NOVEL ASSORTMENT PROBLEMS IN RETAIL OPERATIONS

Oğuz Çetin

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Operations in the Kenan-Flagler Business School.

Chapel Hill
2019

Approved by:

Adam J. Mersereau

Ali K. Parlaktürk

Vinayak Deshpande

Lauren X. Lu

Serhan Ziya

ABSTRACT:

Oğuz Çetin: Novel Assortment Problems in Retail Operations
(Under the direction of Adam J. Mersereau and Ali K. Parlaktürk.)

My thesis extends classical work on retail assortment planning based on the notion that how and where products are displayed impact customer choices. I look for structural insights into such problems modeled using random utility based discrete choice frameworks where the associated mean utility of an alternative is determined not only by the product's inherent features, but also by how it is presented and/or delivered to the customer.

Chapter 2 highlights that promotional displays provide a visibility advantage to both the featured product and its category, but it also has consequences for customer traffic and substitution. Therefore, retailer's optimal choice of product to include in a promotional display depends not only on the product attributes but also on a quantity we call "aisle attractiveness" which is determined by several category-level parameters. The value of the display to a category pivots on whether the display's role is primarily to expand demand for the category or to shape substitution within the category.

Chapter 3 extends the work in Chapter 2 by integrating the possibility of price discounts along with promotional displays, and by incorporating the manufacturer's perspective into the problem. Manufacturers often provide incentives (e.g. per unit discounts on wholesale prices, fixed payments, etc.) to induce retailers to feature their own products by displaying and discounting the retail price, but retailer's responses to such incentives differ across products. Everything else the same, manufacturers of high share and/or high margin products are more likely to acquire retailer's display support. On the other hand, we find that the retailer tends to offer larger discounts to high margin and/or low share products.

In Chapter 4, I consider an assortment problem in a network of stores where transshipments among stores are possible at a certain cost, and customer choice is determined by product attributes and shipment delays. We extend some classical results in the assortment planning literature to the multi-location scenario showing in particular that the structure of an optimal multi-location assortment can be defined in terms of sets of the most popular products. We also show how the coordination of independent retail stores can be achieved.

To my mom and dad

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisors Prof. Adam J. Mersereau and Assoc. Prof. Ali K. Parlaktürk for their guidance and support, both academically and mentally. They always helped me at my difficult times with all their understanding and encouragement. We have spent endless hours in our research meetings in the last five years during which I've learnt how to conduct research independently. This is by far the most precious gain from my doctoral training. I will always appreciate their effort on my academic development.

I am very thankful to Prof. Vinayak Deshpande, Assoc. Prof. Lauren X. Lu, and Prof. Serhan Ziya for volunteering to read and review this dissertation and for their insightful comments. Despite all the inconveniences, they found a way to invest their time and share their knowledge to improve this dissertation.

I have shared lots of great memories over the years at UNC with my friends Naga, Pradeep, Anand, Herbie, Dayton, and RJ. Their friendship is invaluable to me.

I would like to thank two of my close friends in particular, Ali Yılmaz and Onursal Bağırgan, who always cheer me up in our long chats everyday. I came to the U.S. together with two very special individuals, Hüseyin Gürkan and Yunus Emre Kesim, whose friendship and support I cannot appreciate enough. I hope to have their company in all my life.

I would not be able to finish my Ph.D. if I hadn't met my girlfriend, Allie, in my third year. Her presence changed my life entirely because she showed me endless different ways to enjoy it. She also introduced me to Messina's, my favorite family in the U.S. I am so excited for our future together-Allie.

Finally, I would like to express my deepest gratitude to my family. They suffered the most during my Ph.D. far away from home. Their understanding and sacrifice is beyond any measure. Everything I achieve in life is thanks to their endless love, support and belief in me. They are always my biggest motivation.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

I have been interested in best practices in managing demand and supply in retail operations since my senior year in college when I worked for a hard-discount retailer in Turkey. The complexity of operations in a typical retail supply chain including ordering, stocking, presenting, and delivering thousands of SKUs in hundreds of product categories offers a great opportunity for academic research, which motivated me throughout my Ph.D. like many other scholars in operations management.

In this dissertation, the most prominent theme is the customer choice among a set of alternatives in a retail environment, where an alternative means not only a product variant itself, but also how it is presented and/or delivered to the customer, which have profound effects on customer choice. These effects are significant but might sometimes understudied both in practice and in the literature. The focus of this dissertation is the latter aspect, i.e., factors affecting customer choice other than the inherent attributes of the product (that are often determined by the manufacturers); e.g., the location of the product and the category within the store, a secondary display in a more prominent location, the delivery time, the retail price, assortment depth, etc. These are the factors that the retailer can use as levers to shape the overall demand for the category and the demand across products.

Perhaps the most related stream of research to this dissertation is the assortment planning literature, in which modeling the customer choice constitutes the basis of the analysis, and the aim is to control customers' choice across products, or equivalently to control substitutions from one product to the other, by determining the set of products customers choose from, and the level of inventories to prevent stock-outs. However, customers substitute not only because of assortment or inventory related unavailability, but also because of other reasons such as retailer effort to promote a product or to transship a better substitute from another store. These constitute different chapters of this dissertation.

In Chapter 2, we focus on customers' movement within a store. This work has been initiated by our discussions with a convenience store chain, when it became clear that product location decisions are typically made considering customer movements within the store. Moreover, secondary product locations (e.g. endcap displays) are frequently used to exploit high-traffic zones in a store to increase the visibility of certain products. A promotional display expands the overall demand for the category and the demand for the product being featured, but it also has implications on customer visits to regular aisles and therefore demand for other products. We show that the optimal product to promote lies along an "efficient set" of products drawn in terms of product margins and popularities. It also depends on certain category and store features, and whether the display's role is primarily to expand demand for the category or to shape substitution within the category. Our work provides guidance for how retailers can use and value promotional displays effectively.

In Chapter 3, we extend our analysis of the management of promotional displays in two directions. First, we incorporate retail price discounts, which are often supported by promotional displays. Second, we introduce the incentives offered by manufacturers to induce the retailer to promote their own products. We first extend our "efficient set" result from Chapter 2 for when the retailer is allowed to discount the price of the displayed product. We characterize the conditions on the product popularity and the profit margin for when the retailer should support a price discount with a promotional display. We show structure on the retailer's best response (level of pass through and display decision) to a trade deal offer from a manufacturer. Finally, we establish that a fixed payment to the retailer (e.g. display allowance) is better for a manufacturer than offering a per-unit discount on the wholesale price (e.g. scan-back trade deals) if the manufacturer is interested in the retailer's display support. We discuss how our results are related to some existing results in the operations and marketing literatures.

In Chapter 4, we consider a retailer having multiple stores within a close geographical proximity among which the retailer performs transshipment of products. This has implications on stores' assortment decisions. The customers in this setting are choosing not only among product variants, but also between two fulfillment options: (1) from own-store and (2) tranship from another store. Although the operations management literature includes rich work on demand fulfillment in a network of stores, transshipment in the context of single product inventory management, and centralized assortment planning, there has been little focus on assortment decisions in a multi-product and multi-location setting. We consider the centralized and the decentralized versions of the problem and showed how coordination can be achieved in a special case. We also show that some classical results in the assortment planning literature can be extended to the multi-location scenario. For instance, in both the centralized and the decentralized scenarios, the “popular assortment set” result of Ryzin and Mahajan [1999] holds for the union of assortments and for the set of common products in each store, but how exclusive products should be allocated among stores is a challenging key element for each individual store’s performance. Indeed, we show that the allocation of exclusive products is NP-complete in the centralized problem.

We use the (nested) multinomial logit (MNL, hereafter) framework in all chapters of this dissertation to model customer choice among multiple alternatives. MNL is extensively used for estimation purposes [Train, 2009] because the distributional assumptions on the random utility lead to closed-form choice probabilities, which in turn lead to a concave log-likelihood function [McFadden, 1978]. This makes MNL extremely amenable to estimation. On the other hand, the logit profit function itself is not concave in the demand paramaters (e.g., prices), so it is a challenging tool for analytical purposes. Nevertheless, it is also commonly used in analytical work optimizing assortments and product line pricing [Gallego and Topaloglu 2014; Li and Huh 2011; etc.] because it has a strong utility-based foundation and a nice interpretation of customer decision making.

CHAPTER 2: MANAGEMENT AND EFFECTS OF IN-STORE PROMOTIONAL DISPLAYS

2.1 Introduction

There has been considerable research in the operations management literature on the topic of assortment planning (i.e., which products to offer?) and to a lesser extent on shelf-space allocation (i.e., how much space to devote to each product?). However, little attention has been paid to the problem of *where* to place products in the store. This latter problem is both important and complex, as the placement of products in the store is a lever by which the retailer can impact customer traffic patterns and thereby shape demand across products. Retailers' merchandising decisions have significant influence on demand because many customer purchases involve some degree of in-store decision (i.e., whether to buy, or which product to buy). For instance, POPAI [2012] reports that 76% of purchases in supermarkets are either unplanned or planned only up to (but not including) a specific product choice prior to the shopping trip.

In this project we examine an important special case of the product location problem, namely the choice of product to feature in a promotional display. Examples motivating our work include endcap displays in supermarkets and beverage displays near cashier stations in convenience stores. While there is work in the marketing literature seeking to understand customer behavior in stores and to measure the impact of product placement decisions and price and advertising promotions, we are not aware of existing research providing prescriptive insights for the complex decision of which product to feature in a promotional display. Our work seeks to fill this gap.

A display provides a visibility advantage to the featured product over products stocked only in the regular shelf space. We can decompose the impact on demand for products in the category into two effects: (1) the display expands the overall category demand by capturing the attention of some customers who would not otherwise have purchased from the category (the "demand expansion effect") and (2) it may induce some customers to substitute the promoted product for their original preference (the "substitution effect"). We model both effects. We note that both effects can be present in the more traditional assortment planning problem, but the promotional display decision gives the retailer an additional degree of freedom. That is, the promotional display decision allows the manager to modulate the category's visibility and to shape substitution patterns even given a fixed assortment. Furthermore, it can be changed more easily and therefore made more dynamically than the assortment decision.

We assume that the set of products in the category assortment is exogenously given, as are the profit margins and mean utility (or popularity) of each product. This permits us to focus specifically on the effect of the display. All products in a category are co-located in a native "aisle" location, and at most one is selected also to be featured in the promotional display. We assume ample stock of all products.

We employ the nested multinomial logit (NMNL) framework to model customer choice within a category, which we interpret as a two-stage choice process. The customer first encounters the promotional display (if present) and chooses whether to purchase the featured product, visit the category's aisle location, or leave without purchasing from the category. A customer who visits the aisle may choose among the full assortment of products or may choose not to purchase. We assume that a customer who visits the aisle incurs a constant disutility representing the transit cost. The retailer collects the profit margin associated with any purchased product plus a bonus margin from any customer who visits the aisle. This bonus margin captures the expected margin of any impulse buys or complementary products the customer encounters upon visiting the aisle.

Following a review of relevant literature (Section 2.2) and the introduction of our model (Section 2.3), we present our analysis in three parts. In Section 2.4, we examine the choice of product to promote from a given category. We show that the retailer can immediately eliminate some products from consideration based only on margin and popularity characteristics. The remaining products form an "efficient set" within which higher popularity implies lower margin and vice versa. Intuitively, featuring a popular product in the display minimizes the chances of customers leaving empty-handed, while keeping popular, low margin products in the aisle maximizes both margins at the display and additional impulse spending in the aisle. We prove that the latter strategy is optimal when an aisle visit is inherently attractive to customers, which is the case when the aisle assortment is relatively large, when customers are heterogeneous in their preferences, and when the transit cost is low. An example implication is that popular products are good choices for endcap displays in large store footprints where customers incur high transit costs to visit an aisle. Moreover, high margin products, though at a potential popularity disadvantage, work better in a display when the category demand is spread out across multiple products in a large assortment.

In Section 2.5 we analyze the value of the promotional display across categories, assuming that the optimal product from each category would be promoted as described in Section 2.4. A key factor is the strength of the demand expansion effect for a category. We assume that the degree to which the promotional display can increase potential demand differs across categories, as some categories have an impulsive nature while others are more utilitarian. Clearly, the value of the promotional display is greatest when the demand expansion effect is strongest. Moreover, when the demand expansion effect is strong, the value of the display increases in the attractiveness of the aisle (determined by the assortment size, customer heterogeneity, and transit cost); the opposite is true when the demand expansion effect is weak. This implies that the display should be reserved for two classes of product categories: those with high demand expansion upside and an attractive aisle option, and those with stable category-level demand but a less appealing aisle option.

The analyses described so far assume that all products from a category are equally effective at generating awareness for the category when placed in the promotional display; i.e., the demand expansion effect is category-specific but not product-specific. (We note that even under this assumption, products will differ in their ability to generate sales from the display due to the substitution effect.) We believe this assumption is natural, though we are not aware of empirical evidence establishing whether or not demand expansion effects are product-specific. For completeness, in Section 2.6 we consider product-specific demand expansion effects and revisit our results on the choice of product to promote from a category. We require additional assumptions that we specify for our efficient set result to hold. We also demonstrate that the core directional forces identified in Section 2.4 remain in effect, but the forces operate in the directions of new product indices that depend on products' demand expansion abilities.

Throughout the paper, we take the retailer's perspective and we assume that the retailer behaves myopically, focusing on optimizing profits in the current period. However, we recognize that the retailer does not manage promotional displays in a vacuum. In practice there are negotiations over promotional displays between retailers and manufacturers, often resulting in manufacturers offering incentives to retailers to feature certain products. Product popularities may also depend on past promotions due to saturation and stockpiling effects. While we do not model these coordination activities and dynamics directly, we believe that understanding the retailer's incentives with respect to product margins (which can be dynamically influenced by manufacturer discounts) and product popularities (which can be influenced by other types of promotional activities) is an important building block for understanding these coordinations and dynamics.

Proofs of all analytical results are presented in Appendix 1.

2.2 Related Literature

Our study is related to a large literature on assortment planning, where the typical goal is to specify the set of products carried in a store (or in a sales channel) to maximize sales or gross margin under constraints on shelf space or on purchasing budget. A driving factor in this literature is customers' potential to substitute an available product when their original preference is not offered in the assortment, called assortment-based substitution [Kök et al., 2008]. Substitution in our paper arises from the visibility advantage of a product featured in a promotional display.

We employ the nested multinomial logit (NMNL) model to capture the aforementioned substitution effect. The NMNL model is commonly used in empirical choice models as a more flexible extension of the classical multinomial logit model [Train, 2009]. It is also commonly used to model two-stage choice scenarios [Manrai and Andrews, 1998]. An early use of it in the literature on demand for consumer packaged goods is Guadagni and Little [1998], who use it to capture consumers' intertemporal preferences for buying now versus later. Kök and Xu [2011] use it to capture customers' choices among brands then product types, or among product types then brands. Davis et al. [2014] and Gallego and Topaloglu [2014] consider nests as different sales channels or different stores. Aouad and Segev [2015] and Feldman and Topaloglu [2015] model nested consideration sets formed by online search rankings or cutoffs on attributes like quality or price. Our two-stage choice scenario has a spatial interpretation as discussed in Section 2.3. The nests in our setting represent different locations (display or aisle) in the store. Davis et al. [2013] also consider location specific preference weights in a multinomial logit model. They assume exogenously given preference weights, whereas we explicitly model the factors that determine the relative attractiveness of a product in different locations.

A natural constraint in assortment planning is on shelf space. Early research on shelf space allocation develops algorithms for large scale problems [Corstjens and Doyle, 1981; Bultez and Naert, 1988]. More recent studies [Bai et al., 2013; Geismar et al., 2015] incorporate the vertical alignment of products on the shelf which impacts product visibility. Our research does not consider shelf space or vertical alignment, though we consider the visibility difference between the regular shelf space and the promotional display, which arguably provides a more profound visibility advantage to the focal product. Wilkinson et al. [1982] report that the percentage increase in unit sales is between 16% and 39% with expanded shelf space, but between 77% and 243% with secondary displays for the four products they examine.

Customers' impulse spending figures prominently in our problem. Conventional wisdom suggests that exposure to in-store stimuli leads to impulse spending [Inman and Winer, 1998; Hui et al., 2013]. In our case, the promotional display triggers impulse spending on the featured product, but it also diverts some customers from the aisle due to substitution. Our model penalizes the loss in aisle traffic by explicitly capturing the potential impulse spending of customers in the aisle.

A stream of papers in the marketing literature empirically studies how the impact of a promotional display depends on factors such as the demand dispersion within the category [Chevalier, 1975], display's distance to the product's own aisle location [Nordfält and Lange, 2013; Phillips et al., 2015], and to a complementary category's aisle location [Bezawada et al., 2009]. Moreover, Dhar et al. [2001] study how demand expansion of promotional displays can vary across categories. Zhang [2006] models how customers form consideration sets via displays. We incorporate all of these factors in our model. These papers do not consider the optimal product choice for promotional display, and they tend to focus on the impact of external factors (e.g., day of the week, whether the display is assisted by a personnel or a feature ad) on the effectiveness of promotional display. In contrast, we characterize the optimal product and category choice based on their characteristics.

Van Heerde and Neslin [2017] provides a recent review of the marketing literature that empirically studies the effect of *price promotions*. These papers focus on understanding the sales lift of the promoted product resulting from demand expansion (primary) and the substitution (secondary) effects by studying household-level [Gupta, 1988; Chintagunta, 1993; Bell et al., 1999; Pauwels et al., 2002; Van Heerde et al., 2003] and store-level [Van Heerde et al., 2004; Nair et al., 2005] data. Our model considers both primary and secondary effects of *display promotions*. However, different from these studies, we take the retailer's perspective and consider the effects of these promotions on the retailer's profit and hence consider the impact of these promotions on all other products as well. In addition, we take into account the impact of a promotion on customer traffic within the store. The value of driving customers to a particular aisle is a core component of our model.

Cross-category effects of *price promotions* [Ailawadi et al., 2006; Leeflang et al., 2008; Leeflang and Parreño-Selva, 2012] and displays [Bezawada et al., 2009] are well established in the literature. While we do not explicitly model a cross-category effect of promotions, we capture complementarity between categories by accounting for average impulse spending on other categories when a customer visits the aisle.

2.3 Model Setup

We consider a category consisting of $n \geq 2$ products, and we indicate this set of products by $S = \{1, .., n\}$. Each product is characterized by its associated mean utility Y_j and its profit margin m_j. To simplify our exposition, we assume that no pair of products have the same mean utility and profit margin. The retailer has the option of promoting one of the products by featuring it in an in-store promotional display ("display" hereafter) in addition to its regular shelf space in the aisle. We use "aisle" to refer the location in the store permanently allocated to the category. Each product in the category has its own shelf space in the aisle and we do not model any differences in the shelf space allocated to products. We assume that prices are exogenous and are reflected in the margin and popularity parameters.

We employ the NMNL choice framework for two main reasons. First, it is a random utility model of customer choice that is consistent with utility maximizing behavior, that extends the classical MNL model, and that yields closed form choice probabilities [McFadden, 1978]. Second, the choice probabilities in the NMNL model can be decomposed into two stages whereby individuals first choose a nest and then choose an alternative within the chosen nest to maximize their utility

[Train, 2009]. This fits nicely with a natural interpretation of customer choice in our problem as a two-stage decision process as shown in Figure 2.1. (Some of the notations in Figure 2.1 will be defined later in this section, and are included in the figure for completeness.)

In the NMNL model, the utility U_j associated with alternative j in nest k has the following form:

$$U_j = Y_j + W_k + \epsilon_j, \tag{2.1}$$

where Y_j and W_k are alternative-specific and nest-specific deterministic components of the utility, respectively. The random component, ϵ_j, captures the idiosyncratic utility of a customer and enables heterogeneity in customers' choices. The ϵ_j's are assumed to follow a generalized extreme value distribution with the joint c.d.f.

$$\Phi(\epsilon) = \exp\left\{ - \sum_{k=1}^{K} \left(\sum_{j \in B_k} e^{-\epsilon_j/\lambda_k} \right)^{\lambda_k} \right\}, \tag{2.2}$$

where B_k is the set of alternatives in nest k. The difference between the MNL and the NMNL models is that the latter allows correlation between (ϵ_i, ϵ_j) pairs corresponding to alternatives i and j within the same nest, and the parameter $\lambda_k \in (0, 1]$ of the joint c.d.f. $\Phi(\epsilon)$ specifies this correlation structure for nest k. In particular, $(1 - \lambda_k)$ is a measure of the correlation between (ϵ_i, ϵ_j) corresponding to any two alternatives (i, j) within nest k [Train, 2009].

In our choice scenario, each customer makes a decision among three options in stage 1: to purchase from the display, to visit the aisle to see all available products, or to neither purchase nor visit the aisle. These options represent the nests in our NMNL model, and we index these nests by "D" (display), "A" (aisle), and "0" (no-purchase), respectively. We assume that customers incur a "transit cost" if they visit the aisle, which we model as a nest-specific disutility, c, associated with the nest corresponding to the aisle; that is, $W_A = -c$. Note that the nests 0 and D include only one alternative each, implying that λ_0 and λ_D are irrelevant and only λ_A has an impact on the choice probabilities. Hence, we drop the subscript in λ_A and consider λ as the parameter determining the level of correlation among the utilities associated with the alternatives in the aisle.

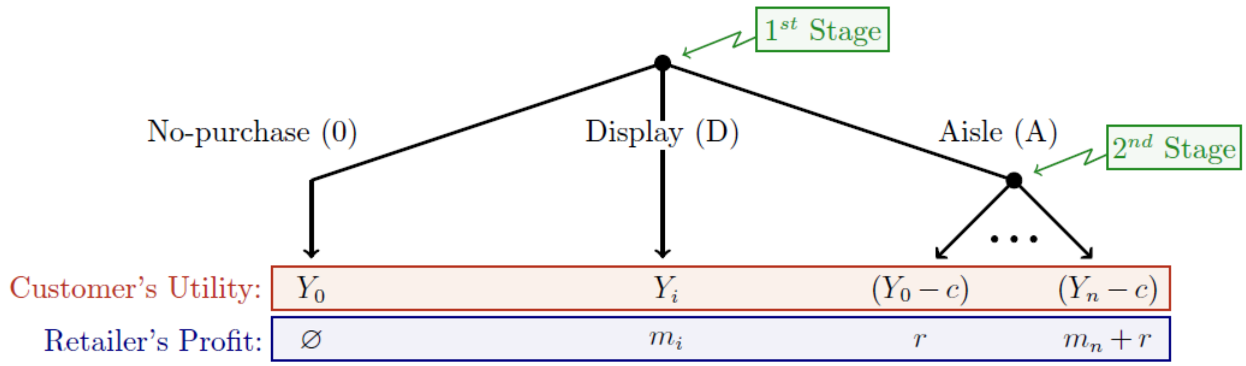


Figure 2.1: Customer choice modeled as a two-stage decision process.

There are $n+1$ alternatives in nest A, corresponding to n product variants and a no-purchase alternative. Therefore, our model assumes that the product being featured in the display is also available in the aisle (matching the grocery and convenience store cases that motivated our study), and it allows for customers not to purchase even after visiting the aisle (which is consistent with customers learning their value realizations for individual aisle products upon reaching the aisle). We note that certain promotional display settings in practice, such as prominent tables in apparel stores, differ from grocery and convenience store settings in that the same product may not be simultaneously available in two locations.

We note that an assumption of the NMNL model is that ϵ_j's corresponding to alternatives in different nests are independent. This may not strictly be the case in our setup because the featured product appears in both the nests D and A, and we would expect a correlation among the utilities for the same product in two locations. We nevertheless use the standard NMNL assumption as an approximation with the benefit of yielding closed form choice probabilities, noting that this approximation may underestimate the demand generated by the display (see Guadagni and Little, 1998 for a similar assumption).

We measured the impact of this independence assumption in a numerical simulation in which customers' utilities for the featured products in nests D and A are correlated. Our simulation model relies on a decomposition of the random error term ϵ_j into nest-specific (ν_k) and product-specific (ξ_j) components such that $\epsilon_j = \nu_k + \lambda\xi_j$ (see Besanko et al., 1998; Berry, 1994 for the same decomposition, and Cardell, 1997 for the distributional properties of ν_k and ξ_j). We capture the aforementioned correlation by fixing a customer's realization of the random term $\xi_{(\cdot)}$ corresponding to the featured product in the two locations (display and aisle). We similarly model correlation between the no purchase option in the display and aisle. To generate random utilities requires sampling from complicated probability distributions. We present some of the technical details and results of the simulation in Appendix 3. In short, we find that assuming independence degrades optimal profits by just 0.5% on average in our simulations and that the directional insights of Section 2.4 continue to hold for most instances when we model these correlations.

Throughout the paper, superscripts denote the product featured in the promotional display. When product $i \in S$ is in the promotional display, the associated first stage probabilities, denoted by P_0^i, P_D^i, and P_A^i are as follows:

$$P_0^i = \mathbb{P}\{\text{no purchase from display and no aisle visit}\} = \frac{e^{Y_0}}{e^{Y_0} + e^{Y_i} + e^{-c+\lambda \cdot I}}, \tag{2.3}$$

$$P_D^i = \mathbb{P}\{\text{purchase from the display}\} = \frac{e^{Y_i}}{e^{Y_0} + e^{Y_i} + e^{-c+\lambda \cdot I}}, \tag{2.4}$$

$$P_A^i = \mathbb{P}\{\text{visit the aisle}\} = \frac{e^{-c+\lambda \cdot I}}{e^{Y_0} + e^{Y_i} + e^{-c+\lambda \cdot I}}, \tag{2.5}$$

where Y_0 is the utility associated with the no-purchase alternative and $I = \ln\left(\sum_{j \in S \cup \{0\}} e^{Y_j/\lambda}\right)$. The quantity I is often called the "inclusive value" of a nest in NMNL models. It links the first and second stages of the decision making process by conveying information about the expected value of the aisle as viewed from the first stage. The expected maximum utility obtained from nest k is $\mathbb{E}_\epsilon\left[\max_{j \in B_k}\{Y_j + \epsilon_j\}\right] = \lambda_k I_k + \gamma$, where $\gamma \approx 0.5772$ is Euler's constant [McFadden, 1978]. Therefore, the expected utility of choosing the aisle option in the first stage is the sum of the expected utility obtained by choosing the best alternative in the aisle ($\lambda \cdot I$) and the transit cost ($-c$). Note that parameter γ is irrelevant in the first stage probabilities because it is common across all nests.

We denote the second stage probability of product j, i.e., the conditional probability of choosing product j given that the customer visits the aisle, by $p_{j|A}$. We have $p_{j|A} = \frac{e^{Y_j/\lambda}}{\sum_{q \in S \cup \{0\}} e^{Y_q/\lambda}}$ for $j \in S \cup \{0\}$. When product i is in the display, the unconditional probability of choosing product j in the aisle is $P_A^i \cdot p_{j|A}$ for $j \in S \cup \{0\}$. Notice that the probability of choosing product j in the aisle is proportional to $e^{Y_j/\lambda}$. Since λ is the same for all products in the aisle, we refer to e^{Y_j} as the "popularity" of product j.

We can interpret the parameter λ as a measure of "customer heterogeneity." In Figure 2.2, we illustrate the second stage choice probabilities in the two extremes, when $\lambda = 10^{-3}$ and $\lambda = 1$, for a set of five fictitious alternatives with fixed mean utilities $\mathbf{Y}$. When $\lambda = 10^{-3}$, the random components of the associated utilities are almost fully correlated, so the alternative associated with the highest mean utility (alternative 1) is chosen with probability close to 1. On the other hand, the random components of the associated utilities are independent when $\lambda = 1$, so the choice probabilities spread out across alternatives in this case.

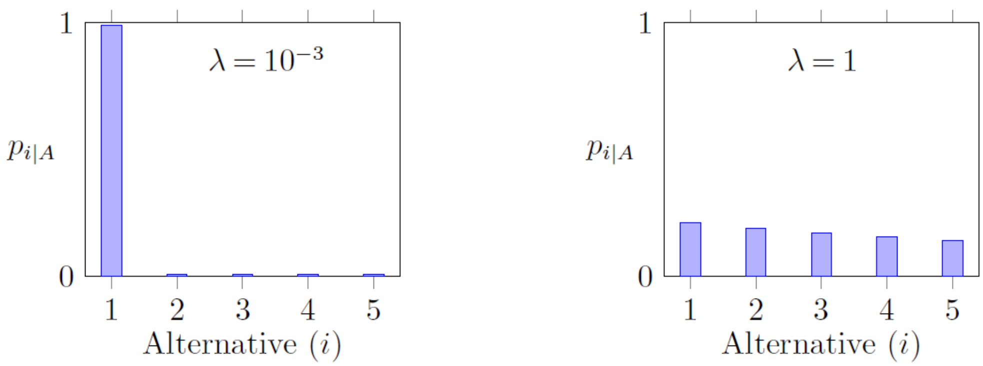


Figure 2.2: Second stage choice probabilities ($p_{i|A}$) when $\lambda = 10^{-3}$ and $\lambda = 1$.

In order to express the retailer's expected profit per customer, we must consider the two sources of the retailer's profit: (1) the customer's potential purchase of one of the products in the focal category, and (2) the impulse spending in case the customer visits the aisle. The impulse spending in the aisle contributes to a key aspect of our model, which is the value of driving customers to the aisle. Indeed, customer aisle visits can result in additional complementary purchases, therefore we assume that the retailer expects to earn r per customer as a result of this unplanned spending. In practice, retailers try to attract customers to the aisles by strategically deciding product locations [Rupp, 2015]. Moreover, Hui et al. [2013] point out that "encouraging customers to travel more of the store may increase unplanned spending by exposing them to more product stimuli". Lastly, POPAI [2012] reports that the percent of basket purchases on impulse increases in the number of aisles visited by a customer.

Let π^i denote the expected profit per customer when the retailer chooses product i to promote in the display. For brevity, we define $\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) = \sum_{j \in S} p_{j|A} \cdot m_j$, which is an average of the margins of all products in the category weighted by the second stage probabilities. We can express π^i as:

$$\pi^i = P_D^i \cdot m_i + P_A^i \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big) \ , \, i \in S. \tag{2.6}$$

The first term of equation (2.6) arises because if product $i \in S$ is featured on the display, an average customer purchases the featured product from the display with probability P_D^i, in which case the retailer collects the profit margin m_i of the focal product. Alternatively, the average customer visits the aisle with probability P_A^i, where she is expected to spend an amount of $\big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)$. Finally, the average customer chooses not to purchase from the category with probability P_0^i, in which case the retailer's profit is 0.

We discussed in the introduction that a promotional display has two primary impacts on demand: it expands the potential overall category demand by evoking the category in customers' minds (the demand expansion effect), and it boosts demand for the focal product at the expense of competitor products and of the no-purchase option (the substitution effect). The model developed so far represents our core model of baseline demand, accounting for the substitution effect only. We will later extend it to reflect the demand expansion effect. Under the assumption that all products from a category are uniformly good at generating demand expansion for the category, the model presented here is sufficient for our study of the choice of product from a category (Section 2.4). Demand expansion effects will become relevant when we consider the value of promotional displays across categories (Section 2.5) and product-specific demand expansion effects (Section 2.6).

Throughout Sections 2.4 and 2.5 we will make the following assumption that the expected impulse spending in the aisle (r) is not too high.

Assumption 2.1. $\max_{i \in S}\{m_i\} > P_A^0 \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)$ *where* $P_A^0 = \dfrac{e^{-c+\lambda \cdot I}}{e^{Y_0} + e^{-c+\lambda \cdot I}}$.

This assumption ensures that promoting some product in S is more profitable than promoting a hypothetical product j with preference weight $e^{Y_j} = 0$ (equivalently, $Y_j = -\infty$.) We will motivate and present a modification to this assumption in Section 2.6.

2.4 Product Choice

As discussed previously, in this section we examine the retailer's choice of which product from a given category to feature in a promotional display. We leave the valuation of the display across categories to Section 2.5.

This section is organized as follows. In Section 2.4.1, we eliminate some of the products from consideration based solely on their product characteristics (specifically, margin and popularity) irrespective of the category parameters r, c, and λ, thereby reducing the full category of products to an "efficient subset" of products that are candidates for the promotional display. In Section 2.4.2, we consider how the retailer's optimal choice of product to promote within the efficient subset changes with respect to the category parameters.

2.4.1 Characterization of the Efficient Set of Products

The problem of characterizing the product that maximizes the overall expected profit per customer when featured in a promotional display is difficult due to its discrete nature. Nevertheless, we can eliminate some products from consideration and identify structure around the retailer's optimal choice.

Table 2.1: The classification of products

	lower margin	higher margin
more popular	traffic builder	winner
less popular	loser	sleeper

We will use the two-by-two classification of products according to their popularities and margins shown in Table 2.1, which is similar to a classification used by Stoops et al. [1988]. For any two products 1 and 2 in a category, we either have a "winner" and a "loser" if $Y_1 \geq Y_2$ and $m_1 \geq m_2$, or a "traffic builder" and a "sleeper" if $Y_1 > Y_2$ and $m_1 < m_2$. A "traffic builder" is a relatively popular product that tends to drive traffic to its location. A "sleeper" product is less popular (i.e., it "sleeps"), but it generates higher returns when sold. We cannot unequivocally say that traffic builders are more attractive to promote than sleepers (or vice versa) in general, but we can make a general comparison between a winner product and a loser product. We generalize the notion of "loser" to the case of $n \geq 2$ products by saying that product j is a "dominated product" if there exists another product i in the category such that $Y_i \geq Y_j$ and $m_i \geq m_j$.

Proposition 2.1. *A dominated product is never optimal to promote.*

After eliminating dominated products from consideration, we call the remaining set of products the "efficient set," and denote it by E. More formally, E can be defined as follows:

$$E = \{i \in S \mid \nexists\, j \in S \text{ s.t. } Y_j \geq Y_i \text{ and } m_j \geq m_i\} \tag{2.7}$$

For any two products in E, one can be considered a traffic builder (higher popularity, lower margin) and the other a sleeper (lower popularity, higher margin) relative to each other.

To refine the efficient set further, we define a product-specific index $\theta_i = \dfrac{e^{Y_i} \cdot m_i}{e^{Y_0} + e^{Y_i}}$. We call θ_i the "profitability index" of product i as it is the expected profit in an MNL model with only product i and the no-purchase alternative in the choice set. As formally stated in Assumption 2.2, we assume hereafter that the profitability index of any product in the category is less than or equal to the expected profit obtained from the aisle, conditional on the customer visiting the aisle.

Assumption 2.2. *Product characteristics* $\mathbf{Y}$ *and* $\mathbf{m}$ *are such that* $\theta_i \leq \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)\ \forall i \in S$.

When this assumption is violated the problem becomes trivial and the retailer would be better off offering product i alone in the display and removing all the other products from the aisle. By using the profitability index, we can further refine the efficient set as stated in Proposition 2.2.

Proposition 2.2. *Consider two products,* i *and* j*, such that* $Y_i < Y_j$*. It is never optimal to promote product* j *if* $\theta_i \geq \theta_j$*.*

Proposition 2.2 says that any promoted product should have a higher profitability index than all the other less popular products in the efficient set. Let E' denote the set of products remaining in the efficient set after refining according to Proposition 2.2, or more formally,

$$E' = \{i \in E \mid \nexists\, j \in E \text{ s.t. } Y_i > Y_j \text{ and } \theta_j \geq \theta_i\} \tag{2.8}$$

Hereafter we order products in E' such that $Y_1 > .. > Y_{|E'|}$ and $m_1 < .. < m_{|E'|}$ w.l.o.g. Therefore, products with low indices can be considered traffic builders relative to products with higher indices (sleepers). Note that we also have $\theta_1 > .. > \theta_{|E'|}$ by Proposition 2.2.

In order to highlight the importance of the retailer's product choice, we compared in a numerical study the optimal expected profit to those of three natural heuristic policies; in particular, policies choosing the most popular product (Y-Policy), the highest margin product (m-Policy), and the product with the highest profitability index (θ-Policy). In Table 2.2, we report the average and maximum optimality gaps over 100 synthetically generated product category instances, each of which includes 15 products. Each product is characterized by its mean utility Y and the profit margin m that are randomly chosen from two uniform distributions with a certain correlation. We considered four different combinations of the impulse spending parameter r and the transit cost parameter c. (The details of the parameter settings are presented in Appendix 2.)

Table 2.2: Percentage optimality gaps corresponding to three heuristic policies

		Y-Policy		m-Policy		θ-Policy	
r	c	Avg.	Max	Avg.	Max	Avg.	Max
Low	Low	7.73	12.16	2.42	5.99	0.50	3.80
Low	High	13.29	22.02	8.06	17.71	0.40	4.05
High	Low	8.07	11.39	0.58	2.29	1.21	5.90
High	High	13.36	20.58	4.34	10.57	0.98	6.99

Note that the Y-Policy performs remarkably worse than the other two policies on average, because it often chooses a product from outside the set E'. The θ-Policy performs the best of the three heuristics on average. This makes sense in light of Proposition 2.2, which suggests that products with higher profitability index θ_i tend to be favored within the set E'. Although performing well on average, the m-Policy and θ-Policy may still lead to losses as large as 18% and 7%, respectively. The m-Policy's worst average performance (8%) is when r is low and c is high, while the θ-Policy's best average performance (0.4%) is achieved in that scenario. In addition, θ-Policy's worst average performance (1.21%) is when r is high and c is low, while the m-Policy's best average performance (0.58%) is achieved in that scenario. This reinforces the importance of understanding the optimal policy as a function of the problem parameters.

2.4.2 Sensitivity Analysis on Retailer's Optimal Decision

The refined efficient set E' is defined only in terms of product popularities and margins, and is independent of values of the category parameters r, c, and λ. However, the optimal choice of product from E' depends on the category parameters, and here we seek to understand these dependencies.

We use $i^*(\cdot) = \underset{i \in S}{\operatorname{argmax}}\{\pi^i(\cdot)\}$ to refer to the retailer's optimal choice of product to promote.

The ordering of products in E' implies that promoting a lower indexed product (traffic builder) yields higher expected profit from the display $(P_D^i \cdot m_i)$, but lower expected profit from the aisle $(P_A^i \cdot (\bar{m}+r))$ compared to a higher indexed product (sleeper). Hence, the retailer's decision among products in E' trades off the display and the aisle profits.

Intuitively, this decision depends on two factors: the aisle's attractiveness for customers, and its profitability for the retailer. As mentioned in Section 2.3, the quantity $\lambda \cdot I(\mathbf{Y}, \lambda)$ shows a customer's expected utility obtained by choosing her best alternative in the aisle. Hence, we call the quantity $e^{-c+\lambda \cdot I(\mathbf{Y},\lambda)}$ the "aisle's attractiveness". Likewise, we call the quantity $\big(\bar{m} + r\big)$ the "aisle profit".

2.4.2.1 Parameters r and c

The transit cost c incurred by customers to reach a category's aisle space varies with several store- and category-specific factors. First, large store footprints are associated with higher search costs because the store size has a direct impact on customers' walking distance within the store to find the products they are looking for [Trivedi et al., 2016; Baumol and Ide, 1956]. Second, Larson et al. [2005] provide empirical evidence that grocery store customers tend to walk along the perimeter of the store, visiting a particular aisle only if a product they are looking for appears in that aisle. Hence, the proximity of the product category to the perimeter of the store is a determinant of how much customers need to deviate from their base trajectory within the store. Third, c may also vary with customers' expectation for convenience; e.g., in a "convenience" store setting c will reflect customers' high disutility to spending extra time in the store.

Similarly, the impulse parameter r is a category-specific parameter that depends on the complementarity of the category with other categories placed in the same aisle. Practitioners consider cross-category sales when deciding floor space allocation, and empirical researchers often consider specific pairs of co-located categories to tease out cross-category effects: e.g., cola and potato chip categories [Bezawada et al., 2009], detergent and fabric softener categories [Manchanda et al., 1999]. Proposition 2.3 shows the *ceteris paribus* impact of the parameters r and c on $i^*(\cdot)$.

Proposition 2.3. *The following statements hold:*

1. *$i^*(r)$ is non-decreasing in the expected impulse spending r.*

2. *$i^*(c)$ is non-increasing in the transit cost c.*

The proposition states that as expected impulse spending r increases, the optimal product to promote moves towards sleeper products. Increasing r increases the aisle profitability, driving the retailer towards featuring sleeper products in the display to keep aisle traffic high. On the other hand, when transit cost c increases, the aisle becomes unattractive to customers and the retailer seeks to emphasize display profits, captured by the immediate profitability θ_i. Recall that by Proposition 2.2, immediate profitability increases with the popularity of the featured product. These intuitions—that high aisle profits drive the retailer to keep traffic builders in the aisle, and that unattractive aisles tend to favor featuring profit-driving products—underly all of the results on product choice to follow.

2.4.2.2 Assortment Size

Next, we consider the effect of assortment size on the retailer's promotion choice. Assortment sizes vary across categories for a given retailer, and retailers vary in their assortment sizes for a given category. A larger assortment implies a more attractive aisle for customers, consistent with the conventional wisdom that customers value variety. In our model, a larger assortment translates to a higher inclusive value $I(\mathbf{Y}, \lambda)$ and hence a more attractive aisle. Specifically, it holds that $I(\mathbf{Y}_S, \lambda) < I(\mathbf{Y}_{S^+}, \lambda)$ for two assortments S and S^+ such that $S \subset S^+$.

Proposition 2.4. *Consider two assortments S and S^+ such that $S \subset S^+$ and $\bar{m}(\mathbf{Y}_S, \mathbf{m}_S) = \bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+})$. Then, we have $\underset{i \in S}{\operatorname{argmax}}\{\pi^i(S)\} \leq \underset{i \in S}{\operatorname{argmax}}\{\pi^i(S^+)\}$.*

We restrict the choice to the set S and assume equal average margins $\bar{m}$ to focus our insight on the impact of assortment breadth.

A recent examination of Target.com reveals that a specific Target store carries 14 brands of liquid laundry detergent and 102 brands of hair shampoo, implying that categories can vary significantly in assortment size. Assuming all else fixed, Proposition 2.4 implies that a relatively higher-margin and lower-popularity product (i.e., a sleeper) will tend to be preferred for the shampoo category due to its large assortment, whereas the opposite is true for laundry detergents.

2.4.2.3 Customer Heterogeneity Parameter λ

The last factor we consider to have an impact on the aisle's attractiveness is the customer heterogeneity parameter λ. As mentioned in Section 2.3, λ specifies the correlation among the random components of the utilities associated with the alternatives in the aisle. In particular, as λ increases in the interval $(0, 1]$, the random utility components become less dependent. This has implications on both the first and the second stage choice probabilities. Since $\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda)$ and $I(\mathbf{Y}, \lambda)$ are functions of λ, both the aisle profit for the retailer and the aisle's attractiveness for the customers depend on the customer heterogeneity.

Lemma 2.1. *The aisle's attractiveness,* $e^{-c+\lambda \cdot I(\mathbf{Y},\lambda)}$*, is increasing in* λ *in the interval* $(0, 1]$*.*

Lemma 2.1 implies that the aisle is more attractive for customers when their preferences are more heterogeneous. In other words, the same assortment in the aisle is appreciated more and attracts more customers when customers are more heterogeneous in their choices. Given Lemma 2.1, we can state the following result on the impact of λ on the retailer's optimal promotional display choice.

Proposition 2.5. *Assume that the average profit margin* $\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda)$ *is non-decreasing in* λ*. Then,* $i^*(\lambda)$ *is non-decreasing in* λ*.*

The intuition here is similar to that of Proposition 2.3 part 2: Promoting a sleeper product makes sense when the aisle is attractive. By Lemma 2.1, the aisle is most attractive for large λ. The assumption on the average profit margin holds if products' profit margins and popularities generally have an inverse relationship within a category. This is likely to be the case in practice because higher profit margin is needed to compensate for lower sales volume, and Vilcassim and Chintagunta [1995] show that optimal retailer profit margins are consistent with this intuition.

Considering its impact on the second stage choice probabilities as shown in Figure 2.2, the customer heterogeneity parameter λ can be considered a proxy for market structure. That is, a small λ leads to a concentrated market in which most customers choose the same product, whereas larger λ leads to a diversified marked in which the customer choice is spread out across multiple product variants. For instance, one might argue that the demand for laundry care products concentrates around a few brands, while demand for hair care products is dispersed, leading to a more attractive aisle relative to the product in the display. In 2016, US market shares of the leading brands in laundry and hair care categories are 27% and 6.5%, respectively [Euromonitor Int., 2017]. This suggests that sleeper products may be appealing to promote for the hair care category, while traffic builders may be good candidates for display in the laundry care category.

2.5 The Value of a Promotional Display to a Category

Section 2.4 considered the choice of product to display from a given category of products. There, the promotional display's primary impact was to shape demand among a set of products. In this section, we look at the value of a promotional display across categories, where a key driver will be the ability of a promotional display to expand the potential demand for a category (i.e., the "demand expansion effect"). We will continue for now to assume that the magnitude of the demand expansion effect is category-specific but not product-specific. We will consider product-specific demand expansion effects in Section 2.6.

Our goal in this section is to understand how the value of assigning a promotional display to a category depends on category characteristics, including the transit cost c, the customer preference heterogeneity as measured by the parameter λ, and the assortment size. A fourth category characteristic that will play a pivotal role is category "expandability," or the extent to which the potential market for a category can be increased by the promotional display. We argue that expandability varies across product groups, as demand for a product type may have an impulsive or utilitarian nature, implying that the demand expandability is large or small, respectively.

We measured demand expandability with the parameter $\beta \geq 0$, which represents the percentage increase in market size when the category is featured on a promotional display. Normalizing the baseline population to 1, we can then write the expected profit from the category when product i is promoted as

$$\Pi^i = \begin{cases} (1+\beta)\pi^i & \text{if } i \in S \\ P_A^0 \cdot (\bar{m} + r) & \text{if } i = 0 \text{ (no-promotion)} \end{cases} \tag{2.9}$$

Observe that when the category is featured in the promotional display, we achieve the same expected profit π^i as in Section 2.4, scaled by $(1+\beta)$ to reflect the expanded market brought by the display. An implicit assumption is that customers in the expanded market behave according to the same choice model as the baseline market. When the category is not displayed, the expected profit is based on the baseline (non-expanded) market size, which we normalize to 1.

We can interpret (2.9) as representing the profits from a set of "aware" customers whose knowledge of the category is not swayed by the promotional display, plus profits from a set (of size β) of "impressionable" customers who consider purchasing from the category only if a promotional display is present. We note that the promotional display still shapes the choices of an "aware" customer in the ways considered in Section 2.4, including enhancing her purchase probability by lowering the transit cost for the displayed product.

Let Δ^i denote the additional expected profit per customer due to featuring product i in a display. More formally,

$$\Delta^i = \Pi^i - \Pi^0 \tag{2.10}$$

Clearly, all else being equal, the value of the display is larger when the demand expansion, or equivalently β, is large. This justifies retailers promoting categories that are mostly purchased impulsively. However, there are other category features that also have impacts on the value of the display. Specifically, the aisle's attractivenes, which was critical in our analysis in Section 2.4.2, will also turn out to have a significant role in determining the value of the display to a category.

Our analysis reveals that β plays a pivotal role determining the net impact of the aisle attractiveness on the value of the display to a category. To clarify, consider for a moment the case with no demand expansion ($\beta = 0$). In this case, we can view the role of the display primarily as shaping substitution among products in the category. If the aisle is highly attractive, the effectiveness of the display in switching customers to the focal product will be limited. Hence, the aisle attractiveness decreases the value of the display. Next, consider the case where the demand expansion is significantly large. Here, a primary role of the display is to advertise the category. An attractive aisle converts more of this awareness into purchases. Hence, aisle attractiveness increases the value of the display in this case.

Similar to the approach we followed in Section 4.2, we will now consider the determinants of the aisle attractiveness for customers – namely the assortment size, customer heterogeneity, and the associated transit cost – and investigate each of their impacts on Δ^i in isolation; that is, holding other dimensions constant. In our online supplement, we discuss a numerical study that shows these results continue to hold even when compared categories differ in more than one dimension, as long as the compared categories are *roughly* similar.

2.5.1 Assortment Size

We first investigate the impact of assortment size on the value of the display. To this end, in this subsection we will consider $\Delta^i = \Delta^i(S)$, $\Pi^i = \Pi^i(S)$, and $\Pi^0 = \Pi^0(S)$ to be functions of the aisle assortments. In Proposition 2.6, we compare the value of the display for two assortments having different sizes:

Proposition 2.6. *Consider two assortments S and S^+ such that $S \subset S^+$ and $\bar{m}(\mathbf{Y}_S, \mathbf{m}_S) \leq \bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+})$. Let $j = \underset{i \in S}{\operatorname{argmax}}\{\Pi^i(S)\}$ and $k = \underset{i \in S}{\operatorname{argmax}}\{\Pi^i(S^+)\}$. There exists a threshold $\beta' \geq 0$ such that:*

1. *$\Delta^j(S) < \Delta^k(S^+)$ for $\beta > \beta'$ and*
2. *$\Delta^j(S) > \Delta^k(S^+)$ for $\beta < \beta'$.*

We note that the threshold β' depends on the problem parameters, including S and S^+, nontrivially. The quantity Δ^i is a function of the probabilities P_A^0, P_A^i, and P_D^i, each of which is a function of the assortment, and it also depends on the average aisle margin $\bar{m}$ which in turn depends on the assortment through the second stage probabilities. In short, Δ^i is a complex function of the assortment S (and of other category parameters, too). The proof of Proposition 2.6 follows from some monotonicity properties of the underlying probabilities, and a more intricate argument using an intermediate quantity in which we fix the average margin. We use related arguments to prove Propositions 2.7 and 2.8 as well.

Proposition 2.6 reveals that the impact of the assortment size pivots on the parameter β. When the demand expansion (β) is large, the value of the display is higher for the larger assortment S^+ because it encourages more customers to make a purchase from the category by offering them a larger variety. In contrast, when the demand expansion is small, the display is a means of controlling brand choice of the baseline population of customers, and a larger assortment makes the aisle more attractive and blunts the power of the display. We can extend Proposition 6 to three assortments as follows:

Corollary 2.1. *Consider $S \subset S^{+} \subset S^{++}$ such that*

$$\bar{m}(\mathbf{Y}_S, \mathbf{m}_S) \leq \bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}) \leq \bar{m}(\mathbf{Y}_{S^{++}}, \mathbf{m}_{S^{++}}).$$

Let $j = \underset{i \in S}{\operatorname{argmax}}\{\Pi^i(S)\}$, $k = \underset{i \in S}{\operatorname{argmax}}\{\Pi^i(S^+)\}$, and $l = \underset{i \in S}{\operatorname{argmax}}\{\Pi^i(S^{++})\}$. There exist $0 \leq \beta_1 \leq \beta_2$ such that

1. $\max\{\Delta^j(S), \Delta^k(S^+)\} < \Delta^l(S^{++})$ *for* $\beta > \beta_2$
2. $\max\{\Delta^j(S), \Delta^l(S^{++})\} < \Delta^k(S^+)$ *for* $\beta_1 < \beta < \beta_2$
3. $\max\{\Delta^k(S^+), \Delta^l(S^{++})\} < \Delta^j(S)$ *for* $\beta < \beta_1$

The corollary shows that the retailer is more likely to pick a product from the larger assortment when demand expansion effect gets stronger. It is straightforward to generalize this result to an arbitrary number of assortments. Similarly, it is straightforward to extend comparisons in Propositions 2.7 and 2.8 to more than two assortments as well.

These results have implications on the retailer's choice of category to promote. Consider, for example, the choice of endcap display between the chips category and the meat jerkies category, both part of the snacks product group in grocery stores. These two product categories are often located in the same aisle in a grocery store, and they can both be expected to have a large category demand expansion. Assuming the chips category has a larger assortment than jerkies, the first statement in Proposition 2.6 suggests that all else being equal, the display should be reserved for the chips category, which is more likely to generate a sale from a customer who is swayed by the display to consider the category.

2.5.2 Customer Heterogeneity

Recall that heterogeneity in preferences implies a more attractive aisle (Lemma 2.1), and a more attractive aisle in turn reduces both the probabilities of no purchase and of purchasing the displayed product. When β is high and the demand expansion effect dominates, it is to the retailer's advantage to increase the proportion of customers making a purchase. When beta is low and the substitution effect dominates, the retailer benefits from promoting a category which generates high profits from the display.

Proposition 2.7. *Let* $0 < \underline{\lambda} < \bar{\lambda} \leq 1$ *and assume that* $\bar{m}(\lambda)$ *is non-decreasing in* λ *for* $\lambda \in [\underline{\lambda}, \bar{\lambda}]$*. Then, there exists a threshold* $\bar{\beta} \geq 0$ *such that:*

1. $\Delta^{i^*(\underline{\lambda})}(\underline{\lambda}) < \Delta^{i^*(\bar{\lambda})}(\bar{\lambda})$ *for* $\beta > \bar{\beta}$ *and*

2. $\Delta^{i^*(\underline{\lambda})}(\underline{\lambda}) > \Delta^{i^*(\bar{\lambda})}(\bar{\lambda})$ *for* $\beta < \bar{\beta}$.

Similar to Proposition 2.5, we study what happens when $\bar{m}$ is non-decreasing in λ. Section 2.4 explains why this is a plausible assumption. Proposition 2.7 suggests that all else being equal, a category for which the demand is spread out across multiple products should be featured in the display when considering product groups with large demand expansions (e.g., snacks, beverages, condiments, etc.). In contrast, displaying a category with a concentrated market structure may be better for product groups with smaller demand expansion relative to the baseline demand (e.g., personal care, household cleaners, paper products, etc.).

2.5.3 Transit Cost

Retailers carefully determine the specific locations of the aisle storage for each category, e.g., popular items are routinely located in the middle of aisles [Rupp, 2015]. This implies that customers incur different transit costs to reach the aisle storage of different categories. When the substitution effect dominates, a category that is easy for customers to reach will benefit least from a promotional display because the display will have limited ability to shape customer demand. When the demand expansion effect dominates, an easy-to-reach category will be best positioned to monetize the increased customer interest arising from demand expansion. These results are formalized as follows:

Proposition 2.8. *Let* $0 \leq \underline{c} \leq \bar{c}$*. There exists a threshold* $\tilde{\beta} \geq 0$ *such that:*

1. $\Delta^{i^*(\underline{c})}(\underline{c}) > \Delta^{i^*(\bar{c})}(\bar{c})$ *for* $\beta > \tilde{\beta}$ *and*

2. $\Delta^{i^*(\underline{c})}(\underline{c}) < \Delta^{i^*(\bar{c})}(\bar{c})$ *for* $\beta < \tilde{\beta}$.

This result is closely related to the specific store layout and the aisle arrangement under consideration. For instance, consider a pallet display reserved for categories belonging to personal care products. Assuming that the baseline demand of personal care products is large relative to the demand expansion, the second part of Proposition 2.8 suggests that the display is more valuable to a category with a less visible shelf storage (e.g., mid-aisle) compared to another category positioned in a more prominent location in the store (e.g., end-aisle).

2.6 Product Choice Under Product-Specific Demand Expansion

We have assumed so far that all the products from a category are equally good at expanding the demand for a category; i.e., the demand expansion effect is constant across products in the same category. In this section we relax this assumption and revisit the choice of product to display from a given category. To that end, define the modified profit function $\tilde{\Pi}^i$ for the expected profit obtained when product i is promoted:

$$\tilde{\Pi}^i = \begin{cases} (1+\beta\phi_i)\pi^i & \text{if } i \in S \\ \Pi^0 & \text{if } i = 0 \text{ (no-promotion)} \end{cases} \tag{2.11}$$

where π^i is the original profit function defined in (2.9) when product i is promoted, and $\phi_i \in [0,1]$ is the product-specific demand expansion parameter. Note that the expected profit functions Π^i and $\tilde{\Pi}^i$ are the same when $\phi_i = 1$ for all $i \in S$, yielding the case we analyzed in Section 2.4. We make the intuitive assumption that $\phi_i \geq \phi_j$ if and only if $Y_i \geq Y_j$, meaning that a more popular product leads to a larger demand expansion when promoted.

The addition of product-specific demand expansion effects complicates our earlier analysis. As the following example shows, Assumption 2.1 is no longer sufficient for Proposition 2.1 to hold.

Example 2.1. $\mathbf{Y} = [\ln(2.5), \ln(1), \ln(0.1)]$, $\mathbf{m} = [0.015, 0.01, 1]$, $\boldsymbol{\phi} = [1, 0.99, .01]$, $\beta = 4$, $r = 0$, $c = 0$, $\lambda = 1$.

Although the second product is dominated by the first product in the above example, it is the optimal product to promote. Since only m_3 is larger than the expected aisle profit $(\bar{m} + r)$, it is better to promote product 3 in the original model. On the other hand, it has a very small demand expansion effect (ϕ_3) in the modified model. This example suggests that a stronger assumption is needed for the dominated product result to hold.

Assumption 2.3. *Let* $\phi_{max} = \max_{i \in S}\{\phi_i\}$*. There exists a product* k *such that* $\tilde{\Pi}^k/(1+\beta\phi_{max}) > \tilde{\Pi}^0$*.*

Recall that Assumption 2.1 requires that promoting some product from the category is more profitable than promoting a vanishingly unpopular product. For a product-independent category expansion effect $1+\beta$, this would imply the existence of a product k such that $\Pi^k/(1+\beta) > \Pi^0$, i.e., promoting product k is more profitable than no-promotion even without the advantage of the demand expansion effect. Assumption 2.3 resembles Assumption 2.1 and implies that displaying product k is more profitable than a hypothetical scenario with no-promotion and the largest possible customer base of size $(1+\beta\phi_{max})$.

Proposition 2.9. *Under Assumption 2.3, a dominated product is never optimal to promote in the case of product-specific demand expansion.*

Proposition 2.9 implies that the main trade-off between popularities and margins of the products in the efficient set is maintained in product-specific demand expansion case.

Our sensitivity results (Propositions 2.3-2.5) also require some modification under product-specific category expansion effects, but our core insights from Section 2.4.2 remain intact. Proposition 2.3, part 1, supported the intuition that as the aisle becomes more profitable to the retailer, the retailer should increase the fraction of customers visiting the aisle, thereby featuring products with lower popularities and (equivalently) higher aisle probabilities P_A^i's. Under product-specific demand expansion effects, as r increases the retailer again seeks to drive more customers to the aisle, but the total quantity of customers visiting the aisle is $P_A^i(1+\beta\phi_i)$ when product i is promoted. Therefore, as r increases it will become increasingly favorable to feature products with higher value of $P_A^i(1+\beta\phi_i)$.

Proposition 2.10. *Under product-specific capacity expansion effects, for* $r_1 > r_2$ *we have* $P_A^{i^*(r_1)}(1+\beta\phi_{i^*(r_1)}) \geq P_A^{i^*(r_2)}(1+\beta\phi_{i^*(r_2)})$*.*

Because P_i^A is decreasing in product popularity Y_i and ϕ_i is non-decreasing in Y_i, how $P_A^i(1+\phi_i)$ index depends on product popularity depends on the relation between Y_i and ϕ_i. Given this relation (e.g., $\phi_i = \frac{e^{Y_i}}{e^{Y_{max}}}$, or $\phi_i = \frac{e^{Y_i}}{\sum_{j\in S} e^{Y_j}}$, $\forall i \in S$), we can characterize a threshold for the demand expansion β such that a higher r leads to promoting a more popular product and vice versa.

The remaining sensitivities analyzed in Section 2.4.2—to c, λ, and assortment size—are consistent with the intuition that as the aisle becomes less attractive to customers, the retailer prioritizes the profitability of the display, there represented by the index θ_i. This intuition remains intact with product-specific demand expansion effects, but the index $\eta_i = \theta_i(1 + \beta\phi_i)$ becomes a more accurate representation of the display profit. Literally, θ_i gives the retailer's profit when product i is featured in the promotional display assuming the aisle is prohibitively unattractive. Though we are not able to prove direct analogs of Propositions 2.3 (part 2), 2.4, and 2.5, it is straightforward to show that as aisle attractiveness vanishes (i.e., $e^{-c+\lambda I} \to 0$) it becomes optimal to promote the product $\underset{i\in S}{\operatorname{argmax}}\{\theta_i\}$ (Recall that the aisle attractiveness $e^{-c+\lambda I}$ increases with λ and assortment size and decreases with c). We note that as $e^{-c+\lambda I} \to +\infty$, it becomes optimal to promote the product $\underset{i\in S}{\operatorname{argmax}}\{\phi_i\}$. That is, when the aisle is very attractive to customers, the display's primary role is to advertise the category, and the retailer should promote the product which provides the most effective advertisement.

2.7 Concluding Remarks

We have focused on the effects of promotional displays, and we leave the integration of other promotional activities such as price discounts and advertising as a future research direction. A given set of price discounts and advertising activities determine an instance of our problem; therefore, our model yields insights into the impact of a promotional display given other promotional activities. However, we have not considered the simultaneous optimization of various promotional activities. There is a rich empirical literature on the effects and interactions of different promotional activities, suggesting broader mechanisms such as increased store traffic, brand switching, store-switching, and stockpiling that could be captured by future analytical models.

Another future research opportunity would be to look at how manufacturers and retailers negotiate and contract on which products to feature in promotional displays. Our model suggests that the true value of a promotional display slot is not one-size-fits-all. This value depends on the characteristics of the product and the store. Furthermore, it must reflect externalities on demands for other products and on other product categories through customer traffic patterns. Understanding this value, which we have studied in this paper, is important for a retailer evaluating manufacturers' bids for space in promotional displays.

CHAPTER 3: TRADE DEALS AND RETAILER PROMOTIONS

3.1 Introduction

In Chapter 2, we took the retailer's perspective and studied the management of promotional displays in vacuum; that is, we excluded the negotiations over promotional displays between retailers and manufacturers, which in practice often result in manufacturers offering incentives to retailers to feature certain products. These incentives are called "*trade promotions*" (or "*trade deals*") in the literature and can take various forms such as per-unit discounts on the wholesale price of units bought ("*off-invoice*") or the units sold ("*scan-back*") by the retailer, or fixed payments (e.g., display allowances). CPG manufacturers in U.S. spent more than 100 billion dollars in trade promotions in 2016, which corresponds to 46% percent of their marketing budget [Cannondale Associates, 2017]. Therefore, it is important to understand how retailers' promotional decisions regarding retail price discounts, promotional displays, etc., are affected by manufacturers' efforts, which play a central role in this chapter.

The integration of potential price discounts in the management of promotional displays is important because retailers often support price discounts with promotional displays to attract customers' attention. Arguably the two most important objectives of a manufacturer in offering a trade deal is to induce retailers to offer a retail price discount and/or to feature the brand on a promotional display in addition to the others such as pushing inventory to the retailer or inducing price fluctuations in the wholesale market [Neslin, 1990]. The synergy between price promotions and displays is documented in the literature [Zhang, 2006]. Our aim in this project is to study this synergy by taking a normative approach and to characterize the products in a category for which the retailer would have a price discount, a display, or both when there is/isn't a trade deal offered by a manufacturer.

Conventional wisdom suggests that brands that cannibalize other brands in the category are not good targets for retailer promotions unless they have higher gross margins than the other brands in the category [Neslin, 1990]. Since trade deals temporarily increase the retailer's effective margin for a product, it might be in the retailer's best interest to promote a product with a relatively low margin via discounts and/or displays depending on the trade deal conditions. We showed in Chapter 2 that the popularity and the profit margin together (in addition to some category features) determine a product's performance when displayed without a price change. In this chapter, we allow retailer to implement price discounts in addition to displays and include trade deals from manufacturers into our analysis.

Manufacturers are concerned about the effectiveness of trade deals for several reasons. The first concern is the retail "*pass through*", which is defined as the percentage of trade promotion dollars that the retailer spends on retail promotions to customers [Ailawadi and Harlam, 2009]. Since the direct control of the retail price by the manufacturer (and therefore the pass through rate in a trade deal) is illegal (Sherman Act), retailers can pocket the trade deal instead of passing through to the customers. This significantly affects profitability of the trade deals for the manufacturer. There are two opposing sides in the literature regarding the impact of product popularity on retailers' pass-through rate: Some claim that the retailer pass through rate is higher for high share products because they lead to a higher promotional lift [Besanko et al., 2005; Pauwels, 2007]. On the other hand, others claim that high share products have higher baseline sales, which leaves less room for additional profits for retailers [Tellis and Zufryden, 1995; Van Heerde and Neslin, 2017].

Second, it is not always in retailer's best interests to support the price discount on a product with a promotional display, because the retailer is interested in the overall category profits and the customer flow (i.e., customer visits in specific aisles) within store rather than boosting the sales of a particular product. In a typical trade deal, the manufacturer requires that the retailer promotes his product in some way to be eligible for trade deal benefits, but the retailer decides on the specific mix of marketing efforts (e.g. retail price discount, display, store flyers etc.), and the retailer's decision is not necessarily the best for the manufacturer. It is shown empirically that high-share and high-margin products receive greater retailer response in practice (higher pass through and display support) when the manufacturer offers a trade deal [Besanko et al., 2005; Pauwels, 2007], but retailers' response to trade deals from a high-margin and low-popularity manufacturer versus a low-margin and high-popularity manufacturer is not analyzed. We aim to shed light into this comparison.

Third, retailers forward-buy when there is a trade deal, and customers forward-buy when the retailer promotes a product, both of which have a long term impact on the performance of promotions. Manufacturers can overcome retailer forward-buying behavior by applying the so called "scan-back" deals, which are based on units sold by the retailer rather than units bought [Drèze and Bell, 2003]. We recognize that the customer forward buying behavior negatively affects both the retailers' and the manufacturers' long term gains from promotions, but we do not capture such inter-temporal effects of promotions in our model and assume a myopic retailer and manufacturers.

Following a review of relevant literature (Section 3.2) and the introduction of our model (Section 3.3), we present our analysis in two main parts. In Section 3.4, we examine a retailer's own effort to promote a product in a given category in the absence of a trade deal. We first consider the retailer's joint pricing problem under the MNL model and show how the well-known equal-margin result (Anderson and de Palma 1992, Besanko et al. 1998, Aydin and Ryan 2000) changes when a product is displayed. Next, we extend our "efficient set" result from Chapter 2 for when the retailer can discount the price of the displayed product. We characterize the condition for when the retailer should support a certain level of price promotion with a display in terms of product popularity, margin and other category features. Finally, we show that the retailer's optimal price discount is higher for a high-margin and low-share product.

In Section 3.5, we focus on trade deals. First, we consider the retailer's response to a given trade deal offer. We show structure on the retailer's best response (level of pass through and display decision) to a trade deal offer from a manufacturer. We find out that higher-margin and lower-popularity products receive greater response from the retailer in a trade deal. A product receives smaller pass through when displayed, which suggests that the manufacturer should not always request display support from the retailer. In fact, we show that if the manufacturer is interested in retailer's display support, a fixed payment to the retailer is better than offering a per-unit discount on the wholesale price.

Proofs of all analytical results are presented in Appendix 5.

3.2 Related Literature

Our study is related to a large stream of research on multi-product price optimization under the MNL model and its variants. Hanson and Martin [1996] show that the objective function in the multi-product pricing problem need not be jointly concave (or even quasi-concave) in the prices, but the transformed problem with market shares as the decision variables is jointly concave [Li and Huh, 2011]. Several studies show that the optimal markup, defined as the price minus cost for each product, is the same for all products (Anderson and de Palma 1992, Besanko et al. 1998, Aydin and Ryan 2000). Aydin and Porteus [2008] show that the objective function is quasi-concave in each price individually, and there exists a unique price vector satisfying the first order conditions. Finally, Gallego and Wang [2014] show that the adjusted nest-level markup is nest-invariant at optimality in the NMNL model. Following this literature, we employ the MNL framework to explicitly model the effect of price on customer choice to study retailer's follow-up decision to pricing, which is the promotion decision involving temporary price discounts and promotional displays. Assuming retailer's optimal pricing as suggested by this literature, we show structure of the retailer's optimal promotional strategy.

There is a large stream of empirical research on estimating and explaining the effects of retailer price discounts on the demand within a product category by using household-level [Gupta, 1988; Chintagunta, 1993; Bell et al., 1999; Pauwels et al., 2002; Van Heerde et al., 2003] and store-level [Van Heerde et al., 2004; Nair et al., 2005] data. Different from these studies, we take a normative approach to find out how the retailer should promote via price discounts and/or promotional displays when a manufacturer incentivizes for its own product to be promoted.

Trade deals received significant attention in the literature because manufacturers spend tremendously on trade deals, but their effectiveness is controversial [Nijs et al., 2010]. Earlier analytical research [Gerstner and Hess, 1991, 1995] points out decentralized pricing decisions in supply chains as the reason for ineffective trade deals, and suggests subsidizing customers (pull promotion) rather than retailers (push promotions) to achieve coordination. Since manufacturers cannot control retail prices, retailer pass through rates vary significantly across products and categories. There is empirical evidence that brands high-share and high-margin products receive higher pass through rates in practice [Besanko et al., 2005; Ailawadi and Harlam, 2009; Pauwels, 2007; Ailawadi and

Harlam, 2009]. However, normative studies [Tellis and Zufryden, 1995; Van Heerde and Neslin, 2017] suggest that high-share products should have smaller pass through rates because they have higher baseline sales and less room for demand increase by promotions. Different from these studies, we focus on the management and effects of promotional displays in addtion to price discounts within the context of manufacturer-retailer interactions.

There are empirical studies in marketing focusing on promotional displays [Chevalier, 1975; Bezawada et al., 2009; Nordfält and Lange, 2013; Phillips et al., 2015]. There is an agreement in this literature that the displays have profound effects on the demand within a product category, the effectiveness of price promotions and the customer traffic flow within a store. These effects depend on several external factors such as the display's location within the store, and the demand structure in the product category, etc. The synergy between display and price discounts has also been studied, but there is no consensus on the sign of the synergy. Gupta [1988] estimates negative interactions between display and price discounts and concludes overlapping effects whereas Papatla and Krishnamurthi [1996] find out magnifying effects. Zhang [2006] shows how customers perceive displays as price cut proxies. These studies are descriptive and do not consider trade deals.

3.3 The Model

We consider a retailer and a category consisting of $n \geq 2$ products which we indicate by $\mathcal{N} = \{1, .., n\}$. Each product has a distinct manufacturer. We denote the regular wholesale and retail prices of product $i \in \mathcal{N}$ by w_i and p_i, respectively as shown in Figure 3.1. We denote the retailer's markup for product i by $m_i = p_i - w_i$. Similar to Chapter 2, the retailer has the option of promoting one of the products by featuring it in an in-store promotional display ("display" hereafter) in addition to its regular shelf space in the aisle. We use "aisle" to refer the location in the store permanently allocated to the category. Each product in the category has its own shelf space in the aisle and we do not model any differences in the aisle space allocated to products.

$$\boxed{\text{Manufacturer } i} \xrightarrow{w_i} \boxed{\text{Retailer}} \xrightarrow{p_i}$$

Figure 3.1: Product i's wholesale price is w_i and the retail price is p_i.

The utility a customer receives from alternative i is

$$U_i = \overbrace{\alpha_i - bp_i}^{Y_i} + \epsilon_i \tag{3.1}$$

where α_i is the mean utility from all the attributes of product i other than the price, and b is the product invariant price sensitivity of customers. Therefore, the net mean utility obtained from product i is $Y_i = \alpha_i - bp_i$. ϵ_i's are assumed to be independent and identically distributed following the standard Gumbel distribution with the c.d.f.

$$\Phi(\epsilon_i) = e^{-e^{-\epsilon_i}}, \forall i \in \mathcal{N} \cup \{0\} \tag{3.2}$$

Although the distribution of ϵ_i's are independent, which implies the standard MNL model (with one nest), we virtually consider the same nest structure we described in Section 2.3. That is, we assume that customers incur a "transit cost" if they choose to visit the aisle, which we model as a nest-specific disutility, c, associated with all the alternatives in the aisle. Moreover, customer aisle visits can result in additional complementary purchases, therefore we assume that the retailer expects to earn r per customer as a result of this unplanned spending. Although the distributional assumption restricts the previous model, we extended it in a different direction now by allowing the retailer to discount the displayed product's price by $\delta \geq 0$. Figure 3.2 demonstrates the customer choice scenario when product i is displayed with a price discount of δ.

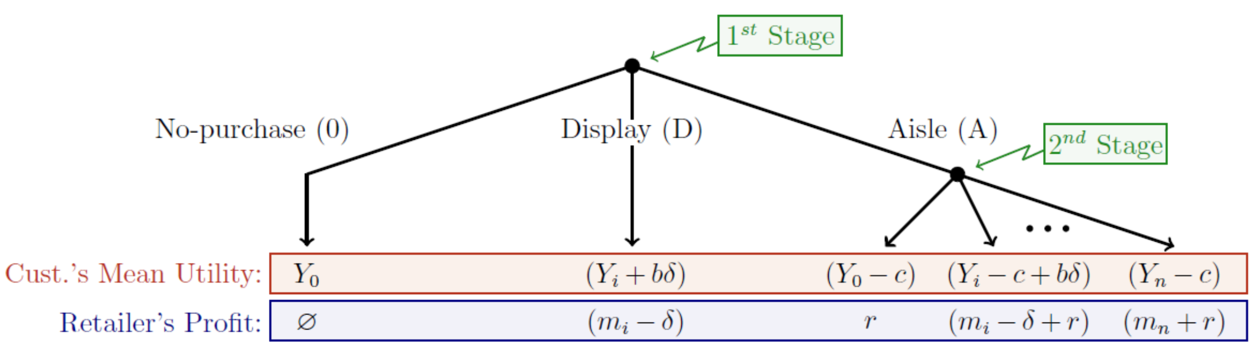


Figure 3.2: Customer choice modeled as a two-stage decision process.

Note that the first stage choice probabilities now depend not only on which product is being featured on a display, but also the level of discount δ. Hence, we need to modify the notation of first stage choice probabilities we used in Chapter 2. We denote retailer's display decision (or action) by $a \in \{D, ND\}$, which correspond to display and no-display, respectively. When product $i \in \mathcal{N}$ is discounted by $\delta \geq 0$, the associated first stage probabilities, depending on the display decision a, are denoted by $P_0^{i,a}(\delta)$, $P_D^{i,a}(\delta)$, and $P_A^{i,a}(\delta)$, which are expressed as follows:

$$P_0^{i,a}(\delta) = \frac{e^{Y_0}}{e^{Y_0} + \mathbb{1}_{\{a=D\}} \cdot e^{Y_i + b\delta} + e^{-c}\left(e^{Y_0} + \sum_{j \in \mathcal{N} \setminus \{i\}} e^{Y_j} + e^{Y_i + b\delta}\right)}, \tag{3.3}$$

$$P_D^{i,a}(\delta) = \frac{\mathbb{1}_{\{a=D\}} \cdot e^{Y_i + b\delta}}{e^{Y_0} + \mathbb{1}_{\{a=D\}} \cdot e^{Y_i + b\delta} + e^{-c}\left(e^{Y_0} + \sum_{j \in \mathcal{N} \setminus \{i\}} e^{Y_j} + e^{Y_i + b\delta}\right)}, \tag{3.4}$$

$$P_A^{i,a}(\delta) = \frac{e^{-c}\left(e^{Y_0} + \sum_{j \in \mathcal{N} \setminus \{i\}} e^{Y_j} + e^{Y_i + b\delta}\right)}{e^{Y_0} + \mathbb{1}_{\{a=D\}} \cdot e^{Y_i + b\delta} + e^{-c}\left(e^{Y_0} + \sum_{j \in \mathcal{N} \setminus \{i\}} e^{Y_j} + e^{Y_i + b\delta}\right)}, \tag{3.5}$$

where Y_0 is the utility associated with the no-purchase alternative. We denote the second stage probability of product j, i.e., the conditional probability of choosing product j given that the customer visits the aisle, by $p^i_{j|A}(\delta)$. Note that the conditional probabilities $p_{j|A}$ for $j \in \mathcal{N}$ does not depend on the displayed product in Chapter 2. On the other hand, since we now allow the displayed product to have a discount (on the display and in the aisle), the conditional probabilities now depend on the displayed product and the discount amount δ as follows:

$$p^i_{j|A}(\delta) = \begin{cases} \dfrac{e^{Y_i + b\delta}}{e^{Y_0} + \sum_{j \in \mathcal{N} \setminus \{i\}} e^{Y_j} + e^{Y_i + b\delta}} & \text{if } j = i, \\ \dfrac{e^{Y_j}}{e^{Y_0} + \sum_{j \in \mathcal{N} \setminus \{i\}} e^{Y_j} + e^{Y_i + b\delta}} & \text{if } j \neq i \end{cases} \tag{3.6}$$

When product i is in the display with a discount of $\delta \geq 0$, the unconditional probability of choosing product j in the aisle is $P^i_A(\delta) \cdot p^i_{j|A}(\delta)$ for $j \in \mathcal{N} \cup \{0\}$. Similar to Chapter 2, we refer to e^{Y_j} as the "popularity" or the "market share" of product j. For brevity, we also define $\bar{m}^i(\mathbf{Y}, \mathbf{m}, \delta) = \sum_{j \in \mathcal{N}} p^i_{j|A}(\delta) \cdot (m_j - \delta \cdot \mathbb{1}_{\{j=i\}})$, which is an average of the margins of all products in the aisle weighted by the second stage probabilities when product i is discounted by δ. Using this notation, we can now express the retailer's profit ($\pi_R^{i,a}(\delta)$) when product i is being discounted by $\delta \geq 0$, and featured by a display ($a = D$ for display) or without a display support ($a = ND$ for no-display) as follows:

$$\pi_R^{i,a}(\delta) = \begin{cases} P_D^{i,D}(\delta) \cdot (m_i - \delta) + P_A^{i,D}(\delta)(\bar{m}^i(\mathbf{Y}, \mathbf{m}, \delta) + r) & \text{if } a = D \\ \\ P_A^{i,ND}(\delta)(\bar{m}^i(\mathbf{Y}, \mathbf{m}, \delta) + r) & \text{if } a = ND \end{cases} \tag{3.7}$$

For ease of exposition in the following sections, we next define two quantities, Σ and M, that are independent of the displayed product and the discount level δ. Σ is the sum of the preference weights of the no-purchase and the aisle alternatives (i.e., a measure of baseline overall demand for the category), and M is the baseline level of profit retailer gets when there is no display ($a = ND$) or discount ($\delta = 0$).

$$\Sigma = e^{Y_0} + e^{-c + \ln\left(\sum_{j=0}^{n} e^{Y_j}\right)} \tag{3.8}$$

$$M = \frac{\Sigma - e^{Y_0}}{\Sigma} \left(\frac{e^{Y_1} m_1 + .. + e^{Y_n} m_n}{e^{Y_0} + e^{Y_1} + .. + e^{Y_n}} + r \right) \tag{3.9}$$

Using these quantities, we can express $\pi_R^{i,a}(\delta)$ in a more tractable format as shown in Equations (A5.1 and A5.2), where the baseline profit of the retailer (M) is separated out from the impact of retailer's promotion. This significantly simplifies our analysis.

3.4 Retailer Induced Promotions

In this section, we study the retailer's problem of which product to promote, how much to discount, and whether to display the product on a promotional display in the absence of a trade deal offered by a manufacturer. The problem of how much to discount a given product is conceptually related to the well-studied multi-product pricing problem under the MNL choice model in the operations management literature, which we call the *retailer's joint pricing problem* and is formulated as follows:

$$\max_{p_1,..,p_n} \sum_{i \in \mathcal{N}} (p_i - w_i) \frac{e^{\alpha_i - bp_i}}{e^{Y_0} + \sum_{j \in \mathcal{N}} e^{\alpha_j - bp_j}} \tag{3.10}$$

It is worth mentioning here that we slightly modified the above problem to capture the dynamics of customer movement within a store and the two-stage customer decision making. To clarify, a customer who decides not to purchase any of the items without visiting the regular aisle space is different in our model from a customer who visits the aisle and then decides not to purchase. This means that our MNL structure is different from the studies mentioned above even in the absence of a promotional display because we have a no-purchase alternative in both stages. The retailer's joint pricing problem in our setting is formulated as follows:

$$\max_{p_1,..,p_n} \sum_{i\in\mathcal{N}}(p_i - w_i)\frac{e^{\alpha_i-c-bp_i}}{e^{Y_0}+\sum\limits_{j\in\mathcal{N}\cup\{0\}} e^{\alpha_j-c-bp_j}} + r\frac{\sum\limits_{j\in\mathcal{N}\cup\{0\}} e^{\alpha_j-c-bp_j}}{e^{Y_0}+\sum\limits_{j\in\mathcal{N}\cup\{0\}} e^{\alpha_j-c-bp_j}} \tag{3.11}$$

The following conjecture states that the equal markups result extends to our setting.

Proposition 3.1. *In Problem (3.11), the following hold*

1. *The markups $m_i = p_i - w_i$ are equal for each product at optimality.*

2. *The optimal markup decreases in r and c, and increases as the assortment expands.*

The first statement in Proposition 3.1 points out a nice structure in the optimal prices of products, and is equivalent of the classical equal margin result in our setting. Specifically, it suggests that the prices should be set such that the markups are equal for each product regardless of their popularities. The level of the optimal markup depends on the total demand, or equivalently the total probability of purchasing one of the items. The second statement in Proposition 3.1 implies that the retailer sets higher prices when an average customer is more likely to purchase one of the items in the category (i.e., when the transit cost c is small and/or the assortment is larger). This is because an average customer's surplus in this case is higher therefore the retailer can charge a higher price. On the contrary, the retailer sets lower prices when the expected impulse spending r is high, because lower prices attract more customers to the aisle.

We haven't considered the retailer's promotional activities (display and price discounts) so far. The retailer can pick at most one of the products to feature on a promotional display with a potential price discount. We assume that only one product in the category is chosen to be promoted at a time. The retailer's promotion problem (in the absence of a trade promotion offered by a manufacturer) is formulated as follows:

$$\max_{i \in \mathcal{N}} \left\{ \max_{\substack{0 \leq \delta \leq m_i \\ a \in \{D, ND\}}} \pi_R^{i,a}(\delta) \right\} \tag{3.12}$$

In words, the retailer chooses a product to promote, decides whether to display the product or not, and how much to discount. We denote an optimal solution to the above problem by (i^*, a^*, δ^*).

Definition 3.1. *Product $j \in \mathcal{N}$ is a "dominated product" if there exists another product $i \in \mathcal{N}$ in the category such that $Y_i \geq Y_j$ and $m_i \geq m_j$.*

Proposition 3.2. *In an optimal solution (i^*, a^*, δ^*), i^* is never a dominated product.*

Proposition 3.2 extends the efficient set result (Proposition 1) of Chapter 2 by allowing price discounts for the displayed product. It implies that a dominating product is always preferred over a dominated product for the same display decision (D or ND) and the level of discount. Note that Proposition 3.2 does not imply that promoting a dominated product (with a display and/or a discount) is worse than no promotion, i.e., the retailer might still be interested in promoting a dominated product. In practice, retailers rotate the promoted product in a category for several reasons such as maintaining good relationships with the manufacturers or the inter-temporal effects of promotions on customer purchasing behavior. This is why the inner maximization problem in (3.12) is a relevant question for a retailer. In words, we are interested in determining the best promotional activity (display or not, how much to discount) for a given product in the assortment, which might potentially be a dominated product. The following sheds light into this question.

Proposition 3.3. *For any product $i \in \mathcal{N}$ and for any $\delta \geq 0$,*

$$\pi_R^{i,D}(\delta) \geq \pi_R^{i,ND}(\delta) \iff m_i \geq M + \delta \left(1 - \frac{e^{Y_i - c}}{\Sigma}\right) + r\frac{e^{Y_i - c}}{\Sigma}(e^{b\delta} - 1). \tag{3.13}$$

Proposition 3.3 implies that the retailer should support a price promotion for a product with a display if and only if the markup for the product is sufficiently high with respect to the baseline total profit M, the discount level δ, and the expected impulse spending r in the aisle. There are immediate implications of this result regarding the inner maximization problem in (3.12): First, assuming $r = 0$ and δ to be the same, a higher margin and/or a higher share product is more likely to perform better when displayed versus not displayed. There is empirical support that such products are more likely to receive display support from the retailer in practice [Chevalier, 1975; Pauwels, 2007]. Moreover, this result explicitly describes how the market share and the profit margin of the product, in addition to other category related features (e.g. M, Σ, r), together determine whether a price discount for a product should be supported by a display. Therefore, it suggests an explanation for opposing results in the empirical literature regarding the synergy between the price promotions and displays [Gupta, 1988; Papatla and Krishnamurthi, 1996].

Assuming that the retailer sets the regular prices of the products in his assortment optimally according to Problem (3.11), Corollary 3.1 implies the optimal promotion strategy for the retailer.

Corollary 3.1. *Suppose the retailer sets the regular prices (p_i, $i \in \mathcal{N}$) optimally, and thus $m_i = m_j = m$ for any $i, j \in \mathcal{N}$. Then,*

1. $i^* = \underset{j \in \mathcal{N}}{\operatorname{argmax}}\{Y_j\}$.
2. $a^* = D$ *for any* $i \in \mathcal{N}$ *and* $\delta \geq 0 \iff m \geq \delta \left(1 + \dfrac{\sum_{j \in \mathcal{N}} e^{Y_j} - e^{Y_i}}{1 + e^c}\right) + r \dfrac{e^{Y_i}(e^{b\delta} - 1)}{1 + e^c}$

Corollary 3.1 points out that the retailer should promote the highest share product (part 1), and it provides hints about the optimal promotional mix. In particular, it points out when a given price discount δ should be supported by a promotional display (part 2), which is indeed a re-statement of the condition in Proposition (3.3). On the other hand, the optimal discount level δ^* is hard to obtain in closed-form. (Currently, we are looking for the necessary and sufficient conditions for δ to be positive. An obvious necessary condition is $m_i \geq M$), but it is not sufficient. It is a difficult analysis because the effective margin of the focal product decreases in δ whereas its purchase probability is increasing, so the net change is analytically hard to measure due to the complexity of purchase probability expressions.)

Conjecture 3.1. *Assume that $Y_i > Y_j$ and $m_i < m_j$ for all $i < j \in \mathcal{N}$. The optimal discount level is higher for a higher share and/or higher margin product, i.e.,* $\underset{0 \leq \delta \leq m_i}{\operatorname{argmax}}\{\pi_R^{i,D}(\delta)\}$ *and* $\underset{0 \leq \delta \leq m_i}{\operatorname{argmax}}\{\pi_R^{i,ND}(\delta)\}$ *are both increasing in i.*

We note that this result is analytically unproven so far, but we have strong numerical support for it. In our numerical simulations, we used the same random category instances (and high-low values of c and r) that we used in Chapter 2, that are synthetically created as described in Appendix 2. The only difference is that we shuffled the popularities and product margins in each category such that $Y_i \geq Y_j$ and $m_i \leq m_j$ for all pairs of $i, j \in \mathcal{N}$, i.e., all the products in the category are in the efficient set. We considered two levels ($b_L = 5, b_H = 10$) of the price sensitivty parameter. In this setting, a total of 84 thousand pairs of products (100 categories x $\binom{15}{2}$ pairs x 2^3 different scenarios of b, c, r) verify this conjecture.

3.5 Manufacturer Induced Promotions

Section 3.4 considered the retailer's own effort to promote a product without any incentive provided by a manufacturer. In this section we consider the retailer's response to a trade deal offered by one of the manufacturers, and discuss its implications for different manufacturers. After formulating the retailer's problem, we show how product and category characteristics affect the optimal pass-through rate and the display decision of the retailer. Then, we show that the retailer tends to pass a smaller portion of the discount to the customers when the product is featured with a promotional display. Finally, we show that a manufacturer is better-off paying the minimum fixed payment that ensures retailer's display support than offering the minimum per-unit discount on the wholesale price that achieves the same.

Suppose manufacturer i offers a discount of δ on its wholesale price, and the retailer decides on the pass-through rate $\alpha \in [0, 1]$ and the display action $a \in \{D, ND\}$. A customer's choice scenario and the retailer's profit associated with each alternative in this case is demonstrated in Figure 3.3. The retailer's problem is formulated as follows:

$$\max_{\substack{0 \leq \alpha \leq 1 \\ a \in \{D, ND\}}} \pi_M^{i,a}(\delta, \alpha) \tag{3.14}$$

where

$$\pi_M^{i,a}(\delta, \alpha) = \begin{cases} P_D^{i,D}(\alpha\delta) \cdot (m_i + (1-\alpha)\delta) + P_A^{i,D}(\alpha\delta)(\bar{\bar{m}}^i(\mathbf{Y}, \mathbf{m}, \delta, \alpha) + r) & \text{if } a = D \\ \\ P_A^{i,ND}(\alpha\delta)(\bar{\bar{m}}^i(\mathbf{Y}, \mathbf{m}, \delta, \alpha) + r) & \text{if } a = ND \end{cases} \tag{3.15}$$

and $\bar{\bar{m}}^i(\mathbf{Y}, \mathbf{m}, \delta, \alpha) = \sum_{j\in\mathcal{N}} p^i_{j|A}(\alpha\delta) \cdot (m_j + (1-\alpha)\delta \cdot \mathbb{1}_{\{j=i\}})$

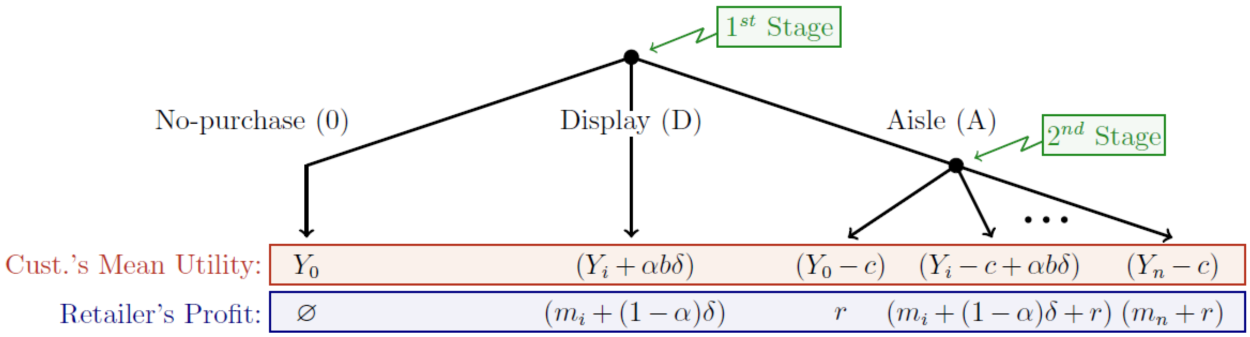


Figure 3.3: Customer choice modeled as a two-stage decision process.

In problem (3.14), the retailer does not choose a product to promote, but responds to a given trade deal offer from a manufacturer. For simplicity, we limit our analysis to the retailer's best response to a trade deal offer from a single manufacturer at a time, but it is difficult to find a closed for solution to the problem (3.14). Hence, we will first consider a simpler version of the problem where the pass-through rate α is exogenously determined, and discuss the insights it provides for certain scenarios in practice.

In practice, it is not always retailer's best interest to support a price discount with a promotional display. Chevalier [1975] and Pauwels [2007] report that high share and high margin products are more likely to be featured on a display when their prices are discounted. The following proposition formally states to which products the retailer should provide a display support given a trade deal of size δ from its manufacturer, and an exogenously determined pass-through rate α:

Proposition 3.4. *Assume $r = 0$. Fix the manufacturer's trade deal offer $\delta \geq 0$ and the retailer's pass through rate $\alpha \in [0, 1]$. The retailer is better-off displaying product i if and only only if $m_i \geq M - \delta(1-\alpha)\left(1 - \frac{e^{Y_i - c}}{\Sigma}\right)$.*

Proposition 3.4 implies that any trade deal offer $\delta > 0$ is sufficient for manufacturer i to guarantee the display support from the retailer if $m_i > M$. The manufacturer of a lower margin product, on the other hand, has to offer a larger trade deal δ to acquire display support from the retailer. Proposition 3.4 is also related to the other merchandising actions the manufacturer can take to induce retailer to display his product. For instance, the manufacturer can choose to incentivize customers directly (e.g., providing coupons to the customers), which can be considered as a special case of our model where $\alpha = 1$. Alternatively, the manufacturer may subsidize a retailer with no pass-through realized (e.g. paying display allowance to an everyday-low-price retailer), which can be considered as the case of $\alpha = 0$ in our model. We consider these two special cases in the following corollary:

Corollary 3.2. *Assume $r = 0$. The following hold:*

$$\pi^i_{M,D}(\delta, 1) > \pi^i_{M,ND}(\delta, 1) \iff m_i \geq M \tag{3.16}$$

$$\pi^i_{M,D}(\delta, 0) > \pi^i_{M,ND}(\delta, 0) \iff m_i \geq M - \delta\left(1 - \frac{e^{Y_i - c}}{\Sigma}\right) \tag{3.17}$$

The first part of Corollary 3.2 points out that a high-margin manufacturer ($m_i > M$) can induce the retailer to display his product by directly subsidizing customers with a deal of any size $\delta > 0$, but a low-margin manufacturer can never achieve it regardless of the size of the deal δ. This implies for manufacturers that directly subsidizing customers is not an effective means of acquiring display support from the retailer. On the other hand, the second part of Corollary 3.2 suggests that even a low-margin manufacturer can get his product displayed with a large enough incentive (display allowance) paid to the retailer.

It is worth mentioning here how the minimum δ a low margin manufacturer ($m_i < M$) has to offer in order to secure the display support from a retailer with zero pass through depends on product and category characteristics. It turns out that the lower the margin the bigger the minimum display allowance should be to compensate the cannibalization of the demand for other products due to displaying a low-margin product. Moreover, the higher the market share ($e^{Y_i - c}/\Sigma$) of a low-margin product, the bigger the minimum display allowance should be, because the display will lead to an even bigger cannibalization within the overall category demand for the retailer in this case. Finally, the bigger the baseline sales (M), and the more attractive the aisle to an average customer (Σ), the larger the minimum display allowance the low-margin manufacturer has to pay.

Our next result relaxes the conditions on the pass through rate α in Corollary 3.2. It provides an answer to the question of when it is profitable for the retailer to display a particular product assuming that he optimally determines the pass through rate α:

Proposition 3.5. *Assume $r = 0$, and $Y_i > Y_j$, $m_i < m_j$ for all $i < j \in \mathcal{N}$. Let $\bar{I}$ be the maximum index $i \in \mathcal{N}$ such that $m_i \geq M - \delta \left(1 - \frac{e^{Y_i - c}}{\Sigma}\right)$.*

1. *For products $1, .., \bar{I}$,*

$$\max_{0 \leq \alpha \leq 1} \{\pi_M^{i,D}(\delta, \alpha)\} \leq \max_{0 \leq \alpha \leq 1} \{\pi_M^{i,ND}(\delta, \alpha)\} \text{ and } \operatorname*{argmax}_{0 \leq \alpha \leq 1} \{\pi_M^{i,ND}(\delta, \alpha)\} = 0. \tag{3.18}$$

2. *For products $\bar{I}, .., |\mathcal{N}|$,*

$$\max_{0 \leq \alpha \leq 1} \{\pi_M^{i,D}(\delta, \alpha)\} \geq \max_{0 \leq \alpha \leq 1} \{\pi_M^{i,ND}(\delta, \alpha)\} \tag{3.19}$$

Proposition 3.5 provides a criterion for whether a retailer, who can optimally solve Problem (3.14), would prefer displaying a product. Note that the criterion is neither for a fixed pass through rate (like in Corollary 3.2), or for any pass trough rate $\alpha \in [0, 1]$, but rather assumes the optimal pass through rate according to Problem (3.14). It implies that for the same trade deal offer δ, the retailer is more likely to display a higher margin and/or less popular product assuming that he is optimally deciding on the pass through rate α for each product. Increasing the trade deal size δ makes it more likely for a retailer to provide a display support to a product, i.e., a manufacturer can get his product displayed with a better trade deal offer.

As mentioned earlier, there is a significant variation in the retailer pass through rate across products within a category. The following result explains how the optimal pass-through rate differs across products in a category:

Conjecture 3.2. *Assume that $Y_i \geq Y_j$ and $m_i \leq m_j$ for all $i < j \in \mathcal{N}$. The optimal pass-through rate is higher for a higher share and/or a higher margin product, i.e., $\operatorname*{argmax}_{0 \leq \alpha \leq 1} \{\pi_M^{i,D}(\alpha)\}$ and $\operatorname*{argmax}_{0 \leq \alpha \leq 1} \{\pi_M^{i,ND}(\alpha)\}$ are both increasing in i.*

Similar to Conjecture 3.1, this result is analytically unproven so far, but we have strong numerical support for it. The numerical simulation setting we described in Section 3.4 verifies this conjecture, too. Note that Conjecture 3.2 assumes a fixed display decision (D or ND) by the retailer and suggests that the retailer should apply a larger pass through to products with higher margin and/or lower-popularity. On the other hand, the display decision also has an impact on the retailer's optimal pass through rate. In particular, the retailer should pass different percentages of the discount δ to the customers in cases when he supports/does not support the promotion with a display. We show that the retailer's optimal pass through is smaller when the product is being featured on a promotional display as formally stated in the following proposition:

Proposition 3.6. *Featuring a product on a display reduces retailer's optimal pass-through rate, or more formally,* $\underset{0\le\alpha\le1}{\operatorname{argmax}}\left\{\pi_M^{i,D}(\delta,\alpha)\right\} \le \underset{0\le\alpha\le1}{\operatorname{argmax}}\left\{\pi_M^{i,ND}(\delta,\alpha)\right\}$ *for any* $i \in \mathcal{N}$ *and* $\delta \ge 0$.

The proof of the above proposition relies on the unimodality of the retailer's profit function $\pi_M^{i,D}(\delta,\alpha)$ in α, which is a modification of the quasi-concavity result by Aydin and Porteus [2008]. Proposition 3.6 implies that the retailer is over-discounting when he doesn't take the display into account. This is intuitive in the following sense: The retailer is passing through a portion of the discount only to increase the demand for the product, which now has a higher profit margin due to the incentive provided by the manufacturer. On the other hand, the display itself increases the demand for the product even without passing the discount to the customers. Hence, the retailer is better-off pocketing the extra profit margin obtained from the trade deal. This result is supported by Nijs et al. [2010], which empirically show that displayed products tend to receive less pass through from the retailers.

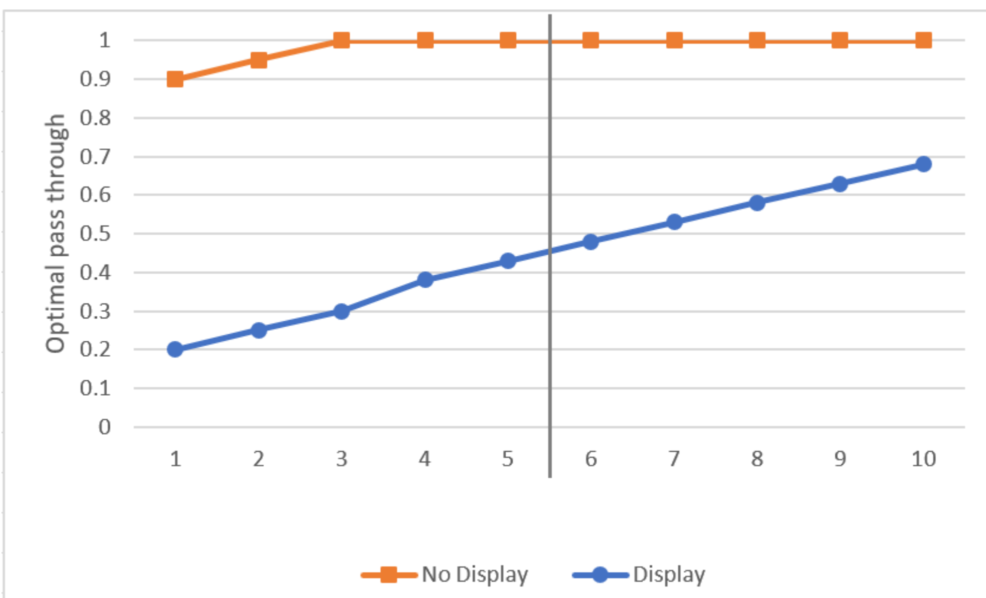


Figure 3.4: Optimal pass through rate for each product for $b = 1, \delta = 2, c = 1, r = 1$

We illustrate the results we discussed so far in Figure 3.4. We consider a product category that consists of 10 products such that $Y_i > Y_j$ and $m_i < m_j$ for $i < j$. When the trade deal δ is offered by one of the manufacturers 1-5 (the ones on the left hand side of the vertical line on the plot), the retailer should not use a promotional display to support the promotion. On the other hand, the promotion should be supported with a display when one of the manufacturers 6-10 (the ones on the right hand side of the vertical line on the plot) offers the same trade deal δ, which verifies Proposition 3.5. For a fixed display decision of the retailer, the optimal pass through rate is increasing as we move towards higher-margin and lower-popularity products, verifying Conjecture 3.2. Finally, for each product the optimal pass through rate is smaller when a promotional displayed is used as suggested by Proposition 3.6.

So far, we took the retailer's perspective and considered the best response of the retailer to a trade deal offer from a manufacturer. We next draw our attention to the manufacturers' incentives. Since the display leads to induce substitution in favor of the focal product at the expense of the sales from the other products in the category, the impact of displaying a product is more significant for the manufacturer than it is for the retailer. We aim to understand how a manufacturer can induce the retailer to feature his product on a display when the retailer decides on the pass-through rate. To that end, we consider two options for the manufacturer to acquire display support from the retailer. The manufacturer can offer a per unit trade deal δ, which we analyzed so far, or the manufacturer can offer a fixed payment F to the retailer. We are interested in understanding which one of these options the manufacturer should choose.

The manufacturer's problem is as follows:

$$\max_{\delta, F \geq 0} \left\{ \left[P_D^{i,D}(\alpha\delta) + P_A^{i,D}(\alpha\delta) p_{j|A}^i(\alpha\delta) \right] (w_i - \delta), \left[P_D^{i,D}(0) + P_A^{i,D}(0) p_{j|A}^i(0) \right] w_i - F \right\} \tag{3.20}$$

$$s.t. \max_{0 \leq \alpha \leq 1} \{\pi_M^{i,D}(\delta, \alpha)\} \geq \max_{0 \leq \alpha \leq 1} \{\pi_M^{i,ND}(\delta, \alpha)\} \tag{3.21}$$

$$\pi_M^{i,D}(0,0) + F \geq \pi_M^{i,ND}(0,0) \tag{3.22}$$

In the above problem, we implicitly assume that the manufacturer's ultimate goal is to maximize his profit by choosing the trade deal amount δ or choosing the fixed payment F while securing the display support from the retailer. It is trivial to separate the problem into two problems each of which is over a single decision variable (δ or F) and has the only constraint corresponding to that decision variable ((3.21) or (3.22), respectively). In fact, securing the display support might be too costly for a manufacturer, in which case he does not offer a trade deal or a fixed payment, however, we assume for our purposes that the manufacturer wants to get the display support from the retailer (there might be a long-term marketing objective such as increasing customer engagement and market share, which is not necessarily profitable in the short term).

We would like to compare manufacturer i's total cost of inducing the retailer to display his product via a per unit discount δ versus a total fixed payment F. The optimal trade deal δ manufacturer should offer to the retailer is difficult to find because it involves in constraint (3.21) the retailer's problem (3.14), to which we could not find a closed form solution. On the other hand, Proposition 3.5 suggests the minimum δ (not necessarily the optimal for Problem (3.20)) manufacturer i has to offer to secure the promotional display. Hence, we can calculate the manufacturer's payoff when he offers the minimum δ that guarantees retailer's display support, and compare it to the minimum fixed payment F that achieves the same.

Proposition 3.7. *The fixed payment F, which makes the retailer indifferent between featuring product i ($m_i < M$) with a promotional display or not, is smaller than the total loss of manufacturer i due to the minimum per unit deal δ that achieves retailer's indifference.*

Proposition 3.7 suggests that manufacturer i has a better payoff when he offers a fixed payment F_{min} compared to when he offers a per unit discount of δ_{min} in order to get the promotional display. Note however that δ_{min} is not the optimal per unit discount manufacturer i could offer, so the optimal per unit discount might potentially be better than the fixed payment.

3.6 Concluding Remarks

We extended our analysis of the management of promotional displays by allowing the retailer to offer price discounts to the customers, and studied the interplay between trade deals and retailer's promotion decisions. We assumed that the retailer promotes one product at a time, which rules out some scenarios, e.g., multiple products being discounted but only one of them being featured on a display etc., but we believe our setting is a common practice among retailers. Our results providence guidance to retailers in their own effort of managing the demand for a category, and in evaluating manufacturers' bids for gaining the retailer's support in favor of their own products. Therefore, our results also have relevance for manufacturers designing trade deal offers.

Our model is a step towards incorporating different levers the retailers use to modulate demand in a category, which have an impact on both the retailer's and the manufacturers' profits. On the other hand, there are other merchandising tools we do not consider in this study such as coupons, store flyers, radio ads, or buy-one-get-one-free deals etc. These would have significant affects especially on customer stockpiling and store-switching behavior. We leave the integration of such marketing efforts and customer reactions as a future research direction. Empirical marketing literature is rich on such retail promotions, but it is hard to incorporate them simultaneously in an analytical model to provide insights and prescriptive results for practitioners.

CHAPTER 4: MULTI-LOCATION ASSORTMENT PLANNING

4.1 Introduction

Retailers strive to satisfy a large spectrum of customer preferences to maximize sales by offering large product assortments. On the other hand, the operational cost of handling too many SKUs, scarce shelf space in stores and warehouses, limited purchasing budgets, and potential cannibalization of profit margins limit retailers' tendency toward product proliferation.

It is crucial for both online and brick-and-mortar retailers to have the right product variant when and where it is demanded by customers. To that end, retailers use different tactics for demand fulfillment that may not require physically holding each stock keeping unit in every location (store or distribution center). One such approach is to move products across different echelons of the supply chain, e.g., *drop-shipping* in online retail and *showrooming* in brick-and-mortar retail. The other way is the so-called *lateral transshipment* [Paterson et al., 2011] which means moving products across locations within the same echelon. The latter is especially advantageous for retail chains with a network of stores because it provides a significant demand fulfillment capability and competitive advantage over online retailers as former Gap CEO, Glen Murphy states, "Some people talk about Amazon with their 100 distribution centers, God bless them. We have 2,600 distribution centers" [Stelzer, 2017].

Lateral transshipments are reported in various retail industries including automotive, apparel, sporting goods, toys, furniture, information technology products, shoes, and spare parts (Çömez et al. 2012, Kukreja et al. 2001, Özdemir et al. 2006, Rudi et al. 2001, Hu et al. 2008). In the automobile dealership industry, "a customer who desires a brand but simply must have a blue car, not the red car on the lot, is almost a daily occurrence for a salesman. Having an inventory of the entire range of models and colors is simply impractical for most dealers. Over time, the practice has arisen of sharing models among dealers. In some cases this is highly formalized with shared web-based inquiry tools; in others, a simple call by the salesman seals the deal" [Zhao and Atkins, 2009]. Similarly, for fashionable items such as apparel and shoes, in many cases a salesperson can find a customer's choice of a product variant in one of the other outlets within the same retail chain and transship it.

The common practices of *pickup in-store* and *ship-from-store* (e.g. Kohl's, JCPenney, Target, etc.) also involve the practice of lateral transshipments. In the case of *pickup in-store*, a product variant that is not available in a certain customer's specified pickup store (either stocked-out or not offered in the assortment) can be shipped from another store that has the product in stock. *Ship-from-store* differs as to where the customer receives the product, but the demand is fulfilled by a store (rather than a distribution center), i.e., the last echelon of the supply chain.

Although the operations management literature includes rich work on single-product demand fulfillment in a network of stores, and centralized assortment planning, there has been little focus on assortment decisions in a multi-product and multi-location setting. In this paper, we consider such an assortment problem that is motivated by our discussions with a retail chain selling auto parts. The retailer typically has multiple stores within a local geographic region, and these local stores can transship products from each other by incurring a transshipment cost. From the retailer's perspective, transshipments enable a larger set of products to be offered to customers, and therefore increases the purchase probability. In addition, it allows an individual store to sell its exclusive products in an expanded set of markets. However, transshipments can be costly due to the logistical and transportation costs and the profit sharing between the sending and the receiving stores. This implies a lower profit margin for the products sold via transshipment, and therefore introduces *cannibalization* of profit obtained from a store's own products.

We employ the multinomial logit (MNL) framework to model customer choice within a category. A customer who visits a particular store may choose among the products included in the store's assortment, in which case the retailer collects the full profit margin. Alternatively, the customer may choose to wait for a product that is available in another store if the retailer offers the transshipment option in that store. In that case, the receiving store incurs the transshipment cost and shares the remaining profit with the sending store. We assume that a customer who waits for the transshipment incurs a constant disutility. Note that a customer's choice set is formed by the store's own assortment, the other stores' assortments, and whether the transshipment is enabled for that particular store. From a customer's perspective, transshipments provide a larger selection of products but can lead to a disutility of waiting.

Following a review of relevant literature (Section 4.2) and the introduction of our model (Section 4.3), we present our analysis in three parts. In Section 4.4, we take a central planner's perspective and examine the optimal assortment and transshipment decisions to maximize the total profits obtained from a network of stores. We show that the centralized problem is NP-Complete even for the case of two stores by using a reduction from the 2-partition problem. We show that it is optimal to transship from other stores if and only if the cost of transshipment is not larger than the expected opportunity cost of a customer leaving without a purchase, which depends only on the store's own assortment depth. Finally, we show structure on the optimal allocation of common and exclusive products across stores, which generalizes some earlier results from the classical assortment planning literature.

In Section 4.5, we consider the decentralized version of the problem with two stores. We discuss the existence and uniqueness of an equilibrium, and compare it to the solution of the centralized problem in terms of the product variety offered to the customers. We show how the coordination of the decentralized system can be achieved for a special case where products have equal popularities in Section 4.6. We assume throughout that transshipment of a product from one location to another is optional, i.e., a store manager (or the central planner) can choose not to offer the transshipment option to his customers for a product stocked in another location. In Section 4.7, we consider the case when transshipments are enforced.

Proofs of all analytical results are presented in Appendix 6.

4.2 Related Literature

4.2.1 Literature on Assortment Planning

Our study is related to a large literature on assortment planning, where the typical goal is to specify the set of products carried in a store (or in a sales channel) to maximize sales or gross margin under constraints on shelf space or on purchasing budget. A driving factor in this literature is customers' potential to substitute an available product when their original preference is not offered in the assortment, called assortment-based substitution [Kök et al., 2015]. (Give examples of papers modeling assortment based substitution in different ways: for MNL Ryzin and Mahajan 1999; Li 2007, for exogenous demand model Smith and Agrawal 2000, for locational choice Gaur and Honhon, 2006, for quality differentiation model Pan and Honhon 2012, etc.)

In practice, however, lateral transshipments (define earlier as *stock movements between locations of the same echolon*) enable retailers to provide customers with their first choice even if they are not included in their assortments. Our work contributes the literature on assortment planning by incorporating the impact of transshipments on optimal assortments.

4.2.2 Lateral Transshipments Among Retailers

There is a large stream of research on lateral transshipments (see Paterson et al. 2011 for a review). This literature mostly focuses on the impact of transshipments on the inventory management of a single product with a few exceptions of multi-product studies [Archibald et al., 1997; Wong et al., 2006; Kranenburg and Van Houtum, 2009]. These papers do not consider substitution, but rather extend the idea of inventory pooling through transshipments into a multi-product setting. Axsäter [2003] interprets unidirectional lateral transshipments as substitution among different quality levels in a single warehouse, but this approach lacks the multi-location aspect we focus on and ignores the disutility of customers' waiting.

4.2.3 Retailer Cooperation/Competition in Transshipment Systems

The above studies on transshipments consider only centralized decision making. Our work is also related with game-theoretical studies on decentralized inventory systems with transshipments [Rudi et al., 2001; Anupindi et al., 2001; Granot and Sošić, 2003; Slikker et al., 2005]. All these studies consider independent inventory decisions in the first stage and demand spillover in the second stage where cooperative transshipment decisions take place after demand is realized. These studies consider single item newsvendor type decisions, whereas we focus on assortment decisions of independent retailers with the availability of the transshipment option. For a model of competing retailers with multiple, substitutable products, see Netessine and Rudi [2003].

4.3 The Choice Model

We consider a network of stores which we indicate by $\mathcal{M} = \{1, .., M\}$, and a product category $\mathcal{N} = \{1, .., N\}$ of all available products variants. Each product is characterized by its associated mean utility Y_j, which is exogenously determined and can be interpreted as a function of product characteristics such as quality, features, price, etc. Following the literature, we assume that the profit margins of all products are the same for the retailer (see Ryzin and Mahajan 1999; Gaur and Honhon 2006) and are normalized to 1. Moreover, we assume that the cost of managing an assortment is linear in the number of products in the assortment, which implies a fixed cost $K \in [0, 1]$ of adding a product variant in an assortment (see Gaur and Honhon 2006; Kurtuluş and Nakkas 2011; Heese and Martínez-de Albéniz 2018; Aydin and Heese 2014).

A customer is equally likely to visit only one of the M stores, or equivalently, each store has an equal-size and exclusive customer population. We denote the assortment in store m by $S_m \subset \mathcal{N}$ and the list of assortments by $\mathbf{S} = (S_1, .., S_M)$. For ease of exposition, we also introduce the following notation for the set of products offered in at least one store ("*union set*"), and the set of products offered in every store ("*common set*"), respectively:

$$\text{Union Set: } \mathcal{S}_\vee = S_1 \cup .. \cup S_M \tag{4.1}$$

$$\text{Common Set: } \mathcal{S}_\wedge = S_1 \cap .. \cap S_M \tag{4.2}$$

In addition to the products in S_m, a customer's choice set in store m includes products that are available in other stores' assortments if store m offers the transshipments of those products. We use the vector $\mathbf{X} = \left[X_{j,m}\right] \in \{0, 1\}^{N \times M}$ to denote store m's binary decision of whether to offer the transshipment of product j from another store (obviously, $X_{j,m} = 0, \forall j \in S_m$ and $\forall j \notin \mathcal{S}_\vee$). The cost of transshipment is $\tau \in [0, 1]$, which is incurred by the retail chain as a whole or the receiving store individually in the centralized and decentralized settings, respectively. The transshipment leads to a disutility of waiting for a customer, which we denote by $c \geq 0$. A customer chooses the product in her choice set that maximizes her utility. The utility obtained from product j in store

m is as follows:

$$U_j = \begin{cases} Y_j + \epsilon_j, & \text{if } j \in S_m \\ Y_j - c + \epsilon_j, & \text{if } j \in \mathcal{S}_\vee \setminus S_m \end{cases} \tag{4.3}$$

We employ the MNL framework to model this customer choice scenario, which is commonly used in the operations management literature. In this model, ϵ_j's are assumed to be i.i.d. random variables following the standard Gumbel distribution. We normalize customer's utility from an outside option to $Y_0 = 0$. Throughout the paper, we call $v_j = e^{Y_j}$, which is the preference weight of product j in the MNL model, the *popularity* of product j because the MNL choice probabilities are proportional to this quantity. Without loss of generality, we index the products in their decreasing order of popularities, i.e., $Y_i \geq Y_j$ for all $i < j \in \mathcal{N}$. We use $\theta = e^{-c} \in (0, 1]$ to denote the discount on the utility of a product when it is transshipped due to customers' disutility of waiting. Using this notation, we express the probability that a customer in store m purchases product $j \in \mathcal{N}$ as:

$$p_j^m(\mathbf{S}, \mathbf{X}) = \begin{cases} \dfrac{v_j}{1 + \sum\limits_{k \in S_m} v_k + \sum\limits_{k \in \mathcal{S}_\vee \setminus S_m} X_{k,m} \theta v_k} & \text{if } j \in S_m, \\ \dfrac{X_{j,m} \theta v_j}{1 + \sum\limits_{k \in S_m} v_k + \sum\limits_{k \in \mathcal{S}_\vee \setminus S_m} X_{k,m} \theta v_k} & \text{if } j \in \mathcal{S}_\vee \setminus S_m \\ 0 & \text{if } j \in \mathcal{N} \setminus \mathcal{S}_\vee \end{cases} \tag{4.4}$$

4.4 The Centralized Problem

The centralized problem (CP, hereafter) is to maximize the total profit obtained from all stores. The decisions to be made for each store are the assortments ($S_m, m \in \mathcal{M}$) and the products to be transshipped from the other stores ($X_{j,m}, j \in \mathcal{S}_\vee, m \in \mathcal{M}$). The CP is formulated as follows:

$$\max_{\substack{S_1,..,S_M \\ X_{j,m}}} \sum_{m \in \mathcal{M}} \left(\sum_{j \in S_m} p_j^m(\mathbf{S}, \mathbf{X}) + (1 - \tau) \sum_{j \in \mathcal{S}_\vee \setminus S_m} p_j^m(\mathbf{S}, \mathbf{X}) - K \cdot |S_m| \right) \tag{4.5}$$

The trade-off in the CP is among (1) the opportunity cost of sending away a customer due to assortment-based substitution, (2) the cost of transshipments among stores, and (3) the fixed cost K of including a product in an assortment. We consider two extreme scenarios to shed light into this trade-off among three cost components. First, consider the case where all the stores have identical assortments and therefore no transshipments occur. This potentially leads to a high total fixed cost of product inclusion, and thus an opportunity for cost reduction by excluding certain products from some of the assortments, and instead relying on transshipments for those products. In the other extreme scenario, suppose that the stores offer mutually exclusive sets of products, that is, a large variety of products are accessible to customers with minimum total fixed cost of product inclusion. This enables the network of stores to minimize the opportunity cost of sending away a customer, but it also requires relying heavily on transshipments. An immediate question here is when transshipments are profitable.

We denote the optimal solution of the CP by $(S_1^C, .., S_M^C; X_{1,1}^C, .., X_{N,M}^C)$. For given assortments, the following lemma characterizes when transshipments are profitable for a store:

Lemma 4.1. *In an optimal solution of the CP,*
$X_{j,m}^C = 1$, $\forall j \in \mathcal{S}_\vee^C \setminus S_m^C$ *and* $\forall m \in \mathcal{M} \iff \tau \leq \dfrac{1}{1 + \sum\limits_{k \in S_m^C} v_k}$.

The above lemma is used as an intermediate result to prove some of the results later and also has a nice interpretation. It implies that transshipping products to a certain store is profitable if and only if the per unit transshipment cost is smaller than the risk of sending away a customer in that store. It also implies that either all or none of the products in $(\mathcal{S}_\vee^C \setminus S_m^C)$ are transshipped to store m from the other stores. Moreover, that decision depends only on store m's own assortment and the transshipment cost τ, i.e., store m's assortment decision by itself implies the optimal decision of whether to transship from the other stores.

We will next show that the CP is a difficult problem if one is interested in constructing an efficient algorithm to solve it optimally. Indeed, we can show that the CP is NP-complete even for two stores via a reduction from the 2-partition problem, which is known to be NP-complete [Karp, 1972].

Proposition 4.1 (NP-completeness)**.** *The centralized problem is NP-complete.*

Although the problem is difficult to solve optimally, the 2-partition problem is well studied in the literature, and there are several approximations with reasonable performance bounds [Graham, 1969; Johnson, 1974]. Our focus is not solution heuristics, rather to come up with insights about the assortment decisions in a multi-location setting with transshipments. To that end, we next characterize structure of the optimal solution of the CP.

Proposition 4.2 (Nested Structure)**.** *In an optimal solution of the CP,*

1. *There exist $k_1 \geq .. \geq k_M = |\mathcal{S}_\wedge^C|$ such that store m's assortment includes products $\{1, .., k_m\}$.*

2. *The products $\{k_1 + 1, .., |\mathcal{S}_\vee^C|\}$ are exclusive, i.e., each of them is offered in only one store.*

The above result extends the classical result of Ryzin and Mahajan [1999], which suggests that the optimal assortment in a single store setting consists of the most popular products. In the case of multiple stores with transshipments, the set of products offered at least in one of the stores ($\mathcal{S}_\vee^C$) is a popular subset of $\mathcal{N}$. Moreover, the set of common products ($\mathcal{S}_\wedge^C$) is a popular subset of $\mathcal{S}_\vee^C$. In Figure 4.1, we demonstrate the optimal assortments and the structure suggested in Proposition 4.2 on a numerical example consisting of 3 stores and 6 products.

	Product Variant					
	1	2	3	4	5	6
Store 1	✓	✓	✓			✓
Store 2	✓	✓	✓		✓	
Store 3	✓	✓		✓		

Figure 4.1: An example of optimal assortments of three stores with $k_1 = 3, k_2 = 3, k_3 = 2$.

Proposition 4.2 also implies that the number of stores offering a product in its assortment is (weakly) decreasing in the popularity of the product. For instance, in the example illustrated in Figure 4.1, the most popular two products are offered in all stores, followed by the third most popular product offered only by two stores, and finally some of the less popular products offered exclusively by a single store. Similar to the result of nested assortments by Davis et al. [2014], the optimal allocation of non exlusive products among stores is nested, i.e., if product j is a non exlusive product offered by store m, it is also offered by store $m' < m$.

Transshipments improve the objective value of the CP because transshipping a product to a store is a decision variable in our formulation, i.e., for any given assortment configuration, transshipments are performed if and only if it is profitable (Lemma 4.1). On the other hand, the impact of transshipments on the optimal assortments, and therefore on the product variety offered to customers is not trivial. The following proposition explains how transshipments affect the optimal assortments.

Proposition 4.3 (Transshipment Effects)**.** *In the centralized problem with two stores, transshipments lead to a smaller common set* ($\mathcal{S}_{\wedge}^{C}$) *and a larger union set* ($\mathcal{S}_{\vee}^{C}$)*.*

Proposition 4.3 reveals that the transshipments allow customers to access a larger set of products. On the other hand, introducing transshipments into the system shrinks the set of common products, which is a subset consisting of most popular products as suggested by Proposition 4.2. Intuitively, adding a product in an assortment has more value for the retailer when transshipment is possible because it allows a product to be sold in multiple markets. By the same reason, transshipment also makes it unnecessary to keep a product in multiple stores.

The proof of Proposition 4.3 uses the single store version of the problem. We denote the optimal solution of the single store version of the problem by S^*_{single}, and characterize it as follows:

$$S^*_{single} = \{1, .., k\} \text{ such that } \frac{v_k}{\left(\sum_{i=0}^{k-1} v_i\right)\left(\sum_{i=0}^{k} v_i\right)} \geq K \tag{4.6}$$

Note that S^*_{single} is the optimal assortment for each store for the case without transshipments. In the proof, we show that it contains the common set ($S^*_{single} \supseteq \mathcal{S}_{\wedge}^{C}$), and is a subset of the union set ($S^*_{single} \subseteq \mathcal{S}_{\vee}^{C}$) of the optimal solution of the case with transshipments.

4.5 The Decentralized Problem

In this section, we consider two independent stores (m and m') which determine their own assortments and whether to transship any products from the other store. When a product is transshipped, the sender receives an $\alpha \in [0, 1]$ portion of the profit obtained from that product whereas the receiver gets the remaining $(1-\alpha)$ portion. The receiver also incurs the transshipment cost τ. The profit store m collects from product j is:

$$\pi_j^m(\mathbf{S}, \mathbf{X}) = \begin{cases} p_j^m(\mathbf{S}, \mathbf{X}) + \alpha p_j^{m'}(\mathbf{S}, \mathbf{X}) - K & \text{if } j \in S_m \\ (1 - \alpha - \tau) \cdot p_j^m(\mathbf{S}, \mathbf{X}) & \text{if } j \in \mathcal{S}_{\vee} \setminus S_m \end{cases} \tag{4.7}$$

The decentralized problem for store m given the decision $(S_{m'}, \mathbf{X}_{m'})$ of store m' is:

$$\max_{S_m, \mathbf{X}_m} \pi^m(\mathbf{S}, \mathbf{X}) = \sum_{j \in (S_m \cup S_{m'})} \pi_j^m(S_m, S_{m'}, \mathbf{X}_m, \mathbf{X}_{m'}) \tag{4.8}$$

We denote the best response of store m as a function of the other stores' assortment and transshipment decisions by $(S_m^*; X_{1,m}^*, .., X_{N,m}^*)$ where we drop the arguments of the best response functions for the ease of exposition. The following lemma on the best response is useful to interpret and prove some of the following results.

Lemma 4.2. *The best response of store m to $S_{m'}$ satisfies the following:*

1. $X_{j,m}^* = 1, \forall j \in S_{m'} \setminus S_m^* \iff \alpha + \tau \leq \dfrac{1}{1 + \sum_{k \in S_m^*} v_k}$.

An evident implication of Lemma 4.2 is that $S_m = S_{m'} = S_{single}^*$ is the unique pure strategy Nash equilibrium for sufficiently large α. Obviously, there is no transshipment in the system in this equilibrium. The receiver's share $(1 - \alpha)$ is small relative to the transshipment cost τ, therefore transshipping a product leads to cannibalization for the receiver and none of the stores transship any products. Therefore, if there is no transshipment in the optimal solution of the centralized problem, it can be implemented in the decentralized system, because it is also an equilibrium.

Note that Lemma 4.2 implies that the transshipment decision of a store depends only on the store's own assortment and the quantity $(\alpha + \tau)$. This allows us to express the payoff function of store m, $\pi^m(\mathbf{S}, \mathbf{X})$, as a function of only S_1 and S_2 as follows:

$$\pi^m(S_1, S_2) \tag{4.9}$$

$$= \begin{cases} \dfrac{\sum_{k\in S_1} v_k}{1+\sum_{k\in S_1} v_k} - K\cdot|S_1| \text{ if } \dfrac{1}{1+\sum_{k\in S_1} v_k} < (\alpha+\tau) \text{ and } \dfrac{1}{1+\sum_{k\in S_2} v_k} < (\alpha+\tau) \\ \dfrac{\sum_{k\in S_1} v_k + (1-\alpha-\tau)\theta \sum_{k\in S_2\setminus S_1} v_k}{1+\sum_{k\in S_1} v_k + \theta \sum_{k\in S_2\setminus S_1} v_k} - K\cdot|S_1| \text{ if } \dfrac{1}{1+\sum_{k\in S_1} v_k} \geq (\alpha+\tau) \text{ and } \dfrac{1}{1+\sum_{k\in S_2} v_k} < (\alpha+\tau) \\ \dfrac{\sum_{k\in S_1} v_k}{1+\sum_{k\in S_1} v_k} + \dfrac{\alpha\theta \sum_{k\in S_1\setminus S_2} v_k}{1+\sum_{k\in S_2} v_k + \theta\sum_{k\in S_1\setminus S_2} v_k} - K\cdot|S_1| \text{ if } \dfrac{1}{1+\sum_{k\in S_1} v_k} < (\alpha+\tau) \text{ and } \dfrac{1}{1+\sum_{k\in S_2} v_k} \geq (\alpha+\tau) \\ \dfrac{\sum_{k\in S_1} v_k + (1-\alpha-\tau)\theta \sum_{k\in S_2\setminus S_1} v_k}{1+\sum_{k\in S_1} v_k + \theta \sum_{k\in S_2\setminus S_1} v_k} + \dfrac{\alpha\theta \sum_{k\in S_1\setminus S_2} v_k}{1+\sum_{k\in S_2} v_k + \theta\sum_{k\in S_1\setminus S_2} v_k} - K\cdot|S_1| \text{ o.w.} \end{cases} \tag{4.10}$$

It is worth mentioning here that neither existence nor uniqueness of a pure strategy Nash Equilibrium is guaranteed in general. (Every finite game has a mixed strategy equilibrium, but we are interested in pure strategy equilibriums for a simple interpretation.) In the following example, there is no pure strategy Nash equilibrium:

Example 4.1. $\mathbf{v} = [1.75, 0.9, 0.58, 0.57, 0.3], K = 0.04, \tau = 0.1, \alpha = 0.085, \theta = 0.75.$

	Product Variant				
	1	2	3	4	5
S_m^C	✓	✓			
$S_{m'}^C$	✓		✓	✓	

Figure 4.2: The optimal solution of the CP.

	Product Variant				
	1	2	3	4	5
$S_m^*(\emptyset)$	✓	✓	✓		
$S_{m'}^*(\{1,2,3\})$	✓			✓	
$S_m^*(\{1,4\})$	✓	✓			
$S_{m'}^*(\{1,2\})$	✓	✓	✓		

Figure 4.3: The series of best responses.

Although the existence of the pure strategy equilibrium is not guaranteed, we can show some structure on the best response and the equilibrium if one exists. The following lemma implies that a best response assortment is the union of a popular subset of products offered by the other store, and a popular subset of products not offered by that store.

Lemma 4.3. *For given $S_{m'}$, the best response of store m satisfies the following:*

1. *Pick $j, k \in S_{m'}$ s.t. $v_k > v_j$. $j \in S_m^* \Rightarrow k \in S_m^*$.*

2. Pick $j, k \in (\mathcal{N} \setminus S_{m'})$ s.t. $v_k > v_j$. $j \in S^*_m \Rightarrow k \in S^*_m$.

The following proposition is a direct consequence of the above observation. We denote a pure strategy Nash equilibrium by $\{S^D_m, S^D_{m'}\}$.

Proposition 4.4 (Equilibrium Structure)**.** *In an equilibrium of the decentralized problem (if it exists),*

1. *The set of products offered in at least one of the stores ($\mathcal{S}^D_\vee$) is a popular subset of $\mathcal{N}$.*

2. *The set of common products ($\mathcal{S}^D_\wedge$) is a popular subset of $\mathcal{S}^D_\vee$.*

Although the solution of the centralized problem and an equilibrium of the decentralized system both have the *popular assortment* structure, they do not necessarily yield the same assortments for the two stores. It is of our interest to compare the assortments in these two different regimes in terms of the product variety offered to the customers. To that end, we consider 3 measures of product variety in the centralized and the decentralized system:

1. The size of the union set: $|\mathcal{S}^C_\vee|$ versus $|\mathcal{S}^D_\vee|$.

2. The size of the common set: $|\mathcal{S}^C_\wedge|$ versus $|\mathcal{S}^D_\wedge|$.

3. The average number of products a customer has an access to: n^C versus n^D.

We define the average number of products a customer has an access to as follows: We assume that an average customer is equally likely to visit each store, and she has access to all the products in the store she visits. In addition, the customer has limited access (scaled by the factor θ) to the products in the other store if the transshipment is available ($X_{(\cdot,\cdot)} = 1$). More formally, we calculate n^C and n^D as follows:

$$n^C = 0.5\left(|S^C_1| + \theta \sum_{i \in S^C_2 \setminus S^C_1} X^C_{1,i}\right) + 0.5\left(\cdot|S^C_2| + \theta \sum_{i \in S^C_1 \setminus S^C_2} X^C_{2,i}\right) \tag{4.11}$$

$$n^D = 0.5\left(|S^D_1| + \theta \sum_{i \in S^D_2 \setminus S^D_1} X^D_{1,i}\right) + 0.5\left(\cdot|S^D_2| + \theta \sum_{i \in S^D_1 \setminus S^C_2} X^D_{2,i}\right) \tag{4.12}$$

We intuitively expect a larger union set in the centralized system. The union set, by definition, consists of the products offered in at least one of the stores. Given arbitrary assortments S_1 and

S_2, the marginal revenue of adding a product to one of these assortments is larger for both stores combined, compared to the marginal revenue of an individual store, whereas the fixed cost K is the same in the centralized and the decentralized systems. Therefore, we expect more products to be added to the union set in the centralized system (for example, if we consider adding products one by one without any common products).

These arguments are formally stated in the following conjecture:

Conjecture 4.1 (Centralized vs. Decentralized)**.** *The optimal assortments in the centralized* (S_1^C, S_2^C) *and the decentralized system* (S_1^D, S_2^D) *satisfy the following:*

1. *The number of products offered in at least one of the stores is larger in the centralized system, i.e.,* $\mathcal{S}_{\vee}^C \supseteq \mathcal{S}_{\vee}^D$. *(There are numerical counter-examples)*
2. *The number of products offered in both stores is smaller in the centralized system, i.e.,* $\mathcal{S}_{\wedge}^C \subseteq \mathcal{S}_{\wedge}^D$. *(No counter-example)*
3. *The average number of products a customer has an access to is larger in the centralized system, i.e.,* $n^C \geq n^D$. *(There are numerical counter-examples)*

We try to numerically verify the results in Conjecture 4.1. In Table A7.2, we report the comparison of the union set, common set, and the average number of products accessible to customers in the centralized and the decentralized systems. In each of the $2^4 = 16$ parameter setting (High/Low for K, θ, τ, α), we consider 100 randomly generated assortments ($\mathcal{N}_{1..100}$) with $|N_k| = 10, \forall k$.

Following is a counter-example to Conjecture 4.1.1.:

Example 4.2. $v_i = 0.685734$ *for* $i = 1,..,10$, $K = 0.05$, $\theta = 0.25$, $\tau = 0.1$, $\alpha = 0.1$. $S_1^C = \{1, 2, 3, 4\}, S_2^C = \{5, 6, 7\}$ *is the optimal solution of the centralized problem.* $S_1^D = \{1, 2, 3, 4\}, S_2^D = \{5, 6, 7, 8\}$ *is the equilibrium of the decentralized system.*

In Example 4.2, it is worth considering why it is profitable for Store 2 to add product 8 in the decentralized setting although it is not profitable in the centralized setting. Adding product 8 to store 2's assortment increases store 2's profits obtained from own-sales and cross-sales in store 1's market. It also increases store 1's profits obtained from transshippments from store 2. On the other hand, it decreases the profit store 1 obtains from transshipping to store 2, which alone makes adding product 8 unprofitable in the centralized setting whereas it is not a concern for store 2 in the decentralized setting.

There are a few observations on Table A7.2 that are worth mentioning here. First, only a few instances violate Conjecture 4.1.1, which requires further understanding and analysis, and Conjecture 4.1.2 is supported in all instances. Second, the parameters (K_L, τ_L, θ_H) incentivize adding exclusive products and perform transshipments in the centralized setting. On the other hand, a high α (which has no impact on the centralized setting) makes transshipments less attractive in the decentralized setting because it shrinks the margin obtained from products transshipped from the other store. Therefore, we expect $\mathcal{S}_\vee^C \supseteq \mathcal{S}_\vee^D$ and $\mathcal{S}_\wedge^C \subseteq \mathcal{S}_\wedge^D$ in this setting, which is verified in Table A7.2. Third, it is quite common that the centralized and the decentralized assortments mismatch, that is why achieving coordination is important.

We would like to shed light into how the coordination in the supply chain can be achieved. To that end, we consider a special case of the problem with equal popularities in Section 4.6.

4.6 Special Case 1: Equal Popularities

In this section, we assume that each product has the same popularity for tractability, i.e., $v_j = v$, $\forall j \in \mathcal{N}$ (see Kurtuluş et al. 2014 for a similar assumption).

4.6.1 The Centralized Problem

In the general version of the centralized problem, it might be optimal to include some common products in the assortments of different stores ($\mathcal{S}_\wedge^* \neq \varnothing$) because (1) including a popular product in an assortment can be more profitable than an exclusive but less popular product even if it can be transshipped from another store, and (2) the product category $\mathcal{N}$ has a small number of product variants and the fixed cost K of including a product in an assortment is small. In the special case of equal popularities, the former is no longer the case, but the latter can still be valid, which is formally stated as follows:

Lemma 4.4. *In an optimal solution of the CP, replacing a common product with an exclusive product is always profitable. This implies that $\mathcal{S}_\wedge^* \neq \varnothing$ only if $\mathcal{S}_\vee^* = \mathcal{N}$. These arguments hold in an equilibrium, too.*

In Figure 4.4, we demonstrate the optimal solution of the CP for two stores and three product variants in different parameter regions. The horizontal dashed line is $\tau = 0.33$, above which a store with two products does not receive transshipments according to Lemma 4.1. This implies the following in region B: In each store, a customer has access to only two products above the dashed line and three products below the dashed line. A similar argument holds for region C, too. This is an important distinction in order to interpret the product variety offered to an average customer in the system.

4.6.1.1 The Impact of K

A quick observation reveals that the total number of products and the number of products in each store are weakly decreasing in the fixed cost K. Similarly, both the union set $\mathcal{S}_\vee$ and the common set $\mathcal{S}_\wedge$ are shrinking as K increases. The product variety offered to an average customer in the system is non-monotone in K. For instance, suppose we start from an initial point in region A above the dashed line and move to the right. The number of products to which an average customer has an access is 3, 2, 2.5, and 2 in the regions A,B,C,D, respectively.

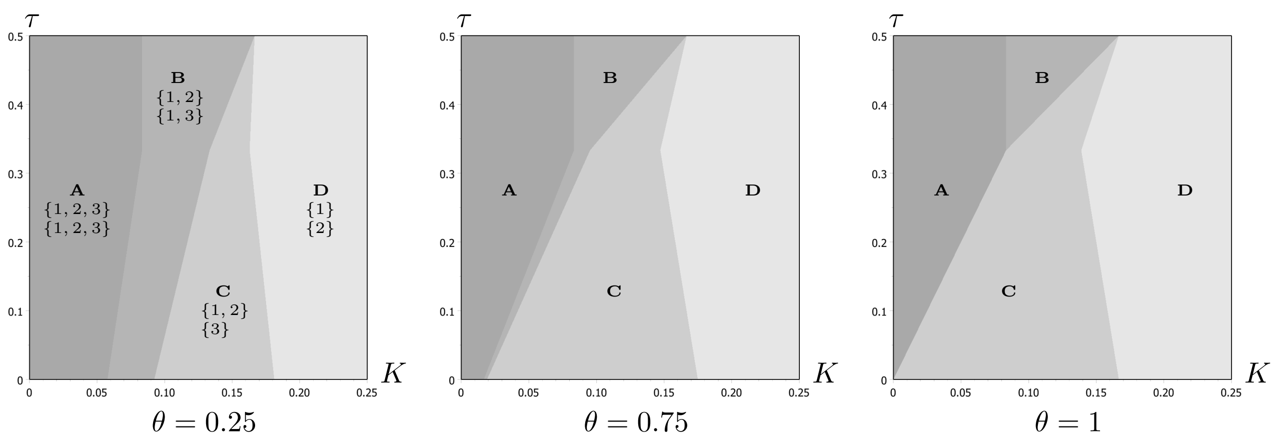


Figure 4.4: The centralized solution for $\mathcal{N} = \{1, 2, 3\}$ and $v_i = 1$, $\forall i \in \mathcal{N}$.

4.6.1.2 The Impact of τ

The transshipment cost τ has more intricate impacts on the assortments which depend on the fixed cost K. For small values of K the total number of products and the number of products in each store are non-decreasing in the transshipment cost τ. Consider the transitions from regions

$B \to A$ and $C \to B$ in upward direction. In these cases, adding a product variant in an assortment is cheap and it decreases the need for transshipments which becomes more costly as τ increases.

On the other hand, when K is sufficiently large, the impact of τ on the number of products is non-monotone. Consider the transitions $C \to D \to C \to B$ in upward direction along the vertical line in Figure 4.4 for $\theta = 1$. In order to clarify this situation, we demonstrate the probabilities of an average customer's purchasing with/without transshipment and not purchasing in Figure 4.5. It is intuitive and easy to show that the $\mathcal{P}\{\text{transshipment}\}$ is weakly decreasing in τ. The decrease in $\mathcal{P}\{\text{transshipment}\}$ can be achieved by either dropping a product ($C \to D$) or adding a product ($D \to C$ and $C \to B$).

4.6.1.3 The Impact of θ

As $\theta \in [0, 1]$ increases, the customers' disutility of waiting gets smaller, i.e., customers gets more patient for waiting. This suggests that the likelihood of transshipment in the system, $\mathcal{P}\{\text{transshipment}\}$, is weakly increasing in θ as shown in Figure 4.5. The total number of products and the number of products in each store are weakly decreasing in θ. Similarly, both the union set $\mathcal{S}_\vee$ and the common set $\mathcal{S}_\wedge$ are shrinking as θ increases.

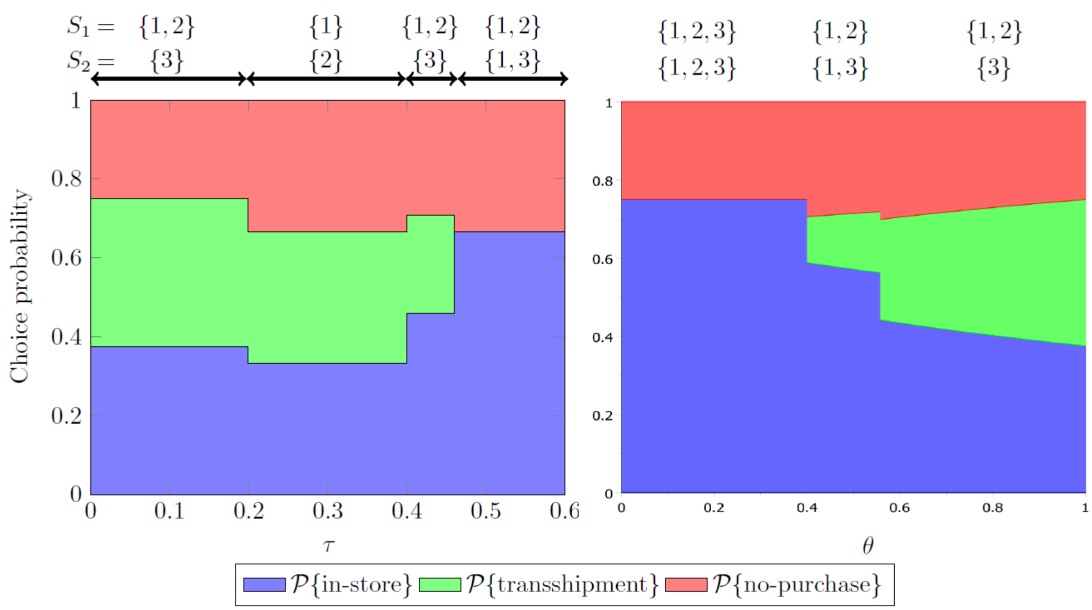


Figure 4.5: The probabilities of an average customer's purchasing with/without transshipment and not purchasing.

4.6.2 The Decentralized Problem

In Example 4.1, we demonstrated that the existence of a pure strategy Nash equilibrium is not guaranteed in general. In the special case of equal popularities, our simulations reveal that there exists a pure strategy Nash equilibrium for any $\alpha \in [0, 1]$, which is formally stated below:

Conjecture 4.2 (Existence)**.** *For any $\alpha \in [0, 1]$, there exists a pure strategy Nash equilibrium.*

Note that it is possible to have multiple pure strategy Nash equilibria even in the case of equal popularities as shown in Example 4.3.

Example 4.3. $v_i = 0.0785 \ \forall i \in \mathcal{N}, K = 0.05, \tau = 0.25, \alpha = 0.01, \theta = 0.75.$

	Product Variant				
	1	2	3	4	5
S_m^C	✓				
$S_{m'}^C$		✓	✓	✓	

Figure 4.6: Equilibrium 1.

	Product Variant				
	1	2	3	4	5
S_m^C	✓	✓			
$S_{m'}^C$			✓	✓	

Figure 4.7: Equilibrium 2.

Note that the above result is analytically unproven yet. It is not an easy task to prove the existence of a pure strategy Nash equilibrium in our setting even in the case of equal popularities. In the literature, there are some sufficient conditions for the existence of pure strategy Nash equilibrium in finite games (finite action sets). For instance, supermodular games (Milgrom and Shannon, 1994), potential games (Monderer and Shapley, 1996), and games with integrally concave payoff funtions (Iimura and Watanabe, 2014) are known to have pure strategy Nash equilibrium. Our payoff function (4.9) is neither supermodular nor integrally concave. Furthermore, we could not construct a potential function in our setting.

An alternative approach is to relax the discreteness of the action set, and consider the continuous version of the decentralized problem as there are some other sufficient conditions for the existence of pure strategy Nash equilibrium in infinite games. For instance, if each payoff function is quasi-concave (Debreu, 1952; Glicksberg, 1952; Fan, 1952), or supermodular (Topkis, 1979; Vives, 1990), or satisfying Nishimura and Friedman's condition (1981), or Ziad's condition (2003), there is a Nash equilibrium in pure strategies. However, our payoff function (4.9) fails to satisfy these conditions, too.

As mentioned in the literature review, coordination of supply chains received significant attention in operations management. In our discussions with practitioners in the auto parts industry, they pointed out the mismatch between the assortments suggested by the headquarters for different outlets/stores of the retailer and the assortments individual store managers determine. This leads to the classic coordination issue, and the following result suggests how the coordination can be achieved:

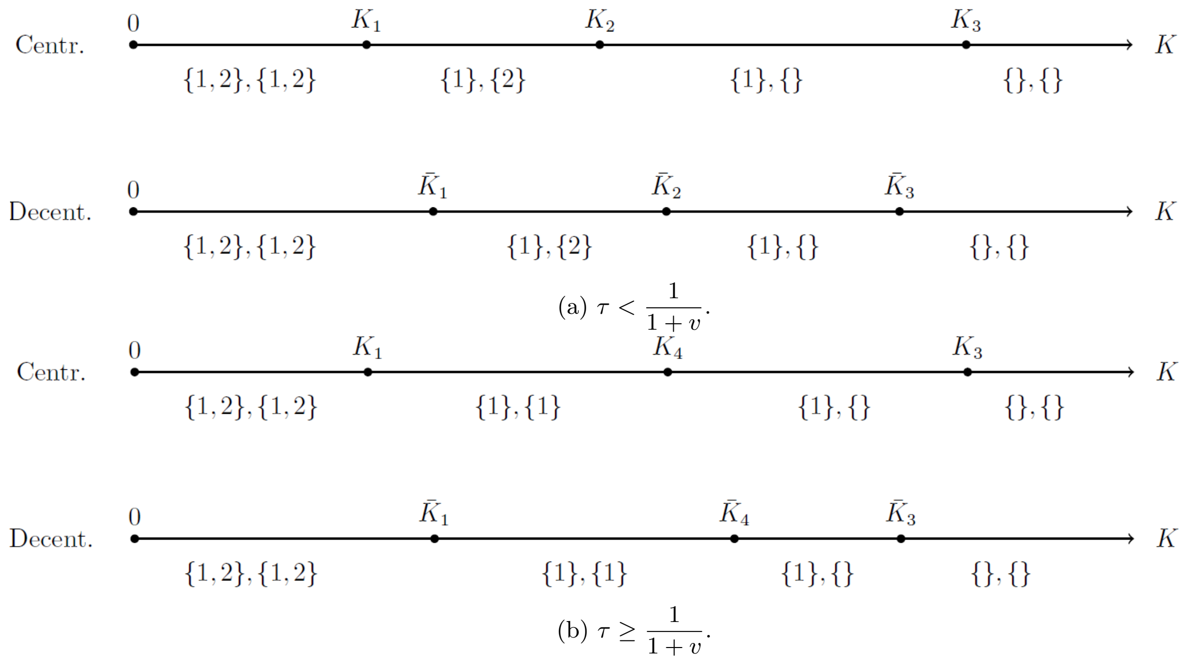
Proposition 4.5 (Coordination)**.** *Assume that there is sufficiently many products in $\mathcal{N}$ such that the stores can always find an exclusive product to include in their assortments. Then, there exists an* $\alpha \in \left[0, \min\left\{\frac{1}{1+|S_m^C|v}, \frac{1}{1+|S_{m'}^C|v}\right\}\right]$ *for which* $\{S_m^C, S_{m'}^C\}$ *is a pure strategy Nash equilibrium.*

We prove this result under the assumption that $|\mathcal{N}|$ is large enough. This implies that there is no incentive for either of the stores to duplicate a product in the assortment of the other store. This simplifies the proof because we can limit our consideration to adding the next exclusive product only (ignoring the possibility of duplicating a product) in the best response functions. We note here that this is not a necessary assumption, but it is sufficient to prove the coordination result. In fact, our numerical simulations reveal that there exists an α that achieves the coordination even when $|\mathcal{N}|$ is small, and stores have common products in their optimal assortments for the centralized problem.

4.6.3 Two Products: $(v_1 = v_2 = v)$

In this section, we consider a further simplified version of the problem with two products $(v_1 = v_2 = v)$. This setting allows us to fully characterize the solution of the CP and the equilibrium of the decentralized system for any parameter set, so we can conclude how centralized/decentralized settings differ in terms of the assortments and product varieties offered to customers. This setting also allows us to verify and observe how coordination of the system can be achieved by setting the profit sharing parameter α appropriately as suggested by Proposition 4.5.

It turns out that the cost of transshipment τ plays a pivotal role on how the assortments are determined in the centralized and the decentralized systems. The solution of the centralized problem and the equilibrium as a function of K when $\tau < \frac{1}{1+v}$ and $\tau \geq \frac{1}{1+v}$ are shown in Figure 4.8a and Figure 4.8b, respectively.



(a) $\tau < \dfrac{1}{1+v}$.

(b) $\tau \geq \dfrac{1}{1+v}$.

Figure 4.8: The solution of the centralized problem and the equilibrium as a function of K and τ.

Potential assortment mismatches between the centralized and the decentralized systems occur in the following regions of K:

1. $\bar{K}_1 - K_1 = \dfrac{\theta\alpha}{(1+v+\theta v)/v}$

2. $\bar{K}_2 - K_2 - = \dfrac{\alpha(1+v)(1+v+\theta v) + [\tau(1+v)-1](1+\theta v)}{(1+v)(1+\theta v)(1+v+\theta v)/v} = \dfrac{\alpha v}{1+\theta v} - \dfrac{\overbrace{[1-\tau(1+v)]}^{>0}v}{(1+v)(1+v+\theta v)}$

3. $K_3 - \bar{K}_3 = \dfrac{\theta(1-\alpha-\tau)}{(1+\theta v)/v}$

4. $\bar{K}_4 - K_4 = \dfrac{\theta\alpha}{(1+\theta v)/v}$

4.7 Special Case 2: Enforced Transshipments

Although it has a well defined structure, the CP is still difficult to solve optimally due to the difficulty of allocating exclusive products. Even for the case of two stores, allocating the exclusive products is a combinatorial problem. The following result sheds light into how the exlusive products should be allocated across stores in the centralized problem:

Proposition 4.6. *The products are allocated across stores evenly in the following sense: For any pair of stores $m, m' \in \mathcal{M}$, the products in $(S_m^C \setminus S_{m'}^C) \cup (S_{m'}^C \setminus S_m^C)$ are allocated such that $\Big| \sum_{j \in S_m^C} v_j - \sum_{j \in S_{m'}^C} v_j \Big|$ is minimized.*

The above result implies that the central solution seeks to find a *fair* allocation of exclusive products across stores. This minimizes the overall transshipment costs while maximizing the stores' total profit. Note that the customers' disutility of waiting for the transshipment is the driving force for this result. When the disutility is negligible ($c = 0$), any allocation of the exclusive products across stores yield the same total profits. This might be the case for certain industries where the customers incur a certain waiting time in either case (transshipped or not). For instance, in the auto parts industry, certain components are transshipped to customers (auto repair shops) only on the next day. In that case, which store transships the component does not make a difference for the customer as long as it is delivered on time.

We note that Proposition 4.6 do not hold when the transshipments are optional as in the rest of our analysis except Section 4.7. Example 4.4 demonstrates how it can be violated if transshipments are not enforced:

Example 4.4. $v_i = 3.417$ *for* $i = 1, .., 5$, $K = 0.025$, $\theta = 0.75$, $\tau = 0.1$.

	Product Variant				
	1	2	3	4	5
S_1^C	✓				
S_2^C		✓			
S_3^C			✓	✓	✓

Figure 4.9: Transshipments are optional.

	Product Variant				
	1	2	3	4	5
S_1^C	✓	✓			
S_2^C	✓	✓			
S_3^C	✓	✓	✓		

Figure 4.10: Transshipments are enforced.

When transshipments are optional, Lemma 4.1 suggests that Store 1 and 2 should receive transshipments from Store 3, but Store 3 should not receive transshipments from the other stores. When transshipments are enforced, the only way to prevent the profit cannibalization of transshipped product (due to the fact that the sender party keeps α portion of the profit) is to duplicate other stores' products, which implies more common products as is the case in Figure 4.10.

Another interesting observation about the enforcement transshipment case is that the equilibrium is not guaranteed to exist even in the case of equal popularities, i.e., Conjecture 4.2 is falsified when transshipments are enforced as demonstrated in Example 4.5:

Example 4.5 (No-Equilibrium). $v_i = 0.69374 \forall i \in \mathcal{N}, K = 0.05, \tau = 0.25, \alpha = 0.1, \theta = 0.75$.

	Product Variant									
	1	2	3	4	5	6	7	8	9	10
S_m^C	✓	✓	✓							
$S_{m'}^C$				✓	✓	✓				

Figure 4.11: The opt. solution of the CP.

	Product Variant									
	1	2	3	4	5	6	7	8	9	10
$S_m^*(\emptyset)$	✓	✓	✓	✓						
$S_{m'}^*(\{1,2,3,4\})$					✓	✓	✓			
$S_m^*(\{5,6,7\})$	✓	✓			✓	✓	✓			
$S_{m'}^*(\{1,2,5,6,7\})$			✓	✓				✓		
$S_{m'}^*(\{3,4,8\})$					✓	✓	✓			

Figure 4.12: The series of best responses.

4.8 Concluding Remarks

We focused on assortment planning in a multi-location setting with transshipments, which to the best of our knowledge is not studied before. Transshipment is a means of demand fulfillment, which is often studied within the context of inventory management. We excluded inventory dimension from our analysis, but included customer choice among a set of alternatives in our analysis of transshipments ,which conforms well with the previous two chapters of this dissertation.

There are several questions that require further analysis. For instance, it is important to quantify the value of lateral transshipments, which by construction is positive in our setting. Similarly, quantifying the profit loss due to allowing stores to modify their assortments independently is also important. The comparison of assortment variety in the centralized and decentralized problems is yet to be further explored. Finally, note that we considered achieving coordination by setting the profit sharing rule (α) between the sender and the receiver stores, but another way to achive it might be to incentivize stores for transshipments by compensating the transshipment cost τ.

APPENDIX 1: PROOFS OF ANALYTICAL RESULTS IN CHAPTER 2

Proof of Proposition 2.1. Suppose we have two products, i and j, such that $Y_i \geq Y_j$ and $m_i \geq m_j$, or equivalently, product i dominates product j. Our aim is to show that product j is never optimal to promote.

Let $K = \underset{i \in S}{\operatorname{argmax}}\{m_i\}$. Note that K is not necessarily a singleton. Hence, let $k = \underset{i \in K}{\operatorname{argmax}}\{Y_i\}$. In words, k is the most popular product among the products having the highest margin. The product k is well defined because by assumption no pair of products have the same popularity and the profit margin. We can assume w.l.o.g. that products i, j, and k are distinct. By definition of product k, the dominated product (j) is different from product k. If products i and k are the same, the following arguments still hold.

We argue that $\max\{\pi^i, \pi^k\} \geq \pi^j$. This implies that there is always a more profitable option (either promoting product i or product k) than promoting product j. Let $M = P_A^0 \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)$. We can write π^x as a convex combination of m_x and M for all $x \in S$ as follows:

$$\pi^x = P_D^x \cdot m_x + (1 - P_D^x) \cdot M \tag{A1.1}$$

We consider two cases separately: $(m_i > M)$ and $(m_i \leq M)$. If $(m_i > M)$, we have

$$\pi^j = P_D^j \cdot m_j + (1 - P_D^j) \cdot M \leq P_D^j \cdot m_i + (1 - P_D^j) \cdot M \leq P_D^i \cdot m_i + (1 - P_D^i) \cdot M = \pi^i \tag{A1.2}$$

The first inequality follows because $m_i \geq m_j$ and the second follows because $P_D^i \geq P_D^j$ and $m_i > M$. Note also that at least one of the above inequalities is strict because we have assumed no two products i and j have $m_i = m_j$ and $Y_i = Y_j$. Therefore we either have $m_i > m_j$ or $P_D^i > P_D^j$, or both. Hence, $\pi^i > \pi^j$ if $m_i > M$. Next, consider the case where $(m_i \leq M)$. By Assumption 2.1, we have $m_k > M$. Since $m_j \leq m_i \leq M < m_k$, it follows from (A1.1) that $\pi^k > \pi^j$.

Proof of Proposition 2.2. We can write π^x as follows:

$$\pi^x = (1 - P_A^x) \cdot \theta_x + P_A^x \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big) \tag{A1.3}$$

Consider two products, i and j, such that $Y_i < Y_j$ and $\theta_i \geq \theta_j$. We have

$$\pi^j = (1 - P_A^j) \cdot \theta_j + P_A^j \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big) < (1 - P_A^i) \cdot \theta_j + P_A^i \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big) \tag{A1.4}$$

$$\leq (1 - P_A^i) \cdot \theta_i + P_A^i \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big) = \pi^i \tag{A1.5}$$

The first inequality follows by Assumption 2.2 and the fact that $P_A^i > P_A^j$. The second inequality follows since $\theta_i \geq \theta_j$.

Proof of Proposition 2.3. Consider two products, $j \in E'$ and $k \in E'$, such that $Y_j > Y_k$.

Part 1: We consider $\pi^j(r)$ and $\pi^k(r)$ as functions of r. Notice that they are linearly increasing in r. Since $P_A^j < P_A^k$, $\pi^k(r)$ increases at a faster rate than $\pi^j(r)$. This implies that $\pi^j(r)$ and $\pi^k(r)$ intersect at a unique value of r above which $\pi^k(r) > \pi^j(r)$. Hence, the optimal choice of product to promote, $i^*(r)$, always switches from a more popular product to a less popular one within the set E' (or stays the same) as r increases. Since the products in E' are indexed in decreasing order of their popularities, $i^*(r)$ is non-decreasing in r.

Part 2: We consider $\pi^j(c)$ and $\pi^k(c)$ as functions of c. We first solve the equation $\pi^j(c) = \pi^k(c)$ for c and argue that the solution is unique, implying that $\pi^j(c)$ and $\pi^k(c)$ intersect at a unique value of c. Some manipulation of the equation $\pi^j(c) = \pi^k(c)$ yields the following:

$$e^{-c+\lambda I} = \frac{e^{Y_0}(e^{Y_j} \cdot m_j - e^{Y_k} \cdot m_k) + e^{Y_j + Y_k} \cdot (m_j - m_k)}{e^{Y_k} \cdot m_k - e^{Y_j} \cdot m_j + (e^{Y_j} - e^{Y_k})\big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)} \tag{A1.6}$$

By continuity and monotonicity of the LHS in c, there exists a unique value of c that satisfies $\pi^i(c) = \pi^j(c)$. Hence, $\pi^j(c)$ and $\pi^k(c)$ intersect only once. Next, we show that $\pi^j(c) > \pi^k(c)$ for sufficiently large c. Since j and k are picked arbitrarily, this is sufficient to show that the the optimal choice of product to promote, $i^*(c)$, always switches from a less popular product to a more popular one within the set E' (or stays the same) as c increases.

In order to show that $\pi^j(c) > \pi^k(c)$ for sufficiently large c, we write π^x as a convex combination of θ_x and $\big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)$ for $x \in S$ as in A1.3. It is straightforward to show that $\lim_{c \to \infty} P_A^x = 0$ for all $x \in S$. Then, in the limit, $\pi^j(c) > \pi^k(c)$, because $\theta_j > \theta_k$ by Proposition 2.2.

Proof of Proposition 2.4. Assume without loss of generality that $\operatorname{argmax}_{i \in S}\{\pi^i(S)\} = j$ and

$\underset{i \in S}{\text{argmax}}\{\pi^i(S^+)\} = k$ where $j \neq k$. (Otherwise, the optimal product is the same for both assortments, and the proposition statement is satisfied with equality.) By Proposition 2.2, both of these products are in the set E' corresponding to assortment S. Our aim is to show that $\theta_j > \theta_k$, which implies that $j < k$ due to the order of products in the set E'.

Fix $M = \bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda) = \bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda)$. Define the following functions of x:

$$\tilde{\pi}^j(x) = P_D^j(x) \cdot m_j + P_A^j(x) \cdot (M + r), \tag{A1.7}$$

$$\tilde{\pi}^k(x) = P_D^k(x) \cdot m_k + P_A^k(x) \cdot (M + r), \tag{A1.8}$$

where $P_D^j(x)$, $P_A^j(x)$, $P_D^k(x)$, $P_A^k(x)$ are the choice probabilities obtained by replacing $\lambda \cdot I(\mathbf{Y}, \lambda)$ with x. After solving $\tilde{\pi}^j(x) - \tilde{\pi}^k(x) = 0$, we get

$$e^{-c+x} = \frac{e^{Y_0}(e^{Y_j} \cdot m_j - e^{Y_k} \cdot m_k) + e^{Y_j + Y_k} \cdot (m_j - m_k)}{e^{Y_k} \cdot m_k - e^{Y_j} \cdot m_j + (e^{Y_j} - e^{Y_k}) \cdot (M + r)} \tag{A1.9}$$

Since the LHS of the above equation is continuous and monotonically increasing in x, $\tilde{\pi}^i(x)$ and $\tilde{\pi}^j(x)$ intersect at a unique value of x, say x'. Furthermore, $x' \in [\lambda \cdot I(\mathbf{Y}_S, \lambda), \lambda \cdot I(\mathbf{Y}_{S^+}, \lambda)]$ because $\tilde{\pi}^j\big(\lambda \cdot I(\mathbf{Y}_S, \lambda)\big) = \pi^j(S) > \pi^k(S) = \tilde{\pi}^k\big(\lambda \cdot I(\mathbf{Y}_S, \lambda)\big)$, and $\tilde{\pi}^k\big(\lambda \cdot I(\mathbf{Y}_{S^+}, \lambda)\big) = \pi^k(S^+) > \pi^j(S^+) = \tilde{\pi}^j\big(\lambda \cdot I(\mathbf{Y}_{S^+}, \lambda)\big)$ by definitions of j and k. This implies that we have $\tilde{\pi}^j(x) > \tilde{\pi}^k(x)$ for all $x < x'$, and $\tilde{\pi}^j(x) < \tilde{\pi}^k(x)$ for all $x > x'$. As $x \to -\infty$, we know that $\tilde{\pi}^j(x) > \tilde{\pi}^k(x)$ iff $\theta_j > \theta_k$ because $\lim_{x \to -\infty} \tilde{\pi}^i(x) = \theta_i$ for all $i \in S$. Hence, we conclude that $\theta_j > \theta_k$.

Proof of Lemma 2.1. It suffices to prove that $\lambda \cdot I(\mathbf{Y}, \lambda)$ is increasing in λ in the interval $(0, 1]$. To that end, we calculate the first derivative of $\lambda \cdot I(\mathbf{Y}, \lambda)$ with respect to λ as follows:

$$\begin{aligned}
\frac{d\left(\lambda \cdot I(\mathbf{Y}, \lambda)\right)}{d\lambda} &= \ln\left(\sum_{k \in S \cup \{0\}} \mathrm{e}^{Y_k/\lambda}\right) - \frac{\sum_{q \in S \cup \{0\}} \dfrac{Y_q}{\lambda} \cdot \mathrm{e}^{Y_q/\lambda}}{\sum_{q \in S \cup \{0\}} \mathrm{e}^{Y_q/\lambda}} \\
&= \frac{\sum_{q \in S \cup \{0\}} \mathrm{e}^{Y_q/\lambda} \left(\ln\left(\sum_{k \in S \cup \{0\}} \mathrm{e}^{Y_k/\lambda}\right) - \dfrac{Y_q}{\lambda}\right)}{\sum_{q \in S \cup \{0\}} \mathrm{e}^{Y_q/\lambda}}
\end{aligned}$$

The denominator is clearly positive. Since $\left(\ln\left(\sum_{k\in S\cup\{0\}} e^{Y_k/\lambda}\right) - \frac{Y_q}{\lambda}\right) > 0$ for all $q \in S \cup \{0\}$ and for all $\lambda \in (0,1]$, the numerator in the above expression is also positive. Hence, we have $\frac{d\,(\lambda\cdot I(\mathbf{Y},\lambda))}{d\lambda} > 0$, implying that $\lambda \cdot I(\mathbf{Y},\lambda)$ is strictly increasing in $\lambda \in (0,1]$.

Proof of Proposition 2.5. Assume without loss of generality that there exists $\lambda' \in (0,1]$ at which the optimal choice of product switches from one product to another. (If λ' does not exist, then the optimal product is the same for all $\lambda \in (0,1]$, and thus we are done.) Suppose these two products are indexed by j and k, or more formally, $\underset{i\in S}{\operatorname{argmax}}\{\pi^i(\lambda'-\epsilon)\} = j$ and $\underset{i\in S}{\operatorname{argmax}}\{\pi^i(\lambda'+\epsilon)\} = k$ for sufficiently small $\epsilon > 0$. By Proposition 2.2, both of these products are in the set E'. Our aim is to show that $\theta_j > \theta_k$, which implies that $j < k$ due to the order of products in the set E'.

Fix $M = \bar{m}(\mathbf{Y},\mathbf{m},\lambda')$. Define $\tilde{\pi}^j(x)$ and $\tilde{\pi}^k(x)$ as in A1.7 and A1.8. Following similar arguments as in A1.9, we can show that $\tilde{\pi}^j(x)$ and $\tilde{\pi}^k(x)$ intersect at a unique value of x. This intersection occurs at $x = \lambda' \cdot I(\mathbf{Y},\lambda')$ because

$$\tilde{\pi}^j\big(\lambda' \cdot I(\mathbf{Y},\lambda')\big) = \pi^j(\lambda') = \pi^k(\lambda') = \tilde{\pi}^k\big(\lambda' \cdot I(\mathbf{Y},\lambda')\big). \tag{A1.10}$$

We have $\lim_{x\to-\infty} \tilde{\pi}^i(x) = \theta_i$ for all $i \in S$. Hence, we have either one of the following cases:

1. $\theta_j > \theta_k$ and therefore $\tilde{\pi}^j(x) > \tilde{\pi}^k(x)$ $\forall x < \lambda' \cdot I(\mathbf{Y},\lambda')$ and $\tilde{\pi}^j(x) < \tilde{\pi}^k(x)$ $\forall x > \lambda' \cdot I(\mathbf{Y},\lambda')$.
2. $\theta_j < \theta_k$ and therefore $\tilde{\pi}^j(x) < \tilde{\pi}^k(x)$ $\forall x < \lambda' \cdot I(\mathbf{Y},\lambda')$ and $\tilde{\pi}^j(x) > \tilde{\pi}^k(x)$ $\forall x > \lambda' \cdot I(\mathbf{Y},\lambda')$.

We claim that case 2 cannot be true, and we show it by contradiction. Suppose that $\theta_j < \theta_k$ and therefore $\tilde{\pi}^j(x) > \tilde{\pi}^k(x)$ for all $x > \lambda' \cdot I(\mathbf{Y},\lambda')$. Then, consider the following equalities:

$$\pi^j(\lambda) = \tilde{\pi}^j\big(\lambda \cdot I(\mathbf{Y},\lambda)\big) + P_A^j\big(\lambda \cdot I(\mathbf{Y},\lambda) \cdot \big(\bar{m}(\mathbf{Y},\mathbf{m},\lambda) - M\big) \tag{A1.11}$$

$$\pi^k(\lambda) = \tilde{\pi}^k\big(\lambda \cdot I(\mathbf{Y},\lambda) + P_A^k\big(\lambda \cdot I(\mathbf{Y},\lambda) \cdot \big(\bar{m}(\mathbf{Y},\mathbf{m},\lambda) - M\big) \tag{A1.12}$$

Since $\bar{m}(\mathbf{Y},\mathbf{m},\lambda)$ is non-decreasing in λ by the proposition assumption, $\big(\bar{m}(\mathbf{Y},\mathbf{m},\lambda) - M\big)$ is non-negative for any $\lambda > \lambda'$ in the above equations. Moreover, $\theta_j < \theta_k$ implies that $Y_j < Y_k$ by definition of E', and therefore $P_A^j\big(\lambda \cdot I(\mathbf{Y},\lambda)\big) > P_A^k\big(\lambda \cdot I(\mathbf{Y},\lambda)\big)$ for any $\lambda \in (0,1]$. Finally, case 2 implies that $\tilde{\pi}^j\big((\lambda'+\epsilon) \cdot I(\mathbf{Y},(\lambda'+\epsilon))\big) > \tilde{\pi}^k\big((\lambda'+\epsilon) \cdot I(\mathbf{Y},(\lambda'+\epsilon))\big)$. As a result, plugging $\lambda = (\lambda'+\epsilon)$

in (A1.11) and (A1.12) reveals that $\pi^j(\lambda' + \epsilon) > \pi^k(\lambda' + \epsilon)$, which contradicts with the definition of j and k. Hence, case 2 cannot be true, implying that $\theta_j > \theta_k$.

Proof of Proposition 2.6. We prove in two steps. First, we show that $\Delta^j(S) - \Delta^k(S^+)$ is decreasing linearly in $\beta \geq 0$, which implies that $\Delta^j(S) < \Delta^k(S^+)$ for some large enough $\beta \geq 0$. Second, we show that $\Delta^j(S) > \Delta^k(S^+)$ when $\beta = 0$. These together imply the existence of the threshold $\beta' \geq 0$ and completes the proof.

1. $\Delta^j(S) - \Delta^k(S^+)$ is decreasing linearly in $\beta \geq 0$.

   We express $\Delta^j(S) - \Delta^k(S^+)$ as follows:

$$
\begin{aligned}
&\Delta^j(S) - \Delta^k(S^+) \\
=&\Pi^j(S) - \Pi^0(S) - \big(\Pi^k(S^+) - \Pi^0(S^+)\big) \\
=&(1+\beta)\big(\pi^j(S) - \pi^k(S^+)\big) - \big(P_A^0(S) \cdot \bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda) - P_A^0(S^+) \cdot \bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda)\big).
\end{aligned}
$$

   We must show that $\pi^j(S) < \pi^k(S^+)$. To that end, consider the following inequalities:

$$
\begin{aligned}
\pi^k(S^+) \geq \pi^j(S^+) &= \big(1 - P_A^j(S^+)\big) \cdot \theta_j + P_A^j(S^+) \cdot \big(\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) + r\big) && \text{(A1.13)} \\
&> \big(1 - P_A^j(S)\big) \cdot \theta_j + P_A^j(S) \cdot \big(\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) + r\big) && \text{(A1.14)} \\
&\geq \big(1 - P_A^j(S)\big) \cdot \theta_j + P_A^j(S) \cdot \big(\bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda) + r\big) = \pi^j(S) && \text{(A1.15)}
\end{aligned}
$$

   where (A1.14) follows from Assumption 2.2 and the fact that $P_A^i(S^+) > P_A^i(S)$ for all $i \in S$, and (A1.15) follows because $\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) \geq \bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda)$.

2. $\Delta^j(S) > \Delta^k(S^+)$ when $\beta = 0$.

   Let $\beta = 0$. By Assumption 2.1, we know that $\Delta^j(S) > 0$, i.e., promoting the optimal product in S is more profitable than no-promotion. If $\Delta^k(S^+) < 0$, then $\Delta^j(S) > 0 > \Delta^k(S^+)$ and we are done. Hence, we consider the case when $\Delta^k(S^+) > 0$.

   Since $\Delta^j(S) \geq \Delta^k(S)$ by the optimality of product j when the assortment is S, showing $\Delta^k(S) > \Delta^k(S^+)$ will be sufficient to prove that $\Delta^j(S) > \Delta^k(S^+)$. To that end, we write

$\Delta^k(S)$ as

$$\begin{aligned}\Delta^k(S) &= \frac{\mathrm{e}^{Y_k}}{\mathrm{e}^{Y_0} + \mathrm{e}^{Y_k} + \mathrm{e}^{-c+\lambda\cdot I(S)}} \cdot m_k \\ &- \left(\frac{\mathrm{e}^{-c+\lambda\cdot I(S)}}{\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)}} - \frac{\mathrm{e}^{-c+\lambda\cdot I(S)}}{\mathrm{e}^{Y_0} + \mathrm{e}^{Y_k} + \mathrm{e}^{-c+\lambda\cdot I(S)}}\right) \cdot (\bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda) + r) \\ &= \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_k} \cdot m_k + \mathrm{e}^{Y_k} \cdot (m_k - \bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda) - r) \cdot \mathrm{e}^{-c+\lambda I(S)}}{(\mathrm{e}^{Y_0} + \mathrm{e}^{Y_k} + \mathrm{e}^{-c+\lambda\cdot I(S)}) \cdot (\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)})}\end{aligned}$$

Fix $M = \bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda)$. Define the function $\bar{\Delta}(S)$ as follows:

$$\bar{\Delta}(S) := \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_k} \cdot m_k + \mathrm{e}^{Y_k} \cdot (m_k - M - r) \cdot \mathrm{e}^{-c+\lambda\cdot I(S)}}{(\mathrm{e}^{Y_0} + \mathrm{e}^{Y_k} + \mathrm{e}^{-c+\lambda\cdot I(S)}) \cdot (\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)})} \tag{A1.16}$$

We will show that $\bar{\Delta}(S)$ decreases as the assortment S expands. To that end, consider the function $f : [I(\mathbf{Y}_S), I(\mathbf{Y}_{S^+})] \to [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S^+)}]$ such that $x = f(I) = \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(\cdot)}$. Notice that x increases monotonically as $I(\cdot)$ increases, or equivalently as the assortment expands. Hence, $f(\cdot)$ is a one-to-one correspondence between the sets $[I(\mathbf{Y}_S), I(\mathbf{Y}_{S^+})]$ and $[\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S^+)}]$. By plugging x in the RHS of (A1.16) we get:

$$\frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_k} \cdot m_k + \mathrm{e}^{Y_k} \cdot (m_k - M - r) \cdot (x - \mathrm{e}^{Y_0})}{(\mathrm{e}^{Y_k} + x) \cdot x} \tag{A1.17}$$

In order to show that $\bar{\Delta}(S)$ decreases as $I(\cdot)$ increases in the interval $[I(\mathbf{Y}_S), I(\mathbf{Y}_{S^+})]$, we will show that (A1.17) is decreasing in x in the interval $[\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S^+)}]$. To that end, we compute the first derivative of (A1.17) with respect to x as:

$$-\frac{\mathrm{e}^{Y_k} x \left[x(m_k - M - r) + 2\mathrm{e}^{Y_0}(M + r)\right] + \mathrm{e}^{2Y_k + Y_0}(M + r)}{(\mathrm{e}^{Y_k} + x)^2 \cdot x^2} \tag{A1.18}$$

The denominator is clearly positive. We will argue that the numerator is also positive for all $x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S^+)}]$. To that end, it suffices to show that

$$\left[x(m_k - M - r) + 2\mathrm{e}^{Y_0}(M + r)\right] > 0 \text{ for all } x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S)}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(S^+)}] \tag{A1.19}$$

Since we consider the case when $\Delta^k(S^+) > 0$ we have the following:

$$\begin{aligned}\Delta^k(S^+) =& \frac{e^{Y_k}}{e^{Y_0} + e^{Y_k} + e^{-c+\lambda \cdot I(S^+)}} \cdot m_k \\ &- \left(\frac{e^{-c+\lambda \cdot I(S^+)}}{e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}} - \frac{e^{-c+\lambda \cdot I(S^+)}}{e^{Y_0} + e^{Y_k} + e^{-c+\lambda \cdot I(S^+)}} \right) \cdot (\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) + r) > 0\end{aligned}$$

Since $\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) \geq M$, by replacing $\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda)$ by M in the above inequality, we get

$$\begin{aligned}& \frac{e^{Y_k}}{e^{Y_0} + e^{Y_k} + e^{-c+\lambda \cdot I(S^+)}} \cdot m_k \\ &- \left(\frac{e^{-c+\lambda \cdot I(S^+)}}{e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}} - \frac{e^{-c+\lambda \cdot I(S^+)}}{e^{Y_0} + e^{Y_k} + e^{-c+\lambda \cdot I(S^+)}} \right) \cdot (M + r) > 0\end{aligned}$$

By plugging $\bar{x} = e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}$ in the above inequality and some algebraic manipulation, we get $m_k > \dfrac{\left(\bar{x} - e^{Y_0}\right)}{\bar{x}} \cdot (M + r)$. This implies that $\bar{x} \cdot m_k - \left(\bar{x} - e^{Y_0}\right) \cdot (M + r) > 0$. Then, we can write

$$\bar{x} \cdot (m_k - M - r) + e^{Y_0} (M + r) > 0 \tag{A1.20}$$

Inequality (A1.20) implies that (A1.19) holds for all $x \leq \bar{x}$. To clarify, note that (A1.19) holds for all x if $(m_k - M - r) \geq 0$. Otherwise, it holds for all $x \leq \bar{x}$ because (A1.20) guarantees it for $x = \bar{x}$, and the LHS of (A1.19) is decreasing in x. As a result, we showed that the expression in (A1.17) is decreasing in $x \in [e^{Y_0} + e^{-c+\lambda \cdot I(S)}, e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}]$, which in turn implies that $\bar{\Delta}(\cdot)$ decreases as $I(\cdot)$ increases in the interval $[I(\mathbf{Y}_S), I(\mathbf{Y}_{S^+})]$ (i.e., as the assortment S expands).

Thus we have $\Delta^j(S) \geq \Delta^k(S) = \bar{\Delta}(S) > \bar{\Delta}(S^+) \geq \Delta^k(S^+)$ when $\beta = 0$.

Proof of Corollary 2.1 Let $\beta'(\cdot)$ denote the threshold β' in Proposition 2.6 as a function of the two assortments being compared. Choose $\beta_1 = \min\{\beta'(S, S^+), \beta'(S, S^{++})\}$ and $\beta_2 = \max\{\beta'(S, S^{++}), \beta'(S^+, S^{++})\}$. Part 1 and 3 follow by definitions of β_1 and β_2. For $\beta > \beta_1$, $\max\{\Delta^k(S^+), \Delta^l(S^{++})\} > \Delta^j(S)$. For $\beta < \beta_2$, $\max\{\Delta^j(S), \Delta^k(S^+)\} > \Delta^l(S^{++})$. Combining two, we have $\max\{\Delta^j(S), \Delta^k(S^+), \Delta^l(S^{++})\} = \Delta^k(S^+)$ for $\beta_1 < \beta < \beta_2$, so part 2 holds.

Proof of Proposition 2.7. Let $i^*(\underline{\lambda}) = j$ and $i^*(\bar{\lambda}) = k$ for ease of exposition. We will follow a similar procedure to the proof of Proposition 2.6 and prove in two steps.

1. $\Delta^j(\underline{\lambda}) - \Delta^k(\bar{\lambda})$ is decreasing linearly in $\beta \geq 0$. It follows the same arguments as the first part of the proof of Proposition 2.6 by replacing S and S^+ with $\underline{\lambda}$ and $\bar{\lambda}$, respectively.

2. $\Delta^j(\underline{\lambda}) > \Delta^k(\bar{\lambda})$ when $\beta = 0$.

   Assume w.l.o.g. that $\exists\ \lambda' \in [\underline{\lambda}, \bar{\lambda}]$ such that $\underset{i \in S}{\operatorname{argmax}}\{\pi^i(\lambda)\} = j$ for all $\lambda \in [\underline{\lambda}, \lambda']$, i.e., the optimal product to promote is j within the sub-interval $[\underline{\lambda}, \lambda']$, and it switches to another product at $\lambda = \lambda'$, denoted by $i^*(\lambda')$. We will show that $\Delta^j(\underline{\lambda}) > \Delta^j(\lambda') = \Delta^{i^*(\lambda')}(\lambda')$, where the equality holds because $\pi^j(\lambda') = \pi^{i^*(\lambda')}(\lambda')$. ($i^*(\lambda')$ and k are not necessarily the same product.) Since the same argument holds for any sub-interval in $[\underline{\lambda}, \bar{\lambda}]$ and because $\Delta^{(\cdot)}(\lambda)$ is continuous, we conclude that $\Delta^j(\underline{\lambda}) > \Delta^k(\bar{\lambda})$. Define the following functions:

$$\underline{\Delta}(\lambda) := \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_j} \cdot m_j + \mathrm{e}^{Y_j} \cdot (m_j - \bar{m}(\lambda') - r) \cdot \mathrm{e}^{-c+\lambda \cdot I(\lambda)}}{(\mathrm{e}^{Y_0} + \mathrm{e}^{Y_j} + \mathrm{e}^{-c+\lambda \cdot I(\lambda)}) \cdot (\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda \cdot I(\lambda)})} \tag{A1.21}$$

$$\bar{\Delta}(\lambda) := \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_j} \cdot m_j + \mathrm{e}^{Y_j} \cdot (m_j - \bar{m}(\underline{\lambda}) - r) \cdot \mathrm{e}^{-c+\lambda \cdot I(\lambda)}}{(\mathrm{e}^{Y_0} + \mathrm{e}^{Y_j} + \mathrm{e}^{-c+\lambda \cdot I(\lambda)}) \cdot (\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda \cdot I(\lambda)})} \tag{A1.22}$$

   Note that $\underline{\Delta}(\lambda) \leq \bar{\Delta}(\lambda)$ for any $\lambda \in [\underline{\lambda}, \bar{\lambda}]$, because $\bar{m}(\lambda)$ is non-decreasing in λ for $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ by the assumption given in the proposition. We will now show that $\bar{\Delta}(\lambda)$ is decreasing in λ for $\lambda \in [\underline{\lambda}, \lambda']$. To that end, consider the function $f : [\underline{\lambda}, \lambda'] \to [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda' \cdot I(\lambda')}]$ such that $x = f(\lambda) = \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda \cdot I(\lambda)}$. Notice that x increases monotonically in λ due to Lemma 2.1. Hence, $f(\lambda)$ is a one-to-one correspondence between the sets $[\underline{\lambda}, \lambda']$ and $[\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda' \cdot I(\lambda')}]$. By plugging x in (A1.22) we get:

$$\bar{\Delta}(x) = \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_j} \cdot m_j + \mathrm{e}^{Y_j} \cdot (m_j - \bar{m}(\underline{\lambda}) - r) \cdot (x - \mathrm{e}^{Y_0})}{(\mathrm{e}^{Y_j} + x) \cdot x} \tag{A1.23}$$

   In order to show that $\bar{\Delta}(\lambda)$ decreases in λ for $\lambda \in [\underline{\lambda}, \lambda']$, we will show that $\bar{\Delta}(x)$ decreases in x for $x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda' \cdot I(\lambda')}]$. The first derivative of (A1.23) with respect to x is:

$$\frac{d\bar{\Delta}(x)}{dx} = -\frac{\mathrm{e}^{Y_j} x \left[x(m_j - \bar{m}(\underline{\lambda}) - r) + 2\mathrm{e}^{Y_0}(\bar{m}(\underline{\lambda}) + r)\right] + \mathrm{e}^{2Y_j + Y_0}(\bar{m}(\underline{\lambda}) + r)}{(\mathrm{e}^{Y_j} + x)^2 \cdot x^2} \tag{A1.24}$$

The denominator is clearly positive. We will argue that the numerator is also positive and so $\frac{d\bar{\Delta}(x)}{dx} < 0$ because of the negative sign in front. To that end, it suffices to show that

$$\left[x(m_j - \bar{m}(\underline{\lambda}) - r) + 2\mathrm{e}^{Y_0}(\bar{m}(\underline{\lambda}) + r)\right] > 0 \text{ for all } x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda}\cdot I(\underline{\lambda})}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda'\cdot I(\lambda')}] \tag{A1.25}$$

By Assumption 2.1, we know that $\Delta^j(\lambda) > 0$ for $\lambda \in [\underline{\lambda}, \lambda']$, i.e., promoting the optimal product is more profitable than no-promotion. It implies the following:

$$\begin{aligned}&\Delta^j(\lambda)\\ &= \frac{\mathrm{e}^{Y_j}\cdot m_j}{\mathrm{e}^{Y_0} + \mathrm{e}^{Y_j} + \mathrm{e}^{-c+\lambda\cdot I(\lambda)}} - \frac{\mathrm{e}^{Y_j}\cdot \mathrm{e}^{-c+\lambda\cdot I(\lambda)}}{(\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(\lambda)})\cdot(\mathrm{e}^{Y_0} + \mathrm{e}^{Y_j} + \mathrm{e}^{-c+\lambda\cdot I(\lambda)})}\cdot(\bar{m}(\lambda) + r) > 0\\ &\text{for all } \lambda \in [\underline{\lambda}, \lambda']\end{aligned}$$

By some algebraic manipulation, we get

$$m_j > \frac{\mathrm{e}^{-c+\lambda\cdot I(\lambda)}}{\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(\lambda)}}\cdot(\bar{m}(\lambda) + r) \;\text{ for all } \lambda \in [\underline{\lambda}, \lambda'] \tag{A1.26}$$

Since $\bar{m}(\underline{\lambda}) \le \bar{m}(\lambda)$ for all $\lambda \in [\underline{\lambda}, \lambda']$, inequality (A1.26) implies

$$m_j > \frac{\mathrm{e}^{-c+\lambda\cdot I(\lambda)}}{\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I(\lambda)}}\cdot(\bar{m}(\underline{\lambda}) + r) \;\text{ for all } \lambda \in [\underline{\lambda}, \lambda'] \tag{A1.27}$$

By plugging x in (A1.27) and some algebraic manipulation, we get

$$x\cdot\big(m_j - \bar{m}(\underline{\lambda}) - r\big) + \mathrm{e}^{Y_0}\big(\bar{m}(\underline{\lambda}) + r\big) > 0 \text{ for all } x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda}\cdot I(\underline{\lambda})}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda'\cdot I(\lambda')}] \tag{A1.28}$$

Inequality (A1.28) implies that (A1.25) holds for all $x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda}\cdot I(\underline{\lambda})}, \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda'\cdot I(\lambda')}]$. Hence, $\bar{\Delta}(x)$ is decreasing in x in that interval. Then we have the following sequence of inequalities:

$$\Delta^j(\underline{\lambda}) = \bar{\Delta}\big(\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda}\cdot I(\underline{\lambda})}\big) > \bar{\Delta}\big(\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda'\cdot I(\lambda')}\big)$$
$$\geq \underline{\Delta}\big(\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda'\cdot I(\lambda')}\big) = \Delta^j(\lambda') = \Delta^{i^*(\lambda')}(\lambda').$$

The first inequality follows as $\bar{\Delta}(x)$ is decreasing in x and $\big(\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\underline{\lambda}\cdot I(\underline{\lambda})}\big) < \big(\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda'\cdot I(\lambda')}\big)$. The second inequality follows by the definition of $\bar{\Delta}(\cdot)$ and $\underline{\Delta}(\cdot)$. Since the same arguments can be made for all such sub-intervals in $[\underline{\lambda}, \bar{\lambda}]$, we conclude that $\Delta^j(\underline{\lambda}) > \Delta^k(\bar{\lambda})$ when $\beta = 1$.

Proof of Proposition 2.8. Let $i^*(\underline{c}) = j$, $i^*(\bar{c}) = k$, and $\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) = M$ for ease of exposition. We will follow a similar procedure to the proof of Proposition 2.6 and prove in two steps.

1. $\Delta^j(\underline{c}) - \Delta^k(\bar{c})$ is increasing linearly in $\beta \geq 0$. It follows the same arguments as the first part of the proof of Proposition 2.6 by replacing S and S^+ with $\underline{c}$ and $\bar{c}$, respectively.

2. $\Delta^j(\underline{c}) < \Delta^k(\bar{c})$ when $\beta = 0$.

   Assume w.l.o.g. that $\exists\, c' \in [\underline{c}, \bar{c}]$ such that $\underset{i\in S}{\operatorname{argmax}}\{\pi^i(c)\} = j\ \forall c \in [\underline{c}, c']$ and $\underset{i\in S}{\operatorname{argmax}}\{\pi^i(c)\} = i^*(c) \neq j$ for $c \in [c', c' + \epsilon]$ for some (small) $\epsilon > 0$. We will show that $\Delta^j(\underline{c}) < \Delta^j(c') = \Delta^{i^*(c')}(c')$, where the equality holds because $\pi^j(c') = \pi^{i^*(c')}(c')$. (Notice that $i^*(c')$ and k are not necessarily the same product.) Since the same argument holds for any sub-interval in $[\underline{c}, \bar{c}]$ and $\Delta^{(\cdot)}(c)$ is continuous, we can conclude that $\Delta^j(\underline{c}) < \Delta^k(\bar{c})$. Define the following function:

$$\bar{\Delta}(c) := \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_j} \cdot m_j + \mathrm{e}^{Y_j} \cdot (m_j - M - r) \cdot \mathrm{e}^{-c+\lambda\cdot I}}{(\mathrm{e}^{Y_0} + \mathrm{e}^{Y_j} + \mathrm{e}^{-c+\lambda\cdot I}) \cdot (\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I})}, \tag{A1.29}$$

   which equals $\Delta^j(c)$ for $\beta = 1$. We will show that $\bar{\Delta}(c)$ is increasing in c for $c \in [\underline{c}, c']$. To that end, consider the function $f : [\underline{c}, c'] \to [\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda\cdot I}, \mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda\cdot I}]$ such that $x = f(c) = \mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda\cdot I}$. Notice that x decreases monotonically in c. Hence, $f(c)$ is a one-to-one correspondence between the sets $[\underline{c}, c']$ and $[\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda\cdot I}, \mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda\cdot I}]$. By plugging x in

(A1.29) we get:

$$\bar{\Delta}(x) = \frac{\mathrm{e}^{Y_0} \cdot \mathrm{e}^{Y_j} \cdot m_j + \mathrm{e}^{Y_j} \cdot (m_j - M - r) \cdot (x - \mathrm{e}^{Y_0})}{(\mathrm{e}^{Y_j} + x) \cdot x} \tag{A1.30}$$

In order to show that $\bar{\Delta}(c)$ increases in c for $c \in [\underline{c}, c']$, we will show that $\bar{\Delta}(x)$ decreases in x for $x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda \cdot I}, \mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda \cdot I}]$. To that end, we compute the first derivative of (A1.30) as:

$$\frac{d\bar{\Delta}(x)}{dx} = -\frac{\mathrm{e}^{Y_j} x \left[x(m_j - M - r) + 2\mathrm{e}^{Y_0}(M + r)\right] + \mathrm{e}^{2Y_j + Y_0}(M + r)}{(\mathrm{e}^{Y_j} + x)^2 \cdot x^2} \tag{A1.31}$$

The denominator is clearly positive. We will argue that the numerator is also positive and so $\dfrac{d\bar{\Delta}(x)}{dx} < 0$ because of the negative sign in front. To that end, it suffices to show that

$$\left[x(m_j - M - r) + 2\mathrm{e}^{Y_0}(M + r)\right] > 0 \text{ for all } x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda \cdot I}, \mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda \cdot I}] \tag{A1.32}$$

It follows from Assumption 2.1 that $m_j \geq \dfrac{\mathrm{e}^{-c+\lambda \cdot I}}{\mathrm{e}^{Y_0} + \mathrm{e}^{-c+\lambda \cdot I}} \cdot (M + r)$ for all $c \in [\underline{c}, c']$. By plugging x in and some algebraic manipulation, we get $x \cdot \big(m_j - M - r\big) + \mathrm{e}^{Y_0}\big(M + r\big) \geq 0$ for all $x \in [\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda \cdot I}, \mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda \cdot I}]$. This implies that (A1.32) holds for all x in the interval $[\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda \cdot I}, \mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda \cdot I}]$. Hence, $\bar{\Delta}(x)$ is decreasing in x in that interval. Then, we can write $\Delta^j(\underline{c}) = \bar{\Delta}\big(\mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda \cdot I}\big) < \bar{\Delta}\big(\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda \cdot I}\big) = \Delta^j(c') = \Delta^{i^*(c')}(c')$. The inequality follows because $\bar{\Delta}(x)$ is decreasing in x, and $(\mathrm{e}^{Y_0} + \mathrm{e}^{-c'+\lambda \cdot I}) < (\mathrm{e}^{Y_0} + \mathrm{e}^{-\underline{c}+\lambda \cdot I})$. Since the same arguments can be made for all such sub-intervals in $[\underline{c}, \bar{c}]$, we conclude that $\Delta^j(\underline{c}) < \Delta^k(\bar{c})$ when $\beta = 1$.

Proof of Proposition 2.9. Suppose we have two products, i and j, such that $Y_i \geq Y_j$ and $m_i \geq m_j$, or equivalently, product i dominates product j. Our aim is to show that product j is never optimal to promote.

Let $\Pi^0 = P_A^0 \cdot \big(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r\big)$. We have $\tilde{\Pi}^x = (1 + \beta\phi_x)\left(P_D^x \cdot m_x + (1 - P_D^x) \cdot \Pi^0\right) \forall x \in S$.

We consider two cases separately: $(m_i > \Pi^0)$ and $(m_i \leq \Pi^0)$. If $(m_i > \Pi^0)$, we have

$$\tilde{\Pi}^j = (1 + \beta\phi_j)\left(P_D^j \cdot m_j + (1 - P_D^j) \cdot \Pi^0\right) \leq (1 + \beta\phi_i)\left(P_D^i \cdot m_i + (1 - P_D^i) \cdot \Pi^0\right) = \tilde{\Pi}^i \tag{A1.33}$$

The inequality holds because $m_i \geq m_j$, $P_D^i \geq P_D^j$, $m_i > \Pi^0$, and $\phi_i \geq \phi_j$, and the inequality is strict because we have assumed no two products i and j have $m_i = m_j$ and $Y_i = Y_j$. Therefore we either have $m_i > m_j$ or $P_D^i > P_D^j$, or both.

Next consider the case where ($m_i \leq \Pi^0$). Since $m_j \leq m_i \leq \Pi^0$, we have $\pi^j \leq \Pi^0$. This also implies that $\tilde{\Pi}^j = (1+\beta\phi_j)\pi^j \leq (1+\beta\phi_{max})\Pi^0 < \tilde{\Pi}^k$ where product k is as defined in Assumption 2.3. Hence, it is never optimal to promote the dominated product j.

Proof of Proposition 2.10. Because of the optimality of $i^*(r_1)$ for $r = r_1$ and of $i^*(r_2)$ for $r = r_2$, we have $\tilde{\Pi}^{i^*(r_1)}(r_1) - \tilde{\Pi}^{i^*(r_2)}(r_1) \geq 0$ and $\tilde{\Pi}^{i^*(r_2)}(r_2) - \tilde{\Pi}^{i^*(r_1)}(r_2) \geq 0$. Adding these inequalities, we get $\tilde{\Pi}^{i^*(r_1)}(r_1) - \tilde{\Pi}^{i^*(r_2)}(r_1) + \tilde{\Pi}^{i^*(r_2)}(r_2) - \tilde{\Pi}^{i^*(r_1)}(r_2) \geq 0$. By substituting using the definition of $\tilde{\Pi}$ and simplifying, we find that the following is an equivalent inequality:

$$\begin{aligned} & r_1 \left[P_A^{i^*(r_1)}(1+\beta\phi_{i^*(r_1)}) - P_A^{i^*(r_2)}(1+\beta\phi_{i^*(r_2)})\right] \\ & +r_2 \left[P_A^{i^*(r_2)}(1+\beta\phi_{i^*(r_2)}) - P_A^{i^*(r_1)}(1+\beta\phi_{i^*(r_1)})\right] \geq 0. \end{aligned}$$

This implies that $(r_1 - r_2)P_A^{i^*(r_1)}(1+\beta\phi_{i^*(r_1)}) \geq (r_1 - r_2)P_A^{i^*(r_2)}(1+\beta\phi_{i^*(r_2)})$ and therefore that $P_A^{i^*(r_1)}(1+\beta\phi_{i^*(r_1)}) \geq P_A^{i^*(r_2)}(1+\beta\phi_{i^*(r_2)})$ because $r_1 > r_2$.

APPENDIX 2: NUMERICAL STUDY REPORTED IN TABLE 2.2

Without loss of generality, we normalized $Y_0 = 0$ in all 100 category instances. For each category instance, we generated 15 products – each of which is characterized by a mean utility (Y) and a profit margin (m). The profit margins are generated randomly following a uniform distribution $U(0.5, 1)$ implying that the highest margin can be at most two times the lowest margin. The mean utilities are generated randomly by following a uniform distribution $U(Y_{min}, Y_{max})$ so that $\frac{e^{Y_{max}} - e^{Y_{min}}}{2} = e^{Y_0}$, and the most popular product has a market share that is at most eight times larger than the least popular product (or equivalently, $\frac{e^{Y_{min}}}{e^{Y_{max}}} = 0.125$). We further assumed that (Y, m) pairs are negatively correlated ($\rho = -0.75$) because we expect the manufacturers to compensate for lower popularity by higher retail margin considering retailer's scarce shelf space.

We considered different levels of customer heterogeneity ($\lambda = 0.01, 0.25, 0.5, 0.75, 1$). We observed that average performances of Y-Policy and m-Policy improve whereas θ-Policy's average performance worsens as λ increases. This is because θ-Policy maximizes the display profit whose relative contribution to the overall profit (display + aisle) is smaller when λ is bigger. For ease of exposition, we reported only the results for the customer heterogeneity parameter $\lambda = 0.75$ in Table 2.2.

The low and high values of the expected impulse spending parameter r are set to 0.1 and 0.3, respectively. This implies that an average customer's expected impulse spending in the aisle is at most 10% and 30% of the highest margin. (Multiplying all product margins and the expected impulse spending parameter r by a constant does not change the optimal product.) The transit cost parameter c is set relative to the expected maximum utility obtained from the aisle, which is $\lambda \cdot I(\mathbf{Y}, \lambda)$. In particular, we set $c_{low} = 0.25 \cdot \lambda \cdot I(\mathbf{Y}, \lambda)$ and $c_{high} = 0.75 \cdot \lambda \cdot I(\mathbf{Y}, \lambda)$. Note that all category instances satisfy Assumptions 2.1 and 2.2 under these settings.

APPENDIX 3: INCORPORATING CORRELATED UTILITIES

A3.1 Methodology and Technical Background

An assumption of the NMNL model is that the random component of the utilities (ϵ_j's) corresponding to alternatives in different nests are independent. This may not strictly be the case in our setup because the featured product appears in both the nests D and A, and we would expect a correlation among the utilities for the same product in two locations. However, the utility of the product in the aisle may not perfectly match the utility of the same product in the display, as the opportunity to compare a product with competitor products may impact a customer's utility for the product.

This implies the need for decomposing the random utility components (ϵ_j's) into nest-specific and alternative-specific components, where the location-specific component captures the heterogeneity in customers' perceived opportunity to compare all available products in the assortment and the product-specific component captures the heterogeneity in customers' preferences over different product attributes. Notice that the NMNL model requires the marginal distribution of ϵ_j's to be Type 1 extreme value distribution. Moreover, ϵ_j's within the same nest are assumed to have a certain correlation specified by the parameter λ. Hence, it is not a straightforward task to decompose ϵ_j's while maintaining the distributional assumptions of NMNL model. Cardell [1997] establishes this decomposition as follows:

Theorem A3.1 (Cardell, 1997)**.** *For $0 < \lambda < 1$ and ξ, a random variable distributed as Type 1 extreme value, there exists a unique distribution, denoted $C(\lambda)$, such that for ν, a random variable, ν and ξ independent, then $\nu + \lambda\xi$ is a random variable distributed as Type 1 extreme value, iff ν is distributed as $C(\lambda)$ where the probability density function of $C(\lambda)$ is $f_\lambda(\nu) = (1/\lambda) \sum_{n=0}^{\infty} \frac{(-1)^n e^{-n\nu}}{n!\Gamma(-\lambda n)}$. The cumulative distribution function of the $C(\lambda)$ does not have a closed-form representation.*

Now consider the following utilities for each alternative in our model:

$$U_0 = Y_0 + \underbrace{\nu_0 + \lambda\xi_0}_{=\epsilon_0} \text{ (no-purchase alternative),} \tag{A3.1}$$

$$U_i = Y_i + \underbrace{\nu_D + \lambda\xi_i}_{=\epsilon_i} \text{ (display alternative when product } i \text{ is displayed),} \tag{A3.2}$$

$$U_j = Y_j - c + \underbrace{\nu_A + \lambda\xi_j}_{=\epsilon_j} \text{ (alternative corresp. to prod. } j \text{ in the aisle).} \tag{A3.3}$$

This decomposition allows us to explicitly model location-specific (ν) and product-specific (ξ) random components of the utility. Note that when $\lambda = 0$, $C(0)$ is the Type 1 extreme value distribution and ϵ_j's corresponding to the alternatives in the aisle nest is fully correlated whereas they are independent when $\lambda = 1$.

In our simulations, each customer is represented by two randomly drawn vectors: $\langle \nu_0, \nu_A, \nu_D \rangle$ and $\langle \xi_0, .., \xi_{15} \rangle$. A customer associates the same realization of the product-specific random utility component (ξ_j) for a product located both in the display and in the aisle. This is how we incorporate correlated utilities for these two alternatives in different nests. We similarly model correlation between the no purchase option in the display and aisle. Note that the nest-specific random components corresponding to these two alternatives (ν_D and ν_A) are different, implying that a customer may still choose purchasing the displayed product upon visiting the aisle.

A3.2 Simulation Parameters

We perform a simulation study to measure the error caused by our approximation. We use the same 100 category instances randomly generated as described in Appendix 2, and we choose the remaining parameters (r, c, λ) to cover a large range of scenarios conforming to our Assumption 1. In particular, we set three levels (low, med, high) for each parameter, with the low values set to the minimum value in their defined range ($r_{min} = 0, c_{min} = 0, \lambda_{min} = 0.01$). We set r_{high} to the maximum value the expected impulse spending in the aisle can take without violating Assumption 1, and c_{high} is calibrated so that 25% of customers visit the aisle in case of no-promotion. Finally, we set the medium values of the parameters to the average of the low and high values. For each category instance and parameter setting, we randomly generate 10,000 customers.

A3.3 Evaluating the Independence Assumption

Table A3.1 shows the impact on expected profits (evaluated with correlation assumption) of promoting the product recommended given the independence assumption compared with promoting the product recommended given the correlation assumption. We note that the average impacts are smaller than 0.5% in all the cells, with the largest average impacts occuring for large λ and large c. This is expected because the product-specific random components (ξ_j's), which cause the deviation between correlated and independent cases, are amplified by larger λ as shown in Equation (A3.1)-(A5.3). We conclude that the typical impact of our independence assumption is small. We note that we occasionally see relatively large impacts for individual instances up to 9%, but these are rare and tend to be largest for large λ and c.

	λ_{low}			λ_{med}			λ_{high}		
	c_{low}	c_{med}	c_{high}	c_{low}	c_{med}	c_{high}	c_{low}	c_{med}	c_{high}
r_{low}	0.002%	0.003%	0.013%	0.114%	0.132%	0.253%	0.077%	0.328%	0.359%
r_{med}	0.005%	0.012%	0.005%	0.164%	0.286%	0.264%	0.196%	0.334%	0.432%
r_{high}	0.006%	0.012%	0.012%	0.054%	0.318%	0.371%	0.093%	0.487%	0.474%

Table A3.1: Average percent loss from promoting the product recommended using the independence assumption compared with the product recommended using the correlation assumption.

We also numerically re-examined our directional insights from Section 2.4.2 (Proposition 2.3-2.5) under the correlation assumption, finding that all of the results are preserved for the vast majority of instances we tried. We omit the details due to space considerations.

APPENDIX 4: SIMULATION SUPPORT FOR SECTION 2.5

We ran simulations to verify the results in Section 2.5 (Propositions 2.6-2.8) when comparing categories that are roughly similar but not identical. In particular, we considered all pairs of the categories we used in our earlier numerical analysis in Section 4 and Section 4.8. Specifically, we obtained $\binom{100}{2} = 4950$ pairs of categories. Then, we considered difference in one parameter (namely assortment size $|S|$, customer heterogeneity λ, and transit cost c) at a time in the following way: For each pair, we set the value of the parameter under consideration high for one of the categories and low for the other one. All the other parameters are set to the medium level as described in Section A.3.2. Proposition 2.6 holds for more than 95% of the instances, whereas Propositions 2.7 and 2.8 hold for all instances. Hence, the choice of category to promote pivots on the demand expansion parameter β as suggested by these results even if the categories are *roughly* similar.

APPENDIX 5: PROOFS OF ANALYTICAL RESULTS IN CHAPTER 3

When product i is being featured with a retailer (R) induced discount of $\delta \geq 0$ and a promotional display (D), the expected profit of the retailer can be written as

$$\pi_R^{i,D}(\delta) = M + \left(m_i + r\frac{e^{b\delta}-1}{e^{b\delta}(1+e^c)-1} - \delta\frac{e^{b\delta}(1+e^c)}{e^{b\delta}(1+e^c)-1} - M\right)\left(\frac{e^{Y_i}\left(\dfrac{e^{b\delta}-1}{e^c}+e^{b\delta}\right)}{\Sigma + e^{Y_i}\left(\dfrac{e^{b\delta}-1}{e^c}+e^{b\delta}\right)}\right) \tag{A5.1}$$

When product i is being featured with a retailer (R) induced discount of $\delta \geq 0$ but without a promotional display (ND), the expected profit of the retailer can be written as

$$\pi_R^{i,ND}(\delta) = M + \left(m_i + r - \delta\frac{e^{b\delta}}{e^{b\delta}-1} - M\right)\left(\frac{e^{Y_i}\left(\dfrac{e^{b\delta}-1}{e^c}\right)}{\Sigma + e^{Y_i}\left(\dfrac{e^{b\delta}-1}{e^c}\right)}\right) \tag{A5.2}$$

When product i is being featured with a manufacturer (M) induced discount of $\delta \geq 0$, a pass through rate of $\alpha \in [0, 1]$, and a promotional display (D), the expected profit of the retailer can be written as

$$\pi_M^{i,D}(\delta, \alpha) =$$

$$M + \left(m_i + r\frac{e^{\alpha b\delta}-1}{e^{\alpha b\delta}(1+e^c)-1} + \delta(1-\alpha)\frac{e^{\alpha b\delta}(1+e^c)}{e^{\alpha b\delta}(1+e^c)-1} - M\right)\left(\frac{e^{Y_i}\left(\dfrac{e^{\alpha b\delta}-1}{e^c}+e^{\alpha b\delta}\right)}{\Sigma + e^{Y_i}\left(\dfrac{e^{\alpha b\delta}-1}{e^c}+e^{\alpha b\delta}\right)}\right) \tag{A5.3}$$

When product i is being featured with a manufacturer (M) induced discount of $\delta \geq 0$, a pass through rate of $\alpha \in [0, 1]$, but without a promotional display (ND), the expected profit of the retailer can be written as

$$\pi_M^{i,ND}(\delta,\alpha) = \begin{cases} M + \left(m_i + r + \delta(1-\alpha)\dfrac{e^{\alpha b\delta}}{e^{\alpha b\delta}-1} - M\right)\left(\dfrac{e^{Y_i}\left(\dfrac{e^{\alpha b\delta}-1}{e^c}\right)}{\Sigma + e^{Y_i}\left(\dfrac{e^{\alpha b\delta}-1}{e^c}\right)}\right) & \text{if } \alpha > 0 \\ M + \delta\dfrac{e^{Y_i - c}}{\Sigma + e^{Y_i - c}} & \text{if } \alpha = 0 \end{cases} \tag{A5.4}$$

Proof of Proposition 3.1. See Theorem 1 in Gallego and Wang [2014]. Our model is a special case of their model with a single nest with the nest coefficient $\gamma = 1$, and product invariant price sensitivity parameter $\beta_i = b$, $\forall i \in \mathcal{N}$.

Proof of Proposition 3.2. The case of no-discount ($\delta = 0$) is shown in the first paper. Hence, we now assume $\delta > 0$. Suppose we have two products, i and j, such that $Y_i \geq Y_j$ and $m_i \geq m_j$, or equivalently, product i dominates product j. Our aim is to show that for any $\delta \geq 0$

1. $\max\{M, \pi^i_{R,D}\} \geq \pi^j_{R,D}$ and

2. $\max\{M, \pi^i_{R,ND}\} \geq \pi^j_{R,ND}$.

The first statement implies that either not performing any promotional activity, in which case the retailer's profit is M, or promoting the dominating product i on the promotional display with a discount of δ, in which case the retailer's profit is $\pi^i_{R,D}$, is better than promoting the dominated product j. The second statement implies a similar argument for the case of a retailer induced discount of $\delta \geq 0$ without a promotional display.

We will prove the first statement, and the second statement follows similarly. We will consider two cases:

1. If $m_i \leq \left(M + \delta\dfrac{e^{b\delta}(1+e^c)}{e^{b\delta}(1+e^c)-1} - r\dfrac{e^{b\delta}-1}{e^{b\delta}(1+e^c)-1}\right)$, Equation A3.1 implies that $M \geq \pi^i_{R,D}$. Since $m_j \leq m_i$, it also implies that $M \geq \pi^j_{R,D}$. Therefore, not performing any promotional activity is better, i.e., $M \geq \max\{\pi^i_{R,D}, \pi^j_{R,D}\}$

2. If $m_i > \left(M + \delta\dfrac{e^{b\delta}(1+e^c)}{e^{b\delta}(1+e^c)-1} - r\dfrac{e^{b\delta}-1}{e^{b\delta}(1+e^c)-1}\right)$, Equation A3.1 implies that $\pi^i_{R,D} > M$. The expression in Equation A3.1 is increasing in both m_i and Y_i, therefore $\pi^i_{R,D} > \pi^j_{R,D}$.

Proof of Proposition 3.3. $\dfrac{\partial(\pi_R^{i,D}(\delta) - \pi_R^{i,ND}(\delta))}{\partial m_i} > 0$. Hence, $\pi_R^{i,D}(\delta)$ and $\pi_R^{i,ND}(\delta)$ intersect at a unique value m_i. Solving $\pi_R^{i,D}(\delta) = \pi_R^{i,ND}(\delta)$ for m_i yields the result after a few steps of algebraic manipulations.

Proof of Corollary 3.1. Follows directly from Proposition 3.2 and 3.3.

Proof of Corollary 3.2. Assume $r = 0$. For any manufacturer induced discount level $\delta > 0$, we have

$$\pi^i_{M,D}(\delta, 1) - \pi^i_{M,ND}(\delta, 1) = \frac{e^{Y_i + b\delta}\Sigma}{(e^{Y_i - c}(e^{b\delta} - 1) + \Sigma)^2 + e^{Y_i + b\delta}\left(e^{Y_i - c}(e^{b\delta} - 1) + \Sigma\right)}(m_i - M) \quad \text{(A5.5)}$$

$$\pi^i_{M,D}(\delta, 0) - \pi^i_{M,ND}(\delta, 0) = \frac{e^{Y_i}\Sigma + \delta(e^{Y_i}\Sigma - e^{2Y_i - c})}{\Sigma(\Sigma + e^{Y_i})}(m_i - M) \quad \text{(A5.6)}$$

It is straightforward to obtain the results given the above differences.

Proof of Proposition 3.4. $\dfrac{\partial(\pi_M^{i,D}(\delta,\alpha) - \pi_M^{i,ND}(\delta,\alpha))}{\partial m_i} > 0$. Hence, $\pi_M^{i,D}(\delta,\alpha)$ and $\pi_M^{i,ND}(\delta,\alpha)$ intersect at a unique value m_i. Solving $\pi_M^{i,D}(\delta,\alpha) = \pi_M^{i,ND}(\delta,\alpha)$ for m_i (when $r = 0$) yields the result after a few steps of algebraic manipulations.

Proof of Proposition 3.5. We will prove it by claiming the following two arguments:

1. In the optimal solution of Problem 3.14, $a^* = ND \Rightarrow \alpha^* = 0$.

   Suppose $\alpha > 0$. We are going to show the following relation which implies the result:

$$\pi_M^{i,ND}(\delta, 0) \leq \pi_M^{i,ND}(\delta,\alpha) \Rightarrow \pi_M^{i,ND}(\delta,\alpha) \leq \pi_M^{i,D}(\delta,\alpha) \quad \text{(A5.7)}$$

   We can write $\pi_M^{i,ND}(\delta, 0)$ as a *constant* function of α as follows:

$$\pi_M^{i,ND}(\delta, 0) = \frac{\Sigma M + e^{Y_i - c}(e^{\alpha b\delta} - 1)\left(M + \delta\left(\dfrac{e^{Y_i - c}}{\Sigma} + \dfrac{1}{e^{e^{\alpha b\delta}-1}}\right)\right)}{\Sigma + e^{Y_i - c}(e^{\alpha b\delta} - 1)} \quad \text{(A5.8)}$$

We can write $\pi_M^{i,ND}(\delta,\alpha)$ as follows:

$$\pi_M^{i,ND}(\delta,\alpha) = \frac{\Sigma M + e^{Y_i-c}(e^{\alpha b\delta}-1)\left(m_i + \delta(1-\alpha)\dfrac{e^{\alpha}b\delta}{e^{\alpha}b\delta - 1}\right)}{\Sigma + e^{Y_i-c}(e^{\alpha b\delta}-1)} \tag{A5.9}$$

Comparing the above two expressions, we have

$$\pi_M^{i,ND}(\delta,0) \le \pi_M^{i,ND}(\delta,\alpha) \iff M - m_i \le \delta \underbrace{\left(1 - \frac{e^{Y_i-c}}{\Sigma} - \alpha\left(1 + \frac{1}{e^{\alpha b\delta}-1}\right)\right)}_{\omega_1} \tag{A5.10}$$

By Proposition 3.4, we know that

$$\pi_M^{i,ND}(\delta,\alpha) \le \pi_M^{i,D}(\delta,\alpha) \iff M - m_i \le \delta \underbrace{\left(1 - \frac{e^{Y_i-c}}{\Sigma}\right)(1-\alpha)}_{\omega_2} \tag{A5.11}$$

Since $\omega_1 < \omega_2$ for $\alpha > 0$, condition in (A5.11) is satisfied whenever condition in (A5.10) is satisfied. Therefore, we have $\pi_M^{i,ND}(\delta,0) \le \pi_M^{i,ND}(\delta,\alpha) \Rightarrow \pi_M^{i,ND}(\delta,\alpha) \le \pi_M^{i,D}(\delta,\alpha)$.

2. If no-display with $\alpha = 0$ is better than display with $\alpha = 0$, it is also better than display with any $\alpha \in [0,1]$ by Proposition 3.4.

   Combining the above two arguments, we conclude that it is optimal to display product i if and only if $m_i \ge M - \delta\left(1 - \frac{e^{Y_i-c}}{\Sigma}\right)$. Moreover, if this condition is not satisfied, the optimal pass through rate is $\alpha^* = 0$.

Proof of Proposition 3.6. Define x such that $1 \le x = e^{\alpha b\delta} \le e^{b\delta}$ for $\alpha \in [0,1]$. Define the following functions:

$$f^{i,D}(x) := M + \left(m_i + r\frac{x-1}{x(1+e^c)-1} + \delta\left(1 - \frac{\ln x}{b\delta}\right)\frac{x(1+e^c)}{x(1+e^c)-1} - M\right) \times \left(\frac{e^{Y_i}\left(\frac{x-1}{e^c} + x\right)}{\Sigma + e^{Y_i}\left(\frac{x-1}{e^c} + x\right)}\right) \tag{A5.12}$$

$$f^{i,ND}(x) := M + \left(m_i + r + \delta\left(1 - \frac{\ln x}{b\delta}\right)\frac{x}{x-1} - M\right)\left(\frac{e^{Y_i}\left(\frac{x-1}{e^c}\right)}{\Sigma + e^{Y_i}\left(\frac{x-1}{e^c}\right)}\right) \tag{A5.13}$$

Note that $f^{i,D}(x) = \pi_M^{i,D}(\delta,\alpha)$ and $f^{i,ND}(x) = \pi_M^{i,ND}(\delta,\alpha)$. Due to the monotonicity of the transformation $x = e^{\alpha b\delta}$, it suffices to show that $\underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,ND}(x)) \geq \underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,D}(x))$.

We prove in two steps. First, we show that $f^{i,D}(x)$ and $f^{i,ND}(x)$ are unimodal where both of them are first monotonically increasing and then monotonically decreasing (the switch occurs at different x values). Second, we show that wherever $f^{i,D}(x)$ attains its maximum, $f^{i,ND}(x)$ is still increasing, implying that the $\underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,ND}(x)) \geq \underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,D}(x))$, which concludes the proof.

Step 1: $f^{i,D}(x)$ and $f^{i,ND}(x)$ are unimodal where both of them are first monotonically increasing and then monotonically decreasing.

The first derivatives of $f^{i,D}(x)$ and $f^{i,ND}(x)$ are as follows:

$$\frac{df^{i,D}(x)}{dx} = \frac{(1+e^c)e^{Y_i-c}}{b\left((x-1)e^{Y_i-c} + e^{Y_i}x + \Sigma\right)^2} \times \underbrace{\left((-b\delta - x + \ln(x) + 1)e^{Y_i-c} - e^{Y_i}x - \Sigma\left(\ln(x) + 1 + (M-\delta-m_i)b\right)\right)}_{\Phi_1} \tag{A5.14}$$

$$\frac{df^{i,ND}(x)}{dx} = \frac{e^{Y_i-c}}{b\left((x-1)e^{Y_i-c} + \Sigma\right)^2} \times \underbrace{\left((-b\delta - x + \ln(x) + 1)e^{Y_i-c} - \Sigma\left(\ln(x) + 1 + (M-\delta-m_i)b\right)\right)}_{\Phi_2} \tag{A5.15}$$

We can write Φ_1 and Φ_2 as follows:

$$\Phi_1 = (e^{Y_i - c} - \Sigma)\ln(x) - (e^{Y_i-c} + e^{Y_i})x + C_1 \tag{A5.16}$$

$$\Phi_2 = (e^{Y_i - c} - \Sigma)\ln(x) - e^{Y_i-c}x + C_2 \tag{A5.17}$$

where C_1 and C_2 are some constants (in x). Note that both Φ_1 and Φ_2 are monotonically decreasing in x. This means, the first derivatives of $f^{i,D}(x)$ and $f^{i,ND}(x)$ change sign at most once. Moreover, if they change sign, they both start positive for $x = 1$ and then become negative for $x = e^{b\delta}$

Step 2: $\underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,ND}(x)) \geq \underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,D}(x))$.

If the first derivatives of $f^{i,D}(x)$ and $f^{i,ND}(x)$ do not change sign in the interval $1 \leq x \leq e^{b\delta}$, the both $f^{i,D}(x)$ and $f^{i,ND}(x)$ are either increasing or decreasing in the entire interval, implying that either $\underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,ND}(x)) = \underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,D}(x)) = 0$ or $\underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,ND}(x)) = \underset{e^{b\delta}\geq x\geq 1}{\operatorname{argmax}}(f^{i,D}(x)) = 1$. In either case, we are done. Now, we consider the case when one of the first derivatives change sign.

After some algebraic manipulation, we get the following identity:

$$f^{i,D}(x) = f^{i,ND}\left(x(1+e^c)\right) - \overbrace{\frac{e^c br - (1+e^c)\ln\left(1+e^c\right)}{b\left(1 + e^c + \dfrac{e^{c-Y_i}\Sigma - 1}{x}\right)}}^{g(x)} \tag{A5.18}$$

Differentiating both sides w.r.t. x yields

$$\frac{df^{i,D}(x)}{dx} = \frac{df^{i,ND}\left(x(1+e^c)\right)}{dx} - \frac{dg(x)}{dx} \tag{A5.19}$$

Since $\dfrac{dg(x)}{dx} > 0\ \forall x \geq 1$, we have $\left.\dfrac{df^{i,ND}\left(x(1+e^c)\right)}{dx}\right|_{x=x_0} > 0$ for any $x_0 \geq 1$ such that $\left.\dfrac{df^{i,D}(x)}{dx}\right|_{x=x_0} = 0$. This means that $f^{i,ND}(x)$ is strictly increasing at $x = x_0(1+e^c)$ where $f^{i,D}(x_0)$ is a local extremum.

Proof of Proposition 3.7. Suppose manufacturer i ($M > m_i$) offered $\delta_{min} = \dfrac{M - m_i}{1 - \dfrac{e^{Y_i - c}}{\Sigma}}$, which is the minimum δ that guarantees the retailer's display support according to Proposition 3.5. In this case, we know that the retailer is indifferent between displaying and not displaying product i, and the optimal pass-through rate is $\alpha^* = 0$. Therefore, the manufacturer's payoff when he offers

δ_{min} is:

$$\frac{e^{Y_i}(1+e^{-c})}{e^{Y_i}+\Sigma}(w_i-\delta_{min}) \tag{A5.20}$$

The minimum fix payment F that guarantees the retailer's display support is $F_{min} = \pi_M^{i,ND}(0,0) - \pi_M^{i,D}(0,0) = M - \dfrac{e^{Y_i}m_i + \Sigma M}{e^{Y_i}+\Sigma}$ at which constraint (3.22) is binding. When the fix payment is $F = F_{min}$, the manufacturer's payoff is:

$$\frac{e^{Y_i}(1+e^{-c})}{e^{Y_i}+\Sigma}w_i - F_{min} = \frac{e^{Y_i}(1+e^{-c})}{e^{Y_i}+\Sigma}w_i - \left(M - \frac{e^{Y_i}m_i+\Sigma M}{e^{Y_i}+\Sigma}\right) \tag{A5.21}$$

Now consider the following series of inequalities:

$$(1+e^{-c}) > 1 - \frac{e^{Y_i-c}}{\Sigma} \tag{A5.22}$$

$$(1+e^{-c})(M-m_i) > \left(1-\frac{e^{Y_i-c}}{\Sigma}\right)(M-m_i) \tag{A5.23}$$

$$\frac{(1+e^{-c})(M-m_i)}{\left(1-\dfrac{e^{Y_i-c}}{\Sigma}\right)} > (M-m_i) \tag{A5.24}$$

$$(1+e^{-c})\delta_{min} > (M-m_i) \tag{A5.25}$$

$$\left(\frac{e^{Y_i}}{e^{Y_i}+\Sigma}\right)(1+e^{-c})\delta_{min} > \left(\frac{e^{Y_i}}{e^{Y_i}+\Sigma}\right)(M-m_i) \tag{A5.26}$$

Inequality (A5.26) implies that the Expression (A5.20) is smaller than Expression (A5.21). Therefore, manufacturer i has a better payoff when he offers a fixed payment F_{min} compared to when he offers per unit discount δ_{min} in order to get the promotional display.

APPENDIX 6: PROOFS OF ANALYTICAL RESULTS IN CHAPTER 4

Proof of Lemma 4.1. ($\Rightarrow$) Assume that $X^C_{j,m} = 1$, $\forall j \in \mathcal{S}^C_\vee \setminus S^C_m$ and $\forall m \in M$. The profit obtained from store m in the optimal solution is:

$$\frac{\sum_{j\in S^C_m} v_j + \sum_{j\in \mathcal{S}^C_\vee \setminus S^C_m} \theta v_j (1-\tau)}{1 + \sum\limits_{k\in S^C_m} v_k + \sum\limits_{k\in \mathcal{S}^C_\vee \setminus S^C_m} \theta v_k} - K \cdot |S^C_m| \tag{A6.1}$$

$$= \left(\frac{\sum\limits_{k\in \mathcal{S}^C_\vee \setminus S^C_m} \theta v_k}{1 + \sum\limits_{k\in S^C_m} v_k + \sum\limits_{k\in \mathcal{S}^C_\vee \setminus S_m} \theta v_k} \right)(1-\tau) + \left(\frac{1 + \sum\limits_{k\in S^C_m} v_k}{1 + \sum\limits_{k\in S^C_m} v_k + \sum\limits_{k\in \mathcal{S}^C_\vee \setminus S^C_m} \theta v_k} \right) \left(\frac{\sum\limits_{k\in S^C_m} v_k}{1 + \sum\limits_{k\in S^C_m} v_k} \right)$$

$$- K \cdot |S^C_m| \tag{A6.2}$$

Notice the convex combination of $(1-\tau)$ and $\left(\frac{\sum\limits_{k\in S^C_m} v_k}{1 + \sum\limits_{k\in S^C_m} v_k} \right)$ in the above expression. Note also that the optimal solution maximizes the weight of $(1-\tau)$ by setting $X^C_{j,m} = 1$, $\forall j \in \mathcal{S}^C_\vee \setminus S^C_m$. This implies $1 - \tau > \frac{\sum\limits_{k\in S^C_m} v_k}{1 + \sum\limits_{k\in S^C_m} v_k}$. Therefore, $\tau \leq \frac{1}{1 + \sum\limits_{k\in S^C_m} v_k}$.

($\Leftarrow$) Similar to the above argument, we can show that if $\tau \leq \frac{1}{1 + \sum\limits_{k\in S^C_m} v_k}$, it is optimal to maximize the weight of $(1-\tau)$ in the above convex combination by setting $X^C_{j,m} = 1$, $\forall j \in \mathcal{S}^C_\vee \setminus S^C_m$ for an arbitrary store $m \in M$.

Proof of Proposition 4.1. Similar to Rusmevichientong et al. [2014] and Feldman and Topaloglu [2015], we use a reduction from the 2-partition problem to show that the centralized problem is NP-complete. To that end, consider the following decision version of the (CP):

CP-Decision Problem: We are given an expected profit threshold T. Is there a solution $\{S_1, .., S_M\}$ and $\{X_{j,m}, j \in \mathcal{N}, m \in \mathcal{M}\}$ that provides an expected profit of T or more in problem (CP).

In the 2-partition problem, we are given positive integers $\{w_1, .., w_N\}$ such that $\sum\limits_{j=1}^{N} w_j = W$. The 2-partition problem asks if the set of integers can be partitioned into two subsets each of which summing up to $W/2$. Using the instance of the 2-partition problem, we define the CP-decision problem as follows: There are 2 stores and N products with preference weights $v_j = w_j, \forall j \in$

$\{1, .., N\}$. The transshipment cost $\tau = 0$, the disutility of waiting $\theta = 1 - \epsilon_1$, the fixed cost $K = \epsilon_2$, and the profit threshold $T = 2 - \frac{(2-\epsilon_1)W + 2}{(1 - \epsilon_1 + 0.25(1-\epsilon_1)^2)W^2 + (2-\epsilon_1)W + 1} - N\epsilon_2$. (Note that in order to complete the proof, we also have to specify ϵ_1 and ϵ_2 in terms of the parameters of the 2-partition problem. We will do it at the end for tractability. ϵ_1 and ϵ_2 can be considered as sufficiently small positive quantities for now.)

We are going to show that there exists a solution to CP with an expected profit of T or more if and only if the set of integers $\{w_1, .., w_N\}$ can be partitioned into two subsets each of which summing up to $W/2$.

Since $\tau = 0$, we have $X_{j,m} = 1$, $\forall j \in \mathcal{S}^*_\vee \setminus S^*_m$ and $\forall m \in M$ in an optimal solution of the CP. By plugging $X_{j,m} = 1$ for all j, m in CP, we get

$$\max_{S_1,S_2} \frac{\sum\limits_{j\in S_1} v_j + \sum\limits_{j\in S_2\setminus S_1} \theta v_j(1-\tau)}{1 + \sum\limits_{j\in S_1} v_j + \sum\limits_{j\in S_2\setminus S_1} \theta v_j} + \frac{\sum\limits_{j\in S_2} v_j + \sum\limits_{j\in S_1\setminus S_2} \theta v_j(1-\tau)}{1 + \sum\limits_{j\in S_2} v_j + \sum\limits_{j\in S_1\setminus S_2} \theta v_j} - K\cdot(|S_1| + |S_2|) \tag{A6.3}$$

Since $K = \epsilon_1 > 0$ is sufficiently small, we have $|S_1 \cup S_2| = \mathcal{N}$ in the optimal solution, i.e., all the products are included in the union set. Hence, we have $K \cdot (|S_1| + |S_2|) = N\epsilon_1$.

Since $\tau = 0$, θ is arbitrarily close to 1, and $K = \epsilon_2 > 0$, we have $|S_1 \cap S_2| = \emptyset$ in an optimal solution of the CP.

The above two arguments imply that an optimal solution of the CP is indeed a 2-partition of the set of products. Let $B = \{\beta : \beta = \sum\limits_{j\in S_1} v_j/W\}$ be the discrete and finite ($|B| = 2^N$) set of feasible allocations, implying that $\sum\limits_{j\in S_1} v_j = W - \sum\limits_{j\in S_2} v_j = \beta W$. In that case, we can rewrite (A6.3) as

$$\max_{\beta\in B}\left(\frac{\beta W + (1-\beta)W\theta(1-\tau)}{1 + \beta W + (1-\beta)W\theta} + \frac{(1-\beta)W + \beta W\theta(1-\tau)}{1 + (1-\beta)W + \beta W\theta} - N\epsilon_2\right) \tag{A6.4}$$

By plugging in $\theta = 1 - \epsilon_1$ and adding/subtracting 2, we get

$$\max_{\beta\in B}\left(2 - \frac{1}{1 + \beta W + (1-\beta)W(1-\epsilon_1)} + \frac{1}{1 + (1-\beta)W + \beta W(1-\epsilon_1)} - N\epsilon_2\right) \tag{A6.5}$$

By rearranging the terms we get

$$\max_{\beta \in B} \left(2 - \frac{(2-\epsilon_1)W + 2}{(1-\epsilon_1)(W^2+W) + W + 1 + (W\epsilon_1)^2(\beta - \beta^2)} - N\epsilon_2 \right) \tag{A6.6}$$

Notice that the denominator of the second term of the above maximization problem is concave quadratic in $\beta \in [0, 1]$. So the maximum is attained for $\beta = 0.5$. When we plug in $\beta = 0.5$, the objective value is equal to the threshold T we specified. Hence, the CP has an objective value T if and only if $\beta = 0.5$, i.e., the set of integers can be partitioned into two subsets each of which summing up to $W/2$.

To complete the proof, we have to specify ϵ_1 and ϵ_2 that guarantee $|S_1 \cup S_2| = \mathcal{N}$ and $|S_1 \cap S_2| = \emptyset$.

$$\epsilon_1 = 1 - \frac{1 - \epsilon_2(1 + 1/w_{max})}{1 + \epsilon_2(1 + w_{max})} \tag{A6.7}$$

$$\epsilon_2 = \frac{2W - w_{min}}{1 + 2W - w_{min}} - \frac{2(W - w_{min})}{1 + 2(W - w_{min})} \tag{A6.8}$$

$K = \epsilon_2$ guarantees that $|S_1 \cup S_2| = \mathcal{N}$, and $\theta = 1 - \epsilon_1$ guarantees that $|S_1 \cap S_2| = \emptyset$.

Proof of Proposition 4.2.1. We are going to show the following: Pick two products $j, k \in \mathcal{N}$ such that $v_k > v_j$. In an optimal solution of the CP, if product j is offered in at least two stores (say, store m and store t), product k must be offered in these two stores, too. This suffices to show Proposition 4.2.1.

Define $V_m = \sum_{j \in S_m} v_j$, and $V_{m'} = \sum_{j \in \mathcal{S}_\vee \setminus S_m} X_{j,m} v_j$. The centralized problem can be rewritten as follows:

$$\min_{S_1,..,S_M} \sum_{m \in M} \frac{1 + \tau\theta V_{m'}}{1 + V_m + \theta V_{m'}} + K \cdot |S_m| \tag{A6.9}$$

We are going to consider all possible cases:

Case 1: $k \in S_t^C$ and $k \notin S_m^C$. We are going to show that replacing product j with k in assortment S_m^C improves the objective value and therefore contradicts with the optimality of $S_1^C, .., S_M^C$. Swapping product j and k in S_m^C leads only to the following updates for store m (other stores are not affected): $V_m \leftarrow V_m + v_k - v_j$ and $V_{m'} \leftarrow V_{m'} - v_k + v_j$. As a result, the numerator (denominator) of the first term in the summation decreases (increases) for those stores, while the second terms do not change. Hence, the objective of the centralized problem is improved.

Case 2: $k \notin S_t^C$ and $k \notin S_m^C$. We are going to show that replacing product j with k in both assortments S_m^C and S_t^C improves the objective value and therefore contradicts with the optimality of $S_1^C, .., S_M^C$.

Case 2.1: Products j and k are included in some stores' assortments other that store m and t. Replacing product j with k in both assortments S_m^C and S_t^C affects only the profits obtained from these stores. Following the arguments in Case 1, it is easy to show the the objective of the centralized problem is improved.

Case 2.2: Product j is included in an assortment of a store other than m and t, but product k is not offered in any other store's assortment.

Case 2.3: Product k is included in an assortment of a store other than m and t, but product j is not offered in any other store's assortment.

Case 2.4: Neither product j nor product k is offered in an assortment of a store other that m and t.

Proof of Proposition 4.2.2. We are going to show that the set of products offered at least in one of the stores ($\mathcal{S}_\vee^C$) is a popular subset of $\mathcal{N}$. This together with Proposition 4.2.1 suffice to prove Proposition 4.2.2.

Define $V_m = \sum_{j \in S_m} v_j$, and $V_{m'} = \sum_{j \in \mathcal{S}_\vee \setminus S_m} X_{j,m} v_j$. The centralized problem can be rewritten as follows:

$$\min_{S_1,..,S_M} \sum_{m \in M} \frac{1 + \tau\theta V_{m'}}{1 + V_m + \theta V_{m'}} + K \cdot |S_m| \tag{A6.10}$$

Assume to the contrary that $\exists j \in \mathcal{S}_\vee^*$ and $k \notin \mathcal{S}_\vee^*$ such that $v_k > v_j$. We argue that swapping product j and k improves the objective function of the centralized problem even keeping the $X_{j,m}$'s fixed for all stores and all products, and therefore contradicts with the optimality of $\{S_1^C, .., S_M^C\}$.

If $j \in \mathcal{S}_\wedge^C$, we showed in Proposition 4.2.1 that swapping j and k improves the objetive function, so we are done.

Suppose $j \notin \mathcal{S}_\wedge^C$. This implies that product j exists in a strict subset of stores' assortments. First consider the stores having product j in their assortments. Swapping j and k leads to the update $V_m \leftarrow V_m + (v_k - v_j)$ for such stores. This decreases the term in summation in (2), and hence improves the objective function. Next consider the stores not having product j in their assortments. If $X_{j,m} = 0$, we set $X_{k,m} = 0$ after replacing j and k, which implies no change for such stores. If $X_{j,m} = 1$, then we know that $\tau < \dfrac{1}{1 + \sum\limits_{k \in S_m} v_k}$ by Lemma 4.1. Hence, we have the following series of inequalities:

$$\tau < \frac{1}{1 + V_m} \tag{A6.11}$$

$$1 - \tau > \frac{V_m}{1 + V_m} \tag{A6.12}$$

$$(1 - \tau)[1 + V_m + \theta V_{m'} - \theta V_{m'}] > V_m \tag{A6.13}$$

$$(1 - \tau)[1 + V_m + \theta V_{m'}] > V_m + (1 - \tau)\theta V_{m'} \tag{A6.14}$$

$$1 - \tau > \frac{V_m + (1 - \tau)\theta V_{m'}}{[1 + V_m + \theta V_{m'}]} \tag{A6.15}$$

$$\tau < \frac{1 + \tau\theta V_{m'}}{[1 + V_m + \theta V_{m'}]} \tag{A6.16}$$

$$[1 + V_m + \theta V_{m'}]\tau\theta(v_k - v_j) < (1 + \tau\theta V_{m'})\theta(v_k - v_j) \tag{A6.17}$$

$$[1 + V_m + \theta V_{m'}][(1 + \tau\theta V_{m'}) + \tau\theta(v_k - v_j)] < (1 + \tau\theta V_{m'})\left([1 + V_m + \theta V_{m'}] + \theta(v_k - v_j)\right) \tag{A6.18}$$

$$\frac{1 + \tau\theta V_{m'} + \tau\theta(v_k - v_j)}{1 + V_m + \theta V_{m'} + \theta(v_k - v_j)} < \frac{1 + \tau\theta V_{m'}}{1 + V_m + \theta V_{m'}} \tag{A6.19}$$

The last inequality above implies that the update $V_{m'} \leftarrow V_{m'} + (v_k - v_j)$, i.e., swapping product $j \in \mathcal{S}_{\vee}^C \setminus S_m^C$ with k, decreases the term in summation in (A6.10). Hence, the objective is improved, which is a contradiction to the optimality of $\{S_1^C, .., S_M^C\}$.

Proof of Proposition 4.3. We will prove the result for the common set and union set in part 1 and part 2, respectively.

Part 1: If transshipments are not allowed in the system, each store is going to solve their own problem (which is the same for all stores) independently. In this case, each store has the assortment $1, .., k$ where k is the maximum indexed product such that $\dfrac{v_k}{\left(\sum\limits_{i=0}^{k-1} v_i\right)\left(\sum\limits_{i=0}^{k} v_i\right)} \geq K$. (Note that the left hand side of this inequality is the increment in store's own demand and the right hand side is the fixed cost of adding a product.) This implies that the optimal common set $\mathcal{S}_{\wedge}^C = \{1, .., k\}$ when transshipments are not allowed in the system. Our aim is to show that the common set includes at most the first k products when transshipments are allowed in the system, which concludes the result.

Suppose we are given assortments S_1 and S_2. Let $V_1 = \sum_{j \in S_1 \setminus S_2} v_j, V_2 = \sum_{j \in S_2 \setminus S_1} v_j$, and $V_{12} = \sum_{j \in S_1 \cap S_2} v_j$. Assume w.l.o.g. that $V_1 \geq V_2$. In addition, define

$$\pi(V_1, V_2, V_{12}) = \begin{cases} \dfrac{V_{12} + V_1 + \theta V_2(1-\tau)}{1 + V_{12} + V_1 + \theta V_2} + \dfrac{V_{12} + V_2 + \theta V_1(1-\tau)}{1 + V_{12} + V_2 + \theta V_1} \text{ if } \tau < \dfrac{1}{1 + V_{12} + V_1} \\ \\ \dfrac{V_{12} + V_1}{1 + V_{12} + V_1} + \dfrac{V_{12} + V_2 + \theta V_1(1-\tau)}{1 + V_{12} + V_2 + \theta V_1} \text{ if } \dfrac{1}{1 + V_{12} + V_1} < \tau < \dfrac{1}{1 + V_{12} + V_2} \\ \\ \dfrac{V_{12} + V_1}{1 + V_{12} + V_1} + \dfrac{V_{12} + V_2}{1 + V_{12} + V_2} \text{ if } \dfrac{1}{1 + V_{12} + V_2} < \tau \end{cases}$$

In words, π is the objective value of the CP for some (S_1, S_2) with no fixed cost of including a product to an assortment $(K = 0)$. For ease of notation, we denote the above three pieces of π by π_1, π_2, π_3, respectively.

The rest of the proof follows two main steps. First, we show that all π_1, π_2, π_3 are concave increasing in V_{12} for any given V_1, V_2. This implies for given V_1, V_2, the optimal decision is to keep adding common products starting from the most popular, and stop at some product when the incremental profit obtained by adding that product (as a common product) is less than K. Second, we show that $\dfrac{\partial^2 \pi_1}{\partial V_{12} \partial V_1} \leq 0, \dfrac{\partial^2 \pi_2}{\partial V_{12} \partial V_1} \leq 0, \dfrac{\partial^2 \pi_3}{\partial V_{12} \partial V_1} \leq 0$. These imply that setting $V_1 = 0$ maximizes the incremental profit obtained by adding a common product. (Similar arguments can be shown for $\dfrac{\partial^2 \pi_1}{\partial V_{12} \partial V_2}, \dfrac{\partial^2 \pi_2}{\partial V_{12} \partial V_2}, \dfrac{\partial^2 \pi_3}{\partial V_{12} \partial V_2}$). Hence, we set $V_1 = 0, V_2 = 0$, and add common products one by one starting from the most popular and stop when the incremental profit is less than K. This procedure yields the maximum number of common products an optimal solution may include.

In order to write the incremental profit of adding product i as a common product, we should be careful about the range of the π function. For instance, for given V_1, V_2, if adding v_i leads to switching from range 1 to range 2 at $V_{12} = V_{12} + \tilde{v}$, we should write the incremental profit as follows:

$$\int_{V_{12}}^{V_{12}+\tilde{v}} \frac{\partial \pi_1(V_1, V_2, v)}{\partial V_{12}} dv + \int_{V_{12}+\tilde{v}}^{V_{12}+v_i} \frac{\partial \pi_2(V_1, V_2, v)}{\partial V_{12}} dv \tag{A6.20}$$

Note that setting $V_1 = 0, V_2 = 0$ pointwise maximizes both integrant in the above expression because $\frac{\partial^2 \pi_1}{\partial V_{12} \partial V_1} \leq 0, \frac{\partial^2 \pi_2}{\partial V_{12} \partial V_1} \leq 0$, and $\frac{\partial^2 \pi_1}{\partial V_{12} \partial V_2} \leq 0, \frac{\partial^2 \pi_2}{\partial V_{12} \partial V_2} \leq 0$.

Part 2: We are going to show that any optimal solution of the CP must include product k (as deined in Part 1) in one of the assortments by assuming the otherwise and showing a contradiction to the optimality of the solution.

Suppose (S_1, S_2) is an optimal solution of the CP and define (V_1, V_2, V_{12}) as in Part 1. We can assume w.l.o.g. that $V_1 \geq V_2$. We can decompose the objective value (except the fixed cost K) into 4 components for this assortment configuration:

$$\text{Store 1's own demand: } \pi_1^{own}(V_1, V_2, V_{12}) \tag{A6.21}$$

$$\text{Store 1 received: } \pi_1^{rcv}(V_1, V_2, V_{12}) \tag{A6.22}$$

$$\text{Store 2's own demand: } \pi_2^{own}(V_1, V_2, V_{12}) \tag{A6.23}$$

$$\text{Store 2 received: } \pi_2^{rcv}(V_1, V_2, V_{12}) \tag{A6.24}$$

These components can be expressed as follows:

$$\pi(V_1, V_2, V_{12}) = \begin{cases} \overbrace{\frac{V_{12} + V_1}{1 + V_{12} + V_1 + \theta V_2}}^{\pi_1^{own}(V_1, V_2, V_{12})} + \overbrace{\frac{\theta V_2 (1-\tau)}{1 + V_{12} + V_1 + \theta V_2}}^{\pi_1^{rcv}(V_1, V_2, V_{12})} + \overbrace{\frac{V_{12} + V_2}{1 + V_{12} + V_2 + \theta V_1}}^{\pi_2^{own}(V_1, V_2, V_{12})} + \overbrace{\frac{\theta V_1 (1-\tau)}{1 + V_{12} + V_2 + \theta V_1}}^{\pi_2^{rcv}(V_1, V_2, V_{12})} & \text{if } \tau < \frac{1}{1 + V_{12} + V_1} \\ \overbrace{\frac{V_{12} + V_1}{1 + V_{12} + V_1}}^{\pi_1^{own}(V_1, V_2, V_{12})} + \overbrace{\frac{V_{12} + V_2}{1 + V_{12} + V_2 + \theta V_1}}^{\pi_2^{own}(V_1, V_2, V_{12})} + \overbrace{\frac{\theta V_1 (1-\tau)}{1 + V_{12} + V_2 + \theta V_1}}^{\pi_2^{rcv}(V_1, V_2, V_{12})} & \text{if } \frac{1}{1 + V_{12} + V_1} < \tau < \frac{1}{1 + V_{12} + V_2} \\ \overbrace{\frac{V_{12} + V_1}{1 + V_{12} + V_1}}^{\pi_1^{own}(V_1, V_2, V_{12})} + \overbrace{\frac{V_{12} + V_2}{1 + V_{12} + V_2}}^{\pi_2^{own}(V_1, V_2, V_{12})} & \text{if } \frac{1}{1 + V_{12} + V_2} < \tau \end{cases}$$

Assume to the contrary that the optimal solution under consideration does not include product k. We are going to show that adding product k in assortment 2 improves the objective value, which contradicts the optimality of the solution.

Assume without loss of generality that $S_1 \cup S_2 = \{1, .., k-1\}$ (The union set has to be a popular subset. If it is smaller that $\{1,..,k-1\}$, we can consider the largest index product in $S_1 \cup S_2$ as product k and all our arguments follow accordingly.) If we add product k in S_2, it leads to the following four changes:

$$\Delta_1^{own} = \pi_1^{own}(V_1, V_2 + v_k, V_{12}) - \pi_1^{own}(V_1, V_2, V_{12}) \tag{A6.25}$$

$$\Delta_1^{rcv} = \pi_1^{rcv}(V_1, V_2 + v_k, V_{12}) - \pi_1^{rcv}(V_1, V_2, V_{12}) \tag{A6.26}$$

$$\Delta_2^{own} = \pi_2^{own}(V_1, V_2 + v_k, V_{12}) - \pi_2^{own}(V_1, V_2, V_{12}) \tag{A6.27}$$

$$\Delta_2^{rcv} = \pi_2^{rcv}(V_1, V_2 + v_k, V_{12}) - \pi_2^{rcv}(V_1, V_2, V_{12}) \tag{A6.28}$$

We will show that (1) $\Delta_1^{own} + \Delta_1^{rcv} \geq 0$, and (2) $\Delta_2^{own} + \Delta_2^{rcv} \geq K$. This suffices because the sum $(\Delta_1^{own} + \Delta_1^{rcv} + \Delta_2^{own} + \Delta_2^{rcv}) \geq K$ implies that adding product k improves the objective value, which contradicts the optimality of (S_1, S_2).

First consider claim (1). If $\dfrac{1}{1+V_{12}+V_1} < \tau$, $\Delta_1^{own} + \Delta_1^{rcv} = 0$ because store 1 does not receive transshipments. If $\dfrac{1}{1+V_{12}+V_1} \geq \tau$,

$$\frac{\partial\left(\pi_1^{own}(V_1, V_2 + v_k, V_{12}) + \pi_1^{rcv}(V_1, V_2 + v_k, V_{12})\right)}{\partial V_2} = \frac{\theta(1-\tau(1+V_1+V_{12}))}{(1+V_1+V_{12}+\theta V_2)^2} \geq 0 \tag{A6.29}$$

Hence, adding product k to S_2, i.e., increasing V_2 by v_k, implies $\Delta_1^{own} + \Delta_1^{rcv} \geq 0$.

Now consider claim (2). If $\dfrac{1}{1+V_{12}+V_2} < \tau$, $\Delta_2^{rcv} = 0$ because store 2 does not receive transshipments, and $\Delta_2^{own} \geq K$ because $V_{12}+V_2 \leq \sum_{i=1}^{k-1} v_i$ and adding product k is profitable by definition of product k even if $V_{12}+V_2 = \sum_{i=1}^{k-1} v_i$. If $\dfrac{1}{1+V_{12}+V_2} \geq \tau$, store 2 receives transshipment before adding product k, but stops doing so after adding product k if $\dfrac{1}{1+V_{12}+V_2+v_k} < \tau$. Therefore, we have to consider different ranges of the function $\pi_2(\cdot)$ seperately. Assuming $\dfrac{1}{1+V_{12}+V_2+\tilde{v}} = \tau$

,we have

$$\Delta_2^{own} + \Delta_2^{rcv} = \int_{V_2}^{V_2+\tilde{v}} \frac{\partial(\pi_2^{own}(V_1, v, V_{12}) + \pi_2^{rcv}(V_1, v, V_{12}))}{\partial v} dv \tag{A6.30}$$

$$+ \int_{V_2+\tilde{v}}^{V_2+v_k} \frac{\partial(\pi_2^{own}(V_1, v, V_{12}))}{\partial v} dv \tag{A6.31}$$

$$= \int_{V_2}^{V_2+\tilde{v}} \frac{\partial\left(\frac{V_{12}+v}{1+V_{12}+v+\theta V_1} + \frac{\theta V_1(1-\tau)}{1+V_{12}+v+\theta V_1}\right)}{\partial v} dv \tag{A6.32}$$

$$+ \int_{V_2+\tilde{v}}^{V_2+v_k} \frac{\partial\left(\frac{V_{12}+v}{1+V_{12}+v+\theta V_1}\right)}{\partial v} dv \tag{A6.33}$$

$$= \int_{0}^{\tilde{v}} \frac{\partial\left(\frac{V_{12}+V_2+v}{1+V_{12}+V_2+v+\theta V_1} + \frac{\theta V_1(1-\tau)}{1+V_{12}+V_2+v+\theta V_1}\right)}{\partial v} dv \tag{A6.34}$$

$$+ \int_{\tilde{v}}^{v_k} \frac{\partial\left(\frac{V_{12}+V_2+v}{1+V_{12}+V_2+v}\right)}{\partial v} dv \tag{A6.35}$$

$$\geq \int_{0}^{\tilde{v}} \frac{\partial\left(\frac{V_{12}+V_2+v+\theta v}{1+V_{12}+V_2+v+\theta v}\right)}{\partial v} dv + \int_{\tilde{v}}^{v_k} \frac{\partial\left(\frac{V_{12}+V_2+v}{1+V_{12}+V_2+v}\right)}{\partial v} dv \tag{A6.36}$$

$$\geq \int_{0}^{v_k} \frac{\partial\left(\frac{\sum_{i=1}^{k-1} v_i + v}{1+\sum_{i=1}^{k-1} v_i + v}\right)}{\partial v} dv \geq K \tag{A6.37}$$

Inequality (A6.36) holds because

$$\frac{\partial\left(\frac{V_{12}+V_2+v}{1+V_{12}+V_2+v+\theta V_1} + \frac{\theta V_1(1-\tau)}{1+V_{12}+V_2+v+\theta V_1}\right)}{\partial v} = \frac{1+V_1\tau\theta}{(1+V_{12}+V_2+v+\theta V_1)^2} \tag{A6.38}$$

$$\geq \frac{1}{(1+V_{12}+V_2+v+\theta V_1)^2} = \frac{\partial\left(\frac{V_{12}+V_2+v+\theta v}{1+V_{12}+V_2+v+\theta v}\right)}{\partial v} \tag{A6.39}$$

Inequality (A6.37) holds because $V_{12} + V_2 + \theta V_1 \leq \sum_{i=1}^{k-1} v_i$.

Proof of Lemma 4.2. ($\Rightarrow$) Assume that $X^*_{j,m} = 1, \forall j \in S^*_{m'} \setminus S^*_m$. The profit obtained from store

m in the equilibrium is:

$$\frac{\sum\limits_{j\in S_m} v_j + \sum\limits_{j\in\mathcal{S}_\vee\setminus S_m} \theta v_j(1-\tau-\alpha)}{1+\sum\limits_{k\in S_m} v_k + \sum\limits_{k\in\mathcal{S}_\vee\setminus S_m} \theta v_k} + \alpha \frac{\sum_{j\in S_m\setminus S_{m'}} \theta v_j X^*_{j,m'}}{1+\sum\limits_{k\in S_{m'}} v_k + \sum\limits_{k\in S_m\setminus S_{m'}} \theta v_k X^*_{k,m'}} - K\cdot|S_m| \tag{A6.40}$$

Notice that the second and third terms in the above expression do not depend on either *tau* or $X^*_{j,m}$, $\forall j\in S^*_{m'}\setminus S^*_m$, and thus irrelevant to our discussion here. The first term can be written as a convex combination of $(1-\tau-\alpha)$ and $\left(\frac{\sum\limits_{k\in S_m} v_k}{1+\sum\limits_{k\in S_m} v_k}\right)$ as follows:

$$\left(\frac{\sum\limits_{k\in\mathcal{S}_\vee\setminus S_m}\theta v_k}{1+\sum\limits_{k\in S_m} v_k + \sum\limits_{k\in\mathcal{S}_\vee\setminus S_m}\theta v_k}\right)(1-\tau-\alpha) + \left(\frac{1+\sum\limits_{k\in S_m} v_k}{1+\sum\limits_{k\in S_m} v_k + \sum\limits_{k\in\mathcal{S}_\vee\setminus S_m}\theta v_k}\right)\left(\frac{\sum\limits_{k\in S_m} v_k}{1+\sum\limits_{k\in S_m} v_k}\right) \tag{A6.41}$$

By the assumption, the best response of store m maximizes the weight of $(1-\tau-\alpha)$ by setting $X^*_{j,m}=1$, $\forall j\in\mathcal{S}^*_\vee\setminus S^*_m$. This implies $1-\tau-\alpha > \frac{\sum\limits_{k\in S_m} v_k}{1+\sum\limits_{k\in S_m} v_k}$. Therefore, $\alpha+\tau \leq \frac{1}{1+\sum\limits_{k\in S^*_m} v_k}$.

($\Leftarrow$) Similar to the above argument, we can show that if $\alpha+\tau\leq\frac{1}{1+\sum\limits_{k\in S^*_m} v_k}$, it is optimal for store m to maximize the weight of $(1-\tau-\alpha)$ in the above convex combination by setting $X^*_{j,m}=1$, $\forall j\in S^*_{m'}\setminus S^*_m$.

Proof of Lemma 4.3.1. Pick $j,k\in S^*_{m'}$ s.t. $v_k>v_j$. Assume to the contrary that $j\in S^*_m$ and $k\notin S^*_m$. We will show that store m is better-off by replacing product j with k.

Define $V_m=\sum\limits_{i\in S^*_m} v_i$, and $V_{m'}=\sum\limits_{j\in S^*_{m'}\setminus S_m} v_j$.

If $(\alpha+\tau)>\frac{1}{1+V_m}$, store m does not transship any product from the other store ($X_{j,m}=0,\forall j\in N$), and replacing product j with k does not change this situation. The own-sales of store m increases because $v_k>v_j$, which implies that store m is better-off by replacing product j with k.

If $(\alpha+\tau)<\frac{1}{1+V_m}$, store m is transshipping product k from the other store. Replacing product j with k will increase store m's own-sales and decrease the amount transshipped from store m'.

Even if store m has to transship product j after replacement, i.e., $(\alpha+\tau) < \dfrac{1}{1+V_m+v_k-v_j}$, we have

$$\frac{A}{B} < \frac{A+(v_k-v_j)(1-\theta(1-\alpha-\tau))}{B+(v_k-v_j)(1-\theta)} \tag{A6.42}$$

where A, B denote the numerator and the denominator of the first term in (A6.40), respectively. This implies that store m is better-off by replacing product j with k.

Proof of Lemma 4.3.2. Pick $j, k \in (\mathcal{N} \setminus S_{m'})$ s.t. $v_k > v_j$. Assume to the contrary that $j \in S_m^*$ and $k \notin S_m^*$. Store m's own-sales (the first term in (A6.40)) increases after replacing product j with k. If store m' is transshipping product j, i.e., $\alpha+\tau \leq \dfrac{1}{1+\sum\limits_{k\in S_m^*} v_k}$, it will transship product k, too, which increases store m's profit obtained by transshipping to store m' (the second term in (A6.40)). If store m' is not transshipping product j, i.e., $\alpha+\tau \geq \dfrac{1}{1+\sum\limits_{k\in S_m^*} v_k}$, then there is no transshipment from store m to m' before or after replacing product j with k in S_m^*.

Proof of Proposition 4.4. Follows directly from Lemma 4.3.1. and Lemma 4.3.2.

Proof of Proposition 4.5. We will consider two cases: 1)$|S_m^C| = |S_{m'}^C|$ and 2) $|S_m^C| = |S_{m'}^C| + 1$. For the two store case with equal popularities, there is no other assortment configuration that can be the optimal solution of the cetralized problem, i.e., the assortment depths should be set as evenly as possible.

<u>Case 1.1:</u> $|S_m^C| = |S_{m'}^C| = b$ and $\tau \geq \dfrac{1}{1+bv}$.

By Lemma 4.1, $X_{j,1}^* = X_{j,2}^* = 0, \forall j \in \mathcal{N}$, i.e., there are no transshipments in the centralized solution. This means the profit obtained from a store is independent from the assortment in the other store. Hence, each store maximizes the profit (without transshipments) by adding the most popular $k = \underset{i\in\mathcal{N}}{\operatorname{argmax}}\left\{\dfrac{v_i}{\mathcal{V}_i(1-\mathcal{V}_i)} \geq K\right\}$ products where $\mathcal{V}_i = \sum\limits_{j=0}^{i} v_j$, because adding the $(k+1)$st product is not profitable. Then, $S_1^* = S_2^* = \{1, 2, .., k\}$.

We would like to show that $(X_{j,1}^* = X_{j,2}^* = 0, \forall j \in \mathcal{N}, S_m^C, S_{m'}^C)$ is an equilibrium for any α. To that end we argue that the best response to S_m^C is $S_{m'}^C$ and vice versa. Neither store 1 nor store 2 is better-off by excluding any of the products from the set $S_m^C = S_{m'}^C = \{1, 2, .., k\}$ because relying on the other store for a product via transshipment is not profitable by Lemma 4.2. Moreover, adding the $(k+1)st$ product is not profitable by definition of k, and because there is no transshipment to the other store.

Case 1.2: $|S_m^C| = |S_{m'}^C| = b$ and $\tau < \dfrac{1}{1+bv}$.

In this case, we want to show the existence of an $\alpha \in \left[0, \dfrac{1}{1+bv} - \tau\right]$ such that $\{S_m^C, S_{m'}^C\}$ is a pure strategy Nash equilibrium. Note that in this equilibrium, equal number of products are offered in each store. Hence, we can interpret the equilibrium as a symmetric one in terms of the number of products offered.

We will fix the number of products offered in store m' to b and search for an α which makes store m best respond by offering b products as well. To that end, we will show the following three arguments, which together imply our result:

1. For $\alpha = 0$, store m's best response to the assortment $S_{m'}^C$ includes $a \leq b$ products.
2. For $\alpha = \dfrac{1}{1+bv} - \tau$, store m's best response to the assortment $S_{m'}^C$ includes $c \geq b$ products.
3. $\dfrac{\pi_b(0) - \pi_x(0)}{\partial \pi_x(\alpha)/\partial \alpha}$ is increasing in x.

Note that if $a = b$ or $c = b$, we are done. Hence, we will now consider the case when $a < b < c$. We can write store m's profit when store m and m' offer x and b products respectively as follows:

$$\frac{xv + \theta bv(1-\alpha-\tau)}{1+xv+\theta bv} + \frac{\alpha\theta xv}{1+bv+\theta xv} - xK \tag{A6.43}$$

1. For $\alpha = 0$, the above profit expression is equal to $\dfrac{xv + \theta bv(1-\tau)}{1+xv+\theta bv} - xK$. Assume to the contrary that store m's best response is offering $a > b$ products. In that case, store m' makes more profit than it does in the solution $\{S_m^C, S_{m'}^C\}$. Moreover, since the total profit $(\pi^m + \pi^{m'})$ is maximized with the solution $\{S_m^C, S_{m'}^C\}$, offering $a > b$ products is less profitable than offering b products.

2. For $\alpha = \dfrac{1}{1+bv} - \tau$, store m' makes the following profit:

$$\frac{bv + \theta xv\left(1 - \dfrac{1}{1+bv}\right)}{1 + bv + \theta xv} + \frac{\left(\dfrac{1}{1+bv} - \tau\right)\theta bv}{1 + xv + \theta bv} - bK \tag{A6.44}$$

$$= \frac{bv}{1+bv} + \frac{\left(\dfrac{1}{1+bv} - \tau\right)\theta bv}{1 + xv + \theta bv} - bK \tag{A6.45}$$

Assume to the contrary that store m's best response is offering $c < b$ products. In that case, store m' makes more profit than it does in the solution $\{S_m^C, S_{m'}^C\}$ because the denominator of the second term is smaller when $x = c < b$. Moreover, since the total profit $(\pi^m + \pi^{m'})$ is maximized with the solution $\{S_m^C, S_{m'}^C\}$, offering $c < b$ products is less profitable than offering b products.

3. Note that $\pi_b^m(\alpha) = \pi_b^m(0)$ for any $\alpha \in [0,1]$. Note also that $\alpha = \dfrac{\pi_b(0) - \pi_x(0)}{\partial \pi_x(\alpha)/\partial \alpha}$ is the α value at which store m is indifferent between offering b or x products. We will show that this expression is strictly increasing in x, which implies that the line $\pi_a(\alpha)$ intersects with the horizontal line $\pi_b(\alpha) = \pi_b^m(0)$ at a smaller α value than $\pi_c(\alpha)$ intersects with it (because $c > a$).

$$\begin{aligned}
&\partial\left(\frac{\pi_b(0) - \pi_x(0)}{\partial \pi_x(\alpha)/\partial \alpha}\right)\Big/\partial x \\
&= \big[(Kb^3\theta^3 + 2Kb^2\theta^2 x + Kb\theta^2 x^2 + 2Kb^2\theta x + Kb\theta x^2 + Kb^3 + (b^2\tau\theta)(1-\theta))v^3 \\
&+ (Kb^2\theta^3 + 2Kb^2\theta^2 + 2Kb\theta^2 x + 4Kb\theta x + K\theta x^2 + 2Kb^2 + (b\tau\theta + b)(1-\theta))v^2 \\
&+ (2Kb\theta^2 + Kb\theta + 2K\theta x + Kb - \theta + 1)v + K\theta\big] \\
&\Big/\big[\theta(1 + (b+x)v)^2(1 + (b\theta + b)v)\big]
\end{aligned}$$

All the terms in the above expression are positive. This means $\pi_x(\alpha)$ intersects with $\pi_b(\alpha)$ at a smaller α value than $\pi_{x'}(\alpha)$ intersects with it for some $x < x'$.

<u>Case 2.1:</u> $|S_m^C| - 1 = |S_{m'}^C| = b$ and $\tau < \dfrac{1}{1+(b+1)v}$.

We will fix the number of products offered in store m' to b and search for an α which makes store m best respond by offering $b+1$ products. In the mean time, we want to make sure store m' best responds to $b+1$ products by b products. To that end, we will show the following three arguments, which together imply our result:

1. For $\alpha = 0$, store m's best response to b products includes $a_m \leq b+1$ products. Similarly, best response of store m' to b products includes $a_{m'} \leq b$ products.
2. For $\alpha = \dfrac{1}{1+(b+1)v} - \tau$, store m's best response to b products includes $c_m \geq b+1$ products. Similarly, best response of store m' to b products includes $c_{m'} \geq b$ products.
3. $\dfrac{\pi_{(b+1)\to b}(0) - \pi_{(b+1)\to(b+1)}(0)}{\partial \pi_{(b+1)\to b}(\alpha)/\partial\alpha} > \dfrac{\pi_{b\to b}(0) - \pi_{b\to(b+1)}(0)}{\partial \pi_{b\to(b+1)}(\alpha)/\partial\alpha}$.

$$\frac{\pi_{(b+1)\to b}(0) - \pi_{(b+1)\to(b+1)}(0)}{\partial \pi_{(b+1)\to b}(\alpha)/\partial\alpha} = \frac{\dfrac{bv+(b+1)\theta v(1-\tau)}{(b+1)\theta v + bv + 1} - bK - \dfrac{(b+1)v + (b+1)\theta v(1-\tau)}{(b+1)\theta v + (b+1)v + 1} + (b+1)K}{\dfrac{\theta(b+1)v}{(b+1)\theta v + bv + 1} - \dfrac{b\theta v}{b\theta v + (b+1)v + 1)}}$$

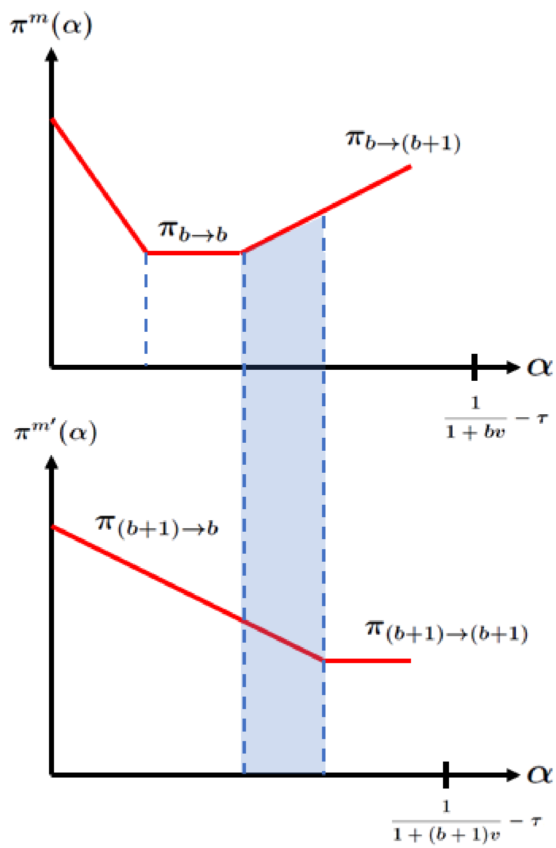
$$\frac{\pi_{b\to b}(0) - \pi_{b\to(b+1)}(0)}{\partial \pi_{b\to(b+1)}(\alpha)/\partial\alpha} = \frac{\dfrac{bv+b\theta v(1-\tau)}{b\theta v + bv + 1} - bK - \dfrac{(b+1)v + b\theta v(1-\tau)}{b\theta v + (b+1)v + 1} + (b+1)K}{\dfrac{\theta(b+1)v}{(b+1)\theta v + bv + 1} - \dfrac{b\theta v}{b\theta v + (b+1)v + 1)}}$$

The difference $\dfrac{\pi_{(b+1)\to b}(0) - \pi_{(b+1)\to(b+1)}(0)}{\partial \pi_{(b+1)\to b}(\alpha)/\partial\alpha} - \dfrac{\pi_{b\to b}(0) - \pi_{b\to(b+1)}(0)}{\partial \pi_{b\to(b+1)}(\alpha)/\partial\alpha}$ is:

$$\frac{v^2\theta\overbrace{\left(b^2\tau\theta^2v^2 + b\tau\theta^2v^2 - b^2\tau v^2 - b\tau v^2 - 2b\tau v + 2b\theta v + 2bv - \tau v + \theta v - \tau + v + 2\right)}^{\phi}}{(b\theta v + bv + 1)(b\theta v + bv + v + 1)(b\theta v + bv + \theta v + 1)(b\theta v + bv + \theta v + v + 1)}$$

We would like to show that $\phi > 0$.

$$\begin{aligned}\phi &= (b^2\theta^2v^2 + b\theta^2v^2 - b^2v^2 - bv^2 - 2bv - v - 1)\tau + \theta v + 2b\theta v + 2bv + v + 2\\ &= \left[(b+1)bv^2(\theta^2-1) - 2bv - v - 1\right]\tau + \theta v + 2b\theta v + 2bv + v + 2\\ &> \left[(b+1)bv^2(\theta^2-1) - bv\right]\tau + \theta v + 2b\theta v + 2bv + v + 1\\ &> -bv\left[(b+1)v + 1\right]\tau + \theta v + 2b\theta v + 2bv + v + 1\\ &> \theta v + 2b\theta v + bv + v + 1 > 0.\end{aligned}$$



<u>Case 2.2:</u> $|S_m^C| - 1 = |S_{m'}^C| = b$ and $\dfrac{1}{1+(b+1)v} < \tau < \dfrac{1}{1+bv}$.

The difference $\dfrac{\pi_{(b+1)\to b}(0) - \pi_{(b+1)\to(b+1)}(0)}{\partial\pi_{(b+1)\to b}(\alpha)/\partial\alpha} - \dfrac{\pi_{b\to b}(0) - \pi_{b\to(b+1)}(0)}{\partial\pi_{b\to(b+1)}(\alpha)/\partial\alpha}$ is:

$$\frac{v\overbrace{\left(-b^2\tau v^2 - b\tau v^2 - 2b\tau v + b\theta v + 2bv - \tau v + \theta v - \tau + v + 2\right)}^{\phi}}{(b\theta v + bv + \theta v + v + 1)(bv+1)(bv+v+1)}$$

We would like to show that $\phi > 0$.

$$
\begin{aligned}
\phi &= (-b^2v^2 - bv^2 - 2bv - v - 1)\tau + b\theta v + 2bv + \theta v + v + 2 \\
&> (-b^2v^2 - bv^2 - bv)\tau + b\theta v + 2bv + \theta v + v + 1 \\
&= -bv((b+1)v + 1)\tau + b\theta v + 2bv + \theta v + v + 1 \\
&> -((b+1)v + 1) + b\theta v + 2bv + \theta v + v + 1 = b\theta v + bv + \theta v > 0.
\end{aligned}
$$

This concludes the result.

Proof of Proposition 4.6. Pick two stores: store 1 and store 2. Suppose we are given the sets $(S_1 \cap S_2)$ and $(S_1 \cup S_2)$. Let $A = (S_1 \cup S_2) \setminus (S_1 \cap S_2)$. Now, consider the problem of allocating the products in A to store 1 and store 2 such that the centralized profit is maximized. Note that this allocation affects only the profit obtained from these two stores as long as $(S_1 \cup S_2)$ is fixed. Moreover, we are going to ignore the products in $\mathcal{S}_\vee \setminus (S_1 \cup S_2)$ to ease exposition, but adding those products into account doesn't change the following arguments.

Define $V_A = \sum_{j \in A} v_j$. Similar to before, define $V_1 = \sum_{j \in S_1 \setminus S_2} v_j$, $V_2 = \sum_{j \in S_2 \setminus S_1} v_j$, and $V_{1,2} = \sum_{j \in S_1 \cap S_2} v_j$. Note that $V_{1,2}$ and the sum of $V_1 + V_2 = V_A$ are fixed. The question is how to allocate V_A into V_1 and V_2 such that the centralized profit is maximized. Let $B = \{\beta : \beta = V_1/V_A\}$ be the discrete and finite ($|B| = 2^{|A|}$) set of feasible allocations, implying that $V_1 = V_A - V_2 = \beta V_A$. The profit obtained from store 1 and store 2 can be written as follows:

$$
\pi(\beta) = \begin{cases}
\dfrac{V_{1,2} + \theta V_A + (1-\theta)\beta V_A}{1 + V_{1,2} + \theta V_A + (1-\theta)\beta V_A} + \dfrac{V_{1,2} + \theta V_A + (1-\theta)(1-\beta) V_A}{1 + V_{1,2} + \theta V_A + (1-\theta)(1-\beta) V_A} \text{ if } \left(\tau < \dfrac{1}{1 + V_{12} + (1-\beta)V_A}\right) \\[2em]
\dfrac{V_{1,2} + \theta V_A + (1-\theta)\beta V_A}{1 + V_{1,2} + \theta V_A + (1-\theta)\beta V_A} + \dfrac{V_{1,2} + (1-\beta) V_A}{1 + V_{1,2} + (1-\beta) V_A} \text{ if } \left(\dfrac{1}{1 + V_{12} + (1-\beta)V_A} < \tau < \dfrac{1}{1 + V_{12} + \beta V_A}\right) \\[2em]
\dfrac{V_{1,2} + \beta V_A}{1 + V_{1,2} + \beta V_A} + \dfrac{V_{1,2} + (1-\beta) V_A}{1 + V_{1,2} + (1-\beta) V_A} \text{ if } \left(\dfrac{1}{1 + V_{12} + \beta V_A} < \tau\right)
\end{cases}
$$

APPENDIX 7: THRESHOLD VALUES OF K IN SECTION 4.6.3

$$K_1 = \frac{\theta\tau + \dfrac{1-\theta}{2v+1}}{(1+v+\theta v)/v}, \qquad K_2 = \frac{1+v+\theta(1-\tau)(1+\theta v)+\theta v^2\tau(1-\theta)}{(1+v)(1+\theta v)(1+v+\theta v)/v},$$

$$K_3 = \frac{\theta(1-\tau) + \dfrac{\theta v+1}{v+1}}{(1+\theta v)/v}, \qquad K_4 = \frac{\theta\tau + \dfrac{1-\theta}{v+1}}{(1+\theta v)/v}$$

$$\bar{K}_1 = \frac{\theta(\alpha+\tau) + \dfrac{1-\theta}{2v+1}}{(1+v+\theta v)/v}, \qquad \bar{K}_2 = \frac{\theta\alpha + \dfrac{\theta v(\alpha+\tau)+1}{\theta v+1}}{(1+v+\theta v)/v},$$

$$\bar{K}_3 = \frac{\theta\alpha + \dfrac{\theta v+1}{v+1}}{(1+\theta v)/v}, \qquad \bar{K}_4 = \frac{\theta(\alpha+\tau) + \dfrac{1-\theta}{v+1}}{(1+\theta v)/v}$$

Table A7.2: Comparison of the union set, common set, and the average number of accessible products in the centralized and decentralzied systems for $|N| = 10$ products.

				Union Set			Common Set			# of Accessible Prod.			
				$\mathcal{S}_\vee^C \supset \mathcal{S}_\vee^D$	$\mathcal{S}_\vee^C \subset \mathcal{S}_\vee^D$	$\mathcal{S}_\vee^C = \mathcal{S}_\vee^D$	$\mathcal{S}_\wedge^C \supset \mathcal{S}_\wedge^D$	$\mathcal{S}_\wedge^C \subset \mathcal{S}_\wedge^D$	$\mathcal{S}_\wedge^C = \mathcal{S}_\wedge^D$	Cent.	Decent.	Equal	No Equ.
K_L	θ_L	τ_L	α_L	33	0	67	0	42	58	31	13	56	0
			α_H	79	0	21	0	82	18	66	25	9	0
		τ_H	α_L	9	0	91	0	12	88	9	5	86	0
			α_H	9	0	91	0	12	88	9	5	86	0
	θ_H	τ_L	α_L	28	1	71	0	74	26	33	45	22	0
			α_H	97	0	3	0	100	0	97	3	0	0
		τ_H	α_L	26	0	74	0	30	70	26	8	66	0
			α_H	26	0	74	0	30	70	26	8	66	0
K_H	θ_L	τ_L	α_L	45	0	55	0	44	56	43	3	54	0
			α_H	81	1	18	0	83	17	75	12	13	0
		τ_H	α_L	30	0	70	0	36	64	26	13	61	0
			α_H	34	0	66	0	37	63	30	12	58	0
	θ_H	τ_L	α_L	21	0	78	0	39	60	21	27	51	1
			α_H	86	1	13	0	97	3	94	5	1	0
		τ_H	α_L	57	2	41	0	63	37	59	8	33	0
			α_H	63	0	37	0	66	34	63	6	31	0

Table A7.3: Comparison of the union set, common set, and the average number of accessible products in the centralized and decentralzied systems for $|N| = 10$ equal-popularity products.

				Union Set			Common Set			# of Accessible Prod.		
				$\mathcal{S}_\vee^C \supset \mathcal{S}_\vee^D$	$\mathcal{S}_\vee^C \subset \mathcal{S}_\vee^D$	$\mathcal{S}_\vee^C = \mathcal{S}_\vee^D$	$\mathcal{S}_\wedge^C \supset \mathcal{S}_\wedge^D$	$\mathcal{S}_\wedge^C \subset \mathcal{S}_\wedge^D$	$\mathcal{S}_\wedge^C = \mathcal{S}_\wedge^D$	Cent.	Decent.	Equal
K_L	θ_L	τ_L	α_L	53	1	46	0	21	79	49	5	46
			α_H	63	16	21	0	60	40	56	23	21
		τ_H	α_L	60	3	37	0	38	62	49	3	48
			α_H	65	0	35	0	65	35	47	0	53
	θ_H	τ_L	α_L	75	5	20	0	17	83	79	5	16
			α_H	76	0	24	0	52	48	91	0	9
		τ_H	α_L	41	1	58	0	30	70	49	1	50
			α_H	62	8	30	0	59	41	69	0	31
K_H	θ_L	τ_L	α_L	23	0	77	0	15	85	18	5	77
			α_H	46	0	54	0	43	57	38	8	54
		τ_H	α_L	30	0	70	0	25	75	26	4	70
			α_H	54	6	40	0	52	48	50	10	40
	θ_H	τ_L	α_L	78	2	20	0	10	90	83	5	12
			α_H	35	9	56	0	30	70	57	12	31
		τ_H	α_L	44	5	51	0	12	88	52	5	43
			α_H	54	0	46	0	50	50	63	0	37

Table A7.4: Comparison of the union set, common set, and the average number of accessible products in the centralized and decentralized systems for $|N| = 10$ products when the transshipments are enforced.

				Union Set			Common Set			# of Accessible Prod.			
				$\mathcal{S}_\vee^C \supset \mathcal{S}_\vee^D$	$\mathcal{S}_\vee^C \subset \mathcal{S}_\vee^D$	$\mathcal{S}_\vee^C = \mathcal{S}_\vee^D$	$\mathcal{S}_\wedge^C \supset \mathcal{S}_\wedge^D$	$\mathcal{S}_\wedge^C \subset \mathcal{S}_\wedge^D$	$\mathcal{S}_\wedge^C = \mathcal{S}_\wedge^D$	Cent.	Decent.	Equal	No Equ.
K_L	θ_L	τ_L	α_L	29	0	71	0	42	58	28	14	58	0
			α_H	23	8	69	0	71	29	14	65	21	0
		τ_H	α_L	4	16	80	0	25	75	4	25	71	0
			α_H	1	45	54	0	56	44	1	57	42	0
	θ_H	τ_L	α_L	25	1	73	0	73	26	25	52	22	1
			α_H	2	26	70	0	96	2	2	96	0	2
		τ_H	α_L	5	35	60	0	66	34	5	64	31	0
			α_H	0	85	15	0	97	3	0	97	3	0
K_H	θ_L	τ_L	α_L	45	0	55	0	44	56	43	3	54	0
			α_H	30	7	63	0	66	34	23	50	27	0
		τ_H	α_L	8	2	90	0	30	70	6	26	68	0
			α_H	5	22	73	0	53	47	2	58	40	0
	θ_H	τ_L	α_L	21	0	78	0	39	60	21	27	51	1
			α_H	1	18	78	0	89	8	1	94	2	3
		τ_H	α_L	19	9	70	0	56	42	19	44	35	2
			α_H	3	41	56	0	95	5	3	94	3	0

REFERENCES

Ailawadi, K. L. and Harlam, B. A. (2009). Findings—retailer promotion pass-through: A measure, its magnitude, and its determinants. *Marketing Science*, 28(4):782–791.

Ailawadi, K. L., Harlam, B. A., Cesar, J., and Trounce, D. (2006). Promotion profitability for a retailer: the role of promotion, brand, category, and store characteristics. *Journal of Marketing Research*, 43(4):518–535.

Anderson, S. and de Palma, A. (1992). Multiproduct firms: A nested logit approach. *Journal of Industrial Economics*, 40(3):261–76.

Anupindi, R., Bassok, Y., and Zemel, E. (2001). A general framework for the study of decentralized distribution systems. *Manufacturing & Service Operations Management*, 3(4):349–368.

Aouad, A. and Segev, D. (2015). Display optimization for vertically differentiated locations under multinomial logit choice preferences.

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

Axsäter, S. (2003). Evaluation of unidirectional lateral transshipments and substitutions in inventory systems. *European Journal of Operational Research*, 149(2):438–447.

Aydin, G. and Heese, H. S. (2014). Bargaining for an assortment. *Management Science*, 61(3):542–559.

Aydin, G. and Porteus, E. L. (2008). Joint inventory and pricing decisions for an assortment. *Operations Research*, 56(5):1247–1255.

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

Bai, R., Van Woensel, T., Kendall, G., and Burke, E. K. (2013). A new model and a hyper-heuristic approach for two-dimensional shelf space allocation. *4OR*, 11(1):31–55.

Baumol, W. J. and Ide, E. A. (1956). Variety in retailing. *Management Science*, 3(1):93–101.

Bell, D. R., Chiang, J., and Padmanabhan, V. (1999). The decomposition of promotional response: An empirical generalization. *Marketing Science*, 18(4):504–526.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Besanko, D., Dubé, J.-P., and Gupta, S. (2005). Own-brand and cross-brand retail pass-through. *Marketing Science*, 24(1):123–137.

Besanko, D., Gupta, S., and Jain, D. (1998). Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science*, 44(11-part-1):1533–1547.

Bezawada, R., Balachander, S., Kannan, P., and Shankar, V. (2009). Cross-category effects of aisle and display placements: a spatial modeling approach and insights. *Journal of Marketing*, 73(3):99–117.

Bultez, A. and Naert, P. (1988). Sh. a.r.p.: Shelf allocation for retailers' profit. *Marketing Science*, 7(3):211–231.

Cannondale Associates (2017). Trade promotion spending and merchandising industry study. Available at http://cadentcg.com/wp-content/uploads/2017-Marketing-Spending-Study.pdf.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 13(2):185–213.

Chevalier, M. (1975). Increase in sales due to in-store display. *Journal of Marketing Research*, pages 426–431.

Chintagunta, P. K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, 12(2):184–208.

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

Corstjens, M. and Doyle, P. (1981). A model for optimizing retail space allocations. *Management Science*, 27(7):822–833.

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. Available at http://www.columbia.edu/~gmg2/logit_const.pdf.

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

Dhar, S. K., Hoch, S. J., and Kumar, N. (2001). Effective category management depends on the role of the category. *Journal of Retailing*, 77(2):165–184.

Drèze, X. and Bell, D. R. (2003). Creating win–win trade promotions: Theory and empirical analysis of scan-back trade deals. *Marketing science*, 22(1):16–39.

Euromonitor Int. (2017). Brand shares. Data retrieved from Passport GMID, Available at http://go.euromonitor.com/passport.html.

Feldman, J. and Topaloglu, H. (2015). Assortment optimization under the multinomial logit model with nested consideration sets. Technical report, Tech. rep., Working Paper.

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

Gallego, G. and Wang, R. (2014). Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62(2):450–461.

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

Geismar, H. N., Dawande, M., Murthi, B., and Sriskandarajah, C. (2015). Maximizing revenue through two-dimensional shelf-space allocation. *Production and Operations Management*, 24(7):1148–1163.

Gerstner, E. and Hess, J. D. (1991). A theory of channel price promotions. *The American Economic Review*, pages 872–886.

Gerstner, E. and Hess, J. D. (1995). Pull promotions and channel coordination. *Marketing Science*, 14(1):43–60.

Graham, R. L. (1969). Bounds on multiprocessing timing anomalies. *SIAM journal on Applied Mathematics*, 17(2):416–429.

Granot, D. and Sošić, G. (2003). A three-stage model for a decentralized distribution system of retailers. *Operations research*, 51(5):771–784.

Guadagni, P. M. and Little, J. D. (1998). When and what to buy: A nested logit model of coffee purchase. *Journal of Forecasting*, 17(3-4):303–326.

Gupta, S. (1988). Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing research*, pages 342–355.

Hanson, W. and Martin, K. (1996). Optimizing multinomial logit profit functions. *Management Science*, 42(7):992–1003.

Heese, H. S. and Martínez-de Albéniz, V. (2018). Effects of assortment breadth announcements on manufacturer competition. *Manufacturing & Service Operations Management*, 20(2):302–316.

Hui, S. K., Inman, J. J., Huang, Y., and Suher, J. (2013). The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing*, 77(2):1–16.

Inman, J. J. and Winer, R. S. (1998). *Where the rubber meets the road: A model of in-store consumer decision making*. Number Tech. Rep. 98–122. Marketing Science Institute Cambridge, MA.

Johnson, D. S. (1974). Approximation algorithms for combinatorial problems. *Journal of computer and system sciences*, 9(3):256–278.

Karp, R. M. (1972). Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer.

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

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

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.

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

Kurtuluş, M. and Nakkas, A. (2011). Retail assortment planning under category captainship. *Manufacturing & Service Operations Management*, 13(1):124–142.

Kurtuluş, M., Nakkas, A., and Ülkü, S. (2014). The value of category captainship in the presence of manufacturer competition. *Production and Operations Management*, 23(3):420–430.

Larson, J. S., Bradlow, E. T., and Fader, P. S. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing*, 22(4):395–414.

Leeflang, P. S., Selva, J. P., Dijk, A. V., and Wittink, D. R. (2008). Decomposing the sales promotion bump accounting for cross-category effects", journal = International Journal of Research in Marketing, volume = 25, number = 3, pages = 201 - 214.

Leeflang, P. S. H. and Parreño-Selva, J. (2012). Cross-category demand effects of price promotions. *Journal of the Academy of Marketing Science*, 40(4):572–586.

Li, H. and Huh, W. T. (2011). Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing & Service Operations Management*, 13(4):549–563.

Li, Z. (2007). A single-period assortment optimization model. *Production and Operations Management*, 16(3):369–380.

Manchanda, P., Ansari, A., and Gupta, S. (1999). The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing science*, 18(2):95–114.

Manrai, A. K. and Andrews, R. L. (1998). Two-stage discrete choice models for scanner panel data: An assessment of process and assumptions. *European Journal of Operational Research*, 111(2):193–215.

McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record*, (673).

Nair, H., Dubé, J.-P., and Chintagunta, P. (2005). Accounting for primary and secondary demand effects with aggregate data. *Marketing Science*, 24(3):444–460.

Neslin, S. A. (1990). Sales promotion: Concepts, methods, and strategies. *NJ: Prentice-Hall: Englewood Cliffs. and Kenneth J. Wisniewski (1989)," Price-Induced Patterns of Competition," Marketing Science*, 8:291–309.

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

Nijs, V., Misra, K., Anderson, E. T., Hansen, K., and Krishnamurthi, L. (2010). Channel pass-through of trade promotions. *Marketing Science*, 29(2):250–267.

Nordfält, J. and Lange, F. (2013). In-store demonstrations as a promotion tool. *Journal of Retailing and Consumer Services*, 20(1):20–25.

Pan, X. A. and Honhon, D. (2012). Assortment planning for vertically differentiated products. *Production and Operations Management*, 21(2):253–275.

Papatla, P. and Krishnamurthi, L. (1996). Measuring the dynamic effects of promotions on brand choice. *Journal of Marketing Research*, 33(1):20–35.

Paterson, C., Kiesmüller, G., Teunter, R., and Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2):125–136.

Pauwels, K. (2007). How retailer and competitor decisions drive the long-term effectiveness of manufacturer promotions for fast moving consumer goods. *Journal of Retailing*, 83(3):297–308.

Pauwels, K., Hanssens, D. M., and Siddarth, S. (2002). The long-term effects of price promotions on category incidence, brand choice, and purchase quantity. *Journal of marketing research*, 39(4):421–439.

Phillips, M., Parsons, A. G., Wilkinson, H. J., and Ballantine, P. W. (2015). Competing for attention with in-store promotions. *Journal of Retailing and Consumer Services*, 26:141–146.

POPAI (2012). 2012 shopper engagement study. Technical report, Point of Purchase Advertising Institute.

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

Rupp, R. (2015). Surviving the sneaky psychology of supermarkets. Available at http://theplate.nationalgeographic.com/2015/06/15/surviving-the-sneaky-psychology-of-supermarkets/.

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.

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

Slikker, M., Fransoo, J., and Wouters, M. (2005). Cooperation between multiple news-vendors with transshipments. *European Journal of Operational Research*, 167(2):370–380.

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

Stelzer, J. (2017). Ship from store: Easier or harder than you think...it depends. *Watson Customer Engagement*.

Stoops, G. T., Pearson, M. M., et al. (1988). Direct product profit: A view from the supermarket industry. *Journal of Food Distribution Research*, 19(2).

Tellis, G. J. and Zufryden, F. S. (1995). Tackling the retailer decision maze: Which brands to discount, how much, when and why? *Marketing Science*, 14(3):271–299.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Trivedi, M., Gauri, D. K., and Ma, Y. (2016). Measuring the efficiency of category-level sales response to promotions. *Management Science*, 63(10):3473–3488.

Van Heerde, H. J., Gupta, S., and Wittink, D. R. (2003). Is 75% of the sales promotion bump due to brand switching? no, only 33% is. *Journal of Marketing Research*, 40(4):481–491.

Van Heerde, H. J., Leeflang, P. S., and Wittink, D. R. (2004). Decomposing the sales promotion bump with store data. *Marketing Science*, 23(3):317–334.

Van Heerde, H. J. and Neslin, S. A. (2017). Sales promotion models. In *Handbook of Marketing Decision Models*, pages 13–77. Springer.

Vilcassim, N. J. and Chintagunta, P. K. (1995). Investigating retailer product category pricing from household scanner panel data. *Journal of Retailing*, 71(2):103–128.

Wilkinson, J. B., Mason, J. B., and Paksoy, C. H. (1982). Assessing the impact of short-term supermarket strategy variables. *Journal of Marketing Research*, pages 72–86.

Wong, H., van Houtum, G.-J., Cattrysse, D., and Van Oudheusden, D. (2006). Multi-item spare parts systems with lateral transshipments and waiting time constraints. *European Journal of Operational Research*, 171(3):1071–1093.

Zhang, J. (2006). An integrated choice model incorporating alternative mechanisms for consumers' reactions to in-store display and feature advertising. *Marketing Science*, 25(3):278–290.

Zhao, X. and Atkins, D. (2009). Transshipment between competing retailers. *IIE Transactions*, 41(8):665–676.