

Management and Effects of In-Store Promotional Displays

Oğuz Çetin, Adam J. Mersereau, and Ali K. Parlaktürk

Kenan-Flagler Business School, University of North Carolina, Chapel Hill, North Carolina 27599, USA,
Oguz.Cetin@kenan-flagler.unc.edu, ajm@unc.edu, Ali.Parlakturk@kenan-flagler.unc.edu

Problem definition: We examine a brick-and-mortar retailer’s choice of which product to include in a promotional display (e.g., an “endcap” display). The display provides a visibility advantage to both the featured product and its category, but it also has consequences for customer traffic and substitution.

Academic / practical relevance: While there has been considerable academic interest in the assortment planning problem (which products to offer?) and in the shelf-space allocation problem (how much space to devote to each product?), little attention has been paid to the problem of where to place products in the store. Promotional display choice can serve as a powerful demand-shaping lever for retailers. A good understanding of this problem can also facilitate a retailer’s negotiations with manufacturers.

Methodology: We develop analytical insights using a problem formulation based on a nested multinomial logit (NMNL) model of customer choice.

Results: When choosing a promotional product from a fixed category, the only possible optimal choices for promotional display lie along an “efficient set” drawn in terms of product popularities and margins. The optimal choice along the frontier depends on a quantity we call “aisle attractiveness” that depends on 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.

Managerial implications: Our work provides guidance for how retailers can use and value promotional displays effectively. We highlight the importance of considering externalities of a display decision on store traffic and demand for other products.

Key words: retail; assortment planning; promotional displays; choice modeling; nested multinomial logit model

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, [POPAL \(2012\)](#) 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) and the introduction of our model (Section 3), we present our analysis in three parts. In Section 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 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 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 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 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 the e-companion.

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 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.

3. Model Setup

We consider a category consisting of $n \geq 2$ products, and we indicate this set of products by $S = \{1, \dots, n\}$. 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 1. (Some of the notations in Figure 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:

$$U_j = Y_j + W_k + \epsilon_j, \quad (1)$$

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.

$$\Phi(\epsilon) = \exp \left\{ - \sum_{k=1}^K \left(\sum_{j \in B_k} e^{-\epsilon_j / \lambda_k} \right)^{\lambda_k} \right\}, \quad (2)$$

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. 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.

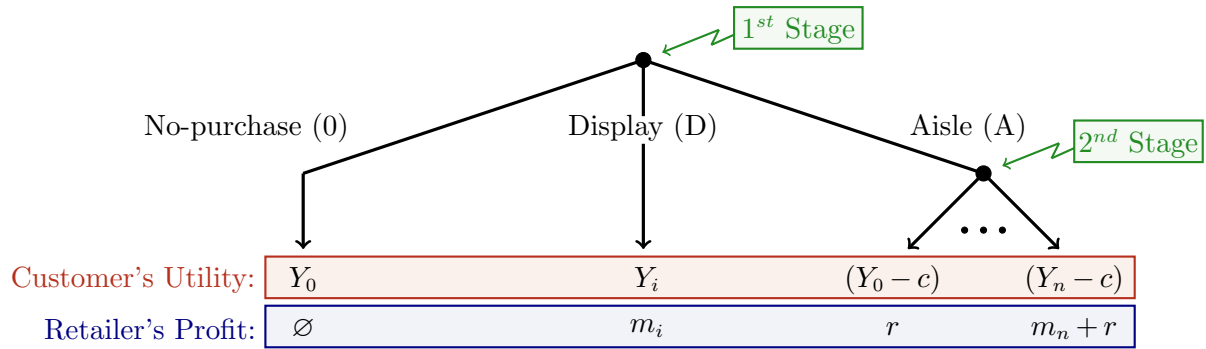


Figure 1 Customer choice modeled as a two-stage decision process.

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 B](#). 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 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:

$$P_0^i = \mathbb{P}\{\text{no purchase from display and no aisle visit}\} = \frac{e^{Y_0}}{e^{Y_0} + e^{Y_i} + e^{-c+\lambda \cdot I}}, \quad (3)$$

$$P_D^i = \mathbb{P}\{\text{purchase from the display}\} = \frac{e^{Y_i}}{e^{Y_0} + e^{Y_i} + e^{-c+\lambda \cdot I}}, \quad (4)$$

$$P_A^i = \mathbb{P}\{\text{visit the aisle}\} = \frac{e^{-c+\lambda \cdot I}}{e^{Y_0} + e^{Y_i} + e^{-c+\lambda \cdot I}}, \quad (5)$$

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 $p_{j|A} = e^{Y_j/\lambda} / \sum_{q \in S \cup \{0\}} e^{Y_q/\lambda}$ 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, 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.

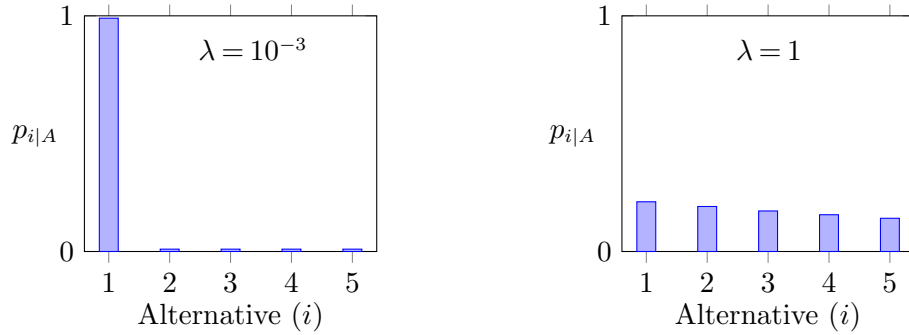


Figure 2 Second stage choice probabilities ($p_{i|A}$) when $\lambda = 10^{-3}$ and $\lambda = 1$ for $\mathbf{Y} = \langle 0.5, 0.4, 0.3, 0.2, 0.1 \rangle$.

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:

$$\pi^i = P_D^i \cdot m_i + P_A^i \cdot (\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r), \quad i \in S. \quad (6)$$

The first term of equation (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 $(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r)$. 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 4). Demand expansion effects will become relevant when we consider the value of promotional displays across categories (Section 5) and product-specific demand expansion effects (Section 6).

Throughout Sections 4 and 5 we will make the following assumption that the expected impulse spending in the aisle (r) is not too high.

ASSUMPTION 1. $\max_{i \in S} \{m_i\} > P_A^0 \cdot (\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r)$ where $P_A^0 = \frac{e^{-c+\lambda \cdot I}}{e^{Y_0} + e^{-c+\lambda \cdot I}}$.

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 6.

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 5.

This section is organized as follows. In Section 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 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.

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 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 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 “sleepers” 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 “sleepers” 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 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:

$$E = \{i \in S \mid \nexists j \in S \text{ s.t. } Y_j \geq Y_i \text{ and } m_j \geq m_i\} \quad (7)$$

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

To refine the efficient set further, we define a product-specific index $\theta_i = \frac{e^{Y_i} \cdot m_i}{e^{Y_0} + e^{Y_i}}$. 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, 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. *Product characteristics \mathbf{Y} and \mathbf{m} are such that $\theta_i \leq (\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r) \forall i \in S$.*

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.

PROPOSITION 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 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, or more formally,

$$E' = \{i \in E \mid \nexists j \in E \text{ s.t. } Y_i > Y_j \text{ and } \theta_j \geq \theta_i\} \quad (8)$$

Hereafter we order products in E' such that $Y_1 > \dots > Y_{|E'|}$ and $m_1 < \dots < m_{|E'|}$ 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 $\theta_1 > \dots > \theta_{|E'|}$ by Proposition 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, 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 A.)

Table 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, 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.

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 $i^*(\cdot) = \arg \max_{i \in S} \{\pi^i(\cdot)\}$ 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 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 $(\bar{m} + r)$ the “aisle profit”.

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 3 shows the *ceteris paribus* impact of the parameters r and c on $i^*(\cdot)$.

PROPOSITION 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, 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.

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 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 $\arg \max_{i \in S} \{\pi^i(S)\} \leq \arg \max_{i \in S} \{\pi^i(S^+)\}$.*

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 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.

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

Lemma 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 1, we can state the following result on the impact of λ on the retailer's optimal promotional display choice.

PROPOSITION 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 3 part 2: Promoting a sleeper product makes sense when the aisle is attractive. By Lemma 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, 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 market 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.

5. The Value of a Promotional Display to a Category

Section 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 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

$$\Pi^i = \begin{cases} (1 + \beta)\pi^i & \text{if } i \in S \\ P_A^0 \cdot (\bar{m} + r) & \text{if } i = 0 \text{ (no-promotion)} \end{cases} \quad (9)$$

Observe that when the category is featured in the promotional display, we achieve the same expected profit π^i as in Section 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 (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 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,

$$\Delta^i = \Pi^i - \Pi^0 \quad (10)$$

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 attractiveness, which was critical in our analysis in Section 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.

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 6, we compare the value of the display for two assortments having different sizes:

PROPOSITION 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 $j = \arg \max_{i \in S} \{\Pi^i(S)\}$ and $k = \arg \max_{i \in S^+} \{\Pi^i(S^+)\}$. 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 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 7 and 8 as well.

Proposition 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 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 $j = \arg \max_{i \in S} \{\Pi^i(S)\}$, $k = \arg \max_{i \in S} \{\Pi^i(S^+)\}$, and $l = \arg \max_{i \in S} \{\Pi^i(S^{++})\}$. 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 7 and 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 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.

5.2. Customer Heterogeneity

Recall that heterogeneity in preferences implies a more attractive aisle (Lemma 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 β is low and the substitution effect dominates, the retailer benefits from promoting a category which generates high profits from the display.

PROPOSITION 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 5, we study what happens when \bar{m} is non-decreasing in λ . Section 4 explains why this is a plausible assumption. Proposition 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.).

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 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 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).

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:

$$\tilde{\Pi}^i = \begin{cases} (1 + \beta\phi_i)\pi^i & \text{if } i \in S \\ \Pi^0 & \text{if } i = 0 \text{ (no-promotion)} \end{cases} \quad (11)$$

where π^i is the original profit function defined in (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 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 1 is no longer sufficient for Proposition 1 to hold.

EXAMPLE 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 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 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 3 resembles Assumption 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 9. Under Assumption 3, a dominated product is never optimal to promote in the case of product-specific demand expansion.

Proposition 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 3-5) also require some modification under product-specific category expansion effects, but our core insights from Section 4.2 remain intact. Proposition 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 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_A^i is decreasing in product popularity Y_i and ϕ_i is non-decreasing in Y_i , how $P_A^i(1 + \beta\phi_i)$ 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 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 3 (part 2), 4, and 5, it is straightforward to show that as aisle attractiveness vanishes (i.e., $e^{-c+\lambda I} \rightarrow 0$) it becomes optimal to promote the product $\arg \max_{i \in S} \{\theta_i\}$ (Recall that the aisle attractiveness $e^{-c+\lambda I}$ increases with λ and assortment size and decreases with c). We note that as $e^{-c+\lambda I} \rightarrow +\infty$, it becomes optimal to promote the product $\arg \max_{i \in S} \{\phi_i\}$. 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.

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.

Acknowledgments

The authors thank Department Editor Brian Tomlin, an anonymous associate editor, and two anonymous reviewers for feedback that resulted in several improvements to the paper. The second author thanks the Sion A. Boney, Jr. Endowment Fund for support. The third author thanks the Benjamin Cone Research Fund for support.

Appendix

A. Parameters of the Numerical Study Reported in Table 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.

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 1 and 2 under these settings.

B. Incorporating Correlated Utilities

B.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 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:

$$U_0 = Y_0 + \underbrace{\nu_0 + \lambda\xi_0}_{=\epsilon_0} \text{ (no-purchase alternative),} \quad (12)$$

$$U_i = Y_i + \underbrace{\nu_D + \lambda\xi_i}_{=\epsilon_i} \text{ (display alternative when product } i \text{ is displayed),} \quad (13)$$

$$U_j = Y_j - c + \underbrace{\nu_A + \lambda\xi_j}_{=\epsilon_j} \text{ (alternative corresp. to prod. } j \text{ in the aisle).} \quad (14)$$

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 $\langle \xi_0, \dots, \xi_{15} \rangle$. 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.

B.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 A, 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.

B.3. Evaluating the Independence Assumption

Table 3 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 occurring 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 (12)-(14). 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 3 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 4 (Proposition 3-5) 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.

References

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

- 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., 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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. and Xu, Y. (2011). Optimal and competitive assortments with endogenous pricing under hierarchical consumer choice models. *Management Science*, 57(9):1546–1563.
- 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. *International Journal of Research in Marketing*, 25(3):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.
- 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.
- Nordfält, J. and Lange, F. (2013). In-store demonstrations as a promotion tool. *Journal of Retailing and Consumer Services*, 20(1):20–25.
- 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.
- Rupp, R. (2015). Surviving the sneaky psychology of supermarkets. Available at <http://theplate.nationalgeographic.com/2015/06/15/surviving-the-sneaky-psychology-of-supermarkets/>.
- 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).
- 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.
- 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.

Electronic Companion to “Management and Effects of In-Store Promotional Displays”

EC.1. Proofs of Analytical Results

Proof of Proposition 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 $K = \arg \max_{i \in S} \{m_i\}$. Note that K is not necessarily a singleton. Hence, let $k = \arg \max_{i \in K} \{Y_i\}$. 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 $M = P_A^0 \cdot (\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r)$. We can write π^x as a convex combination of m_x and M for all $x \in S$ as follows:

$$\pi^x = P_D^x \cdot m_x + (1 - P_D^x) \cdot M \quad (\text{EC.1})$$

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

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

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 1, we have $m_k > M$. Since $m_j \leq m_i \leq M < m_k$, it follows from (EC.1) that $\pi^k > \pi^j$. \square

Proof of Proposition 2. We can write π^x as follows:

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

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

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

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

The first inequality follows by Assumption 2 and the fact that $P_A^i > P_A^j$. The second inequality follows since $\theta_i \geq \theta_j$. \square

Proof of Proposition 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:

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

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 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 $(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) + r)$ for $x \in S$ as in EC.3. It is straightforward to show that $\lim_{c \rightarrow \infty} P_A^x = 0$ for all $x \in S$. Then, in the limit, $\pi^j(c) > \pi^k(c)$, because $\theta_j > \theta_k$ by Proposition 2. \square

Proof of Proposition 4. Assume without loss of generality that $\arg \max_{i \in S} \{\pi^i(S)\} = j$ and $\arg \max_{i \in S} \{\pi^i(S^+)\} = k$ 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, 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 :

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

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

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

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

Since the LHS of the above equation is continuous and monotonically increasing in x , $\tilde{\pi}^j(x)$ and $\tilde{\pi}^k(x)$ 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 $\tilde{\pi}^j(\lambda \cdot I(\mathbf{Y}_S, \lambda)) = \pi^j(S) > \pi^k(S) = \tilde{\pi}^k(\lambda \cdot I(\mathbf{Y}_S, \lambda))$, and $\tilde{\pi}^k(\lambda \cdot I(\mathbf{Y}_{S^+}, \lambda)) = \pi^k(S^+) > \pi^j(S^+) = \tilde{\pi}^j(\lambda \cdot I(\mathbf{Y}_{S^+}, \lambda))$ 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 $x \rightarrow -\infty$, we know that $\tilde{\pi}^j(x) > \tilde{\pi}^k(x)$ iff $\theta_j > \theta_k$ because $\lim_{x \rightarrow -\infty} \tilde{\pi}^i(x) = \theta_i$ for all $i \in S$. Hence, we conclude that $\theta_j > \theta_k$. \square

Proof of Lemma 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:

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

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 $\lambda \in (0, 1]$. \square

Proof of Proposition 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, $\arg \max_{i \in S} \{\pi^i(\lambda' - \epsilon)\} = j$ and $\arg \max_{i \in S} \{\pi^i(\lambda' + \epsilon)\} = k$ for sufficiently small $\epsilon > 0$. By Proposition 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 EC.7 and EC.8. Following similar arguments as in EC.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

$$\tilde{\pi}^j(\lambda' \cdot I(\mathbf{Y}, \lambda')) = \pi^j(\lambda') = \pi^k(\lambda') = \tilde{\pi}^k(\lambda' \cdot I(\mathbf{Y}, \lambda')). \quad (\text{EC.10})$$

We have $\lim_{x \rightarrow -\infty} \tilde{\pi}^i(x) = \theta_i$ 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:

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

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

Since $\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda)$ is non-decreasing in λ by the proposition assumption, the expression $(\bar{m}(\mathbf{Y}, \mathbf{m}, \lambda) - M)$ in the above equations are non-negative for any $\lambda > \lambda'$. Moreover, $\theta_j < \theta_k$ implies

that $Y_j < Y_k$ by definition of E' , and therefore $P_A^j(\lambda \cdot I(\mathbf{Y}, \lambda)) > P_A^k(\lambda \cdot I(\mathbf{Y}, \lambda))$ for any $\lambda \in (0, 1]$. Finally, case 2 implies that $\tilde{\pi}^j((\lambda' + \epsilon) \cdot I(\mathbf{Y}, (\lambda' + \epsilon))) > \tilde{\pi}^k((\lambda' + \epsilon) \cdot I(\mathbf{Y}, (\lambda' + \epsilon)))$. As a result, plugging $\lambda = (\lambda' + \epsilon)$ in (EC.11) and (EC.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 $\theta_j > \theta_k$. \square

Proof of Proposition 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:

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

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

$$\pi^k(S^+) \geq \pi^j(S^+) = (1 - P_A^j(S^+)) \cdot \theta_j + P_A^j(S^+) \cdot (\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) + r) \quad (\text{EC.13})$$

$$> (1 - P_A^j(S)) \cdot \theta_j + P_A^j(S) \cdot (\bar{m}(\mathbf{Y}_{S^+}, \mathbf{m}_{S^+}, \lambda) + r) \quad (\text{EC.14})$$

$$\geq (1 - P_A^j(S)) \cdot \theta_j + P_A^j(S) \cdot (\bar{m}(\mathbf{Y}_S, \mathbf{m}_S, \lambda) + r) = \pi^j(S) \quad (\text{EC.15})$$

where (EC.14) follows from Assumption 2 and the fact that $P_A^i(S^+) > P_A^i(S)$ for all $i \in S$, and (EC.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 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

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

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

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

We will show that $\bar{\Delta}(S)$ decreases as the assortment S expands. To that end, consider the function $f : [I(\mathbf{Y}_S), I(\mathbf{Y}_{S^+})] \rightarrow [e^{Y_0} + e^{-c+\lambda \cdot I(S)}, e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}]$ such that $x = f(I) = e^{Y_0} + e^{-c+\lambda \cdot I(\cdot)}$. 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 $[e^{Y_0} + e^{-c+\lambda \cdot I(S)}, e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}]$. By plugging x in the RHS of (EC.16) we get:

$$\frac{e^{Y_0} \cdot e^{Y_k} \cdot m_k + e^{Y_k} \cdot (m_k - M - r) \cdot (x - e^{Y_0})}{(e^{Y_k} + x) \cdot x} \quad (\text{EC.17})$$

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 (EC.17) is decreasing in x in the interval $[e^{Y_0} + e^{-c+\lambda \cdot I(S)}, e^{Y_0} + e^{-c+\lambda \cdot I(S^+)}]$. To that end, we compute the first derivative of (EC.17) with respect to x as:

$$-\frac{e^{Y_k} x [x(m_k - M - r) + 2e^{Y_0}(M + r)] + e^{2Y_k+Y_0}(M + r)}{(e^{Y_k} + x)^2 \cdot x^2} \quad (\text{EC.18})$$

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

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

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

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

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

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

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

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

Inequality (EC.20) implies that (EC.19) holds for all $x \leq \bar{x}$. To clarify, note that (EC.19) holds for all x if $(m_k - M - r) \geq 0$. Otherwise, it holds for all $x \leq \bar{x}$ because (EC.20) guarantees it for $x = \bar{x}$, and the LHS of (EC.19) is decreasing in x . As a result, we showed that the expression in (EC.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 $\beta = 0$. \square

Proof of Corollary 1 Let $\beta'(\cdot)$ denote the threshold β' in Proposition 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 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 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 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 $\arg \max_{i \in S} \{\pi^i(\lambda)\} = j$ 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:

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

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

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 $f : [\underline{\lambda}, \lambda'] \rightarrow [e^{Y_0} + e^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}, e^{Y_0} + e^{-c+\lambda' \cdot I(\lambda')}]$ such that $x = f(\lambda) = e^{Y_0} + e^{-c+\lambda \cdot I(\lambda)}$. Notice that x increases monotonically in λ due to Lemma 1. Hence, $f(\lambda)$ is a one-to-one correspondence between the sets $[\underline{\lambda}, \lambda']$ and $[e^{Y_0} + e^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}, e^{Y_0} + e^{-c+\lambda' \cdot I(\lambda')}]$. By plugging x in (EC.22) we get:

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

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 $x \in [e^{Y_0} + e^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}, e^{Y_0} + e^{-c+\lambda' \cdot I(\lambda')}]$. The first derivative of (EC.23) with respect to x is:

$$\frac{d\bar{\Delta}(x)}{dx} = -\frac{e^{Y_j} x [x(m_j - \bar{m}(\underline{\lambda}) - r) + 2e^{Y_0}(\bar{m}(\underline{\lambda}) + r)] + e^{2Y_j+Y_0}(\bar{m}(\underline{\lambda}) + r)}{(e^{Y_j} + x)^2 \cdot x^2} \quad (\text{EC.24})$$

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

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

By Assumption 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:

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

By some algebraic manipulation, we get

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

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

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

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

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

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

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

The first inequality follows as $\bar{\Delta}(x)$ is decreasing in x and $(e^{Y_0} + e^{-c+\underline{\lambda} \cdot I(\underline{\lambda})}) < (e^{Y_0} + e^{-c+\lambda' \cdot I(\lambda')})$. 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 $\beta = 1$. \square

Proof of Proposition 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 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 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 $\arg \max_{i \in S} \{\pi^i(c)\} = j \quad \forall c \in [\underline{c}, c']$ and $\arg \max_{i \in S} \{\pi^i(c)\} = i^*(c) \neq j$ 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:

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

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 $f : [\underline{c}, c'] \rightarrow [e^{Y_0} + e^{-c'+\lambda \cdot I}, e^{Y_0} + e^{-\underline{c}+\lambda \cdot I}]$ such that $x = f(c) = e^{Y_0} + e^{-c+\lambda \cdot I}$. Notice that x decreases monotonically in c . Hence, $f(c)$ is a one-to-one correspondence between the sets $[\underline{c}, c']$ and $[e^{Y_0} + e^{-c'+\lambda \cdot I}, e^{Y_0} + e^{-\underline{c}+\lambda \cdot I}]$. By plugging x in (EC.29) we get:

$$\bar{\Delta}(x) = \frac{e^{Y_0} \cdot e^{Y_j} \cdot m_j + e^{Y_j} \cdot (m_j - M - r) \cdot (x - e^{Y_0})}{(e^{Y_j} + x) \cdot x} \quad (\text{EC.30})$$

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 $x \in [e^{Y_0} + e^{-c'+\lambda \cdot I}, e^{Y_0} + e^{-\underline{c}+\lambda \cdot I}]$. To that end, we compute the first derivative of (EC.30) as:

$$\frac{d\bar{\Delta}(x)}{dx} = - \frac{e^{Y_j} x [x(m_j - M - r) + 2e^{Y_0}(M + r)] + e^{2Y_j+Y_0}(M + r)}{(e^{Y_j} + x)^2 \cdot x^2} \quad (\text{EC.31})$$

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

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

It follows from Assumption 1 that $m_j \geq \frac{e^{-c+\lambda \cdot I}}{e^{Y_0} + e^{-c+\lambda \cdot I}} \cdot (M + r)$ for all $c \in [\underline{c}, c']$. By plugging x in and some algebraic manipulation, we get $x \cdot (m_j - M - r) + e^{Y_0}(M + r) \geq 0$ for all $x \in [e^{Y_0} + e^{-c'+\lambda \cdot I}, e^{Y_0} + e^{-\underline{c}+\lambda \cdot I}]$. This implies that (EC.32) holds for all x in the interval $[e^{Y_0} + e^{-c'+\lambda \cdot I}, e^{Y_0} + e^{-\underline{c}+\lambda \cdot I}]$. Hence, $\bar{\Delta}(x)$ is decreasing in x in that interval. Then, we can write $\Delta^j(\underline{c}) = \bar{\Delta}(e^{Y_0} + e^{-\underline{c}+\lambda \cdot I}) < \bar{\Delta}(e^{Y_0} + e^{-c'+\lambda \cdot I}) = \Delta^j(c') = \Delta^{i^*(c')}(c')$. The inequality follows because $\bar{\Delta}(x)$ is decreasing in x , and $(e^{Y_0} + e^{-c'+\lambda \cdot I}) < (e^{Y_0} + e^{-\underline{c}+\lambda \cdot I})$. 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 $\beta = 1$. \square

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

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

$$\tilde{\Pi}^j = (1 + \beta\phi_j)(P_D^j \cdot m_j + (1 - P_D^j) \cdot \Pi^0) \leq (1 + \beta\phi_i)(P_D^i \cdot m_i + (1 - P_D^i) \cdot \Pi^0) = \tilde{\Pi}^i \quad (\text{EC.33})$$

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 3. Hence, it is never optimal to promote the dominated product j . \square

Proof of Proposition 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:

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

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 $r_1 > r_2$. \square

EC.2. Simulation Support for the Results in Section 5

We ran simulations to verify the results in Section 5 (Propositions 6,7,8) 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 Appendix B. 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 Appendix B.2. Proposition 6 holds for more than 95% of the instances, whereas Propositions 7 and 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.