ESSAYS ON ASSORTMENT PLANNING AND INVENTORY MANAGEMENT FOR SUBSTITUTABLE PRODUCTS

by

Ying Cao

APPROVED BY SUPERVISORY COMMITTEE:

____
Dorothée Honhon, Chair

____
Kyle Hyndman

____
Ganesh Janakiraman

____
Elena Katok

ESSAYS ON ASSORTMENT PLANNING AND INVENTORY MANAGEMENT FOR SUBSTITUTABLE PRODUCTS

by

YING CAO, BBA, MS

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

MANAGEMENT SCIENCE

THE UNIVERSITY OF TEXAS AT DALLAS

December 2018

ACKNOWLEDGMENTS

Thank you to my husband, Xianghui Peng, my son, Justice Peng, and my daughter, Grace Peng, for their patient and kind love which makes it possible for me to finish my PhD. Thank you to my parents, my parents in law, and many other family members for their love and support in my life. Thank you to my dissertation committee members, Dr. Kyle Hyndman, Dr. Ganesh Janakiraman and and Dr. Elena Kotok, for guiding me through this process and always being there to provide me insightful advice and encouraging words. A special thank you to my dissertation chair, Dr. Dorothée Honhon, for her guidance, encouragement, and support during my entire PhD study. I would like to extend my deepest gratitude and highest respect to Dr. Honhon.

August 2018

ESSAYS ON ASSORTMENT PLANNING AND INVENTORY MANAGEMENT FOR SUBSTITUTABLE PRODUCTS

Ying Cao, PhD
The University of Texas at Dallas, 2018

Supervising Professor: Dorothée Honhon, Chair

This dissertation consists of three essays which study assortment planning and inventory management of substitutable products motivated by different practical problems.

Chapter 2 considers the assortment planning problem for a retailer who faces customers that buy multiple differentiated products (n-pack) on a store visit. We develop two choice models: the probabilistic choice rule which captures the heterogeneous consumer population choice pattern and maximum choice rule which captures the homogeneous consumer population choice pattern. We find that, under probabilistic choice rule, the optimal assortment is such that it includes a certain number of the most and least popular products. In contrast, under maximum choice rule, the optimal assortment does not have a fixed structure except that it is guaranteed to include the most popular product. We develop an algorithm under maximum choice rule which is shown to have good performance. In addition, we derive the structure of optimal assortment under both choice rules when a retailer ignores key features of n-pack choice model including choice premium and basket shopping behavior. We also conduct a numerical study where we show that ignoring these key features can lead to significant profit loss for a retailer.

Chapter 3 explores the assortment planning for a firm who faces a *two-sided market*. That is, the firm receives revenues from two distinct user groups: the *customers*, who pay for the

products it sells and the *advertisers* who pay to advertise their brand to the customers. We obtain structural properties of the optimal assortment. We also consider the case where the firm is allowed to offer multiple products with the same attractiveness profile and price. In this case, we obtain conditions under which the optimal assortment is made out of distinct products. In addition, we show that ignoring the revenue from the customers or the advertisers, or focusing only on one segment when making product assortment decisions can lead to a significant revenue loss; specifically, we derive the theoretical bound on revenue loss in these situations.

Chapter 4 studies the decision making of an inventory manager who needs to decide order quantities of multiple substitutable products in his store. As such, the decision maker typically checks the sales history of the products. When there is stock-out, the sales history provides inaccurate information because the lost sales are unobservable and the sales from substitution are indistinguishable from first-choice sales, which we refer to as the "double-censoring effect". To study the impact of substitution rate and information amount on decision maker's performance, we design an experiment where subjects need to decide inventory levels for 2 substitutable products in consecutive 30 periods. The experimental data shows that subjects underestimate the demand for high demand product and overestimate the demand for low demand product. Moreover, the bias is worse when there is substitution in fully censored information treatment. Also, when subjects are provided with less information, they tend to order larger quantity in early periods in order to learn demand.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This dissertation studies assortment planning and inventory management of substitutable products motivated by different practical problems.

In Chapter 2, I study the assortment planning problem for a retailer who faces customers who buy multiple differentiated products on a given store visit, which Fox et al. (2017) refer to as buying an"n-pack" of substitutable products. Based on the n-pack selection model, we develop a choice model which we use to calculate product demands given a set assortment. In particular, we consider two different choice rules (the maximum choice rule and probabilistic choice rule) which respectively capture two different consumer choice patterns that arise in practice. We study structural properties of the optimal assortment and explore how the retailer's assortment decision and total profits are impacted when the retailer ignores a key feature of the n-pack choice model, called the"choice premium" which captures the utility that consumers derive from variety in their shopping basket and allows them to hedge against future preference uncertainty. The major results of our study can be summarized as follows. First, we find that, under probabilistic choice rule, the optimal assortment has the following structure: it includes a certain number of the most and least popular products (which we refer to as a "popular-unpopular set"). In contrast, under maximum choice rule, one can only guarantee that the optimal assortment includes the most popular product. These results on the structure of the optimal assortment differ from the well-known "popular set" result from van Ryzin and Mahajan (1999) who show that, when consumer buys at most one unit and their choice is captured by the Multinomial Logit (MNL) model, it is optimal to offer a number of the most popular products (i.e., a "popular set"). Under maximum choice rule, we propose a heuristic method for selecting an assortment and show numerically that it has good performance. Further, we investigate the impact of ignoring the choice premium which drives consumer choice. We find that when the retailer ignores the choice premium, the optimal assortment under probabilistic rule is still a popular-unpopular set, but it can be different from the optimal assortment of the problem when choice premium is considered; while under maximum choice rule, it is always optimal to offer only the most popular product when choice premium is ignored. Finally, we show numerically that ignoring the consumer multi-item purchasing behavior can lead to significant profit losses for the retailer.

In Chapter 3, I consider a firm who must decide on the assortment of products to offer when facing a *two-sided market*. As such, the firm receives revenues from two distinct user groups: the *customers*, who pay for the products it sells and the *advertisers* who pay to advertise their brand to the customers. The customers come from different segments which differ in their preferences for the products. We contribute to the literature on assortment planning by adding the second market dimension to the product selection problem with multiple customer segments. The main results of our study are as follows. First, we show that the assortment problem for a two- sided market with multiple customer segments is complex in that there is no simple structure to the optimal assortment. In particular, we show that a revenue-ordered solution, adapted to the two-sided revenue stream nature of our problem, is not necessarily optimal for the problem in its most general form; however, we provide conditions under which the optimal assortment has a simple ranking-based structure. Further, we show that focusing only on one group of users, that is treating the problem as a one-sided rather than two-sided market, can lead to a significant loss in revenue, and we obtain the theoretical bound for these heuristics. We also investigate a special case wherein the possible products to offer either appeal to all customer segments equally or are tailored to the tastes of one of them.

Chapter 4 is an experimental study on a retailer's inventory management of two substitutable products. Corsten and Gruen (2003) pointed out: *"A retailer that needs to reorder a product will typically examine the sales history of that product. When the item has been out of stock, the sales history data provide inaccurate information to the buyer on what is* *the necessary purchase quantity to meet actual demand. If the out-of-stock has not been detected, then the buying decision will most likely be too low to meet the normal customer demand plus those who delayed purchase until the retailer received additional stock. Alternatively, if the buyer is aware of the OOS situation, then tendency may be to over order, because the buyer is unable to determine the permanent loss of customers caused by the OOS through brand substitution or to the store due to store switching."* For categories with substitutable products, stock-outs lead to a double-censoring effect: lost sales are unobservable and the sales from substitution are indistinguishable from first-choice sales. We design an experiment that includes three treatments and provide subjects with different amount of information regarding demand and sales. In all three treatments, subjects play in the role of a store manager who must make order quantity decisions for two products over 30 consecutive periods in a Newsvendor setting. The two products are substitutable products from the same category and one of them has a lower average demand than the other. The nalysis of the experiment data show that: (1) Subjects have difficulty learning demand. Specifically, they underestimate the demand for high demand product and overestimate the demand for low demand product. (2) When there is substitution, the bias is worse in fully censored information treatment. (3) In fully censored information treatment, subjects tend to order larger quantity than in partially censored information treatment in order to learn demand. (4) In signal treatment, subjects's order quantities for both products are larger than fully censored information treatment and partially censored information treatment.

CHAPTER 2

ASSORTMENT PLANNING FOR n-PACK PURCHASING CONSUMERS

2.1 Introduction

Consumers often purchase multiple items from the same product category on a given shopping trip. Dubé (2004) study some of the top revenue-generating categories in the dry grocery department in U.S. food stores (specifically carbonated soft drinks, ready-to-eat cereals, canned soups and cookies) and find that more than 20% of shopping trips include multiple products being purchased in the same category. Further, 61% of shopping trips for carbonated soft drinks result in the purchase of multiple kinds of products, that is, different flavors with possibly more than one item from each kind. According to Harlam and Lodish (1995), the corresponding number is 74% for yogurt, and is 78% for canned soup. This multi-item purchase behavior can be explained by a desire to avoid multiple trips to the store as well as a taste for variety (Simonson (1990)).

Once at home, consumers consume the products they have purchased over time. At each consumption occasion, their decision of which product to consume is constrained by their available inventory of the product which depends on their initial shopping decision and past consumption decisions. Fox et al. (2017) propose a novel dynamic programming-based consumer choice model which captures the shopping and consumption incentives of consumers. While at the store, consumers select n products for future consumption, referred to as an *n-pack*. The composition of a consumer's chosen n-pack is a function of the expected utility values of the possible product alternatives assuming a forward-looking behavior: the consumer optimizes which product to consume at each consumption occasion, taking into account both his immediate utility and future expected utility. Under some mild assumptions on choice behavior, Fox et al. (2017) show that the overall expected value of an n-pack can be expressed in closed form and includes a *choice premium*, which is the additional expected utility from having the freedom to consume products in any given order.

When making product assortment decisions retailers must take this multi-item shopping behavior into account. Yet, most research on assortment planning has focused until now on models which assume single-unit purchases by all incoming consumers[1]. In this paper, we revisit the assortment planning problem for a single product category when retailer faces multi-item purchasing, so called "n-pack" consumers. Our model uses the expected value derivations from Fox et al. (2017) as inputs into a demand function to calculate the retailer's expected profit from each possible assortment. We consider two possible choice rules: consumers either purchase the maximum value n-pack with probability one (*maximum choice rule*) or they probabilistically choose between each possible n-pack using an attraction-based formula akin to the Multinomial Logit (MNL) model purchase probability equation (*probabilistic choice rule*).

In our base setting we assume that n, i.e., the size of the pack purchased, is identical for all consumers in the population. The set of potential products to offer in the product category is finite and alternatives differ based on a nominal utility (or *popularity*) parameter which is common to the consumer population. As in **?** we assume a fixed selling price for all products and concave inventory costs, which imply that the retailer benefits from economies of scale, thereby preferring to offer fewer products in larger quantities, compared to more variety in smaller quantities. Our model simplifies to the problem considered in **?** when the size of the pack is set equal to 1, that is, when all consumers buy a unique product and their choice is captured by the MNL model.

We summarize our results as follows, First, we find that, under probabilistic choice rule, the optimal assortment has the following structure: it includes a certain number of the most and least popular products (we say the optimal assortment is a "*popular-eccentric set*"). In contrast, under maximum choice rule, one can only guarantee that the optimal assortment

[1]One notable exception of the work by (Cachon and Kök (2007)) who consider the basket shopping behavior across product categories.

includes the most popular product. These results on the structure of the optimal assortment differ from the well known "*popular set*" result from **?** who show that, when consumer buys at most one unit and their choice is captured by the MNL model, it is optimal to offer a subset of the most popular products. Under maximum choice rule, we propose a heuristic method for selecting an assortment and show numerically that it reaches optimality in 88.54% of the problem instances we considered. Further, we show numerically that ignoring the consumer multi-item purchasing behavior can lead to significant profit losses for the retailer - up to 100%. Finally we investigate the impact of ignoring the choice premium which drives consumer choices and find that the retailer may loose up to 100% in profit.

Our paper contributes to the literature in assortment planning in three ways. First, to the best of our knowledge, we provide a first look at the problem of discrete product selection when consumers buy multiple items from the same product category. We do so using a novel model of consumer choice which is well grounded in the theory of shopping and consumption, namely n-pack consumer choice model from Fox et al. (2017). We augment this model by considering two possible choice rules to calculate individual product demand. Second, we highlight a new possible structure for the optimal assortment in a given product category, namely the *popular-eccentric* sets and prove their optimality when consumer choose probabilistically between all possible n-packs. Finally we highlight the need for retailers to incorporate consumers' multi-item shopping behavior and the choice premium they receive when buying more than one unit on each shopping trip.

The rest of this chapter is organized as follows. In §2.2 we review the related literatures. In §2.3 we formulate our model and discuss some of its special cases. We present our analytical results in §2.4. In §2.5 we discuss the impact of choice premium on optimal assortment and profit. In §2.6 we conduct a numerical study. Unless otherwise stated, all proofs are in the Appendix A.

2.2 Literature Review

This paper is mainly correlated with 3 streams of literatures as discussed below: (1) Multi item purchasing decisions for consumers; (2) Assortment planning; (3) Bundling.

Consumers' multi-product shopping behaviors have been studied by a large number of scholars. Simonson (1990) was one of the earliest studies on shopping decisions in presence of preference uncertainty. They showed through three experiments that consumers systematically seek more variety when simultaneously choosing a collection of products for future consumption(s) compared to choosing each product sequentially at the time of consumption. This variety-seeking phenomenon and relevant results have also been generated in a series of research after that (Simonson and Winer (1992), Read and Loewenstein (1995), Read et al. (1999)). In addition to the variety-seeking behavior, Salisbury and Feinberg (2008) pointed out that the diversification in consumers' choice may involve a rational response to some preference uncertainty, for example the brand attractiveness uncertainty. Another related stream of papers study the consumption behavior and its impact on consumers' purchasing decisions. Bown et al. (2003) found through an experimental study that people tend to preserve options from variety for the future consumption, even though this may lead to less desirable outcomes. Guo (2010) developed a structural econometric model for consumers' choice of n-packs and estimated his model on scanner panel data for yogurt purchases where the consumption data was obtained using simulation. Guo showed that allowing for state dependence and consumption flexibility due to preference uncertainty fits the scanner panel data better than nested models, and that consumers make consistent multi-item purchase overtime. Fox et al. (2017) used dynamic programming to determine the optimal consumption policy and the maximum expected value of consuming any n substitutable products selected while shopping (an n-pack). They proposed a canonical model and a generalized model, the later one involves outside option so that an alternative need not be consumed on each consumption occasion.

Planning and management of retail assortment (product variety) have enjoyed substantial research attention from several different angles. The majority of the papers that study assortment decisions assume that each consumer buys at most one product at a store visit. A growing body of work explores relevant operational concerns such as inventory (Aydin and Ryan (2000), Gaur and Honhon (2006), Maddah and Bish (2007), Smith and Agrawal (2000), van Ryzin and Mahajan (1999)), delivery leadtime (Alptekinog? lu and Corbett 2010), substitution upon stockout (Honhon et al. (2010)), and modularity in product design (Hopp and Xu (2005)). Another line of research tackles strategic and competitive aspects such as entry deterrence (Bayus and Putsis (1999)), competitive implications of consumer search (Cachon et al. (2008)), basket-shopping behavior (Cachon and Kök (2007)), brand preference (Kök and Xu (2011)), product satiation (Caro and Martnez-de-Albniz (2012)), and mass customization (Alptekinoğlu and Corbett (2008)). There are few papers which study assortment based on multi-item purchasing customers. One of these is Cachon and Kök (2007) which studies basket shopping behavior across product categories. In this paper, it is assumed that the customer purchase at most one item from each product category considered. This is different from our paper since we study the assortment planning for one product category where customers come to buy multiple item from the category.

2.3 Model

We consider a retailer who selects products to offer in a given product category. Let $\mathcal{M} = \{1, ..., M\}$ denote the set of possible product alternatives to offer and let $S \subseteq \mathcal{M}$ be the assortment chosen by the retailer. All product alternatives in the category are sold at a price p and purchased from a supplier at a cost c. Let $D_j(S)$ denote the expected demand for product j given assortment S and let $\vec{D}(S) = (D_1(S), ..., D_M(S))$ denote the vector of expected demand values. We assume that the retailer incurs inventory holding costs which are captured by a concave function of expected demand as modeled by **?**. Specifically, we assume that total expected inventory costs are equal to $\sigma \sum_{j \in S} [D_j(S)]^\eta$ where $\sigma \geq 0$ and $\eta \in (0, 1]$. The retailer's expected profit is denoted by $\Pi(S) = \pi(\vec{D}(S))$. The retailer's problem is:

$$\max_{S \subseteq \{1,\dots,M\}} \Pi(S) = \pi(\vec{D}(S)) = \sum_{j \in S} (p - c) D_j(S) - \sigma \sum_{j \in S} [D_j(S)]^\eta \tag{2.1}$$

It is easy to verify that total expected profit is convex in $D_j(S)$ for $j \in S$ and increasing in $D_j(S)$ for $D_j(S) \geq \underline{D} = \left(\frac{p-c}{\sigma\eta}\right)^{1/(\eta-1)}$. For most practical problems, we have $p - c > \sigma$ which implies that $\underline{D}$ is a very small quantity; in what follows, we assume that the demand for each product is always greater or equal than $\underline{D}$ so that the retailer's profit is convex increasing in the demand for each product included in the assortment.

Note that, when $\eta = 1$, total expected inventory costs increase linearly with total demand $\sum_{j \in S} D_j(S)$ so that the retailer's problem amounts to a market share maximization problem. In contrast, when $\eta < 1$, the retailer experiences economies of scale in inventory costs due to a pooling effect: for example, an expected demand value of 100 for a single product alternative costs less in inventory costs than two expected demand values of 50 for two distinct product alternatives. The retailer therefore has an incentive to capture demand with as few products as possible in her assortment.

Consumers may buy more than one product from the product category; as in Fox et al. (2017) we refer to the basket of products purchased by a consumer as an *n-pack* where n is the total number of units which are purchased. We assume that all consumers in the population buy the same size basket but that the composition of the n-pack differs from consumer to consumer; an n-pack may contain multiple different product alternatives with one or more units of each kind. Let λ denote the size of the consumer population.

We use the consumer choice model developed by Fox et al. (2017). Specifically, let U_j denote the utility parameter from consuming product alternative j for $j \in \mathcal{M}$. Without loss of generality we assume that the product alternatives are ordered such that $U_1 \geq U_2 \geq \dots \geq$

U_M. The actual utility that consumer i receives from product alternative j is given by $U_j + \epsilon_{ij}$ where the random error terms ϵ_{ij} are independent and assumed to have a standard Gumbel distribution with cumulative distribution function $F(x) = \exp(-e^{-x/\mu - \gamma})$, where γ is Euler's constant (approximately equal to 0.5772). This distribution has a mean equal to zero so that the expected utility from product alternative j is equal to U_j. An n-pack is represented by a $(M \times 1)$ vector $\vec{k} = (k_1, ..., k_M)$ where k_j denotes the number of units of product alternative j, so that $k_j \in \{0, 1, ..., n\}$ and $k_1 + ... + k_M = n$. Let $V(k_1, ..., k_M)$ measure the expected utility or *value* received by the consumer from consuming the n-pack over a period of time consisting of n staggered consumption occasions. As shown in Fox et al. (2017), this value is calculated assuming the consumer is forward-looking in his consumption decisions. This means that he uses dynamic programming to optimize which product to consume at each consumption occasion, taking into account both his immediate utility and future expected utility. Fox et al. (2017) show that V can be calculated as follows:

$$V(k_1, ..., k_M) \quad = \quad \ln \left(\left(\sum_{j=1}^{M} k_j \right)! \right) - \ln \left((k_1)!(k_2!)...(k_M)! \right) + k_1 U_1 + ... + k_M U_M \quad (2.2)$$

As argued by Fox et al. (2017), the first term in this expression reflects "the additional expected utility of having the freedom to consume products in whatever order one chooses" which they refer to as a *choice premium*. For a given n, the term $\ln \left(\left(\sum_{j=1}^{M} k_j \right)! \right) = \ln(n!)$ is a constant effect associated with the pack size and the term $\ln(k_1! \dots k_M!)$ reflects the impact of variety and inventory on consumers' utility from consuming the n-pack. When n is fixed, the choice premium is minimized (and equal to zero) when it contains n units of the same product alternative and is maximized when the quantities are evenly distributed between the available products (as much as possible given the integrality constraints).

If they buy from the retailer, consumers may only buy products which are offered in the her assortment, therefore, given an assortment S, the $\vec{k}$ vectors must satisfy $k_j = 0$ for $j \notin S$. Let $\mathcal{K}^n(S)$ denote the set of all n-packs which satisfy this condition given n and S, i.e., $\mathcal{K}^n(S) = \{\vec{k} : k_j = 0 \text{ for } j \notin S \text{ and } k_1 + ... + k_M = n\}$. The number of possible n-packs given n and S is equal to $\left(\!\binom{|S|}{n}\!\right) = \binom{|S|+n-1}{n} = \frac{(|S|+n-1)!}{n!(|S|-1)!}$, where $|S|$ denotes the size of set S. We assume that consumers face an outside option with attractiveness V_0, which represents the consumers' expected utility from not buying from the retailer's assortment. Let $P^n(\vec{k}, S)$ denote the probability of an n-pack consumer picking the n-pack $\vec{k}$ given assortment S.

Maximum choice rule Under the *maximum choice rule*, each n-pack consumer picks (with probability 1) the n-pack which yields for him the highest expected utility, provided the expected utility it gives him which is higher than that of the outside option. Let $\vec{k}^{*n}(S) = (k_1^{*n}(S), \ldots, k_M^{*n}(S))$ be the n-pack which generates maximum expected value among all the n-pack's in $\mathcal{K}^n(S)$, that is $\vec{k}^{*n}(S) = \arg \max_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k})$. In the rest of this paper, we will call $\vec{k}^{*n}(S)$ the *maximum-value n-pack*. We have:

$$P^n(\vec{k}, S) = \begin{cases} 1 & \text{for } \vec{k} = \vec{k}^{*n}(S). \\ 0 & \text{otw.} \end{cases}$$

Under the maximum choice rule, the demand for product $j \in S$ simplifies to:

$$D_j(S) = \lambda \mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) k_j^{*n}(S) \tag{2.3}$$

where $\mathbb{1}(E)$ is the indicator function for event E.

Probabilistic choice rule Under the *probabilistic choice rule*, we assume that each consumer chooses a specific n-pack with a probability equal to the ratio of its value to the value of al possible n-packs and the outside option. Specifically, the probability of choosing n-pack $\hat{\vec{k}} = (\hat{k}_1, ..., \hat{k}_M)$ is $P^n(\hat{\vec{k}}, S)$ which is given by

$$P^n(\hat{\vec{k}}, S) = \begin{cases} \frac{V(\hat{\vec{k}})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} & \text{for } \vec{k} \in \mathcal{K}^n(S). \\ 0 & \text{otw.} \end{cases} \tag{2.4}$$

Accordingly, each consumer chooses the outside option with probability $\frac{V_0}{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k})+V_0}$.

This choice rule is inspired by the attraction models such as the classic Multinomial Logit (MNL) model.

We summarize our notation in Table 2.1.

Table 2.1. Notation Table

Notation	Definition	Notation	Definition
$\mathcal{M}$	set of possible products	M	number of possible products
n	pack size	S	assortment
λ	population size	p	selling price
c	purchasing cost	σ	inventory cost multiplier
η	inventory cost power	U_j	utility of product j
γ	Euler's constant	V_0	utility of outside option
$\vec{k}$	n-pack vector	$\vec{k}^{*n}(S)$	maximum-value n-pack
V	value of n-pack	V^{*n}	value from the maximum-value n-pack
$\mathcal{K}^n(S)$	set of possible n-packs	$P^n(\vec{k}, S)$	purchase probability for n-pack $\vec{k}$
$D_j(S)$	demand for product j given S	$\vec{D}(S)$	demand vector
$\Pi(S)$	profit function for assortment S	$\pi(\vec{D}(S))$	profit function for demand vector $\vec{D}(S)$

The following example illustrates the differences between the two choice rules.

Example 2.3.1. Let $M = 3$, $V_0 = 3$, $U_1 = 2$, $U_2 = 1.2$ and $U_3 = 1$, $\sigma = 1$, $\eta = 0.8$, $p-c = 10$ and $\lambda = 100$. Suppose the retailer chooses $S = \{1, 2\}$. Table 2.2 presents the possible n-packs and their corresponding value for $n = 1, 2$ and 3. Interestingly, when $n = 3$, n-pack (2,1,0) provides greater value than 3-pack $(3, 0, 0)$ thanks to the choice premium: consumers receive extra utility from consuming different product alternatives, even if it means buying a unit of product with a lower utility, i.e., product 2 instead of product 1. Table 2.3 shows the demand for each product, total demand and retailer's profit under the two choice rules when $n = 1$, $n = 2$ and $n = 3$.

Under probabilistic choice rule, we see that all the products in the assortment always have positive expected demand. In contrast, under maximum choice rule, a product offered in the assortment can receive zero demand and total expected demand for the assortment

Table 2.2. Value of n-packs (maximum-value n-packs in bold characters)

	$\vec{k}$	$V(\vec{k})$
$n=1$	**(1,0,0)**	**2**
	(0,1,0)	1.2
$n=2$	**(2,0,0)**	**4**
	(1,1,0)	3.90
	(0,2,0)	2.4
$n=3$	(3,0,0)	6
	(2,1,0)	**6.30**
	(1,2,0)	5.50
	(0,3,0)	3.6

Table 2.3. Consumer choice under the probabilistic and maximum choice rules

	Probabilistic Choice Rule			Maximum Choice Rule		
	Demand Vector	Total Demand	Retailer Profit	Demand Vector	Total Demand	Retailer Profit
$n=1$	$(32.26, 19.35, 0)$	51.61	489.32	$(0, 0, 0)$	0	0
$n=2$	$(89.47, 65.40, 0)$	154.86	1483.88	$(200, 0, 0)$	200	1930.69
$n=3$	$(147.95, 115.16, 0)$	23.11	2532.07	$(200, 100, 0)$	300.00	2890.88

can be zero when the maximum-value n-pack does not provide more value than the outside option (as is the case when $n = 1$).

2.3.1 Model discussion

In this section we discuss our main modeling assumptions.

The two choice rules we consider should be viewed as two theoretical extremes on the likely spectrum of consumer behavior. Under the probabilistic choice rule, all possible n-pack vectors in $\mathcal{K}^n(S)$ get a positive probability of being purchased, with higher value n-packs receiving a higher probability of purchase. This case captures heterogeneity in the consumer population and can also explain why in practice the same consumer may purchase different sets of products when faced with the same assortment on successive store visits. In contrast, under the maximum choice rule, only the maximum-value n-pack can be purchased by consumers so that product demand is deterministic with all consumers making the same choice from the offered assortment.

In our base model, we have assumed that all consumers buy the same size pack. This simplification of reality allows us to focus on the impact of consumers' shopping basket behavior on a retailer's profits. The chosen n-pack varies with the retailer's assortment in a non-trivial way as illustrated with Example 3.4 above: consumers do not necessarily purchase n units of product with the highest utility. This makes the assortment decision a challenging one to solve, even when n is fixed. Note that, while we assume consumer homogeneity in pack size, our model captures consumer heterogeneity in choice under the probabilistic rule as argued above.

Finally note that we have assumed that the pack size is exogenous to our model and independent of the chosen assortment. In practice, consumers may be led to buy more units of the product by some promotion or an enticing product display. Ultimately the decision of how many products to purchase depends on a consumer's budget, consumption habits, household size, store visit frequency, etc. We abstract from these interesting aspects of the problem in our model so as to focus our attention on the connection between assortment planning and shopping basket size.

2.4 Results

We divide our results into three sections. In §2.4.1 we develop properties of the choice model, more specifically, we provide structural insights on the maximum-value n-pack and corresponding maximum value given a fixed assortment offered by the retailer. In §2.4.2, we study properties of the product demand functions given a fixed assortment and given a consumer choice rule. Finally, in §2.4.3 we consider the assortment optimization problem and provide structural results on the optimal set of products to offer under the probabilistic and maximum choice rules.

2.4.1 Properties of the choice model

In this subsection we present properties of the n-pack choice model as developed by Fox et al. (2017). We focus on the dimensions which are relevant for our analysis and refer the reader to Fox et al. (2017) for more details about the model and its derivation.

Our first result is regarding the structure of the maximum-value n-pack vector.

Lemma 1. The maximum-value n-pack $\vec{k}^{*n}(S)$ is such that $k_i^* \geq k_j^*$ for $i, j \in S$ with $U_i \geq U_j$ for all n.

Given that we have assumed products are ordered such that $U_1 \geq U_2 \geq ... \geq U_M$, this result implies that the maximum-value n-pack is a vector with non-increasing numbers if one ignores the zero values for products which are not included in S. Note that a similar result was noted by Fox et al. (2017) for the special case of $S = \{1, ..., M\}$ (see their Theorem 4(ii)).

Let $\vec{e}^j$ be a vector of size M where all components are equal to zero except the j-th component which is equal to 1. Theorem 4 in Fox et al. (2017) establishes that finding the optimal $\vec{k}^{*n}$ can be done using a greedy procedure as follows:

Algorithm 1 Greedy Algorithm for n-Pack Product Selection

- for $j = 1$, we have $\vec{k}^{*1} = (1, 0, ...0)$.
- for $j = 2 \to n$, $\vec{k}^{*j} = \vec{k}^{*j-1} + \vec{e}^j$ such that $j = \arg\max_{i=1,...,M} V(\vec{k}^{*j-1} + \vec{e}^i) = \arg\max_{i=1,...,M} \{U_i - \ln(k_i^{*j-1} + 1)\}$.

Our next result is regarding the change in the maximum-value n-pack when a product is added to the assortment or when a low-utility product in the assortment is replaced with a high-utility product. In either case, the quantity purchased of the existing products decreases or stays the same.

Lemma 2. Consider assortment $S \subseteq \mathcal{M}$. (i) Given $i \notin S$, we have $k_l^{*n}(S \cup \{i\}) \leq k_l^{*n}(S)$ for all $l \in S$. (ii) Given $j \in S$ and $i \notin S$ such that $U_i \geq U_j$, we have $k_i^{*n}(S \setminus \{j\} \cup \{i\}) \geq k_j^{*n}(S)$ and $k_l^{*n}(S \setminus \{j\} \cup \{i\}) \leq k_l^{*n}(S)$ for all $l \in S, l \neq j$.

To illustrate Lemma 2 (ii), suppose that $S = \{1, 3\}$. What the result says is that, if product 3 is replaced with product 2 which is more popular, customers buying the maximum-value n-pack will buy at least more units of product 2 than they were previously buying of product 3 and cannot buy more units of product 1. In particular, if they buy the same quantity of product 2 as they were buying of product 3, then the quantity of product 1 they buy will remain the same.

Let $V^{*n}(S) = V(\vec{k}^{*n})$ be the value from the maximum-value n-pack, which we refer to as the *n-pack maximum value*. In Proposition 1 we show that $V^{*n}(S)$ increases as (i) n gets larger, (ii) products are added to the assortment, and (iii) a lower-utility alternative in the assortment is replaced with a higher-utility alternative.

Proposition 1. Consider assortment $S \subseteq \mathcal{M}$. (i) $V^{*n}(S)$ is strictly increasing in n. (ii) Given $\hat{S} \subset S$, we have $V^{*n}(\hat{S}) \leq V^{*n}(S)$. (iii) Given $j \in S$ and $i \notin S$ such that $U_i \geq U_j$, we have $V^{*n}(S) \leq V^{*n}(S \backslash \{j\} \cup \{i\})$.

From Proposition 1 (ii), it follows that all consumers derive the most value when the retailer chooses to offer all potential product alternatives, that is, when $S = \mathcal{M}$. However, since the retailer experiences economies of scale in the cost of offering products alternatives, she may choose a smaller assortment which yields a higher profit, as shown in §2.4.3.

2.4.2 Properties of the demand functions

We now explore properties of the expected demand for individual products and of the total expected demand under the maximum and probabilistic choice rules. Unless otherwise stated, our results below apply to both choice rules.

Our first result analyzes the impact on total demand when a product is added to the assortment or when a lower utility product in the assortment is replaced with a higher utility product. In either case, total expected demand increases. Also we show that total expected demand is a nondecreasing function of the utility of each product in the assortment.

Lemma 3. Consider $S \subset \mathcal{M}$. (i) Given $i \notin S$, we have $\sum_{l \in S} D_l(S) \leq \sum_{l \in S \cup \{i\}} D_l(S \cup \{i\})$. (ii) Given $j \in S$ and $i \notin S$ such that $U_i \geq U_j$, $\sum_{l \in S \setminus \{j\} \cup \{i\}} D_l(S \setminus \{j\} \cup \{i\}) \geq \sum_{l \in S} D_l(S)$. (iii) Under probabilistic choice rule, $\sum_{l \in S} D_l(S)$ is nondecreasing and concave in U_j for all $j \in S$ (iv) Under maximum choice rule, for all $j \in S$ there exists a threshold $\overline{U}_j$ such that $\sum_{l \in S} D_l(S) = 0$ for $U_j < \overline{U}_j$ and $n\lambda$ otherwise.

Note that, under maximum choice rule, there are only two extreme situations: either all consumers make a purchase or no one does. Specifically, the total expected demand is equal to zero if the maximum value n-pack does not bring a value which is greater than that of the outside option; otherwise the total expected demand is equal to $n\lambda$, which is the pack size multiplied by the market size. In contrast, under probabilistic choice rule, it is never possible to capture the entire market as there is a positive probability that consumers will choose the outside option (provided its utility V_0 is positive).

Next we explore properties of the expected demand for each product offered in the assortment. First we show that products with a large expected utility receive a larger expected demand.

Lemma 4. Given assortment $S \subseteq \mathcal{M}$ and $i, j \in S$, if $U_i \geq U_j$, then we have $D_i(S) \geq D_j(S)$.

A well-known property of random utility choice models (RUM) is what we refer to as the "*new product demand stealing*" property: for a fixed market size, when a product is added to an assortment, the demand for existing products in the assortment cannot increase (see for example van Ryzin and Mahajan (1999), Gallego et al. (2014), Feng et al. (2017) and

Jagabathula and Rusmevichientong (2016))[2]. We show that this property does not always hold in our setting when customers buy multiple units of the products on the same purchase occasion.

Example 2.4.1. Let $M = 2$, $U_1 = 0.6, U_2 = 0$, $V_0 = 2$, $\lambda = 100$ and $n = 3$. If the retailer offers only product 1, then the maximum-value n-pack is $(3, 0)$ with value $V^* = 1.8$. When she offers both products, the maximum-value n-pack is $(2, 1)$ which gives value $V^* = 2.3$. The purchasing probability and demands for assortment $\{1\}$ and $\{1, 2\}$ are respectively shown in Table 2.4 and Table 2.5.

Table 2.4. Purchase probabilities of n-packs in Example 2.4.1

Assortment	n-packs	Value	Purchase probability of n-packs	
			Maximum rule	Probabilistic rule
$\{1\}$	(3,0)	1.80	0	0.44
$\{1, 2\}$	(3,0)	1.80	0	0.22
	(2,1)	2.30	1	0.28
	(1,2)	1.70	0	0.21
	(0,3)	0	0	0

Table 2.5. Products demands in Example 2.4.1

Assortment	Demand vector	
	Maximum rule	Probabilistic rule
$\{1\}$	(0,0)	(142.11,0)
$\{1, 2\}$	(200,100)	(150,73.05)

Under maximum choice rule, offering only product 1 leads to zero demand since the utility of the outside option is greater than the maximum value. However, offering both products lead to a demand vector of $(200, 100)$ since the maximum value is higher than that

[2]One exception is Wang (2017) which shows the "*new product demand stealing*" property does not hold for MNL model when assuming that the market size is affected by the assortment planning and pricing strategy and the market expansion effect is sufficiently strong.

of the outside option. Hence we see that adding product 2 increases the demand for product 1 from zero to 200.

Under probabilistic choice rule, offering only product 1 gives a demand of 142.11. With both products offered, the vector of product demand is $(150.00, 73.05)$. In other words adding product 2 increases the demand for product 1 from 141.11 to 150.00.

Hence we see that, unlike with RUM models, adding a new product to the assortment may increase demand for the existing products in our model. Under maximum rule, this happens when adding the new product increases the maximum value V^* from consuming an n-pack past the utility of the outside option. Under the probabilistic choice rule, the demand for existing products may increase because adding the new product creates new possible n-packs which combine the new product with the existing ones. It is noteworthy that *new product demand stealing* property of RUM models does not hold in our setting under probabilistic choice rule given that it mimics the MNL model (which is a RUM model) by assigning n-pack purchasing probability proportionally to their value (as previously discussed the MNL model is a special case of our model with probabilistic choice rule obtained when $n = 1$). Further, note that the new product in Example 2.4.1 (Product 2) has zero expected utility. This is another notable difference between our model and the MNL model: under both maximum rule and probabilistic rule, a zero-expected utility product can receive a positive demand and its addition to the assortment generally affects the demand for existing products. In contrast, in the MNL model, zero-expected-utility products always have zero demand and adding them to the assortment does not impact the demand for existing products. Under the probabilistic rule, this is due to the fact that the added product expands the list of possible n-packs to choose from. All these n-packs (except for the one which includes only the new product with zero expected utility) get a positive purchase probability which results in a positive demand for the zero-expected-utility product and impacts the demand for the existing products. And under the maximum choice rule, this happens if it is beneficial to purchase some zero-expected-utility product in the n-pack in order to gain higher value from the choice premium.

Our next lemma establishes that the *new product demand stealing* property holds under maximum choice rule if the existing assortment already captures the entire market. Specifically, adding a new product or replacing a lower-utility product with a higher-utility product weakly decreases the expected demand for existing products.

Lemma 5. Consider assortment $S \subset \mathcal{M}$ such that $\sum_{l \in S} D_l(S) = \lambda n$. Under maximum choice rule, (i) given $i \notin S$, $D_l(S \cup \{i\}) \leq D_l(S)$ for all $l \in S$ (ii) given $j \in S$ and $i \notin S$ such that $U_i \geq U_j$, $D_l(S \setminus \{j\} \cup \{i\}) \leq D_l(S)$ for all $l \in S, l \neq j$.

2.4.3 Assortment optimization

In this section we present properties of the optimal assortment under maximum and probabilistic choice rules. The first set of results apply to both choice rules then we present more advanced results which are specific to each rule.

In the previous section we saw that adding a new product and replacing a product with a more popular product alternative always increases total demand (Lemmas 3 and 22). It does not, however necessarily increase total profits for the retailer as profit depends on the breakdown of demand between the products due to the concave inventory costs. As a result, the optimal assortment may not have a simple structure.

We say that assortment $S \subseteq \mathcal{M}$ is a ***popular assortment*** if it contains a certain number of the products with the highest popularity value, i.e., if it is of the type $\{1, \cdots, j\}$ for $j \in \{1, ..., M\}$ (such sets are also considered in **?**,Cachon et al. (2005) and Bernstein et al. (2011)). Let $\mathcal{P}$ denote the set of popular assortments, that is $\mathcal{P} = \{\{1, ..., j\}, j = 1, ..., M\}$.

But first we consider the case where the retailer's objective is to maximize total demand, a.k.a. total market share, i.e., $\sum_{j=1}^{M} D_j(S)$. Note that this case is equivalent to setting $\eta = 1$

in (2.1), since, in that case, $\Pi(S) = (p - c - \sigma) \sum_{j \in S} D_j(S)$, that is, total profit is linear in total demand.

Proposition 2. Suppose the retailer wishes to maximize total market share. Under the absence of constraint on the size of the assortment, it is optimal to offer every possible product alternative, i.e., $S^* = \{1, \ldots, M\}$. If the size of the assortment is constrained to be at most $m \in \{1, \ldots, M\}$, that is $|S| \leq m$, then the optimal assortment is the popular set $\{1, ..., m\}$.

Proof. The first part of the result follows directly from Lemma 3 as adding extra product alternatives always increases total market share, hence $S^* = \{1, \ldots, M\}$ must be optimal. The second part follows from Lemma 3 because the total demand can be increased by replacing a lower-utility product with a higher-utility product. $\square$

In contrast, when the retailer has strictly concave inventory costs, that is $\eta < 1$, the optimal assortment usually has a different structure. In particular, it may be optimal not to offer all possible products and the optimal assortment may not be a popular set, as examples 2.4.2 and 2.4.3 in the following two sub-sections illustrate.

Our next result is to show that it is always optimal to offer the most popular product, i.e., product 1, in the optimal assortment.

Proposition 3. If the product category is profitable, there always exists an optimal assortment S^* which includes the most popular product, $\{1\} \subseteq S^*$.

We prove the result by showing that, if the assortment does not include the most popular product, replacing its highest utility product with product 1 increases total expected demand and reduce inventory costs as demand gets distributed between the products in a more uneven fashion. Note, however, that adding product 1 to an existing assortment may not always increase the retailer's profit as it may increase inventory costs despite an increase in total expected demand.

Assortment optimization under the probabilistic choice rule

In order to establish the structure of the optimal assortment we first provide a result on the retailer's incremental profit from adding a new product to the assortment. Let $\Pi(S \cup \{i\})$ be the profit of the retailer obtained by adding product $i \in \mathcal{M} \setminus S$ to assortment S.

Lemma 6. Under the probabilistic choice rule, $\Pi(S \cup \{i\})$ is quasiconvex in U_i for all $n \in \mathbb{N}^+$.

Note that Lemma 6 is similar to Lemma 1 from van Ryzin and Mahajan (1999).

We present the following definition which is based on the terminology used by Alptekinoğlu and Grasas (2014) in the context of assortment planning with consumer returns.

Definition 1. We say that assortment $S \subseteq \mathcal{M}$ is a **popular-eccentric assortment** if it contains a certain number of the products with the highest utilities and a certain number of the products with the lowest utilities. Let $\mathcal{E}$ denote the set of popular-eccentric assortments, that is $\mathcal{E} = \{S : S = \{1, \ldots, h, M-r, \ldots, M-1, M\}$ where $0 \leq h \leq M, \ -1 \leq r \leq M-h\}$.

Note that, when $r = -1$, $S = \{1, ..., h\}$ which is a popular assortment; in other words, popular assortments are special cases of popular-eccentric assortments by our definitions. We say a set is a **strict popular-eccentric** assortment if it is popular-eccentric but not popular. For example, if $M = 4$, $\{1, 2\}$ is a popular assortment and therefore also a popular-eccentric assortment; $\{1, 4\}$ and $\{3, 4\}$ are strict popular-eccentric assortments and $\{1, 3\}$ is neither.

Proposition 4. Under the probabilistic choice rule, there always exists an optimal assortment S^* which is a popular-eccentric set. If $n = 1$, there exists an optimal assortment which is a popular set.

Practically speaking, Proposition 4 speeds up the search for the optimal assortment as one only needs to consider popular-eccentric sets of which there are $\frac{M^2+M+2}{2}$. Compared to the total of 2^M assortments, this leads to very significant time savings when M is large.

In the special case where all consumers buy at most one unit of product, that is, when $n = 1$, our model is mathematically equivalent to the MNL model for which **?** prove the optimality of popular sets. The reason why the optimal assortment may not be an popular set when consumers buy packs of size larger than 1 is because adding a product with zero expected utility may impact the demand for existing products as we discussed in the previous section. As a result the retailer's profit may increase when a zero-expected utility product is added to the assortment, i.e. we may have $\Pi(S \cup \{i\}) \geq \Pi(S)$ even if $U_i = 0$. Given that the profit function is quasiconvex in the expected utility of a newly added product, it is therefore possible that the highest increase in expected profit is obtained when adding the product with the lowest expected utility. In contrast, in the MNL model, we always have $\Pi(S \cup \{i\}) = \Pi(S)$ if $U_i = 0$ so either it is optimal to not add any product to the assortment, or it is optimal to add the product with the expected utility.

Other papers have identified settings wherein popular sets fail to be optimal: Alptekinoğlu and Grasas (2014) show that there can exists a popular-eccentric set which is optimal when retailer considers customer returns. Wang and Sahin (2017) numerically show that the optimal assortment can consist of only the least popular products when consumers have a search cost and make choices according to a two-stage consider-then-choose policy. **?** also provide an numerical example where the optimal assortment includes several most popular and least popular products when the cardinality of the assortment is constrained.

We now provide an example where the optimal assortment under probabilistic choice rule is a popular-eccentric set but not a popular set.

Example 2.4.2. Let $M = 3$, $n = 8$, $\lambda = 100$, $p - c = 5$, $\sigma = 4.5$, $\eta = 0.9$, $U = (3, 0.6, 0.1), V_0 = 2$. The optimal assortment under the probabilistic choice rule $S^* = \{1, 3\}$, which is a popular-eccentric set but not a popular set. Table 2.6 compares total demand and profits for three of the possible assortments.

Table 2.6. Comparison of assortments under probabilistic choice rule.

assortment S	Product demand Prod 1	Prod 2	Prod 3	Total demand	Total profits
$\{1\}$	738.46	0	0	738.46	1975.56
$\{1,2\}$	487.90	301.76	0	789.66	1998.85
$\{1,3\}$	521.41	0	266.88	788.29	1999.48

Comparing assortments $\{1,2\}$ and $\{1,3\}$, we see that $\{1,2\}$ yields a higher total demand but lower profits. This is because, due to the concave inventory costs, profits is generally higher when the demand for products is more "uneven". When product 3 is added to an assortment with only product 1, it "steals" less of the demand from product 1 than product 2 would, resulting a more uneven expected demand vector and therefore higher total expected profits.

Proposition 5. Under the probabilistic choice rule, if the size of the assortment is constrained to be equal to $m \in \{1,\ldots,M\}$, that is $|S| = m$, then there exists an optimal assortment which is a popular-eccentric set.

Proof. This result follows directly from Lemma 6. □

Assortment optimization under the maximum choice rule

We start by showing with an example that the optimal assortment under maximum choice rule may not be an popular-eccentric set.

Example 2.4.3. Let $M = 4$, $n = 4$, $\lambda = 10$, $p - c = 10$, $\sigma = 9$, $\eta = 0.5$, $U = (1.6, 1.5, 1, 0.7), V_0 = 7$. The optimal assortment under the maximum choice rule $S^* = \{1,3\}$, which is not a popular-eccentric set. Table 2.7 compares total demand and profits for four of the possible assortments.

Offering only product 1 leads to zero profits as the best 4-pack does not provide more expected utility than the outside option. When the assortment is $\{1,2\}$, the expected-utility-maximizing 4-pack contains 2 units of each product so that consumer demand is evenly split

Table 2.7. Comparison of assortments under maximum choice rule in Example 2.4.3.

Assortment S	$\vec{k}^{*4}(S)$	$V^{*4}(S)$	Product demand				Total demand	Total profits
			Prod 1	Prod 2	Prod 3	Prod 4		
$\{1\}$	$(4,0,0,0)$	6.4	0	0	0	0	0	0
$\{1,2\}$	$(2,2,0,0)$	7.99	20	20	0	0	40	319.50
$\{1,3\}$	$(3,0,1,0)$	7.19	30	0	10	0	40	322.24
$\{1,4\}$	$(3,0,0,1)$	6.89	0	0	0	0	0	0

between the two products and equal to 40 in total. In contrast, when the assortment is $\{1,3\}$, the expected-utility-maximizing 4-packs contains 3 units of product 1 and one unit of product 3, so that the total demand of 40 consumers now divides into 30 for product 1 and 10 for product 2. This uneven split leads to higher profit because of the concave inventory cost. Finally assortment $\{1,4\}$ also leads to zero profits because the best 4-pack, which includes 3 units of product 1 and one unit of product 4, does not provide more expected utility than the outside option.

So in general the optimal assortment under maximum choice rule does not have any specific structure other than the fact that it must include the most popular product by Proposition 3 and that the variety is not larger than n.

Lemma 7. Under the maximum choice rule, there exists an optimal assortment which has a variety equal to or less than n, i.e. $|S^*| \leq n$.

Proof. Let $S \subseteq \mathcal{M}$ be an assortment such that $|S| > n$. Let S' be such that for all $l \in S'$ we have $k_l^{*n}(S) \geq 1$. Thus we have: $S' \subset S$, $|S'| \leq n$ and $\vec{k}^{*n}(S) = \vec{k}^{*n}(S')$. Further we have $D_l(S') = \lambda \mathbb{1}(V(k^{*n}(S') \geq V_0))k_l^{*n}(S') = \lambda \mathbb{1}(V(k^{*n}(S) \geq V_0))k_l^{*n}(S) = D_l(S)$ for all $l \in S'$, and $D_l(S) = 0$ for all $l \in S \setminus S'$, which implies $\Pi(S') = \Pi(S)$. Therefore, for all assortment S which has a variety larger than n, there always exists an assortment with variety less than n and generating same expected profit as S. $\square$

The following result provides a sufficient condition for the optimal assortment under maximum choice rule to be composed only of the most popular product. When this condition is satisfied, all consumers buy the product so that the entire market is captured with only one product.

Proposition 6. If $nU_1 > V_0$ then offering only product 1 is optimal i.e., $S^* = \{1\}$ under the maximum choice rule.

2.5 The Importance of the Choice Premium

As argued by Fox et al. (2017), an important driver of consumer choice is the so-called *choice premium* in (2.2) which comes from buying distinct product alternatives in an n-pack. As Fox et al. (2017) point out: "*The logarithmic terms* $\ln(n!) - \ln((k_1)!(k_2)!\ldots(k_M)!)$ *together reflects the additional expected utility of having the freedom to consume products in whatever order one chooses [...] The term* $\ln(n!)$ *captures the effect of an n-pack's size while the term* $-\ln((k_1)!(k_2)!\ldots(k_M)!)$ *captures the effects of both variety and inventory.*" In this section we study what happens when the retailer ignores this choice premium, that is, assumes that consumer choice is only driven by the utility parameters of the products. Mathematically, the retailer assumes that the expected utility received by the consumer from consuming the n-pack $\vec{k} = (k_1, \ldots, k_M)$ is given by V^I which is equal to:

$$V^I(k_1, ..., k_M) \quad = \quad k_1U_1 + ... + k_MU_M \tag{2.5}$$

Compared to (2.2), the first term, i.e., the choice premium $\ln\left(\left(\sum_{j=1}^{M} k_j\right)!\right) - \ln((k_1)!(k_2)!\ldots(k_M)!)$, is missing.

In the absence of choice premium, having distinct products in an n-pack does not bring additional value to the consumer compared to having only one kind, thus the expected utility of an n-pack consumer is maximized from the pack which includes n units of the highest utility product alternative in S, as stated formally in the Lemma 8:

Lemma 8. In the absence of choice premium, the maximum-value n-pack $\vec{k}^{*n}(S)$ for all $n \in \mathbb{N}^+$ and $S \subseteq \mathcal{M}$ is such that $k^{*n}_{l_{max}} = n$ for $l_{max} = \arg\max_{i \in S} U_i$ and $k^{*n}_l = 0$ for $l \neq l_{max}$. Further, $V^{*n}(S) = nU_{l_{max}}$.

2.5.1 Ignoring the choice premium under maximum choice rule

Using Lemma 8, we show that, in the absence of choice premium, any assortment which includes a product that provides more utility than the outside option when consumed in a pure n-pack is optimal under maximum choice rule.

Proposition 7. In the absence of choice premium, under maximum choice rule, any S such that there exists $j \in S$ with $nU_j \geq V_0$ is optimal. If $nU_1 < V_0$, then offering nothing is optimal.

Proof. Under maximum choice rule, either all consumers buy an n-pack or none of them do. As a result the total demand for a given assortment is either $n\lambda \equiv \overline{D}$ or 0. Consider assortment S such that there exists $j \in S$ with $nU_j \geq V_0$. Let $l_{max} = arg\max\limits_{l \in S} U_l$. According to Lemma 8, the total demand for assortment S is $\sum_{l \in S} D_l = D_{l_{max}}(S) = \lambda n \mathbb{1}(nU_{l_{max}} \geq V_0) = n\lambda$. Then we have $\vec{D}(S) = (0, \ldots, \overline{D}, \ldots, 0) \succeq \vec{D}(S')$ for all $S' \subseteq \mathcal{M}$. Therefore, from Lemma 23, S must achieve maximum profit.

On the other hand, if $nU_l \leq V_0$ for all $l \in \mathcal{M}$, then the total demand $\sum_{l \in S} D_l(S) = D_{l_{max}}(S) = \lambda n \mathbb{1}(nU_{l_{max}} \geq V_0) = 0$ for all $S \subseteq \mathcal{M}$ by Lemma 8, which implies it is optimal to offer nothing. $\square$

In particular, when $nU_1 \geq V_0$, offering only the most popular product, i.e., $S = \{1\}$, achieves maximum expected profit for the retailer.

The following result shows that ignoring the choice premium when consumers choose according to the maximum choice rule can lead to a 100% optimality gap, calculated as $OG = \frac{\Pi(S^*) - \Pi(S^*_I)}{\Pi(S^*)}$ when $\Pi(S^*) > 0$, where S^* and S^{*I} denote an optimal assortment with and without choice premium respectively, and the profit Π is calculated assuming maximum choice rule with choice premium.

Proposition 8. Under maximum choice rule, the optimality gap from ignoring the choice premium is 0 if $nU_1 \geq V_0$, otherwise it is either 0 or 1.

Proof. By Proposition 6, we know that, if $nU_1 \geq V_0$ then $S^* = \{1\}$ is optimal with choice premium. By Proposition 7 it is also optimal without choice premium. Hence, in this case, $OG = 0$. On the other hand, if $nU_1 < V_0$, then offering $\{1\}$ leads to zero profit because the expected utility of the n-pack $(n, 0, ..., 0)$ is less than that of the outside option. Yet, due to the choice premium, it is possible that $V^{*n}(S^*) \geq V_0 > nU_1$. In this case, the optimal expected profit with choice premium is positive, leading to a 100% optimality gap; otherwise the optimal expected profit is zero and so is the optimality gap. □

Below is an example where the optimality gap from ignoring the choice premium under the maximum choice rule is equal to 100%.

Example 2.5.1. Let $\lambda = 10$, $n = 3$, $p - c = 10$, $\sigma = 8$, $\eta = 0.5$, $M = 2$, $U = (1.5, 1)$ and $V_0 = 5$. Table 2.8 compares the expected profit from each possible assortment when the retailer considers and ignores the choice premium.

Table 2.8. Maximum-value 3-Pack and Expected Profit ($V_0 = 5$)

	Considering Choice Premium			Ignoring Choice Premium		
	Maximum-value 3-Pack	Value	Profit	Maximum-value 3-Pack	Value	Profit
$\{1\}$	(3,0)	4.5	0	(3,0)	4.5	0
$\{2\}$	(0,3)	3	0	(0,3)	3	0
$\{1,2\}$	(2,1)	5.1	238.92	(3,0)	4.5	0

Because we have $nU_1 = 4.5 < V_0 = 5$, the retailer who ignores choice premium decides not to offer any product as she believes that no customer can get more expected value than with the outside option. However, if the retailer considers choice premium, she realizes that consumers will buy the $(2,1)$ 3-pack from assortment $\{1,2\}$ because, thanks to the choice premium, it has a value of 5.1 greater than that of outside option. As a result, the optimality gap which results from ignoring the choice premium is 100%.

In other words, under maximum choice rule, it is possible that ignoring the choice premium means underestimating consumers' expected utility and concluding that it is impossible to make money when in fact the product category is profitable if one considers the choice premium.

2.5.2 Ignoring the choice premium under probabilistic choice rule

Under the probabilistic choice rule, we are able to show that the optimal assortment for a retailer who ignores the consumers' choice premium is a popular-eccentric assortment, as is the case with choice premium (as shown in Proposition 4).

Proposition 9. In the absence of choice premium, under the probabilistic choice rule, there exists an optimal assortment which is a popular-eccentric assortment and includes product 1.

Proposition 9 implies that under probabilistic rule the optimal assortment without choice premium has the same structure as the optimal assortment with choice premium. However, the following example shows that optimal assortment in the with or without choice premium may not be the same.

Example 2.5.2. Let $\lambda = 10$, $n = 8$, $p - c = 20$, $\sigma = 19$, $\eta = 0.9$, $M = 3$ and $U = (1.5, 1, 0)$.

Table 2.9 lists the optimal assortments for $V_0 \in \{1.5, 1.7, 8.9\}$. As is shown, it is optimal to offer assortment $\{1,2\}$ across all the three values of V_0 when the retailer considers the choice premium. However the optimal assortment varies with V_0 when the choice premium is ignored: the retailer chooses to offer a smaller assortment (i.e., $\{1\}$) when $V_0 = 1.5$ and a larger size assortment (i.e., $\{1, 2, 3\}$) when $V_0 = 8.9$. Interestingly, when $V_0 = 1.7$, the (unique) optimal assortments in the two cases have the same size but consist of different products: the one when choice premium ignored is the popular-eccentric assortment $\{1, 3\}$.

Table 2.9. Optimal Assortment Comparison under Probabilistic Choice Rule in Example 2.5.2

	Optimal assortment	
V_0	Considering Choice Premium	Ignoring Choice Premium
1.5	$\{1, 2\}$	$\{1\}$
1.7	$\{1, 2\}$	$\{1, 3\}$
8.9	$\{1, 2\}$	$\{1, 2, 3\}$

Example 2.5.2 shows that the optimal assortment when the retailer ignores the choice premium can be either smaller, the same, or larger than the optimal assortment when she considers it.

In our numerical study below we study the optimality gap from ignoring the choice premium under both maximum and probabilistic choice rules.

2.6 Numerical Analysis

The purpose of our numerical study is fourfold. First we study properties of the optimal assortments under both choice rules as a function of the parameters of our model. Second, we propose simple heuristics for the assortment planning problem and investigate their performance, as measured by their optimality gap and computational speed. Third, we measure the value of the consumer choice premium. Fourth, we measure the value of recognizing that consumers buy multiple items at the same time.

2.6.1 Properties of the optimal assortment

To analyze how the optimal assortment varies with the parameters of our model, we conduct a numerical study using 54000 numerical instances which were obtained by varying the model parameters as follows: $\lambda \in \{8, 10, 12\}$, $n \in \{1, 3, 5, 7, 9\}$, $p - c \in \{10, 20\}$, $\eta \in \{0.5.0.75, 0.95, 1.25, 1.5\}$, $\sigma = \beta(p - c)$, where $\beta \in \{0.3, 0.5, 0.7, 0.9\}$. In this numerical study, we fix $M = 6$ and generate 60 utility vectors such that $U_1 + \ldots + U_6 = \frac{A}{B}$, where $A = 12$, $B \in \{1, 10, 100\}$. The idea of using integer A and B to create utility vectors is to vary the distances between utility values and the magnitude of utility levels. Let $\mathbb{A}_i$ be a breakdown of 12 (the value of A) between the 6 products. We consider the following 6 different breakdowns:

$$\mathbb{A}_1 = (11.5, 0.2, 0.15, 0.1, 0.05, 0)$$

$$\mathbb{A}_2 = (11.05, 0.37, 0.26, 0.17, 0.1, 0.05)$$

$$\mathbb{A}_3 = (9.5, 0.9, 0.7, 0.5, 0.3, 0.1)$$

$$\mathbb{A}_4 = (8, 2, 0.9, 0.65, 0.35, 0.1)$$

$$\mathbb{A}_5 = (6, 3, 2, 0.6, 0.3, 0.1)$$

$$\mathbb{A}_6 = (2.25, 2.15, 2.05, 1.95, 1.85, 1.75)$$

For example, when i=3, j=2, the utility vector is $(0.95, 0.09, 0.07, 0.05, 0.03, 0.01)$. As i increases, the distances between product utilities generally become smaller. For each utility vector such that $U_1 + \ldots + U_6 = \frac{\mathbb{A}_i}{B_j}$, a series of the utility of outside option is generated as $\alpha \frac{A}{B_j}$ where $\alpha \in \{0.25, 0.5, 0.75, 1, 1.25\}$.

While we found the optimal assortment to be unique in all the numerical instances we considered under probabilistic choice rule, it is possible that more than one assortments achieve the maximum expected total profit value under maximum choice rule. Among these, some assortments may include products which receive zero demand, which can be removed to obtain an optimal assortment with smaller variety.

In our analysis we report on the *optimal assortment type* and *optimal variety* for each problem instance. The *optimal assortment type* is either (i) a popular assortment (ii) a strict popular-eccentric assortment (i.e., a popular-eccentric set which is not a popular set) or (iii) a non-popular-eccentric assortment.[3] The *optimal assortment variety* for a given problem instance is the cardinality of the smallest optimal assortment of its type.

Table 2.10 summarizes the average optimal variety for different pack size values n under the probabilistic and maximum choice rule. Considering that when $n = 1$, the optimal assortment under both choice rules have extra structure, we compare the average pack sizes for all the cases where $n > 1$. We can see that, under probabilistic rule, when consumers' desired pack size $n > 1$, the average optimal variety decreases as n increases. When customers want to purchase bigger size n-pack, there is a larger number of n-packs available to buy, which weakens the attraction of the outside option. Therefore, with large n, it is more likely for the retailer to achieve optimality by offering a small number of products in the assortment. However, the optimal variety does not show strict monotonicity in n under the maximum choice rule in this numerical study.

Table 2.10. Average Optimal Variety under Different Pack Size

n	Probabilistic Choice Rule	Maximum Choice Rule
1	3.82	1.00
3	4.67	1.61
5	4.18	2.05
7	3.81	2.18
9	3.51	2.04

Table 2.11 shows how average optimal variety changes with the utility vector. We can see that the average size of optimal assortment gets larger as the utility values of candidate products get closer under both choice rules. This is because when there are a few very

[3] When there are multiple optimal assortments of different types for a given problem instance, we classify it using the most restrictive type, that is, if there exists an optimal popular assortment, we will classify the instance as type (i). If there exists a popular-eccentric optimal assortment but no popular optimal assortment, we classify the instance as type (ii). Instances classified as type (iii) must be so that there does not exist an optimal assortment which is a popular-eccentric set.

popular products, it is more likely to achieve optimal profit by including smaller number of products in assortment. Under probabilistic choice rule, the optimal variety decreases in the magnitude of product utilities.

We also reports the average optimal variety as a function of other parameters in Table 2.12. We observe that, as the outside option becomes more attractive, the average optimal variety under probabilistic choice rule increases, and the average optimal variety under maximum choice rule roughly has an upward trend although it does not exhibit strict monotonicity. This positive relationship is due to the fact that the retailer needs to offer larger variety in his assortment in capture enough consumer demand when the outside option is more attractive. It is also shown that the the average optimal variety decreases as the inventory multiplier increases under both choice rules.

Table 2.11. Average Optimal Variety under Product Utility Parameter Values

$\mathbb{A}$	Prob.	Max.	B	Prob.	Max.
$\mathbb{A}_1$	3.67	1.57	1	4.50	1.17
$\mathbb{A}_2$	3.79	1.57	10	4.12	2.06
$\mathbb{A}_3$	3.92	1.66	100	3.39	2.00
$\mathbb{A}_4$	4.00	1.85			
$\mathbb{A}_5$	4.09	1.96			
$\mathbb{A}_6$	4.52	2.86			

Table 2.12. Average Optimal Variety under Other Parameter Values

λ	Prob.	Max.	α	Prob.	Max.	$\frac{\sigma}{p-c}$	Prob.	Max.	η	Prob.	Max.
8	4.01	1.93	0.25	3.38	1.54	0.3	4.90	2.20	0.5	3.25	1.50
10	4.00	1.83	0.5	3.82	1.49	0.5	4.35	1.99	0.75	3.24	1.50
12	4.01	1.78	0.75	4.10	1.75	0.7	3.47	1.50	0.95	4.02	1.50
			1	4.28	2.77	0.9	2.75	1.50	1.25	5.96	3.94
			1.25	4.43	2.75				1.5	5.89	3.61

In Tables 2.13 to 2.17 we report our results regarding the optimal assortment types as a function of the type of utility vector and other parameters. Remember that, we showed in

Section §2.4.3 that the optimal assortment under the probabilistic choice rule is a popular-eccentric assortment (i.e., either popular or strictly popular-eccentric) but the optimal assortment under maximum choice rule can be of any type.

Table 2.13. Structure of the Optimal Assortment under Different Utility Types

$\mathbb{A}$	Probabilistic Choice Rule			Maximum Choice Rule			
	Popular	strictly Popular-eccentric	Other	Popular	strictly Popular-eccentric	Non popular-eccentric	Other
$\mathbb{A}_1$	72.67%	0.87%	26.47%	51.40%	0.00%	1.49%	47.11%
$\mathbb{A}_2$	73.29%	0.67%	26.04%	50.69%	0.80%	1.49%	47.02%
$\mathbb{A}_3$	75.07%	0.09%	24.84%	51.22%	0.80%	1.16%	46.82%
$\mathbb{A}_4$	75.67%	0.00%	24.33%	43.40%	1.60%	0.89%	54.11%
$\mathbb{A}_5$	76.60%	0.00%	23.40%	39.67%	1.60%	1.09%	57.64%
$\mathbb{A}_6$	78.87%	0.00%	21.13%	34.09%	0.00%	0.00%	65.91%

Note: The percentage values in column named ”Other” refer to the instances where it is optimal to offer nothing.

Table 2.14. Structure of the Optimal Assortment under Different B Values

B	Probabilistic Choice Rule			Maximum Choice Rule			
	Popular	strictly Popular-eccentric	Other	Popular	strictly Popular-eccentric	Non popular-eccentric	Other
1	74.46%	0.81%	24.73%	29.30%	0.00%	0.14%	70.56%
10	75.71%	0.00%	24.29%	43.07%	2.40%	2.91%	51.62%
100	75.91%	0.00%	24.09%	62.87%	0.00%	0.00%	37.13%

Note: The percentage values in column named ”Other” refer to the instances where it is optimal to offer nothing.

From Table 2.13, we see that for most problem instances, there exists an optimal assortment which is a popular assortment. Under probabilistic rule, most of the exceptions occur under the utility vector type 1 and 2, where the utility of the most popular product is significantly higher than that of the other ones. And Table 2.17 shows that most of the popular-eccentric optimal assortments under probabilistic rule appear when η is close to 1 (and less than 1). Combining the analysis of Table 2.13 and 2.17, we find that the existence

Table 2.15. Structure of the Optimal Assortment under Different α Values

α	Probabilistic Choice Rule			Maximum Choice Rule			
	Popular	strictly Popular-eccentric	Other	Popular	strictly Popular-eccentric	Non popular-eccentric	Other
0.25	74.31%	0.48%	25.20%	65.13%	0.00%	1.83%	34.87%
0.5	74.83%	0.37%	24.80%	60.07%	0.67%	1.59%	39.93%
0.75	75.39%	0.20%	24.41%	50.46%	0.67%	1.67%	49.54%
1	75.70%	0.19%	24.11%	28.50%	2.67%	0.00%	71.50%
1.25	76.56%	0.11%	23.33%	21.22%	0.00%	0.00%	78.78%

Note: The percentage values in column named "Other" refer to the instances where it is optimal to offer nothing.

Table 2.16. Structure of the Optimal Assortment under Different ratio $\frac{\sigma}{p-c}$

$\frac{\sigma}{p-c}$	Probabilistic Choice Rule			Maximum Choice Rule			
	Popular	strictly Popular-eccentric	Other	Popular	strictly Popular-eccentric	Non popular-eccentric	Other
0.3	93.29%	0.07%	6.64%	54.84%	0.80%	2.46%	41.90%
0.5	82.04%	0.13%	17.82%	47.07%	0.80%	1.35%	50.79%
0.7	65.10%	0.41%	34.49%	39.20%	0.80%	0.13%	59.87%
0.9	61.01%	0.46%	38.53%	39.20%	0.80%	0.13%	59.87%

Note: The percentage values in column named "Other" refer to the instances where it is optimal to offer nothing.

of dominant products and large η are more likely to result in popular eccentric optimal assortment through generating significant impact from inventory cost. When there exists a few dominant products and when η is big, a popular eccentric assortment leads to similar sales profit as a popular assortment because most sales profit is brought by the very popular product(s) for both type of assortments, but the inventory cost of the popular eccentric assortment can be relatively much lower due to the pooling effect. We use Example 2.6.1 to illustrate this.

Example 2.6.1. Let $\lambda = 10$, $n = 5$, $p - c = 20$, $\sigma = 19$, $\eta = 0.9$, $M = 3$, and $V_0 = 0.8$. We consider two different utility vectors. With a equally distanced utility vector $U = (1.5, 1, 0)$, the optimal assortment is popular assortment $\{1, 2\}$ with profit 301.85. And with a utility

Table 2.17. Structure of the Optimal Assortment under Different η Value

η	Probabilistic Choice Rule			Maximum Choice Rule			
	Popular	strictly Popular-eccentric	Other	Popular	strictly Popular-eccentric	Non popular-eccentric	Other
0.5	100.00%	0.00%	0.00%	65.33%	1.33%	0.22%	33.11%
0.75	99.76%	0.24%	0.00%	65.33%	1.33%	0.22%	33.11%
0.95	98.89%	1.11%	0.00%	65.33%	1.33%	0.22%	33.11%
1.25	54.63%	0.00%	45.37%	24.39%	0.00%	3.69%	71.93%
1.5	23.52%	0.00%	76.48%	5.00%	0.00%	0.74%	94.26%

Note: The percentage values in column named "Other" refer to the instances where it is optimal to offer nothing.

vector that contains a dominant product $U = (1.5, 0.1, 0)$, the optimal assortment is popular-eccentric assortment $\{1, 3\}$ with profit 300.7. The demand and profit for assortment $\{1, 2\}$ and $\{1, 3\}$ under the two utility vectors are shown in Table 2.18. As is shown, when $U = (1.5, 1, 0.5)$, the difference between sales profit for popular assortment $\{1, 2\}$ and popular-eccentric assortment $\{1, 3\}$ is big: $974.21 - 961.93 = 12.28$. Although the inventory cost of popular-eccentric assortment $\{1, 3\}$ is lower than that for assortment $\{1, 2\}$ (with difference $672.35 - 661.23 = 11.13$), it is not significant enough to make up the gap in sales profit. While when $U = \{1.5, 0.1, 0\}$, the sales profit for popular assortment $\{1, 2\}$ is only a little higher than popular-eccentric assortment $\{1, 3\}$ (with difference $963.66 - 961.93 = 1.73$), the difference in inventory cost (which equals $663.09 - 661.23 = 1.86$) becomes relatively more significant such that the total profit of $\{1, 3\}$ exceeds the total profit for $\{1, 2\}$.

Table 2.18. Optimal assortment type comparison in Example 2.6.1

U	Assortment	D_1	D_2	D_3	Total Demand	Sales Profit $(p-c)\sum_{l \in S} D_l$	Inventory Cost $\sigma \sum_{l \in S} (D_l)^\eta$	Total Profit
$(1.5, 1, 0)$	$\{1, 2\}^*$	26.24	22.47	–	48.71	974.21	672.36	301.85
	$\{1, 3\}$	32.38	–	15.72	48.10	961.93	661.23	300.70
$(1.5, 0.1, 0)$	$\{1, 2\}$	31.51	16.68	–	48.18	963.66	663.09	300.58
	$\{1, 3\}^*$	32.38	–	15.72	48.10	961.93	661.23	300.70

2.6.2 Heuristics performance

Given our result that the optimal assortment under probabilistic choice rule is a popular-eccentric assortment which includes the most popular product, one needs to compare $\frac{M^2-M+4}{2}$ such assortments to identify the optimal one(s). In contrast, under maximum choice rule, there are 2^{M-1} possible assortments to consider since the only established property is that the optimal assortment must include the most popular product.

In this section, we study the performance of heuristics under both choice rules. Table 2.19 shows the set of heuristics we consider along with their complexity.

Table 2.19. Heuristics

Notation	Name	Description	Complexity
1Only	Most Popular only	$S = \{1\}$	$O(1)$
All	Offer all	$S = \{1, ..., M\}$	$O(1)$
Best_Pop	Best popular assortment	$S = \arg\max_{S' \in \mathcal{P}} \Pi(S')$	$O(M)$
Greedy_Pop_Ecc	Greedy Popular-Eccentric	See Algorithm 2	$O(M)$
Greedy	Greedy	Greedy Algorithm(*)	$O(M^2)$
Best_Pop_Ecc	Best popular-eccentric assortment	$S = \arg\max_{S' \in \mathcal{E}} \Pi(S')$	$O(M^2)$
Pop_Max	Popular-Maxiumum	See Algorithm 3	$O(M^2)$

(*) in each iteration the product which increases profit by the most is added, until all products have been added. The chosen assortment is the one with the highest profit.

Heuristics *1Only*, *All*, *Best_Pop*, *Greedy* and *Best_Pop_Ecc* are fully explained in Table 2.19. Heuristic *(Greedy_Pop_Ecc)* consists in adding products in a greedy fashion in each iteration but choosing only between the most and least popular products not yet included so that the chosen assortment is guaranteed to be a Popular-Eccentric assortment. It is implemented using Algorithm 2 below.

Algorithm 2 Greedy Popular-Eccentric

1: Let $S_1 = \{1\}$. Set $k = 1$.
2: While $k < M$, let i^+ be such that $U_{i^+} = \max_{j \notin S_k} U_j$ and i^- be such that $U_{i^-} = \min_{j \notin S_k} U_j$. If $\Pi(S \cup \{i^+\}) \geq \Pi(S \cup \{i^-\})$ then $S_{k+1} := S \cup \{i^+\}$, otherwise $S := S \cup \{i^-\}$. Set $k := k + 1$.
3: The chosen assortment is $S = \arg\max \{\Pi(S_k); k = 1, ..., M\}$.

Lemma 9. Under probabilistic choice rule, the *Greedy Popular-Eccentric* heuristic (Algorithm 1) is equivalent to the *Greedy* heuristic.

Proof. Let $S_1, \ldots, S_M$ be the assortments obtained from M iterations of the *Greedy Popular-Eccentric* heuristic, and $S'_1, \ldots, S'_M$ be the assortments obtained from M iterations of the *Greedy* heuristic.

Let i_k^+ be such that $U_{i_k^+} = \max_{l \in \mathcal{M} \setminus S_{k-1}} U_l$, and $U_{i_k^-} = \min_{l \in \mathcal{M} \setminus S_{k-1}} U_l$. From the definitions of the two heuristics, we have $S_k = \max_{l \in \{i_k^+, i_k^-\}} \Pi(S_{k-1} \cup \{l\})$, and $S'_k = \max_{l \in \mathcal{M} \setminus S'_{k-1}} \Pi(S'_{k-1} \cup \{l\})$ for $k = 2, \ldots, M$.

We prove that $S_k = S'_k$, $k = 1, \ldots, M$ by induction. After the first iteration, we have $S_1 = S'_1 = \{1\}$. Suppose $S_k = S'_k$ for some k, then according to Lemma 6, $\arg \max_{l \in \{i_{k+1}^+, i_{k+1}^-\}} \Pi(S_k \cup \{l\}) = \arg \max_{l \in \mathcal{M} \setminus S'_k} \Pi(S'_k \cup \{l\})$ which implies $S_{k+1} = S'_{k+1}$ and proves the result. $\square$

We also propose the *Popular-Maxiumum (Pop_Max)* heuristic, presented in Algorithm 3, which works as follows: we add products to the assortment in decreasing order of popularity as long as doing so does not generate a maximum-value n-pack that is more attractive than the outside option. If it does, we instead add the least popular product which would make all consumers buy an n-pack. This heuristic is tailored to the assortment problem under maximum choice rule. The idea is to obtain an assortment which captures the whole market with demands allocated among the products in an unbalanced fashion so as to lower the inventory cost.

Algorithm 3 Popular-Maximum

1: Let $S = \{1\}$. If $V^{*n}(S) \geq V_0$, Stop. Otherwise set $k = 2$ and go to Step 2.
2: If $V^{*n}(S \cup \{k\}) < V_0$, set $S := S \cup \{k\}$, $k = k + 1$ then repeat Step 2. Otherwise, let $j = \max\{l : V^{*n}(S \cup \{l\}) \geq V_0\}$ and set $S := S \cup \{j\}$, Stop.

We use all seven heuristics under maximum choice rule. Under probabilistic choice rule, the Best_Pop_Ecc heuristic is optimal by Proposition 4. Also we prove in Lemma 9 in the Appendix that the *Greedy Popular-Eccentric* and *Greedy* heuristics are equivalent. Therefore, we only consider the first four heuristics listed in Table 2.19 as they have lower complexity than an enumerative search for the optimal assortment.

To evaluate the performance of the heuristics, we use the same numerical study described in previous section. For each numerical instance, we compute the optimality gap of each heuristic as the percentage difference between the optimal expected profit and the expected profit obtained with the assortment chosen by the heuristic. For each heuristic, we report (i) the average computational time in seconds (CPU), (ii) the percentage of instances where the optimal expected profit was obtained (%Opt), (iii)the average optimality gap (Avg_OG), (iv)the standard deviation of optimality gap (Std_OG) and (v) the maximum optimality gap (Max_OG). All problem instances were solved on a 2.3 GHz Intel Core i5 Macbook running Matlab R2017a.

We show the mean and standard deviation of computational time (in 10^{-4} seconds) for the heuristics under different n values in Tables 2.20 and 2.21

Table 2.20. Computational Time under Probabilistic Rule for Different n ($\times 10^{-4}$ seconds)

n	Enumeration	1Only	All	Best_Pop	Greedy_Pop_Ecc
1	17.08 (6.39)	0.03 (0.03)	0.02 (0.02)	3.95 (1.89)	7.37 (3.33)
3	33.68 (10.42)	0.03 (0.01)	0.01 (0.01)	7.93 (1.40)	15.40 (2.23)
5	60.64 (5.41)	0.03 (0.00)	0.01 (0.00)	19.56 (1.29)	38.62 (1.93)
7	130.46 (17.00)	0.04 (0.01)	0.02 (0.01)	53.36 (6.52)	105.00 (13.62)
9	308.66 (47.60)	0.04 (0.03)	0.02 (0.01)	140.14 (25.2)	278.07 (39.17)
All n	103.33 (106.32)	0.04 (0.02)	0.02 (0.01)	41.85 (50.55)	82.67 (99.55)

Table 2.22 presents our results under probabilistic choice rule. The *Best Popular* and *Greedy Popular-Eccentric* algorithms achieve optimality under probabilistic choice rule in most of the instances in our numerical study.

Table 2.23 presents our results regarding the performance of the heuristics under maximum choice rule. The algorithm *Best Popular*, *Greedy Popular-Eccentric*, *Greedy*, *Best Popular Eccentric* and *Popular-Maximum* achieve optimality in most of the cases in our

Table 2.21. Computational Time under Maximum Rule for Different n ($\times 10^{-4}$ seconds)

n	Enumeration	1Only	All	Best_Pop
1	49.56 (22.32)	0.03 (0.05)	0.02 (0.01)	3.80 (2.06)
3	66.92 (8.79)	0.03 (0.01)	0.02 (0.01)	7.78 (1.07)
5	117.49 (5.43)	0.03 (0.01)	0.01 (0.00)	19.45 (1.28)
7	234.83 (26.13)	0.04 (0.01)	0.02 (0.01)	52.92 (9.48)
9	498.86 (76.75)	0.04 (0.02)	0.02 (0.03)	139.99 (24.98)
All n	209.54 (172.46)	0.03 (0.02)	0.02 (0.01)	49.38 (53.08)
n	Greedy_Pop_Ecc	Greedy_Pop_Ecc	Greedy	Best_Pop_Ecc
1	7.19 (2.85)	23.17 (6.44)	10.59 (5.13)	0.75 (0.51)
3	15.12 (1.83)	30.84 (2.89)	20.92 (2.27)	8.65 (11.63)
5	38.41 (2.11)	51.01 (2.63)	50.93 (4.00)	14.41 (22.17)
7	104.71 (16.83)	98.25 (11.08)	122.72 (13.74)	18.59 (26.96)
9	278.83 (43.11)	215.90 (31.42)	301.36 (47.11)	26.71 (46.37)
All n	98.00 (105.12)	90.58 (73.77)	111.41 (111.50)	15.12 (29.08)

Table 2.22. The Performance of Heuristics under Probabilistic Choice Rule

Heuristic	%Opt	Avg_OG	Std_OG	Max_OG
1Only	6.99%	48.51%	28.82%	100%
All	31.86%	4.50%	8.46%	100%
Best_Pop	99.72%	2.36×10^{-7}	5.48×10^{-6}	0.03%
Greedy_Pop_Ecc	99.76%	1.36×10^{-7}	3.34×10^{-6}	0.02%

numerical study, but they all can result in a profit loss up to 100%. Also, *Popular-Maximum* achieves less optimality than the other four algorithms mentioned above, but is faster than the other four heuristics as is shown in Table 2.21.

Table 2.23. The Performance of Heuristics under Maximum Choice Rule

Heuristic	%Opt	Avg_OG	Std_OG	Max_OG
1Only	56.97%	42.06%	48.85%	100%
All	40.11%	6.06%	10.35%	100%
Best_Pop	96.12%	0.96%	8.80%	100%
Greedy_Pop_Ecc	96.98%	0.95%	8.80%	100%
Greedy	97.26%	0.95%	8.80%	100%
Best_Pop_Ecc	97.83%	0.94%	8.80%	100%
Pop_Max 3	88.54%	7.82%	24.80%	100%

2.6.3 The Value of the Choice Premium

As discussed in §2.5 the expected value of an n-pack includes a term referred to as the choice premium which measure the benefits consumers receive from consuming a variety of different products over time. In this section we numerically estimate the optimality gap which results from the retailer ignoring this term, that is, assuming that consumer choice is solely driven by the nominal utilities of the products. We refer to this gap as *the value of the choice premium*. Formally, we calculate $OG = \frac{\Pi(S^*)-\Pi(S_I^*)}{\Pi(S^*)}$ where S^* and S^{*I} denote an optimal assortment with and without choice premium respectively, and the profit Π is calculated assuming maximum choice rule with choice premium. We report the average (*Avg_OG*), standard deviation (*Std_OG*) and maximum value (*Max_OG*) of this optimality gap as well as the percentage of instances where ignoring the choice premium did not cause any profit loss (*%Opt*).

We use the same numerical instances as in the previous section to evaluate this value and present our results in Table 2.24.

Table 2.24. The Value of the Choice Premium

Choice Rule	%Opt	Avg_OG	Std_OG	Max_OG
Probabilistic	61.13%	1.86%	10.35%	100%
Maximum	56.97%	42.06%	48.85%	100%

Under both choice rules, the maximum optimality gap is as high as 100%, which implies that, by ignoring choice premium, the retailer may think the product category is not profitable while in fact it is. Under probabilistic choice rule, the retailer can miss optimality in 38.87% of the instances if he ignores choice premium. Under the maximum choice rule, the retailer achieve optimality in only 56.97% of the examples, and the average profit loss is 42.06%.

Next we analyze how the value of choice premium varies with the parameters of the assortment problem under the probabilistic choice rule. As is discussed in §**??**, the relative

Table 2.25. Value of Choice Premium under Different Magnitude for Utility

B	Probabilistic Choice Rule				Maximum Choice Rule			
	%Opt	Avg_OG	Std_OG	Max_OG	%Opt	Avg_OG	Std_OG	Max_OG
1	90.76%	0.60%	7.66%	100%	94.72%	5.02%	21.60%	100%
10	55.28%	1.53%	10.75%	100%	55.37%	43.26%	48.82%	100%
100	37.59%	3.45%	11.95%	100%	40.51%	58.48%	48.73%	100%

value of choice premium to the direct value from consumption is expected to have significant impact on the optimal assortment when the retailer ignores choice premium. In Table 2.25, we report how the impact of choice premium changes in B. As is seen, the retailer loose much more profit from ignoring choice premium when B is large, that is when the relative value of choice premium is big comparing to the direct utility from consumption.

Table 2.26 exhibits the average optimality gap as a function of other model parameters under probabilistic choice rule. It is easy to show that when $n = 1$, ignoring choice premium does not have any impact on optimal assortment and profit because the choice premium value to consumer is zero in single-item shopping. Thus, we only consider the instances where consumers want to buy multiple items (i.e. $n > 1$) in Table 2.26. We find that, when product utility values are close to each others, the impact of choice premium is big. This happens because a few more or a few less products than the optimal assortment can lead to big loss in total profit due to the fact that the products in optimal assortment contribute almost equally to total demand and total profit. Also, average optimality gap increases in the utility of outside option.

2.6.4 The value of Considering the Consumers' Basket Shopping Behavior

The novelty of the n-pack choice model is that it captures the basket shopping behavior of customers, that is, it recognizes the fact that not only do customers often purchase more than one products from the same category on a unique shopping trip, they often buy multiple kinds of products within the category.

Table 2.26. Average Optimality Gap under Absence of Choice Premium (Probabilistic Choice Rule, $n > 1$)

$\mathbb{A}$	Avg_OG	n	Avg_OG	λ	Avg_OG	α	Avg_OG	$\frac{\sigma}{p-c}$	Avg_OG	η	Avg_OG
$\mathbb{A}_1$	3.35%	3	3.11%	8	2.99%	0.25	2.01%	0.3	1.77%	0.5	1.23%
$\mathbb{A}_2$	3.49%	5	2.19%	10	1.78%	0.5	2.45%	0.5	3.14%	0.75	1.47%
$\mathbb{A}_3$	2.70%	7	2.40%	12	2.43%	0.75	2.50%	0.7	2.81%	0.95	1.03%
$\mathbb{A}_4$	1.99%	9	1.88%			1	2.54%	0.9	2.02%	1.25	5.37%
$\mathbb{A}_5$	1.54%					1.25	2.55%			1.5	15.28%
$\mathbb{A}_6$	1.41%										

In the situation where retailer ignores the basket shopping behavior of consumers, he may understand the consumers' purchasing behavior in different ways. In this section, we explore the impact of basket shopping of consumers by comparing the optimal assortment and profit obtained in the following three ways. (i) The retailer assumes each consumer wants to buy at most one product, i.e., a 1-pack. From van Ryzin and Mahajan (1999), we know that the optimal assortment under the probabilistic rule is a popular assortment in this case. Under the maximum choice rule, it is optimal to offer only the most popular product if the utility of the product is not less than the utility of outside option. (ii) The retailer assumes each customers buys n units of the same product. (iii) The retailer assumes each consumer, rather than buying n products (an n-pack) all at once, visits the store n times and buys at most one product on each visit. It is easy to verify that under the optimal assortment under the probabilistic choice rule in case (ii) and (iii) are also popular assortment, but the optimal assortment obtained in the three scenarios above can be different. In addition, the optimal assortment under maximum choice rule in case (ii) and (iii) are to offer the most popular product if the product category is profitable. However, the total demand and optimal profit under maximum choice rule for all the three scenarios can be different.

We show a summary of the impact of ignoring basket shopping behavior in Table 2.27. As we can see, if the retailer assumes each consumer from the population only buys one item, the gap from optimality under probabilistic choice rule is not very large. However, if he assumes that each consumer buys n units of same products or that each consumer comes n times to buy one product on every shopping trip, the optimality gap can be big. Specifically, when retailer thinks that consumers comes to store before each of the n consumption scenarios to buy only one product, he misses optimality in 77.20% of the examples, and the average gap from optimal profit is as high as 37.30%. Also, under both choice rules, the maximum optimality gap can be 100%, which implies that the retailer can think of the product category to be unprofitable when ignoring basket shopping behaviors when actually the category is profitable.

Table 2.27. The Impact of the Basket Shopping

	Choice Rule	%Opt	Avg_OG	Std_OG	Max_OG
(i)	Probabilistic	99.72%	2.37×10^{-7}	5.48×10^{-6}	0.03%
(ii)		61.95%	8.84%	27.96%	100%
(iii)		22.71%	37.30%	29.04%	100%
(i)	Maximum	56.97%	42.06%	48.85%	100%
(ii)		56.16%	43.45%	49.44%	100%
(iii)		56.16%	43.45%	49.44%	100%

In addition, we show how the impact of basket shopping behavior varies with the model parameters under probabilistic choice rule in Table 2.28 to 2.30. It is obvious that when $n = 1$, the basket shopping behavior does not have any impact on their choice and retailer's assortment, thus in Table 2.28 to 2.30 we only analyze the instances where $n > 1$. In (i), the suboptimality appears only when $B = 1$, while in (ii) and (iii) the average gap from optimal profit increases in B. In (iii) average gap is large when the outside option is more attractive, while in (i) and (ii) the impact from basket shopping is note quite sensitive to the outside option. Also, in (ii) and (iii) the average optimality gap decreases in the power of inventory cost η when $\eta < 1$.

Table 2.28. Average Optimality Gap When Ignoring Basket Shopping Behavior (Method (i), Probabilistic Choice Rule, $n > 1$)

$\mathbb{A}$	Avg_OG	B	Avg_OG	n	Avg_OG	λ	Avg_OG
$\mathbb{A}_1$	9.36×10^{-7}	1	9.25×10^{-7}	3	8.91×10^{-8}	8	3.02×10^{-7}
$\mathbb{A}_2$	8.07×10^{-7}	10	0	5	5.11×10^{-8}	10	3.18×10^{-7}
$\mathbb{A}_3$	1.05×10^{-7}	100	0	7	3.51×10^{-7}	12	3.00×10^{-7}
$\mathbb{A}_4$	0			9	7.85×10^{-7}		
$\mathbb{A}_5$	0						
$\mathbb{A}_6$	0						

α	Avg_OG	$\frac{\sigma}{p-c}$	Avg_OG	η	Avg_OG
0.25	4.16×10^{-7}	0.3	1.18×10^{-8}	0.5	0
0.5	5.71×10^{-7}	0.5	2.19×10^{-8}	0.75	1.84×10^{-7}
0.75	9.74×10^{-8}	0.7	4.43×10^{-7}	0.95	9.35×10^{-7}
1	2.48×10^{-7}	0.9	9.93×10^{-7}	1.25	0
1.25	2.01×10^{-7}			1.5	0

Table 2.29. Average Optimality Gap When Ignoring Basket Shopping Behavior (Method (ii), Probabilistic Choice Rule, $n > 1$)

$\mathbb{A}$	Avg_OG	B	Avg_OG	n	Avg_OG	λ	Avg_OG	α	Avg_OG	$\frac{\sigma}{p-c}$	Avg_OG	η	Avg_OG
$\mathbb{A}_1$	10.98%	1	11.14%	3	12.19%	8	12.45%	0.25	11.22%	0.3	16.42%	0.5	0.66%
$\mathbb{A}_2$	11.21%	10	11.50%	5	13.45%	10	11.10%	0.5	11.40%	0.5	20.62%	0.75	0.43%
$\mathbb{A}_3$	11.41%	100	11.68%	7	9.84%	12	10.72%	0.75	11.34%	0.7	2.57%	0.95	0.06%
$\mathbb{A}_4$	11.35%			9	10.08%			1	11.41%	0.9	0.78%	1.25	50.28%
$\mathbb{A}_5$	11.47%							1.25	11.84%			1.5	99.66%
$\mathbb{A}_6$	12.21%												

Table 2.30. Average Optimality Gap When Ignoring Basket Shopping Behavior (Method (iii), Probabilistic Choice Rule, $n > 1$)

$\mathbb{A}$	Avg_OG	B	Avg_OG	n	Avg_OG	λ	Avg_OG	α	Avg_OG	$\frac{\sigma}{p-c}$	Avg_OG	η	Avg_OG
$\mathbb{A}_1$	46.27%	1	46.44%	3	46.31%	8	49.51%	0.25	30.65%	0.3	52.30%	0.5	41.31%
$\mathbb{A}_2$	47.04%	10	48.09%	5	49.45%	10	47.95%	0.5	42.20%	0.5	53.48%	0.75	40.74%
$\mathbb{A}_3$	48.71%	100	50.22%	7	49.33%	12	47.25%	0.75	50.26%	0.7	42.11%	0.95	39.94%
$\mathbb{A}_4$	49.20%			9	48.03%			1	56.59%	0.9	41.48%	1.25	79.24%
$\mathbb{A}_5$	49.13%							1.25	61.48%			1.5	95.39%
$\mathbb{A}_6$	49.18%												

CHAPTER 3

ASSORTMENT PLANNING FOR A TWO-SIDED MARKET

3.1 Introduction

Many of today's successful businesses involve two distinct groups of users and as such, are referred to as facing a *two-sided market*. For example, the video game industry brings together players who buy the games and video game developers who design them. Credit cards link consumers who use them to pay for their purchases and the merchants who provide this payment option in their store. Typically there is a useful synergy, called the *network effect*, between the two groups of users in a two-sided market as the value of the product or service grows with the size of the two user groups. For example, as more people play with a video game console, more developers are interested in designing games for it, which in turns, leads to more demand from players. The provider of the product or service in a two-sided market typically receives revenues from both groups of users: credit card companies receive interest payments from consumers and swipe fees from merchants. Websites such as Care.com receive subscription payments from individuals looking for a care provider (e.g., nanny or babysitter) and from the individuals who are looking to be hired.

Numerous studies have shown the importance of the pricing dimension in a two-sided market (Rochet and Tirole (2006), Hagiu (2006), Weyl (2010) and Hagiu and Wright (2014)). In some settings it may be optimal to subsidize one group of users, and possibly even give them free access to the product or service, while collecting revenues only from the second group of users. For example, search engine developers provide free access to their website to web-users and collect revenues solely from advertisers.

In some contexts the company offers a portfolio of products or services to its two-sided market: for example, a television network, such as Viacom, owns and operates a large number of television channels which are watched by viewers and on which advertisers promote their brands. The television network collects advertising revenues from advertisers and possibly subscription fees from viewers, based on the type of channel (regular network, cable or premium cable channel). Viewers typically watch only one channel at a time, hence, at any given moment in time, they make a choice from the channel portfolio. Advertisers typically are looking to reach viewers only from specific demographics, such as "women with disposable income aged 18 to 35" (the key demographic for reality television). Offering the right set of channels is key to the success of a television network as the right product offering will bring viewers from the advertisers' key demographics, which in turns will bring more advertising revenues.

The main contribution of our paper to the literature on two-sided markets is that we focus on the assortment question, that is, we study how to select the best set of products or services to offer in order to maximize revenues from both groups of users. At the same time, we contribute to the literature on assortment planning by adding the second market dimension to the product selection problem with multiple customer segments.

The assortment problem for a two-sided market with multiple customer segments is complex in that there is no simple structure to the optimal assortment. In particular, we show that a revenue-ordered solution, adapted to the two-sided revenue stream of our problem, is not necessarily optimal for the problem in its most general form; however we provide conditions under which the optimal assortment has a simple ranking-based structure. Further, we show that focusing only on one group of users, that is treating the problem as a one-sided rather than two-sided market, can lead to a significant loss in revenue.

We also investigate a special case wherein the possible products to offer either appeal to all customer segments equally or are tailored to the tastes of one of them. This question is inspired by the television network example mentioned above and the realization that there exist mainly two types of television channels: the ones which appeal to a general audience, such as the main network channels (NBC, ABC, CBS) and those which are dedicated to a specific demographic, such as many of the cable and premium channels (Cartoon Network, MTV, the Golf channel, etc.). We provide conditions under which it is optimal to diversify the product offering, that is, offer a portfolio of products each dedicated to a specific customer segment and conditions under which it is better to offer a homogenous product offering, that is, offer products which appeal to all customer segments.

The rest of this chapter is organized as follows. In §3.3 we formulate our model and discuss some of its special cases. In §3.4, we give an numerical example to illustrate some properties of the optimal assortment. Our results are presented in §3.5. Unless otherwise stated, all proofs are in the Appendix B.

3.2 Literature Review

Our paper is related to two main streams of research: the literature on assortment planning in operations management journals and the literature on two-sided markets, mostly in marketing journals. In this section we provide a brief review of the most relevant papers in each area.

Because we use an attraction-type model for modeling customer choice, our paper is most related to the papers in the assortment planning literature which consider similar models such as the multinomial logit (MNL) model. Therefore we focus our survey on such paper (for studies on assortment planning using other customer choice models, we refer the reader to Honhon et al. (2010), Honhon et al. (2012), Farias et al. (2012) and Farias et al. (2013)). In their seminal paper, van Ryzin and Mahajan (1999) consider the joint assortment and inventory decisions of a retailer offering differentiated products at the same price in a single-period setting. They show that the optimal assortment contains a certain number of the most popular products. Li (2007) extends their work to products with different profit margins and shows that the optimal assortment contains a certain number of the most profitable products.

Talluri and Van Ryzin (2004) study a single-leg airline seat allocation problem where customers make a choice according to the MNL model. They show that if the parameters of the MNL model are deterministic and known, then the optimal assortment is *revenue-ordered*, that is, it includes a certain number of the products with the highest revenue. Gallego et al. (2011) show that this problem can be formulated and solved as a linear program with linear number of variables and constraints. Rusmevichientong and Topaloglu (2012) propose a robust formulation of the same problem where the true parameters of the MNL model are assumed to be unknown and show that revenue-ordered assortments are still optimal.

Our paper is also related to the work on assortment planning problem with cardinality constraints, that is, when the size of the assortment is limited. Rusmevichientong et al. (2010) extend the work of Talluri and Talluri and Van Ryzin (2004) to include a cardinality constraint on the offered assortment and show that although revenue-ordered assortments are no longer optimal, the optimal assortment can be computed efficiently. Farias et al. (2010) propose a pairwise exchange algorithm for assortment planning with cardinality constraint and show that the algorithm can find an optimal solution when the parameters of MNL model are known in advance. Davis et al. (2013) optimize the assortment subject to a set of totally unimodular constraints and apply this to the assortment planning problem under the MNL choice model with various bounds on the cardinality of the assortment. Gallego and Topaloğlu (2014) study assortment optimization problems where customers make choices according to the nested logit model and there are constraints on the offered variants in each nest. They show that the optimal assortment under cardinality constraints can be found efficiently by solving a linear program.

We develop our model in the context of multiple customer segments and customers of different segments make choices according to different MNL models. This is the so called *mixture of logit* model which can avoid the independence of irrelevant alternatives property of traditional MNL model. Bront et al (2009) consider the assortment problem with multiple customer segments and random model parameters which is referred to as the *mixture of logit* model. They show that the problem NP-hard and propose a mixed integer program formulation of the problem as well as a greedy heuristic. Rusmevichientong et al. (2014) study the same problem and prove that the problem is NP-complete even when there are only 2 customer segments. They show that revenue-ordered optimal assortments are not optimal in general but provide two special cases in which they are. Méndez-Déaz et al. (2010) propose a branch-and-cut algorithm for the assortment problem under a mixture of multinomial logit models which is shown to perform well in the presence of cardinality constraints.

The theory of two-sided market is quite recent and sparse (Chao and Derdenger (2013)). While the definition of a two-sided market remains controversial (Geyl 2010), most of the definitions have some commonalities: they require the existence of two or more distinct user groups, where at least one of them is interested in interacting with another group. These interactions are enabled through the two-sided (or multi-sided) *platform*. The first definition of two-sided market is proposed by Evans (2002) which stresses the cross-network externalities between one or more groups. Rochet and Tirole (2006) propose another definition of two-sided market from the perspective of the price structure's impact on the volume of transactions on the platform. For other general studies on two-sided market, we refer the reader to Rochet and Tirole (2003), Caillaud and Jullien (2003), Rysman (2004), Armstrong (2006), Hagiu (2006), Kaiser and Wright (2006), Evans and Schmalensee (2005), and Weyl (2010)

Rochet and Tirole (2008), Amelio and Jullien (2012), Choi (2010) and Chao and Derdenger (2013) study bundling in the context of two-sided market. Hagiu and Wright (2014) consider the intermediary's choice between functioning as a marketplace or as a reseller where the marketplace serves as a two-sided platform that enables commercial trade between buyers and sellers. To the best of our knowledge, there has not been any research that explores the assortment planning in a two-sided market.

3.3 Model

We consider a firm facing a two-sided market which receives revenues from its two user groups: its *customers* who purchase its products and the *advertisers* who pay to promote their brand to the firm's customers. The firm can choose products to offer from a set of N distinct *product profiles*. The firm's customers come from K different customer segments and each segment has its own preferences for the product profiles which are represented by an attraction model. Specifically, let u_j^k denote the attractiveness of product profile $j = 1, \ldots, N$ for customers from segment $k = 1, \ldots, K$ and let $\mathbf{u}_j = (u_j^1, \ldots, u_j^K)$ denote the attractiveness vector of product profile j. Also let u_0^k denote the attractiveness of the no-purchase option for segment k and let $\mathbf{u}_0 = (u_0^1, \ldots, u_0^K)$ be the attractiveness vector of the no-purchase option.

Let λ denote the total market size and α^k denote the proportion of customers from segment k, such that $\sum_{k=1}^{K} \alpha_k = 1$.

The firm earns a revenue p_j per purchase of a product from profile j by a customer as well as a revenue of δ^k from the advertisers (collectively) per customer of segment k who buys one of its products.

When choosing its assortment of products, the firm is allowed to 'repeat' each product profile up to a certain number of times. Products which have the same profile are sold at the same price p_j and generate the same attractiveness to customers; in practice they may still be somewhat differentiated but the differences are such that they do not impact customer preferences. For example, a television network may include two different kids channels in their portfolio. Let r_j denote the maximum number of *repetitions* of product profile j which are allowed in the assortment. The firm may face a cardinality constraint on size of the chosen assortment; let $\overline{n} \leq +\infty$ denote the maximum number of products which can be offered. Let $\mathbf{x} = (x_1, ..., x_N)$ be the vector which counts the number of repetitions of each product profile in the chosen assortment. For $j = 1, ..., N$, we have $x_j \in \{0, 1, ..., r_j\}$ and $\sum_{j=1}^N x_j \leq \overline{n}$.

In our base setting, which we refer to as the *two-sided market product selection* problem $\mathcal{P}$, the firm chooses the assortment in order to maximize its total revenue (from sales and from advertising), denoted by Π. This problem can be written as follows:

$$
\begin{aligned}
(\mathcal{P}) \qquad \max_{\mathbf{x}\in\mathbb{N}^N} \ \Pi(\mathbf{x}) &= \sum_{j=1}^{N} p_j \sum_{k=1}^{K} \lambda\alpha^k \frac{x_j u_j^k}{\sum_{i=1}^N x_i u_i^k + u_0^k} + \sum_{k=1}^{K} \delta^k \lambda\alpha^k \frac{\sum_{j=1}^N x_j u_j^k}{\sum_{i=1}^N x_i u_i^k + u_0^k} \\
&= \sum_{j=1}^{N}\sum_{k=1}^{K} (p_j + \delta^k)\lambda\alpha^k \frac{x_j u_j^k}{\sum_{i=1}^N x_i u_i^k + u_0^k} \\
\text{s.t.} \qquad & x_j \leq r_j, j = 1, ..., N \\
& \sum_{j=1}^{N} x_j \leq \overline{n}
\end{aligned}
$$

In the first expression for Π, the first term corresponds to the sales revenue (from customers buying the products) and the second term is the advertising revenue (from advertisers). Let $\mathbf{x}^*$ denote the optimal assortment. In what follows we refer to a case of $r_j \geq \overline{n}$ for $j = 1, ..., N$ as a problem with *unlimited repetitions* and to the special case of $r_j = 1$ for $j = 1, ..., N$ as a problem with *no repetitions*.

3.3.1 Special cases

In this section we present special cases of the two-sided market product selection problem.

When $\delta^k = 0$ for $k = 1, ..., K$, the firm does not receive any revenue from the advertisers, only sales revenue from the customers. We refer to this problem as the *one-sided market product selection* problem $\mathcal{P}^o$. This problem is in fact the classic assortment planning problem with multiple customer segments, also referred to as the assortment planning problem with a mixture-of-logits choice model, which is studied in Rusmevichientong et al. (2014). It can be formulated as follows:

$$
\begin{aligned}
(\mathcal{P}^o) \qquad & \max_{\mathbf{x}\in\mathbb{N}^N} \quad && \Pi^o(\mathbf{x}) = \lambda \sum_{j=1}^{N} p_j \sum_{k=1}^{K} \alpha^k \frac{x_j u_j^k}{\sum_{i=1}^N x_i u_i^k + u_0^k} \\
& \text{s.t.} && x_j \le r_j, j = 1, ..., N \\
& && \sum_{j=1}^{N} x_j \le \overline{n}
\end{aligned}
$$

Let $\mathbf{x}^{*o}$ denote the optimal assortment for problem $\mathcal{P}^o$.

Alternatively when $p_j = 0$ for all $j \in \{1, \dots, N\}$, the firm's revenues come solely from the advertisers. In this case, we have:

$$
\begin{aligned}
(\mathcal{P}^f) \qquad & \max_{\mathbf{x}\in\mathbb{N}^N} \quad && \Pi^f(\mathbf{x}) = \sum_{k=1}^{K} \delta^k \lambda \alpha^k \frac{\sum_{j=1}^N x_j u_j^k}{\sum_{i=1}^N x_i u_i^k + u_0^k} \\
& \text{s.t.} && x_j \le r_j, j = 1, ..., N \\
& && \sum_{j=1}^{N} x_j \le \overline{n}
\end{aligned}
$$

In this special case, the firm does not ignore the customers but chooses to fully subsidize their consumption of the product. We refer to this case as the *free product selection* problem $\mathcal{P}^f$ and let $\mathbf{x}^{*f}$ denote the corresponding optimal assortment.

We also define the *single-customer segment product selection* Problem $\mathcal{P}_k^s$ as the special case where only segment k is considered, i.e., such that $\alpha^k = 1$ and $\alpha^l = 0$ for $l \neq k$. This problem can be formulated as:

$$
\begin{aligned}
(\mathcal{P}_k^s) \qquad & \max_{\mathbf{x}\in\mathbb{N}^N} \quad && \Pi_k(\mathbf{x}) = \sum_{j=1}^{N} (p_j + \delta^k) \lambda \alpha^k \frac{x_j u_j^k}{\sum_{i=1}^N x_i u_i^k + u_0^k} \\
& \text{s.t.} && x_j \le r_j, j = 1, ..., N \\
& && \sum_{j=1}^{N} x_j \le \overline{n}
\end{aligned}
$$

Let $\mathbf{x}^{*s_k}$ denote the optimal assortment for problem $\mathcal{P}_k^s$. When there is no cardinality constraint, i.e, when $\overline{n} \ge \sum_{j=1}^N r_j$, this problem reduces to the classical assortment planning problem under the MNL model with a homogenous customer population, which is studied by Li (2007) and Talluri and Van Ryzin (2004).

Finally we also consider the special case where $N = 1$, that is there is only one product profile available to the firm. We refer to this case as the *unique product profile problem.* The firm's only decision variable is the number of repetitions of this product in the assortment. Letting $x \equiv x_1$ and $r \equiv r_1$, the problem can be formulated as:

$$(\mathcal{P}^u) \qquad \max_{x \in \{1,\ldots,\min\{r,\overline{n}\}\}} \qquad \Pi(x) = \sum_{k=1}^{K} \lambda \alpha^k (p + \delta^k) \frac{x u^k}{x u^k + u_0^k}$$

3.4 Illustrative Example

In this section, we use a numerical example to illustrate some key properties of the optimal solution.

Example 3.4.1. Consider a problem with $N = 3$ product profiles with selling prices $(p_1, p_2, p_3) = (23, 98, 93)$. There are two customer segments, i.e., $K = 2$, with attractiveness vectors $\mathbf{u}_1 = (18, 48)$, $\mathbf{u}_2 = (28, 10)$, $\mathbf{u}_3 = (45, 3)$ and $\mathbf{u}_0 = (81, 88)$. Advertising revenues per customer from each segment are equal to $(\delta^1, \delta^2) = (44, 12)$. The total market size $\lambda = 343$ and the proportion of customers from segment 1 is $\alpha^1 = 0.54$ so that the proportion of customers from segment 2 is $\alpha^2 = 0.46$.

Suppose the firm imposes a maximum size of $\overline{n} = 2$ for the assortment and unlimited repetitions of each product profile are allowed, i.e. $r_1 = r_2 = r_3 \geq 2$. Table 3.1 shows the total revenue of all possible assortments with cardinality less or equal to 2, breaking it down between revenue from product sales and from advertising and specifying the number of customers from each segment who buy a product. For example assortment $x = (1, 1, 0)$ corresponds to offering one product from profile 1 and one product from profile 2, while $x = (2, 0, 0)$ corresponds to offering two repetitions of profile 1.

Table 3.1. Revenue allocation in Example 3.4

x	Total revenue (Π)	Revenue from sales	Revenue from advertising	Sales from Segment 1	Sales from Segment 2
$(1,0,0)$	4205	2056	2149	2254	19512
$(0,1,0)$	8522	6237	2285	6748	1773
$(0,0,1)$	9599	6629	2970	9052	547
$(2,0,0)$	6699	3205	3494	3814	2885
$(1,1,0)$	10557	6856	3702	7549	3008
$(1,0,1)$	11737	7481	4257	9470	2267
$(0,2,0)$	13957	10278	3679	10738	3219
$(0,1,1)$	14396	10293	4103	12182	2214
$(0,0,2)$	14398	9993	3346	13339	1059

From Table 3.1, we see that different assortments correspond to very different split of total revenue between sales and advertising. For example, when the assortment is $(1,0,0)$ 49% of the revenue comes from sales, while when the assortment is $(0,1,0)$, 73% of the revenue comes from advertising.

In this example, the optimal assortment is $(0,0,1)$, that is, it is best to 'repeat' product profile 3 twice. Based on the breakdown of profit between sales and advertising revenue, we see that, while other assortments generate higher sales revenues (e.g., $(0,2,0)$) or higher advertising revenues (e.g., $x=(1,0,1)$) the optimal assortment provides the highest total revenues because product 3, which has a relatively high selling price, is also very popular with the most lucrative customer segment 1 from the point of view of advertising.

Next we compare the total revenue obtained with each possible assortment across the different special cases we consider.

As we already saw from Table 3.1, assortment $\{1,2\}$ maximizes total revenue for the two-sided market product selection problem $\mathcal{P}$, that is, when revenues from both customers and advertisers are considered. If instead we ignore revenues from advertisers, the optimal assortment for problem $\mathcal{P}^o$ is $\{1,1\}$, that is, it is optimal to offer two repetitions of product 1 because it has the highest price. When the products are free and revenues come solely from advertisers it is optimal for problem $\mathcal{P}^o$ to offer $\{3,3\}$, because it is very attractive

Table 3.2. Comparison of revenues across problems in Example 3.4

S	Two-sided market ($\mathcal{P}$)	One-sided market ($\mathcal{P}^o$)	Free product ($\mathcal{P}^f$)	Single-segment (1) ($\mathcal{P}_1^s$)	Single-segment (2) ($\mathcal{P}_2^s$)
$\{1\}$	4205.4	2056.2	2149.2	2253.6	**1951.8**
$\{2\}$	8521.7	6237.2	2284.5	6748.3	1773.4
$\{3\}$	**9598.7**	**6629.1**	**2969.6**	**9051.8**	546.9
$\{1,1\}$	6699.1	3205.2	3493.9	3813.8	2885.3
$\{1,2\}$	10557	6855.5	3701.5	7548.6	3008.4
$\{1,3\}$	11737	7480.5	**4256.5**	9469.7	2267.3
$\{2,2\}$	13957	10278	3679	10738	**3219**
$\{2,3\}$	14396	**10293**	4103	12182	2214
$\{3,3\}$	**14398**	9993.2	3345.8	**13339**	1059

to customers from the more profitable segment 2. Finally note that the solutions to the single-segment problems differ from that of the two-sided market problem: $\{1,1\}$ and $\{2,2\}$ are optimal when there are only customers from segments 1 and 2 respectively.

In Table 3.3 we calculate the revenue loss (in percentage) which would be incurred by the firm if instead of solving the two-sided market problem, it used one of the other problems to determine the chosen assortment. Row 1 reports the chosen assortment, Row 2 reports the total revenues obtained from this assortment in a two-sided market and Row 3 calculates the corresponding percentage revenue loss.

Table 3.3. Revenue loss from using incorrect model

	$\mathcal{P}$	$\mathcal{P}^o$	$\mathcal{P}^f$	$\mathcal{P}_1^s$	$\mathcal{P}_2^s$
Optimal Assortment	$\{3,3\}$	$\{2,3\}$	$\{1,3\}$	$\{3,3\}$	$\{2,2\}$
Revenue	14398	14396	11737	14398	13957
Percent revenue loss	0%	0.01%	18.48%	0%	3.06%

From Table 3.3 we see that ignoring the revenues from one user group can lead to a significant revenues loss, especially if one ignores the revenues from sales. We also see that it is important to consider both customer segments as focusing only on one of them also leads to a significant revenue loss.

Now we extend Example 3.4 by increasing the maximum size of chosen assortment to $\overline{n} = 3$ and setting $r_1 = r_2 = r_3 = 3$ (i.e., unlimited repetitions). The revenue for all possible assortments of cardinality 3 in all five problems is shown in Table 3.4.

Table 3.4. Comparison of revenues across problems for cardinality = 3

S	Two-sided market ($\mathcal{P}$)	One-sided market ($\mathcal{P}^o$)	Free product ($\mathcal{P}^f$)	Single-segment (1) ($\mathcal{P}_1^s$)	Single-segment (2) ($\mathcal{P}_2^s$)
$\{1,1,1\}$	8390	3958	4433	4958	3432
$\{1,1,2\}$	11783	7154	4629	8150	3633
$\{1,2,2\}$	11783	7154	4629	8150	3633
$\{1,2,3\}$	15487	10404	5083	12205	3282
$\{1,1,3\}$	12900	7826	5074	9795	3105
$\{1,3,3\}$	15820	10447	**5373**	13250	2570
$\{2,2,2\}$	17792	**13166**	4626	13374	**4418**
$\{2,2,3\}$	**17930**	13019	4911	14350	3580
$\{2,3,3\}$	17787	12669	5118	15159	2628
$\{3,3,3\}$	17380	12116	5264	**15841**	1539

We can see from Tables 3.2 and 3.4 that the optimal for the two-sided market problem $\mathcal{P}$ with $\overline{n} = 3$ is $\{1, 2, 2\}$, that is it is optimal to repeat product profile 2 twice and product profile once.

Let us further extend Example 3.4 by allowing up to 5 products in the chosen assortment and up to five repetitions of each product profile, that is, $\mathcal{N} = \{1^{(1)}, \ldots, 1^{(5)}, \ldots, 3^{(1)}, \ldots, 3^{(5)}\}$. Table 3.5 shows how the optimal assortment for each problem varies with the cardinality constraint.

We see from Table 3.5 that the optimal assortment for Problems $\mathcal{P}_k^s$ is such that it contains only one distinct product profile but this is not the case for problems $\mathcal{P}$ and $\mathcal{P}^o$. We also see that, for all problems, the optimal assortment always contains as many products as allowed. These results are formally proven in Section 3.5.1. The optimal assortment for problem $\mathcal{P}^f$ also contains only one distinct product file in this example, but we will show in Section 3.5.1 that this does not hold in general.

Table 3.5. Optimal assortment and revenue with different cardinality constraint in Example 1 under unlimited repetitions

$\overline{n}$	Two-sided market ($\mathcal{P}$)		One-sided market ($\mathcal{P}^o$)		Free product ($\mathcal{P}^f$)		Single-segment ($\mathcal{P}_1^s$)		Single-segment ($\mathcal{P}_2^s$)	
1	$(0,0,1)$	9599	$(0,0,1)$	9599	$(0,0,1)$	9599	$(0,0,1)$	9599	$(1,0,0)$	4205
2	$(0,0,2)$	14398	$(0,1,1)$	14396	$(0,0,2)$	14398	$(0,0,2)$	14398	$(0,2,0)$	13957
3	$(0,2,1)$	17930	$(0,3,0)$	17792	$(1,0,2)$	15820	$(0,0,3)$	17380	(0,3,0)	17792
4	$(0,4,0)$	18952	$(0,4,0)$	18952	$(1,0,3)$	18436	$(0,0,4)$	19470	$(0,4,0)$	18952
5	$(0,5,0)$	21020	$(0,5,0)$	21020	$(2,0,3)$	18873	$(0,0,5)$	21052	$(0,5,0)$	21020

Note: Under each of the five problems, the first column indicates the optimal assortment, and the second column indicates the revenue of the optimal assortment in the context of the *two-sided market product selection* problem.

Next we study the optimal assortment in Example 3.4 when repetitions are not allowed, that is $\mathcal{N} = \{1,2,3\}$, and we allow up to 3 products in the chosen assortment, that is, $\overline{n} = 1,2,3$.

The optimal assortment and the corresponding total revenue in the context of *two-sided market problem* are shown in Table 3.6.

Table 3.6. Optimal assortment and revenue when no receptions are allowed

$\overline{n}$	Two-sided market ($\mathcal{P}$)		One-sided market ($\mathcal{P}^o$)		Free product ($\mathcal{P}^f$)		Single-segment ($\mathcal{P}_1^s$)		Single-segment ($\mathcal{P}_2^s$)	
1	$(0,0,1)$	9599	$(0,0,1)$	9599	$(0,0,1)$	9599	$(0,0,1)$	9599	$(1,0,0)$	4205
2	$(0,1,1)$	14396	$(0,1,1)$	14396	$(1,0,1)$	11737	$(0,1,1)$	14396	$(1,1,0)$	10557
3	$(1,1,1)$	15487	$(1,1,1)$	15487	$(1,1,1)$	15487	$(1,1,1)$	15487	(1,1,1)	15487

We find from Table 3.6 that for the *two-sided market* problem, the optimal assortment with $\overline{n} = 3$ is $\{1,2\}$, that is, it is optimal to offer only two products even though the maximum size of the chosen assortment is 3. The same holds for the *one-sided market* problem and the *single-segment* problem $\mathcal{P}_1^s$. By comparing Table 3.6 with Table 3.5, we see that for the optimal solution with or without repetitions may differ for a particular problem given a fixed cardinality $\overline{n}$. For example, when $\overline{n} = 3$, the optimal assortment for the *two-sided market* problem with unlimited repetitions is $\{1,2,2\}$, while the optimal assortment with no repetitions is $\{1,2\}$.

3.5 Results

3.5.1 Structural Properties

In this section, we state some structural properties of the *two-sided market* problem and its special cases, namely the *one-sided market* problem, the *free product selection* problem, the *single-customer segment* problem and the unique product profile problem.

Structural Properties of the Two-sided Market Problem

Proposition 10. If an assortment contains only one distinct product profile (possibly with a number of repetitions), adding any number of repetitions of that product profile, when allowed, always increases the total revenue in the *Two-sided market* problem $\mathcal{P}$. Formally, for product profile $j \in \{1, \ldots, N\}$, $\Pi(l\mathbf{e}^j) \geq \Pi(m\mathbf{e}^j)$ for all m and l such that $0 \leq m \leq l \leq \min\{\overline{n}, r_j\}$.

Proposition 10 and Corollary 1 establishes that repeating the product from a single-profile assortment always increases revenue. Proposition 11 shows that the result no longer hold if the initial assortment contains more than one product profile.

Proposition 11. If an assortment contains more than one distinct product profiles, then adding a repetition of a product which is already in the assortment does not always improve the total revenue in the *Two-sided market* problem $\mathcal{P}$.

Proof. Using the parameters from Example 3.4 with $\overline{n} = 4$, we find that:

$$\Pi((1,2,0)) = 14860$$

$$\Pi((2,2,0)) = 14560$$

Therefore, the total revenue of the *two-sided market* problem decreases in this case as we add a repetition of product profile 1 to assortment $\{1,2,2\}$. $\square$

The intuition behind these last results is as follows. We notice that adding a product to an assortment always increases the sales from all segments because the attraction of no-purchase option in each of the segments is weakened. If the assortment contains only one product profile, then adding an repetition of this product results in that the newly captured customers in segment $k \in \{1, \ldots, K\}$ brings in revenue at the same rate (which equals $(p_j + \delta^k)$) as the old customers. Thus the revenue from both sales and advertising increase for all segments. In contrast, when there is more than one distinct product profiles in the assortment, the extra advertising revenues from adding the repetitions of a product could come at the expense of the sales revenues on more lucrative products. If the decrease in revenue from sales exceeds the increase in revenue from advertising, then the total revenue goes down.

Lemma 10. For the *two-sided market* problem $\mathcal{P}$, adding a product with higher price than the highest price among the products already in an assortment always improve the total revenue, that is $\Pi(\mathbf{x} + \mathbf{e}^i) \geq \Pi(\mathbf{x})$ where $p_i \geq \max_{j:x_j>0}\{p_j\}$.

Our next set of results concerns the size of the optimal assortment. In what follows, let $\mathbf{x}^{(n)^*} := \arg\max_{\mathbf{x}\in\mathbb{N}^N, \sum_{j=1}^N x_j = n} \Pi(\mathbf{x})$ be the optimal assortment with cardinality n when infinite repetitions of each product profile are allowed.

Lemma 11. For the *two-sided market* problem $\mathcal{P}$ with unlimited repetitions and finite cardinality constraint, $\Pi(\mathbf{x}^{(l)*}) \geq \Pi(\mathbf{x}^{(m)*})$ for $1 \leq m \leq l \leq \overline{n}$.

Proof. From Lemma 10 one can always increase the profit of assortment $\mathbf{x}$ by adding another repetition of the product with the highest price. Hence, there always exists a product i such that $\Pi(\mathbf{x}^{(n)*} + \mathbf{e}^i) \geq \Pi(\mathbf{x}^{(n)*})$, which implies $\Pi(\mathbf{x}^{(n+1)*}) \geq \Pi(\mathbf{x}^{(n)*} + \mathbf{e}^i) \geq \Pi(\mathbf{x}^{(n)*})$, $n = 1, \ldots, \overline{n}$. Then it follows that $\Pi(\mathbf{x}^{(l)*}) \geq \Pi(\mathbf{x}^{(m)*})$ for $1 \leq m \leq l \leq \overline{n}$. $\square$

Lemma 12. For the *two-sided market* problem $\mathcal{P}$ with no repetitions and finite cardinality constraint, $\Pi(\mathbf{x}^{(l)*}) \geq \Pi(\mathbf{x}^{(m)*})$ does not always hold for $1 \leq m \leq l \leq \overline{n}$.

Proof. Consider an example with 3 products and 2 customer segments, where $(p_1, p_2, p_3) = (15, 10, 3)$, $\mathbf{u}_1 = (3, 1)$, $\mathbf{u}_2 = (1, 5)$, $\mathbf{u}_3 = (2, 10)$ and $\mathbf{u}_0 = (1, 10)$. Advertising revenues per customer from each segment are equal to $(\delta^1, \delta^2) = (1, 10)$. The total market size $\lambda = 400$, which divides equally between the two segments, i.e., $\alpha^1 = \alpha^2 = \frac{1}{2}$. When repetition is not allowed, we have

$$\Pi(\mathbf{x}^{(2)*}) = \Pi((1, 1, 0)) = 3922.5$$
$$\Pi(\mathbf{x}^{(3)*}) = \Pi((1, 1, 1)) = 3875.8$$

in which case $\Pi(\mathbf{x}^{(3)*}) \leq \Pi(\mathbf{x}^{(2)*})$. $\square$

From Lemma 11 we see that, under unlimited repetitions, increasing the size of the assortment and optimally choosing the products, always leads to an increase in revenues. Lemma 12 shows that this result is no longer true when repetitions are not allowed.

We use Lemma 11 to establish that the optimal assortment always has full cardinality under unlimited repetitions, as stated in Proposition 12.

Proposition 12. For the *two-sided market* problem $\mathcal{P}$, the optimal assortment under unlimited repetition and finite cardinality constraint always has cardinality $\overline{n}$. That is, if $r_j \geq \overline{n}$ for $j = 1, ..., N$ and $\overline{n} \leq +\infty$, then $\sum_{j=1}^{N} x_j^* = \overline{n}$.

Proof. According to Lemma 11, $\Pi(\mathbf{x}^{(\overline{n})*}) \geq \Pi(\mathbf{x}^{(m)*})$ for $1 \leq m < \overline{n}$, therefore, $\sum_{j=1}^{N} x_j^* = \overline{n}$. $\square$

We have seen from Table 3.5 in Example 3.4 that the optimal assortment of problems $\mathcal{P}$ may sometimes take the form of repetition of single product but not always. We present this result in Proposition 13 and we will skip the proof.

Proposition 13. For the *two-sided market* problem $\mathcal{P}$ under unlimited repetitions and finite cardinality constraint, the optimal assortments may include more than one different product profile.

Lemma 13. For the *two-sided market* problem $\mathcal{P}$ with unlimited repetitions and finite cardinality constraints, adding product in greedy method does not always generate optimal solution.

Proof. We can see from Example 3.4,

$$\mathbf{x}^{*(2)} = (0, 0, 2)$$

$$\mathbf{x}^{*o(2)} = (0, 1, 2)$$

$$\mathbf{x}^{*(3)} = (0, 2, 1)$$

$$\mathbf{x}^{*o(3)} = (0, 3, 0)$$

Therefore $S^{*(2)} \nsubseteq S^{*(3)}$, $S^{*o(2)} \nsubseteq S^{*o(3)}$, so greedy method does not give optimal solution in this case. $\square$

Structural Properties of Other Problems

We have seem from Proposition 10 that in the *two-sided market* problem, if the assortment contains only one district product file, keep adding that product to the assortment always improve the revenue associated with the assortment. The same result holds in the context of the *one-sided market problem* , the *free product selection* problem and the *single-customer segment product selection* problem as stated in Corollary 1.

Corollary 1. If an assortment contains only one distinct product profile (possibly with a number of repetitions), adding any number of repetitions of that product profile, when allowed, always increases the total revenue in the the *one-sided market problem* $\mathcal{P}^o$, the *free product selection* problem $\mathcal{P}^f$ and the *single-customer segment product selection* problem $\mathcal{P}^s_k$.

Formally, for product profile $j \in \{1, \ldots, N\}$, we have:

$$\Pi^o(l\mathbf{e}^j) \geq \Pi^o(m\mathbf{e}^j)$$

$$\Pi^f(l\mathbf{e}^j) \geq \Pi^f(m\mathbf{e}^j)$$

$$\Pi^s_k(l\mathbf{e}^j) \geq \Pi^s_k(m\mathbf{e}^j)$$

for all m and l such that $0 \leq m \leq l \leq \min\{\overline{n}, r_j\}$, $k = \{1, \ldots, K\}$.

Proof. The results follow directly from Proposition 10 since problems $\mathcal{P}^o$, $\mathcal{P}^f$ and $\mathcal{P}^s_k$ are special cases of problem $\mathcal{P}$. □

Lemma 14 establishes the same result as in Proposition 11 for problems $\mathcal{P}^o$, $\mathcal{P}^f$ and $\mathcal{P}^s_k$.

Lemma 14. If an assortment contains more than one distinct product profile, then adding a repetition of a product which is already in the assortment does not always improve the total revenue in the *one-sided market problem* $\mathcal{P}^o$, the *free product selection* problem $\mathcal{P}^f$ and the *single-customer segment product selection* problem $\mathcal{P}^s_k$, $k = 1, \ldots, K$.

Corollary 2. For the *one-sided market* problem $\mathcal{P}^o$ and the *single-segment problem* $\mathcal{P}^s_k$, adding a product with higher price than the highest price among the products already in an assortment always improve the total revenue, that is $\Pi^o(\mathbf{x} + \mathbf{e}^i) \geq \Pi^o(\mathbf{x})$ and $\Pi^s_k(\mathbf{x} + \mathbf{e}^i) \geq \Pi^s_k(\mathbf{x})$ hold for all $\mathbf{x} \in \mathbb{N}^N$ (or feasible $\mathbf{x}$?) where $p_i \geq \max_{j:x_j>0}\{p_j\}$.

Proof. The results follow directly from Lemma 10 because problems $\mathcal{P}^o$ and $\mathcal{P}^s_k$ are special cases of problem $\mathcal{P}$. □

Corollary 3. For the *one-sided market problem* $\mathcal{P}^o$, the *free product selection* problem $\mathcal{P}^f$ and the *single-customer segment product selection* problem $\mathcal{P}^s_k$, $k = 1, \ldots, K$ with unlimited repetitions and finite cardinality constraint, define

$$\begin{aligned}
\mathbf{x}^{(n)*o} &= \underset{\mathbf{x}\in\mathbb{N}^N, \sum_{j=1}^N x_j = n}{arg\max} \{\Pi^o(\mathbf{x})\} \\
\mathbf{x}^{(n)*f} &= \underset{\mathbf{x}\in\mathbb{N}^N, \sum_{j=1}^N x_j = n}{arg\max} \{\Pi^f(\mathbf{x})\} \\
\mathbf{x}^{(n)*s_k} &= \underset{\mathbf{x}\in\mathbb{N}^N, \sum_{j=1}^N x_j = n}{arg\max} \{\Pi^s_k(\mathbf{x})\}
\end{aligned}$$

Then, for $1 \leq m \leq l \leq \overline{n}$, we always have:

$$\Pi(\mathbf{x}^{*o(l)}) \geq \Pi(\mathbf{x}^{*o(m)})$$

$$\Pi(\mathbf{x}^{*f(l)}) \geq \Pi(\mathbf{x}^{*f(m)})$$

$$\Pi(\mathbf{x}^{*s_k(l)}) \geq \Pi(\mathbf{x}^{*s_k(m)})$$

Proof. The results follow directly from Lemma 11 since problems $\mathcal{P}^o$, $\mathcal{P}^f$ and $\mathcal{P}^s_k$ are special cases of problem $\mathcal{P}$. □

Corollary 4. For the *one-sided market problem* $\mathcal{P}^o$, the *free product selection* problem $\mathcal{P}^f$ and the *single-customer segment product selection* problem $\mathcal{P}^s_k$, $k = 1, \ldots, K$ with unlimited repetitions and finite cardinality constraint, then the optimal assortments always have cardinality $\overline{n}$.

Proof. The results follow directly from Proposition 12 since problems $\mathcal{P}^o$, $\mathcal{P}^f$ and $\mathcal{P}^s_k$ are special cases of problem $\mathcal{P}$. □

When repetitions of products are not allowed, the optimal assortments takes full cardinality only in the *free product selection* problem $\mathcal{P}^f$. We discuss this in Proposition 14.

Proposition 14. For the *free product selection* problem $\mathcal{P}^f$ with finite cardinality constraint where $\overline{n} \leq N$, if repetitions of products are not allowed, then the optimal assortment always have cardinality $\overline{n}$, that is $\mathbf{x}^{*f}$ is such that $\sum_{j=1}^N x_j^{*f} = \overline{n}$.

We can see that the optimal assortment of the *free product selection* problem $\mathcal{P}^f$ always takes full cardinality regardless of whether repetitions are allowed. As we have mentioned, an assortment with higher cardinality always captures a larger number of customers than an assortment with lower cardinality. Since revenue comes solely from advertising in this case, increase in total sales leads to more revenue. Observe the special structure of the *free product selection* problem, we are able to show a result under no cardinality constraint which is stated in Lemma 15.

Lemma 15. For the *free product selection* problem $\mathcal{P}^f$, if $\overline{n} \geq \sum_{j=1}^{N} r_j$, then the optimal assortment $\mathbf{x}^{*f} = (r_1, \ldots, r_N)$.

Intuition for Lemma 15 is that, although the products in the assortment compete with each others on capturing customers, this does not hurt the total revenue because there is no sales revenue in this problem. The whole revenue comes from advertising, the revenue rate of which is the same within each customer segment, and thus always increases as the number of customers in each segment grows. ALSO: Keep adding products will weaken the attraction of no purchase option and thus capture more customer in all segments.

As we have seen in Example 3.4, under unlimited repetitions, the optimal assortment of the *single customer segment* problems takes the form of repeating a product $\overline{n}$ times, now we formally present this result in Proposition 15.

Proposition 15. For the *single segment* problem $\mathcal{P}_k^s$ under unlimited repetitions and finite cardinality constraint, the optimal assortment is $\mathbf{x}^{*s_k} = \overline{n}\mathbf{e}^j$ for some $j \in \{1, ..., N\}$.

We have also seen from Table 3.5 in Example 3.4 that the optimal assortment of problems $\mathcal{P}$, $\mathcal{P}^o$ and $\mathcal{P}^f$ may sometimes take the form of repetition of single product but not always. We present this result in Proposition 16 and we will skip the proof.

Proposition 16. For the *one-sided market* problem $\mathcal{P}^o$ and the *free product selection* problem $\mathcal{P}^f$ under unlimited repetitions and finite cardinality constraint, the optimal assortments may include more than one different product profile.

From Proposition 15, we know that the optimal assortment for the single customer segment problem $\mathcal{P}_s^k$ under unlimited repetitions and cardinality constraint always includes $\overline{n}$ repetitions of the same product profile. However, this product profile may not be the same for $\overline{n}$ and $\overline{n}+1$ as shown in Example 3.4. We state this result in Lemma 16.

Lemma 16. For the *single segment* problem $\mathcal{P}_k^s$ under unlimited repetitions and finite cardinality constraint, the optimal assortment for certain k could be repetitions of different products under different $\overline{n}$.

But under some special conditions, for problem $\mathcal{P}_k^s$ with certain k, the repetitions of a product could dominate the repetitions of another product across all number of repetitions (the conditions is shown in Lemma 17). And this enables us to find a special case where the optimal assortments of problem $\mathcal{P}_k^s$ under unlimited receptions and certain k are repetitions of the same product across all possible values of $\overline{n}$. We present this result in Lemma 18.

Lemma 17. Assume i and j to be two distinct product files in $\{1,\ldots,N\}$. If $p_i \geq p_j$ and $\Pi_k(\mathbf{e}^i) \geq \Pi_k($
$mathbfe^j)$, then $\Pi_k(m\mathbf{e}^i) \geq \Pi_k(m\mathbf{e}^j)$ for $1 \leq m \leq \min\{r^{(i)}, r^{(j)}, \overline{n}\}$.

Proof.

$$\begin{aligned}
& \Pi_k(m\mathbf{e}^i) - \Pi_k(m\mathbf{e}^j) \\
= \;& \frac{m(p_i+\delta^k)u_i^k}{mu_i^k+u_0^k} - \frac{m(p_j+\delta^k)u_j^k}{mu_j^k+u_0^k} \\
= \;& \frac{m(p_i+\delta^k)u_i^k(mu_j^k+u_0^k) - m(p_j+\delta^k)u_j^k(mu_i^k+u_0^k)}{(mu_i^k+u_0^k)(mu_j^k+u_0^k)} \\
= \;& \frac{m^2\left\{u_i^k u_j^k(p_i-p_j) - \frac{u_0^k}{m}\left[(p_j+\delta^k)u_j^k - (p_i+\delta^k)u_i^k)\right]\right\}}{(mu_i^k+u_0^k)(mu_j^k+u_0^k)}
\end{aligned}$$

Given that $p_i \geq p_j$ and $\Pi_k(\mathbf{e}^i) \geq \Pi_k(\mathbf{e}^j)$, we will prove $\Pi_k(m\mathbf{e}^i) \geq \Pi_k(m\mathbf{e}^j)$ by showing that

$$u_i^k u_j^k(p_i-p_j) - \frac{u_0^k}{m}\left[(p_j+\delta^k)u_j^k - (p_i+\delta^k)u_i^k)\right] \geq 0$$

Note that

$$
\begin{aligned}
& \Pi_k(\mathbf{e}^i) \geq \Pi_k(\mathbf{e}^j) \\
\Leftrightarrow\ & \frac{(p_i+\delta^k)u_i^k}{u_i^k+u_0^k} \geq \frac{(p_j+\delta^k)u_j^k}{u_j^k+u_0^k} \\
\Leftrightarrow\ & (p_i+\delta^k)u_i^k(u_j^k+u_0^k) \geq (p_j+\delta^k)u_j^k(u_i^k+u_0^k) \\
\Leftrightarrow\ & u_i^k u_j^k(p_i-p_j) \geq u_0^k\left[(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k\right]
\end{aligned}
$$

Since $p_i \geq p_j$, we have $u_i^k u_j^k(p_i-p_j) \geq 0$. If $(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k \leq 0$:

$$
\begin{aligned}
& u_i^k u_j^k(p_i-p_j) \geq 0 \geq \frac{u_0^k}{m}\left[(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k)\right] \\
\Leftrightarrow\ & u_i^k u_j^k(p_i-p_j) - \frac{u_0^k}{m}\left[(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k)\right] \geq 0
\end{aligned}
$$

If $(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k > 0$:

$$
\begin{aligned}
& u_i^k u_j^k(p_i-p_j) \geq (p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k) \geq \frac{u_0^k}{m}\left[(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k)\right] \\
\Leftrightarrow\ & u_i^k u_j^k(p_i-p_j) - \frac{u_0^k}{m}\left[(p_j+\delta^k)u_j^k-(p_i+\delta^k)u_i^k)\right] \geq 0
\end{aligned}
$$

Therefore, $\Pi_k(\{i^{(1)},\dots,i^{(m)}\}) \geq \Pi_k(\{j^{(1)},\dots,j^{(m)}\})$. □

Lemma 18. For the *single segment* problem $\mathcal{P}_k^s$ under unlimited repetitions, if there exists product $i \in \{1,\dots,N\}$ such that $p_i = \max_{j\in\{1,\dots,N\}}\{p_j\}$ and $\Pi_k^s(\mathbf{e}^i) = \max_{j\in\{1,\dots,N\}}\Pi_k^s(\mathbf{e}^j)$, then the optimal assortments under different cardinality constraints are repetitions of the same product i. That is $\mathbf{x}^{*s_k} = \mathbf{e}^{\overline{n}}$ for all $\overline{n} \leq \infty$.

Proof. The result in this lemma follows directly from Lemma 17 and Proposition 15. □

Lemma 16 indicates that for the *single segment problem* $\mathcal{P}_k^s$ under unlimited repetitions and finite cardinality constraints, optimal assortment can not be obtained through greedy method. And this is also true for problems $\mathcal{P}^o$ and $\mathcal{P}^f$ as we can see from Table 3.5 in Example 3.4. We show this result in Lemma 19.

Lemma 19. For the *one-sided market* problem $\mathcal{P}^o$ and the *free product selection* problem $\mathcal{P}^f$ with unlimited repetitions and finite cardinality constraints, adding product in greedy method does not always generate optimal solution.

Proof. We can see from Example 3.4,

$$\begin{aligned}
\mathbf{x}^{*(2)} &= \{0, 0, 2\} \\
\mathbf{x}^{*o(2)} &= \{0, 1, 2\} \\
\mathbf{x}^{*(3)} &= \{0, 2, 1\} \\
\mathbf{x}^{*o(3)} &= \{0, 3, 0\}
\end{aligned}$$

Therefore $S^{*(2)} \nsubseteq S^{*(3)}$, $S^{*o(2)} \nsubseteq S^{*o(3)}$, so greedy method does not give optimal solution in this case. P^f part: Need to find a numerical example here. $\square$

Proposition 17. For assortment $\mathbf{x} = (x_1, \ldots, x_N)$, if $\frac{\sum_{j=1}^N p_j x_j}{\sum_{j=1}^N x_j} \geq \max\limits_{k \in \{1,\ldots,K\}} \delta_k$, then the revenue generated by this assortment from sales is higher than from advertising: $\Pi^o(\mathbf{x}) \geq \Pi^f(\mathbf{x})$.

Proof. Let $\delta_m = \max\limits_{k \in \{1,\ldots,K\}} \{\delta^k\}$. $\frac{\sum_{j=1}^N p_j x_j}{\sum_{j=1}^N x_j} \geq \delta_m \Rightarrow \sum_{j=1}^N p_j x_j \geq \sum_{j=1}^N \delta_m x_j$, then for an arbitrary assortment $\mathbf{x} = (x_1, \ldots, x_N)$, we have:

$$\begin{aligned}
& \Pi^o(\mathbf{x}) - \Pi^f(\mathbf{x}) \\
= \;& \sum_{k=1}^K \lambda \alpha^k \frac{\sum_{j=1}^N p_j x_j u_j^k}{\sum_{j=1}^N x_j u_j^k + u_0^k} - \sum_{k=1}^K \lambda \alpha^k \frac{\sum_{j=1}^N \delta^k x_j u_j^k}{\sum_{j=1}^N x_j u_j^k + u_0^k} \\
= \;& \lambda \sum_{k=1}^K \lambda \alpha^k \frac{\sum_{j=1}^N (p_j - \delta^k) x_j u_j^k}{\sum_{j=1}^N x_j u_j^k + u_0^k} \\
\geq \;& \lambda \sum_{k=1}^K \lambda \alpha^k \frac{\sum_{j=1}^N (p_j - \delta_m) x_j u_j^k}{\sum_{j=1}^N x_j u_j^k + u_0^k} \\
= \;& \lambda \sum_{k=1}^K \lambda \alpha^k \frac{(\sum_{j=1}^N p_j x_j - \sum_{j=1}^N x_j \delta_m) u_j^k}{\sum_{j=1}^N x_j u_j^k + u_0^k} \\
\geq \;& 0
\end{aligned}$$

□

Corollary 5. If $\min_{j\in\{1,\dots,N\}}\{p_j\} \geq \max_{k\in\{1,\dots,K\}} \delta_k$, then $\Pi^o(\mathbf{x}) \geq \Pi^f(\mathbf{x})$ for all $\mathbf{x} \leq \mathbf{r}$.

Proof. If $\min_{j\in\{1,\dots,N\}}\{p_j\} \geq \max_{k\in\{1,\dots,K\}} \delta_k$, then $\sum_{j=1}^N p_j x_j \geq \sum_{j=1}^N \delta_m x_j$. Then according to Proposition 17, we have $\Pi^o(\mathbf{x}) \geq \Pi^f(\mathbf{x})$. □

3.5.2 Optimality of Revenue-Ordered Assortment

For a one-sided market problem (i.e., no advertisers), Li(2007) and Talluri and van Ryzin(2004) consider the assortment planning model with one customer segment and no cardinality constraint (i.e., $\overline{n} = +\infty$) and prove the optimality of *revenue-ordered assortments*, i.e., they show that it is optimal to offer a certain number of the products with the highest revenues. Rusmevichientong et al. (2014) extend their setting to multiple customer segments and show that, in general, revenue-ordered assortments are no longer optimal. However, they establish that the optimality is preserved under some specially structured cases.

In our setting, revenues come from two distinct user groups, i.e., customers who pay p_j for product j and advertisers who pay δ^k per customer of segment k who buys a product. Hence, there does not exist an obvious definition for revenue-ordered assortments in a two-sided market setting. Ordering assortments based on p_j would amount to ignoring the impact of advertisers.

Despite these considerations, we are able to obtain a set of conditions under which the optimal assortment has a simple ranking-based structure.

Theorem 2. Suppose the N products are such that $u_1^k \leq u_2^k \leq \dots, \leq u_N^k$ and $(p_1+\delta^k)u_1^k \geq (p_2+\delta^k)u_2^k \geq \dots \geq (p_N+\delta^k)u_N^k$ for $k = 1, ..., K$. In this case the optimal assortment for $\overline{n} = +\infty$ is $\mathbf{x}^* = (r_1, r_2, \dots, \hat{r_i}, 0, \dots, 0)$ for some $i \in \{1, ..., N\}$ where $\hat{r_i} \leq r_i$.

3.5.3 Offering Standard vs Specialized Products

In this section we study a special case of the two-sided market problem with two customer segments and three product profiles, i.e. $K = 2$ and $N = 3$ Product 1 (2) is a *specialized* product in that it appeals only to the customers from segment 1 (2). In contrast, product 3 is a *standard* product in that it appeals equally to customers from segments 1 and 2. Specifically, we have $\mathbf{u}_1 = (u_D, 0)$, $\mathbf{u}_2 = (0, u_D)$ and $\mathbf{u}_3 = (u_G, u_G)$, where $u_D, u_G \geq 0$. Also we have $u_0^1 = u_0^2 = u_0$, $p_1 = p_2 = p_3 = p$ and $\alpha_1 = \alpha_2 = \frac{1}{2}$. Without loss of generality we assume that $\delta^1 \geq \delta^2$. We study the optimal assortment with size $\overline{n} \in \{1, 2\}$ and allow unlimited repetitions.

First we consider the problem with $\overline{n} = 1$. The expected revenue from the three possible assortments is:

$$\begin{aligned}
\Pi(\{1\}) &= (p + \delta^1)\frac{1}{2}\lambda\frac{u_D}{u_D + u_0} \\
\Pi(\{2\}) &= (p + \delta^2)\frac{1}{2}\lambda\frac{u_D}{u_D + u_0} \\
\Pi(\{3\}) &= (2p + \delta^1 + \delta^2)\frac{1}{2}\lambda\frac{u_G}{u_G + u_0}
\end{aligned}$$

Since $\delta^1 \geq \delta^2$, we have $\Pi(\{1\}) \geq \Pi(\{2\})$. Hence the optimal assortment is either $\{1\}$ and $\{3\}$. Offering only product 1 means receiving revenues only from customers from segment 1, while offering only product 3 means collecting revenues from customers of both segments. It is easy to see that $\{3\}$ is optimal if $u_G \geq u_D$ since in that case, offering product 3 leads to more sales (and therefore more advertising revenues) from customers of segment 1, as well as a positive revenue from customers of segment 2. When $u_D > u_G$ the tradeoff is between more sales from the more lucrative segment 1 (when offering $\{1\}$) or collecting sales to both segments including from the less profitable segment 2. Formally, the optimal assortment is $\{1\}$ if and only if $\frac{\frac{u_D}{u_D+u_0}}{\frac{u_G}{u_G+u_0}} \geq \frac{2p+\delta^1+\delta^2}{p+\delta^1}$.

Figure 3.5.3 represents the region of optimality for assortment $\{1\}$ and $\{3\}$ as a function of u_G and u_D. In the area above the line, $\{3\}$ is optimal.

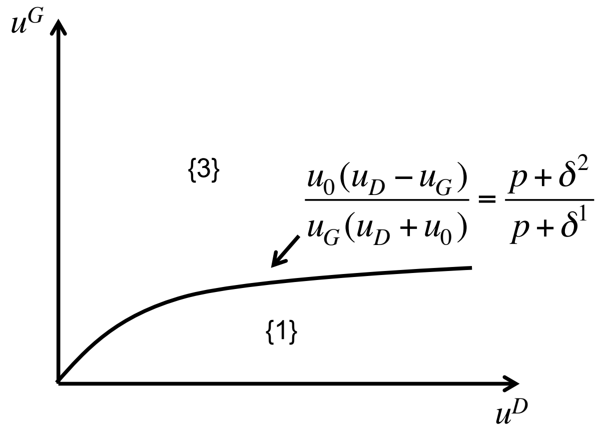

Figure 3.1. Optimality regions for $\overline{n} = 1$ with $p = 1, u_0 = 5, \delta^1 = 10$ and $\delta^2 = 3$

Next let us consider the problem with $\overline{n} = 2$. Because we have assumed that $\delta^1 \geq \delta^2$, it is easy to see that $\Pi(\{1,1\}) \geq \Pi(\{2,2\})$ and $\Pi(\{1,3\}) \geq \Pi(\{2,3\})$. Also we always have $\Pi(\{1,1\}) \geq \Pi(\{1\})$, $\Pi(\{2,2\}) \geq \Pi(\{2\})$ and $\Pi(\{3,3\}) \geq \Pi(\{3\})$. Hence, we only need to consider the following four assortments:

$$
\begin{aligned}
\Pi(\{1,1\}) &= (p+\delta^1)\frac{1}{2}\lambda\frac{2u_D}{2u_D+u_0} \\
\Pi(\{1,3\}) &= (p+\delta^1)\frac{1}{2}\lambda\frac{u_D+u_G}{u_D+u_G+u_0} + (p+\delta^2)\frac{1}{2}\lambda\frac{u_G}{u_G+u_0} \\
\Pi(\{1,2\}) &= (2p+\delta^1+\delta^2)\frac{1}{2}\lambda\frac{u_D}{u_D+u_0} \\
\Pi(\{3,3\}) &= (2p+\delta^1+\delta^2)\frac{1}{2}\lambda\frac{2u_G}{2u_G+u_0}
\end{aligned}
$$

The optimality conditions for the four assortments are shown in Table 3.7 and the optimality regions are represented in Figure 3.2.

Table 3.7. Optimal solution for $\overline{n} = 2$

Optimal assortment	Condition
$\{1,1\}$	$\frac{u_0(u_D-u_G)(u_G+u_0)}{u_G(2u_D+u_0)(u_D+u_G+u_0)} \geq \frac{p+\delta^2}{p+\delta^1}$, $\frac{u_0}{2u_D+u_0} \geq \frac{p+\delta^2}{p+\delta^1}$, $\frac{u_0(u_D-u_G)}{u_G(2u_D+u_0)} \geq \frac{p+\delta^2}{p+\delta^1}$
$\{1,3\}$	$\frac{u_G(2u_D+u_0)(u_D+u_G+u_0)}{u_0(u_D-u_G)(u_G+u_0)} \geq \frac{p+\delta^1}{p+\delta^2}$, $\frac{u_G(u_G+u_0)}{u_D(u_D+u_G+u_0)} \geq \frac{p+\delta^2}{p+\delta^1}$, $\frac{u_D(u_G+u_0)}{u_G(u_D+u_G+u_0)} \geq \frac{p+\delta^2}{p+\delta^1}$
$\{1,2\}$	$\frac{2u_D+u_0}{u_0} \geq \frac{p+\delta^1}{p+\delta^2}$, $\frac{u_D(u_D+u_G+u_0)}{u_G(u_G+u_0)} \geq \frac{p+\delta^1}{p+\delta^2}$, $u_D \geq 2u_G$
$\{3,3\}$	$\frac{u_G(2u_D+u_0)}{u_0(u_D-u_G)} \geq \frac{p+\delta^1}{p+\delta^2}$, $\frac{u_G(u_D+u_G+u_0)}{u_D(u_G+u_0)} \geq \frac{p+\delta^1}{p+\delta^2}$, $2u_G \geq u_D$

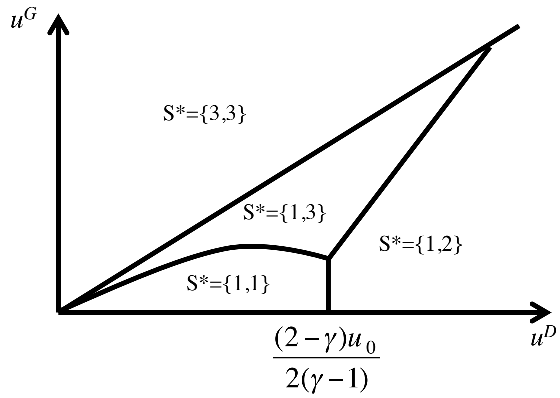

Figure 3.2. Optimality regions for $\overline{n} = 2$ with $p = 1, u_0 = 5, \delta^1 = 10$ and $\delta^2 = 3$

It is interesting to compare the revenue earned by the firm when offering two standard products which appeal to both segments equally, i.e., $S = \{3, 3\}$ or having one product dedicated to each segment, i.e., $S = \{1, 2\}$. As it turns out, offering two specialized products dominates if and only if $u_D \geq 2u_G$, that is, the attractiveness of the dedicated products to their respective segment must be at least twice that of the standard product.

Next we compare the optimal assortments with $\overline{n} = 1$ to $\overline{n} = 2$. If we superimpose the optimality regions from Figures 3.5.3 and 3.2, we obtain Figure 3.3.

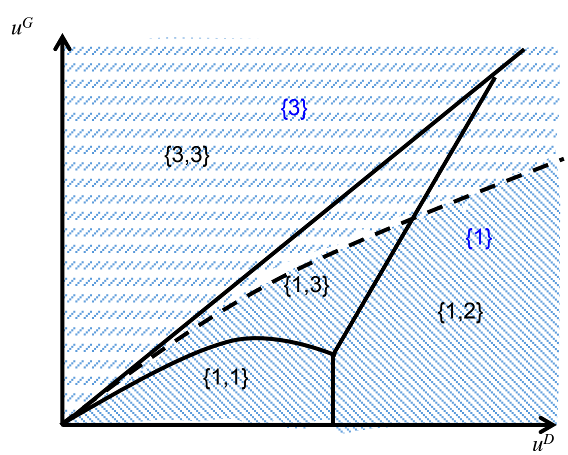

Figure 3.3. Optimality regions for $\overline{n} = 1$ and $\overline{n} = 2$

From Figure 3.3, we see that there exist values of (u_D, u_G) such that the optimal assortment of size 1 is $\{3\}$ and the optimal assortment of size 2 is $\{1, 2\}$, which implies that the firm would have to redesign its assortment, that is, getting rid of product 3 and adding products 1 and 2 when increasing the size of its offering. In other words, this example shows that adding products in greedy fashion to the assortment over time may be suboptimal.

3.6 Bounds

In this section we obtain bounds on the performance of heuristics for the two-sided market problem. First we show that the one-sided market and free-product market solutions can be arbitrarily bad when used in the setting of the two-sided market, which shows that the loss for ignoring one side of the market can be very large.

In the previous example, we see that when one solution led to a 100% optimality gap, the other was optimal. Next we investigate whether it is possible for both solutions to be suboptimal in the two-sided market problem. To do so, we compare $\Pi(x^*)$ to $\max\{\Pi(x^{*o}), \Pi(x^{*f})\}$ where $x^{*o} = \max_x \Pi^o(x)$ and $x^{*f} = \max_x \Pi^f(x)$. Note that for every x we have $\Pi(x) = \Pi^o(x) + \Pi^f(x)$. We first obtain a general bound on the percentage optimality gap measured as: $G^{of} \equiv \frac{\Pi(\mathbf{x}^*) - \max\{\Pi(\mathbf{x}^{*o}), \Pi(\mathbf{x}^{*f})\}}{\Pi(\mathbf{x}^*)}$.

Theorem 3. Under unlimited repetitions, $G^{of} = 0$ if $\frac{(N+\overline{n}-1)!}{(\overline{n})!(N-1)!} \leq 2$ and $G^{of} \leq \frac{1}{2}$ otherwise. Further, this bound is tight.

Proof. According to Proposition 3 and Corollary 4, the number of products in the optimal assortment under unlimited repetitions equals $\overline{n}$ for the *two-sided market problem* $\mathcal{P}$, the *one-sided market problem* $\mathcal{P}^o$ and the *free-product market problem* $\mathcal{P}^f$. Therefore the number of assortments which can be optimal is equal to $\binom{N+\overline{n}-1}{\overline{n}} = \frac{(N+\overline{n}-1)!}{(\overline{n})!(N-1)!}$, which corresponds to the number of ways to choose $\overline{n}$ products from a set of N options with unlimited repetitions allowed.

If $\frac{(N+\overline{n}-1)!}{(\overline{n})!(N-1)!} \leq 2$, there are at most two possible assortments to consider, therefore either x^{*o} or x^*f is optimal; hence we must have $\max\{\Pi(x^{*o}), \Pi(x^{*f})\} = \Pi(x^*)$, which implies that the optimality gap is zero. Now assume that $\frac{(N+\overline{n}-1)!}{(\overline{n})!(N-1)!} \geq 3$. First we show that G^{of} never exceeds $\frac{1}{2}$. We have $\Pi(\mathbf{x}^*) = \Pi^o(\mathbf{x}^*) + \Pi^f(\mathbf{x}^*) \leq \Pi^o(\mathbf{x}^{*o}) + \Pi^f(\mathbf{x}^{*f}) \leq 2\max\{\Pi^o(\mathbf{x}^{*o}), \Pi^f(\mathbf{x}^{*f})\}$. Therefore $G^{of} \leq \frac{1}{2}$.

Finally we show with an example that the bound is tight. Consider an example with $K = 1$, $N = 3$ and $\overline{n} = 1$, which implies that $\frac{(N+\overline{n}-1)!}{(\overline{n})!(N-1)!} = 3$. Let us fix λ, δ and u_0. Given $a, \epsilon \geq 0$, we set the other parameters are follows:

$$\begin{aligned}
u_1 &= \frac{a u_0}{\lambda\delta - a} \\
u_2 &= \frac{(a-\epsilon)u_0}{\lambda\delta(a-\epsilon)} \\
u_3 &= \frac{\epsilon u_0}{\lambda\delta - \epsilon} \\
p_1 &= \frac{\epsilon(u_1 + u_0)}{\lambda u_1} \\
p_2 &= \frac{(a-\epsilon)(u_2 + u_0)}{\lambda u_2} \\
p_3 &= \frac{a(u_3 + u_0)}{\lambda u_3}
\end{aligned}$$

With the above set of parameters, we can verify that the profit for all possible assortments are the values shown in table 3.8:

Table 3.8. Profit for all Possible Assortments (Proof for Theorem 3)

Assortment	Π^o	Π^f	Π
$(1,0,0)$	ϵ	a	$a+\epsilon$
$(0,1,0)$	$a-\epsilon$	$a-\epsilon$	$2a-2\epsilon$
$(0,0,1)$	a	ϵ	$a+\epsilon$

Assuming $\epsilon < \frac{a}{3}$, then $\mathbf{x}^{*o} = (0,0,1), \mathbf{x}^{*f} = (1,0,0)$ and $\mathbf{x}^* = (0,1,0)$. The optimality gap is $\frac{(2a-2\epsilon)-(a+\epsilon)}{(2a-2\epsilon)} = \frac{a-3\epsilon}{2a-2\epsilon}$. This value tends to $1/2$ as $\epsilon \rightarrow 0$.

$\square$

Next we investigate the optimality gap which may result from considering only one customer segment.

As the previous example illustrates, it is possible that none of the solutions to the one-segment problem is optimal in the two-sided market problem. Next we provide a bound on the performance of the best one-segment problem solution. Formally, we compare $\Pi(x^*)$ to $\max_{k=1,\ldots,K} \Pi(x^{*S_k})$ where $x^{*S_k} = \max_x \Pi^{S_k}(x)$. Note that for every x we have $\Pi(x) = \sum_{k=1}^{K} \Pi^{S_k}(x)$. We first obtain a general bound on the percentage optimality gap measured as: $G^s \equiv \frac{\Pi(\mathbf{x}^*) - \max\limits_{k\in\{1,\ldots,K\}} \Pi(\mathbf{x}^{*S_k})}{\Pi(\mathbf{x}^*)}$.

Theorem 4. Under unlimited repetitions, if either of the following two conditions hold: (1) $N \geq K$ and (2) $N \leq \overline{n} + 1$, we have $G^s \leq \frac{\min\{N,K\}-1}{\min\{N,K\}}$ and this bound is tight.

Proof. We show the result for the following two cases.

Case 1: $\min\{N, K\} = N$. Since we have considered the case $K \geq N \geq \overline{n} + 1$ in (b), we consider only the case $K \geq \overline{n} \geq N$ here. Suppose $G^s > \frac{N-1}{N}$, then

$$\begin{aligned}
\Pi(\mathbf{x}^*) &= \sum_{j=1}^{N} \sum_{k=1}^{K} \lambda\alpha^k (p_j + \delta^k) \frac{x_j^* u_j^k}{\sum_{j=1}^{N} x_j^* u_j^k + u_0^k} \\
&> N \max_{k\in\{1,\ldots,K\}} \{\Pi(\mathbf{x}^{*S_k})\} \\
&\geq \sum_{j=1}^{N} \Pi(\overline{n}\mathbf{e}^j) = \sum_{j=1}^{N} \sum_{k=1}^{K} \lambda\alpha^k (p_j + \delta^k) \frac{\overline{n} u_j^k}{\overline{n} u_j^k + u_0^k}
\end{aligned}$$

Since $N \leq \overline{n}$, we have $x_j^* \leq \overline{n}$ for $j = 1, \ldots, N$. Thus there is contradiction in the inequality above because $\frac{x_j^* u_j^k}{\sum_{j=1}^{N} x_j^* u_j^k + u_0^k} \leq \frac{x_j^* u_j^k}{x_j^* u_j^k + u_0^k} \leq \frac{\overline{n} u_j^k}{\overline{n} u_j^k + u_0^k}$ for all $j \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, K\}$. Therefore $G^s \leq \frac{N-1}{N}$.

Now we show that the bound $\frac{N-1}{N}$ is tight. Given $\lambda, \alpha^k, \delta^k, u_0^k,\ k = 1, \ldots, K$ and $a, \epsilon, \xi \geq 0$, we set the other parameters as follows:

$$\begin{aligned}
p_1 &= \frac{(N-1)(a-\xi)a}{\lambda\alpha^1\left[(N-1)a - \overline{n}\xi\right]} - \delta^1 \\
p_j &= \frac{(\overline{n}-1)(a-\epsilon)a}{\lambda\alpha^j\left[(\overline{n}-1)a - \overline{n}\epsilon\right]} - \delta^j \quad j = 2, \ldots, N \\
u_j^j &= \frac{au_0^j}{\overline{n}\left[\lambda\alpha^j(p_j+\delta^j) - a\right]} \quad j = 1, \ldots, N \\
u_1^k &= 0 \quad k = 2, \ldots, N \\
u_1^k &= \frac{\epsilon u_0^k}{\overline{n}\left[\lambda\alpha^k(p_1+\delta^k) - \epsilon\right]} \quad k = N+1, \ldots, K \\
u_j^k &= 0 \quad j = 2, \ldots, N \textit{ and } k = 1, \ldots, K \textit{ where } k \neq j
\end{aligned}$$

With the above set of parameters, we can verify that the profits for all possible assortments are the values shown in Table 3.9:

Table 3.9. Profits for all Possible Assortments (Proof for Thorem 4: Case 1)

Assortment	Π_1	Π_2	$\cdots$	Π_N	Π_{N+1}	Π_{N+2}	$\cdots$	Π_K	Π
$\overline{n}\mathbf{e}^1$	a	0	$\cdots$	0	ϵ	ϵ	$\cdots$	ϵ	$a+(K-N)\epsilon$
$\overline{n}\mathbf{e}^2$	0	a	$\cdots$	0	0	0	$\cdots$	0	a
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\overline{n}\mathbf{e}^N$	0	0	$\cdots$	a	0	0	$\cdots$	0	a
$\sum_{j=1}^N \mathbf{e}^j +$ $(\overline{n}-N)\mathbf{e}^1$	$a-\xi$	$a-\epsilon$	$\cdots$	$a-\epsilon$	b	b	$\cdots$	b	$Na+(N-1)\epsilon+$ $\xi+(K-N)b$

The values b in the table can be calculated because all parameters are set up. Since the exact value of b is not needed in our proof, we omit the calculation but only use the fact that $0 \leq b \leq \epsilon$ which is a result from Proposision 10. Obviously, we have $\mathbf{x}^{*S_k} = \overline{n}\mathbf{e}^k$ for $k = 1, \ldots, N$, $\mathbf{x}^{*S_k} = \overline{n}\mathbf{e}^1$ for $k = N+1, \ldots, K$, and $arg \max\limits_{k\in\{1,\ldots,K\}} \{\Pi(\mathbf{x}^{*S_k})\} = \overline{n}\mathbf{e}^1$. Note that for any assortment $\mathbf{x}$ with cardinality $\overline{n}$, we have $\Pi_k(\mathbf{x}) \leq a$ for $k = 1, \ldots, N$, and $\Pi^{S_k}(\mathbf{x}) \leq \epsilon$ for $k = N+1, \ldots, K$. Assuming $(K-2N+1)\epsilon - \xi \leq (N-1)a \Rightarrow$

$Na + (N-1)\epsilon + \xi + (K-N)b \geq a + (K-N)\epsilon$, then we have

$$\Pi(\mathbf{x}^*) \geq \Pi(\sum_{j=1}^{N} \mathbf{e}^j + (\overline{n} - N)\mathbf{e}^1) = Na + (N-1)\epsilon + \xi + (K-N)b$$

and

$$\Pi(\mathbf{x}^*) \leq \sum_{k=1}^{K} \Pi_k(\mathbf{x}^{*S_k}) = Na + (K-N)\epsilon$$

As $\epsilon, \xi \to 0$, we have:

$$\begin{aligned}
& \lim_{\epsilon\to 0} \frac{Na + (N-1)\epsilon + \xi + (K-N)b - \max_{k\in\{1,\ldots,K\}}\{\Pi(\mathbf{x}^{*S_k})\}}{Na + (N-1)\epsilon + \xi + (K-N)b} \\
= & \lim_{\epsilon\to 0} \frac{Na + (N-1)\epsilon + \xi + (K-N)b - [a + (K-N)\epsilon]}{Na + (N-1)\epsilon + \xi + (K-N)b} \\
= & \frac{N-1}{N} \\
& \lim_{\epsilon\to 0} \frac{Na + (K-N)\epsilon - \max_{k\in\{1,\ldots,K\}}\{\Pi(\mathbf{x}^{*S_k})\}}{Na + (K-N)\epsilon} \\
= & \lim_{\epsilon\to 0} \frac{Na + (K-N)\epsilon - [a + (K-N)\epsilon]}{Na + (K-N)\epsilon} \\
= & \frac{N-1}{N}
\end{aligned}$$

Therefore, $\lim_{\epsilon\to 0} \frac{\Pi(\mathbf{x}^*) - \max_{k\in\{1,\ldots,K\}}\{\Pi(\mathbf{x}^{*S_k})\}}{\Pi(\mathbf{x}^*)} = \frac{N-1}{N}$

Case 2: $\min\{N, K\} = K$. Suppose, $G^s > \frac{K-1}{K}$, then

$$\Pi(\mathbf{x}^*) = \sum_{k=1}^{K} \Pi_k(\mathbf{x}^*) > K \max_{k\in\{1,\ldots,K\}}\{\Pi(\mathbf{x}^{*S_k})\} \geq \sum_{k=1}^{K} \Pi(\mathbf{x}^{*S_k}) = \sum_{k=1}^{K} \Pi_k(\mathbf{x}^{*S_k}) + \sum_{k=1}^{K}\sum_{l\neq k}^{K} \Pi_l(\mathbf{x}^{*S_k})$$

which is contradiction because $\Pi_k(\mathbf{x}^*) \leq \Pi_k(\mathbf{x}^{*S_k})$ for $k = 1, \ldots, K$. Therefore $G^s \leq \frac{K-1}{K}$.

Now we show that the bound $\frac{K-1}{K}$ is tight. Consider an example where $2K \geq N \geq \overline{n} \geq K$. Given $\lambda, \alpha^k, \delta^k, u_0^k,\ k = 1, \ldots, K$ and $a, \epsilon \geq 0$, we set the other parameters as follows:

$$\begin{aligned}
p_j &= \frac{(\overline{n}-2)(a-\xi)a}{\lambda\alpha^j\left[(\overline{n}-2)a - \overline{n}\xi\right]} - \delta^j \quad j = 1, \ldots, \overline{n}-K \\
p_j &= \frac{(\overline{n}-1)(a-\epsilon)a}{\lambda\alpha^j\left[(\overline{n}-1)a - \overline{n}\epsilon\right]} - \delta^j \quad j = \overline{n}-K+1, \ldots, K \\
p_j &= a \quad j = K+1, \ldots, N \\
u_j^j &= \frac{a u_0^j}{\overline{n}\left[\lambda\alpha^1(p_j+\delta^j) - a\right]} \quad j = 1, \ldots, K \\
u_{K+j}^j &= \frac{\epsilon u_0^j}{\overline{n}\left[\lambda\alpha^1(p_j+\delta^j) - \epsilon\right]} \quad j = 1, \ldots, N-K \\
u_j^k &= 0 \quad j = 1, \ldots, K \textit{ and } k \neq j \\
u_{K+j}^k &= 0 \quad j = 1, \ldots, N-K \textit{ and } k \neq j
\end{aligned}$$

With the above set of parameters, we can verify that we have:

Table 3.10. Profits for all Possible Assortments (Proof for Thorem 4: Case 2)

Assortment	Π_1	Π_2	$\cdots$	Π_{N-K}	$\cdots$	$\Pi_{\overline{n}-K}$	$\Pi_{\overline{n}-K+1}$	$\ldots$	Π_K	Π
$\overline{n}\mathbf{e}^1$	a	0	$\cdots$	0	$\cdots$	0	0	$\cdots$	0	a
$\overline{n}\mathbf{e}^2$	0	a	$\cdots$	0	$\cdots$	0	0	$\cdots$	0	a
$\vdots$	$\vdots$	$\vdots$							$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$							$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$							$\vdots$	$\vdots$
$\overline{n}\mathbf{e}^K$	0	0	$\cdots$	0	$\cdots$	0	0	$\cdots$	a	a
$\overline{n}\mathbf{e}^{K+1}$	ϵ	0	$\cdots$	0	$\cdots$	0	0	$\cdots$	0	ϵ
$\overline{n}\mathbf{e}^{K+2}$	0	ϵ	$\cdots$	0	$\cdots$	0	0	$\cdots$	0	ϵ
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\overline{n}\mathbf{e}^N$	0	0	$\cdots$	ϵ	$\cdots$	0	0	$\cdots$	0	ϵ
$\sum_{j=1}^{K}\mathbf{e}^j + \sum_{j=1}^{\overline{n}-K}\mathbf{e}^j$	$a-\xi$	$a-\xi$	$\cdots$	$a-\xi$	$\cdots$	$a-\xi$	$a-\epsilon$	$\ldots$	$a-\epsilon$	$Ka-(\overline{n}-K)\xi$ $-(2K-\overline{n})\epsilon$

Assume that $(\epsilon, \xi) \in \{(y,z) : y \leq a, (2K-\overline{n})\epsilon + (\overline{n}-K)\xi \leq (K-1)a\}$, then we have $\mathbf{x}^{*S_k} = \overline{n}\mathbf{e}^k$ for $k = 1, \ldots, K$. And we note that for any assortment $\mathbf{x}$ with cardinality $\overline{n}$,

$\Pi^{S_k}(\mathbf{x}) \le a$ for $k = 1, \ldots, K$. Therefore

$$\Pi(\mathbf{x}^*) \ge \Pi(\sum_{j=1}^{K} \mathbf{e}^j + \sum_{j=1}^{\overline{n}-K} \mathbf{e}^j) = Ka - (\overline{n} - K)\xi - (2K - \overline{n})\epsilon$$

and

$$\Pi(\mathbf{x}^*) \le \sum_{k=1}^{K} \Pi_k(\mathbf{x}^{*S_k}) = Ka$$

As $\epsilon, \xi \to 0$, we have:

$$\begin{aligned}
& \lim_{\epsilon \to 0} \frac{Ka - (\overline{n} - K)\xi - (2K - \overline{n})\epsilon - \max\limits_{k \in \{1,\ldots,K\}} \{\Pi(\mathbf{x}^{*S_k})\}}{Ka - (\overline{n} - K)\xi - (2K - \overline{n})\epsilon} \\
= \; & \lim_{\epsilon \to 0} \frac{Ka - (\overline{n} - K)\xi - (2K - \overline{n})\epsilon - a}{Ka - (\overline{n} - K)\xi - (2K - \overline{n})\epsilon} \\
= \; & \frac{K-1}{K} \\
& \lim_{\epsilon \to 0} \frac{Ka - \max\limits_{k \in \{1,\ldots,K\}} \{\Pi(\mathbf{x}^{*S_k})\}}{Ka} \\
= \; & \lim_{\epsilon \to 0} \frac{Ka - a}{Ka} \\
= \; & \frac{K-1}{K}
\end{aligned}$$

Therefore, $\lim\limits_{\epsilon \to 0} \frac{\Pi(\mathbf{x}^*) - \max\limits_{k \in \{1,\ldots,K\}} \{\Pi(\mathbf{x}^{*S_k})\}}{\Pi(\mathbf{x}^*)} = \frac{K-1}{K}$

□

CHAPTER 4

A BEHAVIORAL ANALYSIS OF INVENTORY MANAGEMENT FOR SUBSTITUTABLE PRODUCTS

4.1 Introduction

Supply chain managers often need to decide the inventory levels for multiple items. For example, each Walmart store has over 10,000 stock-keeping units (SKUs), and the number of inventory decisions a Walmart store manager has to make on a daily basis is astounding (Boyer and Verma (2009)). In fast-fashion retailing, the store manager also need to make frequent inventory decisions for a large number of products. For example, a store manager of Zara need to decide the stock level for 11,000 articles in a given season on average (Caro and Gallien (2010)).

An supply chain manager who needs to decide order quantities of the products in his store typically checks the sales history of the products. However, the sales of products that are stocked out provide inaccurate information because the full demand is not observable in that situation. In particular, for categories with substitutable products, stock-outs lead to a double-censoring effect: lost sales are unobservable and the sales from substitution are indistinguishable from first-choice sales. The double-censoring effect can lead to a behavioral trap which we refer to as *the Vicious Cycle of Substitution*, as is shown in Figure 4.1. We explain the vicious cycle by assuming a simple scenario, i.e. an inventory manager needs to decide the inventory levels for 2 substitutable products, Product A and Product B, in a product category. If product B is out of stock in a selling period, some of the customers who come to buy B may substitute to A, which results in a large sales of A in that period. By looking at the sales data, the inventory manager may think a lot of consumers want to buy A and as a result choose to stock more of A than its true demand while less B than its true demand in the next period. Then in the next period, there are more customers who want B but find it not available. And the inventory manager will probably find a even larger sales of A for this period, which reinforces his belief in large demand of A while small demand of B. The evolution of this bias will lead to much larger stock level of A than its true demand and much smaller stock level of B than its true demand, which results in decreasing service level overtime.

Inventory decisions for substitutable products using subjective estimates is challenging (Bansal and Moritz (2015)). First, assessing the frequency with which the product substitution flexibility will be used in future period(s) is difficult for decision makers, many of whom resort to ad-hoc analysis or heuristics even in simpler contexts. For example, in the single product newsvendor problem, Schweitzer and Cachon (2000) showed that subjects made suboptimal capacity decisions under uncertain demand, and exhibited systematic biases in ordering behavior. When compared to the single product newsvendor problem, the determination of optimal inventory levels for multiple substitutable products is much more complex and difficult to implement. The additional complexity is due to the potential substitution interactions that must be considered while determining the optimal inventory levels. Typically, the demand of each product is met first using its own available capacity. Any surplus inventory is used for substitution to other products and any unmet demand is met by substitution from other products. Accurately anticipating contingent demand and supply scenarios, their frequency of occurrence, and then incorporating such information into capacity acquisition decisions is a challenging task.

In this paper, we design an experiment to study the impact of substitution on subjects' performance in a newsvendor setting where subjects need to decide inventory levels for two substitutable products.

The rest of this paper is as follows. In §4.2 we review the relevant papers. In §4.3 and §4.4 we propose hypothesis and explain the experimental design. In the last section of this chapter we discuss the experimental resutls.

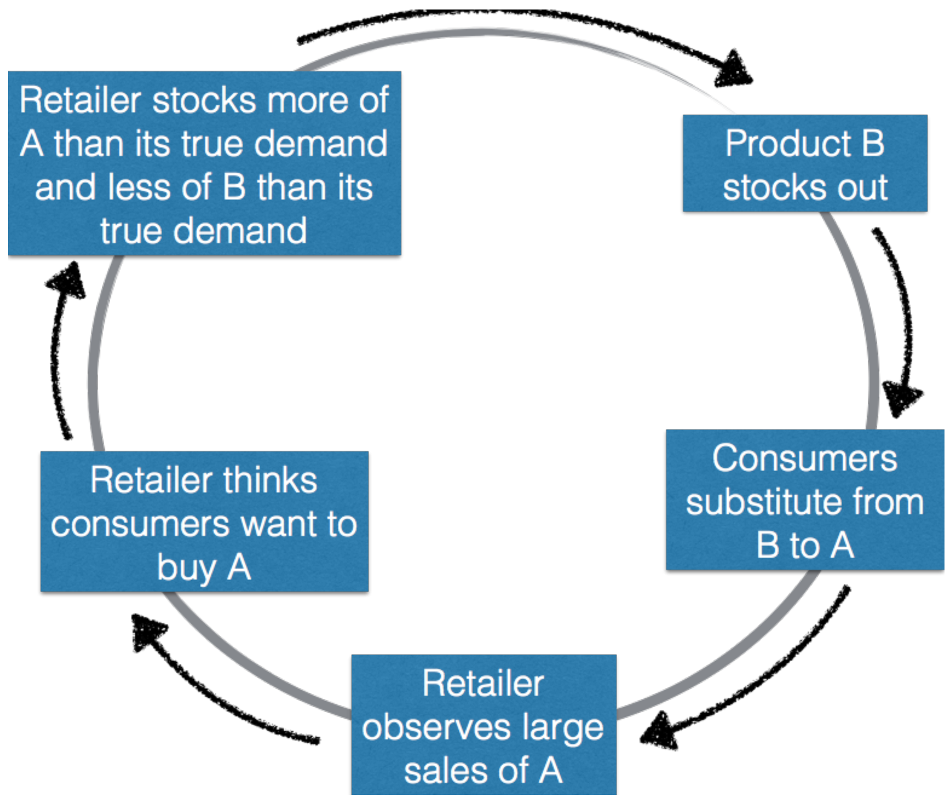

Figure 4.1. The Vicious Cycle of Substitution

4.2 Literature Review

The literatures that are related to this study are discussed below.

The newsvendor problem is a building block for many study on supply chain management and inventory management (Porteus (1990)). In the newsvendor problem, a decision maker must decide the stock level of a perishable product to maximize expected profit for a single selling period. Since the demand for the product is unknown, the decision maker need to evaluate the tradeoff between overordering and underordering. Empirical and experimental study have shown that human individuals make order quantities that systematically deviates from the normative inventory level. Schweitzer and Cachon (2000), were the first work to demonstrate this. In particular, it shows subjects' decisions exhibit pull-to-center effect, where the average order quantities lies in between the mean of the demand distribution and the exptected porfit-maximizing quantity. A large number of studies since then have illustrated the robustness of the pull-to-center result, including Bolton and Katok (2008), Bostian et al. (2008), Benzion et al. (2008), Su (2008), Katok and Wu (2009), Lurie and Swaminathan (2009), Ho et al. (2010), Kremer et al. (2010), Bolton et al. (2012) and Chen et al. (2013).

Most experimental study on newsvendor problem have been focusing on on single newsvendor decision. However, in practice, an inventory manager often need to decide stock level for more than one products. There are a few experimental studies on newsvendor problem which involves more than one newsvendor. Ho et al. (2010) studies newsvendor decisions in multiple locations where they assumes all stores have same price, cost and demand distribution. Ho et al. (2010) validates the "pull-to-center" bias and found that the biases are shown to eliminate the risk-pooling benefit when the demands across store locations are strongly correlated. In contrast, Chen and Li (2016) investigates how people make decisions dealing with 2 unrelated products. One of their main conclusions is that individuals who need to make two newsvendor decisions simultaneously perform worse than making a single newsvendor decision, which is driven by lower levels of learning and stronger demand chasing behavior. Bansal and Moritz (2015) studies how decision makers perform in a 2-product newsvendor setting, but they focus on investigating the behavioral aspects of estimation of the value of substitution. They found that subjects systematically overestimated the monetary value of product substitution and also overestimated sales from substitution.

4.3 Research Hypothesis

To study the bias on demand for substitutable products, we adopt a two-product newsvendor inventory management context in which we assume unidirectional substitution.

A retailer sells two substitutable products 1 and 2, where product 1 has smaller average demand than product 2. The demand for product 1 and 2, denoted D_1, D_2 are random variables mean μ_1, μ_2 and standard deviation σ_1, σ_2. Demand distribution for product i has probability density function $f_i(\cdot)$ with support $[0, \overline{D_i}],\ i = 1, 2$. We assume the demand for the 2 type of products are independent.The selling price for either product is p, and the cost of acquiring a unit of either product is c. The retailer must decide the inventory of product 1 and product 2 in quantities q_1, q_2 at the beggining of the selling period, i.e. before customers come to the store. After the demand for the 2 products are realized, the retailer first uses the available inventory for each product to meet the first choices, and then performs substitution if there is inventory left.

4.3.1 Analytical Model

In this study, we assume the customers only substitute from product 2 to product 1, with substitution rate α. Specifically, α is the fraction of customers who prefer product 2 but are willing to purchase product 1 in the event that product 2 stocks-out. Let S_i be the sales quantity of product $i,\ i = 1, 2$. The the retailer's expected payoff is given by: Total sales for each product are given by:

$$\begin{aligned} \mathbb{E}[\pi(q_1, q_2)] &= \mathbb{E}\left[pS_1 + pS_2 - cq_1 - cq_2\right] \\ &= p\left(\mathbb{E}[S_1] + \mathbb{E}[S_1]\right) - c\left(q_1 + q_2\right) \end{aligned}$$

where

$$\begin{aligned} S_1 &= \min\{D_1 + \alpha(D_2 - q_2)^+, q_1\} \\ S_2 &= \min\{D_2, q_2\}, \end{aligned}$$

TO ADD: nominal inventory levels for 2-product newsvendor problem.

4.3.2 Research Hypotheses

In the experiments, we set selling price for both products to be 10 and cost for both products to be 5. This makes the the 2 products have same profit margin, thus it is not beneficial for the decision maker to deviate from their guess of the true demand for either product. And the 2 products both have critical fractile value that equals 0.5. Further, we chose uniform distribution as the demand distribution for both products under which the mean of demand equals the median of demand. All together, in the last period, it is optimal for the decision maker to make order quantities that equal the mean of the effective demand distribution for either product. In addition, we find through numerical analysis that the profit for decision maker is more sensitive to order quantities under uniform demand distributions compared to bell-shaped distributions (for example, *beta* distribution). This is another reason for us to choose uniform distribution which possibly gives subjects more incentives for demand learning. These setting are for the purpose to motivate subjects to learn the true demand for each product and for us to explore the impact of substitution on subjects' biases in decision making while minimizing pull-to-center effect.

Starting with Schweitzer and Cachon (2000), numerous experimental studies have documented that newsvendor decisions tend to be biased towards the mean demand. Su (2008) explained the phenomenon by the use of the quantal response model. Bolton and Katok (2008) showed that newsvendor performance improves slowly over time with experiences. Inventory decisions for multiple substitutable products are more complicated than single product decision. Given this complexity, we expect the performance of decision makers to be worse. Our basic premise is that subjects are more prone to systematic decision biases on the 2 products. And because the demand distribution for the 2 products are unknown to subjects, the direction of bias on the 2 products should be opposite. Further, in partially censored information treatment and signal treatment where subjects are given either more information or signal, we expect their demand learning process to be different with the fully censored information treatment. Specifically, in fully censored information treatment where least information is given, subjects are most "blind" to substitution information and we expect their bias to be worse when substitution exists. We formalize these premise in the following hypotheses.

Hypothesis 1. Subjects' order quantities are biased when facing 2 products with substitution.

Hypothesis 2. The direction of the bias is for subjects to **over-estimate** demand for the low-demand product and to **under-estimate** demand for the high-demand product.

Hypothesis 3. In fully censored information treatment, subjects' bias is worse when there is substitution.

Hypothesis 4. Subjects will order higher quantities in early periods under **full censored information** than under **partially censored information** in order to facilitate learning the distribution of demand.

Hypothesis 5. When subjects are provided with stock out alarm, the total order quantity for both products are higher.

4.4 Experimental Design

To test the hypotheses, we developed a computer-based decision-making experiment implemented in SoPHIE labs.

In the experiments, subjects play the role of an inventory manager who must decide the order quantities for two types products carried in his store for consecutive 30 periods. We use the newsvendor setting with no inventory carry-over and no salvage value. Before starting the experiment, subjects were provided with written instructions followed by a quiz which tests subjects' understanding of the instructions. In the instruction, we told subjects that the number of customers who prefer product 1 is drawn from an unknown distribution with support $[0, 500]$, and the same for product 2. Subjects were also told that some customers may substitute to the other product when the product they prefer is out of stock, but were not informed of the exact demand distribution and substitution rate. Subjects were required to achieve 100% correctness for the quiz questions in order to start the experiment.

In all the experiments, subjects were shown the average sales of product 1 and 2 of the past periods and a history table which summarize each period from the past. Our experiment consisted of a 3×3 between-subjects design. In the first dimension of our experiment, we vary the substitution rate. The two products are unidirectional substitutable, where the substitution rate from high demand product to low demand product are respectively $0,\ 0.5,\ 1$ in the 3 settings. In the second dimension of our experiment, we vary the information shown in the history table. In the fully-censored information (FC) treatment, subjects were shown the sales quantity for each product and the total profit in each period. In partially-censored (PC) information treatment, in addition to the information from BL treatment, subjects can also see the breakdown of sales for both products: sales from first choice and sales from substitution. The third treatment, stockout-alarm (SOA) treatment, is a signal treatment where subjects were given the same history table as FC treatment except that when there is stockout, the according sales quantity was put in red color between exclamation marks as a warning.

In the experiments, we introduced customer drop-out to motivate subjects to learn the true demand for the two products. In the instruction, subjects were told that if a customer did not get to buy the product he prefered, he may not come back to the store and the demand for that product could as a result shrink in future periods. In all treatments, the drop-out rate for both products is 2% which is not known to subjects. The reason to have a low drop-out rate is to avoid too much demand shrink which could impact the demand learning for subjects. A summary of the expreimental design and number of number of subjects are exhibited in Table 4.1.

After completing all periods of the game, subjects were asked to write down their guess of the true demand for each of the two products. This question allows subjects to earn a bonus based on how close their guess is to the mean of the demand distribution. Subjects also participated in a post experiment survey which contains question about their risk aversion level and the strategy for their order quantity decision making.

Table 4.1. Experimental Decision and Number of Participating Subjects

Substitution rate	Treatment		
	Fully censored (FC)	Partially censored (PC)	Signal (SOA)
$\alpha = 0$	19	20	16
$\alpha = 0.5$	24	13	19
$\alpha = 1$	15	18	21
Total	58	51	56

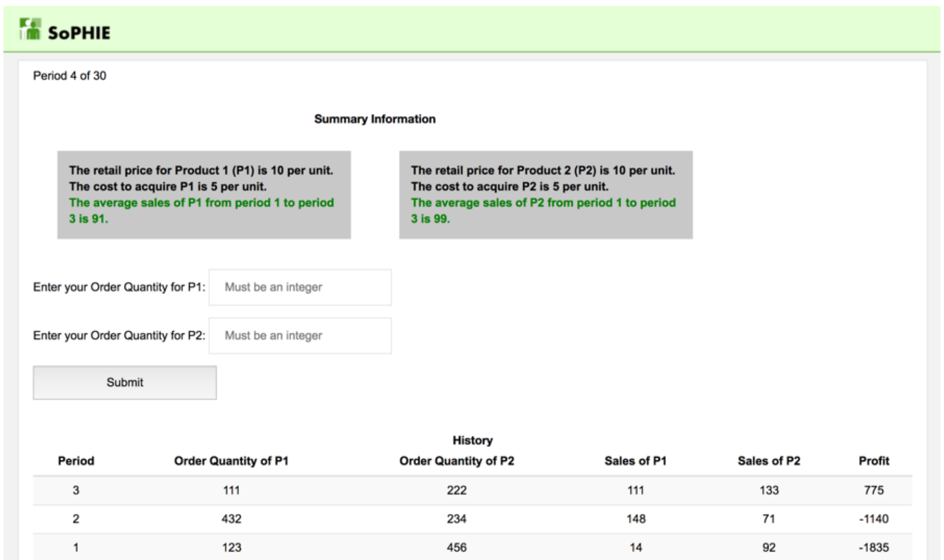

Figure 4.2. The Experiment for Fully Censored Information Treatment

All sessions were conducted at Amazon Mechanical Turk (MTurk) through the SOPHIE-MTurk integration. Figure 4.2, 4.3 and 4.4 respectively show the experiment screen in Sophie for fully censored information treatment, partially censored information treatment and signal treatment.

4.5 Results

In this section we present the results from analysis of the experimental data.

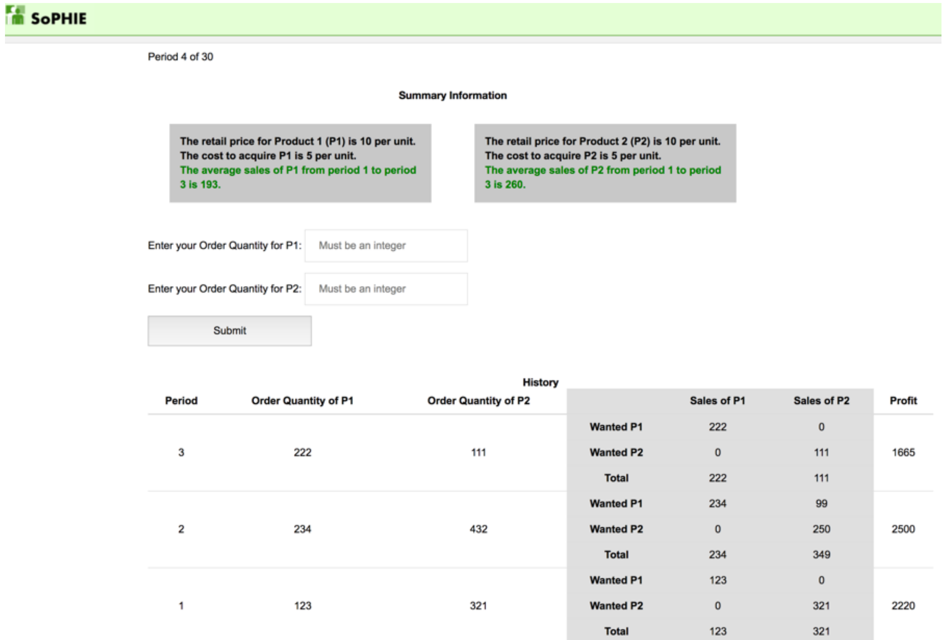

Figure 4.3. The Experiment for Partially Censored Information Treatment

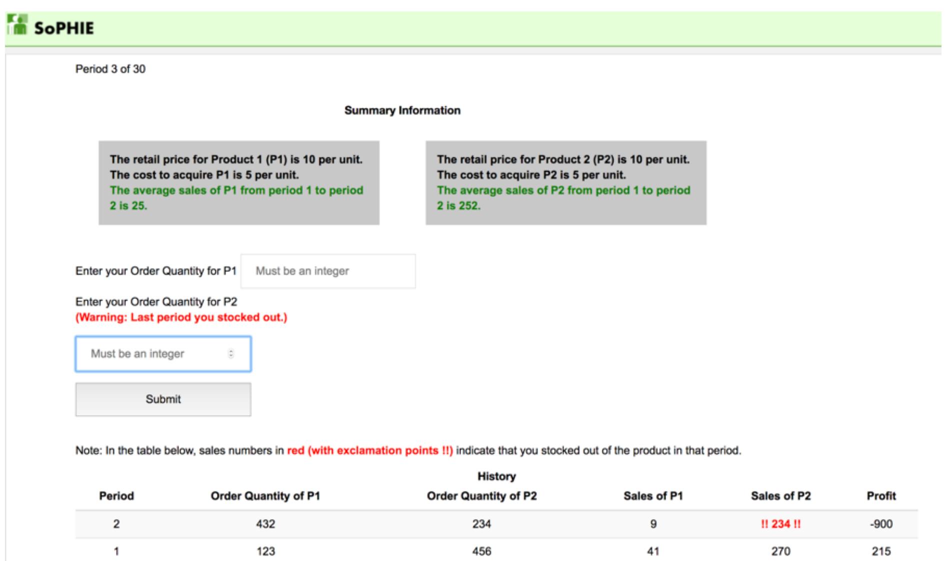

Figure 4.4. The Experiment for Signal Treatment

4.5.1 Existence and Direction of Bias

Given that the critical fractile of both products are 0.5, at the last period (i.e period 30), the subjects are supposed to put order quantities that are very close to their guess of the the average true demand. In the analysis below, unless otherwise indicated, we measure the bias as the deviation from the mean of actual demand distribution in the last period. As is introduced in previous section, in our setting, there is a drop rate that equals to 2%. As such, the demand distribution for either product is not consistent throughout the experiment, which actually happens in practice. We use the following example to illustrate how the demand distribution change is captured when we take lost demand into account.

Example 4.5.1. The original demand distributions are $Unif(0, 100)$ for product 1 , and $Unif(0, 500)$ for product 2. Drop-out rate is 2% for both products. Substitution rate from product 2 to product 1 is 0.5. The seeds for product 1 and 2 at period 1 is $(0.7, 0.6)$, then the realized demands for the 2 products are $(70, 300)$. Suppose the subjects put order quantities $(200, 200)$ in period 1, then the sales quantities are $(120, 200)$. In the sales of product 1, 70 are from first choice, and 50 are from substitution. Since the order quantity for product 2 is less than the true demand, 100 of the customers who prefer product 2 find it out of stock. Among these 100 customers, 50 of them are willing to substitute to product 1 if it is available, and 2 of them will never come back to the store in future periods. So at the beggining of period 2, the total lost demand for product 2 is 2. Suppose the seeds for product 2 in period 2 is 0.5, then the true demand for product 2 in period 2 is $(500 - 2) \times 0.5 = 249$.

Table 4.2 reports the average deviation of order quantity from the mean of actual demand distribution in the last period (period 30). As is shown, subjects' average bias for high demand product is negative in all the three treatments, and the bias is significantly smaller than 0 in fully censored and partially censored treatments. For low demand product, the bias is significantly bigger than zero in all the three treatments. This validates Hypothesis 1

Table 4.2. Deviation from Actual Average Demand in Period 30

Bias	Treatment: Fully censored (FC)	Partially censored (PC)	Signal (SOA)
High demand product	-55.23^{***}	-78.29^{***}	-6.5
	[11.53]	[10.50]	[17.05]
Low demand product	36.61^{***}	21.17^{***}	70.09^{***}
	[7.70]	[6.85]	[14.43]
Total bias	91.84^{***}	99.46^{***}	76.59^{***}
	[12.67]	[10.60]	[13.64]

Note. Standard deviation in brackets. t-test versus actual average demand.
$*$, $**$ and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

and 2. Specifically, given the results that subjects overestimates the demand for low demand product and underestimates the demand for high demand product, we measure the total bias on the 2 products as below:

$$total\ bias\ =\ bias\ of\ low\ demand\ product\ -\ bias\ of\ high\ demand\ product$$

Tables 4.2 shows that subjects are significantly biased on the 2 products in all three treatments. Table 4.3 exhibits the deviation of subjects' guess of average demand from the mean of actual demand distribution at the end of the experiment. The overestimation of low demand product and underestimation of high demand product are significant in most of the treatments. The underestimation of high demand product and overestimation of low demand product can be explained by the double-censoring effect as mentioned earlier. Since subjects do not know the exact demand distribution

4.5.2 Impact of Substitution on Demand Bias

In this section we analyze the impact of substitution on bias.

Table 4.4 and 4.5 shows the bias of high demand product and low demand product for each treatment under different substitution rate. We can see from Table 4.4 subjects

Table 4.3. Deviation of Average Demand Prediction from Actual Average Demand

Bias	Treatment		
	Fully censored (FC)	Partially censored (PC)	Signal (SOA)
High demand product	-82.55^{***}	-101.87^{***}	-66.27^{***}
	[9.40]	[8.53]	[10.57]
Low demand product	12.14^{***}	7.54	22.41^{**}
	[4.19]	[6.60]	[9.51]
Total bias	94.69^{***}	109.41^{***}	88.68^{***}
	[8.90]	[6.18]	[7.74]

Note. Standard deviation in brackets. t-test versus actual average demand.
$*$, $**$ and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

over estimates the demand for high demand product under all 3 substitution levels under fully censored information treatment and partially censored information treatment. Under fully censored information treatment where there is least information shown to subjects, the bias for high demand product is worse when there is substitution, i.e. when $\alpha > 0$. In contrast, under partially censored information treatment where the breakdown between sales is presented to subjects, the bias for high demand product is worst when there is no substitution. This suggests that when the sales of substitution is shown to subjects while in fact the 2 products are not substitutable, the subjects can get hurt from too much information. From Table 4.5 we can see that the overestimation of low demand product under fully censored information is also worse when there is substitution. Under partially censored information treatment, the bias for low demand product is -0.5, which implies that the subjects learns well from the sales details and almost not biased when substitution rate is 0. While when there is substitution, even when offering sales breakdown between first choice and substitution, the subjects have much larger bias for the low demand product. Table 4.6 reports the total bias under different substituton rate. When the total bias is measured as the bias for low demand product minus the bias for high demand product, the total bias is worse under all three treatments when there is subsitution.

Table 4.4. Deviation from Actual Average Demand in Period 30 (High Demand Product)

Substitution rate	Treatment: Fully censored (FC)	Partially censored (PC)	Signal (SOA)
$\alpha = 0$	-22.37	-84.35^{***}	45.09
	[21.43]	[21.94]	[29.77]
$\alpha = 0.5$	-82.75^{***}	-82.73^{***}	-20.58
	[17.28]	[13.73]	[26.24]
$\alpha = 1$	-49.17^{**}	-65.67^{***}	-33.07
	[19.38]	[20.26]	[30.32]

Note. Standard deviation in brackets. t-test versus actual average demand.
$*$, $**$ and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4.5. Deviation from Actual Average Demand in Period 30 (Low Demand Product)

Substitution rate	Treatment: Fully censored (FC)	Partially censored (PC)	Signal (SOA)
$\alpha = 0$	1.26	-0.5	82.72^{**}
	[4.66]	[9.01]	[36.91]
$\alpha = 0.5$	47.48^{***}	22.33^{**}	49.74^{***}
	[12.04]	[9.83]	[9.20]
$\alpha = 1$	62.53^{***}	43.49^{***}	78.89^{***}
	[18.41]	[15.63]	[25.51]

Note. Standard deviation in brackets. t-test versus actual average demand.
$*$, $**$ and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4.6. Total Deviation from the Sum of Actual Average Demand in Period 30

Substitution rate	Treatment: Fully censored (FC)	Partially censored (PC)	Signal (SOA)
$\alpha = 0$	23.63	83.83^{***}	37.63^{*}
	[20.76]	[25.44]	[27.65]
$\alpha = 0.5$	130.23^{***}	105.06^{***}	70.32^{***}
	[17.89]	[10.05]	[20.73]
$\alpha = 1$	111.70^{***}	109.36^{***}	111.95^{***}
	[19.53]	[19.49]	[21.35]

Note. Standard deviation in brackets. t-test versus actual average demand.
$*$, $**$ and $***$ denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4.7. t-Test for Deviation in Period 30: Impact of Substitution

Treatment	Product	Substitution rate ($\alpha = 0$)	Substitution rate ($\alpha > 0$)	Difference	
FC	High demand	-22.37	-70.46	48.09	$t = 1.91,\ p - value = 0.0322$
	Low demand	1.26	52.99	-51.72	$t = -4.65,\ p - value = 0.0000$
	Total bias	23.63	123.45	-99.82	$t = -4.05,\ p - value = 0.0001$
PC	High demand	-84.33	-75.42	-8.91	$t = -0.36,\ p - value = 0.3361$
	Low demand	-0.5	31.49	-31.99	$t = -2.54,\ p - value = 0.0070$
	Total bias	83.83	106.90	-23.08	$t = -0.84,\ p - value = 0.2031$
SOA	High demand	45.09	-27.14	72.23	$t = 2.01,\ p - value = 0.0264$
	Low demand	82.72	65.04	17.68	$t = -0.45,\ p - value = 0.3297$
	Total bias	37.63	92.18	-54.55	$t = -1.73,\ p - value = 0.0477$

Note. Two-sample t-test.

Further, we show the two sample t-test for bias with substitution and bias without substitution. Under fully censored information treatment, the bias for both high demand product and low demand product are significantly worse when there is substitution compared to when there is not substitution. Under partially censored information treatment, the bias for low demand product is significantly worse when there is substitution.

4.5.3 Order Quantity

Table 4.8 shows the average order quantities of all 30 periods. The order quantities are overall higher under signal treatment. Compared to signal treatment, the stock out in the other two treatments are not stressed. In case of a stock out subjects will understand they could have sold more of that product but they may not inflate these sales value enough.

Table 4.9 reports the average order quantity in the first 5 periods. Under fully censored information treatment, when subjects are provided with less information than partially censored information treatment, subject's average order quantities for both products are higher than partially censored information treatment. Further, the difference for high demand product is significant.

Figure 4.5 shows the average order quantity of high demand product and low demand product though the 30 periods for each of the 3 treatments. We see the average order

Table 4.8. Average Order Quantities for Low Demand Product and Hight Demand Product

Substitution rate	Treatment		
	Fully censored (FC)	Partially censored (PC)	Signal (SOA)
$\alpha = 0$	(60.21, 161.57)	(48.58, 124.52)	(132.26, 233.35)
$\alpha = 0.5$	(95.87, 141.39)	(85.20, 133.96)	(94.04, 173.61)
$\alpha = 1$	(100.18, 145.24)	(94.04, 119.77)	(120.96, 171.16)
All α	(85.65, 148.75)	(75.95, 126.80)	(115.05, 189.76)

Note. Note. The first number parenthesis is the average demand for low demand product, and the second number is the average demand for high demand product.

Table 4.9. t-Test for Average Order Quantity in the First 5 Periods: Fully-Censored Treatment vs. Partially-Censored Treatment

Product	Treatment		Difference	
	FC	PC		
High demand product	162.79	148.99	13.79	$t = 1.55,\ p - value = 0.0611$
Low demand product	111.36	104.26	7.10	$t = 1.05,\ p - value = 0.1477$
Total order quantity	274.15	253.25	20.90	$t = 1.57,\ p - value = 0.0528$

Note. Two-sample t-test.

quantity of high demand product under all the 3 treatments increase sharply at the early periods while the average order quantity of low demand product under all the 3 treatments show an obvious drop, which implies that the subjects start to realize the 2 products have different average demand at very early periods. In addition, the average order quantity for high demand product under signal treatment is all way closer to the average demand than fully censored information treatment and partially censored information treatment. While for low demand product, the average order quantity under partially censored information treatment is closer to average demand than the other two treatments.

4.5.4 Impact of treatment

In this section we compare subjects performance under different treatment.Table 4.10 exhibits the comparison between fully censored information treatment and partially censored

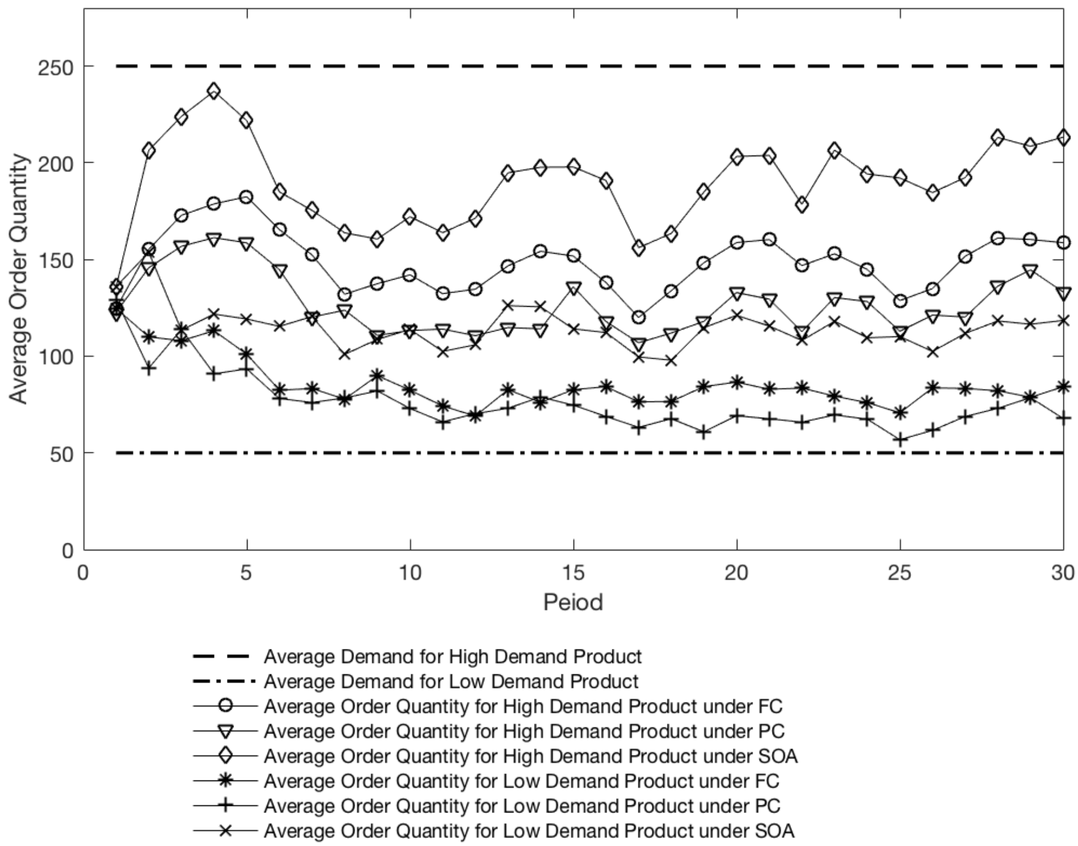

Figure 4.5. Average Order Quantity

information treatment. When subjects are provided with the details of sales, i.e. the. breakdown between first choice and substitution, the bias for high demand product is significantly worse but the bias for low demand product is significantly better. The average total bias value on the 2 products is larger in partially censored information treatment than in fully censored information treatment.

Table 4.11 compares fully censored information treatment and signal treatment. In signal treatment, the underestimation of high demand product is better while the overestimation of low demand product is worse, and the total bias on the two products is better. The data in Table 4.12 implies that the total bias on the two products is significantly better in signal treatment than in partially censored information treatment.

Table 4.10. t-Test for Deviation in Period 30: Fully-Censored Treatment vs. Partially-Censored Treatment

Bias	Treatment FC	Treatment PC	Difference	
High demand product	-55.23	-78.29	23.05	$t = 1.48,\ p - value = 0.0709$
Low demand product	36.61	21.17	15.44	$t = 1.50,\ p - value = 0.0683$
Total bias	91.84	99.46	-7.62	$t = -0.46,\ p - value = 0.3228$

Note. Two-sample t-test.

Table 4.11. t-Test for Deviation in Period 30: Fully-Censored Treatment vs. Signal Treatment

Bias	Treatment FC	Treatment SOA	Difference	
High demand product	-55.23	-6.5	-48.73	$t = -2.37,\ p - value = 0.0099$
Low demand product	36.61	70.09	-33.48	$t = -2.05,\ p - value = 0.0218$
Total bias	91.84	76.59	15.25	$t = 0.82,\ p - value = 0.2072$

Note. Two-sample t-test.

Table 4.12. t-Test for Deviation in Period 30: Partially-Censored Treatment vs. Signal Treatment

Bias	Treatment PC	Treatment SOA	Difference	
High demand product	-78.29	-6.5	-71.79	$t = -3.59,\ p - value = 0.0003$
Low demand product	21.17	70.09	-48.92	$t = -3.07,\ p - value = 0.0015$
Total bias	99.46	76.59	22.87	$t = 1.32,\ p - value = 0.0942$

Note. Two-sample t-test.

Figure 4.6 shows the average total bias under all the three treatments. Same as §**??**, the total bias is measured by bias of high demand product minus the bias of low demand product, where bias for each product is the difference between order quantity and mean of true demand. As is shown, the total bias under all the three treatments has an obvious drop in the first 6 periods. The total bias under signal treatment is smaller than the other 2 treatments in most of the periods. And the total bias when subjects are provided the breakdown in sales between first choice and substitution, the total bias is higher than when they are not offered this information in most of the periods.

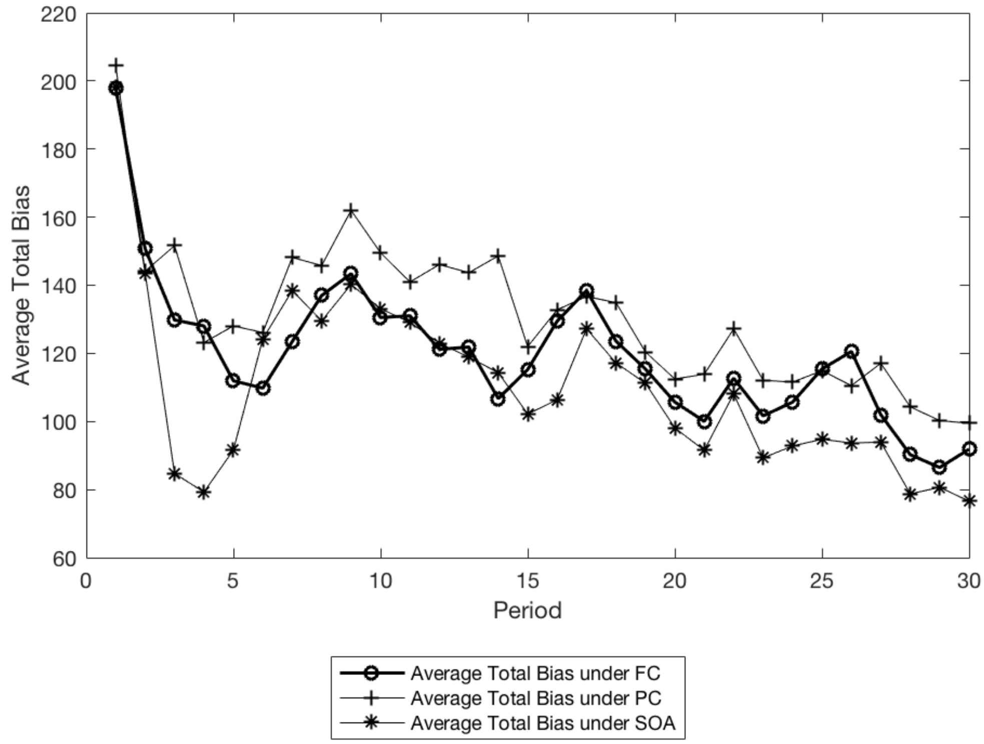

Figure 4.6. Total Bias

CHAPTER 5

CONCLUSION

5.1 Conclusion of Chapter 2

In Chapter 2, we study the assortment planning problem for a single product category when retailer faces multi-item purchasing, so called "n-pack" consumers as introduced by Fox et al (2017). With some mild assumptions on consumer behavior, Fox et al (2017) develop a choice model to obtain the expected value from consuming an n-pack, which we use as an input into our demand function for given assortment. We study the structure of the optimal assortment under two choice rules: consumers either purchase the maximum value n-pack with probability one (maximum choice rule) or they probabilistically choose between each possible n-pack using an attraction-based formula akin to the Multinomial Logit (MNL) model purchase probability equation (probabilistic choice rule). In addition, we explore how the retailer's assortment decision and total profits are impacted when the retailer ignores a key feature of the n-pack choice model, called "choice premium". Specifically, this choice premium captures the utility that consumers derive from variety in their shopping basket which allows them to hedge against future preference uncertainty. Moreover, we investigate the impact on the retailer?s optimal assortment and profits when he ignores the multi-item shopping behavior of consumers, i.e. when he assumes all consumers buys at most one unit from the product category on a store visit.

We find the structure of optimal assortment under the two choice rules. Under probabilistic choice rule, the optimal assortment is a "popular-eccentric set". In contrast, under maximum choice rule, one can only guarantee that the optimal assortment includes the most popular product. Under maximum choice rule, we propose a heuristic method for selecting an assortment and show numerically that it reaches high optimality. We also show that the optimal assortment under the probabilistic rule is a popular-eccentric set but can be different from the optimal assortment when considering choice premium. And the optimal assortment under maximum choice rule when choice premium is ignored contains only the most popular product. The numerical study shows that ignoring key features of consumer choices including choice premium and basket shopping can lead to significant profit loss.

5.2 Conclusion of Chapter 3

In this paper we analyze the assortment planning problem of a firm facing a two-sided market which receives revenues per product sold from its customers and revenues from promotion campaigns by some advertisers. Customers come from different segments which differ in their preferences for the products. The firm is allowed to offer multiple products with the same attractiveness profile and price, that is, the firm is allowed to *repeat* products in the chosen assortment. We show that the optimal assortment does not generally have a simple structure. When unlimited repetitions are allowed, we show that it may still be optimal to offer distinct product profiles in the assortment when there are multiple customer segments. In contrast, the optimal assortment contains only one product profile if there is only one customer segment. For a special case with 3 product profiles and 2 customer segments, we provide conditions under which the firm would choose to offer standard products which appeal to all segments as opposed to a dedicated products for each segment. We show that in general, the optimal assortment is not *revenue-ordered*, that is, it does not include a certain number of the products with the highest revenue. However, we provide conditions under it has a simple ranking-based structure. We also consider special cases wherein the firm ignores the revenue received from the advertisers (one-sided market) or the revenue received from the customers (free product). We show that the optimal solution to these problems can lead to a significant revenue loss in the two-sided market problem, therefore it is important for the firm to consider both streams of revenues when making assortment decisions.

In our future research we plan to further analyze the structure of the optimal assortment in the two-sided market problem and its special cases. In particular we are interested in seeing when a greedy heuristic is optimal. We also plan to propose a number of heuristic policies for choosing an assortment and study their performance numerically.

In our model we have assumed that the product selling prices are fixed. This assumption allows us to focus on the assortment dimension, which, to our knowledge has not been studied previously in the context of a two-sided market. Studying the interaction between pricing and assortment is a interesting direction for future research. We have also assumed that the firm must choose from a discrete set of possible product profiles, which have pre-set attractiveness levels to the different customer segments. In our future research we plan on studying the product *design* problem in which the firm is directly choosing the attractiveness levels of its products under some cost constraints.

5.3 Conclusion of Chapter 4

A retailer who needs to decide order quantities of the products in his store typically checks the sales history of the products. However, the sales of products that are stocked out provide inaccurate information because the full demand is not observable in that situation. In particular, for categories with substitutable products, stock-outs lead to a double-censoring effect: lost sales are unobservable and the sales from substitution are indistinguishable from first-choice sales.

We design an experiment to study the impact of substitution on subjects? performance in a newsvendor setting where subjects need to decide inventory levels for two substitutable products. Our experiment consisted of a 3×3 between-subjects design. In the first dimension of our experiment, we vary the substitution rate. In the second dimension of our experiment, we vary the information shown in the history table. In the fully-censored information (FC) treatment, subjects were shown the sales quantity for each product and the total profit in each period. In partially-censored (PC) information treatment, in addition to the information from BL treatment, subjects can also see the breakdown of sales for both products: sales from first choice and sales from substitution. The third treatment, stockout-alarm (SOA) treatment, is a signal treatment where subjects were given an alarm whenever stock out happened. The analysis of experiment data show that when deciding inventory levels for 2 substitutable products, subjects underestimate the demand for high demand product and overestimate the demand for low demand product. Moreover, the bias is worse when there is substitution in fully censored information treatment. Comparison between treatments shows that in fully censored information treatment where subjects are provided with less information compared to partially censored information treatment, subjects tend to order larger quantity in early periods in order to learn demand. Also, when given alarm on stock out, subjects's order quantities for both products are larger than the other treatments.

APPENDIX A

PROOFS FOR CHAPTER 2

Lemma 20. We have $V(k_1,...,k_M) \leq V(k_1,...,k_j+1,...,k_M)$ for all $j=1,...,M$.

Proof. From (2.2), we have $V(k_1,\ldots,k_j+1,\ldots,k_M) - V(k_1,...,k_M) = \ln\left(\frac{(n+1)!}{k_1!\ldots(k_j+1)!\ldots k_n!}\right) - \ln\left(\frac{n!}{k_1!\ldots k_j!\ldots k_n!}\right) + U_j + \gamma = \ln\frac{n+1}{k_j+1} + U_j + \gamma$ which is non-negative since $U_j \geq 0$ and $k_j \leq n$. $\square$

Lemma 21. For a given $S \subseteq \mathcal{M}$, consider n-packs $\vec{k} = (k_1,\ldots,k_M)$ and $\vec{k}' = (k_1',...,k_M') \in \mathcal{K}^n(S)$. If for $i,j \in S$ with $i<j$ (so that $U_i \geq U_j$), we have $k_i = k_j' \geq k_j = k_i'$ and $k_l' = k_l$ for all $l \neq i,j$ then $V(\vec{k}) \geq V(\vec{k}')$.

Proof.

$$\begin{aligned}
V(\vec{k}) - V(\vec{k}') &= \ln\left(\frac{n!}{k_1!\ldots k_M!}\right) + k_1U_1 + \ldots + k_iU_i + \ldots + k_jU_j + \ldots + k_MU_M - n\gamma \\
&\quad - \ln\left(\frac{n!}{k_1!\ldots k_M!}\right) + k_1U_1 + \ldots + k_jU_i + \ldots + k_iU_j + \ldots + k_MU_M - n\gamma \\
&= (k_i - k_j)(U_i - U_j) \\
&\geq 0
\end{aligned}$$

where the inequality follows from the fact that $i<j$ implies $U_i \geq U_j$. $\square$

Corollary 6. For a given assortment $S \subseteq \mathcal{M}$, $k_j^{*n}(S)$ is non-decreasing in n for all $j \in S$.

Proof. This result directly follows from the fact that the greedy procedure stated in Algorithm 1 is optimal. $\square$

Proof of Lemma 1. This result directly follows from 21.

Proof of Proposition 1. Part (i): for any $j \in \mathcal{S}$:

$$V^{*n+1}(S) = V(\vec{k}^{*n+1}) \geq V(k_1^{*n}, \ldots, k_j^{*n} + 1, \ldots, k_M^{*n}) \geq V(\vec{k}^{*n}) = V^{*n}(S).$$

where the second inequality comes from Lemma 20.

Part (ii): we have

$$\begin{aligned} V^{*n}(\hat{S}) &= V(\vec{k}^{*n}(\hat{S})) = \max_{\substack{(k_1,\ldots,k_M):\sum_{i:i\in\hat{S}} k_i = n \\ k_i = 0, i \notin \hat{S}}} V(k_1, \ldots, k_M) \\ &\leq \max_{\substack{(k_1,\ldots,k_M):\sum_{i:i\in S} k_i = n \\ k_i = 0, i \notin S}} V(k_1, \ldots, k_M) = V(\vec{k}^{*n}(S)) = V^{*n}(S). \end{aligned}$$

Part (iii):

$$\begin{aligned} V^{*n}(S) = V(\vec{k}^{*n}(S)) &= \ln\left(\frac{n!}{(k_1^{*n}(S))! \cdots (k_j^{*n}(S))! \cdots (k_M^{*n}(S))!}\right) \\ &\quad + k_1^{*n}(S)U_1 + \cdots + k_j^{*n}(S)U_j + \cdots + k_M^{*n}(S)U_M + n\gamma \\ &\leq \ln\left(\frac{n!}{(k_1^{*n}(S))! \cdots (k_j^{*n}(S))! \cdots (k_M^{*n}(S))!}\right) \\ &\quad + k_1^{*n}(S)U_1 + \cdots + k_j^{*n}(S)U_i + \cdots + k_M^{*n}(S)U_M + n\gamma \\ &\leq V^{*n}(S') \end{aligned}$$

Proof of Lemma 2. (i) The proof is by induction on n. Let $S = \{l_1, ..., l_m\}$, such that $U_{l_1} \geq U_{l_2} \geq ... \geq U_{l_m}$. First we show the result holds for $n = 1$. From Corollary 1, the maximum-value 1-pack with assortment S is such that $k_{l_1}^{*1}(S) = 1$ and $k_{l_r}^{*1}(S) = 0$ for $r = 2, \ldots, m$. There are two cases: (i) $U_i < U_{l_1}$ or (ii) $U_i \geq U_{l_1}$. In Case (i), the maximum-value n-pack with S' is such that $k_{l_1}^{*1}(S') = 1$, $k_i^{*1}(S') = 0$ and $k_{l_r}^{*1}(S') = 0$ for $r = 2, \ldots, m$. In Case (ii), it is such that $k_i^{*1}(S') = 1$ and $k_{l_r}^{*1}(S') = 0$ for $r = 1, \ldots, m$. In both cases, we have $k_{l_r}^{*1}(S') \leq k_{l_r}^{*1}(S)$ for $r = 1, \ldots, m$.

Next we show that if the result holds for n it also holds for $n+1$. From Algorithm 1, there exists $j \in S'$ such that $k_j^{*n+1}(S') = k_j^{*n}(S') + 1$ and $k_x^{*n+1}(S) = k_x^{*n}(S')$ for $x = 1, ..., M$ and

$x \neq j$ and it is such that $U_j - \ln(k_j^{*n}(S')+1) \geq U_x - \ln(k_x^{*n}(S')+1)$ for $x = 1, ..., M$ and $x \neq j$. From the induction hypothesis there are two cases: (i) j is such that $k_j^{*n}(S') < k_j^{*n}(S)$ or (ii) j is such that $k_j^{*n}(S') = k_j^{*n}(S)$. In Case (i), we have $k_j^{*n+1}(S') \leq k_j^{*n}(S') + 1 \leq k_j^{*n}(S) \leq k_j^{*n+1}(S)$ so the result holds for $n+1$. In Case (ii), we show that $k_j^{*n+1}(S) = k_j^{*n}(S) + 1$ which would then imply that $k_l^{*n+1}(S') \leq k_l^{*n+1}(S)$ for all $l \in S$.

Suppose not, then let t be such that $k_{l_t}^{*n+1}(S) = k_{l_t}^{*n}(S) + 1$ and $k_x^{*n+1}(S) = k_x^{*n}(S)$ for $x = 1, ..., M$ and $x \neq l_t$. This implies that $U_{l_t} - \ln(k_{l_t}^{*n}(S) + 1) > U_j - \ln(k_j^{*n}(S) + 1) = U_j - \ln(k_j^{*n}(S') + 1) \geq U_{l_t} - \ln(k_{l_t}^{*n}(S') + 1)$ where the first and third inequality come from Algorithm 1 for assortments S and S' respectively and the second one is from the definition of Case (ii). However, from the induction hypothesis, we know that $k_{l_t}^{*n}(S) \geq k_{l_t}^{*n}(S')$, which implies that $U_{l_t} - \ln(k_{l_t}^{*n}(S) + 1) \leq U_{l_t} - \ln(k_{l_t}^{*n}(S') + 1)$. Hence we have a contradiction.

(ii) The proof for this part is similar to the proof for part (i) and therefore omitted.

Proof of Lemma 3. (i) First we show that $\sum_{l \in S} D_l(S) \leq \sum_{l \in S \cup \{i\}} D_l(S \cup \{i\})$. Under the probabilistic decision rule, we have

$$\begin{aligned}
\sum_{l \in S} D_l(S) &= \sum_{l \in S} \lambda \left(\sum_{\vec{k} \in \mathcal{K}^n(S)} k_l \frac{V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
&= \lambda \sum_{l \in S} \left(\sum_{\vec{k} \in \mathcal{K}^n(S)} k_l \frac{V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
&= \lambda \left(\sum_{\vec{k} \in \mathcal{K}^n(S)} \frac{V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \left(\sum_{l \in S} k_l \right) \right) \\
&= \lambda n \left(\frac{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
&< \lambda n \left(\frac{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + \sum_{\substack{\vec{k} \in \mathcal{K}^n(S \cup \{i\}) \\ k_i > 0}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + \sum_{\substack{\vec{k} \in \mathcal{K}^n(S \cup \{i\}) \\ k_i > 0}} V(\vec{k}) + V_0} \right) \\
&= \lambda n \left(\frac{\sum_{\vec{k} \in \mathcal{K}^n(S \cup \{i\})} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S \cup \{i\})} V(\vec{k}) + V_0} \right) \\
&= \sum_{l \in S \cup \{i\}} D_l(S \cup \{i\})
\end{aligned}$$

Under the maximum choice rule, we have

$$\begin{aligned}
\sum_{l \in S} D_l(S) &= \sum_{l \in S} \lambda k_l^{*n}(S) \mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) \\
&= \lambda \sum_{l \in S} k_l^{*n}(S) \mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) \\
&= \lambda n \mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) \\
&\leq \lambda n \mathbb{1}(V(\vec{k}^{*n}(S \cup \{i\})) \geq V_0) \\
&= \sum_{l \in S \cup \{i\}} D_l(S \cup \{i\})
\end{aligned}$$

where the inequality follows from $V(\vec{k}^{*n}(S \cup \{i\})) \geq V(\vec{k}^{*n}(S))$ by Proposition 1.

(ii) Second, we show that $\sum_{l\in S} D_l(S) \leq \sum_{l\in S\setminus\{j\}\cup\{i\}} D_l(S\setminus\{j\}\cup\{i\})$. Denote assortment $S' = S \setminus \{j\} \cup \{i\}$. Under the probabilistic decision rule, we have

$$\begin{aligned}
\sum_{l\in S} D_l(S) &= \sum_{l\in S} \lambda \left(\sum_{\vec{k}\in\mathcal{K}^n(S)} k_l \frac{V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
&= \lambda \sum_{\vec{k}\in\mathcal{K}^n(S)} \frac{V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k}) + V_0} \left(\sum_{l\in S} k_l \right) \\
&= \lambda n \left(\frac{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
&= \lambda n \left(\frac{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j>0}} V(\vec{k})}{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j>0}} V(\vec{k}) + V_0} \right) \\
&\leq \lambda n \left(\frac{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i>0}} V(\vec{k})}{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i>0}} V(\vec{k}) + V_0} \right) \\
&= \lambda n \left(\frac{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i>0}} V(\vec{k})}{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i>0}} V(\vec{k}) + V_0} \right) \\
&= \lambda n \left(\frac{\sum_{\vec{k}\in\mathcal{K}^n(S')} V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S')} V(\vec{k}) + V_0} \right) \\
&= \sum_{l\in S'} D_l(S')
\end{aligned}$$

Under the maximum choice rule, we have

$$\begin{aligned}
\sum_{l\in S} D_l(S) &= \sum_{l\in S} \lambda k_l^{*n}(S) \mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) \\
&= \lambda n \mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) \\
&\leq \lambda n \mathbb{1}(V(\vec{k}^{*n}(S')) \geq V_0) \\
&= \sum_{l\in S'} D_l(S')
\end{aligned}$$

where, in both cases, the inequality follows from $V(\vec{k}^{*n}(S')) \geq V(\vec{k}^{*n}(S))$ by Proposition 1.

(iii) Next we prove that the total demand $\sum_{l\in S} D_l(S)$ is nondecreasing and concave in U_j for all $j \in S$ under probabilistic choice rule. Let $A^n = \sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k})$. Then total demand under the probabilistic choice rule can be written as:

$$\sum_{l\in S} D_l(S) = \lambda n \left(\frac{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) = \lambda n \left(\frac{A^n}{A^n + V_0} \right)$$

For $n \in \mathbb{N}^+$ and for all $j \in S$, A^n is positive and linear increasing in U_j, therefore we have $\frac{\partial A^n}{\partial U_j} \geq 0$ and $\frac{\partial^2 A^n}{\partial (U_j)^2} = 0$. Thus:

$$\begin{aligned}
\frac{\partial \left[\sum_{l\in S} D_l(S)\right]}{\partial U_j} &= \lambda n \frac{\frac{\partial A^n}{\partial U_j} V_0}{\left[A^n + V_0\right]^2} \geq 0 \\
\frac{\partial^2 \left[\sum_{l\in S} D_l(S)\right]}{\partial (U_j)^2} &= \lambda n \frac{-2\left(\frac{\partial A^n}{\partial U_j}\right)^2 V_0}{\left[A^n + V_0\right]^3} \leq 0
\end{aligned}$$

which shows that $\sum_{l\in S} D_l(S)$ is increasing and concave in U_j for all $j \in S$.

(iv) Finally we prove the the last part of the result. We have $\sum_{l\in S} D_l(S) = \lambda n \mathbb{1}(V^{*n}(S) \geq V_0)$. By Proposition 1 (iii), we know that $V^{*n}(S)$ is increasing in U_j for all $j \in S$. This means for each $j \in S$ there exists a value $\overline{U}_j$ such that total demand is zero for $U_j < \overline{U}_j$ and equal to λn for$U_j \geq \overline{U}_j$.

Proof of Lemma 4 . Under the probabilistic decision rule,

$$\begin{aligned}
& D_i(S) - D_j(S) \\
= \; & \lambda \left(\sum_{\vec{k} \in \mathcal{K}^n(S)} k_i \frac{V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
& - \lambda \left(\sum_{\vec{k} \in \mathcal{K}^n(S)} k_j \frac{V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \right) \\
= \; & \lambda \sum_{q_i=0}^{n} \sum_{q_j=0}^{n-q_i} q_i \frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i=q_i, k_j=q_j}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \\
& - \lambda \sum_{q_i=0}^{n} \sum_{q_j=0}^{n-q_i} q_j \frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i=q_i, k_j=q_j}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \\
= \; & \lambda \sum_{q_i=0}^{n} \sum_{q_j=0}^{n-q_i} (q_i - q_j) \frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i=q_i, k_j=q_j}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \\
= \; & \lambda \sum_{q_i-q_j=-n}^{n} (q_i - q_j) \frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i=q_i, k_j=q_j}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \\
= \; & \lambda \sum_{q=-n}^{n} q \frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i-k_j=q}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \\
= \; & \lambda \sum_{q=1}^{n} q \left[\frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i-k_j=q}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} - \frac{\sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i-k_j=-q}} V(\vec{k})}{\sum_{\vec{k} \in \mathcal{K}^n(S)} V(\vec{k}) + V_0} \right] \\
\geq \; & 0
\end{aligned}$$

The last inequality holds because according to Lemma 21, for all $i = 1, \ldots, n$:

$$\begin{aligned}
& \sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i-k_j=q}} V(\vec{k}) - \sum_{\substack{\vec{k} \in \mathcal{K}^n(S) \\ k_i-k_j=-q}} V(\vec{k}) \\
= \; & \sum_{\substack{\vec{k}, \vec{k}' \in \mathcal{K}^n(S) \\ k_l = k'_l \forall l \neq i,j \\ k_i - k_j = k'_j - k'_i = q}} \left[V(\vec{k}) - V(\vec{k}') \right] \\
\geq \; & 0
\end{aligned}$$

Under the maximum choice rule,

$$D_i(S) - D_j(S) = \lambda(k_i^{*n}(S) - k_j^{*n}(S))\mathbb{1}(V(\vec{k}^{*n}(S)) \geq V_0) \geq 0$$

where the inequality comes from Corollary 1.

Proof of Lemma 5. First we prove result (i).For $l \in S$, we have

$$\begin{aligned} D_l(S) &= \lambda k_l^{*n}(S)\mathbb{1}(V^{*n}(S) \geq V_0) \\ &= \lambda k_l^{*n}(S) \\ &\geq \lambda k_l^{*n}(S \cup \{i\}) \\ &= \lambda k_l^{*n}(S \cup \{i\})\mathbb{1}(V^{*n}(S \cup \{i\}) \geq V_0) = D_l(S \cup \{i\}) \end{aligned}$$

where the second equality is because $\sum_{l \in S} D_l(S) = \lambda n$ implies that $V^{*n}(S) \geq V_0$, the inequality follows from Lemma 2 (ii) and the second to last equality is because by Proposition 1 (ii), $V^{*n}(S \cup \{i\}) \geq V^{*n}(S)$ which implies that $V^{*n}(S \cup \{i\}) \geq V_0$. Result (ii) follows directly from Lemma 2 (ii).

Lemma 22. Consider $S \subseteq \mathcal{M}$ and $S' = S \backslash \{j\} \cup \{i\}$ such that $U_i \geq U_j$, we have $D_i(S') \geq D_j(S)$.

Proof. Under the maximum choice rule, we have

$$\begin{aligned} D_j(S) &= \lambda \mathbb{1}(V^{*n}(S) \geq V_0)k_j^{*n}(S) \\ &\leq \lambda \mathbb{1}(V^{*n}(S') \geq V_0)k_j^{*n}(S) \\ &\leq \lambda \mathbb{1}(V^{*n}(S') \geq V_0)k_i^{*n}(S') = D_i(S') \end{aligned}$$

where the first inequality follows from Proposition 1 (iii) as $V^{*n}(S) \leq V^{*n}(S')$ and the second inequality follows from Lemma 2 (ii) as $k_i^{*n}(S') \geq k_j^{*n}(S)$.

Under the probabilistic choice rule, we have

$$
\begin{aligned}
D_j(S) &= \lambda\left(\sum_{\vec{k}\in\mathcal{K}^n(S)}\frac{V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S)}V(\vec{k})+V_0}k_j\right)\\
&\leq \lambda\left(\sum_{\vec{k}\in\mathcal{K}^n(S')}\frac{V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S')}V(\vec{k})+V_0}k_i\right)\\
&= D_i(S')
\end{aligned}
$$

where the inequality follows from the fact that $\sum_{\substack{\vec{k}\in\mathcal{K}^n(S')\\ k_i=t}} V(\vec{k}) \geq \sum_{\substack{\vec{k}\in\mathcal{K}^n(S)\\ k_j=t}} V(\vec{k})$ for all $t = 1,\dots,n$ by Lemma 21. $\square$

Definition 5. We say vector $\vec{x} = (x_1, ..., x_M)$ **weakly majorizes** vector $\vec{y} = (y_1, ..., y_M)$, denoted by $\vec{x} \succ^W \vec{y}$, if $\sum_{i=1}^{j} x_{[i]} \geq \sum_{i=1}^{j} y_{[i]}$ for $j = 1, ..., M$, where $x_{[i]}$ denotes the i-th highest component of vector $\vec{x}$.

Note that an important property of the weak majorization order is that if $\vec{x} \succ^W \vec{y}$ and $\vec{x}' \succ^W \vec{y}'$ then $\vec{x} + \vec{x}' \succ^W \vec{y} + \vec{y}'$.

Lemma 23. If $\vec{D}(S) \succ^W \vec{D}(S')$ then $\Pi(S) \geq \Pi(S')$.

Proof. Marshall et al. (2011) use a stronger definition of majorization as follows: they say vector $\vec{x} = (x_1, ..., x_M)$ *majorizes* vector $\vec{y} = (y_1, ..., y_M)$, denoted by $\vec{x} \succ_M \vec{y}$, if $\sum_{i=1}^{M} x_i = \sum_{i=1}^{M} y_i'$ and $\sum_{i=1}^{j} x_{[i]} \geq \sum_{i=1}^{j} y_{[i]}'$ for $j = 1, ..., M-1$. Further, they prove the following result: if a function $g(\cdot)$ is convex and $\vec{x} \succ_M \vec{y}$ then $\sum_{j=1}^{M} g(x_i) \geq \sum_{j=1}^{M} g(y_i)$ (also see Lemma 2 in **?**).

Let $\Delta = \sum_{l=1}^{M} D_l(S) - \sum_{l=1}^{M} D_l(S')$, $\delta_1 = \min\left\{D_{[1]}(S) - D_{[1]}(S'), \Delta\right\}$ and $\delta_j = \min\left\{\sum_{l=1}^{j} D_{[l]}(S) - \sum_{l=1}^{j} D_{[l]}(S') - \sum_{l=1}^{j-1}\delta_l, \Delta - \sum_{l=1}^{j-1}\delta_l\right\}$ for $j = 2, \dots, M$. We construct a demand vector $\hat{\vec{D}} = (\hat{D}_1, \dots, \hat{D}_M)$ such that $\hat{D}_j = D_{[j]}(S) - \delta_j \geq 0$ for $j = 1, \dots, M$. It is easy to verify that $\hat{\vec{D}} \succ_M \vec{D}(S')$. Setting $g(x_i) = (p-c)x_i - \sigma(x_i)^\eta$ in (2.1), it is easy to show that $g(x_i)$ is convex in x_i for $i = 1, \ldots, M$, therefore $\pi(\hat{\vec{D}}) = \sum_{j=1}^{M} g(\hat{D}_j) \geq \pi(\vec{D}(S')) = \sum_{j=1}^{M} g(D_j(S'))$. Finally, from our assumption that the profit function $\pi(\vec{D})$ is increasing in D_j for all $j = 1, \ldots, M$, we have $\Pi(S) = \pi(\vec{D}(S)) \geq \pi(\hat{\vec{D}}) \geq \pi(\vec{D}(S')) = \Pi(S')$. $\square$

Proof of Proposition 3 Let $S^* = \{l_1, ..., l_m\}$ such that $U_{l_1} \geq U_{l_2} \geq ... \geq U_{l_m}$. Suppose (contradiction) that $1 \notin S^*$. Consider alternative assortment $S' = S^* \setminus \{l_1\} \cup \{1\}$. We prove that $\vec{D}(S') \succ^W \vec{D}(S^*)$.

First we show the result under maximum choice rule. From Lemma 4 we have that $D_{l_1}(S^*) \geq D_{l_2}(S^*) \geq ... \geq D_{l_m}(S^*)$ and $D_1(S') \geq D_{l_2}(S') \geq ... \geq D_{l_m}(S')$ therefore, to prove that $\vec{D}(S') \succ^W \vec{D}(S^*)$, we show that $\sum_{j=1}^{r} D_{l_j}(S) \leq D_1(S') + \sum_{j=2}^{r} D_{l_j}(S')$ for $r = 1, ..., m$. Using (2.3), for $r = 1, ..., m$,

$$
\begin{aligned}
\sum_{j=1}^{r} D_{l_j}(S^*) &= \lambda \mathbb{1}(V^{*n}(S^*) \geq V_0) \sum_{j=1}^{r} k_{l_j}^{*n}(S^*) \\
&\leq \lambda \mathbb{1}(V^{*n}(S') \geq V_0) \sum_{j=1}^{r} k_{l_j}^{*n}(S^*) \\
&\leq \lambda \mathbb{1}(V^{*n}(S') \geq V_0) \left(k_1^{*n}(S') + \sum_{j=2}^{r} k_{l_j}^{*n}(S') \right) \\
&= D_1(S') + \sum_{j=2}^{r} D_{l_j}(S')
\end{aligned}
$$

where the first inequality is from Proposition 1 (iii) as we have that $V^{*n}(S') \geq V^{*n}(S^*)$. And the second inequality is from Lemma 2 (ii): because we have $k_1^{*n}(S') \geq k_{l_1}^{*n}(S^*)$ and $k_{l_j}^{*n}(S') \leq k_{l_j}^{*n}(S^*)$ for $j = 2, \ldots, m$, we must have $\sum_{j=1}^{r} k_{l_j}^{*n}(S^*) \leq k_1^{*n}(S') + \sum_{j=2}^{r} k_{l_j}^{*n}(S)$ for $r = 1, ..., m$. This proves that $\vec{D}(S') \succ^W \vec{D}(S^*)$, from which we get the result using Lemma 23.

Under probabilistic rule, by Lemma 22 we have $D_1(S') \geq D_{l_1}(S^*)$. And for $r = 2, \ldots, m$ we have:

$$\begin{aligned}
D_1(S') + \sum_{j=2}^{r} D_j(S') &= \lambda \frac{\sum_{\vec{k}\in\mathcal{K}^n(S')}(k_1 + \sum_{j=2}^{r} k_j)V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S')} V(\vec{k}) + V_0} \\
&= \lambda \frac{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S') \\ k_1=0}} (\sum_{j=2}^{r} k_j)V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S') \\ k_1>0}} (k_1 + \sum_{j=2}^{r} k_j)V(\vec{k})}{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S') \\ k_1=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S') \\ k_1>0}} V(\vec{k}) + V_0} \\
&= \lambda \frac{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S^*) \\ k_{l_1}=0}} (\sum_{j=2}^{r} k_j)V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S') \\ k_1>0}} (k_1 + \sum_{j=2}^{r} k_j)V(\vec{k})}{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S^*) \\ k_{l_1}=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S') \\ k_1>0}} V(\vec{k}) + V_0} \\
&\geq \lambda \frac{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S^*) \\ k_{l_1}=0}} (\sum_{j=2}^{r} k_j)V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S^*) \\ k_{l_1}>0}} (k_{l_1} + \sum_{j=2}^{r} k_j)V(\vec{k})}{\sum_{\substack{\vec{k}\in\mathcal{K}^n(S^*) \\ k_{l_1}=0}} V(\vec{k}) + \sum_{\substack{\vec{k}\in\mathcal{K}^n(S^*) \\ k_{l_1}>0}} V(\vec{k}) + V_0} \\
&= \lambda \frac{\sum_{\vec{k}\in\mathcal{K}^n(S^*)}(k_{l_1} + \sum_{j=2}^{r} k_j)V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S^*)} V(\vec{k}) + V_0} \\
&= \sum_{j=1}^{r} D_j(S^*)
\end{aligned}$$

where the inequality follows from the fact that $V(\vec{k}) \geq V(\vec{k}')$ for all $\vec{k}$ and $\vec{k}'$ such that $k_1 = k'_{l_1}$, $k_{l_j} = k'_{l_j}$ for $j = 2, \ldots, m$, and $k_l = k'_l = 0$ for $l \notin S^* \cup S'$ by Lemma 21. So we have $\vec{D}(S') \succ^W \vec{D}(S^*)$, from which we get the result using Lemma 23.

Proof of Lemma 6. When adding a product indexed i to an assortment S, the total profit is:

$$\begin{aligned}
\Pi(S \cup \{i\}) &= (p - c) \sum_{l \in S\cup\{i\}} D_l(S \cup \{i\}) - \sigma \sum_{l \in S\cup\{i\}} [D_l(S \cup \{i\})]^\eta \\
&= (p - c) \sum_{j \in S\cup\{i\}} \lambda \left(\sum_{\vec{k}\in\mathcal{K}^n(S\cup\{i\})} k_j \frac{V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S\cup\{i\})} V(\vec{k}) + V_0} \right) \\
&\quad -\sigma \sum_{j \in S\cup\{i\}} \left[\lambda \left(\sum_{\vec{k}\in\mathcal{K}^n(S\cup\{i\})} k_j \frac{V(\vec{k})}{\sum_{\vec{k}\in\mathcal{K}^n(S\cup\{i\})} V(\vec{k}) + V_0} \right) \right]^\eta
\end{aligned}$$

For $j \in S \cup \{i\}$, define

$$
\begin{aligned}
A_j(U_i) &= \sum_{\vec{k} \in \mathcal{K}^n(S \cup \{i\})} k_j V(\vec{k}) = \sum_{\vec{k} \in \mathcal{K}^n(S \cup \{i\})} k_j \left[\ln \left(\frac{n!}{k_1! \ldots k_M!} \right) + \sum_{l=1}^{M} k_l U_l + n\gamma \right] \\
B(U_i) &= \sum_{\vec{k} \in \mathcal{K}^n(S \cup \{i\})} V(\vec{k}) + V_0 = \sum_{\vec{k} \in \mathcal{K}^n(S \cup \{i\})} \left[\ln \left(\frac{n!}{k_1! \ldots k_M!} \right) + \sum_{l=1}^{M} k_l U_l + n\gamma \right] + V_0
\end{aligned}
$$

It is easy to verify that $B(U_i)$ and $A_j(U_i)$, $j \in S \cup \{i\}$, are linear functions in U_i. Further, we have:

$$
\begin{aligned}
\Pi(S \cup \{i\}) &= (p-c)\lambda \sum_{j \in S \cup \{i\}} \frac{A_j(U_i)}{B(U_i)} - \sigma \lambda^\eta \sum_{j \in S \cup \{i\}} \left[\frac{A_j(U_i)}{B(U_i)} \right]^\eta \\
&= \frac{(p-c)\lambda \sum_{j \in S \cup \{i\}} A_j(U_i) - \sigma \lambda^\eta \sum_{j \in S \cup \{i\}} [A_j(U_i)]^\eta [B(U_i)]^{1-\eta}}{B(U_i)} \qquad \text{(A.1)}
\end{aligned}
$$

Define $g(U_i) = (p-c)\lambda \sum_{j \in S \cup \{i\}} A_j(U_i) - \sigma \lambda^\eta \sum_{j \in S \cup \{i\}} [A_j(U_i)]^\eta [B(U_i)]^{1-\eta}$ so that $\Pi(S \cup \{i\}) = \frac{g(U_i)}{B(U_i)}$. According to [Greenberg and Pierskalla (1971)], $\Pi(S \cup \{i\})$ is quasiconvex in U_i if the following two conditions hold: (i) $B(U_i)$ is linear in U_i and $B(U_i) \geq 0$ for all $U_i \geq 0$; (ii) $g(U_i)$ is convex in U_i. Condition (i) holds as $V(\vec{k})$ is linear in U_i. We will prove $g(U_i)$ is convex in U_i by showing that $g''(U_i) \geq 0$. We have:

$$
\begin{aligned}
g'(U_i) &= (p-c)\lambda \sum_{j \in S \cup \{i\}} \frac{\partial A_j(U_i)}{\partial U_i} \\
&\quad -\sigma \lambda^\eta \left[\sum_{j \in S \cup \{i\}} \eta [A_j(U_i)]^{\eta-1} \frac{\partial A_j(U_i)}{\partial U_i} \right] [B(U_i)]^{1-\eta} \\
&\quad -\sigma(1-\eta)\lambda^\eta \left[\sum_{j \in S \cup \{i\}} [A_j(U_i)]^\eta \right] [B(U_i)]^{-\eta} \frac{\partial B(U_i)}{\partial U_i}
\end{aligned}
$$

And we have:

$$\begin{aligned}
g''(U_i) &= \sigma\eta(1-\eta)\lambda^\eta \left[\sum_{j\in S\cup\{i\}} (A_j(U_i))^{\eta-2}\left(\frac{\partial A_j(U_i)}{\partial U_i}\right)^2\right][B(U_i)]^{1-\eta} \\
&\quad -2\sigma\eta(1-\eta)\lambda^\eta\left[\sum_{j\in S\cup\{i\}} [A_j(U_i)]^{\eta-1}\frac{\partial A_j(U_i)}{\partial U_i}\right][B(U_i)]^{-\eta}\frac{\partial B(U_i)}{\partial U_i} \\
&\quad +\sigma\eta(1-\eta)\lambda^\eta\left[\sum_{j\in S\cup\{i\}} (A_j(U_i))^{\eta}\right][B(U_i)]^{-\eta-1}\left[\frac{\partial B(U_i)}{\partial U_i}\right]^2 \\
&= \sigma\eta(1-\eta)\lambda^\eta[B(U_i)]^{-\eta-1}\sum_{j\in S\cup\{i\}}[A_j(U_i)]^\eta \cdot \\
&\quad \left[\left(\frac{\partial A_j(U_i)}{\partial U_i}\frac{B(U_i)}{A_j(U_i)}\right)^2+\left(\frac{\partial B(U_i)}{\partial U_i}\right)^2-2\left(\frac{\partial A_j(U_i)}{\partial U_i}\frac{B(U_i)}{A_j(U_i)}\right)\left(\frac{\partial B(U_i)}{\partial U_i}\right)\right] \\
&= \sigma\eta(1-\eta)\lambda^\eta[B(U_i)]^{-\eta-1}\sum_{j\in S\cup\{i\}}[A_j(U_i)]^\eta\left[\frac{\partial A_j(U_i)}{\partial U_i}\frac{B(U_i)}{A_j(U_i)}-\frac{\partial B(U_i)}{\partial U_i}\right]^2 \geq 0
\end{aligned}$$

So we have $g''(U_i) \geq 0$, which implies that $g(U_i)$ is convex in U_i. Thus $\Pi(S\cup i) = \frac{g(U_i)}{f(U_i)}$ is quasiconvex in U_i.

Proof of Proposition 4. Suppose S^* is not a popular-eccentric set, then there exist products $j \in S^*$ and $i, l \in \mathcal{M} \setminus S^*$ such that $U_l < U_j < U_i$. From Lemma 6, we must have either $\Pi((S^* \setminus \{j\}) \cup \{l\}) > \Pi(S^*)$ or $\Pi((S^* \setminus \{j\}) \cup \{i\}) > \Pi(S^*)$ or both, which contradicts that S^* is optimal. Therefore, S^* is popular-eccentric set. When $n = 1$, $D_j(S) = \lambda\left(\frac{U_j}{\sum_{l\in S} U_l + V_0}\right)$. In comparison the MNL model would predict a demand for product j equal to $D_j(S) = \lambda\left(\frac{e^{U_j}}{\sum_{l\in S} e^{U_l} + e^{V_0}}\right)$. **?** prove the optimality of popular sets under the MNL model. Given that the exponential function preserves the attractiveness order, their result continues to hold in our setting for the special case of 1-packs.

Proof of Proposition 6. An upper bound on total demand (which is valid under both choice rules) is equal $\overline{D} = \lambda n$, which is achieved when all consumers make a purchase. Let $S^* = \{1\}$, then $k^{*n} = (n, 0, ..., 0)$ and $V^{*n}(S^*) = nU_1 > V_0$. As a result $D(S^*) = (n\lambda, 0, ..., 0)$

and total demand $\sum_{i \in S} D_i(S^*) = D_1(S^*) = n\lambda = \overline{D}$. Now consider any other vector S. We must have $\vec{D}(S^*) \succ_W \vec{D}(S) = (D_1, ..., D_M)$ and therefore from Lemma 23, we have $\Pi(S^*) = \pi(D(S^*)) \geq \pi(D(S)) = \Pi(S)$.

Proof of Proposition 4. Suppose S^* is not a popular-eccentric set, then there exist products $j \in S^*$ and $i, l \in \mathcal{M} \setminus S^*$ such that $U_l < U_j < U_i$. From Lemma 6, we must have either $\Pi((S^* \setminus \{j\}) \cup \{l\}) > \Pi(S^*)$ or $\Pi((S^* \setminus \{j\}) \cup \{i\}) > \Pi(S^*)$ or both, which contradicts that S^* is optimal. Therefore, S^* is popular-eccentric assortment.

Proof of Lemma 8. When there is no choice premium, the value of any n-pack $\vec{k}(S)$ is:

$$\begin{aligned} V(\vec{k}(S)) &= \sum_{l \in S} (k_l) U_l + n\gamma \\ &\leq n(U_{l_{max}} + \gamma) = V(\vec{k}^{*n}(S)) \end{aligned}$$

Proof of Proposition 9. It is easy to show that Lemma 6 continues to hold in the absence of choice premium so the proof is identical to that of Proposition 4.

APPENDIX B

PROOFS FOR CHAPTER 3

Proof of Proposition 10: For m and l such that $0 \leq m \leq l \leq \min\{\overline{n}, r_j\}$:

$$\begin{aligned}\Pi(l\mathbf{e}^j) &= \sum_{k=1}^{K} \lambda\alpha^k (p_j + \delta^k) \frac{l u_j^k}{l u_j^k + u_0^k} \\ &\geq \sum_{k=1}^{K} \lambda\alpha^k (p_j + \delta^k) \frac{m u_j^k}{m u_j^k + u_0^k} \\ &= \Pi(m\mathbf{e}^j)\end{aligned}$$

The inequality above follows from the fact that $\frac{x}{x+z} \geq \frac{y}{y+z}$ for all $x \geq y$.

Proof of Lemma 10: Consider an assortment with repetition vector $\mathbf{x}$. Let i be the product profile such that such that $p_i \geq \max_{j:x_j \geq 1}\{p_j\}$. Let $\mathbf{e}^i$ denote a vector of length N where the i-th element is equal to 1 and all others are equal to zero.

$$\begin{aligned}&\Pi(\mathbf{x}+\mathbf{e}^i) - \Pi(\mathbf{x}) \\ = &\sum_{k=1}^{K} \lambda\alpha^k \frac{\sum_{j\neq i}^{N}(p_j+\delta^k)x_j u_j^k + (p_i+\delta^k)(x_i+1)u_i^k}{\sum_{j\neq i}^{N} x_j u_j^k + (x_i+1)u_i^k + u_0^k} - \sum_{k=1}^{K} \lambda\alpha^k \frac{\sum_{j\neq i}^{N}(p_j+\delta^k)x_j u_j^k + (p_i+\delta^k)x_i u_i^k}{\sum_{j\neq i}^{N} x_j u_j^k + x_i u_i^k + u_0^k} \\ = &\sum_{k=1}^{K} \lambda\alpha^k \frac{(p_i+\delta^k)u_i^k u_0^k + (p_i+\delta^k)u_i^k \left(\sum_{j\neq i}^{N} x_j u_j^k\right) - \sum_{j\neq i}^{N}(p_j+\delta^k)x_j u_j^k u_i^k}{(\sum_{j\neq i}^{N} x_j u_j^k + (x_i+1)u_i^k + u_0^k)(\sum_{j\neq i}^{N} x_j u_j^k + x_i u_i^k + u_0^k)}\end{aligned}$$

where for $k = 1, \ldots, K$

$$\begin{aligned}&(p_i+\delta^k)u_i^k \left(\sum_{j\neq i}^{N} x_j u_j^k\right) - \sum_{j\neq i}^{N}(p_j+\delta^k)x_j u_j^k u_i^k \\ = \; &u_i^k \left[\sum_{j\neq i}^{N}(p_i+\delta^k)x_j u_j^k - \sum_{j\neq i}^{N}(p_j+\delta^k)x_j u_j^k\right] \\ = \; &u_i^k \sum_{j\neq i}^{N}(p_i - p_j)x_j u_j^k \\ \geq \; &0\end{aligned}$$

The last inequality holds because $p_i \geq p_j$ for all j such that $x_j \geq 1$. Therefore, $\Pi(\mathbf{x} + \mathbf{e}^i) \geq \Pi(\mathbf{x})$.

Proof of Proposition 14: Let $\mathbf{x}^{*f}$ be the optimal assortment for the *free-product* problem under no repetition. Suppose $\sum_{j=1}^N x_j^{*f} = n < \overline{n}$, consider a product $\hat{i} \in \{i : x_J^{*f} = 0\}$, we have

$$\begin{aligned}
\Pi^f(\mathbf{x}^{*f} + \mathbf{e}^{\hat{i}}) - \Pi^f(\mathbf{x}^{*f}) &= \sum_{k=1}^K \delta^k \lambda \alpha^k \frac{\sum_{j=1}^N x_j^{*f} u_j^k + u_{\hat{i}}^k}{\sum_{j=1}^N x_j^{*f} u_j^k + u_{\hat{i}}^k + u_0^k} - \sum_{k=1}^K \delta^k \lambda \alpha^k \frac{\sum_{j=1}^N x_j^{*f} u_j^k}{\sum_{j=1}^N x_j^{*f} u_j^k + u_0^k} \\
&= \sum_{k=1}^K \delta^k \lambda \alpha^k \left(\frac{\sum_{j=1}^N x_j^{*f} u_j^k + u_{\hat{i}}^k}{\sum_{j=1}^N x_j^{*f} u_j^k + u_{\hat{i}}^k + u_0^k} - \frac{\sum_{j=1}^N x_j^{*f} u_j^k}{\sum_{j=1}^N x_j^{*f} u_j^k + u_0^k} \right)
\end{aligned}$$

where for $k = 1, \ldots, K$, we have

$$\frac{\sum_{j=1}^N x_j^{*f} u_j^k + u_{\hat{i}}^k}{\sum_{j=1}^N x_j^{*f} u_j^k + u_{\hat{i}}^k + u_0^k} - \frac{\sum_{j=1}^N x_j^{*f} u_j^k}{\sum_{j=1}^N x_j^{*f} u_j^k + u_0^k} \geq 0$$

Therefore, $\Pi^f(\mathbf{x}^{*f} + \mathbf{e}^{\hat{(i)}}) \geq \Pi^f(\mathbf{x}^{*f})$, which contradicts that $\mathbf{x}^{*f}$ is optimal.

Proof of Proposition 15: Consider the *single-segment* problem $\mathcal{P}^{S_k}$, $k \in \{1, \ldots, K\}$, we assume that $\overline{n}\mathbf{e}^j = \underset{i \in \{1,\ldots,N\}}{arg \max} \{\Pi_k^s(\overline{n}\mathbf{e}^i)\}$. Suppose $\mathbf{x}^{*s_k} \neq \overline{n}\mathbf{e}^j$, then there exists product $i \neq j$ such that $\Pi_k^s((\overline{n}-1)\mathbf{e}^j + \mathbf{e}^i) > \Pi_k^s(\overline{n}\mathbf{e}^j)$. Then,

$$\begin{aligned}
& \Pi_k^s((\overline{n}-1)\mathbf{e}^j + \mathbf{e}^i) > \Pi_k^s(\overline{n}\mathbf{e}^j) \\
\Leftrightarrow \quad & \lambda \alpha^k \frac{(\overline{n}-1)(p_j + \delta^k) u_j^k + (p_i + \delta^k) u_i^k}{(\overline{n}-1) u_j^k + u_i^k + u_0^k} > \lambda \alpha^k \frac{\overline{n}(p_j + \delta^k) u_j^k}{\overline{n} u_j^k + u_0^k} \\
\Leftrightarrow \quad & \left[(\overline{n}-1)(p_j + \delta^k) u_j^k + (p_i + \delta^k) u_i^k\right] \left(\overline{n} u_j^k + u_0^k\right) > \left[\overline{n}(p_j + \delta^k) u_j^k\right] \left[(\overline{n}-1) u_j^k + u_i^k + u_0^k\right] \\
\Leftrightarrow \quad & \overline{n} u_i^k u_j^k (p_i - p_j) + u_0^k \left[(p_i + \delta^k) u_i^k - (p_j + \delta^k) u_j^k\right] > 0 \qquad \text{(B.1)}
\end{aligned}$$

But we know that

$$
\begin{aligned}
& \Pi_k^s(\overline{n}\mathbf{e}^j) \geq \{\Pi_k^s(\overline{n}\mathbf{e}^i) \\
\Leftrightarrow \quad & \lambda\alpha^k \frac{\overline{n}(p_j+\delta^k)u_j^k}{\overline{n}u_j^k+u_0^k} \geq \lambda\alpha^k \frac{\overline{n}(p_i+\delta^k)u_i^k}{\overline{n}u_i^k+u_0^k} \\
\Leftrightarrow \quad & \left[\overline{n}(p_j+\delta^k)u_j^k\right]\left(\overline{n}u_i^k+u_0^k\right) \geq \left[\overline{n}(p_i+\delta^k)u_i^k\right]\left(\overline{n}u_j^k+u_0^k\right) \\
\Leftrightarrow \quad & \overline{n}u_i^k u_j^k(p_j-p_i) + u_0^k\left[(p_j+\delta^k)u_j^k - (p_i+\delta^k)u_i^k\right] \geq 0 \\
\Leftrightarrow \quad & \overline{n}u_i^k u_j^k(p_i-p_j) + u_0^k\left[(p_i+\delta^k)u_i^k - (p_j+\delta^k)u_j^k\right] \leq 0
\end{aligned}
$$

which contradicts inequality (B.1). Therefore $\mathbf{x}^{*s_k} = \overline{n}\mathbf{x}^j$.

Proof of Lemma 15: Suppose $\mathbf{x}^{*f} \neq (r_1, \ldots, r_N)$, then $x_j^{*f} \leq r_j$ for $j = 1, \ldots, N$ and $\sum_{j=1}^N x_j^{*f} < \sum_{j=1}^N r_j$. Then we have:

$$
\begin{aligned}
& \Pi^f(\mathbf{x}^{*f}) - \Pi^f((r_1, \ldots, r_N)) \\
= \quad & \sum_{k=1}^K \delta^k\lambda\alpha^k \frac{\sum_{j=1}^N x_j^{*f}u_j^k}{\sum_{j=1}^N x_j^{*f}u_j^k + u_0^k} - \sum_{k=1}^K \delta^k\lambda\alpha^k \frac{\sum_{j=1}^N r_j u_j^k}{\sum_{j=1}^N r_j u_j^k + u_0^k} \\
= \quad & \sum_{k=1}^K \delta^k\lambda\alpha^k \left(\frac{\sum_{j=1}^N x_j^{*f}u_j^k}{\sum_{j=1}^N x_j^{*f}u_j^k + u_0^k} - \frac{\sum_{j=1}^N r_j u_j^k}{\sum_{j=1}^N r_j u_j^k + u_0^k}\right) \\
< \quad & 0
\end{aligned}
$$

which is contradiction. The last inequality above follows from the fact that $\sum_{j=1}^N x_j^{*f}u_j^k \leq \sum_{j=1}^N r_j u_j^k$. Therefore $\mathbf{x}^{*f} = \{r_1, \ldots, r_N\}$.

Proof of Theorem 2: Consider the product selection problem with only customer segment k and a fixed cardinality, that is:

$$
\max_{\substack{S\subseteq\mathcal{N};\\ |S|=l}} \Pi_k(S) \quad = \quad \sum_{j\in S}(p_j+\delta^k)\lambda\alpha^k \frac{u_j^k}{\sum_{i\in S} u_i^k + u_0^k} \tag{B.2}
$$

We claim that the solution to this problem for all $k \in \{1, ..., K\}$ is a set $S_{(l)}$ for some $l \in \{1, ..., N\}$. Suppose (contradiction) that this is not the case, i.e., assume that there exists a product $m < n$ such that $m \notin S^*$ and $n \in S^*$. Let $T = S^* \backslash \{n\}$. We have:

$$\begin{aligned}
\Pi_k\left(T \cup \{m\}\right) &= \sum_{j \in T}(p_j + \delta^k)\lambda\alpha^k \frac{u_j^k}{\sum_{i \in T} u_i^k + u_m^k + u_0^k} + (p_m + \delta^k)\lambda\alpha^k \frac{u_m^k}{\sum_{i \in T} u_i^k + u_m^k + u_0^k} \\
&= \lambda\alpha^k \frac{\sum_{j \in T}(p_j + \delta^k)u_j^k + (p_m + \delta^k)u_m^k}{\sum_{i \in T} u_i^k + u_m^k + u_0^k} \\
&\geq \lambda\alpha^k \frac{\sum_{j \in T}(p_j + \delta^k)u_j^k + (p_n + \delta^k)u_n^k}{\sum_{i \in T} u_i^k + u_n^k + u_0^k} \\
&= \Pi_k\left(S^*\right)
\end{aligned}$$

where the inequality comes from the two conditions on the product ordering. This proves the claim.

Since for every customer segment k, the solution to (B.2) is $S_{(l)}$ then the solution to the product selection problem with K customer segment and a cardinality constraint of $|S| = l$ is also $S_{(l)}$. Finally, the result follows from the fact that the solution to the unconstrained problem can be obtained by considering all possible values of $l \in \{1, ..., N\}$.

REFERENCES

Amelio, A., and Jullien, B., 2012. Tying and freebies in two-sided markets. *International Journal of Industrial Organization*, 30(5), 436-446.

Alptekinoğlu, A. and Corbett, C.J., 2008. Mass customization vs. mass production: Variety and price competition. *Manufacturing & Service Operations Management*, 10(2), 204-217.

Alptekinoğlu, A. and Corbett, C.J., 2010. Leadtime-variety tradeoff in product differentiation. *Manufacturing & Service Operations Management*, 12(4), 569-582.

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

Armstrong, M., 2006. Competition in two?sided markets. The RAND Journal of Economics, 37(3), 668-691.

Aydin, G. and Ryan, J.K., 2000, April. Product line selection and pricing under the multinomial logit choice model. In Proceedings of the 2000 MSOM conference.

Bayus, B.L. and Putsis Jr, W.P., 1999. Product proliferation: An empirical analysis of product line determinants and market outcomes. *Marketing Science*, 18(2), 137-153.

Bansal, S. and Moritz, B., 2015. Perceived versus actual value of product substitution flexibility: An experimental investigation. *Journal of Operations Management*, 38, 56-70.

Benzion, U., Cohen, Y., Peled, R. and Shavit, T., 2008. Decision-making and the newsvendor problem: an experimental study. *Journal of the Operational Research Society*, 59(9), 1281-1287.

Bernstein, F., Kk, A. G., and Xie, L. 2011. The role of component commonality in product assortment decisions. *Manufacturing* & *Service Operations Management*, **13**(2), 261-270.

Bolton, G.E. and Katok, E., 2008. Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. Manufacturing & Service Operations Management, 10(3), pp.519-538.

Bolton, G.E., Ockenfels, A. and Thonemann, U.W., 2012. Managers and students as newsvendors. *Management Science*, 58(12), 2225-2233.

Bostian, A.A., Holt, C.A. and Smith, A.M., 2008. Newsvendor ?pull-to-center? effect: Adaptive learning in a laboratory experiment. Manufacturing & Service Operations Management, 10(4), pp.590-608.

Bown, N.J., Read, D. and Summers, B., 2003. The lure of choice. *Journal of Behavioral Decision Making*, 16(4), 297-308.

Boyer, K. and Verma, R., 2009. Operations and supply chain management for the 21st century. Nelson Education.

Bront, J. J. M., Méndez-Díaz, I., and Vulcano, G., 2009. A column generation algorithm for choice-based network revenue management. *Operations Research*, 57(3), 769-784.

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

Cachon, G.P. and 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. and Xu, Y., 2008. On the effects of consumer search and firm entry in a multiproduct competitive market. *Marketing Science*, 27(3), 461-473.

Caillaud, B. and Jullien, B., 2003. Chicken & egg: Competition among intermediation service providers. *RAND journal of Economics*, 309-328.

Carrillo, J. D., 2006. Platform competition: The role of multi-homing and complementors.

Caro, F. and Gallien, J., 2010. Inventory management of a fast-fashion retail network. *Operations Research*, 58(2), 257-273.

Caro, F. and Martnez-de-Albniz, V., 2012. Product and price competition with satiation effects. *Management Science*, 58(7), 1357-1373.

Chao, Y., and Derdenger, T., 2013. Mixed bundling in two-sided markets in the presence of installed base effects. *Management Science*, 59(8), 1904-1926.

Chen, K. and Li, S., 2016. The Behavioral Traps in Making Multiple, Simultaneous, Newsvendor Decisions. 2016.

Chen, L., Kök, A.G. and Tong, J.D., 2013. The effect of payment schemes on inventory decisions: The role of mental accounting. *Management Science*, 59(2), 436-451.

Choi, J. P., 2010. Tying in two-sided markets with multi-homing. *The Journal of Industrial Economics*, 58(3), 607-626.

Corsten, D, and Gruen, T.. 2003. Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International Journal of Retail & Distribution Management*, 31(12), 605-617.

Davis, J., Gallego, G., and Topaloglu, H., 2013. Assortment planning under the multinomial logit model with totally unimodular constraint structures. Department of IEOR, Columbia University.

Dubé, J. 2004. Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks. *Marketing Science*, **23**(1), 66-81.

Evans, D. S., 2002. The antitrust economics of two-sided markets. Available at SSRN 332022.

Evans, D. S., and Schmalensee, R., 2005. The industrial organization of markets with two-sided platforms (No. w11603). National Bureau of Economic Research.

Farias, V., Jagabathula, S. and Shah, D., 2011. Assortment optimization under general choice. arXiv preprint arXiv:1108.3596.

Farias, V. F., Jagabathula, S., and Shah, D., 2012. Sparse choice models. In Information Sciences and Systems (CISS), 2012 46th Annual Conference, IEEE, 1-22.

Farias, V. F., Jagabathula, S., and Shah, D., 2013. A nonparametric approach to modeling choice with limited data. *Management Science*, 59(2), 305-322.

Feng, G., Li, X., and Wang, Z. 2017. On the Relation Between Several Discrete Choice Models. *Operations Research*.

Fox, E., Norman, L., and Semple, J. 2017. Choosing an n-Pack of Substitutable Products. *Management Science*.

Gallego, G., Ratliff, R. and 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.

Gallego, G., Ratliff, R., and Shebalov, S., 2011. A general attraction model and an efficient formulation for the network revenue management problem. Techinal Report, Columbia University, New York, NY.

Gallego, G., and Topaloğlu, H., 2014. Constrained assortment optimization for the nested logit model. *Management Science*, 60(10), 2583-2601.

Honhon, D., Jonnalagedda, S., and Pan, X. A. (2012). Optimal algorithms for assortment selection under ranking-based consumer choice models. *Manufacturing & Service Operations Management*, 14(2), 279-289.

Gallego, G., Ratliff, R., and Shebalov, S., 2011. A general attraction model and an efficient formulation for the network revenue management problem. Techinal Report, Columbia University, New York, NY.

Gallego, G., and Topaloğlu, H., 2014. Constrained assortment optimization for the nested logit model. *Management Science*, 60(10), 2583-2601.

Gaur, V. and Honhon, D., 2006. Assortment planning and inventory decisions under a locational choice model. *Management Science*, 52(10), 1528-1543.

Guo, L., 2010. Capturing consumption flexibility in assortment choice from scanner panel data. *Management Science*, 56(10), 1815-1832.

Hagiu, A., 2006. Pricing and commitment by two-sided platforms. *RAND Journal of Economics*, 720-737.

Hagiu, A., and Wright, J., 2014. Marketplace or reseller?. *Management Science*, 61(1), 184-203.

Harlam, B. A., and Lodish, L. M. 1995. Modeling consumers' choices of multiple items. *Journal of Marketing Research*, 404-418.

Honhon, D., Gaur, V. and Seshadri, S., 2010. Assortment planning and inventory decisions under stockout-based substitution. *Operations research*, 58(5), 1364-1379.

Honhon, D., Jonnalagedda, S., and Pan, X. A. (2012). Optimal algorithms for assortment selection under ranking-based consumer choice models. *Manufacturing* & *Service Operations Management*, 14(2), 279-289.

Hopp, W.J. and Xu, X., 2005. Product line selection and pricing with modularity in design. *Manufacturing & Service Operations Management*, 7(3), 172-187.

Ho, T.H., Lim, N. and Cui, T.H., 2010. Reference dependence in multilocation newsvendor models: A structural analysis. *Management Science*, 56(11), 1891-1910.

Jagabathula, S., and Rusmevichientong, P. 2016. The Limit of Rationality in Choice Modeling: Formulation, Computation, and Implications.

Julllien, B., and Caillaud, B., 2003. Chicken and Egg: Competition among Intermediaton Service Providers. *The RAND Journal of Economics*, Jg, 34, 521-552.

Kaiser, U., and Wright, J., 2006. Price structure in two-sided markets: Evidence from the magazine industry. *International Journal of Industrial Organization*, 24(1), 1-28.

Katok, E. and Wu, D.Y., 2009. Contracting in supply chains: A laboratory investigation. *Management Science*, 55(12), 1953-1968.

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

Kremer, M., Minner, S. and Van Wassenhove, L.N., 2010. Do random errors explain newsvendor behavior?. *Manufacturing & Service Operations Management*, 12(4), 673-681.

Li, Z. (2007). A Single Period Assortment Optimization Model. *Production and Operations Management*, 16(3), 369-380.

Lurie, N.H. and Swaminathan, J.M., 2009. Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes*, 108(2), 315-329.

Maddah, B. and Bish, E.K., 2007. Joint pricing, assortment, and inventory decisions for a retailer's product line. *Naval Research Logistics*, 54(3), 315-330.

Marshall, A. W., I. Olking, B. Arnold. 2011. Inequalities: Theory of Majorization and Its Applications. Second edition. Springer Series in Statistics. Springer, New York.

Méndez-Déaz, I., Miranda-Bront, J. J., Vulcano, G., and Zabala, P. (2010). A Branch-and-Cut Algorithm for the Latent Class Logit Assortment Problem. *Electronic Notes in Discrete Mathematics*, 36, 383-390.

Porteus, E.L., 1990. Stochastic inventory theory. Handbooks in operations research and management science, 2, 605-652.

Read, D. and Loewenstein, G., 1995. Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, 1(1), 34.

Read, D., Loewenstein, G. and Kalyanaraman, S., 1999. Mixing virtue and vice: Combining the immediacy effect and the diversification heuristic. *Journal of Behavioral Decision Making*, 12(4), 257-273.

Rochet, J. C., and Tirole, J., 2003. Platform competition in two-sided markets. *Journal of the European Economic Association*, 990-1029.

Rochet, J. C., and Tirole, J., 2006. Two-sided markets: a progress report. *The RAND Journal of Economics*, 37(3), 645-667.

Rochet, J. C., and Tirole, J., 2008. Tying in two-sided markets and the honor all cards rule. *International Journal of Industrial Organization*, 26(6), 1333-1347.

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

Rusmevichientong, P., and Topaloglu, H., 2012. Robust assortment optimization in revenue management under the multinomial logit choice model. *Operations Research*, 60(4), 865-882.

Rusmevichientong, P., Shmoys, D., Tong, C., and Topaloglu, H., 2014. Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management*, 23(11), 2023-2039.

Rysman, M., 2004. Competition between networks: A study of the market for yellow pages. *The Review of Economic Studies*, 71(2), 483-512.

Salisbury, L.C. and Feinberg, F.M., 2008. Future preference uncertainty and diversification: The role of temporal stochastic inflation. *Journal of Consumer Research*, 35(2), 349-359.

Schweitzer, M.E. and Cachon, G.P., 2000. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, 46(3), 404-420.

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

Simonson, I.. 1990. The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, 150-162.

Simonson, I. and Winer, R.S., 1992. The influence of purchase quantity and display format on consumer preference for variety. *Journal of Consumer Research*, 19(1), 133-138.

Su, X., 2008. Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management*, 10(4), 566-589.

Talluri, K., and Van Ryzin, G., 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1), 15-33.

Wang, R.. 2017. Consumer Choice and Market Expansion: Modeling, Optimization and Implementation.

Weyl, E. G., 2010. A price theory of multi-sided platforms. *The American Economic Review*, 1642-1672.

van Ryzin G., & S. Mahajan. 1999. On the RelationShip Between Inventory Costs and Variety Benefits in Retail Assortments. *Management Science*, **45**(11), 1496-1509.

BIOGRAPHICAL SKETCH

Ying Cao is a doctoral student in Operations Management at the Naveen Jindal School of Management, The University of Texas at Dallas. She received her bachelor's in Industrial Engineering and master' degree in Management Science and Engineering from Nanjing University of Aeronautics and Astronautics. She also holds a master's degree in Supply Chain Management from The University of Texas at Dallas. Her research interests include retail operations, behavioral operations and supply chain management. She will be joining the Black School of Business, Penn State University Erie as a tenure track Assistant Professor in Project and Supply Chain Management in Fall 2018.

CURRICULUM VITAE

Ying Cao

August 20, 2018

Contact Information:

Department of Project and Supply Chain Management
Black School of Business
Penn State Erie, The Behrend College
4701 College Dr.
Erie, PA 16563 U.S.A.

Phone: (814) 898-6710
Fax: (814) 898-6223
Email: yxc673@psu.edu

Educational History:

B.B.A, Industrial Enginnering, Nanjing University of Aeronautics and Astronautics, 2010
M.S., Management Science and Engineering, Nanjing University of Aeronautics and Astronautics, 2013
M.S., Supply Chain Management, University of Texas at Dallas, 2015
Ph.D., Management Science, University of Texas at Dallas, 2018

Research Interests:

Retail Operations, Behavioral Operations, Supply Chain Management

Research

Ying Cao, Dorothée Honhon, "Assortment Planning with n-Pack Purchasing Consumers", working paper.
Ying Cao, Dorothée Honhon, Sridhar Seshadri, "Assortment Planning in a Two-Sided Market", working paper.
Ying Cao, Dorothée Honhon, Kyle Hyndman, "A Behavioral Analysis of Inventory Management for Substitutable Products", working paper.

Teaching Experience

OPRE 3310 Operations Management (Undergraduate), University of Texas at Dallas, Summer 2018

Professional Honors and Awards:

MIT Professional Education - Short Programs, Discrete Choice Analysis: Predicting Demand and Market Shares, 2017
WORMS (WOmen in OR/MS Childcare Travel Fund at INFORMS, 2016
INFORMS Doctoral Student Consortium, 2014

Professional Services:

Reviewer for MSOM (Manufacturing and Service Operations Management) Conference, 2018.
Session Chair (for 2 sessions) in Supply Chain and Revenue Management Track, DSI Annual Meeting, DC, 2017.
Session Chair in Supply Chain and Revenue Management Track, DSI Annual Meeting, DC, 2018 (Scheduled).

Affliations:

Decision Science Institute (DSI): Member
Institute for Operations Research and the Management Sciences (INFORMS): Member
Manufacturing and Service Operations Management Society (MSOM): Member
Production and Operations Management Society (POMS): Member
The Forum for Women in OR/MS (WORMS): Member