t is

$$\begin{aligned}
\mathbb{E}[D_{t+1}] &= \sum_j \left[\frac{v_j \overline{a_j q_j}}{\sum v_i \overline{a_i}} \cdot \frac{\sum_{i \neq j} v_i \overline{a_i q_i} + v_j (\overline{a_j} + 1 - c_j) \overline{q_j}}{\sum_{i \neq j} v_i \overline{a_i} + v_j (\overline{a_j} + 1 - c_j)} \right] \\
&+ \left[1 - \frac{\sum_i v_i \overline{a_i q_i}}{\sum_i v_i \overline{a_i}} \right] \cdot \frac{\sum_i v_i \overline{a_i q_i}}{\sum_i v_i \overline{a_i}} \\
&= \sum_j \left[\frac{v_j \overline{a_j q_j}}{\sum v_i \overline{a_i}} \cdot \frac{\sum_i v_i \overline{a_i q_i} + v_j (1 - c_j) \overline{q_j}}{\sum_i v_i \overline{a_i} + v_j (1 - c_j)} \right] \\
&+ \left[1 - \frac{\sum_j v_j \overline{a_j q_j}}{\sum_i v_i \overline{a_i}} \right] \cdot \overline{\lambda}
\end{aligned}$$

We need to prove that

$$\mathbb{E}[D_{t+1}] \geq \mathbb{E}[D_t], \tag{A.10}$$

which is equivalent to show, using Equation A.9, that

$$\sum_j \left[\frac{v_j \overline{a_j q_j}}{\sum v_i \overline{a_i}} \cdot \frac{\sum_i v_i \overline{a_i q_i} + v_j (1 - c_j) \overline{q_j}}{\sum_i v_i \overline{a_i} + v_j (1 - c_j)} \right] + [1 - \overline{\lambda}] \cdot \overline{\lambda} \geq \overline{\lambda}$$

Rearranging the terms, the proof obligation becomes

$$\frac{1}{\sum_i v_i \overline{a_i}} \sum_j \left[\frac{v_j^2 \overline{a_j q_j} (1 - c_j)}{\sum_i v_i \overline{a_i} + v_j (1 - c_j)} \left(\overline{q_j} - \overline{\lambda} \right) \right] \geq 0$$

or, equivalently,

$$\sum_j \left[\frac{v_j^2 \overline{a_j} \overline{q_j}(1-c_j)}{\sum_i v_i \overline{a_i} + v_j(1-c_j)} \left(\overline{q_j} - \overline{\lambda}\right) \right] \geq 0. \tag{A.11}$$

Let $k = \max\{i \in N | (\overline{q_i} - \lambda) \geq 0\}$, i.e., the largest $k \in N$ such that $q_k \geq \lambda$. By separating the sum into positive and negative terms, we obtain

$$\begin{aligned}
&\sum_j \left[\frac{v_j^2 \overline{a_j} \overline{q_j}(1-c_j) \left(\overline{q_j} - \overline{\lambda}\right)}{\sum_i v_i \overline{a_i} + v_j(1-c_j)} \right] = S^+ + S^- \quad \text{where} \\
&S^+ = \sum_{j=1}^{k} \left[\frac{v_j \overline{q_j}(1-c_j)}{\sum_i v_i \overline{a_i} + v_j(1-c_j)} \overline{a_j} v_j (\overline{q_j} - \overline{\lambda}) \right], \\
&S^- = \sum_{j=k+1}^{n} \left[\frac{v_j \overline{q_j}(1-c_j)}{\sum_i v_i \overline{a_i} + v_j(1-c_j)} \overline{a_j} v_j (q_j - \overline{\lambda}) \right]
\end{aligned}$$

By definition of k, all the terms in S^+ are positive and the terms in S^- are negative. Now, by definition of k and $\overline{q_i} = \frac{q_i}{(1-c_i)}$, we have

$$\begin{aligned}
&\forall i \leq k: \ (1-c_i) \leq \frac{q_i}{\overline{\lambda}}, \\
&\forall i > k: \ (1-c_i) \geq \frac{q_i}{\overline{\lambda}}.
\end{aligned} \tag{A.12}$$

We now compute a lower bound for S^+ and S^-. For S^+, using Equation A.12 for $j \leq k$, we have

$$\begin{aligned}
&\frac{v_j \overline{q_j}(1-c_j)}{\sum_i v_i \overline{a_i} + v_j(1-c_j)} \leq \frac{v_j q_j}{\sum_i v_i \overline{a_i} + v_j \frac{q_j}{\overline{\lambda}}} \\
&\leq \overline{\lambda} \cdot \frac{v_j q_j}{\overline{\lambda} \sum_i v_i \overline{a_i} + v_j q_j} \leq \overline{\lambda} \cdot \frac{v_k q_k}{\overline{\lambda} \sum_i v_i \overline{a_i} + v_k q_k}
\end{aligned} \tag{A.13}$$

The last inequality follows by $v_i \geq v_k$ and $q_i \geq q_k$ (using Theorem 11) and the following property: For all $c > 0$ and $x \geq y \geq 0$,

$$\frac{x}{c+x} \geq \frac{y}{c+y} \Leftrightarrow (c+y)x \geq (c+x)y \Leftrightarrow cx \geq cy \Leftrightarrow x \geq y.$$

For S^-, consider the following expression for $j > k$:

$$\overline{\lambda}\left(\sum_i v_i \overline{a_i}\right) \underbrace{[v_j q_j - v_k q_k]}_{\geq 0} + v_k q_k \underbrace{[v_j q_j - \overline{\lambda} v_j (1-c_j)]}_{\geq 0}$$

Here the first term is greater or equal than zero because $v_j \geq v_k$ and $q_j \geq q_k$ using Theorem 11 again. The second term is also greater than zero because it can be lower-bounded (using Equation A.12) by:

$$v_j q_j - \overline{\lambda} \frac{v_j q_j}{\overline{\lambda}} = 0.$$

Hence,

$$\begin{aligned}
&\overline{\lambda}(\sum_i v_i \overline{a_i})[v_j q_j - v_k q_k] + v_k q_k [v_j q_j - \overline{\lambda} v_j (1-c_j)] \geq 0 \\
&v_j q_j [\overline{\lambda}(\sum_i v_i \overline{a_i}) + v_k q_k] \geq \overline{\lambda} v_k q_k [\sum_i v_i \overline{a_i} + v_j(1-c_j)] \\
\Leftrightarrow\ & \frac{v_j q_j}{\sum_i v_i \overline{a_i} + v_j(1-c_j)} \geq \frac{\overline{\lambda} v_k q_k}{\overline{\lambda} \sum_i v_i \overline{a_i} + v_k q_k}
\end{aligned} \tag{A.14}$$

Putting together Equations A.13 and A.14 gives us a lower bound to $S^+ + S^-$:

$$S^+ + S^- = \frac{\overline{\lambda} v_k q_k}{\overline{\lambda} \sum_i v_i \overline{a_i} + v_k q_k} \cdot \sum_{i=1}^{n} v_i \overline{a_i} (\overline{q_i} - \overline{\lambda}) \tag{A.15}$$

Now, by definition of $\overline{\lambda}$,

$$\overline{\lambda} = \frac{\sum_{i=1}^{n} v_i \overline{a_i q_i}}{\sum_{i=1}^{n} v_i \overline{a_i}} \Leftrightarrow \sum_{i=1}^{n} v_i \overline{a_i} (\overline{q_i} - \overline{\lambda}) = 0.$$

which implies that

$$\frac{\overline{\lambda} v_k q_k}{\overline{\lambda} \sum_i v_i \overline{a_i} + v_k q_k} \cdot \sum_{i=1}^{n} v_i \overline{a_i} (\overline{q_i} - \overline{\lambda}) = 0$$

concluding the proof. □

Relationship with the Cascade Model: Observe that the quality ranking over the continuation quality orders the products in decreasing order of $\frac{q_i}{1-c_i}$ value, which is exactly the adjusted ecpm from Aggarwal et al. [2008]; Kempe and Mahdian [2008b] with all the revenues set to 1. Obviously, the quality ranking (when the continuation probabilities preserve the quality rank in the continuation model), and hence the adjusted ecpm ranking, are not the best rankings to show to an incoming participant (the performance ranking is), but our results show that they have nice asymptotic properties.

A.5 Experimental Results

This section report computational results to highlight the theoretical analysis. The computational results use settings that model the MusicLab experiments discussed in Salganik et al. [2006]; Krumme et al. [2012]; Abeliuk et al. [2015]. As mentioned in the introduction, MusicLab is a trial-offer market where participants can try a song and then decide to download it. The experiments use an agent-based simulation to emulate MusicLab. Each simulation consists of N steps and, at each iteration t,

1. we simulate selecting a song i according to the probabilities $p_i(\sigma, d)$, where σ is the ranking proposed by the policy under evaluation and d is the social influence signal.

2. with probability q_i, the sampled song is downloaded, in which case the simulator increases the social influence signal for song i, i.e., $d_{i,t+1} = d_{i,t} + 1$. Otherwise, $d_{i,t+1} = d_{i,t}$, and if the continuation model is used, the simulation goes back to Step 1 with probability c_i and advances to the next step otherwise.

Every T iterations, a new list σ is computed using one of the ranking policies. The experimental setting, which aims at being close to the MusicLab experiments, considers 50 songs and simulations with 20,000 steps. The songs are displayed in a single column. The analysis in Krumme et al. [2012] indicated that participants are more likely to try songs higher in the list. More precisely, the visibility decreases with the list position, except for a slight increase at the bottom positions. We use the first setting for qualities, appeals and visibilities from Abeliuk et al. [2015], where the quality and the appeal are chosen independently according to a Gaussian distribution normalized to fit between 0 and 1. In addition, the experiments consider 12 different continuation probabilities, varying ρ and the power r as shown in Figure A.3. The results were obtained by averaging $W = 100$ simulations.

Parameters	P-rank	Q-rank	D-rank	R-rank
$\rho = 0.1, r = 0$	5.3%	4.9%	5.8%	7.5%
$\rho = 0.1, r = 0.25$	4.1%	4.3%	4.9%	5.2%
$\rho = 0.1, r = 1$	2.2%	2.4%	2.6%	2.5%
$\rho = 0.1, r = 2$	1.4%	1.4%	0.2%	1%
$\rho = 0.5, r = 0$	30.6%	31%	38.3%	51.8%
$\rho = 0.5, r = 0.25$	24.2%	24.6%	28.4%	33.6%
$\rho = 0.5, r = 1$	13.4%	13.2%	14.8%	12.2%
$\rho = 0.5, r = 2$	7.2%	7.3%	6.2%	4.6%
$\rho = 0.9, r = 0$	67.3%	67.7%	93.9%	143.3%
$\rho = 0.9, r = 0.25$	51.6%	52.1%	65.2%	79.9%
$\rho = 0.9, r = 1$	26.6%	26.8%	28.5%	24.2%
$\rho = 0.9, r = 2$	13.6%	13.7%	12.2%	8.3%

Table A.1: Improvement in market efficiency (in percentage) for the continuation model

Parameters	P-RANK	Q-RANK	D-RANK	R-RANK
$\rho = 0.5, r = 0.25$	13776.1	13804.1	12000.1	9393.8
$\rho = 0.5, r = 1$	12579.0	12565.5	10643.7	7885.0
$\rho = 0.9, r = 0.25$	16784.7	16840.9	15435.7	12680.8
$\rho = 0.9, r = 1$	14041.4	14059.1	11926.5	8741.7

Table A.2: Market Efficiency in the continuation model.

Table A.1 presents results on market efficiency (i.e., the number of downloads) for the trial and offer market with continuation. The most interesting message from these results, is the observation that the popularity and random rankings improve more than the performance and quality rankings, unless the quality has less impact ($r = 2$, because the continuation is decreasing in r) in the continuation. This can be explained by the fact that the continuation provides a way to correct a potentially weak ranking. However, as indicated in Table A.2, this correction is not enough to bridge the gap with the performance and quality rankings.

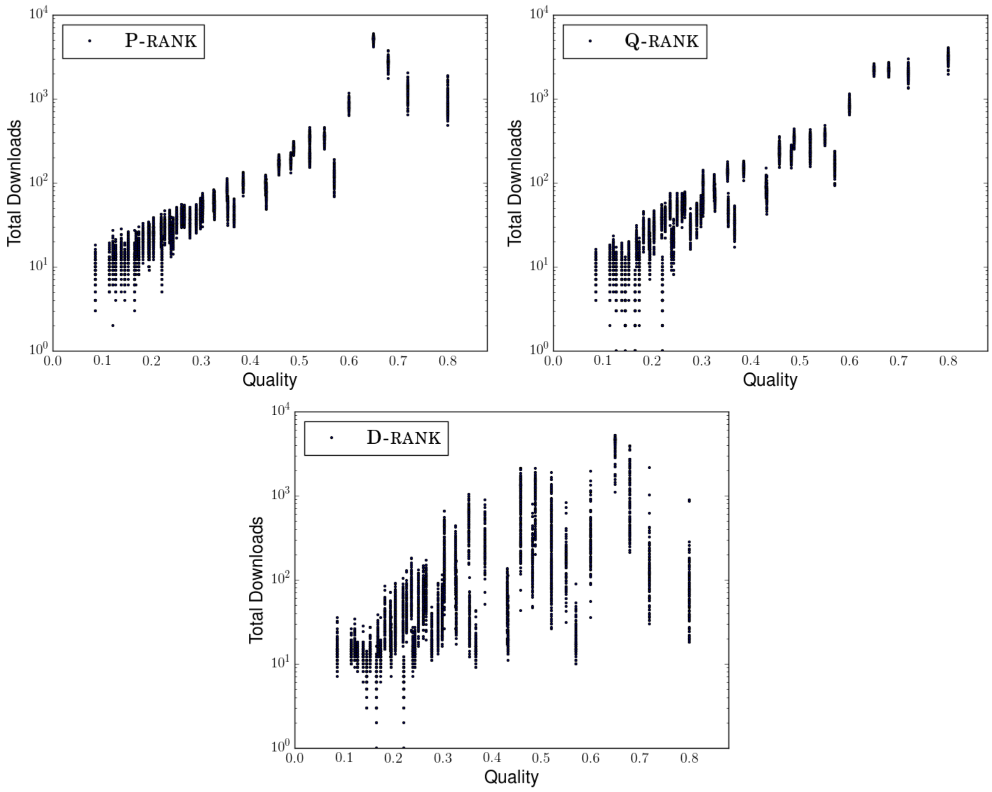


Figure A.4: The Distribution of Downloads Versus Song Qualities for $\rho = 0.9$, $r = 1$. The songs on the x-axis are ranked by increasing quality from left to right. Each dot is the number of download of a product in one of the 100 experiments.

Figure A.4 depicts experimental results on the predictability of the market under the continuation model under various ranking policies. The figure plots the number of downloads of each song for 100 experiments. In the plots, the songs are ranked by increasing quality from left to right on the x-axis. Each dot in the plot shows the number of downloads of a song in one of the 100 experiments. The results are essentially unchanged when moving from the traditional to a continuation multinomial logit model. The popularity ranking still exhibits significantly more unpredictability than the performance and quality rankings and the continuations are not able to compensate for the inherent unpredictability.

A.6 Conclusion and Final Remarks

Motivated by applications in online markets, this chapter generalises the ubiquitous multinomial logit model to a setting that allows market participants to sample multiple products before deciding whether to purchase or not. We showed that trial-offer markets with continuation can be reduced to the original trial-offer model, transferring many fundamental properties of ranking policies to a more general setting. In particular, the quality ranking still benefits from position bias and social influence. Moreover, under a general class of continuation functions, the quality ranking is also preserved and the market reaches the same asymptotic equilibrium. Experimental results shows that the continuation model compensates for some of the weaknesses of the popularity ranking by boosting its market performance more than the quality and performance ranking, unless the continuation probability depends too strongly on quality. A potential line of future research, is to generalise these results further to hierarchical trial-offer markets, or analyse the effect of other continuation functional forms.

Missing Proofs

This Appendix contains the missing proofs from the main text.

Lemma 6. *Fixed-Price policy can be arbitrarily bad for the Joint Assortment and Pricing problem under the Threshold Luce model.*

Proof. Consider $N+1$ products, with product one having $u > 0$ utility and $a_0 = 1$. For all the remaining N products let their utility to be: αu, with $\alpha < 1$ such that in presence of product one, all the rest of the products are ignored for threshold t. The optimal revenue if we consider a fixed price strategy is [Li and Huh, 2011; Wang, 2012]:

$$R' = W(\exp(u-1))$$

Because no matter what fixed price we select, the N lower utility products are completely ignored and the first product is the only one contributing to the revenue, and this is the best revenue that we can achieve given that. Now, let us consider the optimal revenue obtained with the strategy described in Theorem (10)

$$R^* = W\left(\left[\frac{(1+t)+N}{1+t}\right]\cdot\exp\left(\frac{(1+t)(u-1)+N(\alpha u-1)+N\ln(1+t)}{(1+t)+N}\right)\right), \quad \text{(B.1)}$$

let us find an explicit relation between R' and R^*. Starting from equation (B.1):

$$\begin{aligned}
R^* =& W\left(\left[\frac{(1+t)+N}{1+t}\right]\cdot\exp\left(\frac{(1+t)(u-1)+N(\alpha u-1)+N\ln(1+t)}{(1+t)+N}\right)\right)\\
R^* =& W\left(\left[\frac{(1+t)+N}{1+t}\right]\cdot\exp\left((u-1)\cdot\frac{(1+t)+N\alpha}{(1+t)+N}+\frac{N\ln(1+t)}{(1+t)+N}\right)\right)\\
R^* =& W\left(\left[\frac{(1+t)+N}{1+t}\right]\cdot\exp\left((u-1)+\frac{N}{(1+t)+N}\cdot\underbrace{(\ln(1+t)-(u-1)\cdot(1-\alpha))}_{\Gamma}\right)\right)
\end{aligned}$$

We know that the Lambert function is concave, increasing and unbounded [Corless et al., 1996; Li and Huh, 2011]. With this in mind, let u be such that Γ is greater or equal than zero (for example, setting $u = 1.9$, $\alpha = 0.5$ and $t = 0.5$, makes $\Gamma > 0$ and product 1 dominates the rest of the products), this is:

$$\frac{\ln(1+t)}{1-\alpha} + 1 \geq u. \tag{B.2}$$

Using this, we have:

$$R^* \geq W\left(\left[\frac{(1+t)+N}{1+t}\right] \cdot \exp\left(u-1\right)\right) \tag{B.3}$$

Where the argument of the Lambert function is exactly the same as R^*, but multiplied by a constant factor larger than one and depending on N. Putting everything together, we have:

$$R^* \geq W\left(\left[\frac{(1+t)+N}{1+t}\right] \cdot \exp\left(u-1\right)\right) \geq R' \tag{B.4}$$

The expression in the middle can be arbitrarily larger than R' by letting N tend to infinity, and so is R^*. Thus, the fixed price policy can be arbitrarily bad under the TLM. $\square$

Conditions where same price policy is optimal: Is a known result that the same price policy is optimal when consumers follows the Multinomial Logit with equal price sensitivities. For the case of differentiated price sensitivities, the adjusted mark-up $p_i - c_i - \frac{1}{\beta_i}$ is constant across products, where p_i, c_i and β_i are the price, cost and price sensitivity of product i respectively. As we showed in Section4, in some cases, the same price policy is optimal. Can we characterize when this happens? The answer is yes, and we formalize it in the following proposition.

Proposition 13. *Same price policy is optimal, for the Threshold Luce Model with equal price sensitivities if and only if:*

$$\max_{i \in X} u_i - \min_{i \in X} u_i \leq \ln(1+t) \iff \max_{i,j \in X}\left(u_i - u_j\right) \leq \ln(1+t) \tag{B.5}$$

The proof relies on the fact that the ratio between attractiveness for equal price sensitivities, only depends on the intrinsic utilities.

Proof. For equal price sensitivities, if we have same price p for all products, then the Threshold ratio for any two products can be written as:

$$\begin{aligned}
\frac{a_i(p)}{a_j(p)} &\leq 1+t \\
\frac{\exp(u_i - p)}{\exp(u_j - p)} &\leq 1+t \\
\exp(u_i - p - (u_j - p_j)) &\leq 1+t \\
u_i - u_j &\leq \ln(1+t)
\end{aligned}$$

This last equation holds for all $i, j \in X$ if and only if we have $\max_{i,j \in X} \left(u_i - u_j\right) \leq \ln(1+t)$, which is the condition expressed on the proposition. $\square$

Bibliography

Abeliuk, A.; Berbeglia, G.; Cebrian, M.; and Van Hentenryck, P., 2015. The benefits of social influence in optimized cultural markets. *PLoS ONE*, 10, 4 (04 2015), e0121934. (cited on pages 14, 93, 103, and 104)

Abeliuk, A.; Berbeglia, G.; Cebrian, M.; and Van Hentenryck, P., 2016. Assortment optimization under a multinomial logit model with position bias and social influence. *4OR*, 14, 1 (2016), 57–75. (cited on pages 14, 15, 16, 17, and 64)

Abeliuk, A.; Berbeglia, G.; Van Hentenryck, P.; Hogg, T.; and Lerman, K., 2017. Taming the unpredictability of cultural markets with social influence. In *Proceedings of the 26th International Conference on World Wide Web*, 745–754. International World Wide Web Conferences Steering Committee. (cited on pages 14 and 15)

Aggarwal, G.; Feldman, J.; Muthukrishnan, S.; and Pal, M., 2008. Sponsored search auctions with markovian users. *CoRR*, abs/0805.0766 (2008). (cited on page 103)

Ahumada, A. and Ülkü, L., 2018. Luce rule with limited consideration. *Mathematical Social Sciences*, 93 (2018), 52 – 56. (cited on pages 11 and 44)

Alptekinoğlu, A. and Semple, J. H., 2016. The exponomial choice model: A new alternative for assortment and price optimization. *Operations Research*, 64, 1 (2016), 79–93. (cited on pages 6, 17, and 18)

Anderson, S. P.; Depalma, A.; and Thisse, J.-F., 1992. *Discrete Choice Theory of Product Differentiation*. The MIT Press. ISBN 026201128X. (cited on pages 3, 5, and 18)

Aouad, A.; Farias, V. F.; and Levi, R., 2015. Assortment optimization under consider-then-choose choice models. *Available at SSRN 2618823*, (2015). (cited on pages 11 and 17)

Aouad, A. and Segev, D., 2015. Display optimization for vertically differentiated locations under multinomial logit choice preferences. *Available at SSRN 2709652*, (2015). (cited on page 17)

Bandyapadhyay, S., 2014. A variant of the maximum weight independent set problem. *CoRR*, abs/1409.0173 (2014). (cited on pages 53 and 55)

Belonax Jr, J. A. and Mittelstaedt, R. A., 1978. Evoked set size as a function of number of choice criteria and information variability. *Advances in consumer research*, 5, 1 (1978). (cited on page 10)

Ben-Akiva, M. and Lerman, S., 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press series in transportation studies. MIT Press. ISBN 9780262022170. (cited on page 5)

Berbeglia, G., 2016. Discrete choice models based on random walks. *Operations Research Letters*, 44, 2 (2016), 234–237. (cited on pages 6, 47, and 62)

Berbeglia, G., 2018. The generalized stochastic preference choice model. *arXiv preprint arXiv:1803.04244*, (2018). (cited on page 11)

Berbeglia, G.; Garassino, A.; and Vulcano, G., 2018. A comparative empirical study of discrete choice models in retail operations. *Available at SSRN 3136816*, (2018). (cited on page 4)

Berbeglia, G. and Joret, G., 2017. Assortment optimisation under a general discrete choice model: A tight analysis of revenue-ordered assortments. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17 (Cambridge, Massachusetts, USA, 2017), 345–346. ACM, New York, NY, USA. (cited on pages 8 and 17)

Berry, S.; Levinsohn, J.; and Pakes, A., 1995. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, (1995), 841–890. (cited on page 18)

Besbes, O. and Sauré, D., 2016. Product assortment and price competition under multinomial logit demand. *Production and Operations Management*, 25, 1 (2016), 114–127. (cited on pages 3 and 18)

Blanchet, J.; Gallego, G.; and Goyal, V., 2016. A markov chain approximation to choice modeling. *Operations Research*, 64, 4 (2016), 886–905. (cited on pages 6, 47, and 62)

Block, H. D. and Marschak, J., 1960. Random orderings and stochastic theories of responses. *Contributions to probability and statistics*, 2 (1960), 97–132. (cited on pages 2, 5, and 7)

Bodea, T. and Ferguson, M., 2014. *Segmentation, revenue management and pricing analytics*. Routledge. (cited on pages 2 and 4)

Börsch-Supan, A., 1990. On the compatibility of nested logit models with utility maximization. *Journal of Econometrics*, 43, 3 (1990), 373–388. (cited on page 18)

Brady, R. L. and Rehbeck, J., 2016. Menu-dependent stochastic feasibility. *Econometrica*, 84, 3 (2016), 1203–1223. (cited on pages 10, 11, and 12)

Broder, A.; Kumar, R.; Maghoul, F.; Raghavan, P.; Rajagopalan, S.; Stata, R.; Tomkins, A.; and Wiener, J., 2000. Graph structure in the web. *Comput. Netw.*, 33, 1-6 (Jun. 2000), 309–320. (cited on page 13)

Bront, J. J. M.; Méndez-Díaz, I.; and Vulcano, G., 2009. A column generation algorithm for choice-based network revenue management. *Operations Research*, 57, 3 (2009), 769–784. (cited on page 16)

Camion, P., 1965. Characterization of totally unimodular matrices. *Proceedings of the American Mathematical Society*, 16, 5 (1965), 1068–1073. (cited on page 52)

Catalano, M.; Lo Casto, B.; and Migliore, M., 2008. Car sharing demand estimation and urban transport demand modelling using stated preference techniques. (2008). (cited on page 5)

Cattaneo, M. D.; Ma, X.; Masatlioglu, Y.; and Suleymanov, E., 2017. A random attention model. *arXiv preprint arXiv:1712.03448*, (2017). (cited on pages 11 and 12)

Caves, R. E., 2000. *Creative industries: Contracts between art and commerce*. 20. Harvard University Press. (cited on page 13)

Cha, M.; Kwak, H.; Rodriguez, P.; Ahn, Y.-Y.; and Moon, S., 2007. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 1–14. ACM. (cited on page 13)

Chen, N.; Gallego, G.; and Tang, Z., 2019. The use of binary choice forests to model and estimate discrete choice models. *ArXiv*, abs/1908.01109 (2019). (cited on pages 12 and 13)

Chen, Y.-C. and Mišić, V. V., 2019. Decision forest: A nonparametric approach to modeling irrational choice. *arXiv preprint arXiv:1904.11532*, (2019). (cited on pages 12 and 13)

Chernev, A., 2003a. Product assortment and individual decision processes. *Journal of Personality and Social Psychology*, 85, 1 (2003), 151. (cited on page 9)

Chernev, A., 2003b. When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *Journal of consumer Research*, 30, 2 (2003), 170–183. (cited on page 9)

Chernev, A.; Böckenholt, U.; and Goodman, J., 2015. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25, 2 (2015), 333–358. (cited on pages 7, 9, 22, and 25)

Chiang, W.-C.; Chen, J. C.; Xu, X.; et al., 2007. An overview of research on revenue management: current issues and future research. *International journal of revenue management*, 1, 1 (2007), 97–128. (cited on pages 2 and 4)

Corless, R. M.; Gonnet, G. H.; Hare, D. E. G.; Jeffrey, D. J.; and Knuth, D. E., 1996. On the lambertw function. *Advances in Computational Mathematics*, 5, 1 (Dec 1996), 329–359. (cited on pages 81 and 107)

Craswell, N.; Zoeter, O.; Taylor, M.; and Ramsey, B., 2008. An experimental comparison of click position-bias models. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*, WSDM '08 (Palo Alto, California, USA, 2008), 87–94. ACM, New York, NY, USA. (cited on page 14)

Daganzo, C., 1979. *Multinomial probit: the theory and its application to demand forecasting*. Academic Press, New York. (cited on pages 6 and 17)

Dai, J.; Ding, W.; Kleywegt, A.; Wang, X.; and Zhang, Y., 2014. Choice based revenue management for parallel flights. *Available at SSRN 2404193*, (2014). (cited on page 11)

Davis, J.; Gallego, G.; and Topaloglu, H., 2013. Assortment planning under the multinomial logit model with totally unimodular constraint structures. *Department of IEOR, Columbia University.*, (2013). (cited on pages 15, 16, 17, 52, 64, and 89)

Davis, J. M.; Gallego, G.; and Topaloglu, H., 2014. Assortment optimization under variants of the nested logit model. *Operations Research*, 62, 2 (2014), 250–273. (cited on page 16)

De los Santos, B.; Hortaçsu, A.; and Wildenbeest, M. R., 2012. Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, 102, 6 (2012), 2955–80. (cited on page 10)

De Vany, A. and Walls, W. D., 1999. Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of cultural economics*, 23, 4 (1999), 285–318. (cited on page 13)

Debreu, G., 1960. Review of R. D. Luce Individual Choice Behavior. *American Economic Review*, 50 (1960), 186–188. (cited on pages 5, 7, 8, and 22)

Dong, L.; Kouvelis, P.; and Tian, Z., 2009. Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11, 2 (2009), 317–339. (cited on page 17)

Doyle, J. R.; O'Connor, D. J.; Reynolds, G. M.; and Bottomley, P. A., 1999. The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases. *Psychology and Marketing*, 16, 3 (1999), 225–243. (cited on pages 7, 8, and 22)

Dutta, R., 2018. Gradual Pairwise Comparison and Stochastic Choice. Technical report. (cited on page 12)

Echenique, F. and Saito, K., 2018. General Luce model. *Economic Theory*, (Sep 2018), 1–16. (cited on pages xi, 3, 9, 11, 43, 44, 45, 46, and 87)

Echenique, F.; Saito, K.; and Tserenjigmid, G., 2018. The perception-adjusted Luce model. *Mathematical Social Sciences*, 93 (2018), 67 – 76. (cited on pages xi, 3, 9, 17, 21, 22, 23, 24, 25, and 87)

Erdem, T.; Swait, J.; and Louviere, J., 2002. The impact of brand credibility on consumer price sensitivity. *International journal of Research in Marketing*, 19, 1 (2002), 1–19. (cited on page 18)

Even, S., 1979. *Graph Algorithms*. W. H. Freeman & Co., New York, NY, USA. ISBN 0716780445. (cited on page 51)

Farias, V. F.; Jagabathula, S.; and Shah, D., 2013. A nonparametric approach to modeling choice with limited data. *Management Science*, 59, 2 (2013), 305–322. (cited on pages 6 and 7)

Farias, V. F.; Jagabathula, S.; and Shah, D., 2017. Building optimized and hyper-local product assortments: A nonparametric choice approach. *Available at SSRN 2905381*, (2017). (cited on page 7)

Feldman, J. and Topaloglu, H., 2015. Bounding optimal expected revenues for assortment optimization under mixtures of multinomial logits. *Production and Operations Management*, 24, 10 (2015), 1598–1620. (cited on page 16)

Feng, G.; Li, X.; and Wang, Z., 2017. On the relation between several discrete choice models. *Operations research*, 65, 6 (2017), 1516–1525. (cited on page 12)

Ford, D. R. and Fulkerson, D. R., 2010. *Flows in Networks*. Princeton University Press, Princeton, NJ, USA. ISBN 0691146675, 9780691146676. (cited on pages 50 and 51)

Ford, L. R., 1957. Solution of a ranking problem from binary comparisons. *The American Mathematical Monthly*, 64, 8 (1957), 28–33. (cited on page 5)

Gallego, G. and Li, A., 2017. Attention, consideration then selection choice model. Working paper, available at SSRN: 2926942. (cited on page 17)

Gallego, G.; Li, A.; Truong, V.-A.; and Wang, X., 2016. Approximation algorithms for product framing and pricing. Technical report, Submitted to Operations Research (2016). (cited on page 17)

Gallego, G.; Ratliff, R.; and Shebalov, S., 2015. A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research*, 63, 1 (2015), 212–232. (cited on pages 6 and 53)

Gallego, G. and Stefanescu, C., 2009. Upgrades, upsells and pricing in revenue management. *Available at SSRN 1334341*, (2009). (cited on pages 3 and 18)

Gallego, G. and Topaloglu, H., 2014. Constrained assortment optimization for the nested logit model. *Management Science*, 60, 10 (2014), 2583–2601. (cited on page 16)

Gallego, G. and Wang, R., 2014. Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62, 2 (2014), 450–461. (cited on page 18)

Garey, M. R. and Johnson, D. S., 1979. *Computers and intractability : a guide to the theory of NP-completeness*. W. H. Freeman, first edn. ISBN 0716710455. (cited on pages 16 and 64)

Gensch, D. H., 1985. Empirically testing a disaggregate choice model for segments. *Journal of Marketing Research*, 22, 4 (1985), 462–467. (cited on page 5)

Gigerenzer, G. and Goldstein, D. G., 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103, 4 (1996), 650. (cited on page 10)

Gilbride, T. J. and Allenby, G. M., 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, 23, 3 (2004), 391–406. (cited on page 10)

Greene, W. H. and Hensher, D. A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37, 8 (2003), 681 – 698. (cited on page 6)

Grötschel, M.; Lovász, L.; and Schrijver, A., 1981. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1, 2 (1981), 169–197. (cited on page 51)

Guadagni, P. M. and Little, J. D. C., 1983. A logit model of brand choice calibrated on scanner data. *Marketing Science*, 2, 3 (1983), 203–238. (cited on page 5)

Hanson, W. and Martin, K., 1996. Optimizing multinomial logit profit functions. *Management Science*, 42, 7 (1996), 992–1003. (cited on page 17)

Hauser, J. R., 1978. Testing the accuracy, usefulness, and significance of probabilistic choice models: An information-theoretic approach. *Operations Research*, 26, 3 (1978), 406–421. (cited on page 10)

Hauser, J. R.; Ding, M.; and Gaskin, S. P., 2009. Non-compensatory (and compensatory) models of consideration-set decisions. 2009 Sawtooth Software Conference Proceedings. (cited on page 17)

Hauser, J. R. and Wernerfelt, B., 1990. An evaluation cost model of consideration sets. *Journal of consumer research*, 16, 4 (1990), 393–408. (cited on page 10)

Hausman, J. and McFadden, D., 1984. Specification tests for the multinomial logit model. *Econometrica*, 52, 5 (1984), 1219–40. (cited on page 5)

Haynes, G. A., 2009. Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology and Marketing*, 26, 3 (2009), 204–212. (cited on pages 7 and 22)

Herne, K., 1997. Decoy alternatives in policy choices: Asymmetric domination and compromise effects. *European Journal of Political Economy*, 13, 3 (1997), 575–589. (cited on pages 3, 7, and 43)

Hirsch, P. M., 1972. Processing fads and fashions: An organization-set analysis of cultural industry systems. *American journal of sociology*, 77, 4 (1972), 639–659. (cited on page 13)

Hopp, W. J. and Xu, X., 2005. Product line selection and pricing with modularity in design. *Manufacturing & Service Operations Management*, 7, 3 (2005), 172–187. (cited on pages 3 and 18)

Huber, J.; Payne, J. W.; and Puto, C., 1982. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, 9, 1 (1982), 90–98. (cited on pages 2, 3, and 43)

Huh, W. T. and Li, H., 2015. Pricing under the nested attraction model with a multistage choice structure. *Operations Research*, 63, 4 (2015), 840–850. (cited on page 18)

Hummel, P. and McAfee, R. P., 2014. Position auctions with externalities and brand effects. *CoRR*, abs/1409.4687 (2014). (cited on page 64)

Iyengar, S. S. and Lepper, M. R., 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology*, 79, 6 (2000), 995. (cited on pages 7, 8, and 22)

Jagabathula, S., 2014. Assortment optimization under general choice. *Available at SSRN*, (2014), 1598–1620. (cited on page 16)

Jagabathula, S. and Rusmevichientong, P., 2017. A nonparametric joint assortment and price choice model. *Management Science*, 63, 9 (2017), 3128–3145. (cited on page 10)

Jagabathula, S. and Rusmevichientong, P., 2018. The limit of rationality in choice modeling: Formulation, computation, and implications. *Management Science*, 65, 5 (2018), 2196–2215. (cited on pages 13 and 87)

Kalra, T.; Mathew, R.; Pal, S. P.; and Pandey, V., 2017. *Maximum Weighted Independent Sets with a Budget*, 254–266. Springer International Publishing, Cham. ISBN 978-3-319-53007-9. (cited on pages 53 and 54)

Karush, W., 1939. *Minima of Functions of Several Variables with Inequalities as Side Constraints*. Master's thesis, Department of Mathematics, Univiserty of Chicago. (cited on page 77)

Kempe, D. and Mahdian, M., 2008a. A cascade model for externalities in sponsored search. In *Internet and Network Economics*, 585–596. Springer. (cited on pages 14 and 64)

Kempe, D. and Mahdian, M., 2008b. A cascade model for externalities in sponsored search. In *Proceedings of the 4th International Workshop on Internet and Network Economics*, WINE '08 (Shanghai, China, 2008), 585–596. Springer-Verlag, Berlin, Heidelberg. (cited on page 103)

Kök, A.; Fisher, M. L.; and Vaidyanathan, R., 2005. Assortment planning: Review of literature and industry practice. In *Retail Supply Chain Management*. Kluwer. (cited on pages 1, 4, and 15)

Kouvelis, P.; Xiao, Y.; and Yang, N., 2015. PBM competition in pharmaceutical supply chain: Formulary design and drug pricing. *Manufacturing & Service Operations Management*, 17, 4 (2015), 511–526. (cited on page 18)

Krumme, C.; Cebrian, M.; Pickard, G.; and Pentland, S., 2012. Quantifying social influence in an online cultural market. *PLOS ONE*, 7, 5 (05 2012), e33785. (cited on pages 14, 92, 103, and 104)

Lea, A. M. and Ryan, M. J., 2015. Irrationality in mate choice revealed by túngara frogs. *Science*, 349, 6251 (2015), 964–966. (cited on page 8)

Lerman, K. and Hogg, T., 2014. Leveraging position bias to improve peer recommendation. *PLOS ONE*, 9, 6 (06 2014). (cited on page 14)

Li, G.; Rusmevichientong, P.; ; and Topaloglu, H., 2015. The d-level nested logit model: Assortment and price optimization problems. *Operations Research*, 63, 2 (2015), 325–342. (cited on page 16)

Li, H. and Huh, W. T., 2011. Pricing Multiple Products with the Multinomial Logit and Nested Logit Models: Concavity and Implications. *Manufacturing & Service Operations Management*, 13, 4 (October 2011), 549–563. (cited on pages 18, 70, and 107)

Lim, A.; Rodrigues, B.; and Zhang, X., 2004. Metaheuristics with local search techniques for retail shelf-space optimization. *Management science*, 50, 1 (2004), 117–131. (cited on page 14)

Liu, N.; Ma, Y.; and Topaloglu, H., 2018. Assortment optimization under the multinomial logit model with sequential offerings. Technical report. (cited on pages 22 and 89)

Lorenz, J.; Rauhut, H.; Schweitzer, F.; and Helbing, D., 2011. How social influence can undermine the wisdom of crowd effect. *Proceedings of the national academy of sciences*, 108, 22 (2011), 9020–9025. (cited on page 13)

Luce, R. D., 1959. *Individual Choice Behavior*. John Wiley and Sons. (cited on pages 4, 5, and 92)

Maldonado, F.; Van Hentenryck, P.; Berbeglia, G.; and Berbeglia, F., 2018. Popularity signals in trial-offer markets with social influence and position bias. *European Journal of Operational Research*, 266, 2 (2018), 775–793. (cited on pages 14 and 15)

Manzini, P. and Mariotti, M., 2014. Stochastic choice and consideration sets. *Econometrica*, 82, 3 (2014), 1153–1176. (cited on pages 10, 12, and 17)

Masatlioglu, Y.; Nakajima, D.; and Ozbay, E. Y., 2012. Revealed attention. *American Economic Review*, 102, 5 (2012), 2183–2205. (cited on page 10)

McCausland, W. J. and Marley, A., 2013. Prior distributions for random choice structures. *Journal of Mathematical Psychology*, 57, 3–4 (2013), 78 – 93. (cited on page 8)

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, (1974), 105–142. (cited on page 5)

McFadden, D., 1978. Modeling the choice of residential location. *Transportation Research Record*, , 673 (1978). (cited on pages 2, 4, and 5)

McFadden, D. and Train, K., 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15, 5 (2000), 447–470. (cited on page 6)

McGill, J. I. and van Ryzin, G. J., 1999. Revenue management: Research overview and prospects. *Transportation Science*, 33, 2 (1999), 233–256. (cited on pages 1 and 4)

Megiddo, N., 1979. Combinatorial optimization with rational objective functions. *Mathematics of Operations Research*, 4, 4 (1979), 414–424. (cited on pages 43, 51, and 52)

Méndez-Díaz, I.; Miranda-Bront, J. J.; Vulcano, G.; and Zabala, P., 2014. A branch-and-cut algorithm for the latent-class logit assortment problem. *Discrete Applied Mathematics*, 164 (2014), 246–263. (cited on page 16)

Mirsky, L., 1971. A dual of Dilworth's decomposition theorem. *The American Mathematical Monthly*, 78, 8 (1971), 876–877. (cited on page 51)

Möhring, R. H., 1985. *Graphs and Order: The Role of Graphs in the Theory of Ordered Sets and Its Applications*, chap. Algorithmic Aspects of Comparability Graphs and Interval Graphs, 41–101. Springer Netherlands, Dordrecht. ISBN 978-94-009-5315-4. (cited on pages 43, 50, and 51)

Natarajan, K.; Song, M.; and Teo, C.-P., 2009. Persistency model and its applications in choice modeling. *Management Science*, 55, 3 (03 2009), 453–469. Copyright - Copyright Institute for Operations Research and the Management Sciences Mar 2009; Document feature - Equations; Graphs; Tables; ; Last updated - 2016-04-02; CODEN - MNSCDI. (cited on page 12)

Natenzon, P., 2019. Random choice and learning. *Journal of Political Economy*, 127, 1 (2019), 419–457. (cited on page 8)

Negahban, S.; Oh, S.; and Shah, D., 2012. Iterative ranking from pair-wise comparisons. In *Advances in Neural Information Processing Systems*, 2474–2482. (cited on page 5)

NEWMAN, J. P.; FERGUSON, M. E.; GARROW, L. A.; AND JACOBS, T. L., 2014. Estimation of choice-based models using sales data from a single firm. *Manufacturing & Service Operations Management*, 16, 2 (2014), 184–197. (cited on page 5)

NEWMAN, M. E., 2005. Power laws, pareto distributions and Zipf's law. *Contemporary physics*, 46, 5 (2005), 323–351. (cited on page 13)

PAYNE, J. W., 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational behavior and human performance*, 16, 2 (1976), 366–387. (cited on page 10)

PAYNE, J. W.; BETTMAN, J. R.; AND LUCE, M. F., 1996. When time is money: Decision behavior under opportunity-cost time pressure. *Organizational behavior and human decision processes*, 66, 2 (1996), 131–152. (cited on page 10)

PAYNE, S. L., 1951. The art of asking questions. (1951). (cited on page 14)

PRAS, B. AND SUMMERS, J., 1975. A comparison of linear and nonlinear evaluation process models. *Journal of Marketing Research*, (1975), 276–281. (cited on page 10)

RATLIFF, R. M.; VENKATESHWARA RAO, B.; NARAYAN, C. P.; AND YELLEPEDDI, K., 2008. A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings. *Journal of Revenue and Pricing Management*, 7, 2 (2008), 153–171. (cited on page 5)

RAYFIELD, W. Z.; RUSMEVICHIENTONG, P.; AND TOPALOGLU, H., 2015. Approximation methods for pricing problems under the nested logit model with price bounds. *INFORMS Journal on Computing*, 27, 2 (2015), 335–357. (cited on page 18)

ROBERTS, J. H. AND LATTIN, J. M., 1991. Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 28, 4 (1991), 429–440. (cited on page 10)

RUSMEVICHIENTONG, P.; SHEN, Z.-J. M.; AND SHMOYS, D. B., 2010. Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations research*, 58, 6 (2010), 1666–1680. (cited on pages 5 and 15)

RUSMEVICHIENTONG, P.; SHMOYS, D.; TONG, C.; AND TOPALOGLU, H., 2014. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23, 11 (2014), 2023–2039. (cited on page 16)

RUSMEVICHIENTONG, P. AND TOPALOGLU, H., 2012. Robust assortment optimization in revenue management under the multinomial logit choice model. *Operations Research*, 60, 4 (2012), 865–882. (cited on page 16)

RUSSO, J. E. AND ROSEN, L. D., 1975. An eye fixation analysis of multialternative choice. *Memory & Cognition*, 3, 3 (1975), 267–276. (cited on page 12)

SALGANIK, M. J.; DODDS, P. S.; AND WATTS, D. J., 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, 5762 (2006), 854–856. (cited on pages 93 and 103)

SALGANIK, M. J. AND WATTS, D. J., 2008. Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market. *Social Psychology Quarterly*, 71, 4 (2008), 338–355. (cited on page 14)

SCHWARTZ, B., 2004. *Paradox of Choice: Why More is Less*. HarperCollins. (cited on pages 7, 9, 22, and 25)

SETHI-IYENGAR, S.; HUBERMAN, G.; AND JIANG, W., 2004. How much choice is too much? contributions to 401 (k) retirement plans. *Pension design and structure: New lessons from behavioral finance*, 83 (2004), 84–87. (cited on page 9)

SHAH, A. M. AND WOLFORD, G., 2007. Buying behavior as a function of parametric variation of number of choices. *Psychological science*, 18, 5 (2007), 369. (cited on page 9)

SHUM, H. AND TROTTER, L., 1996. Cardinality-restricted chains and antichains in partially ordered sets. *Discrete Applied Mathematics*, 65, 1 (1996), 421 – 439. First International Colloquium on Graphs and Optimization. (cited on page 55)

SIMONSON, I. AND TVERSKY, A., 1992. Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, 29, 3 (Aug 01 1992), 281. Last update - 2013-02-23. (cited on pages 3, 7, 8, 22, and 87)

ŞIMŞEK, A. S. AND TOPALOGLU, H., 2018. An expectation-maximization algorithm to estimate the parameters of the markov chain choice model. *Operations Research*, 66, 3 (2018), 748–760. (cited on page 6)

SONG, J.-S. AND XUE, Z., 2007. Demand management and inventory control for substitutable products. (cited on page 17)

STRAUSS, A. K.; KLEIN, R.; AND STEINHARDT, C., 2018. A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research*, 271, 2 (2018), 375–387. (cited on pages 1, 2, and 4)

TALLURI, K. AND VAN RYZIN, G., 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50, 1 (2004), 15–33. (cited on pages 2, 5, 15, and 35)

THURSTONE, L. L., 1927. A law of comparative judgment. *Psychological review*, 34, 4 (1927), 273. (cited on page 5)

TRAIN, K., 2003. *Discrete Choice Methods with Simulation*. SUNY-Oswego, Department of Economics. (cited on page 5)

TVERSKY, A., 1972a. Choice by elimination. *Journal of Mathematical Psychology*, 9, 4 (1972), 341 – 367. (cited on pages 7 and 10)

TVERSKY, A., 1972b. Elimination by aspects: a theory of choice. *Psychological Review*, 79, 4 (1972), 281–299. (cited on pages 7, 8, 10, and 22)

TVERSKY, A. AND KAHNEMAN, D., 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185, 4157 (1974), 1124–1131. (cited on pages 2, 3, and 10)

TVERSKY, A. AND SIMONSON, I., 1993. Context-dependent preferences. *Management Science*, 39, 10 (1993), 1179–1189. (cited on pages 7, 43, and 87)

VAN HENTENRYCK, P.; ABELIUK, A.; BERBEGLIA, F.; MALDONADO, F.; AND BERBEGLIA, G., 2016. Aligning popularity and quality in online cultural markets. In *Tenth International AAAI Conference on Web and Social Media*. (cited on pages 14, 98, 100, and 101)

VAN RYZIN, G. AND VULCANO, G., 2015. A market discovery algorithm to estimate a general class of nonparametric choice models. *Management Science*, 61, 2 (2015), 281–300. (cited on pages 5 and 7)

VAN RYZIN, G. AND VULCANO, G., 2017. Technical note—an expectation-maximization method to estimate a rank-based choice model of demand. *Operations Research*, (2017). (cited on page 89)

VULCANO, G.; VAN RYZIN, G.; AND CHAAR, W., 2010. Om practice—choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing & Service Operations Management*, 12, 3 (2010), 371–392. (cited on pages 1 and 5)

WANG, R., 2012. Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters*, 40, 6 (2012), 492 – 497. (cited on pages 70 and 107)

WANG, R., 2019. Consumer choice with consideration set: Threshold Luce model. *Available at SSRN 3383802, June 1st*, (2019). (cited on pages 18, 63, 88, and 89)

WANG, R. AND SAHIN, O., 2018. The impact of consumer search cost on assortment planning and pricing. *Management Science*, 64, 8 (2018), 3649–3666. (cited on pages 11, 17, 18, 72, 83, 84, and 88)

WATTS, D. J., 2011. *Everything is obvious: How common sense fails us*. Crown Pub. (cited on page 14)

WILLIAMS, H. C., 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and planning A*, 9, 3 (1977), 285–344. (cited on pages 5 and 16)