MULTI-LOCATION ASSORTMENT OPTIMIZATION UNDER LEAD TIME EFFECTS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

INDUSTRIAL ENGINEERING

By

Utku Karaca

August 2018

Multi-Location Assortment Optimization Under Lead Time Effects
By Utku Karaca
August 2018

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Alper Şen(Advisor)

Oya Karaşan

Mehmet Rüştü Taner

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

MULTI-LOCATION ASSORTMENT OPTIMIZATION UNDER LEAD TIME EFFECTS

Utku Karaca
M.S. in Industrial Engineering
Advisor: Alper Şen
August 2018

We have investigated the assortment planning problem for an online retailer that has multiple fulfillment centers to maximize its expected profit. Each fulfillment center is responsible for a customer segment which has its own customer profile, and each customer segment's demand is governed by a multinomial logit model (MNL), resulting in a mixtures of MNL (MMNL) model. A demand is primarily met by the responsible fulfillment center, if available. However, if a product is not available in the responsible fulfillment center, the demand can be met by fulfillment centers in other regions at an additional shipping cost paid by the firm. The shipping cost depends on the distance between regions, so it varies by origin and destination. We assume that each customer has access to the entire assortment in all fulfillment centers. To solve this problem, different from the literature, we have formulated the problem using a conic quadratic mixed integer programming approach. Later, the conic formulation is strengthened with valid inequalities. We have provided a numerical study to test the performance of our formulation against other formulations. Results show that our conic formulation together with the valid inequalities delivers outstanding performance compared to others in the literature. We also validated our approach using data from a local chain that operates in Northwestern part of Turkey.

Keywords: online retailing, multi-location assortment optimization, MMNL model, conic integer programming.

ÖZET

TESLİM SÜRESİ ETKİSİ ALTINDA ÇOK KONUMLU ÜRÜN ÇEŞİDİ EN İYİLEMESİ

Utku Karaca
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Alper Şen
August 2018

Bu çalışmada, beklenen karını en çoklama amacıyla, çoklu dağıtım merkezlerine sahip olan bir elektronik perakendeci için ürün çeşidi en iyilemesi problemi ele alınmıştır. İncelenen problemde, her dağıtım merkezi bir müşteri sınıfından sorumludur ve her sınıf kendi müşteri profiline sahiptir. Her müşteri sınıfının ürün tercihlerinin ayrı bir MNL tüketici seçimi modeline göre davrandığı kabul edilmiştir. Çok müşteri sınıflı MNL tüketici seçimi modelleri literatürde MNL karışım (MMNL) modelleri olarak geçmektedir. Müşteriden gelen bir talep öncelikle bulunduğu bölgedeki dağıtım merkezinden karşılanmaktadır. Eğer belirli bir bölgedeki dağıtım merkezinde bir ürün yoksa, bu ürüne ait talep diğer bölgelerdeki dağıtım merkezlerinden de karşılanabilmektedir. Bu durumda ortaya çıkan ilave sevkiyat masrafı şirket tarafından karşılanmaktadır. Sevkiyat masrafı dağıtım merkezinin ve teslimat noktasının konumlarına bağlı olarak değişmektedir. Her müşteri bütün dağıtım merkezlerinde bulunan ürünler arasından tercih yapabilmektedir. Literatürün aksine, tanımlanan bu problemi çözmek için konik ikinci dereceden karışık tamsayılı programlama yaklaşımı kullanılarak bir formülasyon geliştirilmiştir. Daha sonra bu formülasyon geçerli eşitsizlikler ile kuvvetlendirilmiştir. Önerilen formülasyonun performansını literatürde bulunan diğer performanslar ile karşılaştırmak için sayısal bir çalışma gerçekleştirilmiştir. Çalışma sonucunda önerilen geçerli eşitsizlikler ile kuvvetlendirilmiş konik formülasyon literatürdeki önerilen formülasyonlara kıyasla üstün performans gösterdiği gözlemlenmiştir. Son olarak geliştirilen yaklaşım Kuzeybatı Türkiye'de faaliyet gösteren yerel bir perakende zincirinden alınan verilerle doğrulanmıştır.

Anahtar sözcükler: e-perakendecilik, çok konumlu ürün çeşidi en iyilemesi, MMNL model, konik tamsayılı programlama.

Acknowledgement

First and foremost, I would like to thank my advisor Assoc. Prof. Alper Şen for his invaluable support, understanding and encouragement throughout my master's study. I feel extremely lucky to have the opportunity to work under his supervision.

I would like to thank Prof. Oya Karaşan and Assoc. Prof. Mehmet Rüştü Taner for devoting their time to read and review my thesis and for their valuable comments.

I am indebted to Ali İhsan Akbaş for his true friendship in my life for many years and shepherding my thesis to its final form.

I am grateful to the Department of Industrial Engineering for the unique experiences throughout my undergraduate and graduate study.

I am extremely grateful to my dearest friends Doğancan Demirtaş, Eren Togay and Umut Gölbaşı with whom I shared lots of sleepless nights and unforgettable moments. Also, I am indebted to Beyza Yılmaz for simultaneous coffee talks and her lasting support. I want to thank Merve Bolat for her invaluable support along with her encouragement during our countless talks and walks together in the campus. I would also like to thank Hale Erkan, Başak Erman and Harun Avcı for their endless support in our shared journey of the master's study with plenty of compelling, beautiful and unforgettable memories. Lastly, I would like to give my special thanks to Damla Akoluk. I would not achieve any of these things without her endless support and love.

I am deeply grateful to my parents Dilek Karaca, Başak Karaca and Süleyman Karaca for their invaluable support, encouragement and understanding throughout my life. Without their existence, it is not possible to imagine any of the things I have done so far; my love for them is something that words cannot explain.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The main objective of a retailer is to maximize its revenue which is directly related with the set of products that are served to its customers, namely its assortment. From customers' perspective, in general, purchase behavior changes with the different assortment offerings even though they come for a certain item. For instance, a customer comes to a store and wants to buy a new television. He has some definite preferences about the attributes of the television such as the screen size, sharpness rate, operating system, resolution, and price. He/She is looking for a television that has these features with a price not exceeding his/her budget. He/She is going to buy a product which exceeds his/her threshold utility in the assortment if any. If there is no such alternative in the store, the customer leaves without purchasing and checks products in other stores. Therefore, for the retailer, the following question needs to be analyzed and answered carefully: What should the assortment be in order to maximize the revenue? This problem can be applied not only to retail sector but also to many other contexts such as online advertising and social security [1].

Expanding the product variety as much as possible seems the best solution for retailers. Quelch and Kenny [2] have presented a report on the product variety and have stated that this variety increased 16% per year whereas the shelf space increased only 1.6% per year between 1985 and 1992. The growth has been lopsided, and this would create a problem for choosing products among available options. This increases the complexity of the problem. Also, in the report, they have pointed out the inventory holding cost and have stated that increasing variety means an increase on inventory levels not only for fast products but also for slow ones, causing problems especially if the products are perishable. Additionally, when the variety increases, a trade-off between filling the shelf space with most expensive but less popular products, or relatively less profitable and more popular ones, arises. Fisher and Vaidyanathan [3] have stated this problem with a real life example in retail domain. Therefore, assortment selection is one of the most difficult and significant decisions for retailers [4], and this problem is a difficult combinatorial problem in nature because of the many potential products to offer [5].

Assortment planning is an emerging field of academic study. In general, academic approach to the decision problem is based on the mathematical formulation of an optimization problem to find optimal assortment with optimal inventory levels. One can see either Kök et al. [6] or Mahajan and van Ryzin [7] for an extensive literature review.

Cachon et al. [8] have stated that a common approach to assortment selection process starts with fitting a consumer choice model to observed sales data since the key input in most assortment models is a consumer choice model [9]. In assortment planning literature, there are three commonly used customer choice models; locational choice model, exogenous demand model and utility based models, mainly multinomial logit model (MNL). Before explaining each choice model, we need to introduce a notation for all choice models.

N: The set of all products, $N = \{1, 2, \ldots, n\}$,
S : A subset of products carried by a retailer, $S \subset N$,
π_j: The price of product j.

Perhaps the most popular discrete choice model is the multinomial logit model. This model is widely used in the literature, especially in economics and marketing [10, 6]. Every customer visiting a store gets a utility U_j from product j. This utility can be decomposed into two parts, deterministic component u_j and random component ε_j:

$$U_j = u_j + \varepsilon_j.$$

It is assumed that the random component ε_js are independent and identically distributed (henceforth i.i.d.) random variables with Gumbel (or double exponential) distribution [11] which has cumulative distribution function (cdf)

$$F(x) = e^{-e^{-(\frac{x}{\mu}+\gamma)}}$$

where γ is Euler's constant (0.5772...) and μ is a scale parameter. The mean and the variance of the distribution are:

$$E[\varepsilon_j] = 0, \quad Var[\varepsilon_j] = \frac{\mu^2\pi^2}{6}.$$

In the multinomial logit model (MNL), the probability that a customer chooses product j from assortment $S \cup \{0\}$, where $\{0\}$ stands for the possibility of no-purchase, is defined as

$$p_j(S) = \frac{\upsilon_j}{\upsilon_0 + \sum_{k \in S} \upsilon_k} \tag{1.1}$$

where $\upsilon_j = e^{\frac{u_j}{\mu}}$, and υ_0 is the base utility, i.e., utility of no-purchase option for the customer [12].

The wide usage of the MNL model is due to the fact that it is "analytically tractable, relatively accurate and can be estimated easily" [13]. Even though the MNL model is easy to use with other attributes and easy to estimate in industrial applications, it has two main deficiencies. The first and the most major one is the Independence of Irrelevant Alternatives (IIA) property which holds if the ratio of choice probability of any two alternatives is independent. An explanatory example, known as "blue bus/red bus paradox", has been given by Debrue [14]. In this example, first, an individual is given two alternatives, bus and car with the same probability, 1/2, to go to work. In the second situation, as an alternative, another colored bus is introduced, resulting with the alternatives of blue bus, red bus and car. It is known that the individual is indifferent for bus alternatives since both alternative have the same utility to him/her. As an observer, the probability of choosing bus or car is expected to remain same as in the first situation, 1/2. However, in the MNL model, the probabilities are calculated as 1/3 and 2/3 for choosing car and bus, respectively, since the alternatives are car, blue bus and red bus. Therefore, IIA property does not hold in practice since there are products among alternatives that can be grouped. The second deficiency is related with substitution between different products. In the MNL model, it is impossible to have two different products with the same penetration and different substitution rates. Due to the IIA property, one needs to use the MNL with caution.

A pioneering work in the area by Guadagni and Little [15] have studied estimating market share of brands with the marketing attributes such as advertising and promotion. After this study, the MNL model and many other MNL models, which are lately introduced, are widely used in the marketing literature. Another well-known application is done by Ben-Akiva and Lerman's [16] on the estimation of travel demand.

A generalization of the MNL model is Nested Logit (NL) and is introduced by Ben-Akiva and Lerman [17] to alleviate the IIA property. In NL model, the process of choosing a product in an assortment consists of two steps. When a customer arrives to the system, first, she/he chooses a subset according to a MNL model over subsets and then selects a product from that subset, again, using a MNL model. To overcome IIA property, one needs to create disjoint subsets. A key problem with the NL model is that the user needs to know key attributes of products and customers' hierarchical selection process due to the IIA property.

For this study, we consider the assortment optimization problem under the MNL model since our alternatives belong to the same subgroup, and we do not need any other partitioning process. The model will be explained in detail in the following chapters.

The exogenous demand model is another widely used model in the literature, mostly in inventory management for substitutable products. In this model, it is assumed that each individual has a favorite product, and the individual buys that product if available. If that product is not in the assortment, the individual will buy the second favorite product with the probability of δ, or the individual will leave without buying anything with the probability $1-\delta$. If the second favorite product is not available, the same procedure continues until the individual comes up with a decision. Comparing to the MNL model, the exogenous demand model has more degrees of freedom. Although the model enables consumers for any type of substitution such as adjacent substitution and one-product substitution, in practice, the estimation of parameters is difficult. Smith and Agrawal [18], and Kök and Fisher [19] have used the exogenous demand model for assortment planning problems.

The last commonly used discrete choice model, namely locational choice model, has been introduced by Hotelling [20], extended by Lancaster [21, 22], and extensively discussed in Anderson et al. ([12], §4). In this model, each product is seen as a bundle of its characteristics and can be shown as a vector in characteristics space. Each individual is represented by her/his ideal preferences of m attributes $y \in \mathbb{R}^m$, and each product j is represented by a point z_j in the same characteristics space $\mathbb{R}^m$. Recalling that π_j is the price of alternative j, a customer's utility is defined as

$$U_j = C - \pi_j - g(y, z_j)$$

where C is a positive constant and $g : \mathbb{R}^m \to \mathbb{R}$ is a metric for measuring distance between the consumer's ideal point y and product j's location z_j in $\mathbb{R}^m$. In the selection process, the consumer chooses the product with the maximum utility, i.e., the one that has the smallest U_j. If that product is not available, then the consumer moves to the second highest one. In the locational choice model, substitution can happen between similar products whereas in the MNL model, consumers can substitute to any product in the assortment. Therefore, IIA property does not hold, and the rate of substitution between products can be controlled. Gaur and Honhon [23], McBride and Zufryden [24], and Kohli and Sukumar [25] have used the locational choice model in their assortment planning study.

Another important issue about retailing is fulfillment decisions. For this study, we primarily consider an online retailing setting environment, also known as e-tailing. There are certain reasons behind this assumption. First of all, in each year, online retailing increases its share in all retailing sales [26]. In 2017, online retail sales reached $453 billion in the U.S. which is approximately 16% higher than the previous year, and the forecast for annual volume in 2020 is made as around $3 trillion [27]. The situation is similar for Turkey. The Internet penetration in Turkey has reached 58%, and the volume of e-commerce sales has reached 30.8 billion Turkish Liras in 2016, indicating 13% and 25% increase, respectively [28]. Secondly, cost of providing online sales channel can reach, on average, 6% [29] in whole budget of a company. When the larger companies are considered, the amount would be huge. Therefore, providing additional sales channel should be analyzed to see whether it is feasible or not. Supporting the difficulty of the problem, recent PwC report [30] shows that the greatest challenge for executives facing in providing an omni-channel experience for their customers is budget constraints. The numbers show the potential of the Internet retailing.

In old fashion retail, each customer visits stores, generally supplied by distributors, and sees the assortment, i.e., shelves in the store. Then the customer decides to buy a product or to leave the store with no purchase because of the availability of products. This availability depends on the existing shelf space and so, it has a direct relationship with the pyhsical space. In this type of retail, there is no lead time for customers since the customers are physically in a store where products are offered.

In online retailing, there are many fundamental differences. First of all, demand is met from warehouses (henceforth fulfillment centers) which can be located in any region. Unlike the older type, there is less limitation on the physical space since products are not stored in stores, enabling companies to carry wider assortment. It increases the chance of selling a product for a company, consequently, its revenues. Secondly, companies create a distribution network and can carry region-based inventory depending on customers profile in the region, so that this specialization increases the customer satisfaction. Online retailing also decreases the operational cost since it enables the company to fulfill demand from any fulfillment center whereas in the traditional one, every available product needs to be, physically, in the store. In this work, we are going to consider an online retail company and work on an assortment problem by considering the perspective mentioned above.

In this study, the problem is to find the optimal assortments for an online retail company's fulfillment centers in different areas under lead time effects. To maximize total revenue, the company needs to select products to serve its customers. For each fulfillment center, there is a constraint on the number of products in the assortment, i.e., capacity constraint. It is assumed that each fulfillment center has its own customer profile, i.e., the number of customer segments are equal to the number of fulfillment centers, and each customer in a segment served by the same fulfillment center shows same purchasing behavior, namely, they are in the same segment. For the first part of the study, each group's demand is considered as an independent MNL, so that each group has different preferences. Moreover, a customer's demand can be met from other regions. For this case, the preference of a product is decreased since there would be a lead time for that product. We assume that lead time decreases the preferences of products. If there is only one fulfillment center, then we need to consider the problem as a single-location assortment problem with multiple customer segments, which is equivalent to Mixed Multinomial Logit (MMNL) model. Rusmevichientong et al. [31] have solved assortment optimization problem under the MNL model in an efficient way, however, for the MMNL model, there is no efficient solution, and its NP-hardness is proved by Bront et al. [32] and Rusmevichientong et al. [33].

For this study, first, we will formulate the assortment problem under the MMNL model as a mathematical problem. After, we will reformulate the problem as a mixed integer linear program and develop a conic formulation which is extended from the pioneering work of Şen et al. [34]. To strengthen the formulation, McCormick estimators are used, resulting in significantly less time for solving the problem. The new formulation is used on a real sales transaction data with different settings of the problem such as capacity and no purchase preferences.

Later, first, we will present a numerical study for our proposed formulations on four regions with forty products and give analysis about their performances under different cases. Furthermore, in Chapter 5, we will present a case study on real data which is given by a local chain with four stores in the Northwestern part of Turkey. In our study, we will mainly focus on the data from the two closest stores. Although the data is from traditional retailing stores, it can be viewed as an e-tailing setting due to the following reasons: First, the distance between the two stores is only 150 meters, which allows a customer to purchase a product from the other store in case the product is not available in the store which he/she arrived in the first place. Obviously, this leads a disutility for the customer which we model explicitly. Second, there are no limitations regarding the movement of products between the two stores, resulting that the firm can select the locations for both storing and serving its products. Hence, the company can create its own distribution network and carry region-based inventory depending on the profiles of its customers.

In the case study, we will fit MNL models for two customer segments where each customer segment can see the assortment of two stores and buy products from both stores. Using the fitted MNL models, we will determine the preferences of products, and based on the preferences, we will propose assortments for two stores.

The remainder of the study is organized as follows: in §2, assortment planning and choice model estimation literature is reviewed. In §3, mathematical representation of assortment optimization is shown. §4 gives a numerical study of the formulations in §3. §5 provides details of a case study and a suggestion on assortment planning with the help of parameters. Finally, §6 concludes the thesis and points out possible research areas.

Chapter 2

Literature Review

In this chapter, we provide a review of literature on both assortment optimization and parameter estimation under the MNL model and its extensions. Kök et al.[6], Karampatsa et al. [35], and Mahajan and van Ryzin [7] have done a comprehensive literature review in assortment planning. In marketing and economics literature, the MNL model has been widely used, however, the applications are relatively new in assortment optimization problem. In the literature, main focus has been on the single-location setting, whereas in some applications, the assortment planning problem needs to consider multiple locations simultaneously. Eventually, to explain customer behavior empirically or to launch choice-based inventory models, one has to be able to estimate a choice model. In the related literature, this estimation is generally done by constructing a likelihood function and trying to find its maximizers.

In the following sections, assortment planning literature under the variations of MNL model for both single and multi-location setting is presented. Furthermore, the literature on parameters estimation is also reviewed.

2.1 Assortment Planning

2.1.1 Single Location

2.1.1.1 Under MNL Model

Initally, van Ryzin and Mahajan [36] have formulated the assortment planning problem under a MNL model of consumer choice. In their model, the retailer maximizes total utility of the assortment under the newsvendor model. They have assumed that the revenue and cost of all products are the same, i.e., $r_j = r, c_j = c$ for all $j \in S$, and the products are ordered according to their preferences, i.e., $\upsilon_1 \geq \upsilon_2 \geq \cdots \geq \upsilon_n$. The probability that a consumer chooses product $j \in S$ is

$$p_j(S) = \frac{\upsilon_j}{\sum\limits_{k \in S \cup \{0\}} \upsilon_k},$$

where S is the set of products in the assortment, and the preferences are defined as $\upsilon_j = e^{u_j/\mu}$. They have assumed that the assortment does not have substitutes, and customers select their first choice if available, otherwise, they do not substitute with the second choice, meaning that if the first choice is not available, the sale is lost. They have defined a rule of adding product j to the assortment. If the product j makes more profit than the sum of the profit losses when it is added, then it is better to include product j in the assortment. They have come up with a structural result that the optimal solution can be found either by adding the next highest preference product to the assortment or not. In other words, if the current assortment consists of products $\{1, 2\}$, the model would check only the revenue of an assortment consisting products $\{1, 2, 3\}$ since the next highest preference product is the product 3. Therefore, they have defined a popular set $P = \{\varnothing, \{1\}, \{1, 2\}, \{1, 2, 3\}, \ldots, \{1, 2, \ldots, N\}\}$ and stated that the profit maximizer assortment is one of the elements of P, decreasing the number of feasible solutions from 2^{N+1} to $|N + 1|$. Additionally, they have shown that it is more profitable to carry deeper assortment with sufficiently high price and sufficiently high no-purchase preference.

Later, the same problem is also studied by Mahajan and van Ryzin [13], however, unlike the previous paper, they have considered dynamic substitution. They have examined the structural properties of the expected profit function. They have shown that under broad assumptions on the demand process, total sales of each product are concave in their own inventory levels, and the marginal value of an additional unit of the given product is decreasing in the inventory levels of all other products. They have shown that the expected profit function is not concave, indicating that finding global optimal solution is difficult. To overcome this, they have proposed a stochastic gradient algorithm with a convergence guarantee under mild conditions and compared the algorithm with heuristics available in the literature. Finally, they have concluded that, under substitution, the retailer should stock relatively more of popular alternatives and relatively less of unpopular ones than what would be optimal under a traditional newsboy problem.

Chong et al. [37] have presented an empirically based modeling framework for executives to evaluate the revenue and lost sales inference of alternative category assortments. The new framework is needed because of the increasing complexity of managing a category assortment which is caused by the increased product turnover and proliferation rates in many of the categories. Along with a local improvement heuristic, the modeling framework creates a variant category assortment with higher revenue. They have validated their framework with shopping trips and purchase records. Additionally, they have provided a numerical study and concluded a profit improvement of up to 25.1%.

Cachon et al. [8] have extended van Ryzin and Mahajan's model [36] in a way that examines the impacts of consumer choice under search, enabling that if customer can find whatever he/she looks for, he/she would not buy that product and search it in other stores. Including the previous "search" case, they have studied three versions of the assortment problem. Their analysis have indicated that the decision of adding a product to an assortment needs a consideration regarding not only direct cost and revenue of that product but also the foreseen indirect benefit of an extended assortment that keeps consumer from searching other stores. Among their numerical results under different settings, they have shown that adding an unprofitable product to an assortment is optimal since it could prevent consumer from searching. They have tested the no-search model and concluded that it performs well with the categories having many alternatives. However, under overlapping assortment model, broadening the assortment decreases the value of search since this lessens the potential number of new alternatives that the customer may observe if she/he chooses to search. This significant impact cannot be captured by the developed model. They have concluded that the reduction in an assortment should be done carefully when the market enables consumer search.

Li [38] has considered the problem for a single period with dissimilar cost parameters. They have taken the store traffic as a continuous random variable and found the structure of the optimal assortment. They have defined a measure called profit rate to calculate the profitability of each product and shown that the optimal assortment consists of a few of the highest profit rate products. They have also analyzed the discrete store traffic case and stated that finding the optimal solution is difficult. For this case, with the inspiration of the continuous case, they have proposed a profit rate heuristic which can attain optimal solution with the setting where cost parameters are equal and demand is distributed normally. They have concluded that measuring the profitability of each product is important under the random demand and existence of cannibalization.

Miller et al. [39] have developed a methodology for selecting the optimal assortment for rarely purchased products. The idea incorporates consumer heterogeneity and uses integer programming formulation for the assortment selection problem. They have assessed their methodology by using real data of a national retail chain.

Rusmevichientong et al. [31] have considered static and dynamic capacitated assortment optimization problem. As a capacity constraint, they have determined a limit on the number of different products that a retailer can carry in the assortment. For the static part, they have assumed the parameters of the MNL model are known and developed a profit-maximizing algorithm to find the optimal assortment. For the dynamic one, they have presented an adaptive policy which takes the parameters as unknown, estimates them from the past data and optimizes the profit at the same time. They have also tested their policy with online retailer data and concluded that their policy works well.

Rusmevichientong and Topaloglu [40] have formulated a robust assortment optimization problem under the MNL model. Assuming that the parameters of the model are not known, they have proposed a set of possible parameters, called uncertainty set. The optimization model maximizes the worst-case expected revenue over the uncertainty set of parameters. Both static and dynamic cases are taken into consideration. For the static case, there are no inventory decisions to be made, whereas in the dynamic case, there exists a limited initial inventory that have to be allocated over time. They have characterized the optimal policy for both cases and obtained the operational insights. They also presented a family of uncertainty sets, enabling the decision maker to manage the trade-off between increasing the average revenue and protecting himself against the worst-case scenario. Finally, by conducting a numerical study, they have compared their methods with available methods in the literature and concluded that their method provides better performance for the worst case, especially when there is significant uncertainty in the parameter values.

Davis et al. [41] have studied the assortment problem with constraints on cardinality, location of products on the shelves and their precedences. All variations are modeled separately, and these models can be reformulated by using a set of totally unimodular constraints. They have shown that solving fractional binary problem can be transformed into a linear program.

Topaloğlu [42] has formulated a nonlinear program to solve an assortment selection and stocking problem jointly, where the decision variables are the quantity and the duration time on the shelves of each product offered in the assortment, over a finite selling horizon. The formulation is difficult to solve because the number of decision variables increases exponentially as the number of possible products increases, and the objective function is not concave. They have used the structure of the MNL model to reformulate the nonlinear program which consists of a decomposable objective function and linearly growing decision variables. The proposed model is solved through a dynamic program which needs discretizing the state variable, resulting that they have studied an integer program that closely tracks the solution as another approximation method. Their proposed reformulation and approximate solution provide insights for offer sets.

One of the recent work on the assortment planning is done by Goyal et al. [43]. They have considered a single-period joint assortment planning and inventory control problem under dynamic substitution with stochastic demand. They have shown the NP-hardness of the problem along with the hardness of approximating within a factor better than $1 - 1/e$. They have proposed a scheme, called polynomial-time approximation scheme (PTAS), to solve the problem with any level of accuracy in an efficient way. The algorithm guarantees the near-optimal performance. They have supplied their findings with a numerical study and concluded that assortments with a relatively small number of alternatives can achieve the most of the probable revenue.

The last study in this section that we are going to present is by Dzyabura and Jagabathula [44]. In this study, they have considered a firm which offers both online and offline channels to its customers. Since the offline channel enables customers to see and experience products, it has an impact on the customer decisions. Hence, they have studied assortment of the offline channel that maximizes the profit of the firm across both channels. When they have modeled the consumer demand, they have used the MNL model and have included the effect of the physical evaluation on preferences. They have proved the NP-hardness of the problem and derived the optimal solutions for some special cases. For the general problem, they have provided near-optimal approximations. Finally, they have presented an empirical study and concluded that the offline assortment planning can increase the expected revenue of online sales up to 40%.

2.1.1.2 Under Nested Logit Model

Kök and Xu [45] have studied assortment and pricing problems, jointly, for a category under the nested logit model. The objective is to find the optimal assortment and corresponding prices maximizing the expected revenue. In the paper, they have considered two different structures of decision. First one is called brand-primary model, and in this model, customers, first, choose a brand, then select a product from that brand. Second one is called type-primary model where the order of choice is vice versa, i.e., first, customers choose a product type, then select a brand for that product type. They have reported different optimal assortments for each model. They have concluded that the executives need to understand the hierarchical choice process of their customers in order to create the right set of products and prices. Finally, they have extended the structural properties of the optimal solution defined by van Ryzin and Mahajan [36].

Alptekinoğlu and Grasas [46] have used nested logit model for the assortment planning problem for a set of horizontally differentiated products. Additionally, they have considered the customer returns. They have analyzed different settings of return policies. When the return rates are high or returning is prohibited, the optimal set of products includes only the most popular products, consistent with the results of the van Ryzin and Mahajan [36]. On the other hand, when the return rates are low and the policy is relatively strict, they have stated that it is optimal to select a mixture of eccentric and the most popular ones. Finally, they have shown empirical results to support their findings.

Davis et al. [5] have studied assortment planning problem under nested logit model. They have proved that the problem is solvable when the consumers make a purchase from the selected nest for sure and the dissimilarity parameters of each nest satisfies certain conditions. If one of the assumptions does not hold, then the problem becomes NP-hard. To handle with the NP-hardness, they have developed parsimonious collections of candidate assortments with worst-case performance guaranteed. Finally, they have found an upper bound on the optimal expected revenue.

Gallego and Topaloğlu [47] have considered the assortment planning under nested logit model with cardinality and space constraints. They have formulated the problem as an LP model and shown that under the cardinality constraints, the optimal solution can be found efficiently. However, with space constraints, the problem becomes NP-hard. To overcome this, they have provided a performance-guaranteed assortment. Finally, they have solved the joint assortment planning and pricing problems with nests efficiently.

Li and Rusmevichientong [48] have proposed a simple and fast greedy algorithm for finding optimal assortment under the two-level nested logit model. The algorithm, in each iteration, takes out at most one product from each nest to compute an optimal solution. They have also provided a necessary and sufficient condition for an optimal assortment.

In the work of Rodríguez and Aydın [49], pricing and assortment planning decisions for a manufacturer, which has dual sales channels (direct and through retailer), is studied with a consideration of inventory costs. They have modeled the consumer demand under nested logit model. They have provided many managerial insights for the pricing decisions of the manufacturer's direct channel. They have found that alternatives with highly uncertain demand would carry a lower wholesale price. Moreover, they have studied the case that there exists a conflict in between assortment preferences of both manufacturer and retailer.

Feldman and Topaloğlu [50] have studied the problem under the same conditions with Gallego and Topaloğlu [47]. They have found an exact solution method for the assortment optimization problem with cardinality constraints. They have also provided an approximation algorithm to solve the model with space constraints. Finally, they have supplied their findings with a comparison of their algorithm and the upper bounds on the optimal expected revenue calculated by a linear program.

Taking the prices as fixed, Li et al. [51] have considered the problem under d-level nested logit model where d is an arbitrary number. They have provided an algorithm, which has the complexity of $O(dn \log n)$, to solve the problem. For the pricing part, assuming the assortment is fixed, they have proposed an algorithm that produces a sequence of prices converging to a stationary point iteratively.

2.1.1.3 Under Mixtures of MNL Model

In general, especially in online retailing, customers do not follow the same MNL model since their profiles are not the same. Web pages serve to many markets, resulting that there is often customer heterogeneity in terms of preferences. Hence, single MNL model cannot explain different customer profiles, and customers need to be grouped into segments. For each customer segment, a MNL model can be used to express its preferences. Therefore, mixtures of MNL model (MMNL) is used when there exists customer heterogeneity. The model is introduced by Cardell and Dunbar [52] and Boyd and Mellman [53]. The MMNL model does not have limitations of the MNL model, and it is appropriate and is used for our study.

Bront et al. [32] and Méndez Díaz et al. [54] have formulated the assortment planning problem as a mixed integer linear program where the consumer choices are governed by the mixtures of MNL (MMNL) model. Both papers have shown the NP-hardness of the problem. To deal with this, Bront et al. [32] have provided a greedy heuristic algorithm, and Méndez Díaz et al. [54] have proposed a branch and cut algorihm using valid inequalities. The proposed algorithm is computationally efficient and provides near-optimal solutions.

Rusmevichientong et al. [33] have focused on the revenue-ordered assortments and analyzed the cases in which they are optimal. For the revenue-ordered assortment, when the number of consumer groups and products are low, they have presented an approximation guarantee of $\min\{G, \lceil n/2 \rceil\}$, where n is the number of products and G is the possibly mean vector of utilities. However, they have given an approximation guarantee of $e \log(er_1/r_n)$, where r_i is the revenue of ith product when the number of consumer groups and products are high. They have supplied their findings with numerical experiments.

Feldman and Topaloğlu [55] have studied the capacitated assortment optimization problem under the MMNL model. They have proved that even if the capacity contsraints are not tight, the problem is NP-hard. To handle the NP-hardness of the problem, they have provided a fully polynomial time approximation scheme.

A recent work is done by Şen et al. [34]. They have studied on the single-location assortment planning problem under the MMNL model. The problem has been reformulated, for the first time, as a conic quadratic mixed integer program. With the new formulation, they can solve large instances in shorter times optimally. They have also provided McCormick estimators to strengthen the conic formulation. Finally, they have presented a comparison between their model and other formulations in the literature and concluded that their model is the best among the compared ones. In this thesis, we will extend the single-location conic quadratic formulation to the multi-location setting.

2.1.2 Multiple Locations

Kök et al. [6] have reviewed the related literature and stated that the majority of the studies have been done for single-location setting. There are some recent works that examine assortment optimization for multi-location setting, and this is an open area for research. In this thesis, we are considering a multi-location setting for the problem.

Aydın and Hausman [56] have studied the assortment optimization problem where the consumer choices are governed by the MNL model in a decentralized supply chain with a supplier and a retailer. They have shown that because of the lower profit margins than the vertically integrated supply chain, the retailer prefers deeper assortment than the optimal assortment of the supply chain, which creates coordination problems. To solve this, the supplier can arrange paying a fee per product to the retailer, resulting both sides having more profit. Finally, they have presented an example that is from a grocery industry.

Singh et al. [57] have analyzed the impact of product variety problem for different supply chain structures. At first, they have considered a traditional channel where the retailer owns the inventory, governs its stocks and decides the assortment and stock levels. Another structure is drop-shipping channel where the wholesaler owns the inventory, governs its stocks, decides the assortment and stock levels, and ships the products directly to the customers. For this setting, the wholesaler imposes an extra cost for bearing the inventory risk. They have also provided the structure of the optimal solution in both channels under a single threshold policy. The main inference is that it is cost-efficient to carry larger assortment for wholesaler in the drop-shipping channel than the retailer in the traditional channel. Moreover, in the drop-shipping channel, the relative margin of the wholesaler increases with the number of retailers. Finally, they have analyzed a single firm's assortment and inventory level decisions where each customer's order is fulfilled from either one of the retail locations or central warehouse, and shown that it is cost-efficient to stock most preferred alternatives in both retail stores and the warehouse.

2.2 Parameter Estimation

2.2.1 With Panel Data

After Guadagni and Little [15], many papers in marketing literature have studied estimation of the parameters of the MNL model to investigate the marketing related variables on demand. In those papers, generally, panel data gathered from loyalty card shopping data has been used. In the MNL model, the deterministic component u_j is calculated based on observable attributes of each alternative, and to do this, we can use *linear in attributes* model. Let $x_j = (x_{j1}, x_{j2}, \ldots, x_{jm})$ be the vector of attributes of product j, and $\beta = (\beta_1, \ldots, \beta_m)$ be the coefficient vector for all attributes.

$$v_j = e^{\frac{u_j}{\mu}} \quad u_j = \beta^T x_j, \quad j = 0, 1, \ldots, n.$$

Let y_{ij} be the indicator variable, and it takes 1 if product j is chosen by individual i. By using formula (1.1), one can find the choice probabilities. To find the maximum likelihood estimators for vector β, the likelihood function can be written as the following [16]:

$$L(\beta) = \prod_i \prod_j \left[\frac{e^{\beta^T x_j}}{\left(\sum_{k \in S_i} e^{\beta^T x_k}\right)} \right]^{y_{ij}}$$

where S_i is the assortment shown to customer i.

For computational reasons, we need to take the logarithm of the likelihood function. So, the log-likelihood function can be written as follows:

$$\log(L(\beta)) = \ell(\beta) = \sum_i \sum_j \log \left[\frac{e^{\beta^T x_j}}{\sum_k e^{\beta^T x_k}} \right] y_{ij}$$

$$= \sum_i \sum_j y_{ij} \left[\beta^T x_j - \log \left(\sum_{k \in S_i} e^{\beta^T x_k} \right) \right]$$

where S_i is the set of alternatives for customer i. The global concavity of the log-likelihood function is proved by McFadden [58], enabling the use of any type of nonlinear optimization technique to find maximizer β. Mahajan and van Ryzin [7] have stated that "under relatively general conditions, the MLE estimator of β is consistent, asymptotically efficient and asymptotically normally distributed." As seen in the formulation, the required computation power for estimation process does not increase with the number of variants but with number of observable attributes [59]. Therefore, with this approach, it is possible to estimate parameters of the MNL model. We refer our reader to Chiang [60], Bucklin and Gupta [61], Chintagunta [62], and Chong et al. [37] for the extensions of the model.

2.2.2 With Sales Transaction Data

For estimating parameters of the MNL model, we need complete data, however, in general, the data consists of only sales data. Sales transactions data are the information recorded from transactions, namely sales, given in each period, for a given product. With this data, we cannot observe non-buyers, therefore, this data can be called as an incomplete data, resulting that the approach described in previous section cannot be used.

To handle the incompleteness, Talluri and van Ryzin [63] have developed an *Expectation-Maximization* algorithm [64, 65] to estimate parameters of the choice model and arrival rates of customers with missing data where missingness comes from no-purchase option. Their algorithm uses only sales transaction data, starts with an arbitrary feasible estimates and iteratively calculates conditional expectation of likelihood function and finds the maximizer values of parameters. Vulcano et al. [66] have implemented the EM algorithm for sales transaction data to estimate demand under the MNL model. Chong et al. [37] and Kök and Fisher [19] also have applied this algorithm for estimation part of the assortment optimization.

Vulcano et al. [66] have studied the problem under the same settings and developed an EM algorithm to estimate parameters of the model. The difference of their approach is that they have treated observed demand as incomplete realizations of primary demand whereas Talluri and van Ryzin [63] utilize the data to estimate arrival rate and parameters at the same time.

Newman et al. [67] have proposed a two-step estimation method for the parameters of multinomial logit model. They have presented their work as an alternative work to EM algorithm developed by Talluri and van Ryzin [63]. Their proposed algorithm consists of decomposition of the log-likelihood function into marginal and conditional component. The algorithm works faster than the EM algorithm computationally and the method performs as well as the EM algorithm in terms of log-likelihood value. Moreover, the two-step method provides consistent estimates for parameters and can integrate with the price and other product attributes.

Abdallah and Vulcano [68] have studied the demand estimation under the MNL model. They have proposed a Minorization-Maximization (MM) algorithm which only requires market share and sales transaction data to maximize the likelihood function and guaranteed the convergence when there is a unique maximizer point for the likelihood function. They have also provided a variation of the algorithm which requires only sales transaction data for the demand estimation. To measure their goodness of fit, they have compared their MM method with the well-known EM methods [63, 66] and two-step approach [67] based on the log-likelihood function, prediction accuracy and revenue estimates. Moreover, they have shown that their proposed MM algorithm demonstrates better performance than EM and two-step approach algorithms.

Fadıloğlu et al. [69] have developed a mathematical model for assortment optimization. They have used sales transactional data of a local chain to perform assortment planning. Their model have mainly focused on the efficiency of the shelf usage and increased the profitability of the local chain. They have concluded that with their model, the profitability of the assorment has substantially increased after the reorganization of the product list.

A final study we have reviewed on parameter estimation is the Keane and Wasi's work [70]. In this study, the authors have attempted to estimate parameters of the discrete choice model with large choice sets. They have shown that random subsets can be used to decrease computational need when the choice set is large enough. They also provide an application on real data where the whole choice set has 60 options and the randomly generated subsets have 10 and 20 options. They conclude that subsets do not create significant bias in the estimation of parameters.

2.3 Our Contribution

The literature about the assortment optimization problems under variants of logit models are presented in this chapter. Most of the assortment optimization studies consider a single-location setting. However, our work focuses on the assortment planning problem for multi-location setting where each location has a separate MNL model. We have extended the Şen et al. [34] model and studied the setting of different shipping costs between fulfillment centers and lead time effects. Our objective is to find the assortment that maximizes the expected profit of the firm considering the transportation costs among different regions. We have also done a case study which includes both parameter estimation and assortment planning jointly.

Chapter 3

Formulation

3.1 Problem Definition

We study multi-location capacitated assortment optimization problem under dissimilar customer preferences. Let M denote the set of customer segments, and N denote the set of all possible products. Each customer segment has a separate MNL model, i.e., each customer segment has its own preferences for each product in the assortment and no-purchase preferences. Therefore, in multi-location setting, all demand can be modeled as a mixtures of MNL model. In online retail setting, each fulfillment center is located in different area and each of them has a customer segment. Although each of them has an assigned customer segment, they can serve any demand from other regions/segments as well provided that an additional transportation cost is incurred. Taking customer preferences, capacity constraints and shipping costs into consideration, the retailer needs to decide which products to offer in each location. In this thesis, the above setting is considered, and finding the optimal assortment for every fulfillment center is the primary concern.

For the interregional transportation, consider the following matrix Φ where the entry ϕ_{ij} shows the j^{th} preferred fulfillment center for the i^{th} region. The matrix below constitutes an example preference matrix of a system of three fulfillment centers:

$$\Phi = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$$

For instance, the first preference of each region is itself, meaning that if a fulfillment center can meet a demand from its region, then the demand is supplied from there. This is so, however, assuming that a demand for a product is originated from region 2, and the product is not available in region 2's fulfillment center. Therefore, since the second preference of region 2 is region 1, the product is sent by fulfillment center 1 if available. According to the preference matrix, the preferences of the first, the second and the third fulfillment centers are {1,2,3},{2,1,3} and {3,2,1}, respectively.

For our model setting, we assume that with a decrease on the preference order, the lead time increases, hence, the preference of a customer for that particular product decreases. Li [71], Li and Lee [72], and So and Song [73] have considered the lead time of a product as a discounting measure on the preference. Fisher et al. [74] have studied the value of rapid delivery in online retailing and reported that the decrease in lead time yields higher revenue. Therefore, we define each product's preference based on the location of that product and also the demand location, i.e., v_{ikj} where (i, k, j) corresponds to the demand location i, the stored location k and the product j, respectively.

For the mixtures of MNL model (MMNL), the assortment problem for a single location is NP-hard [32, 33]. Therefore, heuristics and approximations are widely developed to solve this problem in the literature [75, 76]. For this problem, we propose a mathematical formulation to the find optimal solution. First, we formulate the problem using a mixed integer linear programming approach. Then, we provide a conic quadratic mixed integer programming formulation for the problem, and add valid McCormick inequalities to strengthen the formulation.

3.2 Baseline Model

3.2.1 Sets & Parameters

Sets

M : Set of regions for customer segments and fulfillment centers.

N : Set of products.

Parameters

λ_i : Probability that the demand originates from region $i \in M$.

υ_{ikj} : Preference of product j in region i when it is served from region $k \in M$.

υ_{i0} : No purchase preference in customer class $i \in M$.

π_j : Unit revenue of product $j \in N$.

κ_i : Capacity of fulfillment center $i \in M$.

τ_{ikj} : Additional cost of transporting product $j \in N$ from region $i \in M$ to $k \in M$, 0 if $i = k$.

3.2.2 Decision Variables

$$z_{ikj} = \begin{cases} 1 & \text{if product } j \in N \text{ for region } i \in M \text{ is served from } k \in M, \\ 0 & \text{otherwise.} \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{if product } j \in N \text{ is stored in region } i \in M, \\ 0 & \text{otherwise.} \end{cases}$$

3.2.3 Objective Function

We can decompose the objective function into two; revenue and expenditure (addtional transportation cost).

3.2.3.1 Revenue

Expected revenue of a customer located in region $i \in M$ is given by:

$$\frac{\sum_{j \in N} \sum_{k \in M} \pi_j z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}}. \tag{3.1}$$

If we sum over the all customer segments, then the total expected revenue from a customer becomes

$$\sum_{i \in M} \lambda_i \left[\frac{\sum_{j \in N} \sum_{k \in M} \pi_j z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right]. \tag{3.2}$$

3.2.3.2 Transportation Cost

The expected total transportation cost can be written as follows:

$$\sum_{i \in M} \lambda_i \left[\frac{\sum_{j \in N} \sum_{k \in M} \tau_{ikj} z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right]. \tag{3.3}$$

3.2.3.3 Objective Function

When we sum the revenue and the expected transportation cost, the objective function becomes

$$\sum_{i \in M} \left[\lambda_i \frac{\sum_{j \in N} \sum_{k \in M} \pi_j z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right] - \sum_{i \in M} \lambda_i \left[\frac{\sum_{j \in N} \sum_{k \in M} \tau_{ikj} z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right]. \tag{3.4}$$

If we simplify equation (3.4), the objective function can be written as:

$$\max \quad \sum_{i \in M} \lambda_i \left[\frac{\sum_{j \in N} \sum_{k \in M} (\pi_j - \tau_{ikj}) z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right]. \tag{3.5}$$

3.2.4 Constraints

$$\sum_{k \in M} z_{ikj} \leq 1, \qquad \forall i \in M, \forall j \in N, \tag{3.6}$$

$$\sum_{j \in N} x_{ij} \leq \kappa_i, \qquad \forall i \in M, \tag{3.7}$$

$$x_{kj} \geq z_{ikj}, \qquad \forall i, k \in M, \forall j \in N, \tag{3.8}$$

$$x_{ij} \in \{0, 1\}, \qquad \forall i \in M, \forall j \in N, \tag{3.9}$$

$$z_{ikj} \in \{0, 1\}, \qquad \forall i, k \in M, \forall j \in N. \tag{3.10}$$

The first constraint (3.6) ensures that each demand is supplied from at most one fulfillment center. The second constraint (3.7) stands for the capacity constraint, limiting the number of different products in region i by κ_i. The third constraint (3.8) states that if product j in region k is shipped to a customer in region i, then it needs to be stored at region k. The last two ones (3.9, 3.10) declare the binary variables.

3.3 Mixed Integer Linear Programming Formulation

The Mixed Integer Linear Programming Formulation (MILP) capacitated assortment optimization model is formulated by Bront et al. [32], Méndez-Díaz et al. [54], and Şen et al. [34]. Therefore, we can reformulate our base model in a linear manner. First, we need to linearize the objective function by introducing a new variable. Let

$$y_i = \frac{1}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}}.$$

So that the objective function (3.5) becomes

$$\max \quad \sum_{i \in M} \sum_{j \in N} \sum_{k \in M} \lambda_i (\pi_j - \tau_{ikj})(z_{ikj} y_i \upsilon_{ikj}). \tag{3.11}$$

There still exists a nonlinear term which is $z_{ikj}y_i$. In order to linearize the term, let $t_{ikj} = z_{ikj}y_i$, a bilinear term, and to linearize the term, we can add the following constraints

$$t_{ikj} - y_i \leq M - Mz_{ikj}, \qquad 0 \leq t_{ikj} \leq y_i \qquad \text{and} \qquad t_{ikj} \leq Mz_{ikj}$$

where M is a sufficiently large number. We can replace M with an upper bound on $y := \dfrac{1}{\upsilon_{i0}}$. With these changes, the final MILP formulation becomes:

$$\max \qquad \sum_{i\in M}\sum_{j\in N}\sum_{k\in M} \lambda_i(\pi_j - \tau_{ikj})t_{ikj}\upsilon_{ikj}$$

$$\text{subject to} \tag{3.12}$$

$$\sum_{k\in M} z_{ikj} \leq 1, \qquad \forall i \in M, j \in N, \tag{3.13}$$

$$\sum_{j\in N} x_{ij} \leq \kappa_i, \qquad \forall i \in M, \tag{3.14}$$

$$x_{kj} \geq z_{ikj}, \qquad \forall i, k \in M, \forall j \in N, \tag{3.15}$$

$$y_i\upsilon_{i0} + \sum_{j\in N}\sum_{k\in M} t_{ikj}\upsilon_{ikj} = 1, \qquad \forall i \in M, \tag{3.16}$$

$$\upsilon_{i0}(y_i - t_{ikj}) \leq 1 - z_{ikj}, \qquad \forall i, k \in M, \forall j \in N, \tag{3.17}$$

$$\upsilon_{i0}t_{ikj} \leq z_{ikj}, \qquad \forall i, k \in M, \forall j \in N, \tag{3.18}$$

$$0 \leq t_{ikj} \leq y_i, \qquad \forall i, k \in M, \forall j \in N, \tag{3.19}$$

$$z_{ikj} \in \{0, 1\}, \qquad \forall i, k \in M, \forall j \in N, \tag{3.20}$$

$$x_{ij} \in \{0, 1\}, \qquad \forall i \in M, \forall j \in N, \tag{3.21}$$

$$y_i \geq 0, \qquad \forall i \in M, \tag{3.22}$$

$$t_{ikj} \geq 0, \qquad \forall i, k \in M, \forall j \in N. \tag{3.23}$$

The first three constraints (3.13, 3.14, 3.15) are the same with the base model. Constraint (3.16) is for the new variable introduced because of the linearization of the objective function. The succeeding three constraints (3.17,3.18,3.19) are for the linearization of the nonlinear term. In the formulation, z_{ikj} does not have to be defined as a binary variable, since the formulation will pick the closest location for location i.

In the next section, the poor performance of the MILP model is shown via numerical results. When the capacity constraints for fulfillment centers are tight, the problem cannot be solved for large instances. This problem is also stated in Bront et al. [32], Méndez-Díaz et al. [54], and Feldman and Topaloglu [50]. To overcome this, we will present a conic quadratic mixed integer problem formulation which is an efficient alternative to the MILP formulation. In the following section, the transformation of the base model to conic quadratic MIP form will be presented, and additional valid inequalities will be shown to strengthen the conic quadratic MIP formulation.

3.4 Conic Formulation

Lobo et al. [77] have defined second order cone programming, where the problem parameters are $A_i \in \mathbb{R}^{(n_i-1)\times n}$, $b_i \in \mathbb{R}^{n_i-1}$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, and $f \in \mathbb{R}^n$, as follows:

$$
\begin{aligned}
\min \quad & \boldsymbol{f}^T \boldsymbol{x} \\
\text{subject to} \quad & \\
& \|\boldsymbol{A}_i \boldsymbol{x} + \boldsymbol{b}_i\| \leq \boldsymbol{c}_i^T \boldsymbol{x} + d_i, \qquad i = 1, \ldots, N,
\end{aligned}
$$

where $x \in \mathbb{R}^n$ is the decision variable of the optimization problem. In the formulation, $\|\cdot\|$ refers to the L2 or Euclidean norm, i.e., $\|u\| = (u^T u)^{1/2}$. Considering the convex objective function and constraints, which define a convex set, this problem is called a convex programming problem. One can see Alizadeh and Goldfarb [78], and Lobo et al. [77] for detailed second order cone programming reviews.

Convex linear programs (LP), quadratic programs (QP), quadratically constrained programs (both LP and QP), problems with hyperbolic constraints, and non-linear convex optimization problems can be reformulated as a second order cone program. For hyperbolic constraints, we can use second order cone programming approach in the following way [77]:

$$\left\| \begin{bmatrix} 2w \\ x-y \end{bmatrix} \right\| \leq x+y \qquad \Longleftrightarrow \qquad w^T w \leq xy, \quad x \geq 0, \quad y \geq 0$$

For our conic reformulation, we use this hyperbolic constraint approach in our model. The approach above can be used only in convex minimization problems. Hence, we need to reformulate our objective function.

3.4.1 Objective Function

The objective function is

$$\max \sum_{i \in M} \lambda_i \left[\frac{\sum_{j \in N} \sum_{k \in M} (\pi_j - \tau_{ikj}) z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right].$$

Let $\delta_{ikj} = \pi_j - \tau_{ikj}$ be the profit of a product in location i, and $\overline{\delta}_i = \max_{j \in N, k \in M} \delta_{ikj}$ be the most profitable product in location i.

Our primary goal is to maximize the revenue, however, for conic formulation, we try to minimize the gap between the maximum possible revenue and realized revenue, i.e., deviation from the ideal case. Therefore, we can represent our problem as follows:

$$\min \sum_{i \in M} \lambda_i \left[\overline{\delta}_i - \frac{\overline{\delta}_i \upsilon_{i0}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} - \frac{\sum_{j \in N} \sum_{k \in M} (\overline{\delta}_i - \delta_{ikj}) z_{ikj} \upsilon_{ikj}}{\upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}} \right]. \tag{3.24}$$

The first term $\sum_{i \in M} \lambda_i \overline{\delta}_i$ corresponds to the highest revenue which can be possible selling only the most profitable products. One can check for the correctness of the objective function by summing the linear and conic objectives, resulting in the first term which is only a constant.

By adding $y_i = \dfrac{1}{\upsilon_{i0} + \sum\limits_{j \in N} \sum\limits_{k \in M} z_{ikj}\upsilon_{ikj}}$ and $t_{ikj} = z_{ikj} y_i$, the objective function becomes

$$\min \quad \sum_{i \in M} \lambda_i \bar{\delta}_i \upsilon_{i0} y_i + \sum_{i \in M} \sum_{j \in N} \sum_{k \in M} \lambda_i (\bar{\delta}_i - \delta_{ikj}) t_{ikj} \upsilon_{ikj}. \tag{3.25}$$

3.4.2 Constraints

$$\sum_{k \in M} z_{ikj} \le 1, \qquad \forall i \in M, j \in N, \tag{3.26}$$

$$\sum_{j \in N} x_{ij} \le \kappa_i, \qquad \forall i \in M, \tag{3.27}$$

$$x_{kj} \ge z_{ikj}, \qquad \forall i \in M, k \in M, j \in N, \tag{3.28}$$

$$y_i \ge \frac{1}{\upsilon_{i0} + \sum\limits_{j \in N} \sum\limits_{k \in M} z_{ikj}\upsilon_{ikj}}, \qquad \forall i \in M, \tag{3.29}$$

$$t_{ikj} \ge z_{ikj} y_i, \qquad \forall i, k \in M, \forall j \in N, \tag{3.30}$$

$$x_{ij} \in \{0, 1\}, \qquad \forall i \in M, \forall j \in N, \tag{3.31}$$

$$0 \le z_{ikj} \le 1, \qquad \forall i, k \in M, \forall j \in N, \tag{3.32}$$

$$y_i \ge 0, \qquad \forall i \in M, \tag{3.33}$$

$$t_{ikj} \ge 0, \qquad \forall i, k \in M, \forall j \in N. \tag{3.34}$$

$$\tag{3.35}$$

Let $w_i = \upsilon_{i0} + \sum\limits_{j \in N} \sum\limits_{k \in M} z_{ikj}\upsilon_{ikj}$. So, constraints (3.29) and (3.30) can be written in the cone form:

$$y_i w_i \ge 1, \qquad t_{ikj} w_i \ge z_{ikj}^2$$

Constraint (3.29) is going to be satisfied in the optimal solution as an equality. Constraint (3.30) can be written in the cone form because z_{ikj} is a binary decision variable, so that it always equals its square. Therefore, we change constraints with the following form:

$$y_i w_i \ge 1, \qquad \forall i \in M \tag{3.36}$$

$$t_{ikj} w_i \geq z_{ikj}^2, \qquad \forall i,k \in M, j \in N \tag{3.37}$$

To strengthen the continuous relaxation of the formulation, we also add the following constraint:

$$\upsilon_{i0} y_i + \sum_{k \in M} \sum_{j \in N} \upsilon_{ikj} z_{ikj} \geq 1 \tag{3.38}$$

3.4.3 Final Conic Quadratic MIP Formulation with different shipping costs

$$\min \quad \sum_{i \in M} \lambda_i \bar{\delta}_i \upsilon_{i0} y_i + \sum_{i \in M} \sum_{j \in N} \sum_{k \in M} \lambda_i (\bar{\delta}_i - \delta_{ikj}) t_{ikj} \upsilon_{ikj}$$

$$\text{subject to} \tag{3.39}$$

$$\sum_{k \in M} z_{ikj} \leq 1, \qquad \forall i \in M, j \in N, \tag{3.40}$$

$$\sum_{j \in N} x_{ij} \leq \kappa_i, \qquad \forall i \in M, \tag{3.41}$$

$$x_{kj} \geq z_{ikj}, \qquad \forall i,k \in M, j \in N, \tag{3.42}$$

$$w_i = \upsilon_{i0} + \sum_{j \in N} \sum_{k \in M} z_{ikj} \upsilon_{ikj}, \qquad \forall i \in M, \tag{3.43}$$

$$y_i w_i \geq 1, \qquad \forall i \in M, \tag{3.44}$$

$$t_{ikj} w_i \geq z_{ikj}^2 \qquad \forall i,k \in M, \forall j \in N, \tag{3.45}$$

$$\upsilon_{i0} y_i + \sum_{k \in M} \sum_{j \in N} \upsilon_{ikj} z_{ikj} \geq 1 \qquad \forall i,k \in M, \forall j \in N, \tag{3.46}$$

$$z_{ikj} \in \{0,1\}, \qquad \forall i,k \in M, \forall j \in N, \tag{3.47}$$

$$x_{ij} \in \{0,1\}, \qquad \forall i \in M, \forall j \in N, \tag{3.48}$$

$$y_i \geq 0, \qquad \forall i \in M, \tag{3.49}$$

$$t_{ikj} \geq 0, \qquad \forall i,k \in M, \forall j \in N, \tag{3.50}$$

$$w_i \geq 0, \qquad \forall i \in M. \tag{3.51}$$

In this case, z_{ikj} should be defined as a binary variable because of the conic formulation. In order to make z_{ikj} a continuous variable, one can add the following inequality to prevent the partial shipment between regions, i.e., to force z_{ikj} to be 1 if only if k is the location closest to i that carries product j and 0 otherwise.

$$z_{i\Phi_{iw}j} \geq x_{\Phi_{iw}j} - \sum_{k=1}^{w-1} x_{\Phi_{ik}j}, \qquad \forall i,k \in M, \forall j \in N. \tag{3.52}$$

3.4.4 Strengthening the Formulation with McCormick Inequalities

We can strengthen the formulation with valid McCormick inequalities [79] for the bilinear term t. To do this, we will determine lower and upper bounds on

$$y_i = \frac{1}{\upsilon_{i0} + \sum_{k\in M}\sum_{j\in N} \upsilon_{ikj} z_{ikj}}.$$

The following proposition states these bounds:

Proposition 1 *The following bounds on variables y_i , $i \in M$, are valid:*

$$y_i^u = \frac{1}{\upsilon_{i0}} \geq y_i \tag{3.53}$$

$$y_i^l = \frac{1}{\upsilon_{i0} + \sum_{k=1}^{|M|} \sum_{j=\boldsymbol{K}_{[k-1]_i}+1}^{\boldsymbol{K}_{[k]_i}} \upsilon_{i[k]_i[j]}} \leq y_i \tag{3.54}$$

where $\boldsymbol{K}_{[k]_i} = \sum_{\ell=1}^{k} \kappa_{[\ell]_i}$, $[k]_i$ is the kth closest region to ith region and $\upsilon_{i[k]_i[j]}$ is the jth largest of preferences in the kth closest region to region i. The lower bound (3.54) is attained when the denominator is maximum. This is possible when a customer has no purchase possibility (υ_{i0}) and the first $\kappa = \sum_{i\in M} \kappa_i$ highest preference value products are available.

Proposition 2 *The following conditional bounds on variables y_i , $i \in M$ are valid:*

$$z_{ikj} = 0 \Rightarrow \begin{cases} y^u_{i|z_{ikj}=0} = \frac{1}{\upsilon_{i0}} & \geq y_i \\ y^l_{i|z_{ikj}=0} = \frac{1}{\upsilon_{i0} + f_{i|z_{ikj}=0}} & \leq y_i \end{cases}$$

$$z_{ikj} = 1 \Rightarrow \begin{cases} y^u_{i|z_{ikj}=1} = \frac{1}{\upsilon_{i0} + \upsilon_{ikj}} & \geq y_i \\ y^l_{i|z_{ikj}=0} = \frac{1}{\upsilon_{i0} + f_{i|z_{ikj}=1}} & \leq y_i \end{cases}$$

where $f_{i|z_{ikj}=0}$ and $f_{i|z_{ikj}=1}$ are defined as follows:

$$f_{i|z_{ikj}=0} = \max \sum_{w \in M} \sum_{l \in N} \upsilon_{iwl} x_{wl} \tag{3.55}$$

s.t.

$$x_{kj} = 0$$

$$\sum_{l \in N} x_{wl} \leq \kappa_w \qquad \forall w \in M$$

$$f_{i|z_{ikj}=1} = \max \sum_{w \in M} \sum_{l \in N} \upsilon_{iwl} x_{wl} \tag{3.56}$$

s.t.

$$x_{kj} = 1$$

$$\sum_{l \in N} x_{wl} \leq \kappa_w \qquad \forall w \in M$$

One can easily solve the maximization problems (3.55) and (3.56) in a straight forward manner.

In Proposition 2, when the product j in region i for a customer in region k is not available, i.e., $z_{ikj} = 0$, the lower bound of y_i can be obtained by adding the first κ highest preference products. The upper bound is straightforward. When $z_{ikj} = 1$, the lower bound is calculated with the $\kappa - 1$ highest preference products without product j. This is because when j is in the assorment, we have 1 less capacity to fill. For this case, since the product j is in the assortment, also, we need to add it for the upper bound.

With propositions (1, 2), for our bilinear term t_{ikj} in the formulation, we can add the following valid McCormick inequalities:

$$t_{ikj} \leq y^u_{i|z_{ikj}=1} z_{ikj}, \qquad \forall i, k \in M, j \in N \tag{3.57}$$

$$t_{ikj} \geq y^l_{i|z_{ikj}=1} z_{ikj}, \qquad \forall i, k \in M, j \in N \tag{3.58}$$

$$t_{ikj} \geq y_i - y^u_{i|z_{ikj}=0}(1 - z_{ikj}), \qquad \forall i, k \in M, j \in N \tag{3.59}$$

$$t_{ikj} \leq y_i - y^u_{i|z_{ikj}=0}(1 - z_{ikj}), \qquad \forall i, k \in M, j \in N \tag{3.60}$$

Chapter 4

Numerical Study

In this chapter, we conduct a numerical study based on the formulations given in Chapter 3 to test the conic programming approach and McCormick estimators' effectiveness. We compare the results of the conic and mixed integer linear formulations on a set of parameters. The impacts of different capacity and no-purchase preference are discussed.

The parameter set is generated as follows. The set consists of 4 customer segments/regions and 40 products, i.e., $|M| = 4, |N| = 40$. The probability of a customer coming from customer segment i is equal for all $i \in M$, i.e., the probability of a demand originated from customer segment i is $\lambda_i = \frac{1}{4}$. The price of the products are randomly generated from a continuous uniform distribution between 5 and 6. All customer segments have the same price for all products. The cost of shipping products are determined based on the distances among the regions. The shipping cost matrix can be seen in Table 4.1. In the table, there are two different cost settings (separated by /) which are used in the numerical studies.

If any product is shipped to another region, then it causes a decrease in its preference. Thus, v_{ikj} is decreased with respect to its distance to the current region. It is not product based. Decrease percentages can be viewed in Table 4.2.

Table 4.1: Cost Matrix

τ_{ik}	**Region 1**	**Region 2**	**Region 3**	**Region 4**
Region 1	0	2 / 1.0	1 / 0.5	3 / 1.5
Region 2	2 / 1.0	0	3 / 1.5	1 / 0.5
Region 3	1 / 0.5	3 / 1.5	0	2 / 1.0
Region 4	3 / 1.5	1 / 0.5	2 / 1.0	0

In the table, two different γ sets are presented (separated by /) and used in the numerical studies.

Table 4.2: γ Matrix

γ_{ik}	**Region 1**	**Region 2**	**Region 3**	**Region 4**
Region 1	1	0.90 / 0.70	0.95 / 0.90	0.85 / 0.60
Region 2	0.90 / 0.70	1	0.85 / 0.60	0.95 / 0.90
Region 3	0.95 / 0.90	0.85 / 0.60	1	0.90 / 0.70
Region 4	0.85 / 0.60	0.95 / 0.90	0.90 / 0.70	1

The preferences, υ_{ikj}, are defined in the following way:

$$\upsilon_{ij} \sim U(0,1) \quad \forall i \in M, j \in N,$$

$$\upsilon_{ikj} = \gamma_{ik}\upsilon_{ij},$$

where γ_{ik} is a discount factor for the preferences and decreases as the lead time increases, meaning that the preference of a product decreases as the lead time of that product increases. For the numerical study, no-purchase preferences are taken the same for all customer segments and three different no-purchase preferences are considered (10, 20, 50). For the capacities of the fulfillment centers, again, three different scenarios are studied where $\sum_{i=1}^{4} \kappa_i = \kappa \in (20, 32, 40)$, and in each study, all fulfillment centers are assumed to have equal capacities. As a result, for the numerical study, we have observed 4 different parameter settings (cost and discount factor) with 9 combinations of capacity and no-purchase pairs and 5 different replications, resulting in 180 instances.

To observe the impact of our formulation, instances are solved on the commercial solver Gurobi 64bit 7.5.1 on Windows 10 with the default settings of

Gurobi, except that for the continuous relaxation of conic programs, the outer-approximation approach is used. The specifications of the used computer are Intel Core i7-4710MQ 2.50 GHz processor, 16 GB memory and Windows 10 operating system. For each scenario, the time limit is set as 1800 seconds.

4.1 Results

In this section, we are going to present our numerical studies and interpret the results. For each parameter setting, we have organized a table showing the average results of the optimization problem. Each table starts with the changing parameters (v_0, κ) pair. In the "Assort/bind" column, the average of the size of the assortment and the number of instances (out of 5) that have full size assortment are showed. Each table represents the performances of four different formulations (MILP, MILP+MC, CONIC, CONIC+MC). Each formulation has two columns and four entries which are "Root Gap", "End Gap", "Time/#" and "Nodes". The "Root Gap" column gives the percentage gap between the objective value of the continuous relaxation (z_{root}) and the objective value of the optimal integer solution (z_{opt}), i.e., Root Gap $= 100 \times |\frac{(z_{opt} - z_{root})}{z_{opt}}|$. Similarly, the "End Gap" represents the percentage gap between the best lower bound at termination (z_{bb}) and the objective value of the optimal integer solution (z_{opt}), i.e., End Gap $= 100 \times |\frac{(z_{opt} - z_{bb})}{z_{opt}}|$. Therefore, the "End Gap" value equals to 0 if an instance is solved within the time limit. The second column of each formulation shows both "Time/#" and "Nodes" corresponding to the average solution time, if applicable, number of optimally solved instances, and the number of nodes explored, respectively. However, if an instance cannot be solved with any of the methods within the time limit, the best integer solution is taken as a solution.

The tables are arranged according to their parameter settings:

(I). $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma = (1.0, 0.95, 0.90, 0.85)$,

(II). $\tau = (0.0, 0.5, 1.0, 1.5)$, $\gamma = (1.0, 0.95, 0.90, 0.85)$, (III). $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma = (1.0, 0.90, 0.70, 0.60)$,

(IV). $\tau = (0.0, 0.5, 1.0, 1.5)$, $\gamma = (1.0, 0.90, 0.70, 0.60)$,

where τ is the cost of shipping to regions and γ is the discount factor settings for the regions. The parameter setting τ is applied to all regions. The first entry of τ shows the cost of shipping a product to the closest (including itself) region. Considering Setting (I), the cost of shipping a product from a region to the second or third closest region is set as 1.0 or 2.0, respectively. Furthermore, shipping a product to the most distant region from any region is set as 3.0. The parameter setting γ is, again, applied to all regions. The first entry of the setting states that the preference of a product does not decrease if it is shipped within the same region. However, in Setting (I), if a product is sent from the second closest region, its preference is decreased to 0.95% of its original preference. In each table, each instance's field, which consists of two rows, gives a summary of five samples of a (υ_0, κ) pair.

Table 4.3: Results for (I): Sample averages for problems with the setting $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma = (1.0, 0.95, 0.90, 0.85)$.

			MILP		MILP + MC		CONIC		CONIC + MC	
υ_0	κ	Assort bind	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes
10	20	20.0	60.79%	– / 0	9.94%	– / 0	15.98%	– / 0	2.46%	621.17 / 3
		5	42.72%	1,247,987	6.25%	434,269	2.73%	950,693	0.15%	1,060,082
	32	27.2	44.21%	– / 0	7.32%	– / 0	19.80%	– / 0	3.03%	– / 0
		0	34.62%	686,440	5.14%	592,123	1.69%	862,623	1.05%	1,161,374
	40	30.6	35.00%	– / 0	5.15%	– / 0	20.29%	– / 0	2.61%	– / 0
		0	27.55%	399,324	3.55%	654,429	1.20%	984,782	0.90%	1,087,537
20	20	20.0	51.25%	– / 0	6.73%	– / 0	8.99%	– / 0	0.82%	44.95 / 5
		5	26.36%	3,144,312	3.69%	636,278	0.63%	1,747,253	0.00%	68,425
	32	31.4	60.59%	– / 0	4.18%	– / 0	11.75%	– / 0	0.79%	466.91 / 5
		3	42.41%	1,149,441	2.08%	441,247	0.39%	1,730,207	0.01%	459,995
	40	36.4	55.69%	– / 0	2.54%	– / 0	12.43%	496.09 / 3	0.53%	93.09 / 5
		0	42.34%	447,242	0.99%	566,168	0.05%	879,568	0.01%	83,623
50	20	20.0	21.56%	– / 0	3.81%	1,016.24 / 3	3.90%	99.45 / 5	0.15%	5.45 / 5
		5	8.15%	3,699,325	0.53%	865,504	0.01%	153,473	0.00%	3,036
	32	32.0	30.58%	– / 0	1.79%	1,027.08 / 1	5.20%	66.85 / 5	0.12%	6.61 / 5
		5	18.17%	2,369,983	0.42%	684,782	0.01%	88,913	0.01%	3,222
	40	38.2	34.50%	– / 0	0.67%	138.39 / 5	5.61%	11.87 / 5	0.07%	4.39 / 5
		0	23.16%	1,918,076	0.01%	44,694	0.01%	22,108	0.01%	758

Table 4.3 presents the results of the case where $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma =$

$(1.0, 0.95, 0.90, 0.85)$, e.g., if a product is shipped from a region to the most distant region, it would cost 3 units and that product's preference is decreased to 85% of its original preference value. In the table, when the MILP formulation is considered, the root gap reaches from 21.56% to 60.79%. Moreover, with the MILP formulation, there is no instance that can be solved within the time limit. The remaining gaps at the termination are considerably large which can reach from 8.15% to 42.72%. Adding the McCormick estimators to the MILP formulation (MILP+MC) ends up a significant improvement on the performance of the MILP formulation. The formulation is able to solve only 9 instances which have $\upsilon_0 = 50$. The improvement can also be seen in the root and end gap values. The maximum root gap value is decreased to 9.94% and the maximum end gap value is decreased to 6.25% which are still beyond acceptable.

With the CONIC formulation, 18 instances can be solved within the time limit, and when the unsolved ones are considered, end gap rates are significantly better than the MILP formulation, especially in the instances of larger no-purchase preference and capacities. Strengthening the CONIC formulation with McCormick estimators (CONIC+MC) creates the best model in terms of obtained results. With the CONIC+MC formulation, 33 instances can be solved within the time limit, and the root gaps are the smallest ones among the formulations.

In Table 4.4, the discount factor set is specified as $\gamma = (1.0, 0.90, 0.70, 0.60)$, meaning that customers are more sensitive to the lead time of products. For the setting, MILP formulation cannot solve any instance similar to Setting (I), and what is worse is that the maximum root gap and end gap are increased to 66.33% and 52.22%, respectively. Strengthening the MILP formulation with McCormick estimators (MILP+MC) leading to better results with respect to the MILP ones. The maximum root and end gap have declined to 6.93% and 5.37%, respectively, which are less than Setting (I). According to the results of the Table 4.3 and 4.4, the MILP formulation has difficulty in solving Setting (II). However, adding McCormick estimators to the MILP formulation works better for Setting (II), meaning that the estimators leading a more efficient cuts than Setting (I).

The CONIC formulation performs better but is still unable to solve instances

Table 4.4: Results for (II): Sample averages for problems with the setting $\tau = (0.0, 0.5, 1.0, 1.5)$, $\gamma = (1.0, 0.95, 0.90, 0.85)$.

			MILP		MILP + MC		CONIC		CONIC + MC	
υ_0	κ	Assort bind	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes
10	20	20.0	61.25%	– / 0	6.87%	– / 0	11.05%	– / 0	1.54%	317.55 / 5
		5	47.73%	1,516,260	5.37%	545,318	1.42%	1,207,219	0.01%	453,216
	32	32.0	39.92%	– / 0	3.08%	– / 0	14.50%	– / 0	1.18%	742.86 / 3
		5	37.23%	998,980	2.09%	438,219	0.92%	1,457,892	0.13%	1,026,552
	40	37.2	33.43%	– / 0	1.61%	– / 0	15.74%	1,180.57 / 1	0.54%	144.43 / 5
		1	31.12%	1,200,267	0.84%	465,088	0.18%	1,878,811	0.01%	69,265
20	20	20.0	53.45%	– / 0	6.93%	– / 0	6.00%	885.79 / 1	0.55%	26.35 / 5
		5	29.32%	3,021,645	4.97%	795,920	0.28%	1,925,793	0.01%	43,754
	32	32.0	66.33%	– / 0	2.67%	– / 0	8.02%	1,210.15 / 2	0.44%	103.53 / 5
		5	49.01%	1,481,809	1.86%	546,370	0.14%	1,980,747	0.01%	136,644
	40	39.0	59.69%	– / 0	0.59%	– / 0	8.78%	98.40 / 5	0.13%	10.80 / 5
		1	52.22%	1,045,621	0.19%	413,918	0.01%	119,655	0.01%	1,991
50	20	20.0	22.01%	– / 0	4.95%	– / 0	2.52%	58.02 / 5	0.10%	3.25 / 5
		5	9.88%	3,956,481	2.31%	996,835	0.01%	91,749	0.00%	1,267
	32	32.0	31.38%	– / 0	2.18%	– / 0	3.40%	35.94 / 5	0.08%	4.41 / 5
		5	20.02%	2,571,540	1.17%	736,863	0.01%	51,863	0.01%	2,781
	40	39.4	35.55%	– / 0	0.40%	229.72 / 3	3.73%	775.20 / 2	0.03%	4.25 / 5
		3	25.70%	1,987,478	0.05%	236,320	0.02%	2,505,765	0.01%	433

where $\upsilon_0 = 10, \kappa \in (20, 32)$. The maximum root gap and end gap are 15.74% and 1.42% which is significantly better than the MILP formulation. Including McCormick estimators in the CONIC formulation (CONIC+MC) works best. The CONIC+MC formulation can solve 42 out of 45 instances within time limit. The maximum root gap is decreased to 1.54% and the maximum end gap is 0.13%.

The average results of Setting (III) can be seen in Table 4.5. The maximum root and end gaps are 61.70% and 42.57% , respectively, for the MILP formulation. No instances can be solved within the time limit, and the results at the termination has quite large remaining gaps. Adding the McCormick estimators to the MILP formulation (MILP+MC) improves the performance of the MILP formulation by decreasing the maximum root gap to the 8.76% and end gap to the 5.46%. The MILP+MC formulation can solve 8 instances within the time limit, especially the instances where $\upsilon_0 = 50$.

The CONIC formulation performs better than the MILP formulation in terms of root gap (max. 16.11%) and end gap (max. 2.56%). The formulation can solve

Table 4.5: Results for (III): Sample averages for problems with the setting $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma = (1.0, 0.90, 0.70, 0.60)$.

			MILP		MILP + MC		CONIC		CONIC + MC	
υ_0	κ	Assort bind	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes
10	20	20.0	61.70%	– / 0	8.76%	– / 0	12.66%	– / 0	1.83%	344.88 / 5
		5	39.34%	1,216,542	5.46%	474,674	2.56%	729,110	0.01%	396,058
	32	28.0	45.57%	– / 0	6.32%	– / 0	15.78%	– / 0	2.29%	– / 0
		0	36.97%	846,588	4.27%	626,117	1.82%	768,192	0.59%	1,237,485
	40	31.0	35.92%	– / 0	4.43%	– / 0	16.11%	– / 0	2.01%	– / 0
		0	25.64%	371,582	2.89%	783,841	1.28%	880,077	0.48%	1,135,556
20	20	20.0	46.58%	– / 0	5.91%	– / 0	6.87%	– / 0	0.58%	20.51 / 5
		5	21.88%	3,187,744	2.93%	686,718	0.51%	1,224,196	0.00%	19,263
	32	31.4	56.82%	– / 0	3.97%	– / 0	9.03%	– / 0	0.66%	203.86 / 5
		2	39.66%	1,764,573	2.09%	500,986	0.53%	1,495,241	0.01%	233,660
	40	35.8	55.69%	– / 0	2.51%	– / 0	9.43%	– / 0	0.48%	92.46 / 5
		0	42.57%	1,180,486	1.14%	633,486	0.24%	1,682,886	0.01%	64,805
50	20	20.0	19.18%	– / 0	3.09%	556.53 / 4	2.89%	76.69 / 5	0.09%	2.92 / 5
		5	6.25%	3,503,959	0.15%	602,077	0.01%	96,186	0.00%	604
	32	32.0	27.19%	– / 0	1.84%	449.58 / 1	3.85%	121.37 / 5	0.10%	4.91 / 5
		4	14.84%	2,370,772	0.32%	754,673	0.01%	173,929	0.01%	1,515
	40	37.6	30.70%	– / 0	1.03%	598.53 / 3	4.08%	20.19 / 5	0.08%	5.74 / 5
		0	18.80%	1,963,729	0.09%	428,430	0.01%	19,530	0.01%	1,323

15 instances, however, it has difficulty in solving instances where $\upsilon_0 \in (10, 20)$ regardless of the capacity constraint. The CONIC formulation strengthened with McCormick estimators (CONIC+MC) shows the best performance. The maximum root gap and end gap have declined to 2.29% and 0.59%, respectively. 35 instances can be solved optimally with the CONIC+MC formulation. For Setting (III), the instances where $\upsilon_0 = 10, \kappa \in (32, 40)$ cannot be solved with any formulations.

Table 4.6 presents the average results of the setting in which $\tau = (0.0, 0.5, 1.0, 1.5)$ and $\gamma = (1.0, 0.90, 0.70, 0.60)$. For this setting, the shipping costs are decreased, and customers are more sensitive to lead times of products. For the MILP formulation, again, no instances can be solved, and the maximum root and end gap are found as 63.93% and 47.87%, respectively. The results are not satisfactory because of the remaining gaps at the termination. With McCormick estimators in the MILP formulation (MILP+MC), the root and end gaps are decreased to 6.51% and 4.93%. However, only one instance of the parameter setting ($\upsilon_0 = 50$ and $\kappa = 20$) can be solved within the time limit.

Table 4.6: Results for (IV): Sample averages for problems with the setting $\tau = (0.0, 0.5, 1.0, 1.5)$, $\gamma = (1.0, 0.90, 0.70, 0.60)$.

υ_0	κ		MILP		MILP + MC		CONIC		CONIC + MC	
		Assort bind	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes	Root Gap End Gap	Time/# Nodes
10	20	20.0	63.93%	– / 0	6.51%	– / 0	8.49%	– / 0	7.44%	61.91 / 5
		5	45.42%	1,749,339	4.93%	539,131	1.13%	853,079	0.00%	72,864
	32	32.0	43.05%	– / 0	3.26%	– / 0	11.35%	– / 0	3.44%	429.11 / 4
		0	39.95%	978,030	2.27%	433,571	1.11%	1,196,520	0.04%	603,385
	40	36.2	35.39%	– / 0	1.96%	– / 0	12.29%	– / 0	1.53%	544.53 / 5
		0	32.94%	1,162,921	1.18%	504,839	0.78%	1,420,559	0.01%	395,885
20	20	20.0	47.73%	– / 0	6.19%	– / 0	4.45%	369.63 / 2	7.18%	11.53 / 5
		5	24.09%	3,085,350	3.85%	861,695	0.11%	1,133,729	0.00%	8,075
	32	32.0	61.09%	– / 0	2.89%	– / 0	5.73%	666.02 / 1	2.30%	30.75 / 5
		2	43.12%	1,738,567	1.93%	553,989	0.18%	1,719,477	0.01%	32,322
	40	38.8	60.99%	– / 0	1.29%	– / 0	5.49%	1,274.02 / 3	1.05%	19.58 / 5
		0	47.87%	1,345,300	0.71%	482,634	0.06%	1,831,602	0.01%	4,841
50	20	20.0	19.52%	– / 0	3.78%	1,755.50 / 1	8.01%	16.91	5.67%	2.38 / 5
		5	7.40%	3,719,310	0.87%	1,122,968	0.00%	14,656 / 5	0.00%	224
	32	32.0	27.82%	– / 0	2.09%	– / 0	2.62%	24.98	1.16%	3.48 / 5
		4	16.51%	2,457,795	0.88%	719,786	0.00%	28,652 / 5	0.00%	573
	40	39.0	31.62%	– / 0	0.93%	– / 0	2.11%	9.08	0.57%	3.42 / 5
		0	21.21%	1,943,073	0.24%	575,308	0.00%	3,947 / 5	0.01%	13

Unlike the linear formulations, the CONIC formulation shows better performance than the MILP formulation. 21 instances, consisting $\upsilon_0 \in (20, 50)$ settings, can be solved within the time limit. The maximum root and end gaps declined to 12.29% and 1.13%, respectively. Adding the McCormick estimators to the CONIC formulation (CONIC+MC) can solve 44 instances out of 45, and works best in terms of all fields in the table. With the stronger formulation, the maximum root gap is decreased to 7.44% and the maximum end gap is declined to 0.04%. The optimal solutions can be found with less number of nodes compared to the MILP formulation with a success rate of up to 99.99%.

The comparison among scenarios can be seen in Table 4.7, recalling the scenarios: The tables are arranged according to their parameter settings:

(I). $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma = (1.0, 0.95, 0.90, 0.85)$,

(II). $\tau = (0.0, 0.5, 1.0, 1.5)$, $\gamma = (1.0, 0.95, 0.90, 0.85)$,

(III). $\tau = (0.0, 1.0, 2.0, 3.0)$, $\gamma = (1.0, 0.90, 0.70, 0.60)$,

Table 4.7: Average optimal assortment sizes and expected revenue for each parameter setting.

Scenario	υ_0	κ	$\|S\|$	Expected Revenue	Scenario	υ_0	κ	$\|S\|$	Expected Revenue
(I)	10	20	20.0	2.34	(II)	10	20	20.0	2.64
		32	27.2	2.76			32	32.0	3.07
		40	30.6	2.94			40	37.2	3.23
	20	20	20.0	1.59		20	20	20.0	1.79
		32	31.4	1.98			32	32.0	2.22
		40	36.4	2.14			40	39.0	2.38
	50	20	20.0	0.81		50	20	20.0	0.97
		32	32.0	1.08			32	32.0	1.21
		40	38.2	1.20			40	39.4	1.34
(III)	10	20	20.0	2.29	(IV)	10	20	20.0	2.53
		32	28.0	2.72			32	32.0	2.98
		40	31.0	2.92			40	36.2	3.15
	20	20	20.0	1.52		20	20	20.0	1.69
		32	31.4	1.92			32	32.0	2.11
		40	35.8	2.09			40	38.8	2.29
	50	20	20.0	0.76		50	20	20.0	0.84
		32	32.0	1.02			32	32.0	1.13
		40	37.6	1.15			40	39.0	1.26

(IV). $\tau = (0.0, 0.5, 1.0, 1.5)$, $\gamma = (1.0, 0.90, 0.70, 0.60)$.

The optimal solutions can differ for each setting in terms of assortment sizes. For all parameter setting, and all combinations of no-purchase preference values, optimal assortments use the full size when $\kappa = 20$. Decreasing the cost of shipping increases the average size of the assortments. Moreover, when $\upsilon_0 \geq 20$, as customers become more sensitive to the lead times of products, the average optimal assortment size would decrease. However, when $\upsilon_0 = 10$, we cannot jump to this conclusion, so there is no certain behavior for this case. When we analyze the expected revenue per customer values, if the cost of shipping a product is decreased, the expected revenue of a customer would increase, and as the customers become more sensitive to the lead time of a product, the expected revenue per customer decreases.

Chapter 5

Case Study

In Chapter 3, we proposed a new formulation for the assortment optimization problem and showed its efficiency with randomly generated numerical results in Chapter 4. In this chapter, we will analyze a case study using real transactional data from a grocery store. First we will fit two MNL models to the data to estimate parameters of the attributes of the products. Then, we will solve an assortment optimization problem with different settings of no-purchase preference v_0 and discounting factor γ.

5.1 Grocery Store Data

The data is collected from a local grocery store chain with four stores in Northwestern Turkey. The chain sells many products in many categories in its four stores. For this study, we only consider two of the stores, one of them is small and the other one is large. We consider those two stores because they are geographically close enough to allow for one store's customers to also shop in the other store if necessary, enabling us to study the effect of distance (lead time) using our models. The small store can carry a smaller assortment due to capacity constraints. We assume that if a customer cannot find a product in one of the stores, then she/he might be willing to go to the other store for that product (with an obvious decrease in his/her utility). This assumption is a realistic assumption since the distance between those two stores is less than 150 meters and a customer can reach a product within a reasonable amount of time if the product is not available in the store she/he previously visited but available in the other store. There are many categories or products on sale in these two stores. For this study, we focus on the detergent category which has 36 different products. The chain has provided us with daily sales transaction and inventory status data covering for about two years, from January 2016 to December 2017. When we analyze the data, we have observed that the detergent category has been offered in the small store only for the last 6 months of the data. Therefore, we have decided to use the last 4 months of the data in our analysis. In that time interval, there were 9748 sales where 584 sales were in the small store, 9164 sales were in the large store. The company also supplied its inventory records, so that we can determine the actual assortment when a customer arrives to any of these stores.

5.2 Estimation of Attributes

With the detergent sales transactional and inventory data, first, we have determined the attributes of each product, then fit two MNL models based on the choices of customers. For each product, we determined that the following attributes are relevant for consumer choice: price, size, powder/liquid, brand, for white, black or colored clothes, and perfume. The Table 5.1 shows the attributes and corresponding values/ranges.

Table 5.1: Attributes of products.

Attribute	Possible Value/Range
Price(TL)	Between 4.49 and 23.90 TL
Brand	Alo, Ariel, Bingo, Dalan, Dalin, Etimatik, H.Sakir, Omo, Persil, Perwoll, Tursil
Size(kg)	(0.5, 8.0)
Powder/Liquid	(0: Powder, 1: Liquid)
White Clothes	(0: Not for White Clothes, 1: For White Clothes)
Black Clothes	(0: Not for Black Clothes, 1: For Black Clothes)
Colored Clothes	(0: Not for Colored Clothes, 1: For Colored Clothes)
Perfume	(0: Not Perfume, 1: Perfume)

The deterministic part of a customer's utility on each product can be written as

$$
\begin{aligned}
u_j = \beta^T X_j = \;& \beta_0 + \beta_{PRICE} x_{j1} + \beta_{ARIEL} x_{j2} + \beta_{BINGO} x_{j3} + \beta_{DALAN} x_{j4} + \\
& \beta_{DALIN} x_{j5} + \beta_{ETIMATIK} x_{j6} + \beta_{HSAKIR} x_{j7} + \beta_{OMO} x_{j8} + \\
& \beta_{PERSIL} x_{j9} + \beta_{PERWOLL} x_{j10} + \beta_{TURSIL} x_{j11} + \beta_{size} x_{j12} + \\
& \beta_{POWDER} x_{j13} + \beta_{WHITE} x_{j14} + \beta_{BLACK} x_{j15} + \beta_{COLORED} x_{j16} + \\
& \beta_{PERFUME} x_{j17},
\end{aligned}
$$

where $\beta_k = (\beta_{PRICE}, \ldots, \beta_{PERFUME})$ is the parameter vector for the MNL model and $X_j = (x_{j1}, \ldots, x_{j17})$ is the product attribute vector for product $j, j = 1, \ldots, 36$

In order to estimate the parameters of the attributes, i.e., vector β, we can use the likelihood theory. Our likelihood function, which based on the MNL model, can be written as the follows:

$$
L(\beta) = \prod_{i=1}^{n} \frac{e^{X_j \beta}}{\sum_{k \in S_i} e^{X_k \beta}}, \tag{5.1}
$$

where $i = 1, 2, \ldots, n$ is the customers, S_i is the assortment for customer i as he/she arrives at the store. For computational reasons, we can take the logarithm of the likelihood function (5.1) which yields to

$$
\ell(\beta) = \sum_{i=1}^{n} \log \left(\frac{e^{X_j \beta}}{\sum_{k \in S_i} e^{X_k \beta}} \right) = \sum_{i=1}^{n} \left(X_j \beta - \log \sum_{k \in S_i} e^{X_k \beta} \right). \tag{5.2}
$$

The log-likelihood function $\ell(\beta)$ is concave [13]. Therefore, we can solve the maximization problem $\max_{\beta} \ell(\beta)$. To analyze data and solve this problem, we used R (version 3.5.1) [80], and as a solver, we used optim() function. Originally, optim heuristically performs minimization, however there is an option for maximization. As solving methods, optim has different options which are Nelder and Meat [81], BFGS, conjugate gradients [82] and L-BFGS-B [83]. One can see for detailed discussion of methods in [84].

For our study, we have used BFGS method, also known as a variable metric algorithm, which is a quasi-Newton method. The method evaluates function values and its gradients to reach the global optima. Since the function works heuristically, there is a risk for convergence to local optimum point. To handle this issue, we have started 50 different initial points and performed optimization. When we have analyzed the results, each element in vector β differs less than 0.001. For evaluating the preferences of each product, we have used the average of the 50 results as a β vector. Table 5.2 shows vector β. We determined a baseline product whose brand is Alo and type is liquid.

Table 5.2: MNL Model Estimates for Small and Large Stores

Store / β	Small	Large
β_0	0.0891	0.0891
Price	-0.1134	-0.1134
Brand:Ariel	0.5420	0.3804
Brand:Bingo	1.5436	0.6310
Brand:Dalan	-0.2689	-0.4605
Brand:Dalin	0.5257	0.0292
Brand:Etimatik	0.3550	-0.1705
Brand:H.Sakir	0.0885	-0.1786
Brand:Omo	-0.0381	0.0933
Brand:Persil	1.7048	0.6940
Brand:Perwoll	0.2605	0.2610
Brand:Tursil	0.6587	0.0016
Size	-0.0291	0.1239
Powder	-0.3307	0.2686
White Clothes	-0.1624	-0.2422
Black Clothes	0.1217	0.2323
Colored Clothes	0.2254	-0.1346
Perfume	-0.1462	0.1425

In the table, the coefficient of each stores' MNL model is shown, including the base utility β_0. We can interpret that as the price goes up, the utility of the product decreases. Moreover, the utility increases if a product belongs to Ariel. Same interpretations can be done via Table 5.2.

For the goodness of the model, we perform a Likelihood Ratio Test (LRT) where we test our proposed model with the null model. In the LRT, the likelihood ratio $\mathcal{D}$ is calculated as

$$\mathcal{D} = -2\log\frac{L_0}{L_1}$$

where L_0 is the likelihood value of the null model and L_1 is the likelihood value of the proposed model. The computed value is approximated to the χ^2 distribution. If the computed value is larger than the χ^2 value with significance level $100(1-\alpha)\%$ and k degrees of freedom, we reject the null model. Also, it is possible to verify the model with looking to the p-value of the model where the p-value is calculated based on the LRT. We reject the null model if the p-value of the proposed model is less than significance level α. For our study, we determined our significance level as a 1%, performed the LRT for MNL models we have proposed, and concluded two p-values $(3.8621\text{e-}10, 1.7284\text{e-}10)$ for the small and large stores, respectively, which are almost zero. Therefore, we reject null model for the two stores, concluding that we can use the model above for our general choice model problem.

5.3 Numerical Results

In the previous section, we have estimated each attribute's coefficient, so that we can compute the estimates of each product's preferences. In this section, we will solve an assortment optimization problem with the calculated preferences of the products under different no-purchase preference υ_0 and discount parameter γ.

For the numerical studies, we have five different values for the no-purchase parameter $\upsilon_0 = (1, 5, 10, 20, 50)$ and four different values for the discount factor $\gamma = (0.25, 0.50, 0.75, 0.90)$. Moreover, these parameters are taken as the same for both stores, i.e.,

$$\upsilon_{i0} = \upsilon_0, \quad \forall i = (1, 2); \qquad \gamma_{ik} = \gamma, \quad \forall i, k = (1, 2),$$

where $1, 2$ correspond to the small and large stores, respectively. Therefore, there are $5 \times 4 = 20$ combinations. Additionally, for the small store, we have analyzed two different capacity constraints $\kappa_1 = (5, 10)$.

In the case study, our aim is to find optimal assortment that maximizes the expected revenue per customer for each location at the same time. The objective function is the following

$$\max \left[\lambda_1 \frac{\sum_{j\in S_1} r_j \upsilon_{1j} + \sum_{j\in S_2/S_1} r_j \upsilon_{1j}\gamma}{\upsilon_0 + \sum_{j\in S_1} \upsilon_{1j} + \sum_{j\in S_2/S_1} \upsilon_{1j}\gamma} + \lambda_2 \frac{\sum_{j\in S_2} r_j \upsilon_{2j} + \sum_{j\in S_1/S_2} r_j \upsilon_{2j}\gamma}{\upsilon_0 + \sum_{j\in S_2} \upsilon_{2j} + \sum_{j\in S_1/S_2} \upsilon_{2j}\gamma} \right],$$

where γ is a discount factor for lead time, (λ_1, λ_2) are the probability of customer coming to the store $(1, 2)$, respectively, and υ_0 is the no-purchase preference. For the study, the parameters (λ_1, λ_2) are taken as $(\frac{584}{9748}, \frac{9164}{9748})$ which is the ratio of the sales in the original data. All instances are solved on the commercial solver Gurobi 64bit 7.5.1 on Windows 10 with the default settings. The specifications of the used computer are Intel Core i7-4710MQ 2.50 GHz processor, 16 GB memory and Windows 10 operating system. For each scenario, the time limit is determined as 1800 seconds.

We have analyzed five different cases: the performances of the formulations when $\kappa_1 = 10$, the CONIC+MC formulation results when $\kappa_1 = 10$, the heatmap of the expected revenue per customer, the heatmap of the cross-shopping rate of customers in Store 1, the benefit of the enabling transfers between stores and the CONIC+MC formulation results when $\kappa_1 = 5$.

The results can be seen in Table 5.3. The first two columns represent the no-purchase preference and discount factor. Following columns consist of four different formulations for the optimization problem. For each model, we have reported the number of iterations, the number of nodes to explore and run time to solve that instance. Moreover, if an instance cannot be solved in 1800 seconds, we have reported the end gap which is computed as $100 \times \frac{|(z_{\text{opt}} - z_{\text{bb}})|}{|z_{\text{opt}}|}$, where z_{bb} is the value of best bound at termination and z_{opt} is the value of the optimal integer solution.

As in the literature and in §4, the MILP formulation performs poorly when the discount factor is high. Even though the problem is relatively small, for 8 instances, the MILP formulation cannot solve the problem within the time limit. The excessive branching caused by weak relaxation leads to unsatisfactory performance. When the McCormick estimators are added to the MILP formulation, performances of all instances are increased considerably. The number of explored nodes and the solving time is decreased significantly. However, for the cases where no-purchase preference is low and the discount factor is high, the model has difficulty for obtaining an optimal solution and requires deeper node exploration.

Table 5.3: Comparison table for formulations on real data.

v_0	γ	CONIC			CONIC + MC			MILP			MILP + MC		
		Iteration	Nodes	Time	Iteration	Nodes	Time	Iteration	Nodes	Time	Iteration	Nodes	Time
1	0.25	25,728	4,835	1.773	2,221	98	1.531	10,249,133	176,047	180.537	6,376,727	300,823	80.345
	0.50	98,072	19,185	2.772	45,824	9,592	2.664	79,842,104	1,606,987	0.13%	142,674,452	6,628,367	– 0.02%
	0.75	421,627	107,554	4.926	148,395	24,826	4.175	79,198,511	1,829,461	0.18%	94,222,919	3,271,480	– 0.05%
	0.90	2,212,614	550,469	22.671	1,067,458	209,768	18.796	69,180,892	1,434,592	0.21%	95,476,487	3,518,950	– 0.08%
5	0.25	29,244	6,466	1.360	684	1	1.156	40,459	2,083	8.908	3,097	207	1.221
	0.50	292,937	62,043	3.829	3,145	340	1.537	573,784	23,482	17.644	114,173	6,851	7.063
	0.75	10,695,629	1,355,951	106.526	78,878	10,066	3.922	4,341,107	136,455	56.951	23,845,783	848,781	41.457
	0.90	92,415,899	13,526,948	830.701	3,826,044	430,560	80.157	162,353,355	5,110,531	0.12%	133,187,753	4,359,522	– 0.01%
10	0.25	34,213	4,922	1.958	816	1	0.986	121,943	5,925	2.860	1,562	30	1.328
	0.50	383,219	37,430	9.932	1,148	1	1.299	262,388	12,207	6.879	26,015	2,001	1.749
	0.75	4,882,004	346,903	99.256	70,823	12,787	3.184	4,352,260	211,763	68.748	4,479,461	281,046	31.023
	0.90	9,010,629	1,848,166	187.238	4,341,807	632,605	74.104	161,857,295	7,026,211	0.12%	13,5335,029	6,792,799	– 0.02%
20	0.25	15569	2949	0.985	951	1	0.874	7,311	467	1.094	1,977	75	1.406
	0.50	183,959	24,716	4.705	1,085	99	1.135	65,418	3,908	2.149	3,623	474	1.430
	0.75	5,279,543	769,327	79.386	8,336	1,895	1.354	10,231,853	916,154	95.316	173,121	19,855	15.717
	0.90	3,303,273	267,865	124.954	387,807	88,001	10.082	31,383,716	2,384,759	0.15%	9,094,926	1,144,282	– 0.02%
50	0.25	1,059	1	0.812	434	1	0.531	1,841	15	1.471	1,098	1	0.907
	0.50	9,630	1,681	0.901	1,250	82	0.891	4,905	1,010	1.249	1,404	16	1.113
	0.75	10,648	4,270	1.350	2,852	702	0.929	7,415	1,450	1.259	4,470	109	1.169
	0.90	2,110,828	272,487	15.986	795,546	72,488	7.572	5,829,440	869,587	25.257	1,159,201	143,408	12.940

In contrast to the linear formulations, the conic formulations are stronger. The optimal solution is reached with a significantly less number of nodes and number of iterations. When the no-purchase preferences are small, the conic formulation performs better. As the discount factor increases, the conic formulation reaches the optimal solution in a longer time. We note that as the no-purchase preference increases, the MILP and the conic formulation show almost the same performance.

When the McCormick estimators are added to the conic formulation (CONIC+MC), the results improve significantly. All instances are solved in a shorter time and with a less number of iterations. The reason for this improvement is the usage of valid McCormick estimators and conic constraints, resulting in a tight formulation.

Table 5.4: Study results on real data when $\kappa_1 = 10$.

υ_0	γ	π	π_1	π_2	$\lvert S\rvert$	$\lvert S_1\rvert$	$\lvert S_2\rvert$	$\lvert S_1 \setminus S_2\rvert$	$\lvert S_2 \setminus S_1\rvert$	$1 \to 2$	$2 \to 1$	FR_1	FR_2
1	0.25	15.1300	16.0447	15.0717	19	10	19	0	9	5.11%	0.00%	86.15%	90.09%
	0.50	15.2349	16.0447	15.1833	19	10	19	0	9	9.72%	0.00%	86.83%	90.09%
	0.75	15.3300	16.0447	15.2844	19	10	19	0	9	11.89%	0.00%	87.14%	90.09%
	0.90	15.3832	16.0447	15.3411	19	10	19	0	9	27.67%	0.00%	87.62%	90.09%
5	0.25	9.9082	11.7702	9.7895	25	10	25	0	15	5.61%	0.00%	56.97%	70.13%
	0.50	10.1975	11.7061	10.1013	27	10	27	0	17	11.86%	0.00%	59.82%	71.76%
	0.75	10.4530	11.7396	10.3710	26	10	26	0	16	16.73%	0.00%	61.64%	70.91%
	0.90	10.5962	11.8064	10.5191	24	10	24	0	14	16.88%	0.00%	61.71%	69.22%
10	0.25	6.9871	8.7826	6.8756	27	10	24	3	17	4.99%	1.54%	40.59%	54.50%
	0.50	7.3707	9.0781	7.2618	30	10	30	0	20	12.74%	0.00%	45.44%	59.67%
	0.75	7.6664	9.0781	7.5764	30	10	30	0	20	17.97%	0.00%	48.71%	59.67%
	0.90	7.8322	9.1106	7.7507	29	10	29	0	19	18.09%	0.00%	48.78%	58.74%
20	0.25	4.4639	5.5737	4.3932	31	10	24	7	21	3.93%	3.73%	29.31%	37.36%
	0.50	4.8447	6.4590	4.7421	31	10	30	1	21	8.89%	1.15%	31.14%	43.01%
	0.75	5.1659	6.4944	5.0813	32	10	32	0	22	14.70%	0.00%	34.60%	46.11%
	0.90	5.3336	6.4944	5.2596	32	10	32	0	22	17.14%	0.00%	36.46%	46.11%
50	0.25	2.1777	2.8652	2.1339	31	10	23	8	21	1.84%	1.92%	14.39%	19.27%
	0.50	2.4233	3.1955	2.3756	33	10	26	7	23	5.08%	3.33%	17.00%	23.26%
	0.75	2.6684	3.3141	2.6315	35	10	25	10	25	9.37%	6.89%	20.75%	26.69%
	0.90	2.8103	3.5797	2.7612	35	10	29	6	25	10.75%	4.42%	21.67%	28.24%

We have reported the statistics regarding the optimal assortments in Table 5.4. The first two columns represent the no-purchase preference and discount factor which are inputs for the optimization problem. The rest of the columns are the results of these settings. The third column shows the expected revenue per customer in the whole system, and the fourth and the fifth columns represent the small (Store 1) and the large (Store 2) stores' expected revenues per customer, respectively. The number of products offered in the two stores are reported in the sixth column, and the seventh and the eighth columns represent the sizes of the assortments in first and second stores, respectively. The ninth column represents the number of products that are in store 1, but not in store 2. The tenth column is similar for store 2. The following two columns represent the percent of customers who originally come for one store but shop at the other store for their purchase, namely cross-shopping rates. The last two columns show the fill rate for each store considering those customers who shopped at a different store than their intention, i.e., the percent of customers who made a purchase in either store.

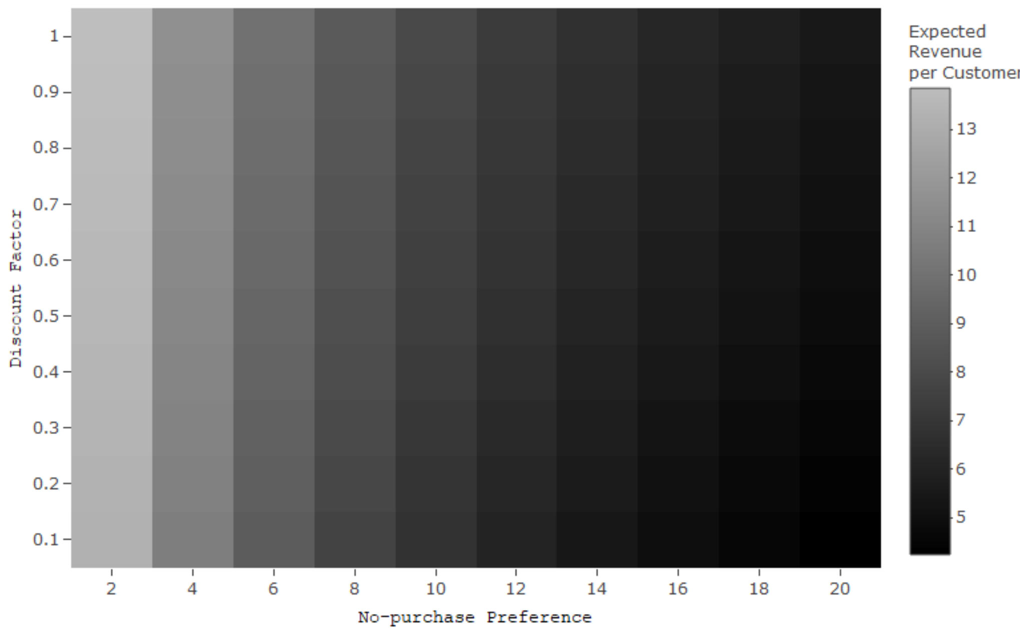
In the table, the expected revenue of a customer goes up as the discount factor goes up and no-purchase preference goes down. The assortments of the stores become diverse as the no-purchase preference increases in order to minimize the probability of the lost sales. For the small store, the assortment is always full-sized, and for the large store, the assortment size increases as the no-purchase preference is increased. When the no-purchase preference is low, the small store's assortment is usually the subset of the large store's assortment. Moreover, as the discount factor grows, the cross-shopping rate of customers in store 1 increases. When the discount factor is increased, the purchase preferences become less affected by the lead time, resulting in a higher probability of purchase. Additionally, since the probability of selling a product decreases as the no-purchase preference increases, customers prefer to switch stores less.

In Figure 5.1, the heat-map of the expected revenue per customer is represented. In the figure, the expected revenue per customer increases as the no-purchase preference is decreased, likewise in Table 5.4. However, it can be inferred that, as the no-purchase preference decreases, the discount factor seems to have no significant effect on customer behavior since the expected revenue per customer remains the same.

The heat-map of the cross-shopping rate of customers in Store 1 can be seen in Figure 5.2. The maximum rate is obtained when the no-purchase preference is as low as 2 and the discount factor is equal to 1, meaning that when the stores are close enough, customers would be more likely to switch their stores to buy a product.

In Table 5.5, we have analyzed the case where capacity of the small store is

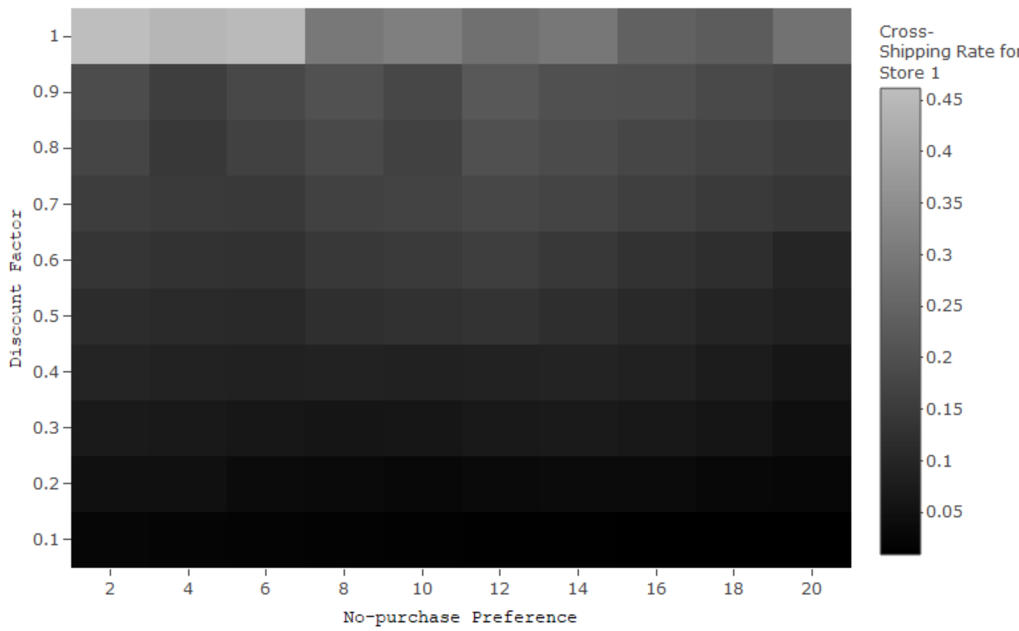
Figure 5.1: Heat-map of expected revenue with respect to the no-purchase preferences and discounting factors.



specified as 5, i.e., $\kappa_1 = 5$. The columns of the table are similar to the Table 5.4.

In all instances, the optimal model utilizes the full capacity of the small store. When we analyze the revenues, we can clearly see that the difference between the expected revenues from the two stores increased in comparison to Table 5.4. We can expect a decrease on the π value since the capacity is decreased to 5. As the no-purchase preference increases, the number of unique products in the optimal assortment is increased in order to maximize the probability of selling a product. Similarly, as the discount factor increases, the model tries to expand the assortment since increasing discount factor means that customers are more likely to switch their stores when the product they look for is not available. The cross-shopping rates of customers decreases as the no-purchase preference increases, i.e., the cross-shopping rate of customers in the small store is inversely proportional with the no-purchase preference. However, the cross-shopping rate increases if the discount factor increases, i.e., the cross-shopping rate of customers in the small store is directly proportional with the discount factor. Also, the fill rate of the stores are directly proportional with the discount factor, and inversely

Figure 5.2: Heat-map of cross-shopping rate of customers in Store 1 with respect to the no-purchase preferences and discounting factors.



proportional with the no-purchase preference. Additionally, comparing to Table 5.4, cross-shopping rates are increased for all no-purchase preference values, and decreasing the capacity of the small store results an increase in the number of different products in the stores' assortments. However, in general, the firm carries a smaller assortment when $\gamma \leq 0.50$. The fill rate of the small store is lessened when the capacity is decreased, whereas the decrease in the fill rate of the large store occurs only in the case where $\gamma = (0.50, 0.75)$.

In Table 5.6, we study the effect of ignoring cross-shopping customers. In the right side of Table 5.6, we report what happens when the firm plans its assortment independently in each store, but then allows its customers to cross-shop. In the last column, we present the percentage loss in the expected revenue per customer.

The results show that considering the cross-shopping customers in assortment planning can have a significant effect on expected revenues and fill rates. When the no-purchase preference is low, ignoring cross-shopping customers in assortment optimization may decrease the expected revenue from customers in store 2

Table 5.5: Study results when capacity of the small store $\kappa_1 = 5$.

υ_0	γ	π	π_1	π_2	$\lvert S\rvert$	$\lvert S_1\rvert$	$\lvert S_2\rvert$	$\lvert S_1 \setminus S_2\rvert$	$\lvert S_2 \setminus S_1\rvert$	$1 \to 2$	$2 \to 1$	FR_1	FR_2
1	0.25	14.4659	16.0156	14.3671	20	5	19	1	15	13.60%	1.10%	83.50%	90.11%
	0.50	14.8622	16.0171	14.7886	20	5	20	0	15	30.42%	0.00%	84.72%	90.42%
	0.75	15.1730	16.0171	15.1192	20	5	20	0	15	39.61%	0.00%	86.74%	90.42%
	0.90	15.3262	16.0447	15.2804	19	5	19	0	14	54.55%	0.00%	87.26%	90.09%
5	0.25	8.8697	10.9655	8.7388	24	5	19	5	19	9.78%	4.93%	51.39%	64.67%
	0.50	9.6387	11.7061	9.5069	27	5	27	0	22	20.22%	0.00%	56.82%	71.76%
	0.75	10.2078	11.7061	10.1123	27	5	27	0	22	27.55%	0.00%	60.78%	71.76%
	0.90	10.5020	11.7396	10.4231	26	5	26	0	21	30.59%	0.00%	61.43%	70.91%
10	0.25	6.0104	8.3743	5.9013	25	5	20	5	20	6.99%	3.55%	34.87%	49.11%
	0.50	6.8307	8.6372	6.7278	30	5	25	5	25	18.27%	5.96%	42.77%	57.26%
	0.75	7.4430	9.0781	7.3388	30	5	30	0	25	23.85%	0.00%	46.67%	59.67%
	0.90	7.7456	9.1107	7.6586	29	5	29	0	24	26.21%	0.00%	48.33%	58.74%
20	0.25	3.6841	6.0675	3.6677	28	5	23	5	23	6.36%	2.20%	22.87%	36.95%
	0.50	4.4035	6.1910	4.3304	31	5	26	5	26	13.14%	4.09%	28.45%	41.36%
	0.75	4.9869	6.4396	4.8942	32	5	30	2	27	19.41%	1.72%	33.62%	45.80%
	0.90	5.2668	6.4944	5.1886	32	5	32	0	27	22.42%	0.00%	36.10%	46.11%
50	0.25	1.7168	3.1776	1.7206	30	5	25	5	25	3.20%	1.11%	10.84%	20.50%
	0.50	2.1397	3.3221	2.1093	32	5	27	5	27	7.14%	2.14%	14.46%	23.17%
	0.75	2.5336	3.4546	2.4931	34	5	29	5	29	12.00%	3.09%	18.94%	26.15%
	0.90	2.7572	3.6079	2.7169	35	5	30	5	30	14.61%	3.60%	21.34%	28.31%

(hence the total expected revenue). The cost of ignoring is higher when discount factor is high and no-purchase preference is high. Moreover, when the no-purchase preference is very high, ignoring the cross-shopping customers affects the expected revenue from customers in both stores.

Table 5.6: The effect of ignoring cross-shopping in assortment planning.

		Transfers are considered					Transfers are ignored					
υ_0	γ	π	π_1	π_2	FR_1	FR_2	π	π_1	π_2	FR_1	FR_2	Δ_π
1	0.25	15.1300	16.0447	15.0717	86.15%	90.09%	15.0138	16.0447	14.9482	85.41%	90.09%	0.77%
	0.50	15.2349	16.0447	15.1833	86.83%	90.09%	15.0138	16.0447	14.9482	85.41%	90.09%	1.45%
	0.75	15.3300	16.0447	15.2844	87.14%	90.09%	15.0138	16.0447	14.9482	85.41%	90.09%	2.06%
	0.90	15.3832	16.0447	15.3411	87.62%	90.09%	14.9535	16.0447	14.8840	82.89%	90.09%	2.79%
5	0.25	9.9082	11.7702	9.7895	56.97%	70.13%	9.5875	11.7702	9.4484	54.41%	70.13%	3.24%
	0.50	10.1975	11.7061	10.1013	59.82%	71.76%	9.5836	11.7061	9.4484	54.41%	71.76%	6.02%
	0.75	10.4530	11.7396	10.3710	61.64%	70.91%	9.5766	11.7396	9.4388	53.93%	70.91%	8.38%
	0.90	10.5962	11.8064	10.5191	61.71%	69.22%	9.5806	11.8064	9.4388	53.93%	69.22%	9.58%
10	0.25	6.9871	8.7826	6.8756	40.59%	54.50%	6.6241	8.6528	6.4948	37.47%	53.79%	5.20%
	0.50	7.3707	9.0781	7.2618	45.44%	59.67%	6.6495	9.0781	6.4948	37.47%	59.67%	9.78%
	0.75	7.6664	9.0781	7.5764	48.71%	59.67%	6.6495	9.0781	6.4948	37.47%	59.67%	13.26%
	0.90	7.8322	9.1107	7.7507	48.78%	58.74%	6.6515	9.1107	6.4948	37.47%	58.74%	15.07%
20	0.25	4.4639	5.5737	4.3932	29.31%	37.36%	4.0926	5.2230	4.0205	26.42%	34.94%	8.32%
	0.50	4.8447	6.4590	4.7421	31.14%	43.01%	4.1558	6.4298	4.0109	24.42%	42.35%	14.22%
	0.75	5.1659	6.4944	5.0813	34.60%	46.11%	4.1545	6.4944	4.0054	23.32%	46.11%	19.58%
	0.90	5.3336	6.4944	5.2596	36.46%	46.11%	4.1545	6.4944	4.0054	23.32%	46.11%	22.11%
50	0.25	2.1777	2.8652	2.1339	14.39%	19.27%	1.9570	2.6317	1.9140	12.79%	17.69%	10.13%
	0.50	2.4233	3.1955	2.3756	17.00%	23.26%	1.9654	2.8175	1.9111	12.56%	20.61%	18.90%
	0.75	2.6684	3.3141	2.6315	20.75%	26.69%	1.9407	2.4042	1.9111	12.56%	21.26%	27.27%
	0.90	2.8103	3.5797	2.7612	21.67%	28.24%	1.9741	3.0917	1.9029	12.24%	24.92%	29.75%

Chapter 6

Conclusion

In this thesis, we focused on an assortment optimization problem in an online retailing environment. This type of retailing rises all over the world, and in Turkey, the transaction size of the online retailing is almost doubling every year. Regardless of the company type (local or global), every company puts its efforts into making its online operation's management better. One of the keys for such an achievement is the fulfillment decisions in online retailing that are significantly different than traditional retailing. In online retailing, unlike the traditional retailing, retailers own a network consisting of fulfillment centers and can meet a customer demand from any of the fulfillment centers, meaning that every retailer has to decide from where to meet each demand.

Planning the assortment at each fulfillment center has a great importance for online retailers since it affects the purchasing behavior of its customers. Hence, decisions should be made cautiously, and also, those decisions need to consider the cost of shipping since a demand can be met from another region's fulfillment center. In this setting, our study addresses the profit maximization problem of an online retailer by selecting each fulfillment center's assortment. It is assumed that for each fulfillment center, there is an associated customer segment and each customer segment's demand is governed by its own MNL model. Since our study considers many fulfillment centers, the general consumer choice model can be classified as MMNL model. In our problem setting, we also take the capacity constraints of fulfillment centers into consideration, i.e., we put a limit on the number of products that can be carried in the assortment of each fulfillment center.

In our problem setting, customers have access to the entire assortment in all fulfillment centers. However, for the products located in distant regions, the preferences are decreased directly in proportion to the distances, meaning lower probability of sales for products that are carried in distant regions. Therefore, the demand can be satisfied from any of the fulfillment centers. If a product is shipped from another region, additional shipping cost occurs which is undertaken by the retailer. The cost of shipping is not fixed and depends on the distance among regions.

First, we have provided a non-linear formulation for the assortment planning problem, called baseline formulation. The formulation is linearized and reformulated as an mixed integer linear problem (MILP). Moreover, by expanding the conic formulation of Şen et al. [34], we have developed a conic formulation for this multi-location problem. To make the formulation stronger, we have proposed McCormick estimators.

We have presented a numerical study with two sets of parameters where each combination has five randomly generated samples. In each sample, there are four customer segments and forty products. We have examined all formulations (MILP, MILP+MC, CONIC, CONIC+MC) on each instance. We have observed that CONIC+MC formulation (conic formulation strengthened with McCormick estimators) performs the best and decreases the run times significantly for all instances in comparison to linear formulations. Although the CONIC formulation expresses an improvement in all instances, it still remains incapable of solving all instances. The MILP formulation cannot solve any instance, whereas, the MILP+MC formulation works better and can reach the optimal solution for some of the instances.

Additionally, we have tested our formulations on the real data which is provided by a local retail chain. The detergent sales data over two years from two stores is analyzed, and two MNL models are fitted one for each store. With the determined coefficients of attributes, we have calculated each product's preference values, and used them as an input for the assortment optimization formulations. We tested our formulation for the case where the small store has a limited capacity. Supporting the results of the numerical study in Chapter 4, our CONIC+MC formulation shows the best performance. Moreover, we have analyzed the cases where the transfer between stores is ignored. For the case in which the transfers are ignored, we have concluded that the revenue of the chain decreases, especially in the large store.

Our results, in general, show that conic formulations can be very useful in solving large instances of multi-location assortment planning problems, and that ignoring the multi-location nature of the problem may lead to significant losses for the firms.

Bibliography

[1] D. Bertsimas and V. V. Mišic, "Data-driven assortment optimization," working paper, MIT Sloan School, 2015.

[2] J. A. Quelch and D. Kenny, "Extend profits, not product lines," *Harvard Business Review*, vol. 72, pp. 153—-164, 1994.

[3] M. Fisher and R. Vaidyanathan, "Which products should you stock?," *Harvard Business Review*, vol. 90, pp. 108–118, 2012.

[4] S. A. Smith, "Optimizing retail assortments for diverse customer preferences," in *Retail supply chain management*, pp. 183–205, Springer, 2008.

[5] J. M. Davis, G. Gallego, and H. Topaloglu, "Assortment optimization under variants of the nested logit model," *Operations Research*, vol. 62, no. 2, pp. 250–273, 2014.

[6] A. G. Kök, M. L. Fisher, and R. Vaidyanathan, "Assortment planning: Review of literature and industry practice," in *Retail Supply Chain Management*, pp. 99–153, Springer, 2008.

[7] S. Mahajan and G. J. van Ryzin, "Retail inventories and consumer choice," in *Quantitative Models for Supply Chain Management*, pp. 491–551, Springer, 1999.

[8] G. P. Cachon, C. Terwiesch, and Y. Xu, "Retail assortment planning in the presence of consumer search," *Manufacturing & Service Operations Management*, vol. 7, no. 4, pp. 330–346, 2005.

[9] O. Besbes and D. Sauré, "Product assortment and price competition under multinomial logit demand," *Production and Operations Management*, vol. 25, no. 1, pp. 114–127, 2016.

[10] J. Yates and D. F. Mackay, "Discrete choice modelling of urban housing markets: A critical review and an application," *Urban Studies*, vol. 43, no. 3, pp. 559–581, 2006.

[11] E. Gumbel, "Statistics of extremes," *Columbia University Press, New York*, 1958.

[12] S. P. Anderson, A. De Palma, and J. F. Thisse, *Discrete choice theory of product differentiation.* MIT Press, 1992.

[13] S. Mahajan and G. Van Ryzin, "Stocking retail assortments under dynamic consumer substitution," *Operations Research*, vol. 49, no. 3, pp. 334–351, 2001.

[14] G. Debreu, "Review of r. d. luce individual choice behavior," *American Economic Review*, vol. 50, pp. 186–188, 1960.

[15] P. M. Guadagni and J. D. Little, "A logit model of brand choice calibrated on scanner data," *Marketing Science*, vol. 2, no. 3, pp. 203–238, 1983.

[16] M. E. Ben-Akiva and S. R. Lerman, *Discrete choice analysis: Theory and application to travel demand*, vol. 9. MIT Press, 1985.

[17] M. E. Ben-Akiva, *Structure of passenger travel demand models.* PhD thesis, Massachusetts Institute of Technology, 1973.

[18] S. A. Smith and N. Agrawal, "Management of multi-item retail inventory systems with demand substitution," *Operations Research*, vol. 48, no. 1, pp. 50–64, 2000.

[19] A. G. Kök and M. L. Fisher, "Demand estimation and assortment optimization under substitution: Methodology and application," *Operations Research*, vol. 55, no. 6, pp. 1001–1021, 2007.

[20] H. Hotelling, "Stability in competition," *Economic Journal*, vol. 39, pp. 41–57, 1929.

[21] K. J. Lancaster, "A new approach to consumer theory," *The Journal of Political Economy*, vol. 74, no. 2, pp. 132–157, 1966.

[22] K. Lancaster, "Socially optimal product differentiation," *The American Economic Review*, vol. 65, no. 4, pp. 567–585, 1975.

[23] V. Gaur and D. Honhon, "Assortment planning and inventory decisions under a locational choice model," *Management Science*, vol. 52, no. 10, pp. 1528–1543, 2006.

[24] R. D. McBride and F. S. Zufryden, "An integer programming approach to the optimal product line selection problem," *Marketing Science*, vol. 7, no. 2, pp. 126–140, 1988.

[25] R. Kohli and R. Sukumar, "Heuristics for product-line design using conjoint analysis," *Management Science*, vol. 36, no. 12, pp. 1464–1478, 1990.

[26] J. R. Nicholson, "New insights on retail e-commerce (esa issue brief #04-17)," *Economics and Statistics Administration, U.S. Department of Commerce*, Retrieved from https://www.esa.gov/reports/new-insights-retail-e-commerce, 2017.

[27] A. Fareeha, "Decade in review: E-commerce sales vs. total retail sales," *Digital Commerce 360*, Retrieved from https://www.digitalcommerce360.com/article/e-commerce-sales-retail-sales-ten-year-review/, 2018.

[28] O. Kantarcı, M. Özalp, C. Sezginsoy, O. Özaşkınlı, and C. Cavlak, "Dijitalleşen dünyada ekonominin itici gücü: E-ticaret," *TÜSIAD*, 2016.

[29] J. Acimovic, *Lowering outbound shipping costs in an online retailing environment by making better fulfillment and replenishment decisions*. PhD thesis, Massachusetts Institute of Technology, 2012.

[30] PwC, "Total retail 2017: 10 retailer investments for an uncertain future," tech. rep., PwC, http://www.pwc.com/gx/en/industries/assets/total-retail-2017.pdf, 2017. Accessed June, 2018.

[31] P. Rusmevichientong, Z.-J. M. Shen, and D. B. Shmoys, "Dynamic assortment optimization with a multinomial logit choice model and capacity constraint," *Operations Research*, vol. 58, no. 6, pp. 1666–1680, 2010.

[32] J. J. M. Bront, I. Méndez-Díaz, and G. Vulcano, "A column generation algorithm for choice-based network revenue management," *Operations Research*, vol. 57, no. 3, pp. 769–784, 2009.

[33] P. Rusmevichientong, D. Shmoys, C. Tong, and H. Topaloglu, "Assortment optimization under the multinomial logit model with random choice parameters," *Production and Operations Management*, vol. 23, no. 11, pp. 2023–2039, 2014.

[34] A. Şen, A. Atamtürk, and P. Kaminsky, "Technical note: A conic integer optimization approach to the constrained assortment problem under the mixed multinomial logit model," *Operations Research*, vol. 66, no. 4, pp. 994–1003, 2018.

[35] M. Karampatsa, E. Grigoroudis, and N. F. Matsatsinis, "Retail category management: A review on assortment and shelf-space planning models," in *Operational Research in Business and Economics*, pp. 35–67, Springer, 2017.

[36] G. v. Ryzin and S. Mahajan, "On the relationship between inventory costs and variety benefits in retail assortments," *Management Science*, vol. 45, no. 11, pp. 1496–1509, 1999.

[37] J.-K. Chong, T.-H. Ho, and C. S. Tang, "A modeling framework for category assortment planning," *Manufacturing & Service Operations Management*, vol. 3, no. 3, pp. 191–210, 2001.

[38] Z. Li, "A single-period assortment optimization model," *Production and Operations Management*, vol. 16, no. 3, pp. 369–380, 2007.

[39] C. M. Miller, S. A. Smith, S. H. McIntyre, and D. D. Achabal, "Optimizing and evaluating retail assortments for infrequently purchased products," *Journal of Retailing*, vol. 86, no. 2, pp. 159–171, 2010.

[40] P. Rusmevichientong and H. Topaloglu, "Robust assortment optimization in revenue management under the multinomial logit choice model," *Operations Research*, vol. 60, no. 4, pp. 865–882, 2012.

[41] J. Davis, G. Gallego, and H. Topaloglu, "Assortment planning under the multinomial logit model with totally unimodular constraint structures," *Working Paper*, 2013.

[42] H. Topaloglu, "Joint stocking and product offer decisions under the multinomial logit model," *Production and Operations Management*, vol. 22, no. 5, pp. 1182–1199, 2013.

[43] V. Goyal, R. Levi, and D. Segev, "Near-optimal algorithms for the assortment planning problem under dynamic substitution and stochastic demand," *Operations Research*, vol. 64, no. 1, pp. 219–235, 2016.

[44] D. Dzyabura and S. Jagabathula, "Offline assortment optimization in the presence of an online channel," *Management Science*, vol. 64, no. 6, pp. 2767–2786, 2018.

[45] A. G. Kök and Y. Xu, "Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models," *Management Science*, vol. 57, no. 9, pp. 1546–1563, 2011.

[46] A. Alptekinoğlu and A. Grasas, "When to carry eccentric products? optimal retail assortment under consumer returns," *Production and Operations Management*, vol. 23, no. 5, pp. 877–892, 2014.

[47] G. Gallego and H. Topaloglu, "Constrained assortment optimization for the nested logit model," *Management Science*, vol. 60, no. 10, pp. 2583–2601, 2014.

[48] G. Li and P. Rusmevichientong, "A greedy algorithm for the two-level nested logit model," *Operations Research Letters*, vol. 42, no. 5, pp. 319–324, 2014.

[49] B. Rodríguez and G. Aydın, "Pricing and assortment decisions for a manufacturer selling through dual channels," *European Journal of Operational Research*, vol. 242, no. 3, pp. 901–909, 2015.

[50] J. Feldman and H. Topaloglu, "Bounding optimal expected revenues for assortment optimization under mixtures of multinomial logits," *Production and Operations Management*, vol. 24, no. 10, pp. 1598–1620, 2015.

[51] G. Li, P. Rusmevichientong, and H. Topaloglu, "The d-level nested logit model: Assortment and price optimization problems," *Operations Research*, vol. 63, no. 2, pp. 325–342, 2015.

[52] N. S. Cardell and F. C. Dunbar, "Measuring the societal impacts of automobile downsizing," *Transportation Research Part A: General*, vol. 14, no. 5-6, pp. 423–434, 1980.

[53] J. H. Boyd and R. E. Mellman, "The effect of fuel economy standards on the u.s. automotive market: an hedonic demand analysis," *Transportation Research Part A: General*, vol. 14, no. 5-6, pp. 367–378, 1980.

[54] I. Méndez-Díaz, J. J. Miranda-Bront, G. Vulcano, and P. Zabala, "A branch-and-cut algorithm for the latent-class logit assortment problem," *Discrete Applied Mathematics*, vol. 164, pp. 246–263, 2014.

[55] J. Feldman and H. Topaloglu, "Technical note: Capacitated assortment optimization under the multinomial logit model with nested consideration sets," *Operations Research*, vol. 66, no. 2, pp. 380–391, 2018.

[56] G. Aydın and W. H. Hausman, "Supply chain coordination and assortment planning," *Working paper. University of Michigan*, 2003.

[57] P. Singh, H. Groenevelt, and N. Rudi, "Product variety and supply chain structures," *Working paper. University of Rochester*, 2005.

[58] D. McFadden, "Conditional logit analysis of qualitative choice behavior," *Frontiers in Economics*, 1974.

[59] P. S. Fader and B. G. Hardie, "Modeling consumer choice among skus," *Journal of Marketing Research*, vol. 33, pp. 442–452, 1996.

[60] J. Chiang, "A simultaneous approach to the whether, what and how much to buy questions," *Marketing Science*, vol. 10, no. 4, pp. 297–315, 1991.

[61] R. E. Bucklin and S. Gupta, "Brand choice, purchase incidence, and segmentation: An integrated modeling approach," *Journal of Marketing Research*, vol. 29, no. 2, pp. 201–215, 1992.

[62] P. K. Chintagunta, "Investigating purchase incidence, brand choice and purchase quantity decisions of households," *Marketing Science*, vol. 12, no. 2, pp. 184–208, 1993.

[63] K. Talluri and G. Van Ryzin, "Revenue management under a general discrete choice model of consumer behavior," *Management Science*, vol. 50, no. 1, pp. 15–33, 2004.

[64] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (methodological)*, vol. 39, pp. 1–38, 1977.

[65] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*, vol. 382. John Wiley & Sons, 2007.

[66] G. Vulcano, G. Van Ryzin, and R. Ratliff, "Estimating primary demand for substitutable products from sales transaction data," *Operations Research*, vol. 60, no. 2, pp. 313–334, 2012.

[67] J. P. Newman, M. E. Ferguson, L. A. Garrow, and T. L. Jacobs, "Estimation of choice-based models using sales data from a single firm," *Manufacturing & Service Operations Management*, vol. 16, no. 2, pp. 184–197, 2014.

[68] T. Abdallah and G. Vulcano, "Demand estimation under the multinomial logit model from sales transaction data," *Working Paper*, 2016.

[69] M. M. Fadıloğlu, O. E. Karaşan, and M. Ç. Pınar, "A model and case study for efficient shelf usage and assortment analysis," *Annals of Operations Research*, vol. 180, no. 1, pp. 105–124, 2010.

[70] M. P. Keane, N. Wasi, *et al.*, "Estimation of discrete choice models with many alternatives using random subsets of the full choice set: With an application to demand for frozen pizza," tech. rep., Economics Group, Nuffield College, University of Oxford, 2012.

[71] L. Li, "The role of inventory in delivery-time competition," *Management Science*, vol. 38, no. 2, pp. 182–197, 1992.

[72] L. Li and Y. S. Lee, "Pricing and delivery-time performance in a competitive environment," *Management Science*, vol. 40, no. 5, pp. 633–646, 1994.

[73] K. C. So and J.-S. Song, "Price, delivery time guarantees and capacity selection," *European Journal of Operational Research*, vol. 111, no. 1, pp. 28–49, 1998.

[74] M. Fisher, S. Gallino, and J. Xu, "The value of rapid delivery in online retailing," *Working Paper*, 2016.

[75] G. Berbeglia and G. Joret, "Assortment optimisation under a general discrete choice model: A tight analysis of revenue-ordered assortments," *arXiv preprint arXiv:1606.01371*, 2015.

[76] S. Jagabathula, "Assortment optimization under general choice," *Working Paper*, 2014.

[77] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1-3, pp. 193–228, 1998.

[78] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Mathematical Programming*, vol. 95, no. 1, pp. 3–51, 2003.

[79] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part 1–convex underestimating problems," *Mathematical Programming*, vol. 10, no. 1, pp. 147–175, 1976.

[80] R. Ihaka and R. Gentleman, "R: a language for data analysis and graphics," *Journal of Computational and Graphical Statistics*, vol. 5, no. 3, pp. 299–314, 1996.

[81] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1967.

[82] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964.

[83] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM Journal on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, 1995.

[84] R Core Team, *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2018.