To my family

OPTIMAL ASSORTMENT PLANNING UNDER CAPACITY CONSTRAINT: SINGLE AND MULTI-FIRM SYSTEMS USING TRANSSHIPMENTS

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

ECEM CEPHE

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF SCIENCE

THE DEPARTMENT OF
MANAGEMENT
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

August 2016

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Decision Science and Operations Management.

Prof. Erdal Erel
Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Decision Science and Operations Management.

Assoc. Prof. Nagihan Çömez Dolgan
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Decision Science and Operations Management.

Assoc. Prof. Alper Şen
Examining Committee Member

Approval of the Institute of Economics and Social Sciences

Title and Name
Director
Prof. Dr. Halime Demirkan

ABSTRACT

OPTIMAL ASSORTMENT PLANNING UNDER CAPACITY CONSTRAINT:
SINGLE AND MULTI-FIRM SYSTEMS USING TRANSSHIPMENTS

Cephe, Ecem
M.S., Department of Management
Supervisor: Prof. Dr. Erdal Erel

August 2016

To survive in today's competitive market, firms need to meet customer expectations by offering high quality products with high variety. However, there might be some physical and financial constraints limiting the assortment size. The process of finding the optimal product assortment by considering both the benefits of a large assortment as well as the costs and limits of it is known as assortment planning. In this thesis, assortment planning is analyzed under predetermined assortment capacity limits for two cases. First, a single firm's optimal assortment problem is studied to maximize its profits. Second, assortment planning problem of a system of multiple firms is investigated jointly, where firms are interacting through product sharing, called transshipments. Transshipments are known to increase product availability, thus decreasing stock-outs. Transshipments have been always utilized as an emergency demand satisfaction action in support of inventory management in the literature. Differently, in this study, transshipments are evaluated in advance of inventory management while making the assortment planning of firms. In both problems, demand is defined to have an exogenous model, where each customer has a predetermined preference for each product from the potential set. Proportional demand substitutions are also allowed from an out of assortment product to others.

The results on the optimal assortment of a single firm are used as a benchmark to the optimal assortments of multiple firms communicating through transshipments. By relying on proven optimality results, it is shown how easily optimal assortments can be obtained compared to a full enumeration. Extensive numerical analyses are reported on the performances of the heuristic algorithm and sensitivity of optimal assortments to system parameters.

Keywords: Assortment Planning, Capacity Constraint, Exogenous Demand Model, Substitution, Transshipment.

ÖZET

KAPASİTE LİMİTLERİ ALTINDA ÜRÜN ÇEŞİTLİLİĞİ PLANLAMASI: TEKLİ VE TRANSFER-SATIŞ KULLANAN ÇOKLU BAYİ SİSTEMLERİ

Cephe, Ecem
Yüksek Lisans, İşletme Bölümü
Tez Yöneticisi: Prof. Dr. Erdal Erel

Ağustos 2016

Günümüzün rekabetçi ortamında bayiler, müşteri beklentilerini karşılayabilmek için kaliteli ve çeşitli ürün sunmalıdırlar ancak fiziksel ve finansal kısıtlar ürün çeşitliliğini sınırlayabilir. Geniş ürün portfolyosu sunmanın avantajı, maliyeti ve limitlerini dikkate alarak en iyi ürün çeşitliliğini bulma süreci ürün çeşitliliği planlaması olarak bilinir. Bu tezde, ürün çeşitliliği planlaması, önceden belirlenmiş kapasite limitleri altnda iki farklı durum için çalışılmıştır. İlki, tek bayili sistemlerde karı eniyilemek için ürün çeşitliliği planlamasıdır. İkincisi, transfer-satış politikası ile ürün paylaşımının olduğu çoklu bayili sistemlerin ürün çeşitliliği planlamasıdır. Transfer-satışlar ürün bulunurluğunu arttırarak stok yetersizliğini azaltma amacıyla kullanılır. Literatürde, transfer-satış acil durumlardaki talepleri karşılamada stok yönetimini destekleyici bir aksiyon olarak görülmektedir. Literatürden farklı olarak, bu çalışmada transfer-satış, stok planlamasından önce ürün çeşitliliği planlaması aşamasında ele alınmıştır. Çalışılan her iki modelde de müşteri tercihinin bilindiği harici talep modeli kullanılmaktadır. Tek bayili sistemlerin ürün çeşitliliği planlamasından elde edilen sonuçlar, transfer-satış politikasının kullanıldığı çoklu bayili sistemlerin ürün çeşitliliği planlamasına referans oluşturmaktadır. En iyi ürün çeşitliliğinin özellikleri üzerine elde edilen sonuçlar kullanılarak tüm sayım metoduna kıyasla daha kolay bir şekilde en iyi ürün çeşitliliğine ulaşılabildiği gösterilmektedir. Sayısal analizler ile sezgisel algoritmanın performansı ve en iyi ürün çeşitliliğinin model parametrelerine hassasiyeti test edilmiştir.

Anahtar Kelimeler: Harici Talep Modeli, İkame Ürün, Kapasite Limiti, Transfer-Satış, Ürün Çeşitliliği Planlaması.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Erdal Erel for his support during my reserach. I am thankful to his valuable feedback, worthwhile guidance and immense knowledge.

I am also grateful to Assoc. Prof. Nagihan Çömez Dolgan who always supports and inspires me with her precious instructions. Her guidance helped me in all time of my graduate study and throughout the development of this thesis. I would like to thank Assoc. Prof. Alper Şen for accepting to be in my thesis committee and his valuable feedback for the study.

I would like to state my special thanks to my fellows Tiniç, Shahid and Burçak, my dear schoolmates Ecehan, Merve, Gökçe and my dear colleagues Naz, Özge, Büşra, Mustafa for their valuable support through the process of researching and writing this thesis.

Finally, I must express my sincere and profound indebtedness to my dear family and to Emre Yazganoğlu for their steady encouragement. This accomplishment would not have been possible without them.

Last but not least, I am grateful to the Scientific and Technological Council of Turkey (TUBITAK) and Assoc. Prof. Nagihan Çömez Dolgan for supporting this work by TÜBİTAK grant 110M488.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. Definition and Background of Assortment Planning

In a rapidly expanding world through online markets with countless number of available sellers for a potential buyer, sellers feel the pressure of differentiating themselves to obtain a competitive edge. To be competitive, either a firm should offer innovative products or provide many products, or both at the same time. When innovation requires a complex research and development and long term investment, generally firms may prefer catching customers with different products. The term assortment refers to the set of products offered by a firm at each point in time. Assortment planning is the process of deciding (i) the variety or breadth (number of categories), and (ii) depth (number of products in a category) of the assortment as well as determining (iii) the corresponding inventory levels for each offered product. Offering the variety comes with its relevant complexities and costs as well. So, the aim of assortment planning is to offer the optimal variety of products to customers to maximize the total profit from sales with respect to given costs and limitations of this variety.

Briesch, Chintagunta, and Fox (2009) report that customers' store choice decisions can be more sensitive to assortments than to prices. Firms need to devote effort to make periodical assortment planning considering changing customer preferences over time and seasons and the launch of new products to the market (Kök, Fisher, & Vaidyanathan, 2015). The assortment planning decision is quite complex due to several trade-offs it contains between a rich and limited assortment. On one hand, while a rich assortment increases the customer traffic with higher variety, a narrow assortment with a lower depth can help customers to decide easier, so that the probability of purchase increases (Mantrala et al., 2009). From the firm's perspective, while a rich assortment can build a good image for the competition, the management of stocks gets more difficult with the broadened assortment. When there are more products offered, the demand per product decreases that leads to higher coefficient of variation per product so higher overage costs. Besides, firms may be constrained with space and budget requirements of the offered variety. Therefore, a balanced methodology should be followed so that both firms can increase their profitability and customer expectations can be met.

When customers visit a firm, they typically demand a specific product. The unavailability of the product may lead them to consider either leaving the firm without any purchase or switch to another product (Mahajan & Van Ryzin, 2001). The act of switching to an alternative product when the favorite product is unavailable is known as substitution. The substitution can occur because either the product is not available in the assortment, called as assortment-based substitution, or there is a shortage in the inventory of the product, called stock-out-based substitution (Mahajan & Van Ryzin, 2001). Corsten and Gruen (2004) report that almost half of the customers may tend to switch to a different product when their favorite is not available. The existence of customers' substitution behavior can complicate the firms' assortment choices further.

Depending on the type of the operating system of the firm, some of the above concerns can become more vital or be irrelevant. For the firms working in make-to-stock environment, the inventory management is a crucial part of the assortment planning. The firm needs to consider the inventory space limitations and the financial and operational expenses of inventory management of the selected assortment such as overage costs (Li, 2007). In following, stock-out-based substitution becomes relevant. In make-to-order systems, as no significant inventories are kept for the final products, the assortment planning mainly contains two levels, determining categories and the depth of each category. Besides, only assortment-based substitutions should be considered. Alptekinoğlu and Grasas (2014) show how assortment decisions differ in make-to-order and make-to-stock systems.

1.2. Assortment Planning in Transshipment Systems

According to Corsten and Gruen (2004), customers who do not prefer to substitute their favorite product with another when they face with a stock-out may instead prefer to visit another store with the hope of obtaining from there. They report that among those surveyed, nearly one third mention that they can search for the product in another store. Thus, from the perspective of the initially visited firm, such a behavior of the customer may mean a lost sale. However, companies who are aware of such a customer tendency may turn a possible lost sale into an opportunity by collaborating with others in the same echelon.

Two or more companies may agree to share their inventories by sending goods from the one who has stock to another, which is out or short of stock. The transfer of a good between two companies in the same echelon is called a *transshipment*. The system of companies using transshipments can be called an inventory sharing or transhipment system. Transshipments are often utilized in various retail sectors such as automotive, apparel, sports goods, furniture, and shoes as well as in after-sales services for spare parts in airline and automotive sectors, and also between production facilities (Kukreja, Schmidt, & Miller, 2001), (Narus &

Anderson, 1996), (Özdemir, Yücesan, & Herer, 2006), (Rudi, Kapur, & Pyke, 2001) and (Kranenburg & Van Houtum, 2009).

Transshipments enable a firm to fulfill his customer expectations via other firms and prevent customers from leaving the firm with no purchase. Generally transshipments are quicker and cheaper alternatives to emergency orders to the suppliers, which are usually located at a further distance and reluctant to send small size and irregular orders. Therefore, nearby peer companies are preferred to get quick inventory as transshipments. On the other hand, the firm who sends transshipments gets the opportunity of utilizing her excess inventory, which may be prone to diminish and cause overstocking cost. Transshipments can be sent before a company actually stocks out in expectation of future shortage, called preventive transshipments or after the company stocks out, while there are customers waiting to be satisfied, called lateral transshipments.

The commonly investigated question about transshipments is their effects on regular inventory replenishment decisions. While it might be expected to observe decreased inventory levels with the use of transshipments as a risk pooling strategy, Çömez, Stecke, and Çakanyıldırım (2012) and Yang and Schrage (2009) report that transshipments may lead to even increased stocks. Another question about the exercise of transshipments is that, when transshipment policy is not present, the decision of when to send/receive transshipments should be endogenously made depending on the cost of transshipment, demand pattern and overstocking cost of sending party, and stock-out cost of the receiving party. It is shown in the literature that even in centrally managed systems, it can be better to holdback the inventory of a requested firm for his own future demand and not to send a transshipment to a stocked-out firm, when the transshipments are costly and time consuming (Çömez, Stecke, & Çakanyildirim, 2012). In decentralized and individually managed systems, how to split the gain from the transshipment between the receiver and the sender is also a question to be answered. The payment can be in terms of either a unit transshipment price to be paid per transshipment (Hu, Duenyas, & Kapuscinski, 2007) or allocation of the total profit gain from all demand satisfied through transshipments by the end of a selling season (Sošic, 2006).

Almost all studies on transshipment systems, except Akçay and Tan (2008), Tan and Akçay (2014) and Dağ (2015) consider transshipments for stock-outs that occur due to mismanaged inventories. It can be called as stock-out-based cooperation. In all these studies, it is assumed that all firms cooperating through transshipments offer the same set of products, which is mostly considered as one, a single common product. As customer substitutions may be exercised both stock-out-based and assortment-based, transshipments can be utilized for out of assortment products as well. Especially, when assortment costs are high and/or assortment size is constrained, a company may need to operate with a limited assortment, which is composed of most popular and/or most profitable products. Still, customer demand for the out of assortment products might be worth satisfying. In such a case, cooperating with another firm that keeps out of assortment products can increase profits despite the transshipment costs.

As the availability of transshipments affect the regular inventory replenishments in case of stock-out-based transshipments, when assortment-based transshipments are allowed, assortment planning should consider the possible transshipments. Under a centrally managed system, it might be optimal to keep a product in the assortments of only some of the firms so that in case a need at other firms, transshipments can be exercised. If the transshipments are cost-free, it is expected that it is enough to keep each product in the assortment of only one firm and use transshipments in case of a need. Otherwise, assortments of all firms become intertwined and requires the solution of a more complex problem then the assortment planning of a single firm.

1.3. Overview of This Thesis

This thesis aims to study optimal assortment planning for make-to-order firms under assortment capacity constraint when customers' substitution behavior between products are also explicitly considered. The existence of substitution among products complicates the assortment planning problem, thus invalidates the use of simple greedy algorithms. Our aim is first to shed light on the properties of optimal assortments to understand the effects of substitution and other system variables on the choice of products in the assortment. Then by using optimality properties to introduce algorithms to solve assortment planning problems in most efficient way optimally or close-to optimality.

Assortment planning is done either to maximize net profits by accounting for assortment costs or to maximize revenues under assortment constraints. Assortment cost per product kept can be incurred because of fixed costs of material handling, labor cost for merchandize presentation, record keeping and reordering (Smith & Agrawal, 2000) or warehousing, monitoring, personnel, computer time, and shelving (Dukes, Geylani, & Srinivasan, 2009). While it might be more challenging to compute fixed assortment cost per product, constraints on assortment size often rise in practice inevitably. Physical space availability is one of the most common constraints on the size of assortments, especially in a brick-and-mortar environment. For a make-to-stock firm, space limitation can be more restrictive as significant inventories can be kept for the products offered. Space limitation can be in terms of both the shelf space for visualization of the products and also the stocking space for the inventory kept. Most retail firms especially in downtown areas or premium malls work with no back room for inventory such as Tesco grocery stores in Europe (Scdigest.com 2007). For a make-to-order firm, space availability can still be constraining as for the products in the assortment, sample models might be required to be kept. Besides, for each type of product to be offered and produced, special equipments might be needed that occupy space. If all products have symmetric or similar space requirements, then the space constraint can be reduced to cardinality constraint, which limits the number of products in the assortment.

Budget constraint can be another significant limitation on the choice of assortments. When a product is decided to be offered to customers, the assortment cost can be both due to fixed costs of offering the product such as handling and replenishment and also periodic operational costs such as inventory costs. Often, if a product is offered in the assortment, a certain amount of inventory should be guaranteed to be kept that has a non-zero inventory investment cost. Thus, all products in the assortment can be interacting through the total budget that is available for the assortment. If all products have symmetric or similar budget requirements, then the budget constraint can be reduced to a cardinality constraint.

We first study the assortment problem for a single firm and then for a system of firms, which are interacting through sharing their products called transshipments. We consider firms that operate under a make-to-order system, thus inventory decisions are not relevant to assortment planning. The model can be also applied to make-to-stock or retail firms for which stocking decision is not constraining the assortment decision. For these firms, either the inventory may be low such as for the case of the slow-moving goods or all inventory is not carried on the shelves, but in a depot with limited facings on the shelves (Kök, Fisher, & Vaidyanathan, 2008).

For the single firm assortment problem, if a customer can find his favorite product, he purchases. If he cannot find his favorite product, he can decide to substitute with his second favorite product with a certain probability or may leave with no purchase. If the second favorite product is not available, customer does not search for a second substitute. It is shown that the optimal assortment is in the popular set, i.e., includes most preferred products, and the capacity is always fully utilized if products have different preference rates, but equal marginal profits. When products also vary in their marginal profits, the optimal assortment cannot be obtained with a greedy algorithm. So some optimality properties are shown such as an upper limit on the substitution rate for a pair of products below which a more profitable product is always preferred than the less profitable one. It is proved that in general when the substitution rate is smaller it is more probable that the assortment capacity is fully utilized and optimal assortment is composed of most profitable products. When assortment capacity is high, capacity utilization may decrease and the optimal assortment may include some less profitable products and exclude some more profitable ones.

For a multi-firm system, customers may have different experiences as follows. If a customer can find his favorite product at the visited firm, then he leaves by purchasing the good. If his favorite product is not available at the visited firm, but available at another firm in the system, then a transshipment is initiated for the customer, which incurs a transshipment cost for the system with the same revenue from the customer. If the favorite product does not exist in neither visited nor another firm's assortment, then the customer can make a substitution with his second favorite product with a certain probability. The second favorite product is first searched within the assortment of the visited firm, then in another firm's assortment for transshipment if it is not available. Thus, transshipment is always initiated first before a customer looks for a substitution. Because there is always a probability for the customer not to look for a substitution, which may lead to a loss sale for the system.

The resulting problem is to find the best assortments at all firms so as to maximize the profit coming from direct sales of these firms, substituted demands, and also transshipped products with respect to the capacity constraint of each firm. It is shown that the common assortment of all firms is in the popular set if all the products have equal marginal profits. However, in contrast to a single firm system, even under equal product profit margins, the individual assortment of each firm might not be in the popular set, as well as the total assortment. Moreover, assortment capacities of some firms might not be fully used. The reason is that by keeping some products out of the assortment, costly transshipments for these products can be avoided. Instead, the system is forced to use demand substitution for the available products. Then a limit on substitution rate for each particular firm is defined, below which the firm's assortment capacity is fully utilized at the optimality. It is shown that capacity utilization of a firm may increase as the products become more profitable compared to transshipment cost and also demand rate of the firm increases. Moreover, it is proved that if all firms have the same assortment limit, there is no incentive for under utilization. Thus all firms fully utilize their capacities.

We contribute to the literature on single firm assortment planning by analyzing an exogenous demand model with unequal demand and profit parameters under a capacity constraint. We are able to show that while the assortment capacity is fully utilized by the most preferred products when profit margins are equal, if it is not so, counter intuitively both a higher margin and more preferred product can be put out of the assortment. The result is due to existence of substitution where the absence of a popular and high margin product can bring higher profit by directing its potential customers to even higher margin substitutes. This is an example of previously observed "bait and switch" event, but at a higher degree. We also introduce limits on the substitution rate below which the firm can fully utilize its assortment capacity and more profitable products are preferred over less profitable ones.

We contribute to the currently limited literature on multi-firm assortment planning by jointly studying optimal assortment planning of multiple firms that are cooperating through transshipments. To our knowledge, this study is the first attempt to study properties of optimal assortments for the firms using transshipments and having assortment capacities. We can show that even if all products have equal margins, the optimal assortments of firms may not be obtained with a greedy algorithm. Transshipment costs lead the central planner to avoid some popular products not to utilize excessive transshipments for these products. Moreover, some firms' capacities can be left underutilized for the same reason especially when substitution rate is high enough. However, when substitution rate is low, assortment capacities can be fully utilized not to lose customers.

The rest of the thesis is organized as follows. In Chapter 2, we present a review of studies in the literature, which are related to our work. In Chapter 3, we model and solve the constrained assortment problem of a single firm. We also provide the results of numerical analyses to further explain and illustrate the properties of optimal assortment. We describe and study the multi-firm assortment planning problem under transshipments in Chapter 4. Finally in Chapter 5 we conclude with our final remarks by summarizing our findings and obtained insights with a follow-up research perspective.

CHAPTER 2

LITERATURE REVIEW

Our study aims to open a new window into assortment planning problem by considering the collaboration opportunity through transshipment while the assortment decision is made. Thus, it is related to two well studied topics, assortment planning and transshipment systems. Studies on transshipment systems are relatively new compared to assortment planning and it has been mainly subject of supply chain management. Assortment planning has been studied by both operations management and marketing researchers. For both topics, researchers analyze different trade-offs by considering various problem characteristics aiming to obtain easily usable and optimal, if possible, solution procedures for the practice. In this chapter, we present a survey of the related literature first on assortment planning, then transshipment systems, and finally on assortment cooperation models, which is quite limited.

2.1. Assortment Planning Studies

The extensive literature on assortment planning dates back to 1950's where Sadowski (1959) is probably the first to call "assortment problem" (Pentico, 2008). Past studies differ in terms of the model characteristics they consider such as consumer demand model, demand substitution pattern, inclusion of inventory level decisions, and existence of assortment capacity or not. Mantrala et al. (2009) differentiate assortment studies according to the resulting trade-offs in optimization from the consideration of these characteristics from consumers', retailers' and environmental perspectives.

Bernstein, Kök, and Xie (2015) categorize the earlier work mainly according to demand models they use: the multinomial logit, exogenous demand, and locational choice models, which greatly affect the other problem characteristics that can be modeled such as substitution type. Multinomial Logit (MNL) is one of those most commonly used customer behavior models in assortment problems due to its robustness in practical estimation. Besides, it allows easy incorporation of other decisions such as pricing into the customer demand model. MNL is a utility-based model assuming that each customer visiting the store associates a utility with each product, where the utility can be decomposed into two parts, the deterministic and a random component. Customers making a decision from a discrete set of products are utility maximizing individuals. There are two main critics about MNL model. One of them is the Independence from Irrelevant Alternatives (IIA). This property states that for a customer, the ratio of choice probabilities for two products is independent on other available choices in the overall set. It means that omitting a product from the model will change the parameter estimates of all the remaining items at the same relative rate, which cannot be always correct. Yet, it is commonly applied in the area of market research, economics, logistics etc. A second weakness about the MNL model is the modeling of substitutions, which is quite restricted.

In an exogenous demand model, the demand for each product is ex-ante specified for all possible products, so does not depend on selected assortment. Besides, substitution behavior of customers are also predefined independent of the choice of assortment set. When the most favorite product of a customer is not available, either because of stock-out or being out of assortment, with a predetermined probability, the demand is substituted with the second favorite product, which is not necessarily within the available assortment. The number of substitutions can be fixed to a certain number or it can continue until an available product can be reached. The main shortcoming of exogenous demand model is that there is no underlying consumer behavior model defining demand rates, thus it requires significant data collection for application.

Locational choice model is another demand model that is utilized in assortment studies, relatively less compared to two other models. It has been firstly proposed by Lancaster (1975) as an extended study of Working and Hotelling (1929). Products are regarded as bunches of attributes and consumer choices are defined according to these attributes. The specific attribute that identifies each product is its location. Consumers are supposed to prefer the products in the nearest location and substitute with the products in the second nearest location and so forth. Although its concept in modeling the demand is the same as MNL, i.e. the utility-based approach, it does not require IIA property. Besides, substitution can be controlled by the firm by offering products in neighborhood location to be substitutes.

Generalized Attraction Demand Model (GAM) has been introduced by Gallego, Ratliff, and Shebalov (2014) as a customer demand model, which is a generalized model that may reduce to MNL or exogenous demand models as special cases. MNL ignores the consumer search option when the first choice is unavailable and this yields to the overestimation of recaptured demand. On the contrary, exogenous demand model ignores the switching option from direct demand and hence, yielding underestimation of overall demand. In addition to the direct attraction values, GAM considers switching attraction values as well. Thus, GAM is a more flexible method and captures demand dependencies in more detail.

In this thesis, we prefer to review the literature according to the existence of capacity limit on the assortment. Thus, in Chapter 2.1.1, we first review assortment planning studies that have objectives of maximizing profit. In Chapter 2.1.2, constrained assortment problems are reviewed. For an extensive literature survey on assortment planning, we refer to Kök et al. (2008), Pentico (2008) and Misra (2008).

2.1.1 Uncapacitated Assortment Planning Studies

Cachon, Terwiesch, and Xu (2005) study an assortment optimization problem where the consumers search for the best product to fit their expectations that may lead to lost sales for the retailer is explicitly considered. In previous studies, the consumers decision not to buy a product from the assortment set is defined by a no-purchase option, which is assumed to be independent of the assortment. In this study, consumer search is explicitly modeled where there is a cost of search for the customer and the retailer's assortment influences the search decision. The paper studies whether it is important for an assortment planning process to explicitly account for consumer search. They show that when there is a low chance that the same product is carried in assortments of other retailers, which is called the independent assortment search model, no-search model performs well. However, if there is a limited pool of products for other retailers to include, which is called the overlapping assortment search model, the retailer's assortment should be broadened to reduce the value of search for consumers. Thus, in such a case, assuming no-search may result in assortments that are less profitable than the optimal assortment. In general, the no-search assumption performs better as search costs increase because consumers tend to search less.

Kök and Fisher (2007) study an assortment planning problem to find an algorithm to offer the best assortment with the highest expected gross profit in which consumers may substitute their first favorite product in case of unavailability. They contribute to the literature by presenting a real data to show how assortment planning works in practice. Additionally, they develop a method to estimate parameters of the model when sales summary data is available. Lastly, they enhance an iterative heuristic to find structural properties of the assortment such as deciding the priority order of the products in the assortment.

Li (2007) presents a single-period joint assortment and inventory optimization problem with MNL choice model. The paper extends the model of van Ryzin and Mahajan (1999) by allowing unequal cost and preference parameters and presenting an exact procedure to find the optimal assortment for the continuous store traffic. It is shown that the optimal assortment should include the first few items that have the highest profit rate which is a particularly defined metric taking profit margin, overage cost and demand variation into account. When the store traffic is a discrete random variable, demand for each individual variant in the assortment is split according to the incremental splitting rule. As it is more complicated, the optimization method for continuous traffic is used as an heuristic for the discrete demand model.

Schön (2010) presents an approach to deal with product line selection problem of a firm chasing a personalized or group pricing strategy which has been popular recently, especially in e-business. The purpose of product line selection problem is to decide how many products to offer, how to differentiate them along key factors, and how to price selected products. The practice of charging individual product prices distinguishes the paper from the other studies on product line selection problem. The model is applied to a data set determining a product line for an IT infrastructure provider, Alpha and the existing product yields with personalized and uniform pricing are calculated. The profit improvement of price discrimination is found to range from 3% to 20%.

Davis, Gallego, and Topaloglu (2014) aim to find an assortment to offer where customers are governed by nested logit model. Relaxing one of the assumptions makes the problem NP-hard so that the authors provide set of candidate assortments with worst case performance guarantee. Additionally, they present a convex program whose optimal objective value is an upper bound on the optimal expected revenue in order to test the optimality gap. The main contribution of the study is the classification of the complexity of the assortment problem with respect to the magnitude of the nest dissimilarity and presence or absence of the no purchase alternative within a nest. So as to solve the assortment optimization problem, four different cases are examined. In the first case, dissimilarity parameter of the nests is taken as smaller than one and it is assumed that customers always make a purchase within the selected nest. In the second case, it is assumed that dissimilarity parameter can take any value but customers always make a purchase within the selected nest. The third case takes dissimilarity parameter as smaller than one but the customers may leave a chosen nest without any purchase. Lastly, in the fourth case there is no restriction on the dissimilarity parameter and the purchase behavior.

Alptekinoğlu and Grasas (2014) study the effect of return policy on product assortment decision. Two basic operational environments, make-to-order and make-to-stock, are analyzed separately as they have different structures of optimal assortment. Nested multinomial logit is used as a demand model to show the consumer's preference and the post-purchase behavior is specified as either keeping or returning the product. They conclude that with a strict return policy where the refund is sufficiently low, it is optimal to have a mix of the most popular and most eccentric products. However, with a lenient return policy where the refund is high, it is preferable to carry the most popular products in both make-to-order and make-to-stock environments. When the fixed cost for variety is neglected, strict return policy case for make-to-order environment differs from make-to-stock such that the optimal assortment is only composed of most eccentric products. Additionally, they show that it is possible to offer variety with more lenient return policy. In summary, they conclude that retailers should carefully consider return policies when determining their assortment.

2.1.2 Capacitated Assortment Planning Studies

In following, we provide a review of assortment planning problems subject to certain constraints such as cardinality, i.e., number of products, and space consumption.

Smith and Agrawal (2000) focus on product variety in the retailing environment to analyze the impact of retail assortments on inventory management and customer service. They present an exogenous demand model to observe the effects of substitution and a methodology for deciding item inventory levels in order to maximize total expected profit subject to resource constraints. Different from the other multi-item retail inventory systems, random substitution pattern is also allowed. As a result of illustrative examples, they show that substitution effects lead to reduction in the optimal number of items to stock when fixed costs are present. Even when the fixed costs are neglected, it is not always optimal to offer all items in the assortment when items have different profit margins. Lastly, when items have non-identical demand rates it might not be optimal to stock the most popular item.

Fadıloglu, Karaşan, and Pınar (2010) provide an optimization model to have the optimal product mix to eliminate the product pollution on the shelf so as to maximize the profit. The optimum SKU list is found by deciding for the SKU that needs to be eliminated from the list to prevent product pollution. It is assumed that customers who cannot find the SKU they look for may substitute the SKU for another one within product categories. There is a constraint to ensure that ratio of demand corresponding to undeleted SKUs to the total demand is not less than minimum conservation ratio. They contribute to the literature by proposing an optimization model that gives the best product mix with minimal data set, therefore the model can be easily used by any retailer that keep track of its sales. The proposed model is applied to the shampoo product class at two regional supermarket chains and tested with real data.

Misra (2008) analyzes a joint assortment and price competition of retailers which offer exclusive products with MNL demand under the display constraint. He carries out an empirical study to investigate the impact of competition on assortment size and prices and observes that reduction in the assortment size leads to a reduction in the total profit. Besides, competition between retailers provokes increase in the total number of products offered in the assortment and reduction in the prices. The results gathered from the study focus on the best response analysis. Although the methodology is beneficial for retailers to decide their assortments across multiple categories, it does not give an equilibrium or uniqueness results.

Rusmevichientong, Shen, and Shmoys (2009) deal with an assortment optimization problem under the nested logit choice model to maximize the total profit. The problem is formulated as an integer programming problem that reduces to a knapsack problem. They come up with a polynomial time approximation scheme for the sum of ratios optimization problem with a capacity constraint and a fixed number of product groups.

Rusmevichientong, Shen, and Shmoys (2010) study an assortment optimization problem in which a retailer chooses an assortment of products to maximize the profit under capacity constraint. Both static models where customer preferences are known and dynamic models where parameters are inferred by offering different assortments over time and observing the customer selections are formulated. They develop a geometric nonrecursive polynomial time algorithm under multinomial choice model, thus structural properties of the static algorithm are exploited for the dynamic optimization problem. The formulation assumes that each customer independently subscribes to the same choice model; nonetheless, this assumption is not appropriate when there are various markets and the customers within the market are heterogeneous. This study contributes to the literature by discovering the previously unknown relationship between customer preference and the optimal assortment in a capacitated setting.

Wang (2012) focuses on an assortment and price optimization problem where a retailer chooses an assortment of competing products and determines their prices so as to maximize the total expected profit subject to a capacity constraint. MNL model is preferred as a demand model. The author shows that the size of an optimal assortment problem is equal to capacity, which is different from the traditional capacitated assortment problem where the prices or margins are fixed. When prices are fixed, some low-margin products can be set out of the assortment to direct customers to those high-margin ones. However, when the prices are determined along with the assortment, the assortment capacity can be fully utilized with selected prices. Wang (2013) also explores an assortment optimization with a capacity constraint, but unlike the previous studies, uses a generalized attraction model. The author provides an efficient algorithm to find the optimal assortment in polynomial time which is more influential than linear programs and guarantees the optimal feasible solution and helps to establish a time threshold structure for a dynamic capacitated assortment problem.

Davis et al. (2014) examine assortment optimization problems under the nested logit model where the products are arranged in nests and each product in each set with a fixed revenue. Feasible set of products are analyzed under cardinality and space constraints to maximize the expected revenue. They are able to solve a linear program to obtain an assortment that provides the largest expected revenue with cardinality constraint. However, the problem with the space constraint is NP-hard, so it is solved by a tractable linear program to have an assortment with a certain performance guarantee. The study contributes to the literature by examining expected revenue in the form of a fraction of two non-linear functions. So the results are more general when it is compared with MNL modeling.

2.2. Transshipment Planning Studies

Inventory management is one of the earliest and highly investigated topics of operations research/management. It aims to determine the timing and quantity of inventory intake to match the available stock with the demand. The main stream inventory management investigates the regular inventory purchase decisions and assumes that any unmatch between these purchased quantities and the random demand will result in demand dissatisfaction costs such as lost sales or backorders. As the cost of unsatisfying demand increases and firms become more customer-centric, the need and effort for alternative demand satisfaction methods raise. Making emergency orders from the supplier (Moinzadeh & Nahmias, 1988), using multiple suppliers, some as regular for low price and some as quick emergency suppliers despite higher prices (Tomlin, 2006), splitting outstanding orders for partial shipments (Thomas & Tyworth, 2006), and expediting outstanding orders by either fastening production and/or transportation at an extra cost (Cheng & Duran, 2004) are some of the alternative demand satisfaction methods. When the supplier is located far away and/or cannot provide quick and cheap fixes to regular inventory shipments, the inventory sharing between the same echelon members such as between retailers is started to be commonly used recently as an alternative demand satisfaction method. Inventory sharing is called as "inventory pooling", "dealer trade", or "transshipment" in the literature.

The basic question on transshipments is their interaction with regular inventory replenishment decisions. Thus, earlier studies on transshipments model predetermined transshipment policies and focus on the effects of these transshipments on replenishment decisions. Transshipment policies can be preset such as at the end of each inventory review period, after all demand is realized and before it is satisfied. Transshipments are made so that excessive inventory at one firm is sent to another to satisfy excessive demand as much as possible called "complete pooling". Among these complete pooling studies, Robinson (1990) model a finite horizon periodic review system and show that a stationary base stock policy is optimal for regular replenishments in a centralized two-location transshipment system, but can be used as a good heuristic for a multi-location setting. Rudi et al. (2001) analyze a single period system of two independent retailers and show that there is a coordinating transshipment price that can lead retailers to order at system-optimal quantities at the beginning of the period.

There is another line of studies, which also preset transshipment policies, not to complete pooling but to some partial pooling policy, which is not necessarily optimal. Grahovac and Chakravarty (2001) use a one-for-one replenishment policy, i.e., whenever a demand is realized, immediately a replenishment order is sent to the supplier, along with a transshipment policy in which a transshipment can happen after each demand. When the firm is out of stock to satisfy a demand, the firm either places a regular order if his net stock is above a predetermined level K, or an emergency order first to the distributor, or to another firm for a transshipment if the net stock is not above K. If a transshipment request is made, the requested firm can share a product if she has more than $K+1$ items. Thus, for both requesting and accepting a transshipment, the same threshold level is used.

Zhao, Deshpande, and Ryan (2005) define an (S, K) policy for transshipments and replenishments such that S is the order-up-to level for inventory replenishment and K is the threshold transshipment level, above which transshipment requests would be accepted by a requested retailer. Zhao, Deshpande, and Ryan (2006) extend both Grahovac and Chakravarty (2001) and Zhao et al. (2005) by considering a base-stock replenishment policy, a threshold level for sending transshipment requests, and another threshold for filling requests.

Another important question on transshipment systems is how to set an optimal or a close-to-optimal transshipment policy when transshipments can be initiated before all demand uncertainty is resolved. Otherwise, complete pooling policy is trivially the optimal one. The transshipment policy should be defining when to ask for a transshipment, from whom to ask (for more than two-firm systems), how to respond to a transshipment request, and how to share the profit from transshipments when firms are decentrally owned.

For a centrally managed two-firm model, T. Archibald, Sassen, and Thomas (1997) allow multiple transshipments during a replenishment cycle in a continuous time model with independent Poisson demands. They consider an emergency order when a transshipment is not available. They obtain optimal threshold levels of time to manage transshipments such that only if the remaining time to next replenishment is below the threshold time, a transshipment should be initiated and an emergency order from the manufacturer should be obtained otherwise.
T. Archibald (2007) and T. W. Archibald, Black, and Glazebrook (2009) extend the problem to a multi-firm system by introducing a heuristic policy using respectively, transshipment cost and transshipment cost along with the expected value of the inventory, to make a transshipment decision based on the results from a two-firm system. Çömez, Stecke, and Çakanyildirim (2012) work on a centralized inventory sharing system of two retailers that are replenished periodically, the

unsatisfied demand is backordered, and there is a non-negligible transshipment time. The optimal transshipment policy is shown to be characterized by hold-back inventory levels that are non-increasing in the remaining time until the next replenishment, which is opposite to what is obtained by T. Archibald et al. (1997) for a lost sale environment.

The investigation of optimal transshipment policies for decentralized systems is relatively more demanding. Çömez, Stecke, and Çakanyıldırım (2012) study a system of competing retailers for a single season. The requested retailer has the incentive of not sending a transshipment because of the demand overflow probability between the retailers. The paper provides a flexible sharing mechanism that is regulated by dynamic inventory holdback levels that change during the season. For the systems with more than two retailers, the optimal solution of a two-retailer system is used to develop a heuristic policy. Çömez-Dolgan and Fescioğlu-Unver (2015) study transshipment policies among multiple retailers. They aim to decide which retailer will be asked to take place in transshipment. Various heuristic policies are examined, according to the availability of inventory level information by the requesting retailer and the transshipment respond policy of the requested retailer. They conclude that the performances of the heuristic policies are satisfactory especially in large systems and it might be better for a retailer to select the requested retailer randomly if there is no information symmetry about inventory levels or salvage prices of others.

2.3. Assortment Cooperation Studies

Assortment planning literature is mostly concentrated on single product category of a single firm as it is already complicated enough because the set of products in an assortment are intertwined through either total assortment budget, total assortment capacity, substitution among products, or all at the same time. The literature on joint assortment planning of several product categories and/or multiple firms is very scarce. Kök et al. (2015) mention the opening in this stream of assortment planning research. Multi-category assortment planning is important as consumers' demand in one category may be related to the availability of products in another category called basket effect. Multi-firm assortment planning is needed to account for inventory competition (Netessine & Rudi, 2003) and also cooperation (Çömez, Stecke, & Çakanyildirim, 2012).

Among these limited number of studies, Cachon and Kök (2007) analyze assortment planning for multiple product categories at a retailer. They show how individual category managers at a retailer can make suboptimal price and variety decisions for each category ignoring basket shopping consumers who can purchase from multiple categories. It is shown that category managers choose less variety and higher prices than those at system optimality.

Dukes et al. (2009) consider a game-theoretic setting of a manufacturer and two retailers, where manufacturer offers two products in his product line. They show that if one retailer has the channel power to determine its assortment first, then it can strategically offer a shallower assortment by carrying only the popular product. This makes the other retailer to carry both products inducing higher assortment costs, which leads to relaxed price competition for the commonly carried product. Moreover, when the manufacturer has the channel power, he prefers to carry both products.

Besbes and Sauré (2016) analyze the equilibrium behavior for product assortment and price competition in a duopoly of retailers under display constraints. They model two different settings, assortment only competition where prices are exogenously fixed, and joint assortment with price competition. For the assortment competition, there are two specific cases, one with the exclusive products where the sets of available products do not overlap and with common products. They analyze the existence of equilibrium solutions individually where there is only assortment or both price and assortment competition. They state that retailers may not want to offer the maximal number of products but the ability of modifying prices enable them to use the capacity fully.

Kök and Xu (2011) study an assortment planning and pricing problem for a product category with two brands. They model two different settings, a brand-primary model where consumers choose a brand first, then the product type in the brand and a type-primary model where consumers choose a product type first, then decide the brand. They consider both centrally and decentrally managed systems of two brands. They show that for the brand-primary model, both the centrally and decentrally obtained assortments for each brand include the most popular product types within each brand. With the type-primary choice model, the overall assortment includes a set of most popular products. They emphasize the importance of making an assortment planning process that is aligned with the hierarchical choice process of targeted customers.

There are three studies, which are relatively closer to our study. Akçay and Tan (2008) study the assortment cooperation of independent producers to offer a combined set of products to their customers. They investigate the characteristics of firms and their products that would facilitate a beneficial cooperation. First, they model firms each offering a single product and the essential conditions are analyzed for the beneficial cooperation. Second, firms with overlapping products are modeled and the benefit of the cooperation is evaluated with respect to some key dimensions such as product market shares, profit margins, substitution effect among products, firm sizes, and firm assortment. Tan and Akçay (2014) also deal with assortment-based cooperation for two firms, but with a constraint on production capacities. They present a continuous time Markov chain model both for centralized and decentralized firms. They identify a case where firms use the benefit of assortment cooperation and evaluate the impact of key problem parameters such as substitution behavior of customers. They show that assortment-based cooperation is always beneficial if two firms are symmetric. Both of these studies aim to understand the effects of assortment-based transshipments on system profits under given assortments ignoring assortment planning decisions. Exceptionally, Dağ (2015) study the assortment planning problem of a system of multiple firms using transshipments. Different than our study, they do not define assortment capacities, but assortment costs. They prove that the assortments of firms should be nested, i.e., the assortment of a firm should be the subset of another firm's assortment whose market size is equal or larger.

CHAPTER 3

CONSTRAINED ASSORTMENT PLANNING OF A FIRM

3.1. Synopsis

In this chapter, we deal with an assortment planning problem of a make-to-order firm whose product portfolio can include a limited number of products due to a capacity constraint, which can be called as a cardinality constraint. Consumers in the market have certain demand rates for the products possible to be offered by the firm. The goal is to maximize the total profit of the firm from the sales with respect to the capacity limitation. The assortment problem is first analyzed for products all having equal profit margins, but specific customer preferences. It is shown that in such a case, the optimal assortment can be obtained by a simple greedy algorithm. When profit margins are not identical, the best assortment cannot be obtained by a greedy algorithm under most generic problem parameters. So, by the investigation of the objective function, some properties of the optimal assortment are obtained that can simplify the computations to obtain the optimal solution. In Chapter 3.2, the analytical model is defined in detail. The properties of optimal assortment are shown in Chapter 3.3 and the decrease in computational efforts due to these properties is explained in Chapter 3.4. Finally, the results of numerical analyses on optimal assortments are discussed in Chapter 3.5.

3.2. Problem Definition and Model

We study the assortment planning problem of a firm, who needs to decide the portfolio of products to offer out of a possible set. Each product in the possible set has a predetermined customer preference and a profit margin. The total assortment size is restricted by a cardinality capacity constraint due to physical and/or financial limitations. The products to be chosen in the assortment are related due to this capacity constraint as well as the substitution tendency of customers for unavailable products. The objective is to determine the assortment that will maximize the total profit from this assortment subject to the capacity constraint.

The total expected number of customers in the market is denoted by λ. The possible set of all potential products is N. The demand for each product in N is known given that customer preferences are independent of offered assortment. This demand model is known as exogenous demand model in the literature (Smith & Agrawal, 2000), (Yücel, Karaesmen, Salman, & Türkay, 2009), (Kök et al., 2015). Each product i has a certain probability of being the first choice of a visiting customer denoted by α_i such that $0 < \alpha_i \leq 1$ and $\sum\limits_{i \in N} \alpha_i = 1$.

An arriving customer pays r_i to the firm to buy product i, if his most favorite product i is in the assortment. If product i is not offered by the firm, then the customer may substitute with another product with probability $\theta \leq 1$. The substitution from an unavailable product can be defined in different ways such as randomly, to any other product with the same probability, adjacently, to neighboring products according to some attributes, or proportionally, according to demand rates of other possible products (Kök et al., 2015). Here, we define proportional substitutions. Given that the customer decided to substitute his first choice with a second one, the probability that the customer substitutes product i with product j is given by

$$\delta_{ij} = \left(\frac{\alpha_j}{1 - \alpha_i}\right)$$

where δ_{xx}= 0. If the firm has the second favorite product j, then the demand of the customer is satisfied. Otherwise, the customer leaves the system without any purchase.

In this thesis, we do not consider inventory decisions. So, stock-out-based substitutions are out of scope for this study. Consequently, the term substitution refers to only assortment-based substitution. This form of demand substitution is also known as static substitution and considered in the literature by other assortment studies (Besbes & Sauré, 2016), (Kök & Fisher, 2007), (Akçay & Tan, 2008), (Yücel et al., 2009). It implicitly assumes that any demand for the products existing in the assortment can be satisfied. Moreover, an unsatisfied customer is assumed to attempt a substitution only once, from her most favorite product to a second favorite one. A single substitution assumption is often used in the literature. Smith and Agrawal (2000) find it reasonable as it enables analytical tractability and results in simple closed form relationships among substitution probabilities. Kök (2003) also shows that allowing multiple substitution behavior does not bring a significant contribution to a model because it is possible to approximate multiple substitution behavior with a single substitution by increasing the substitution probability (Akçay & Tan, 2008).

The objective of the firm is to maximize total expected system profit $\Pi(.)$ subject to the capacity constraint. For this purpose, the firm should decide the best assortment S, which is the product portfolio to be offered. The total number of products in the assortment is denoted by $|S|$, where $|X|$ is the cardinality of a set X. There is a capacity constraint for the firm which is denoted by C. Then the assortment planning problem is as follows.

$$\max \Pi(S) = \lambda \sum_{i \in S} \left(\alpha_i r_i + \theta \sum_{j \notin S} \alpha_j \delta_{ji} r_i \right)$$

$$s.t., \ |S| \leq C$$

The profit function is composed of the expected profit from the direct sales of the firm for customers' first choice demands and the expected profit from the substitution of the first choice product from the firm's assortment. The notation used for a single firm setting is summarized in Table 3.1.

Table 3.1: Notation for single firm assortment planning problem

Parameters	
N	The set of all products
λ	Expected number of customers in the market
C	Assortment capacity of the firm
α_i	Demand rate(probability) for product i, $i \in N$
r_i	Profit margin for one unit of product i, $i \in N$
θ	Rate(probability) of substitution
δ_{ij}	Substitution probability from product i to product j, $i, j \in N$
Variable	
S	Assortment of the firm

3.3. Properties of Optimal Assortments

In this chapter, we first analyze the assortment planning problem under symmetric product profit margins ($r_i = r$), where products only differ in their customer demand rates α_i as in Alptekinoğlu and Grasas (2014) and Cachon et al. (2005). For products having identical profit margins, we can show that optimal assortment of the firm includes a set of most popular products. Thus, the optimal assortment is called to be in the *popular (assortment) set* P. Popular assortment is the set of products in descending order according to their purchase probabilities (Kök & Xu, 2011). The most popular product i, having the largest α_i is indexed by i=1 and accordingly remaining ones are sorted in decreasing order of α_i. Popular set also includes the null set that is $P = \{\{\}, \{1\}, \{1,2\}, .., \{1,2,..,N\}\}$. The finding is stated by Theorem 1.

Theorem 1. *When all products have the equal profit margin $r_i = r$, for $i \in N$, the optimal assortment of the firm*
(i) is in the popular set and (ii) fully utilizes the capacity C.

According to Theorem 1, at the optimal, the firm will fully utilize its assortment capacity and include only most popular products. Wang (2012) shows that the assortment capacity would be fully used when prices are endogenous. We can show that when profit margins are exogenously set at the same value, the assortment capacity would be also fully utilized as there is no incentive of directing customers from one product to another one by excluding some products from the assortment. As a result of Theorem 1, the optimal assortment can be easily obtained by including C number of products with the highest demand probability.

In a more general setting, where products may differ in their both purchase probability α_i and also profit margin r_i, it is crucial to consider not only the demand rate, but rather the expected profitability of each product, defined by $\alpha_i r_i$, which can be also called as *profit rate*. Let the product indices are set such that $\alpha_x r_x \geq \alpha_y r_y \ \forall \ x \leq y$, x, $y \in N$. The product x is called as *dominant* over product y. Let $a(i)$ denote the index of the product which has the i^{th} largest demand rate, such that $\alpha_{a(1)} \geq \alpha_{a(2)} ... \geq \alpha_{a(N)}$. Lemma 1 explains the priority order of products as a function of demand rates and profit margins.

Lemma 1. *Product x has always a higher priority than a less dominant product y, for $x \leq y$, to be in the optimal assortment*
(i) if $\alpha_x \leq \alpha_y$ (where $r_x \geq r_y$ will also hold trivially)
(ii) or when $\alpha_x > \alpha_y$, if θ does not exceed the threshold level $\bar{\theta}_{xy}$ such that

$$\bar{\theta}_{xy} = \frac{(\alpha_x r_x - \alpha_y r_y)(1-\alpha_x)(1-\alpha_y)}{(\alpha_x - \alpha_y)\sum_{i=1}^{|C-1|}\alpha_i r_i - \alpha_x\alpha_y(r_x - r_y) - (\alpha_x r_x - \alpha_y r_y)((1-\alpha_x)(1-\alpha_y)\sum_{i=C+1}^{|N|}\frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \alpha_x\alpha_y)} \quad (3.1)$$

If a product x is dominant over y ($\alpha_x r_x \geq \alpha_y r_y$), three cases are possible for the relationships between demand rates and profit margins. Case (i) $\alpha_x \leq \alpha_y$ and $r_x \geq r_y$: here as both the profit rate and the profit margin of product x is higher, it is easy to understand that x is more preferable for the assortment. It is advantageous for the system for product x to have a small demand rate, because then higher demand rate products with lower margins can be set out of the assortment and their demands can be directed to high margin products through substitution. This is a realization of “bait and switch”, where it is aimed to satisfy the demand for a low margin item by offering a higher margin substitute although the high margin product has lower preference by the customers.

Case (ii) $\alpha_x > \alpha_y$ and $r_x \leq r_y$: according to (3.1), x has a priority only if the substitution rate does not exceed the limit. If the substitution rate is above the threshold, product y can be preferred as it has a higher margin, despite the lower demand rate as a “bait and switch” tactic. The firm may benefit from bait and switch when substitution rate θ is higher as $1 - \theta$ denotes the probability of losing an unsatisfied customer and decreases with θ.

Case (iii) $\alpha_x > \alpha_y$ and $r_x \geq r_y$: this is the most interesting case. Although product x has a higher margin and also higher popularity, which directly leads to dominance of x over y, it is still not guaranteed that x has a higher priority than y

to be in the optimal assortment. In fact, the result can be illustrated by an example. Let $|N| = 4$ with $\alpha_1 = 0.4$, $\alpha_2 = 0.3$, $\alpha_3 = 0.2$, $\alpha_4 = 0.1$, $r_1 = 5.1$, $r_2 = 6$, $r_3 = 5$, $r_4 = 9$, $\theta = 0.9$, and assortment capacity $C = 3$. For this setting, the optimal assortment of the firm is $S = \{2, 3, 4\}$ with a total expected profit of $\Pi(S) = 592$. Thus, product 1, which has both a higher margin and higher demand rate than product 3 is out of the optimal assortment, while 3 is in. This is similar to realization of bait and switch, where a high margin substitute for a low margin one is offered. However, in this example, the demand of both a high margin and high popularity item is preferred to be substituted by a lower margin and lower popularity one. When the same example is repeated with $\theta = 0.8$ by keeping everything else the same, the optimal assortment of the firm becomes $S = \{1, 2, 4\}$. So, the result can be explained by the high demand rate of product 1, which also brings a high substitution potential to other available products when it is out of assortment. Particularly, when the substitution probability is high enough, a highly popular product, which is also high margin, can be set out of the assortment to benefit from the popularity of the product to direct customers to other even higher margin products. Li (2007) also shows a similar behavior but explains the situation due to high overage cost and resulting high demand variability of the high margin and high demand rate product where substitution is not allowed. However, in our model bait and switch can result due to high substitution rate.

Note that the limit (3.1) on the substitution rate is a sufficient condition, but not necessary. For the above example, $\theta_{13} = 0.63$. Thus, for any $\theta \leq \theta_{13} = 0.63$, product 1 is always preferred over product 3 in the optimal assortment. However, it is reported above that when $\theta = 0.8$, the optimal assortment is $S = \{1, 2, 4\}$. So, product 1 is preferred instead of 3 even if θ is above the threshold θ_{13}.

Theorem 2. *(i) If all products satisfy the relationships $r_x \geqslant r_y$ and $\alpha_x \leqslant \alpha_y$ for $x \leqslant y$, then the optimal assortment is composed of some number of most dominant* *products.*

(ii) Otherwise, it is more probable for the optimal assortment to include only some number of most dominant products, when θ and/or capacity C are smaller.

Theorem 2i states that if all products can be sorted such that for each product pair in the possible set N, a dominant product has a lower demand rate (which also means that it has a higher margin due to dominance definition), then the optimal assortment can be obtained by selecting a certain number of most dominant products. Thus, the optimal assortment can be obtained by a greedy algorithm by adding one product to the assortment at every iteration until $|S| = C$ and checking for the improvement in the total profit.

Theorem 2ii sheds light on the effects of substitution rate and assortment capacity on the structure of optimal assortment. It states that when the substitution rate is high, there is incentive to exclude some highly dominant products out of assortment. When the statement is evaluated together with Lemma 1, it can be concluded that a dominant product with a higher demand rate can be left out of the assortment to benefit from its demand rate to direct customer to other higher profit margin products given the high substitution rate. Similarly, when the assortment capacity is high, there can be more products in the assortment, which increases the probability of keeping high margin substitutes in the assortment for the out-of-assortment products. Thus, the optimal assortment may not be in the dominance set anymore.

The proof of Theorem 2ii on the effect of assortment capacity C has an additional insight as well. The proof is based on the result that as the capacity C increases, the marginal benefit of adding product x over product y decreases, which may pass through zero only once from positive to negative. This implies that once it is shown that it is more profitable to include product y over x, for $y \geq x$ in dominance order, it is never optimal to include x in the optimal by excluding y as the assortment is enlarged. This result is important in developing successful greedy heuristic algorithms.

While it is shown that the assortment capacity is fully utilized when all product margins are identical, it might not be the case when products also differ in profit margins. The reason is again for bait and switch strategy, i.e., to direct customers to high margin products by excluding low margin ones from the assortment. Theorem 3 proves that the firm may take a risk of shallowing the assortment and expect customers to substitute their demands with high margin ones when the substitution rate is high. It also introduces an upper limit on the substitution rate, below which the assortment capacity will always be fully utilized at optimality.

Theorem 3. *(i) The assortment capacity of the firm is always fully utilized at optimality if θ does not exceed the threshold level $\bar{\theta}_c$ such that*

$$\bar{\theta}_c = \min_{x \in N} \left\{ \bar{\theta}_x = \frac{-r_x}{r_x \sum\limits_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \frac{1}{1-\alpha_x} \sum\limits_{i=1}^{C-1} \alpha_i r_i} \mid \bar{\theta}_x > 0 \right\}.$$

(ii) Capacity utilization is non-increasing in θ and C.

We further elaborate Theorem 2 and Theorem 3 with a numerical example. Table 3.2 reports the optimal assortment sets for a problem instance run with different assortment capacities C and substitution rates θ, where $\alpha_i = \{0.27, 0.21, 0.19, 0.16, 0.13\}$, $r = (20, 15, 10, 10, 9)$, $\forall i = 1..5$, and $\lambda = 100$. The result shows how non-popularity and capacity usage are affected with respect to θ and C.

Table 3.2: Illustrative example of a single firm assortment.

C\θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	1	1	1	1	1	1	1	1	1	1
2	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,2	1,2
3	1,2,3	1,2,3	1,2,3	1,2,3	1,2,3	1,2,3	1,2,3	1,2,3	1,2,4	1,2,4
4	1,2,3,4	1,2,3,4	1,2,3,4	1,2,3,4	1,2,3,4	1,2,3,4	1,2,4,5	1,2,3,4	1,2,4,5	1,2,4
5	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5	1,2,4,5	1,2,3,4	1,2,4,5	1,2,4

Table 3.2 indicates that as θ increases both capacity usage (1) and product dominance (4) are non-increasing. Moreover, as C increases, capacity usage (2) and product dominance (3) are non-increasing too.

3.4. Notes on Complexity

The optimal assortment for a single firm can be found by full enumeration of all possible assortments. Thus, the number of assortments to evaluate and compare is

$$\sum_{i=0}^{C} \binom{N}{i}.$$

In this chapter, we show how optimal assortment properties obtained in Chapter 3.3 lead to a reduction in the number of necessary iterations to have the optimal solution.

When all products have symmetric margins, by Theorem 1, the optimal assortment can be easily obtained. The most popular C products should be selected for the optimal assortment. Therefore, there is no associated complexity.

When products may also differ in profit margins, Theorem 2i states that the optimal assortment is composed of some number of most dominant products if all products satisfy the relationships $r_x \geqslant r_y$ and $\alpha_x \leqslant \alpha_y$ for $x \leqslant y$. Accordingly, the number of iterations to find the optimal set decreases to C+1.

Where the optimal assortment size can be any number not above C.

Theorem 3i is about capacity utilization and asserts that the assortment capacity of the firm is always fully utilized at optimality as long as θ does not exceed the threshold level $\bar{\theta}_c$. Note that $\bar{\theta}_c$ is a sufficient condition, but not necessary. Thus, if the substitution rate is small enough, number of iterations to obtain the optimal solution is the capacity-combination of the set of all products, N as follows.

$$\binom{N}{C}.$$

If the conditions defined by Theorem 2i and Theorem 3i hold simultaneously, the result is straightforward. The optimal assortment is composed of C number of most dominant products. Hence, the result should be found directly.

In order to illustrate the reduction in complexity with the use of above optimality properties, we report the number of iterations to be evaluated to reach the optimal set for different number of products and assortment capacities. Table 3.3 presents the number of assortments to be evaluated under full enumeration, in the case when either Theorem 2i or Theorem 3i, but not both, hold.

As specified by Table 3.3, for a single firm with 10 products and a capacity limit of 5, it is possible to decrease the number of candidate assortments by 99,06% from 638 to 6 knowing that the optimal set is comprised of some number of most dominant products since $r_x \geqslant r_y$ and $\alpha_x \leqslant \alpha_y$ for $x \leqslant y$. Fully utilization of capacity at optimality leads to a 60,50% reduction in the number of candidate assortments from 638 to 252. In a scenario in which 10 of the products satisfy the dominance rule and the capacity is fully utilized, the optimal set is equal to 5 most dominant products.

Table 3.3: The number of candidate assortments to be evaluated to obtain the optimal solution

$\lvert N \rvert$	C	Full Enumeration	By Theorem 2i	By Theorem 3i
10	5	638	6	252
10	7	968	8	120
11	5	1024	6	462
11	7	1816	8	330
12	5	1586	6	792
12	7	3302	8	792
12	9	4017	10	220
13	6	4096	7	1716
13	8	7099	9	1287
13	10	8100	11	286

3.5. Computational Results

In this chapter, we present numerical analyses supporting the analytical properties we obtained for the optimal assortment of a single firm in Chapter 3.3. We try to observe the sensitivity of the optimal assortment to certain problem parameters and we present our heuristic approach respectively in Chapter 3.5.1 and Chapter 3.5.2.

Firstly, in order to do the sensitivity analysis we create randomly generated problem instances by changing the observed parameter in every 10 problems and calculate the average performance over each 10 problem instances. Second, we compare our approximate solution reached by heuristic model to the optimal assortment and test its performance. Heuristic model finds the approximate solution by including best number of most dominant products according to the dominance rule which is the multiplication of the profit margin with the customer demand rate. We analyze the gap between heuristic approach and full enumeration. Throughout the numerical analyses, we try to use a broad range of model inputs so as to increase the robustness of the results. Each randomly selected parameter is obtained according to Uniform distribution with the ranges summarized in Table 3.4. In all problem instances, it is set as $\lambda = 100$.

Table 3.4: Uniform distributions of randomly generated parameters

θ	α_i	r_i
U(0.3,1)	U(0,1)	U(1,10)

3.5.1 Sensitivity of Optimal Assortments

In this chapter, we analyze how the system inputs; probability of substitution θ, symmetry in purchase probability of products α_i and capacity of the assortment C affect the optimal assortment. To test the effect of each parameter, we first create 10 problem instances for a single firm where the number of products $|N| = 10$. For each problem set, we use randomly generated parameters with Uniform distributions by using the ranges mentioned in Table 3.4, except the related parameter. We set the analyzed parameter to a certain value for all 10 problem instances, obtain the optimal results for 10 instances, and calculate the average values for the measured performance metrics over 10 instances. We repeat the same 10-problem analysis several times, by setting the analyzed parameter to a different value each time.

All problem instances are solved to obtain the optimal assortments. Then performance metrics are obtained with the optimal assortments. Particularly, we use following metrics to measure the effects of system inputs.

- Average percent change in total profit ($\Delta\Pi$%)
- Average percentage of direct sales (DS%)
- Average percentage of sales for the substituted demand (SS%)
- Average capacity utilization (%).

In a single firm system, total revenue is composed of direct customer sales if the demanded product is available in the firm and sales for the substituted demand if the demand is met by the second favorite product. According to the objective function:

Direct Sales(DS) : $\lambda \sum_{i \in S} \alpha_i r_i$
Substitution Sales(SS): $\lambda \sum_{i \in S} \theta \sum_{j \notin S} \alpha_j \delta_{ji} r_i$.

Then the percentage of direct sales is DS/(DS+SS)*100. Similarly, SS%=SS/(DS+SS)*100. Thus, DS% + SS% = 100%.

The capacity utilization ratio for a problem instance is the ratio of actual number of products to be kept in the optimal assortment to the assortment capacity C. Then it is reported in percentage by multiplying with hundred. By using the total profit values of the first 10 instances as a basis, we measure the percent change in profit ($\Delta\Pi$%) for each problem instance as the concerned parameter value is increased from its basis value. For each fixed value of the concerned parameter over 10 problem instances, the average values of these four metrics are obtained and reported. With the help of this methodology, we aim to eliminate the bias selection effects of other parameters by reporting average performances over 10 problem instances generated randomly.

The sensitivity analysis is conducted with asymmetric profit margins r_i. According to Theorem 1, when all products have the same profit margin $r_i = r$, the optimal set is in the popular set and the capacity is always fully utilized. Thus, the optimal assortment is well-defined. Additionally, we keep the capacity of the system fixed as $C = 7$ to be able to report the change in capacity utilization while observing the effect of θ and α_i.

Firstly, we report the computational results for the sensitivity of optimal assortment to probability of substitution θ. 20 different values of θ are tested so that Table 3.5 reports the results of 200 problem instances in total. Table 3.5 illustrates that as θ increases total profit of the system increases as expected. Profit contribution from direct sales decreases whereas profit due to substituted demand increases. As θ increases capacity utilization declines. This pattern is consistent with Theorem 2, which states that as the substitution probability θ gets higher, a highly popular product can be set out of the assortment to direct customers to less popular but more profitable products with higher r_i. Also it may be sufficient to put fewer number of, but more profitable products into the assortment expecting customers to substitute. Thus, profit from direct sales decreases and profit from substituted demand increases.

Table 3.5: Sensitivity of optimal solution to θ for a single firm with $|N| = 10$ and $C = 7$

θ	Avg $\Delta\Pi$%	Avg DS %	Avg SS %	Avg Capacity Utilization %
0.3000	0.00	94.33	5.67	100
0.3452	0.85	93.54	6.46	100
0.3736	1.39	93.05	6.95	100
0.4105	2.21	91.03	8.97	100
0.4473	3.03	90.31	9.69	100
0.4842	3.88	89.26	10.74	100
0.5210	4.74	88.06	11.94	100
0.5578	5.63	87.10	12.90	100
0.5947	6.54	86.38	13.62	100
0.6315	7.44	85.66	14.34	100
0.6684	8.36	84.57	15.43	100
0.7052	9.30	83.68	16.32	100
0.7420	10.26	80.84	19.16	100
0.7789	11.31	80.08	19.92	100
0.8157	12.38	78.62	21.38	98.6
0.8526	13.47	77.88	22.12	98.6
0.8894	14.60	75.76	24.24	97.1
0.9262	15.80	73.63	26.37	97.1
0.9560	16.79	73.01	26.99	97.1
0.9631	17.02	72.87	27.13	97.1

Secondly, we investigate the effect of product purchase probabilities α_i as they get more asymmetric. The asymmetry is measured by the variance among α_i, for $i \in N$, denoted by Var(α_i). We first set $\alpha_1 = \alpha_2 = ... = \alpha_{10}$, which is the totally symmetric case, then increase the asymmetry among α_i values for each of every 10 problem instances for a total of 150 problems. Table 3.6 shows that as purchase probabilities get more asymmetric, total profit of the system and profit from direct sales increase whereas profit from substituted demand decreases. Capacity is always fully utilized among the observed scenarios. As α_i gets asymmetric, more popular products start to take place in the assortment so that loss of customer becomes infrequent. This leads to an increase in direct sales and decrease in substituted sales. In order to meet customer demand by direct sales of favorite products it is more likely to use capacity fully.

Table 3.6: Sensitivity of optimal solution to variance of α_i for a single firm with $|N| = 10$ and $C = 7$

Var(α_i)	Avg $\Delta\Pi$%	Avg DS%	Avg SS%	Avg Capacity Utilization%
0.0000	0.00	86.92	13.08	100
0.0009	3.99	87.43	12.57	100
0.0011	4.42	87.42	12.58	100
0.0013	4.77	87.56	12.44	100
0.0015	5.10	87.60	12.40	100
0.0017	5.43	87.63	12.37	100
0.0020	5.60	88.41	11.59	100
0.0023	6.40	87.70	12.30	100
0.0026	6.89	87.73	10.27	100
0.0029	7.34	87.82	12.18	100
0.0033	7.67	87.91	12.09	100
0.0036	8.16	89.55	10.05	100
0.0040	8.40	89.69	10.31	100
0.0042	8.58	89.99	10.01	100
0.0044	8.86	89.98	10.02	100

Lastly, we explore the effect of assortment capacity. We increase capacity gradually from $C = 1$ up to $C = 10$ and run 10 different problem instances at each capacity level so that Table 3.7 shows the results of 100 problem instances. It is observed that as capacity C increases, total profit of the system, profit from both direct and substituted demand sales increase until a certain capacity level. But there is a higher effect on direct sales, that is why percentage-wise, direct sales percentage increases, while substituted demand sales decreases. After a certain capacity level, capacity utilization decreases, as strategically some products are left out of the assortment consistent with Theorem 3. Moreover, although average capacity utilization may decrease, as the total capacity increases, there may be the same number of or more products in the assortment. That is why the average percentage of sales from direct sales tends to increase.

Table 3.7: Sensitivity of optimal solution to capacity C for a single firm with $|N| = 10$

C	Avg $\Delta\Pi$%	Avg DS%	Avg SS%	Avg Capacity Utilization%
1	0.00	71.26	28.74	100
2	69.13	74.33	25.67	100
3	119.19	77.28	22.72	100
4	157.07	80.34	19.66	100
5	181.28	83.41	16.59	100
6	198.88	86.64	13.36	100
7	209.57	89.29	10.71	100
8	216.02	90.62	9.38	97.5
9	219.28	91.33	8.67	91.1
10	220.22	91.49	8.51	85.0

3.5.2 Performance of Heuristic Algorithms

In this chapter, relying on Theorem 2, we introduce a heuristic algorithm to obtain the optimal assortment. We obtain an assortment for the firm by including best number of most dominant products, where dominance is obtained by the higher expected profitability of a product, which is the multiplication of the profit margin with the customer demand rate. The search for the best assortment of all dominant sets is done by adding one more product to the assortment from the dominant set at each iteration and checking the improvement in the total profit at every iteration till $|S| = C$. As not all optimal assortments may be composed of most dominant product according to Theorem 2, the gap between the profit from heuristic and optimal assortments, which is obtained by full enumeration, is evaluated.

Table 3.8 presents the results for 100 random problem instances evaluated at each of the specified number of products and capacity limits. All problem sets are generated by using Table 3.4. Table 3.8 reports the number of instances, in which capacity utilization is different between the optimal and heuristic assortments out of 100 problems. Moreover, the number of problem instances with none-dominant set of optimal assortment is given, which is the case where the heuristic and optimal assortments are not overlapping. Additionally, we illustrate average values of profit gaps between heuristic and optimal assortments over 100 problems. $\Delta\Pi\%$ for a problem instance denotes the percent decrease in total profit caused by the heuristic solution compared to optimal profit.

For a single firm with 6 products and a capacity limit of 3, out of 100 there are only 6 instances that have different number of products and 37 instances having non-overlapping products in heuristic assortment compared to optimal one. It can be seen that for the same capacity of 6 products, as the number of products N increases capacity utilization gap between heuristic and optimal solutions decreases. But for the same potential set N, as the capacity C increases, the capacity usage gap between heuristic and optimal solutions increases. This may be due to first Theorem 3, which states that the capacity utilization tend to decrease with the capacity limit. Thus, this decrease in optimal solution might not be

Table 3.8: Performance of the heuristic solution with respect to optimal solution

$\|N\|$	C	# of Instances with difference in capacity utilization	# of Instances with none-dominant set of optimal assortment	Avg $\Delta\Pi\%$	Avg $\Delta DS\%$	Avg $\Delta SS\%$
6	3	6	37	1.786	3.071	11.389
6	6	48	58	1.866	6.653	26.477
7	3	6	32	1.512	1.821	6.829
7	6	47	65	2.158	4.890	22.555
7	7	57	63	2.016	5.043	22.938
8	3	2	20	0.677	0.925	3.287
8	6	51	88	10.348	15.353	98.164
8	8	57	67	1.744	4.933	23.098
9	3	2	27	0.955	1.176	4.011
9	6	38	71	2.228	4.206	18.358
9	9	65	79	2.346	5.561	28.932
10	3	0	27	1.132	1.360	4.592
10	6	23	67	3.500	5.043	19.786
10	9	73	92	10.278	13.494	86.287
10	10	74	88	1.919	5.456	34.180

caught by the heuristic solution leading to a higher capacity usage gap. Second, as indicated by Theorem 2, as the capacity increases, it is more probable for the optimal assortment to be none-dominant set, which can be also observed in Table 3.8. The increasing gap between heuristic and optimal solutions in terms of the choice of right products can also lead the difference in the number of products chosen to increase. However, interestingly, for a given N, the percent profit decrease due to heuristic solution does not have a monotone behavior in capacity C. Similarly, for a given C, the profit gap is not monotone in N.

CHAPTER 4

CONSTRAINED ASSORTMENT PLANNING OF MULTIPLE FIRMS USING TRANSSHIPMENTS

4.1. Synopsis

In this chapter, we present an assortment planning problem of multiple firms each having a limited assortment capacity. Consumers in the market have certain demand rates for the products that are existing in the market. After assortments of all firms are determined, if a demanded product is not offered by the visited firm, it is allowed that the product can be transshipped from another firm's assortment, if available. So, firms can benefit from assortment-based cooperation through transshipments. If the demanded product does not exist in any firm's assortment, then the customer may prefer to substitute its demand with another product, which can be either directly satisfied by the visited firm or through transshipment from another assortment or lost if neither firm keeps it. Each transshipment incurs a cost, so decreases the net profit from a customer sales. The aim is to determine optimal assortments for all firms in the system so that the total system profit of multiple firms from demand satisfaction can be maximized subject to assortment capacity constraints of the firms.

It is shown that even under symmetric product profit margins, the optimal assortment may not be obtained according to decreasing product popularity order. Besides, it might be optimal not to utilize some assortment capacities due transshipment costs and demand substitution behavior. Then we study conditions that lead firms to utilize their assortment capacities fully. The problem and analytical model are defined in Chapter 4.2 in detail. The properties of optimal assortment are shown in Chapter 4.3. Finally, the results of numerical analyses on optimal assortments are discussed in Chapter 4.4.

4.2. Problem Definition and Model

In a multiple make-to-order firm system using transshipments, when a potential customer of the firm asks for a product, there are several possible scenarios that we need to consider. Firstly, the product can be directly supplied if it is available in the assortment of the visited firm. Otherwise, the product can be supplied by a transshipment if it is offered by any other firm in the system. When the product is not offered by any firm, then the customer may switch from his most favorite product to a second favorite one, which can either be satisfied directly by the visited firm, or by a transshipment from another firm, or lost due to unavailability in the system. We assume that transshipment option is utilized, if available, before encouraging a customer to substitute her demand, as it is assumed that substitution happens with a certain probability, which might be less than one. We consider a single substitution attempt and only assortment-based substitutions without any inventory considerations as in Chapter 3.

Similar to Chapter 3, customer choice is modeled by the exogenous demand model. Let λ denote the total expected number of customers in the market. Firm m has a market share of q_m among all M firms in the system such that $0 < q_m \leq 1$ and

$\sum_{m=1}^{M} q_m = 1$. The set of all products that can be offered by a firm is denoted by N. Each product i has a certain probability of being the first choice of a visiting customer denoted by α_i such that $0 < \alpha_i \leq 1$ and $\sum_{i \in S} \alpha_i = 1$. There is a predetermined assortment capacity for each firm which is denoted by C_m for firm m.

Let firm m's assortment be $S_m \subseteq S$. The overall set of all products that will be offered by the set of M firms is called as *union assortment* and denoted by S such that $S = (S_1 \cup S_2 \cup ... \cup S_M)$. The set of all products that will be offered by each and every firm is called as *common assortment* and denoted by $\bar{S}$ such that $\bar{S} = (S_1 \cap S_2 \cap ... \cap S_M)$.

For the analysis of multi-firm assortment planning problem, we consider products with identical profit margins, $r_i = r$. Even with identical profit margins, the analysis of the optimal assortments is quite challenging. Thus, we believe that as being the first study on the assortment planning of multiple firms using transshipments with assortment capacities, the managerial insights to be obtained with this problem is very important without allowing further difficulty in analytically tractability.

An arriving customer of firm m pays r to the firm to buy the product, if his most favorite product i is in the assortment of firm m, i.e., $i \in S_m$. If product i is not offered by firm m, but offered by firm l, then by incurring a transshipment cost τ, the demand can be satisfied through firm l. As it is a centrally managed system, the transshipment cost can be paid by either party, sender or the receiver. If the demanded product i is not offered by any firm, then the customer may decide to substitute her first choice with a second one with probability θ. The probability that the customer substitutes product i with product j is $\delta_{ij} = \alpha_j/(1 - \alpha_i)$, where δ_{xx}= 0. If firm m has the second favorite product j, then the customer gets it directly from firm m. Otherwise, if it is available by any other firm l, then the demand is satisfied via transshipment. Finally, if the product j is not available in the overall assortment S, the customer leaves the firm without a purchase. The notation used for multiple firms setting is summarized in Table 4.1.

Table 4.1: Notation for multi-firm assortment planning problem

Parameters	
N	The set of all products
M	Number of firms in the system
λ	Expected number of customers in the market
q_m	Fraction of customer visiting firm m, $m \in \{1..M\}$
C_m	Assortment capacity of firm m, $m \in \{1..M\}$
α_i	Purchase rate(probability) for product i, $i \in N$
r	Profit margin for one unit of product
τ	Transshipment cost per product
θ	Rate(probability) of substitution
δ_{ij}	Substitution probability from product i to product j, $i, j \in N$
Variables	
S_m	Assortment of firm m, $m \in \{1..M\}$

Let $\Pi(S_1, S_2, ..., S_M)$ be the total expected profit of M firms.

$$\begin{aligned}
\max \Pi(S_1, S_2, ..., S_M) \;\; &= \;\; \lambda \sum_{m=1}^{M} q_m \sum_{i \in S_m} \big(\alpha_i r + \theta r \sum_{j \notin S} \alpha_j \delta_{ji}\big) \\
&\quad + \lambda \sum_{m=1}^{M} q_m \sum_{i \in S \setminus S_m} (r - \tau)\big(\alpha_i + \theta \sum_{j \notin S} \alpha_j \delta_{ji}\big) \qquad (4.1) \\
s.t., |S_m| \leq C_m, &\quad m \in \{1..M\}.
\end{aligned}$$

The objective function is composed of, respectively, the revenue from the direct satisfaction of customers' first choice demands, substitution of the first choice demand with the visited firm's assortment, satisfaction of the first choice demand by transshipment, and the substitution of the first choice demand from another firm's assortment via transshipment. Constraints indicate the predetermined assortment capacities of firms.

4.3. Properties of Optimal Assortments

Recall that in Chapter 3 we define *popular (assortment) set* as a group of products in a decreasing order of purchase probabilities where the most popular product i, having the largest α_i is indexed by i=1 and accordingly remaining ones are sorted in decreasing order of α_i (Kök & Xu, 2011). In this chapter, we show that when all products have identical profit margins, the common assortment $\bar{S}$ of all firms is in the popular set at optimality. Thus, the common assortment will include a number of most popular products.

Theorem 4. *The common assortment $\bar{S}$ is in the popular set.*

For an individual single firm, when all products have the same profit margin r, the optimal assortment is always in the popular set and there is no incentive to underutilize the assortment capacity according to Theorem 1. In a multi-firm setting with transshipments allowed among firms, while the common assortment is in the popular set, the complete individual assortment of a firm, thus the union assortment might not be. Moreover, the assortment capacities might not be fully utilized. This can be illustrated by an example.

Lets consider a three-firm system where market shares of firms are such that $q_1 = 0.7$, $q_2 = 0.2$, $q_3 = 0.1$. Product purchase probabilities are $\alpha_i = \{0.33, 0.28, 0.22, 0.17\}$, product margin is $r = 8$, substitution probability is $\theta = 0.8$, transshipment cost is $\tau = 5$, and $\lambda = 100$. The assortment capacities are $(2, 3, 4)$, respectively, for firms 1 to 3. In this particular setting, optimal firm assortments and the distribution of total profit is given Table 4.2.

Table 4.2: Optimal solution for a sample problem with $M = 3$ with $|N = 4|$

S_1	1,2	Own sales	528.8
S_2	1,2,4	Transshipments	35.7
S_3	1,2,4	Substitution from own stock	119.3
$\bar{S}$	1,2	Substitution by transshipment	8.1
S	1,2,4	Total Profit	691.9

Table 4.2 shows that while the optimal assortment of firm 1 is in the popular set, assortments of firms 2 and 3 are not. Moreover, assortment capacity of firm 3, which is four, is not fully used.

Optimal assortments of firm 2 and 3 are not in the popular set, because keeping more popular products in their assortments, there will be more transshipments for these popular products to those firms who are not able to keep these products. As the capacity of firm 1 is only two, she is assigned products 1 and 2. As she cannot keep product 3 because of capacity limit, any demand to this product will be satisfied by transshipments when any other firm in the system keeps it. Thus, it is system-optimal to keep a less popular product 4 instead of 3 so that there will less transshipments for product 4, decreasing transshipment costs. As there is no opportunity for providing product 3, there will be more substitution that will be satisfied within own stores or by transshipments. If firms 2 and 3 are forced to have assortments in the popular set such as $S_2 = S_3 = \{1, 2, 3\}$, by keeping S_1 the same, the profit distribution will change such that profit from own sales and transshipments will increase to 540.8 and 46.2, respectively, while the profit from substitutions from own stock and by transshipment will decrease to 88.6 and 7.8, respectively.

The reason behind underutilization of a firm's assortment capacity under symmetric profit margins is also due to use of transshipments, which have non-negligible costs. When a product is included in the assortment of a firm, but not in the common assortment, then at every firm, which does not hold this product, the demand for the product is satisfied by transshipment. If this particular product is not held by any firm in the system, then instead of a transshipment, a substitution can be initiated to satisfy the demand of the product, where it is possible that the substituted product may be in the assortment of the visited firm. Thus, substitution can lead to a full-price revenue by the direct demand satisfaction from the inventory of the visited firm, where the transshipment leads to a full-price minus transshipment cost earning. If transshipments are not forced by the system, but selectively used according to a need, which is out of scope of current study, then the result on the assortment planning could be different.

Given that transshipments and substitutions are alternatives in demand satisfaction of the products, which are not hold in the assortment of a visited firm, when substitutions are more probable, transshipments can be better risked to avoid transshipment cost. Thus, assortment capacity can be underutilized not to pay transshipment cost when a substitution is more probable. Theorem 5 introduces a threshold level on substitution level for each firm to fully utilize its assortment capacity.

Theorem 5. *(i) The assortment capacity of firm m is fully utilized at optimality if θ does not exceed the threshold level $\bar{\theta}_c^m$ such that*

$$\bar{\theta}_c^m = 1 - \frac{\tau}{r}(1 - q_m).$$

(ii) Capacity utilization is non-decreasing in r and q_m.

Theorem 5 also confirms Theorem 1 as $q_m = 1$ for a single firm system results in $\theta_c^m = 1$, where $\theta \leq 1$ by definition. Note that $\theta \leq \bar{\theta}_c^m$ is a sufficient condition, but not necessary. So, when $\theta > \bar{\theta}_c^m$, the capacity of firm m can be still fully utilized.

Adding a totally new product to the system may not always be profitable depending on the substitution probability. As long as it is lower than the specified threshold level, $\bar{\theta}_c^m$, it makes sense to add the product to the assortment of firm m. Marginal change in total profit due to this addition is directly proportional with the profit margin of the product. Whereas, if the product that is to be added is only new for the specific firm m, but already existing in the system, there is nothing to worry, it can be added once again to the system. Expectedly, marginal change in total profit due to this addition is directly proportional with the fraction of customers visiting firm m. This property implies that there is no point in underutilization of the capacity as the addition of the product cannot decrease total profit. Based upon this property, Theorem 6 associates cardinality of an assortment to its assortment capacity.

Theorem 6. *(i) If the assortment capacities of two firms m and l are equal, $C_m = C_l$, at optimality, they hold the same number of products, i.e., $|S_m^*| = |S_l^*|$.*
(ii) When all firms have the same assortment capacity C such that $C_1 = ... = C_M = C$, then at optimality, all firm capacities are fully utilized, i.e., $|S_1^*| = ... = |S_m^*| = ...|S_M^*| = C$.

Theorem 6 states that if the assortment capacities of all firms are equal to C, then at optimality, they all hold the same number of products, which is equal to their common assortment limit C. Thus, their capacities are fully utilized as well. However, this property is partially valid for any two firms m and l having the same assortment limit, $C_m = C_l$. Although they are finally assigned equal number of products, the number of products might be less than their assortment limit, $C_m = C_l$. So, their capacities are not necessarily fully utilized. Recall the three-firm example discussed earlier in this chapter with the optimal solution given in Table 4.2. All remaining the same, the assortment capacities are now set as $(2, 4, 4)$, respectively, for firms 1 to 3. At optimality, the assortment sets are $S_1 = \{1, 2\}$ and $S_2 = S_3 = \{1, 2, 4\}$ with a total profit of $\pi = 691.9$. Although, capacities of firms 2 and 3, $C_2 = C_3 = 4$, allow to hold the remaining out-of-assortment product 4, the cost of potential transshipments of product 4 to firm 1 restricted this product to be included in the assortment. Thus, while $|S_2^*| = |S_3^*| = 3$, their capacities are still underutilized.

By Theorem 1, we know that the single firm utilizes its capacity fully when all products have equal profit margin. Thus, above result indicates that in case of allowed transshipments, it might be better not to utilize assortment capacities of some of the firms, thus narrowing the assortments compared to an independent system without transshipments. The reason is that transshipments are costly and instead of these costly transshipments, the unsatisfied demand can be expected to be substituted. So, by Theorem 5, if the substitution probability would be zero or any number less than or equal to $1 - \tau/r$, even under transshipments, each firm's capacity would be fully utilized. This result is almost the opposite of that in Besbes and Sauré (2016), where the competition between two companies leads one to offer a broader set of products compared to when it is operating as a monopolist. In our model, collaboration through transshipments can actually narrow the offerings.

Compared to Dağ (2015), constraints on assortment sizes invalidates a nice property on optimal assortments. With a similar model, the study shows that optimal assortments of firms are nested such that the assortment of a firm is a subset of another one with equal or higher market size. In our constrained assortment planning model, even if the assortment capacities are equal for any two or all firms, the assortment of the firm with smaller market size is not necessarily subset of higher market size one. In Dağ (2015), existence of nested property allows the analysis of individual firm assortments in more details, which is not the case in our constrained optimization model. Therefore, the structure of optimal assortments are further investigated through numerical analysis in Chapter 4.4.

4.4. Computational Results

In this chapter, we present the results of extensive numerical analyses to support and extend the analytical properties we obtained on optimal assortments of multiple firms. We first aim to observe the sensitivity of the optimal assortments to problem parameters.

Similar to sensitivity analysis of a single firm, throughout the numerical analysis, we use a broad range for the model inputs to increase the robustness of the results. Unless stated otherwise, for each problem instance each parameter is generated according to Uniform distribution with the ranges summarized in Table 4.3. $\lambda = 100$ is set for all problem instances.

Table 4.3: Uniform distributions of randomly generated parameters

θ	α_i	r_i	q_m	τ
U(0.3,1)	U(0,1)	U(1,10)	U(0,1)	U(0,1)

4.4.1 Sensitivity of Optimal Assortments

In this chapter, we analyze how the system inputs; probability of substitution θ, symmetry in purchase probability of products α_i and transshipment cost τ affect the optimal assortments. To test the effect of each parameter, we create 10 problem instances for multiple firms where $M = 3$ and the number of products $|N| = 5$. For each problem set, we use randomly generated parameters with Uniform distributions by using the ranges mentioned in Table 4.3, except the concerned parameter. We set the analyzed parameter to a certain value for all 10 problem instances, obtain the optimal result for 10 instances, and calculate the average values for the measured performance metrics over 10 instances. We repeat the same 10 problem analysis several times, by setting the analyzed parameter to a different value each time.

All problem instances are solved by full enumeration to obtain the optimal assortments. Then performance metrics are obtained with the optimal assortments. Particularly, we use following metrics to measure the effects of system inputs.

- Average percent change in total profit ($\Delta\Pi\%$)
- Average percentage of direct sales (DS%)
- Average percentage of sales through transshipment, both for the first and second-choice demands ($TS\%$)
- Average percentage of sales for the substituted demand from stock (SS%)
- Average number of products in total assortment S ($|S|$)
- Average number of different products in total assortment S
- Average number of products in the common assortment $\bar{S}$ ($|\bar{S}|$)
- Average capacity utilization for each firm (%).

In a multiple firm system, total revenue is composed of the revenue from the direct satisfaction of customers' first choice demands, satisfaction of the first choice demand by transshipment, substitution of the first choice demand with the visited firm's assortment, and the substitution of the first choice demand from another firm's assortment via transshipment. According to the objective function,

Transshipment Sales(TS): $\lambda \sum_{m=1}^{M} q_m \sum_{i \in S \setminus S_m} (r-\tau)\big(\alpha_i + \theta \sum_{j \notin S} \alpha_j \delta_{ji}\big)$

Substitution Sales(SS): $\lambda \sum_{m=1}^{M} q_m \sum_{i \in S_m} \theta r \sum_{j \notin S} \alpha_j \delta_{ji}$

Direct Sales(DS): $\lambda \sum_{m=1}^{M} q_m \sum_{i \in S_m} \alpha_i r$

Therefore, for a multi-firm system DS% + TS% + SS%=100%. The capacity utilization ratio for a firm m is the ratio of actual number of products to be kept in the optimal assortment for this firm to the assortment capacity C_m. Then it is reported in percentage by multiplying with hundred.

We present average percent change in total profit ($\Delta\Pi$%) by using the total profit values of the first 10 instances as a basis. For each fixed value of the concerned parameter over 10 problem instances, the average values of the metrics defined above are obtained and reported. With the help of this methodology, we aim to eliminate the bias selection effects of other parameters by reporting average performances over 10 problem instances generated randomly. The sensitivity analysis is conducted with symmetric profit margins r and we keep the capacity levels for three firms fixed at $C_1 = 3, C_2 = 5, C_3 = 2$.

Firstly, we evaluate the sensitivity of optimal assortment to probability of substitution θ for a multiple firm system. 20 different values of θ are tested so that Table 4.4 reports the results of 200 problem instances in total. Table 4.4 shows that as θ increases total profitability of the system increases expectedly, profit contribution from direct sales and transshipment decrease. On the other hand, profit contribution from sales for the substituted demand increases.

Moreover, reported by Table 4.5, the total number and the number of different products in S decrease while the number of products in the common assortment $\bar{S}$ stays the same as θ increases. The reason is that as customer substitutions are more probable, the assortment can be narrowed down. Accordingly, the capacity utilization of the second firm, whose capacity limit is the highest with 5, decreases while the others always use full capacity. The analysis supports the claim that transshipments can be better risked to avoid transshipment cost as substitution probability increases. It is also better to underutilize capacity of S_2 in order not to pay transshipment cost when θ is high enough.

Table 4.4: Sensitivity of total profit to θ for a $M = 3$ firm system with $|N| = 5$ and $C_1 = 3, C_2 = 5, C_3 = 2$

θ	Avg $\Delta\Pi\%$	Avg $DS\%$	Avg $TS\%$	Avg $SS\%$
0.3000	0.000	93.614	6.165	0.221
0.3452	0.075	92.766	6.165	0.445
0.3736	0.207	92.020	6.165	1.815
0.4105	0.409	91.078	6.165	2.757
0.4473	0.660	90.699	6.165	3.135
0.4842	0.923	90.464	6.165	3.370
0.5210	1.220	89.639	6.165	4.195
0.5578	1.526	89.381	6.165	4.454
0.5947	1.070	88.555	5.805	5.640
0.6315	2.239	88.252	5.805	5.942
0.6684	2.604	87.954	5.805	6.241
0.7052	2.969	87.661	5.805	6.534
0.7420	3.380	87.007	4.987	8.007
0.7789	3.830	85.936	4.878	9.186
0.8157	4.309	85.055	4.358	10.587
0.8526	4.888	83.277	4.358	12.365
0.8894	5.562	81.355	2.607	16.038
0.9262	6.286	80.816	2.607	16.577
0.9560	6.875	80.134	2.574	17.292
0.9631	7.017	80.031	2.574	17.396

Secondly, we consider the sensitivity of optimal assortment to purchase probabilities α_i. We measure the effect of α_i, as they get more asymmetric. To reflect the asymmetry, we use the variance among α_i, for $i \in N$, denoted by $\text{Var}(\alpha_i)$. We first set $\alpha_1 = \alpha_2 = ... = \alpha_{10}$, which is the symmetric case and then increase the asymmetry among α_i values for a total of 150 problems.

Table 4.6 shows that as purchase probability of products α_i gets more asymmetric, total revenue of the system increases. The contribution profit from direct sales increases, whereas profit from sales through transshipment decreases. The profit contribution from sales for the substituted demand is almost constant. Thus, there is clear switch from transshipments to direct sales. The reason is that as product preferences α_i get more asymmetric, more popular products are placed in the

Table 4.5: Sensitivity of optimal solution to θ for a $M = 3$ firm system with $|N| = 5$ and $C_1 = 3, C_2 = 5, C_3 = 2$

θ	Avg $\|S\|$	Avg # of Different Prod. in S	Avg $\|\bar{S}\|$	Capacity utilization of S_1	Capacity utilization of S_2	Capacity utilization of S_3
0.3000	9.9	4.9	2	100%	98%	100%
0.3452	9.7	4.7	2	100%	94%	100%
0.3736	9.6	4.6	2	100%	92%	100%
0.4105	9.5	4.5	2	100%	90%	100%
0.4473	9.3	4.3	2	100%	86%	100%
0.4842	9.3	4.3	2	100%	86%	100%
0.5210	9.2	4.2	2	100%	84%	100%
0.5578	9.2	4.2	2	100%	84%	100%
0.5947	9.1	4.1	2	100%	82%	100%
0.6315	9.1	4.1	2	100%	82%	100%
0.6684	9.1	4.1	2	100%	82%	100%
0.7052	9.1	4.1	2	100%	82%	100%
0.7420	9.0	4.0	2	100%	80%	100%
0.7789	8.9	3.9	2	100%	78%	100%
0.8157	8.8	3.8	2	100%	76%	100%
0.8526	8.8	3.8	2	100%	76%	100%
0.8894	8.5	3.5	2	100%	70%	100%
0.9262	8.5	3.5	2	100%	70%	100%
0.9560	8.5	3.5	2	100%	70%	100%
0.9631	8.5	3.5	2	100%	70%	100%

assortment so that loss of customer becomes less frequent. This leads to an increase in direct sales. While the following effect in a single firm system was the decrease in substituted sales, in a transshipment system, the placement of more popular products decreases the transshipment sales.

Lastly, we measure the effect of transshipment cost τ for multiple firm system. Table 4.8 shows that as τ increases, profit from direct sales increases while the contribution of transshipments decreases as would be expected. The transshipments are substituted by the demand substitutions. Table 4.9 illustrates the change in assortment combination as τ increases. It seems that both the size of

Table 4.6: Sensitivity of total profit to variance of α_i for a $M = 3$ firm system with $|N| = 5$ and $C_1 = 3, C_2 = 5, C_3 = 2$

Var(α_i)	Avg $\Delta\Pi$%	Avg DS%	Avg TS%	Avg SS%
0.00002	0.0000	81.56	27.95	8.08
0.00025	0.3069	82.63	23.14	10.91
0.00030	0.1780	82.77	22.91	10.86
0.00036	0.0468	82.91	22.67	10.81
0.00042	0.0839	83.06	22.44	10.76
0.00049	0.2141	83.20	22.21	10.71
0.00056	0.3440	83.34	21.98	10.66
0.00064	0.4733	83.49	21.75	10.61
0.00072	0.6022	83.63	21.52	10.56
0.00081	0.7307	83.77	21.30	10.51
0.00090	0.8588	83.91	21.07	10.46
0.00100	0.9864	84.06	20.84	10.41
0.00110	1.1136	84.20	20.62	10.36
0.00116	1.2770	84.25	20.51	10.36
0.00121	1.2403	84.34	20.40	10.31

total assortment and the variety in total assortment are decreased with τ. By narrowing the total assortment, some of the possible transshipments are prevented to decrease transshipment costs. So the second firm's capacity is under utilized when τ is very high. But the number of products in the common set which is composed of most popular products stays the same, which is not related to transshipment cost as these products are transshipped.

Table 4.7: Sensitivity of total profit to variance of α_i for a $M = 3$ firm system with $|N| = 5$ and $C_1 = 3, C_2 = 5, C_3 = 2$

Var(α_i)	Avg $\lvert S\rvert$	Avg # of Different Prod. in S	Avg $\lvert\bar{S}\rvert$	Capacity utilization of S_1	Capacity utilization of S_2	Capacity utilization of S_3
0.00002	9.6	4.5	2	100 %	92 %	100 %
0.00025	9.5	4.5	2	100 %	90 %	100 %
0.00030	9.5	4.5	2	100 %	90 %	100 %
0.00036	9.5	4.5	2	100 %	90 %	100 %
0.00042	9.5	4.5	2	100 %	90 %	100 %
0.00049	9.5	4.5	2	100 %	90 %	100 %
0.00056	9.5	4.5	2	100 %	90 %	100 %
0.00064	9.5	4.5	2	100 %	90 %	100 %
0.00072	9.5	4.5	2	100 %	90 %	100 %
0.00081	9.5	4.5	2	100 %	90 %	100 %
0.00090	9.5	4.5	2	100 %	90 %	100 %
0.00100	9.5	4.5	2	100 %	90 %	100 %
0.00110	9.5	4.5	2	100 %	90 %	100 %
0.00116	9.5	4.5	2	100 %	90 %	100 %
0.00121	9.5	4.5	2	100 %	90 %	100 %

This is no change in Avg $|S|$, Avg # of Different Prod. in S, Avg $|\bar{S}|$ and capacity utilizations of S_1, S_2 and S_3.

Table 4.8: Sensitivity of total profit to τ for a $M = 3$ firm system with $|N| = 5$ and $C_1 = 3, C_2 = 5, C_3 = 2$

τ	Avg $\Delta\Pi\%$	Avg $DS\%$	Avg $TS\%$	Avg $SS\%$
0.0303	0.0000	84.19	15.80	0.00
0.0411	-0.0327	84.22	15.77	0.00
0.0909	-0.1837	84.35	15.64	0.00
0.0911	-0.1843	84.35	15.64	0.00
0.1297	-0.3016	84.45	15.54	0.00
0.1605	-0.3954	84.53	15.46	0.00
0.2981	-0.8168	84.88	15.11	0.00
0.3284	-0.9100	84.96	15.03	0.00
0.3375	-0.9381	84.98	15.01	0.00
0.3648	-1.0223	85.06	14.93	0.00
0.5169	-1.4941	85.45	14.54	0.00
0.5658	-1.6444	85.10	13.75	1.05
0.6039	-1.7520	85.19	13.66	1.05
0.6945	-2.0088	85.40	13.44	1.05
0.7184	-2.0767	85.46	13.39	1.05
0.8074	-2.3305	85.67	13.18	1.06
0.8367	-2.4143	85.74	13.11	1.06
0.9313	-2.6859	85.97	12.88	1.06
0.9596	-2.7675	86.04	12.82	1.06
0.9888	-2.8517	86.11	12.75	5.31

Table 4.9: Sensitivity of optimal solution to τ for a $M = 3$ firm system with $|N| = 5$ and $C_1 = 3, C_2 = 5, C_3 = 2$

τ	Avg $\|S\|$	Avg # of Different Prod. in S	Avg $\|\bar{S}\|$	Capacity utilization of S_1	Capacity utilization of S_2	Capacity utilization of S_3
0.0303	10.0	5.0	2	100%	100%	100%
0.0411	10.0	5.0	2	100%	100%	100%
0.0909	10.0	5.0	2	100%	100%	100%
0.0911	10.0	5.0	2	100%	100%	100%
0.1297	10.0	5.0	2	100%	100%	100%
0.1605	10.0	5.0	2	100%	100%	100%
0.2981	10.0	5.0	2	100%	100%	100%
0.3284	10.0	5.0	2	100%	100%	100%
0.3375	10.0	5.0	2	100%	100%	100%
0.3648	10.0	5.0	2	100%	100%	100%
0.5169	10.0	5.0	2	100%	100%	100%
0.5658	9.8	4.8	2	100%	96%	100%
0.6039	9.8	4.8	2	100%	96%	100%
0.6945	9.8	4.8	2	100%	96%	100%
0.7184	9.8	4.8	2	100%	96%	100%
0.8074	9.8	4.8	2	100%	96%	100%
0.8367	9.8	4.8	2	100%	96%	100%
0.9313	9.8	4.8	2	100%	96%	100%
0.9596	9.8	4.8	2	100%	96%	100%
0.9888	9.8	4.8	2	100%	96%	100%

CHAPTER 5

CONCLUSIONS

In this thesis, assortment planning is studied for two different systems under capacity limits. First, a single firm's optimal assortment problem is studied to maximize his profit. Second, assortment planning problem of a system of multiple firms is considered in which product sharing is possible via practice of transshipment. In both problem cases, customer demand is defined with exogenous demand model, where each customer has a predetermined preference for each product from the potential set. Proportional demand substitutions are also taken into account to explain the customer behavior when the demanded product cannot be satisfied. The results on the optimal assortment of a single firm are used to elaborate on the optimal assortments of multiple firms communicating through transshipments.

Our analytical study presents that the optimal assortment of a single firm is in the popular set and fully utilizes the assortment capacity when all products have symmetric profit margins. Therefore, the optimal assortment can be easily obtained by including the products with the highest demand probability to fill up the capacity. When products have asymmetric profit margins, it is necessary to examine expected profitability of each product, the profit rate. If all products can be nicely sorted with respect to both their profit margins and demand probabilities, the optimal assortment is composed of some number of most dominant (profitable) products. Otherwise, some number of highly dominant products are left out of assortment when substitution rate is high, which increases the probability of keeping high margin substitutes. Knowing that the optimal assortment is composed of some number of most dominant products, it is possible to decrease the number of possible assortments by 99% for a single firm problem with 10 products and a capacity limit of 5. Additionally, capacity is not always fully utilized when product profit margins are asymmetric due to bait and switch strategy. We introduce an upper limit on the substitution rate, below which the capacity is always fully utilized at optimality. This can lead to a reduction in the number of possible assortments by 60% for a single firm problem with 10 products.

For multiple firms, we consider products with identical profit margins owing to the complexity of the problem. We show that the common assortment is in the popular set. However, different than a single firm system, assortment capacities might not be fully utilized because of transshipment cost. We obtain an upper limit on the substitution rate for each firm such that below this threshold, the capacity of the firm is fully utilized at optimality. Moreover, we show that if the assortment capacities of two firms are equal then the cardinality of their optimal assortments are equal, i.e., they keep the same number of products, even if they are not at their capacity limit. When all firms have the same capacity limit, they all fully utilize their capacities at optimal solution. In summary, we show that collaboration through transshipment may result in a narrower product set.

We conduct numerical analyses so as to support analytical properties we obtained for the optimal assortment for both single and multiple firm problems. We aim to observe the sensitivity of the optimal assortment with regard to some problem parameters and analyze how the system inputs affect the optimal assortment. Finally, relying on the optimality properties obtained analytically, we introduce heuristic assortment policy for single firm test its performances numerically.

To our knowledge, this study contributes to the single firm assortment planning literature by allowing both product specific demand rates and profit margins. Moreover, we contribute to the very limited literature on assortment planning of multiple firms using transshipments by introducing the assortment capacities. The current study can be extended in different aspects such as allowing asymmetric profit margins for multiple firms. Also, by relying the optimality properties obtained in this study, various heuristic assortment algorithms can be defined and compared.

BIBLIOGRAPHY

Akçay, Y., & Tan, B. (2008). On the benefits of assortment-based cooperation among independent producers. *Production and Operations Management*, *17*(6), 626–640.

Alptekinoğlu, A., & Grasas, A. (2014). When to carry eccentric products? optimal retail assortment under consumer returns. *Production and Operations Management*, *23*(5), 877–892.

Archibald, T. (2007). Modelling replenishment and transshipment decisions in periodic review multilocation inventory systems. *Journal of the Operational Research Society*, *58*(7), 948–956.

Archibald, T., Sassen, S., & Thomas, L. (1997). An optimal policy for a two depot inventory problem with stock transfer. *Management Science*, *43*(2), 173–183.

Archibald, T. W., Black, D., & Glazebrook, K. D. (2009). An index heuristic for transshipment decisions in multi-location inventory systems based on a pairwise decomposition. *European Journal of Operational Research*, *192*(1), 69–78.

Bernstein, F., Kök, A. G., & Xie, L. (2015). Dynamic assortment customization with limited inventories. *Manufacturing & Service Operations Management*, *17*(4), 538–553.

Besbes, O., & Sauré, D. (2016). Product assortment and price competition under multinomial logit demand. *Production and Operations Management*, *25*(1), 114–127.

Briesch, R. A., Chintagunta, P. K., & Fox, E. J. (2009). How does assortment affect grocery store choice? *Journal of Marketing Research*, *46*(2), 176–189.

Cachon, G. P., & Kök, A. G. (2007). Category management and coordination in retail assortment planning in the presence of basket shopping consumers. *Management Science*, *53*(6), 934–951.

Cachon, G. P., Terwiesch, C., & Xu, Y. (2005). Retail assortment planning in the presence of consumer search. *Manufacturing & Service Operations Management*, *7*(4), 330–346.

Cheng, L., & Duran, M. A. (2004). Logistics for world-wide crude oil transportation using discrete event simulation and optimal control. *Computers & chemical engineering*, *28*(6), 897–911.

Çömez, N., Stecke, K. E., & Çakanyildirim, M. (2012). In-season transshipments among competitive retailers. *Manufacturing & Service Operations Management*, *14*(2), 290–300.

Çömez, N., Stecke, K. E., & Çakanyıldırım, M. (2012). Multiple in-cycle transshipments with positive delivery times. *Production and Operations Management*, *21*(2), 378–395.

Çömez-Dolgan, N., & Fescioğlu-Unver, N. (2015). Managing transshipments in a multi-retailer system with approximate policies. *Journal of the Operational Research Society*, *66*(6), 947–964.

Corsten, D., & Gruen, T. W. (2004). Stock-outs cause walkouts. *Harvard Business Review*, *82*(5), 26–28.

Dağ, H. (2015). *Assortment planning in transshipment systems* (Unpublished doctoral dissertation). Bilkent University.

Davis, J. M., Gallego, G., & Topaloglu, H. (2014). Assortment optimization under variants of the nested logit model. *Operations Research*, *62*(2), 250–273.

Dukes, A. J., Geylani, T., & Srinivasan, K. (2009). Strategic assortment reduction by a dominant retailer. *Marketing Science*, *28*(2), 309–319.

Fadıloglu, M. M., Karaşan, O. E., & Pınar, M. Ç. (2010). A model and case study for efficient shelf usage and assortment analysis. *Annals of Operations Research*, *180*(1), 105–124.

Gallego, G., Ratliff, R., & Shebalov, S. (2014). A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research*, *63*(1), 212–232.

Grahovac, J., & Chakravarty, A. (2001). Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items. *Management Science*, *47*(4), 579–594.

Hu, X., Duenyas, I., & Kapuscinski, R. (2007). Existence of coordinating transshipment prices in a two-location inventory model. *Management Science*, *53*(8), 1289–1302.

Kök, A. G. (2003). Management of product variety in retail operations.. (Dissertations available from ProQuest)

Kök, A. G., & Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, *55*(6), 1001–1021.

Kök, A. G., Fisher, M. L., & Vaidyanathan, R. (2008). Assortment planning: Review of literature and industry practice. In *Retail supply chain management* (pp. 99–153). Springer.

Kök, A. G., Fisher, M. L., & Vaidyanathan, R. (2015). Assortment planning: Review of literature and industry practice. In *Retail supply chain management* (pp. 175–236). Springer.

Kök, A. G., & Xu, Y. (2011). Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models. *Management Science*, *57*(9), 1546–1563.

Kranenburg, A., & Van Houtum, G. (2009). A new partial pooling structure for spare parts networks. *European Journal of Operational Research*, *199*(3), 908–921.

Kukreja, A., Schmidt, C. P., & Miller, D. M. (2001). Stocking decisions for low-usage items in a multilocation inventory system. *Management Science*, *47*(10), 1371–1383.

Lancaster, K. (1975). Socially optimal product differentiation. *The american economic review*, *65*(4), 567–585.

Li, Z. (2007). A single-period assortment optimization model. *Production and Operations Management*, *16*(3), 369.

Mahajan, S., & Van Ryzin, G. (2001). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, *49*(3), 334–351.

Mantrala, M. K., Levy, M., Kahn, B. E., Fox, E. J., Gaidarev, P., Dankworth, B., & Shah, D. (2009). Why is assortment planning so difficult for retailers? a framework and research agenda. *Journal of Retailing*, *85*(1), 71–83.

Misra, K. (2008). Understanding retail assortments in competitive markets. *Northwestern University*.

Moinzadeh, K., & Nahmias, S. (1988). A continuous review model for an inventory system with two supply modes. *Management Science*, *34*(6), 761–773.

Narus, J. A., & Anderson, J. C. (1996). Rethinking distribution: adaptive channels. *Harvard Business Review*, *74*(4), 112.

Netessine, S., & Rudi, N. (2003). Centralized and competitive inventory models with demand substitution. *Operations Research*, *51*(2), 329–335.

Özdemir, D., Yücesan, E., & Herer, Y. T. (2006). Multi-location transshipment problem with capacitated transportation. *European Journal of Operational Research*, *175*(1), 602–621.

Pentico, D. W. (2008). The assortment problem: A survey. *European Journal of Operational Research*, *190*(2), 295–309.

Robinson, L. W. (1990). Optimal and approximate policies in multiperiod, multi-location inventory models with transshipments. *Operations Research*, *38*(2), 278–295.

Rudi, N., Kapur, S., & Pyke, D. F. (2001). A two-location inventory model with transshipment and local decision making. *Management science*, *47*(12), 1668–1680.

Rusmevichientong, P., Shen, Z.-J. M., & Shmoys, D. B. (2009). A ptas for capacitated sum-of-ratios optimization. *Operations Research Letters*, *37*(4), 230–238.

Rusmevichientong, P., Shen, Z.-J. M., & Shmoys, D. B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations research*, *58*(6), 1666–1680.

Ryzin, G. v., & Mahajan, S. (1999). On the relationship between inventory costs and variety benefits in retail assortments. *Management Science*, *45*(11), 1496–1509.

Sadowski, W. (1959). A few remarks on the assortment problem. *Management Science*, *6*(1), 13–24.

Schön, C. (2010). On the optimal product line selection problem with price discrimination. *Management Science*, *56*(5), 896–902.

Smith, S. A., & Agrawal, N. (2000). Management of multi-item retail inventory systems with demand substitution. *Operations Research*, *48*(1), 50–64.

Sošic, G. (2006). Transshipment of inventories among retailers: Myopic vs. farsighted stability. *Management science*, *52*(10), 1493–1508.

Tan, B., & Akçay, Y. (2014). Assortment-based cooperation between two make-to-stock firms. *IIE Transactions*, *46*(3), 213–229.

Thomas, D. J., & Tyworth, J. E. (2006). Pooling lead-time risk by order splitting: a critical review. *Transportation Research Part E: Logistics and Transportation Review*, *42*(4), 245–257.

Tomlin, B. (2006). On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science*, *52*(5), 639–657.

Wang, R. (2012). Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters*, *40*(6), 492–497.

Wang, R. (2013). Assortment management under the generalized attraction model with a capacity constraint. *Journal of Revenue and Pricing Management*, *12*(3), 254–270.

Working, H., & Hotelling, H. (1929). Applications of the theory of error to the interpretation of trends. *Journal of the American Statistical Association*, *24*(165A), 73–85.

Yang, H., & Schrage, L. (2009). Conditions that cause risk pooling to increase inventory. *European Journal of Operational Research*, *192*(3), 837–851.

Yücel, E., Karaesmen, F., Salman, F. S., & Türkay, M. (2009). Optimizing product assortment under customer-driven demand substitution. *European Journal of Operational Research*, *199*(3), 759–768.

Zhao, H., Deshpande, V., & Ryan, J. K. (2005). Inventory sharing and rationing in decentralized dealer networks. *Management Science*, *51*(4), 531–547.

Zhao, H., Deshpande, V., & Ryan, J. K. (2006). Emergency transshipment in decentralized dealer networks: When to send and accept transshipment requests. *Naval Research Logistics (NRL)*, *53*(6), 547–567.

APPENDIX: Proofs of Lemmas/Theorems

Proof of Theorem 1. **(i)** Consider adding either product x or y to an existing assortment S such that $\alpha_x \geqslant \alpha_y$. Let $\vartheta_x(S)$ denote the marginal benefit of adding product x to an existing assortment S assuming that $|S| \leq C-1$, i.e., the assortment capacity is not exceeded by adding product x. We can represent the increase in profit by adding x as follows.

$$\vartheta_x(S) = \Pi(S \cup \{x\}) - \Pi(S) = \lambda r \alpha_x + \lambda r \theta \sum_{i \notin S \cup \{x\}} \alpha_i \delta_{ix} - \lambda r \theta \sum_{i \in S} \alpha_x \delta_{xi}.$$

Next consider adding product y to assortment S.

$$\vartheta_y(S) = \Pi(S \cup \{y\}) - \Pi(S) = \lambda r \alpha_y + \lambda r \theta \sum_{i \notin S \cup \{y\}} \alpha_i \delta_{iy} - \lambda r \theta \sum_{i \in S} \alpha_y \delta_{yi}.$$

When we replace δ_{ik} with its open form $\frac{\alpha_k}{1-\alpha_i}$, the difference between marginal profits of two alternative assortments is obtained as follows.

$$\begin{aligned}
&\vartheta_x(S) - \vartheta_y(S) \\
&= \lambda r(\alpha_x - \alpha_y) + \lambda r \theta \Big[\alpha_x \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} - \alpha_y \sum_{i \notin S \cup \{y\}} \frac{\alpha_i}{1-\alpha_i} - \sum_{i \in S} \alpha_i \big(\frac{\alpha_x}{1-\alpha_x} - \frac{\alpha_y}{1-\alpha_y}\big)\Big] \\
&= \lambda(\alpha_x - \alpha_y) r + \lambda r \theta \Big[\sum_{i \notin S \cup \{x,y\}} \frac{\alpha_i(\alpha_x - \alpha_y)}{1-\alpha_i} - \frac{\alpha_x \alpha_y (\alpha_x - \alpha_y)}{(1-\alpha_x)(1-\alpha_y)} - \sum_{i \in S} \frac{\alpha_i(\alpha_x - \alpha_y)}{(1-\alpha_x)(1-\alpha_y)}\Big] \\
&= \lambda(\alpha_x - \alpha_y) r \Big[1 - \frac{\theta(\alpha_x \alpha_y + \sum_{i \in S} \alpha_i)}{(1-\alpha_x)(1-\alpha_y)} + \theta \sum_{i \notin S \cup \{x,y\}} \frac{\alpha_i}{1-\alpha_i}\Big]. \qquad (5.1)
\end{aligned}$$

Consider the first and the second terms inside the square brackets in (5.1). We can show that

$$
\begin{aligned}
1 &\geq \frac{\theta(\alpha_x\alpha_y + \sum\limits_{i\in S}\alpha_i)}{(1-\alpha_x)(1-\alpha_y)} \\
1-\alpha_x-\alpha_y+\alpha_x\alpha_y &\geq \theta(\alpha_x\alpha_y+\sum_{i\in S}\alpha_i) \\
1-\alpha_x-\alpha_y-\theta\sum_{i\in S}\alpha_i &\geq (\theta-1)\alpha_x\alpha_y \qquad (5.2)
\end{aligned}
$$

As $1-\alpha_x-\alpha_y \geq \sum\limits_{i\in S}\alpha_i$ holds by definition and given that $\theta \leq 1$, the left hand side of (5.2) is non-negative, while the right hand size is non-positive. Thus, (5.2) is always satisfied and the sum of the terms inside square brackets in (5.1) is non-negative. So, for $\alpha_x \geqslant \alpha_y$, it is never better to include y rather than x. Hence the optimal assortment of a single firm is in the popular set when all products have the same profit margin.

(ii) Consider the marginal benefit of adding product x to an existing assortment S. After replacing δ_{ik} with its open form $\frac{\alpha_k}{1-\alpha_i}$, $\delta_x(S)$ is rearranged as follows.

$$
\begin{aligned}
\vartheta_x(S) &= \lambda r\alpha_x + \lambda r\alpha_x\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - \lambda r\theta\frac{\alpha_x}{1-\alpha_x}\sum_{i\in S}\alpha_i \\
&= \lambda r\alpha_x\Big(1+\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i} - \theta\frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i\Big). \qquad (5.3)
\end{aligned}
$$

Consider the first and the third terms inside the parenthesis in (5.3). We can show that

$$
\begin{aligned}
1 &\geq \theta\frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i \\
1 &\geq \alpha_x+\theta\sum_{i\in S}\alpha_i.
\end{aligned}
$$

The result holds given that $\theta \leq 1$, $\sum\limits_{i\in N}\alpha_i$=1, and $S\cup\{x\}\subseteq N$. So, it is always profitable to add a product to the assortment as long as there is available capacity.

Hence, capacity is fully utilized at the optimal assortment when all products have the same profit margin. □

Proof of Lemma 1. **(i)** Consider adding product x to an existing assortment S. The marginal profit of adding x is as follows. Recall that $\vartheta_x(S)$ denotes the marginal benefit of adding product x to an existing assortment S assuming that $|S| \leq C-1$, i.e., the assortment capacity is not exceeded by adding product x.

$$\vartheta_x(S) = \lambda\alpha_x r_x + \lambda r_x\theta \sum_{i \notin S\cup\{x\}} \alpha_i\delta_{ix} - \lambda\theta \sum_{i\in S} \alpha_x\delta_{xi}r_i.$$

Now consider adding product y to the existing assortment S.

$$\vartheta_y(S) = \lambda\alpha_y r_y + \lambda r_y\theta \sum_{i \notin S\cup\{y\}} \alpha_i\delta_{iy} - \lambda\theta \sum_{i\in S} \alpha_y\delta_{yi}r_i.$$

When we replace δ_{ik} with its open form and after some rearrangement, the difference in marginal profits of two alternative assortments is as follows.

$$\begin{aligned}\vartheta_x(S) - \vartheta_y(S) \quad &= \quad \lambda(\alpha_x r_x - \alpha_y r_y) + \lambda\theta(\alpha_x r_x - \alpha_y r_y) \sum_{i\notin S\cup\{x,y\}} \frac{\alpha_i}{1-\alpha_i} \\ &\quad +\lambda\theta\Big(\frac{\alpha_y}{(1-\alpha_y)} \sum_{i\in S\cup\{x\}} \alpha_i r_i - \frac{\alpha_x}{(1-\alpha_x)} \sum_{i\in S\cup\{y\}} \alpha_i r_i\Big). \qquad (5.4)\end{aligned}$$

In (5.4), the first term is the profit difference from direct customer demands, the second one is due to the substitution from out of assortment products, and the last term is the profit difference between the substitution from product y to other existing products and the substitution from x to others. First and second terms in (5.4) are non-negative by the dominance of x over y, i.e., $\alpha_x r_x \geq \alpha_y r_y$. Third term is non-negative as $\alpha_x r_x \geq \alpha_y r_y$ and $\alpha_y \geq \alpha_x$. Thus, the expected marginal profit of adding product x rather than y to an existing assortment is non-negative under given conditions.

(ii) Reconsider the difference in marginal benefits of adding product x and y to an existing assortment S stated by (5.4). (5.4) can be rearranged as follows.

$$\begin{aligned}\vartheta_x(S) - \vartheta_y(S) &= \lambda(\alpha_x r_x - \alpha_y r_y) + \lambda\theta(\alpha_x r_x - \alpha_y r_y)\sum_{i \notin S \cup \{x,y\}} \frac{\alpha_i}{1-\alpha_i} \\ &\quad -\lambda\theta\frac{(\alpha_x - \alpha_y)\sum_{i\in S}\alpha_i r_i - \alpha_x\alpha_y(r_x(1-\alpha_x) - r_y(1-\alpha_y))}{(1-\alpha_x)(1-\alpha_y)}. \quad (5.5)\end{aligned}$$

(5.5) obtains its minimum value with the minimum value of $\sum_{i \notin S \cup \{x,y\}} \frac{\alpha_i}{1-\alpha_i}$. Given a capacity of C, the minimum set of $i \notin S \cup \{x, y\}$ is obtained when $|S| = C - 1$. Recall that $a(i)$ denote the index of the product which has the i^{th} largest product preference, such that $\alpha_{a(1)} \geq \alpha_{a(2)} ... \geq \alpha_{a(N)}$. Then $\sum_{i \notin S \cup \{x,y\}} \frac{\alpha_i}{1-\alpha_i}$ has the minimum value of $\sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}}$. Moreover, as $\sum_{i\in S}\alpha_i r_i$ gets larger, (5.5) gets smaller. (5.5) obtains its maximum value when the most dominant products are in S such that $\sum_{i\in S}\alpha_i r_i = \sum_{i=1}^{|C-1|}\alpha_i r_i$, as $|S| \leq C-1$ should hold to be able to consider adding at least one more product to S. Thus, we can state that

$$\begin{aligned}\vartheta_x(S) - \vartheta_y(S) &\geq \lambda(\alpha_x r_x - \alpha_y r_y)(1 + \theta\sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}}) \\ &\quad -\lambda\theta\frac{(\alpha_x - \alpha_y)\sum_{i=1}^{|C-1|}\alpha_i r_i - \alpha_x\alpha_y(r_x(1-\alpha_x) - r_y(1-\alpha_y))}{(1-\alpha_x)(1-\alpha_y)}.\end{aligned}$$

For this minimum marginal profit difference to be non-negative, it should hold that

$$\theta \leq \frac{(\alpha_x r_x - \alpha_y r_y)(1-\alpha_x)(1-\alpha_y)}{(\alpha_x - \alpha_y)\sum_{i=1}^{|C-1|}\alpha_i r_i - \alpha_x\alpha_y(r_x - r_y) - (\alpha_x r_x - \alpha_y r_y)((1-\alpha_x)(1-\alpha_y)\sum_{i=C+1}^{|N|}\frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \alpha_x\alpha_y)},$$

which is called as $\bar{\theta}_{xy}$. □

Proof of Theorem 2. (i) It directly follows from Lemma 1i.

(ii) Consider the equation $\vartheta_x(S) - \vartheta_y(S)$ in (5.5). We know that the sum of first two terms is always non-negative by definition. The last term is a linear function of θ. Thus, if the last term, which has minus sign in front, is positive, and the total is positive, (5.5) linearly decreases in θ, which can pass through zero from positive to negative at most once. So, there is no chance that (5.5) can switch from negative to positive as θ increases. Next, firstly note that if $r_x \geq r_y$ and $\alpha_x \leq \alpha_y$ do not hold at the same time, then either $r_x \geq r_y$ and $\alpha_x \geq \alpha_y$ or $r_x \leq r_y$ and $\alpha_x \geq \alpha_y$ from the definition of dominance relation between x and y for $x \leq y$. So let $\alpha_x \geq \alpha_y$. As $|S|$ increases, while the second term, which is always non-negative, decreases, the third term, which has a minus sign in front, increases. Thus (5.5) decreases in total. Therefore, as the capacity C increases, $|S|$ can increase, so the marginal benefit of adding product x over product y decreases, which may pass through zero only once from positive to negative. □

Proof of Theorem 3. (i) Consider the change in profit by adding one more product, let say x, to an existing assortment set S.

$$\begin{aligned}
\vartheta_x(S) &= \lambda\alpha_x r_x + \lambda r_x\theta \sum_{i\notin S\cup\{x\}} \alpha_i\delta_{ix} - \lambda\theta\sum_{i\in S}\alpha_x\delta_{xi}r_i \\
&= \lambda\alpha_x r_x + \lambda\alpha_x r_x\theta \sum_{i\notin S\cup\{x\}} \frac{\alpha_i}{1-\alpha_i} - \lambda\theta\frac{\alpha_x}{1-\alpha_x}\sum_{i\in S}\alpha_i r_i \\
&= \lambda\alpha_x\Big[r_x + \theta\Big(r_x \sum_{i\notin S\cup\{x\}} \frac{\alpha_i}{1-\alpha_i} - \frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i r_i\Big)\Big]. \qquad (5.6)
\end{aligned}$$

Let the value of θ that makes $\vartheta_x(S) = 0$ be called θ_{Sx} such that

$$\theta_{Sx} = \Big(\frac{-r_x}{r_x \sum_{i\notin S\cup\{x\}} \frac{\alpha_i}{1-\alpha_i} - \frac{1}{1-\alpha_x}\sum_{i\in S}\alpha_i r_i}\Big).$$

For $\theta \leq \theta_{Sx}$, adding product x to current assortment S is non-hurting. Thus, product x can be added. If $\theta_{Sx} \leq 0$, then independent of θ, product x is profitable to be included in the assortment. For $\theta_{Sx} > 0$, over all possible S sets, θ_{Sx} obtains its smallest value for the minimum value of $\sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i}$ and maximum value of $\sum_{i \in S} \alpha_i r_i$. Given a capacity of C, the minimum set of $i \notin S \cup \{x\}$ is obtained when $|S| = C - 1$. Recall that $a(i)$ denote the index of the product which has the i^{th} largest product preference, such that $\alpha_{a(1)} \geq \alpha_{a(2)} ... \geq \alpha_{a(N)}$. Then $\sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i}$ has the minimum value of $\sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}}$. Moreover, $\sum_{i \in S} \alpha_i r_i$ has the maximum value when $|S| = C - 1$ and it is $\sum_{i=1}^{C-1} \alpha_i r_i$, such that $\alpha_1 r_1 \geq \alpha_2 r_2 ... \geq \alpha_N r_N$ by definition. As a result, θ_{Sx} achieves its minimum non-negative value, which is denoted by $\bar{\theta}_x$, such that

$$\bar{\theta}_x = \max \left\{ 0, \frac{-r_x}{r_x \sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \frac{1}{1-\alpha_x} \sum_{i=1}^{C} \alpha_i r_i} \right\}.$$

Among all non-zero $\bar{\theta}_x$, the minimum value of $\bar{\theta}_x$ over all x is denoted by $\bar{\theta}_c$, which is the critical substitution level to add any product i, $i \in N$. If $\theta \leq \bar{\theta}_c$, it is always profitable to add a product to any given assortment as long as the capacity limit is not exceeded. $\bar{\theta}_c$ can be formally defined as

$$\begin{aligned} \bar{\theta}_c &= \min_{x \in N} \left\{ \bar{\theta}_x \mid \bar{\theta}_x > 0 \right\} \\ &= \min_{x \in N} \left\{ \bar{\theta}_x > 0 \mid \bar{\theta}_x = \frac{-r_x}{r_x \sum_{i=C+1}^{|N|} \frac{\alpha_{a(i)}}{1-\alpha_{a(i)}} - \frac{1}{1-\alpha_x} \sum_{i=1}^{C-1} \alpha_i r_i} \right\}. \end{aligned}$$

(ii) If $\theta > \bar{\theta}_c$, it is possible that (5.6) can be negative if the equation in parenthesis is negative. If it is the case, as θ increases $\delta_x(S)$ decreases, which may lead to under utilization of the capacity. On the other hand, for a fixed θ, if the term in parenthesis in (5.6) decreases, $\delta_x(S)$ decreases as well. This may happen as the set of set $i \notin S \cup \{x\}$ decreases, which may be the result of an increase in capacity C.

□

Proof of Theorem 4. Consider adding product x to every firm's assortment. Let

$\bar{S}$ denote the current common assortment. Then let the total marginal profit of adding product x to the common assortment be denoted as $\vartheta_x(\bar{S})$ such that $\vartheta_x(\bar{S}) = \Pi(S_1 \cup \{x\}, .., S_M \cup \{x\}) - \Pi(S_1, .., S_M)$.

$$\begin{aligned}\vartheta_x(\bar{S}) \;&=\; \lambda\alpha_x r\sum_{m=1}^{M} q_m + \lambda r\theta\sum_{m=1}^{M} q_m \sum_{i\notin S\cup\{x\}} \alpha_i\delta_{ix} \\ &\quad -\lambda r\theta\sum_{m=1}^{M} q_m \sum_{i\in S_m} \alpha_x\delta_{xi} - \lambda(r-\tau)\theta\sum_{m=1}^{M} q_m \sum_{i\in S\setminus S_m} \alpha_x\delta_{xi}. \end{aligned} \tag{5.7}$$

Now consider adding product y to each firm's assortment, where the current assortment is $(S_1, .., S_M)$. The total marginal profit of adding product y to the common assortment is $\vartheta_y(\bar{S})$ and defined as follows.

$$\begin{aligned}\vartheta_y(\bar{S}) \;&=\; \lambda\alpha_y r\sum_{m=1}^{M} q_m + \lambda r\theta\sum_{m=1}^{M} q_m \sum_{i\notin S\cup\{y\}} \alpha_i\delta_{iy} \\ &\quad -\lambda r\theta\sum_{m=1}^{M} q_m \sum_{i\in S_m} \alpha_y\delta_{yi} - \lambda(r-\tau)\theta\sum_{m=1}^{M} q_m \sum_{i\in S\setminus S_m} \alpha_y\delta_{yi}. \end{aligned}$$

After replacing δ_{ij} with its open form $\frac{\alpha_j}{1-\alpha_i}$, the difference between marginal benefits of two alternative assortments can be written as follows.

$$\begin{aligned}&\vartheta_x(\bar{S}) - \vartheta_y(\bar{S}) \\ &= \lambda r(\alpha_x-\alpha_y)\sum_{m=1}^{M} q_m \;-\lambda(r-\tau)\theta\sum_{m=1}^{M} q_m \sum_{i\in S\setminus S_m}\Big(\frac{\alpha_x\alpha_i}{1-\alpha_x} - \frac{\alpha_y\alpha_i}{1-\alpha_y}\Big) \\ &+\lambda r\theta\sum_{m=1}^{M} q_m\Big[\sum_{i\notin S\cup\{x,y\}}\frac{\alpha_i(\alpha_x-\alpha_y)}{1-\alpha_i} - \frac{\alpha_x\alpha_y(\alpha_x-\alpha_y)}{(1-\alpha_x)(1-\alpha_y)} - \sum_{i\in S_m}\frac{\alpha_i(\alpha_x-\alpha_y)}{(1-\alpha_x)(1-\alpha_y)}\Big] \\ &= \lambda r\theta(\alpha_x-\alpha_y)\sum_{m=1}^{M} q_m\Big[\frac{1}{\theta} + \sum_{i\notin S\cup\{x,y\}}\frac{\alpha_i}{1-\alpha_i} - \frac{\sum_{i\in S_m}\alpha_i + \sum_{i\in S\setminus S_m}\alpha_i + \alpha_x\alpha_y}{(1-\alpha_x)(1-\alpha_y)} \\ &\qquad\qquad + \frac{\tau\sum_{i\in S\setminus S_m}\alpha_i}{r(1-\alpha_x)(1-\alpha_y)}\Big]. \end{aligned} \tag{5.8}$$

Consider the first and third terms inside the square brackets in (5.8). We can show that

$$\begin{aligned}
\frac{1}{\theta} \geq \frac{\sum_{i\in S_m} \alpha_i + \sum_{i\in S\setminus S_m} \alpha_i + \alpha_x\alpha_y}{(1-\alpha_x)(1-\alpha_y)} &= \frac{\sum_{i\in S} \alpha_i + \alpha_x\alpha_y}{(1-\alpha_x)(1-\alpha_y)} \\
1-\alpha_x-\alpha_y-\theta\sum_{i\in S}\alpha_i \geq -(1-\theta)\alpha_x\alpha_y. & \qquad (5.9)
\end{aligned}$$

As $1-\alpha_x-\alpha_y \geq \sum_{i\in S}\alpha_i$ and $\theta \leq 1$ hold by definition, left-hand side of (5.9) is non-negative, while the right-hand side is non-positive. Thus, (5.8) is non-negative. Hence, for $\alpha_x \geq \alpha_y$, it is never better to include y rather than x. As a result, we can claim that the common assortment of multiple firms $\bar{S}$ is in the popular set. □

Proof of Theorem 5. (i) Consider the change in profit of adding product x to the existing assortment set S_m of firm m. Let the marginal change in total profit by adding product x to S_m be denoted as $\vartheta_x(S_m)$ such that $\vartheta_x(S_m) = \Pi(S_1, .., S_m \cup \{x\}, ..., S_M) - \Pi(S_1, .., S_m, .., S_M)$. Given that product x is newly added so $x \notin S_m$, there are two possible cases. Either $x \in S$, i.e., product x already kept by at least one of the other firms or $x \notin S$, i.e., product x is not in the assortment of any firm. Consider the first case, where $x \notin S_m$, but $x \in S$.

Case 1: $x \notin S_m$**,** $x \in S$

$$\begin{aligned}
\vartheta_x(S_m) &= \lambda\alpha_x\tau q_m + \lambda\tau\theta q_m \sum_{i\notin S}\alpha_i\delta_{ix} \\
&= \lambda\alpha_x\tau q_m(1+\theta\sum_{i\notin S}\frac{\alpha_i}{1-\alpha_i}). \qquad (5.10)
\end{aligned}$$

(5.10) is clearly non-negative. Thus, if product x is in the assortment of any one of the firms, adding the same product to the assortment of another firm does never hurt the total system profit. This result holds under our constrained assortment problem, but may not with the unconstrained profit maximization model.

Case 2: $x \notin S$

$$\vartheta_x(S_m)$$
$$= \lambda\alpha_x r q_m + \lambda\alpha_x(r-\tau)\sum_{l=1,l\neq m}^{M} q_l + \lambda r\theta q_m \sum_{i\notin S\cup\{x\}} \alpha_i\delta_{ix} + \lambda(r-\tau)\theta \sum_{l=1,l\neq m}^{M} q_l \sum_{i\notin S\cup\{x\}} \alpha_i\delta_{ix}$$
$$-\lambda r\theta q_m \sum_{i\in S_m} \alpha_x\delta_{xi} - \lambda(r-\tau)\theta q_m \sum_{i\in S\setminus S_m} \alpha_x\delta_{xi} - \lambda r\theta \sum_{l=1,l\neq m}^{M} q_l \sum_{i\in S_l} \alpha_x\delta_{xi}$$
$$-\lambda(r-\tau)\theta \sum_{l=1,l\neq m}^{M} q_l \sum_{i\in S\setminus S_l} \alpha_x\delta_{xi}. \tag{5.11}$$

Dividing (5.11) by $\lambda\alpha_x$ and simplifying further, we obtain

$$\vartheta_x(S_m)/(\lambda\alpha_x)$$
$$= rq_m + (r-\tau)(1-q_m) + \theta(r-\tau(1-q_m))\sum_{l=1}^{M} q_l \sum_{i\notin S\cup\{x\}} \alpha_i$$
$$-\frac{r\theta}{1-\alpha_x}\sum_{l=1}^{M} q_l(\sum_{i\in S_l}\alpha_i + \sum_{i\in S\setminus S_l}\alpha_i) + \frac{\tau\theta}{1-\alpha_x}\sum_{l=1}^{M} q_l \sum_{i\in S\setminus S_l}\alpha_i$$
$$= (r-\tau+\tau q_m)\big(1+\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}\big) - \frac{r\theta}{1-\alpha_x}\sum_{i\in S}\alpha_i + \frac{\tau\theta}{1-\alpha_x}\sum_{l=1}^{M} q_l \sum_{i\in S\setminus S_l}\alpha_i. \tag{5.12}$$

(5.12) obtains its smallest value when $S = N \setminus \{x\}$ and $S_1 \equiv S_2 \equiv .... \equiv S_M.$, such that

$$\vartheta_x(S_m)/(\lambda\alpha_x) \geq r - \tau + \tau q_m - r\theta = r(1-\theta) - \tau(1-q_m). \tag{5.13}$$

For (5.13) to be non-negative, it should hold that

$$\theta \leq 1 - \frac{\tau}{r}(1-q_m) = \bar{\theta}_C^m.$$

Hence, if $\theta \leq \bar{\theta}_c^m$, it is always profitable to add a product to any given assortment of firm m as long as the capacity limit is not exceeded.

(ii) First consider Case 1 in (i). (5.10) is independent of r and non-decreasing in q_m. Next consider Case 2. By rearranging the terms, (5.12) can be rewritten as follows.

$$r\Big[1+\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}-\frac{\theta}{1-\alpha_x}\sum_{i\in S}\alpha_i\Big]-\tau\Big[1+\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}\Big]$$
$$q_m\Big[\tau+\tau\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}\Big]+\frac{\tau\theta}{1-\alpha_x}\sum_{l=1}^{M}q_l\sum_{i\in S\setminus S_l}\alpha_i. \tag{5.14}$$

Consider the first square brackets in (5.14), which is the multiplier of r. As the maximum value of $\sum_{i\in S}\alpha_i$ can be $1-\alpha_x$, so

$$1+\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}-\frac{\theta}{1-\alpha_x}\sum_{i\in S}\alpha_i\geq 1+\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}-\theta\geq 0,$$

where the result follows by definition, $\theta\leq 1$. Thus, $\vartheta_x(S_m)$ is non-decreasing in r as the multiplier of r is non-negative in $\vartheta_x(S_m)$.

Next, consider the last two terms in (5.14). When these two terms are rearranged, we obtain

$$q_m\Big[\tau+\tau\theta\sum_{i\notin S\cup\{x\}}\frac{\alpha_i}{1-\alpha_i}+\frac{\tau\theta}{1-\alpha_x}\sum_{i\in S\setminus S_m}\alpha_i\Big]+\frac{\tau\theta}{1-\alpha_x}\sum_{l=1,l\neq m}^{M}q_l\sum_{i\in S\setminus S_l}\alpha_i. \tag{5.15}$$

As $\sum_{l=1}^{M}q_l=1$ by definition, as q_m increases by ϵ, $1-q_m$ decreases by ϵ. Let's assume that q_m increases by ϵ and q_l is decreased by ϵ. Then the total change in (5.15) is as follows.

$$
\begin{aligned}
&\epsilon\Big[\tau + \tau\theta \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} + \frac{\tau\theta}{1-\alpha_x} \sum_{i \in S \setminus S_m} \alpha_i\Big] - \epsilon\frac{\tau\theta}{1-\alpha_x} \sum_{i \in S \setminus S_l} \alpha_i \\
&\geq \epsilon\Big[\tau + \tau\theta \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} + \frac{\tau\theta}{1-\alpha_x} \sum_{i \in S \setminus S_m} \alpha_i\Big] - \epsilon\frac{\tau\theta}{1-\alpha_x}(1-\alpha_x) \\
&= \epsilon\Big[\tau(1-\theta) + \tau\theta \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} + \frac{\tau\theta}{1-\alpha_x} \sum_{i \in S \setminus S_m} \alpha_i\Big] \\
&\geq 0.
\end{aligned}
$$

Thus, as q_m increases, $\vartheta_x(S_m)$ does not decrease, which concludes the proof. $\square$

Proof of Theorem 6. **(i)** Assume that $C_m = C_l$ and $|S_m| < |S_l|$ at optimal solution. We know from (5.10) that when a product x, which is $\in S$, but $\notin S_m$ is added to the assortment of firm m, then the total system profit cannot decrease. So we can always get a solution with an equal or higher profit by increasing $|S_m|$ such that $|S_m| = |S_l|$, by adding the products $i \in S_l \setminus S_m$ to the assortment of S_m. So this contradicts with $|S_m| < |S_l|$ at optimality and proves that $|S^*_m| = |S^*_l|$ should hold.

(ii) We know from (i) that when $C_1 = ... = C_m = ... = C_M$, then it must hold that $|S^*_1| = ... = |S^*_m| = ...|S^*_M|$ at optimality. Let's assume that all capacity usages are $C-1$ at optimality, i.e., $|S_1| = ... = |S_m| = ...|S_M| = C-1$. We can show that a better or equivalent solution can be obtained by increasing the number of products at each firm to C.

Consider two cases; $S \equiv \bar{S}$ and $S \not\equiv \bar{S}$.

Case 1: $S \equiv \bar{S}$

So all firms have the same assortment, $S_1 \equiv S_2 ... \equiv S_m ... \equiv S_M$. Then the marginal benefit of adding a product $x \notin S$ to all assortments is $\vartheta_x(\bar{S})$ as defined by (5.7). By replacing δ_{ij} with its open form and using $\sum_{m=1}^{M} q_m = 1$, $\vartheta_x(\bar{S})$ can be rewritten as follows.

$$\begin{aligned}
\vartheta_x(\bar{S}) &= \lambda\alpha_x r + \lambda\alpha_x r\theta \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} \\
&\quad - \frac{\lambda\alpha_x r\theta}{1-\alpha_x} \sum_{m=1}^{M} q_m \sum_{i \in S_m} \alpha_i - \frac{\lambda\alpha_x (r-\tau)\theta}{1-\alpha_x} \sum_{m=1}^{M} q_m \sum_{i \in S \setminus S_m} \alpha_i \\
&= \lambda\alpha_x \Big(r + r\theta \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} - \frac{r\theta}{1-\alpha_x} \sum_{m=1}^{M} q_m \big(\sum_{i \in S_m} \alpha_i + \sum_{i \in S \setminus S_m} \alpha_i \big) \\
&\quad + \frac{\tau\theta}{1-\alpha_x} \sum_{m=1}^{M} q_m \sum_{i \in S \setminus S_m} \alpha_i \Big) \\
&= \lambda\alpha_x \Big(r + r\theta \big(\sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} - \frac{1}{1-\alpha_x} \sum_{i \in S} \alpha_i \big) + \frac{\tau\theta}{1-\alpha_x} \sum_{m=1}^{M} q_m \sum_{i \in S \setminus S_m} \alpha_i \Big) \\
&= \lambda\alpha_x \Big(r \big(1 - \frac{\theta}{1-\alpha_x} \underbrace{\sum_{i \in S} \alpha_i}_{\leq 1-\alpha_x} \big) + r\theta \sum_{i \notin S \cup \{x\}} \frac{\alpha_i}{1-\alpha_i} + \frac{\tau\theta}{1-\alpha_x} \sum_{m=1}^{M} q_m \sum_{i \in S \setminus S_m} \alpha_i \Big) \\
&\geq 0.
\end{aligned}$$

So, adding product x to each firm's assortment would never hurt the current total profit $\pi(S)$. As $(S_1 \cup \{x\}, ..., S_M \cup \{x\})$ is not worse off than $(S_1, ..., S_M)$, $|S_1| = ... = |S_m| = ...|S_M| = C - 1$ cannot be the optimal solution.

Case 2: $S \not\equiv \bar{S}$

So all firms don't have the same assortment. Given that $|S_1| = ... = |S_m| = ...|S_M| = C - 1$, if $S_l \not\equiv S_m$, by adding product x, which satisfies $x \in S_l$ and $x \notin S_m$, to S_m and product y, which satisfies $y \in S_m$ and $y \notin S_l$, to S_l would not decrease the total profit by the non-negativity of $\vartheta_x(S_m)$ from (5.10). As a result, it becomes $|S_l \cup \{y\}| = |S_m \cup \{x\}| = C$. Similar additions to each firm's assortment can be made to obtain a solution where each firm will have a capacity usage of C number of products. Hence, this completes the proof. $\square$