Retail Assortment Planning in the Presence of Consumer Search*

Gérard P. Cachon
cachon@wharton.upenn.edu
opim.wharton.upenn.edu/~cachon

Christian Terwiesch
terwiesch@wharton.upenn.edu

Yi Xu
xuyi@wharton.upenn.edu

Operations and Information Management
The Wharton School · University of Pennsylvania
Philadelphia PA, 19104

December 20, 2002

*The authors would like to thank Erica Plambeck for her helpful comments. An electronic copy of this paper is available from the first author's webpage. Financial support from the Wharton e-Business Initiative (WeBI) is gratefully ackowledged.

Abstract

Consumers often know what kind of product they wish to purchase, but do not know which specific variant best fits their needs, so consumers shop around. As a result, a consumer may find an acceptable product in one store but nevertheless purchase nothing, opting to continue searching for an even better product. We study three versions of the retail assortment problem: a traditional, no-search, version that does not explicitly consider consumer search and two versions that implement two different consumer search models. In each of the three versions we use the multinomial logit to model consumer choice. Our analysis suggests the retailer's decision to add a product to an assortment should not only consider the direct costs and revenues of the product, but also anticipate the indirect benefit an extended assortment has in preventing consumer search. For example, we show it may even be optimal to add an unprofitable product to an assortment so as to prevent consumer search. We also test whether the no-search version of assortment planning could perform well in practice. In particular, we presume a retailer implements no-search assortment planning with the following iteration: an assortment is chosen, sales data are collected, a consumer choice model is fitted to the data, a new assortment is chosen, etc. We show that this iteration can lead to a heuristic equilibrium: there exists an assortment that is optimal given the previously observed sales data and the sales data observed with that assortment recommends the same assortment. We demonstrate that the heuristic equilibrium assortment is never deeper than optimal and often contains fewer variants than optimal. The profit loss from the heuristic equilibrium assortment can be substantial: in some extreme cases the no-search assortment planning model recommends killing the category (i.e., stocking no products) even though the optimal assortment is profitable. We conclude that retailers should be careful when considering a reduction in their assortment if they sell in a market characterized by consumer search.

1 Introduction

Consumers often know what kind of product they wish to purchase, but do not know which specific model or variant best fits their needs. Consider, for example, a consumer shopping for a new digital camera. Upon inspecting the cameras in the assortment of a retail store, the consumer might be able to assess the utility associated with each of the present cameras, yet would face uncertainty about those outside the store's assortment. Hence, even if she finds an acceptable camera in the current store, exceeding her utility associated with not buying a camera at all, she may nevertheless continue her search at other retailers. Traditional assortment planning models do not account for search behavior explicitly. This paper studies whether it is important to do so.

A traditional approach to assortment planning begins by fitting a consumer choice model to observed sales data. The multinomial logit model (MNL) is an intuitive, frequently used and successfully applied consumer choice model (see Anderson, de Palma and Thisse 1992). With that model a deterministic base utility is estimated for each product variant, including a faux variant that represents the consumer's utility from not purchasing any variant. A consumer's realized utility for each variant equals the variant's base utility plus a random shock. The consumer chooses the variant that has the highest realized utility, which may be the "no-purchase" variant.[1] After the development of the consumer choice model, an assortment is chosen that balances the sales gains of a deeper assortment with the additional operational costs of a deeper assortment (e.g., van Ryzin and Mahajan 1999 and Mahajan and van Ryzin 2001)

In addition to the traditional no-search model, we study two models that incorporate the influence of search into the MNL model of the consumer choice process. In each search model

[1] In many applications the consumer's purchase incidence decision (whether to purchase a variant in the category or not) is separated from the consumer's brand choice decision (which variant to purchase conditional that some variant is purchased). In our model these decisions are integrated via the creation of a faux variant to represent the no-purchase decision. From an analytical perspective it does not matter in our model whether these decisions are separated or integrated. However, the structure of these decisions can influence the estimation procedure.

the consumer still does not make a purchase if the no-purchase variant has the highest realized utility. However, the consumer may not make a purchase even if there is an acceptable product in the retailer's assortment (i.e., there is some variant that has a higher utility than the no-purchase utility) because the consumer may choose to continue her search (at some other retailer) for an even better variant.

Whether a consumer chooses the search option depends on a number of factors: the realized utility of the best variant at the retailer, the cost of search and the consumer's expectation regarding how much more utility she may be able to gain with search outside the retailer's assortment. It is with respect to the latter that our two search models differ. Roughly speaking, with the *independent assortment* search model a consumer expects that the retailer's assortment is unique, i.e., if the search option is chosen then the consumer expects to observe different variants by searching. It follows that the consumer's expected value from search is independent of the retailer's assortment (because search yields different variants no matter which variants the retailer carries). Examples that might reasonably fit this setting include antique dealers and jewelry stores.

The second search model is referred to as the *overlapping assortment* model. In this case there is a limited number of products available in the market, as in the digital camera example. As a result, expanding the retailer's assortment now reduces the value of search because it reduces the potential number of new variants the consumer could observe if search is chosen. This is an important effect that is not captured in the no-search model: a deeper assortment increases the likelihood a consumer purchases some variant because it reduces the value of search.

With each of the three models (no-search and the two search models) the retailer's objective is to choose an assortment that maximizes expected profit. Expected revenue for each variant in the assortment depends on which variants are in the assortment. Expected cost for each variant is concave and increasing in the variant's demand rate to reflect economies of scale in managing inventory, shelf space and transportation.

Even though the traditional assortment planning model does not account for consumer search, it still may perform reasonably well in practice. To explore this issue, we presume a retailer conducts multiple iterations with assortment planning: an initial assortment is chosen, sales data is collected with the initial assortment, an updated assortment is implemented, sales data is collected with the updated assortment, etc. For example, the retailer may begin with a relatively narrow assortment and slowly increase the number of variants. As a new variant is added the retailer is able to observe the sales for that variant as well as the incremental sales for the category. Alternatively, the retailer may begin with all of the potential products in a category (which is sometimes called depth testing) and then slowly decrease the number of variants. We find that this process can converge to a stable assortment, i.e., iterative application of the assortment planning model leads to a heuristic equilibrium: the equilibrium assortment is optimal (according to the no-search model) given the sales data observed with that assortment. (See Cachon and Kok 2002 for a precise definition of a heuristic equilibrium.) Although an equilibrium assortment is stable, we find that the no-search assortment planning model chooses an assortment with fewer variants than optimal. In a numerical study we find that the profit loss (measured as a percentage of the optimal profit) is relatively small in the search model with an independent assortment, but can be substantial in the search model with an overlapping assortment. In particular, we find that the no-search model can lead a retailer to close down a category (carry zero variants) even though there exists profitable assortments.

The remainder of this paper is organized as follows. The next section reviews the related literature. Section 3 describes the model. Section 4 studies the structure of the optimal assortment for the three assortment planning models. In section 5, we evaluate the performance of the no-search assortment model. Section 6 provides a guideline on the estimation of the parameters associated with consumer search. Conclusions and future research are provided in Section 7.

2 Literature Review

This research is most closely related to assortment planning papers that balance the trade off between the expanded revenue of a deeper assortment and the higher operational costs of a deeper assortment: e.g., Smith and Agrawal (2000), Smith (2002), Chong, Ho and Tang (2001), van Ryzin and Mahajan (1999) and Mahajan and van Ryzin (2001). Smith and Agrawal (2000) as well as Smith (2002) do not implement the MNL model for consumer choice, whereas the other three papers do. Chong, Ho and Tang (2001) divided the consumer choice model into two parts, a purchase incidence model and a brand selection model, whereas this research collapses those two parts into one. van Ryzin and Mahajan (1999) study two types of demand uncertainty: in the first type each consumer's utility is an independent random draw from an known distribution, whereas in the second type all consumers have the same utility drawn once from the same distribution. This research assumes the former (heterogenous utility across consumers). Mahajan and van Ryzin (2001) assume consumers purchase their most preferred variant among the variants that are available in stock (assuming it exceeds the no-purchase variant), whereas the other papers and our model do not consider consumer switching behavior in response to stockouts. Finally, the major point of differentiation is that none of those papers considers the impact of consumer search on the estimation of the consumer choice model or the resulting assortment decision. Aydin and Hausman (2002) use a model similar to van Ryzin and Mahajan (1999), but their analysis of the assortment problem focuses on supply chain coordination.

There is an extensive literature in economics that considers consumer search among retailers over price, e.g., Stigler (1961), Salop and Stiglitz (1977) and Stahl (1989). In our model consumers do not search for a lower price but rather, for a better product. Anderson and Renault (1999), Stahl (1982) and Wolinsky (1983, 1984) incorporate consumer search for a better product, but they assume each retailer carries only a single product. Weitzman (1979) and Mogan (1983) study optimal consumer search strategies but do not investigate how consumer search influences retail assortment planning.

Fisher and Rajaram (2000) study a model for merchandise testing to calibrate a consumer choice model (i.e., a sales forecast for each store and each product), but they do not consider which products to stock, nor do they account for consumer search.

In the context of a different model (a newsvendor model with clearance pricing) Cachon and Kok (2002) define the concept of a heuristic equilibrium. Roughly speaking, a heuristic is in equilibrium if the actions it recommends are optimal given the inputted data, and the inputted data are the expected outcomes given the recommended actions. In our context the heuristic is the assortment planning model and the inputted data are the estimates of the consumer choice model. The concept of a heuristic equilibrium is relevant in this setting because the assortment chosen influences the outcome data (each variant's market share, including the no-purchase variant) and the outcome data influences the estimated utilities for the choice model used to choose the assortment. We study the properties of the heuristic equilibrium for our model in Section 5. Armony and Plambeck (2002) also study a heuristic equilibrium, but in a different setting.

While there have been many successful applications of the MNL in marketing (e.g. Guadagni and Little 1983), it should be noted that the MNL is not a perfect model of consumer choice. For example, purchase incidence always increases with the MNL model when new variants are added, but there is empirical evidence indicating that purchase incidence can actually decrease as the assortment is expanded in part due to the added confusion an extremely deeper assortment imposes on a consumer (e.g., Huffman and Kahn 1998). Furthermore there is evidence that consumer choice is influenced by how an assortment is presented and organized for consumers (e.g., Simonson 1999, and Hoch, Bradlow and Wansink 1999). That issue is not captured by the MNL model.

3 Model

We consider a risk neutral retailer who sells a product with n possible variants to risk neutral consumers. Let $N = \{1, 2, ..., n\}$ be the set of possible variants and let S be the subset of variants in the retailer's assortment. In addition to the actual product variants, we create variant 0, a faux variant, to represent the no-purchase option, i.e., if a consumer chooses variant 0 then they choose to not purchase any product.

We use the multinomial logit (MNL) to model consumers' utilities across variants. (See Mahajan and van Ryzin 1998 for a review of the assumptions, limitations, characteristics and properties of the MNL model). With this model, let U_i be a consumer's utility associated with product variant i, where U_0 is the utility of the no-purchase variant: $U_i = (u_i - p_i) + \zeta_i$, where u_i is a constant, identical across consumers, p_i is the market price of variant i ($p_0 = 0$) and the random variable ζ_i has a zero mean Gumbel distribution. We refer to $(u_i - p_i)$ as variant i's expected net utility. Without loss of generality, the variants are labeled in decreasing net utility order: $u_i - p_i \geq u_j - p_j$ for all $i < j$. Let $F(x)$ be the distribution function of ζ_i,

$$F(x) = \exp\left[-\exp\left(-\left(\frac{x}{\mu} + \gamma\right)\right)\right],$$

where γ is Euler's constant and μ is a scale parameter. Let $f(x)$ be its density function. The realizations of ζ_i are independent across consumers, so ζ_i creates consumer heterogeneity.

A consumer's product choice depends on which of the three models is considered. Nevertheless, with each model there is some probability that a consumer chooses variant i if the assortment is S. Let $q_i(S)$ be that probability, where recall that $q_0(S)$ is the no-purchase probability. Without loss of generality, the consumer population is normalized to 1, so $q_i(S)$ is also variant i's demand. Section 3.1 evaluates $q_i(S)$ for each of the three models.

Let c_i be the retailer's purchase cost of variant i. Define m_i as the net profit margin of variant i, $m_i = p_i - c_i$. We assume the profit margins are monotonically increasing in the variants' net utility, i.e., $m_j \geq m_k$ if $u_j - p_j \geq u_k - p_k$.[2] In addition to each variant's purchase cost, let $c(q_i(S))$ be the operational costs associated with including variant i into the retailer's assortment, i.e., the shelf space, holding, handling and transportation costs of stocking variant i. To reflect economies of scale in operational costs, such as is common in the EOQ or newsvendor models, assume $c(\cdot)$ is concave and increasing.[3] Finally, expected profit of a variant $i \in S$ is

$$\pi_i(S) = m_i q_i(S) - c(q_i(S)),$$

where $m_i q_i(S)$ is the expected net revenue of variant i. The retailer's objective is to choose an assortment S to maximize expected profit:

$$\max_{S \subseteq N} \; \pi(S) = \sum_{i \in S} \pi_i(S). \tag{1}$$

The superscripts $\{m, si, so\}$ are used to identify notation associated with the no-search model, the search model with independent assortment and the search model with overlapping assortment respectively. For example, $q_i^m(S)$ is variant i's demand with the no-search (traditional MNL) model and assortment S and $\pi_i^{si}(S)$ is variant i's profit with search and independent assortment.

3.1 Evaluating demand

This section evaluates $q_i^m(S)$, $q_i^{si}(S)$ and $q_i^{so}(S)$, i.e., variant i's demand with assortment S and the traditional no-search model, the search model with independent assortment and the

[2] With monotone margins it is shown that the optimal assortment (in two of the three models) has the nice property that it belongs to a relatively small set of assortments. That property is not assured without this assumption.

[3] Because the demand rate has been normalized to 1, the cost function should reflect this normalization.

search model with overlapping assortment.

The traditional no-search model assumes the following consumer choice process: the consumer visits the retailer and observes the utility realizations for the no-purchase variant and each variant in the assortment S; if the maximum utility of the variants in the assortment, $U_{\max} = \max_{i \in S} U_i$ is less than the no-purchase utility, U_0, then the consumer purchases nothing; otherwise, the consumer purchases the most preferred variant in the assortment. Note that the consumer does not consider the possibility of continuing her search at some other retailer to find a product variant with an even higher utility.

The result for the traditional MNL model is well known (Chapter 2 in Anderson, de Palma and Thisse 1992):

$$q_i^m(S) = \frac{v_i}{\sum_{j \in S} v_j + v_0}, \text{ for } i = 0 \text{ and } i \in S, \tag{2}$$

where $v_j = \exp((u_j - p_j)/\mu)$ is referred to variant j's "preference".

Now consider the two search models. As already mentioned, search is influenced by the consumers' cost of search as well as their expectation regarding the value of search. In the construction of a consumer search model there is clearly a choice to be made between a rich model and a parsimonious model. A rich model reflects many possible nuances involved in the consumer search process. For example, there surely is heterogeneity in consumer search costs and a consumer's expectation regarding the value of search can be quite complex: it could depend on the number of previous retailers the consumer visited, whether the consumer can return to the current retailer (which we refer to as recall) and the consumer's belief regarding how many new variants and which variants the consumer could sample via search. We believe that incorporating all of these features into a search model would render the model analytical intractable and unimplementable (e.g., it would not be clear how all of the necessary input parameters could be empirically estimated in practice). Hence, we construct two parsimonious models that we hope capture the first order effects introduced by the presence of consumer search. The models represent two extreme interpretations regarding a consumer's expectation for the value of search.

With the independent assortment model a consumer expects to receive utility $U_r = u_r + \zeta_r$ if the consumer chooses to search, where u_r is a constant common to all consumers and ζ_r is a zero mean Gumbel random variable.[4] A consumer's utility cost for search is b. Therefore, the consumer choice process is as follows: a consumer observes the realizations of the utility of the S variants in the retailer's assortment and the no-purchase utility; the consumer surely does not purchase a variant from the retailer if $U_{\max} < U_0$; if $U_{\max} > U_0$, then the consumer either purchases her most preferred variant from the retailer or the consumer chooses to search (thereby incurring the search cost b). From the retailer's perspective the no-purchase variant is equivalent to the search option: in either case the consumer does not purchase from the retailer. From the consumer's perspective they are different: with the search option the consumer decides to forgo an acceptable variant ($U_{\max} > U_0$) for the chance of earning an even higher utility with search. ($U_{\max}$ is observed by the consumer when choosing whether to search, but the realization of U_r is not yet observed, so a realization $U_r < U_{\max}$ is possible.)

Theorem 1 *The demand for variant i in the independent assortment search model is*

$$q_i^{si}(S) = q_i^m(S)\left(1 - H(\bar{U}, S)\right), \text{ for } i \in S, \tag{3}$$

where $\bar{U} = u_r - b$, $H(\bar{U}, S) = \exp\left(-\lambda\left(v_0 + \sum_{j\in S} v_j\right)\right)$ *and* $\lambda = \exp\left[-\left(\bar{U}/\mu + \gamma\right)\right]$.

Proof. For notational convenience, let $y = U_{\max}$. Search is worthwhile if and only if,

$$\int_{y-u_r}^{\infty} (u_r + x - y) f(x)dx - \int_{-\infty}^{y-u_r} (y - u_r - x) f(x)dx \geq b,$$

where the first term is the expected gain from search, the second term is the expected loss from search and the third term is the cost of search. After rearranging terms, and recognizing $E[\zeta_r] = 0$, the above simplifies to $u_r - b \geq y$: a consumer benefits from search if $U_{\max} < \bar{U}$ where $\bar{U} = u_r - b$.

It remains to evaluate $q_i^{si}(S)$. Variant i is chosen if $U_i = U_{\max}$ (it is the best variant in the assortment), $U_i > \bar{U}$ (its utility is above the search threshold) and $U_i > U_0$ (its utility is better than no-purchase). Let ϕ be the realization of ζ_i. Thus, variant i is chosen if $\phi \geq \bar{U} - (u_i - p_i)$ ($U_i > \bar{U}$) and $U_i > \max\{U_{\max}, U_0\}$, which has probability

[4] There is an alternative interpretation for this model. Suppose search allows the consumer to sample from another set of possible variants, S', with the utility of each variant in S' having a Gumbel distribution. Then the maximum utility from this set is also a Gumbel distribution which can be modeled as having a base net utility u_r.

$\Pi_{j=0,j\in S/i}F\left((u_i-p_i)+\phi-(u_j-p_j)\right)$. Overall, the probability variant i is selected, $q_i^{si}(S)$, is

$$q_i^{si}(S)=\int_{\bar{U}-(u_i-p_i)}^{\infty}f(\phi)\prod_{\substack{j=0\\ j\in S/i}}F\left((u_i-p_i)+\phi-(u_j-p_j)\right)d\phi.$$

Following the process of deriving choice probabilities in the traditional MNL, substitute the CDF of the Gumbel distribution and conduct the change of variables $\delta=\exp\left[-\left(\phi/\mu+\gamma\right)\right]$ and $v_j=\exp((u_j-p_j)/\mu)$ to obtain,

$$q_i^{si}(S)=\int_0^{\alpha}\exp\left[-\delta\left(\frac{v_0+\sum_{j\in S}v_j}{v_i}\right)\right]d\delta$$

where $\alpha=\exp\left[-\left(\left(\bar{U}-u_i\right)/\mu+\gamma\right)\right]$. The above integral simplifies to

$$\begin{aligned} q_i^{si}(S) &= \frac{v_i}{v_0+\sum_{j\in S}v_j}\left\{1-\exp\left[-\left(v_0+\sum\nolimits_{j\in S}v_j\right)\exp\left[-\left(\bar{U}/\mu+\gamma\right)\right]\right]\right\} \\ &= q_i^m(S)\left(1-H(\bar{U},S)\right). \end{aligned}$$

■

According to Theorem 1, the demand for variant i with the independent search model is a fixed fraction, $\left(1-H(\bar{U},S)\right)$, of the demand for variant i without search, $q_i^m(S)$. Furthermore, the search threshold, $\bar{U}$, is independent of the store's current assortment, S, i.e., expanding the retailer's assortment has no impact on the threshold utility to prevent search because the retailer's assortment does not impact the consumer's expectation of the value of search.

From (2) we see that with the no-search model there exists a cannibalization effect: variant i's demand decreases as additional variants are added to the assortment. However, it is not so clear that the cannibalization effect always exists in the independent assortment model: while $q_i^m(S)$ decreases as the assortment expands, the search adjustment factor, $(1-H(\bar{U},S))$, increases as the assortment expands.[5] According to the next theorem, the former dominates the latter, so the cannibalization effect does exist even in the independent assortment model.

Theorem 2 *For all S and S^+ such that $i\in S\subset S^+$, $q_i^{si}(S)>q_i^{si}(S^+)$: variant i's demand decreases as the assortment is expanded in the independent assortment model.*

---

[5] The maximum utility from the assortment is stochastically increasing in the breadth of the assortment, so it is more likely that the threshold $\bar{U}$ is exceeded.

Proof. Define function $T(\omega) = \omega\left(1 - \exp\left(-\frac{\lambda}{\omega}\right)\right)$ where λ is a positive constant. The first derivative of $T(\omega)$ is,

$$T'(\omega) = 1 - \left(1 + \frac{\lambda}{\omega}\right)\exp\left(-\frac{\lambda}{\omega}\right).$$

$T(\omega)$ is an increasing function on the interval $[0, \infty)$ because it follows from the Taylor expansion of $\exp(\lambda/\omega)$ that $1 + \lambda/\omega < \exp(\lambda/\omega)$ for $\omega \geq 0$ and $\lambda \geq 0$. From (3), we have

$$q_i^{si}(S) = v_i T\left(\left(\sum_{k \in S} v_k + v_0\right)^{-1}\right).$$

Because $\sum_{k \in S} v_k + v_0 < \sum_{k \in S^+} v_k + v_0$ and $T(\omega)$ is increasing, it follows that $q_i^{si}(S) > q_i^{si}(S^+)$.■

In our second search model we imagine a market with a limited number of potential variants, e.g., digital cameras. As a result, the value of search to a consumer may very well depend on the retailer's assortment: as the retailer deepens his assortment a consumer may lower her expected value of search because search provides value only if new variants are discovered. To be specific, define $\bar{S}$ to be the variants outside the retailer's assortment, $\bar{S} = N - S$. The consumer choice process is now as follows: the consumer observes the no-purchase utility, U_0, and the utility for each variant in the retailer's assortment; the consumer surely does not purchase if $U_0 > U_{\max}$; if $U_{\max} > U_0$ then the consumer either purchases the highest utility variant in S or chooses to search. If search is chosen the consumer incurs the cost b but then has the opportunity to purchase the best variant in the entire set N, i.e., the consumer finds another retailer that has an assortment which includes both S and $\bar{S}$. In this model the retailer is effectively competing against some full assortment retailer (or a combination of retailers that carry the full assortment). As in the independent assortment model, there is a single search cost, b, for all consumers.

We refer to the second search model as the overlapping assortment model because the retailer's assortment overlaps with the assortment of the outside alternative. To the extent that the independent assortment model represents a worst case for the search option (the consumer can end up with a less desirable variant if search is chosen), this model represents the best case for the search option (the consumer can only find a more desirable product by searching). We suspect other search models are qualitatively a mixture of these two extremes, and they are analytically more cumbersome.[6]

[6] For example, the following search model creates a combinatorical mess: if a consumer

Theorem 3 *Variant i's demand in the overlapping assortment model is*

$$q_i^{so}(S) = q_i^m(S)\left(1 - H(\bar{U}(S), S)\right), \text{ for } i \in S,$$

where H is defined as in Theorem 1, and $\bar{U}(S)$ is the unique solution to

$$\int_{\bar{U}(S)}^{\infty} (\bar{y} - \bar{U}(S)) w(\bar{y}, S) d\bar{y} = b, \tag{4}$$

where $w(\bar{y}, S)$ is the density function of $\max_{i \in \bar{S}} U_i$ (the maximum utility observed in set $\bar{S} = N - S$).

Proof. Let $\bar{y}$ be the realized maximum utility from the set $\bar{S}$ and let $w(\bar{y}, S)$ be its density function. (Because the realizations of ζ_i are independent for the products in $\bar{S}$, it is straightforward to evaluate $w(\bar{y}, S)$). Again, for notational convenience let $y = U_{\max}$. Thus, the consumer searches if

$$\int_{y}^{\infty} (\bar{y} - y) w(\bar{y}, S) d\bar{y} \geq b : \tag{5}$$

the consumer is assured of at least y utility, so the first term is the expected incremental gain over y from search. The left hand side of (5) is decreasing in y, so there exists a unique threshold utility, $\bar{U}(S)$, such that

$$\int_{\bar{U}(S)}^{\infty} (\bar{y} - \bar{U}(S)) w(\bar{y}, S) d\bar{y} = b.$$

A consumer searches if and only if $U_{\max}$ is less than the threshold $\bar{U}(S)$. Note that the same holds in the independent assortment model except the threshold is independent of the assortment in that model. Therefore, the analysis to determine $q_i^{so}(S)$ follows the approach in Theorem 1 to determine $q_i^{si}(S)$.$\blacksquare$

As in the independent assortment model, variant i's demand in the overlapping assortment equals a fraction of the demand predicted by the MNL model, $q_i^m(S)$. However, the key difference between the two search models is that the search utility threshold in the independent assortment model is fixed, $\bar{U}$, whereas it depends on the assortment in the overlapping

searches then she observes the variants in set S'', where S'' only partially overlaps with set S (hence, the consumer might not find her most preferred item in set S) and only partially overlaps with set $\bar{S}$ (even after search there may be variants that are not observed). This model also complicates the assortment decision, not only must a retailer decide whether to include a variant in the assortment based on the variant's utility, the retailer must consider whether the variant is carried by other retailers and whether that fact is correctly included in consumers' expectations.

assortment model, $\bar{U}(S)$. As a result, although the cannibalization effect exists in the no-search and in the independent assortment models, it can be shown that the cannibalization effect does not always exist in the overlapping search model: given $i \in S$ and $S \subset S^{+}$, it is possible that variant i's demand *increases* as the assortment expands from S to S^{+}, i.e., $q_i^{so}(S) < q_i^{so}(S^{+})$. In that case, not only does an expanded assortment generate incremental revenue from the added variants, the added variants actually can increase the demand and efficiency of the variants already in the assortment.

4 The Optimal Assortment

The retailer's problem is to find the assortment that maximizes expected profit. To better understand this problem, consider how profit is impacted by adding variant j to an assortment S to create assortment $S_j = S \cup \{j\}$. On the positive side, variant j generates incremental demand. Let $q_i(v_j)$ be the demand for variant i when variant j is added to the assortment S. (Note that $q_i(S)$ is variant i's demand with assortment S and $q_i(v_j)$ is variant i's demand with assortment S_j.) Hence, $q_j(v_j)$ is variant j's incremental demand. Hopefully the addition of variant j generates a positive incremental profit. Let $\pi_i(v_j)$ be the profit of product i when variant j is added to the assortment

$$\pi_i(v_j) = m_i q_i(v_j) - c(q_i(v_j)).$$

(As with variant i's demand rate, we use the notation $\pi_i(v_j)$ for variant i's profit with assortment S_j, in contrast to $\pi_i(S)$ which is variant i's profit with assortment S.) However, a potential negative with the introduction of variant j is the reduction in demand for each of the variants in assortment S, the cannibalization effect, thereby reducing the profit on those variants for two reasons: they have less revenue and they are less cost efficient. Define $L(v_j)$ as the change in the profit of the variants in S when variant j is introduced:

$$L(v_j) = \sum_{i \in S} \pi_i(S) - \sum_{i \in S} \pi_i(v_j)$$

If $\pi_j(v_j) > L(v_j)$, then adding variant j to the assortment S would increase the retailer's profit.

The number of possible assortments, $2^N - 1$, increases rapidly in N. As a result, it is desirable to avoid full enumeration in the hunt for the optimal assortment. Of course, some assortments are intuitively more reasonable than others: e.g., one suspects that the

x variants, $x \leq n$, with the highest net utility would be preferred over the x variants with the lowest net utility. Indeed, it may seem reasonable that the optimal assortment with x variants should be the assortment with the x highest net utility variants. Unfortunately, that is not true in general: due to the cannibalization effect, conditional on adding one more variant to the assortment, it may be better to add an unpopular variant (which does not reduce the demand of the variants in the current assortment too much) rather than a popular variant. However, we show that in the no-search model and in the independent assortment model the following statement is true: if an assortment with x variants is optimal, then the assortment contains the x highest net utility variants. In that case, the optimal assortment is included in the following set, which we will call the popular assortment set:

$$P = \{\{\}, \{1\}, \{1, 2\}, ..., \{1, ..., n\}\},$$

where recall the variants are labeled in decreasing order of net utility ($u_i - p_i$ is decreasing in i) Unfortunately, that nice property does not hold in general with the overlapping assortment model.

4.1 Traditional MNL

van Ryzin and Mahajan (1999) consider the assortment planning problem in the no-search model with a particular cost structure derived from the newsvendor model. They find that the optimal assortment is indeed in P (the set of popular variant assortments). They obtain this result by showing that for *any* assortment S (which need not include the most popular variants) the retailer's profit function is quasi-convex in the net-utility of the variant added to S. Therefore, if a variant is added to assortment S then it should be the most popular variant (highest net utility), but a variant is only added if it increases the retailer's total profit. As a result, if an assortment with x variants is optimal, then it must include the x most popular variants: if it is profitable to add v_k to the assortment, $v_k < v_j$, then, because the retailer's profit function is quasi-convex in the net utility of the variant added to the assortment, it is even better to add v_j to the assortment. This section generalizes their results to include any concave increasing cost function.

Theorem 4 *The function $h^m(v_j) = \pi_j^m(v_j) - L^m(v_j)$ is quasi-convex in v_j on the interval $[0, \infty)$.*

Proof. Let $V_S = v_0 + \sum_{i \in S} v_i$. From differentiation,

$$h^{m\prime}(v_j) = \frac{\left[mV_S - c'\left(\frac{v_j}{v_j+V_S}\right)V_S\right] - \left[m\sum_{i\in S} v_i - \sum_{i\in S} c'\left(\frac{v_i}{v_j+V_S}\right)v_i\right]}{(v_j+V_S)^2} \tag{6}$$

Because $v_j/(v_j+V_S)$ is increasing in v_j and $c(\cdot)$ is concave, the first bracketed term in the numerator of (6) is increasing in v_j. Because $v_i/(v_j+V_S)$ is decreasing in v_j, the second bracketed term in the numerator of (6) is decreasing in v_j. Therefore, the numerator is increasing in v_j. From $(v_j + V_S^2) > 0$, the sign of $h^{m\prime}(v_j)$ is determined by the sign of the numerator in (6). Because the numerator is increasing, there is at most one v_j such that $h^{m\prime}(v_j) = 0$. Hence, $h^m(v_j)$ is quasi-convex in v_j.■

4.2 Independent assortment

In the independent assortment model a consumer's search threshold, $\bar{U}$, is independent of the assortment, but the adjustment factor, $1-H(\bar{U},S)$, is not independent of S: as a variant is added to the assortment the adjustment factor increases, i.e., consumers are less likely to continue search if their most preferred variant is chosen from a deeper assortment. While this adds a significant analytical complication, it nevertheless continues to hold that the retailer's profit is quasi-convex in the net utility of a variant added to any assortment S. Hence, the optimal assortment is still within the set of popular assortments, P.

Theorem 5 *The function $h^{si}(v_j) = \pi_j^{si}(v_j) - L^{si}(v_j)$ is quasi-convex in v_j on the interval $[0,\infty)$.*

Proof. Differentiate $h^{si}(v_j)$,

$$\begin{aligned} h^{si\prime}(v_j) &= mq_j^{si\prime}(v_j) - c'(q_j^{si}(v_j))q_j^{si\prime}(v_j) + \sum_{i\in S}\left[mq_i^{si\prime}(v_j) - c'(q_i^{si}(v_j))q_i^{si\prime}(v_j)\right] \qquad (7) \\ &= J(v_j)f(v_j) + N(v_j)g(v_j) \end{aligned}$$

where

$$\begin{aligned} \bar{H}(v_j) &= 1 - H(\bar{U}, S_j), \\ V_S &= v_0 + \sum_{i\in S} v_i, \\ J(v_j) &= \frac{\bar{H}(v_j)V_S + \bar{H}'(v_j)v_j\,(v_j+V_S)}{(v_j+V_S)^2}, \end{aligned}$$

$$
\begin{aligned}
f(v_j) &= m - c'\left(\frac{\bar{H}(v_j)v_j}{v_j+V_S}\right), \\
N(v_j) &= \frac{\bar{H}'(v_j)\,(v_j+V_S) - \bar{H}(v_j)}{(v_j+V_S)^2}, \\
g(v_j) &= m\sum\nolimits_{i\in S} v_i - \sum\nolimits_{i\in S} c'\left(\frac{\bar{H}(v_j)v_i}{v_j+V_S}\right)v_i.
\end{aligned}
$$

We wish to show that there exists at most one v_j such that $h^{si\prime}(v_j)=0$. $J(v_j)>0$ because each term is positive. It can be shown that $N(v_j)<0$, $\forall v_j>0$. By a similar argument as in Theorem 4, $f(v_j)$ is increasing and $g(v_j)$ is decreasing. Unfortunately, neither J nor N are monotone in v_j, so we need additional results to prove quasi-convexity.

Because $J(v_j)>0$ and $N(v_j)<0$, if $h^{si\prime}(v_j)=0$ then it must be that $f(v_j)$ and $g(v_j)$ have the same sign: either $f(v_j)<0$ and $g(v_j)<0$, or $f(v_j)>0$ and $g(v_j)>0$. The following statement is stronger: for all $\{v_j^1,v_j^2\}$ such that $v_j^1<v_j^2$ and $h'(v_j^1)=h(v_j^2)=0$, it holds that $f(v_j^1)f(v_j^2)>0$, i.e., it is never the case that $f(v_j^1)$ and $f(v_j^2)$ have different signs. That statement is validated by contradiction. Suppose $f(v_j^1)<0$ and $g(v_j^1)<0$, while $f(v_j^2)>0$ and $g(v_j^2)>0$: in that case $g(v_j^2)>g(v_j^1)$, which contradicts $g(v_j)$ is decreasing. Suppose $f(v_j^1)>0$ and $g(v_j^1)>0$, while $f(v_j^2)<0$ and $g(v_j^2)<0$: in that case $f(v_j^2)<f(v_j^1)$, which contradicts $f(v_j)$ is increasing.

Given that the sign of $f(v_j)$ is the same for all v_j such that $h^{si\prime}(v_j)=0$, there are two cases to consider to prove that there exists at most one v_j such that $h^{si\prime}(v_j)=0$: either $f(v_j)<0$ or $f(v_j)>0$.

Case (1): $f(v_j)<0$ and $g(v_j)<0$. The proof for this case applies the same technique as in Theorem 4: we rewrite $h^{si\prime}(v_j)$ as the product of a positive term and an increasing term. Unfortunately, this approach does not work with the next case, $f(v_j)>0$.

We have

$$
\begin{aligned}
h^{si\prime}(v_j) &= \frac{e^{-\lambda(v_j+V_S)}}{(v_j+V_S)^2}\left[D(v_j)f(v_j) - K(v_j)g(v_j)\right] && (8)\\
&= \frac{e^{-\lambda(v_j+V_S)}}{(v_j+V_S)^2}D(v_j)\left(-f(v_j)\right)\left[\frac{g(v_j)}{f(v_j)}\frac{K(v_j)}{D(v_j)} - 1\right], && (9)
\end{aligned}
$$

where

$$
\begin{aligned}
D(v_j) &= e^{\lambda(v_j+V_S)} - 1 + \frac{\lambda v_j}{V_S}\left(v_j+V_S\right) \\
K(v_j) &= e^{\lambda(v_j+V_S)} - 1 - \lambda\left(v_j+V_S\right)
\end{aligned}
$$

Both $D(v_j)$ and $K(v_j)$ are positive and increasing in v_j. Furthermore, $K(v_j)/D(v_j)$ is increasing in v_j. Let $\overline{v}_j$ be the smallest v_j such that $h^{si\prime}(v_j) = 0$. Then $g(v_j)/f(v_j)$ is increasing in v_j for all $v_j \geq \overline{v}_j$ such that $f(v_j) < 0$ and $g(v_j) < 0$ (because $g(v_j)$ is decreasing and $f(v_j)$ is increasing in v_j). Therefore,

$$\frac{g(v_j)}{f(v_j)}\frac{K(v_j)}{D(v_j)} - 1$$

is increasing for all $v_j \geq \overline{v}_j$ such that $f(v_j) < 0$ and $g(v_j) < 0$. Furthermore,

$$D(v_j)f(v_j) - K(v_j)g(v_j) > 0$$

for any $v_j \geq \overline{v}_j$ such that $f(v_j) > 0$ and $g(v_j) < 0$. Hence, there does not exist a $v_j \geq \overline{v}_j$ such that $h^{si\prime}(v_j) = 0$.

Case (2): $f(v_j) \geq 0$ and $g(v_j) \geq 0$. We are unable to rewrite (8) in a multiplicative form of a positive term and an increasing term in this case (since $-f(v_j) < 0$). However, our desired result follows if it can be shown that for all v_j such that $h^{si\prime}(v_j) = 0$, then $h^{si\prime\prime}(v_j) > 0 :$ if $h^{si\prime}(v_j)$ is increasing for every v_j such that $h^{si\prime}(v_j) = 0$, then there can be at most a single v_j such that $h^{si\prime}(v_j) = 0$. Differentiate:

$$\begin{aligned} h^{si\prime\prime}(v_j) &= mq_j^{si\prime\prime}(v_j) - c''(q_j^{si}(v_j))\left(q_j^{si\prime}(v_j)\right)^2 - c'(q_j^{si}(v_j))q_j^{si\prime\prime}(v_j) \\ &\quad + \sum\nolimits_{i\in S}\left[mq_i^{si\prime\prime}(v_j) - c''(q_i^{si}(v_j))\left(q_i^{si\prime}(v_j)\right)^2 - c'(q_i^{si}(v_j))q_i^{si\prime\prime}(v_j)\right]. \end{aligned}$$

Because $c(\cdot)$ is concave, $c''(q_j^{si}(v_j))\left(q_j^{si\prime}(v_j)\right)^2 \leq 0$ and $c''(q_i^{si}(v_j))\left(q_i^{si\prime}(v_j)\right)^2 \leq 0$. It follows that

$$\begin{aligned} h^{si\prime\prime}(v_j) &\geq mq_j^{si\prime\prime}(v_j) - c'(q_j^{si}(v_j))q_j^{si\prime\prime}(v_j) + \\ &\quad \sum\nolimits_{i\in S}\left[mq_i^{si\prime\prime}(v_j) - c'(q_i^{si}(v_j))q_i^{si\prime\prime}(v_j)\right] = \tilde{J}(v_j)f(v_j) + \tilde{N}(v_j)g(v_j) \end{aligned}$$

where

$$\begin{aligned} \tilde{J}(v_j) &= \frac{2V_S\left[\bar{H}'(v_j)\left(v_j + V_S\right) - \bar{H}(v_j)\right] + \bar{H}''(v_j)v_j\left(v_j + V_S\right)^2}{\left(v_j + V_S\right)^3}, \\ \tilde{N}(v_j) &= \frac{\left[\bar{H}''(v_j)\left(v_j + V_S\right)^2 - 2\bar{H}'(v_j)\left(v_j + V_S\right) + 2\bar{H}(v_j)\right]}{\left(v_j + V_S\right)^3}. \end{aligned}$$

Thus, it is sufficient to show that

$$\widetilde{J}(v_j)f(v_j) + \tilde{N}(v_j)g(v_j) > 0. \tag{10}$$

From $\tilde{J}(v_j) \leq 0$ (because $\bar{H}''(v_j) \leq 0$), $f(v_j) \geq 0$, $g(v_j) \geq 0$ and

$$h^{si\prime}(v_j) = J(v_j)f(v_j) + N(v_j)g(v_j) = 0,$$

(10) holds if

$$\frac{-\tilde{N}(v_j)}{\tilde{J}(v_j)} > \frac{-N(v_j)}{J(v_j)}. \tag{11}$$

Simplifying and rearranging terms in (11) yields

$$2 - 2\exp[-\lambda(v_j + V_S)] - \lambda(v_j + V_S)\exp[-\lambda(v_j + V_S)] - \lambda(v_j + V_S) < 0, \tag{12}$$

which holds for all $v_j + V_s > 0$.■

4.3 Overlapping Assortment

The overlapping assortment model is more complex than the independent assortment model because now the consumer's search threshold, $\bar{U}(S)$, is decreasing in the depth of the assortment. As a result, it is no longer always true that the retailer's profit is quasi-convex in the net utility of an added variant. This occurs because the cannibalization effect is not always present in the overlapping search model: it may be better to add an unpopular variant to reduce search without causing much cannibalization of existing demand than it is to add a popular variant. Hence, the optimal assortment is not guaranteed to be within the set of popular assortments, P. While we did find a scenario in which the optimal assortment is outside P, that did not occur in any of the scenarios in our extensive numerical study, which suggests that restricting the search for the optimal assortment to P is reasonable for a wide range of parameters.[7]

The optimal assortment in this model differs from the other two models qualitatively in two other respects. First, in the no-search and independent assortment models every variant in the optimal assortment must earn a strictly positive profit: due to the cannibalization effect, an assortment's total profit increases if all non-positive profit variants in the assortment are removed. The same does not hold in the overlapping assortment model, i.e., it may be optimal to have negative profit variants in the assortment so as to decrease consumer search.

Second, in the overlapping assortment model, a retailer has the possibility to carry the full assortment (i.e., all n potential variants), which allows him to prevent consumer search entirely. In fact, according to the following theorem, there exists a sufficiently low search

[7] The parameters with this scenario include $m_1 = m_2 = m_3 = 2$, $u_0 = 3.4$, $u_1 = 4.1$, $u_2 = 0.7$, $u_3 = 0.6$, $b = 0.122$ and the cost function is $c(q) = 0.5q^{1/2}$. The optimal assortment is $\{1, 3\}$.

cost such that the full assortment is optimal for the retailer. In such cases, the profitability of an individual variant is dominated by the retailer's desire to prevent consumer search.

Theorem 6 *In the overlapping assortment model, for any given set of preference, $\{v_0, v_1, ..., v_n\}$, if the full assortment (all n variants) yields a positive profit, then there exists a threshold consumer search cost $\bar{b}$ such that the full assortment is optimal for all $b \leq \bar{b}$.*

Proof. We need to show that every assortment $S \neq \{1, 2, ..., n\}$ leads to less profit than the full assortment, which is assumed to generate a positive profit. The expected profit of assortment S is

$$\begin{aligned} \pi(S) &= \sum_{i \in S} \left[m_i q_i^{so}(S) - c(q_i^{so}(S)) \right] \\ &= \sum_{i \in S} \left[m_i q_i^{so}(S) \left(1 - H(\bar{U}(S), S)\right) - c\left(q_i^{so}(S)\left(1 - H(\bar{U}(S), S)\right)\right) \right] \end{aligned} \tag{13}$$

From the search rule specified in Theorem 3, the search threshold $\bar{U}(S)$ is monotonically decreasing in the search cost b. Furthermore, $\lim_{b \to \infty} \bar{U}(S) = -\infty$ and $\lim_{b \to 0} \bar{U}(S) = \infty$. Therefore, from (13), there exists a sufficiently low b such that $H(\bar{U}(S), S)$ is close enough to 1 to ensure that $\pi(S)$ is less than the full assortment profit for any S.■

5 Implementation of the No-Search Model

Implementation with an assortment planning model begins with the estimation of the model's necessary input parameters. In the case of the traditional no-search model these include each variant's preference (v_i), margin (m_i), and the operational cost function, $c(\cdot)$. An advantage of the no-search model is that it does not require the retailer to estimate any search parameters, which may be one reason why a retailer may choose the no-search model for assortment planning. Alternatively, a retailer may adopt the no-search model because she is unaware of how to incorporate consumer search into the assortment planning process. No matter the reason, given the empirical and theoretical parsimony of the no-search model, the objective of this section is to determine the extent to which the traditional no-search model yields good assortment decisions even if search does influence consumer choice. In other words, we want to identify conditions in which the no-search model is robust to mis-specifications of the consumer choice process. The first subsection considers this issue analytically and the subsequent subsection considers it numerically.

5.1 Heuristic Equilibrium with the No-Search Model

Suppose a retailer implements the traditional no-search model for assortment planning even though consumer choice is influenced by search. To obtain analytical results we assume consumer choice is actually described by either the independent assortment model or the overlapping assortment model.

As determined earlier, according to the no-search model the optimal assortment is contained within the popular assortment set, P. As a result, it is quite useful for the retailer to know the ranking of the variants, i.e., which is variant 1 (the most popular variant), which is variant 2 (the second most popular variant), etc. If the ranking of the variants is known, then the assortment decision can be simplified from "which variants" to "how many variants": if x variants are chosen then they are the x most popular variants, $\{1, 2, ..., x\}$. Based on this observation, we hereafter simplify notation by using x to represent S, where $S = \{1, 2, ...x\}$.

We assume the retailer begins the assortment planning process with a "depth-test": the full assortment is placed in the store and sales data are collected to be used to estimate each variant's preference. After the depth-test the product variants are labeled from most popular to least popular. Variant 0 is always the no-purchase variant even if $v_0 < v_1$. We assume the depth-test is sufficiently accurate to identify the relative ranking of the products without error.

Using the data from the depth-test the retailer then adopts the optimal assortment according to the no-search model. Additional sales data are collected and then a new assortment may be chosen. This iterative process of assortment planning stops when the retailer obtains an assortment that is optimal (again, according to the no-search model) given the inputted preferences, and the sales data observed with the assortment are consistent with the inputted preferences.

To formalize this iterative assortment planning process, let $d_i(x)$ be variant i's observed demand with assortment $x \in [0, n]$, where the demands have been normalized so that

$$\sum_{i=0}^{x} d_i(x) = 1. \tag{14}$$

Note that the no-purchase variant, 0, is included in the summation in (14). Let $\hat{v}_i(x)$ be the estimate of variant i's preference using the sales data from assortment x. The "^" superscript is used to emphasize that $\hat{v}_i(x)$ is an estimate of variant i's preference in contrast to variant i's actual preference, v_i, which is not directly observable to the retailer. The retailer obtains $\hat{v}_i(x)$ by solving a system of $x+1$ equations which includes (14) and the following x equations:

$$d_i(x) = \frac{\hat{v}_i(x)}{\sum_{j=0}^{x} \hat{v}_j(x)}, \qquad i \in [0, x]$$

The solution to this set of linear equations is a one dimensional space as opposed to a single point. Hence, one variant's preference should be fixed to some arbitrary constant. Without loss of generality, fix the preference for variant 1, say $\hat{v}_1(x) = 1$.

In practice there could be sampling error in demand, hence $d_i(x)$ may not exactly equal expected demand with assortment x, $q_i^{si}(x)$ or $q_i^{so}(x)$. Therefore, sampling error would propagate into the estimate of each variant's preference, and in turn that error could influence the assortment decision. While that is an interesting issue (for example, one could imagine choosing the duration of the data collection period as an optimization problem), we shall bypass this issue by assuming $d_i(x)$ indeed does equal variant i's expected demand with assortment x. As a result, the retailer obtains the estimate $\hat{v}_i(x)$ every time the assortment x is implemented. This is a conservative assumption in the sense that it favors the performance of the iterative algorithm.

The assortment x may not be optimal according to the set of preferences estimated with the assortment x. Therefore, the retailer will wish to make an adjustment to the assortment. Let $\hat{v}(x)$ be the set of preferences estimated with assortment x. We use the data from the depth-test to include in $\hat{v}(x)$ the preferences of the variants not included in $x : \hat{v}(x) = \{\hat{v}_0(x), ..., \hat{v}_x(x), \hat{v}_{x+1}(n), ..., \hat{v}_n(n)\}$. As described in footnote 8, this does not create an issue. Let $x(\hat{v})$ be a correspondence that returns the optimal assortment according to the no-search model given the inputted preferences. (Because the optimal assortment is in P, it is sufficient for $x(\hat{v})$ to return an integer in the range $[0, n]$.) Furthermore, let x^t be the t^{th} assortment chosen by the retailer, let $\hat{v}^t$ be the preference estimates from the t^{th} assortment and define the depth-test assortment as $t = 1$. Hence, the iterative assortment planning process begins with $x^1 = n$ and $\hat{v}^1 = \hat{v}(n)$. Subsequent iterations have $x^{t+1} = x(\hat{v}^t)$ and $\hat{v}^{t+1} = \hat{v}(x^t)$. A heuristic equilibrium of this process is an assortment-preference pair, $\{x^*, \hat{v}^*\}$, such that $x^* = x(\hat{v}^*)$ and $\hat{v}^* = \hat{v}(x^*)$: the assortment is optimal given the estimated preferences and the estimated preferences are observed given the assortment. If the iterative process reaches a heuristic equilibrium, then the retailer makes no further modification to the assortment. (See Cachon and Kok 2002 for a more general definition of a heuristic equilibrium, and see Armony and Plambeck 2002 for an application of a heuristic equilibrium in a different setting.) We are interested in whether the optimal assortment, x^o, is a heuristic equilibrium, and if it is not, then do the existing heuristic equilibria have too many or too few variants? Some preliminary results are needed before answering those questions.

A well known result from the no-search model is that the relative demand of any two variants, including the no-purchase variant, is independent of the assortment, i.e., if i and j are two different variants in an assortment, then $q_i^m(S)/q_j^m(S)$ is constant for all S that contain i and j. This is referred to as the Independence of Irrelevant Alternatives (IIA) property. The IIA property also holds in both the independent and overlapping assortment models between any two product variants,

$$\frac{q_i^{si}(x)}{q_j^{si}(x)} = \frac{v_i}{v_j}, \text{ for } i, j \neq 0,$$

$$\frac{q_i^{so}(x)}{q_j^{so}(x)} = \frac{v_i}{v_j}, \text{ for } i, j \neq 0.$$

Therefore, given that the preference of variant 1 is held fixed, iterative assortment planning correctly estimates the preferences of the product variants under any assortment, i.e., for all x and all $0 < i < j \leq n$,

$$\frac{\hat{v}_i(x)}{\hat{v}_j(x)} = \frac{v_i}{v_j}.$$

In other words, given our assumption that $d_i(x)$ equals expected demand with assortment x, the retailer is able to correctly infer from the data the relative preference of any two product variants.[8] If that result included the no-purchase variant, then the retailer would correctly infer the preference of all variants and the iterative assortment planning process would converge to the optimal assortment, i.e., $x^* = x^o$. But that is not the case.

Using the result from section 3, we see in the independent assortment model that the relative preference of the no-purchase variant to any of the product variants is not independent of the assortment:

$$\frac{q_0^{si}(x)}{q_i^{si}(x)} = \frac{q_0^m(x) + \sum_{j \leq x} q_j^m(x) H(\bar{U}, x)}{q_i^m(x)\left(1 - H(\bar{U}, x)\right)}.$$

---

[8] As a result, $\hat{v}_x(x) = \hat{v}_x(n)$, which is why $\hat{v}_{x+1}(n)$ can be included in $\hat{v}(x)$ without distorting the results.

and the same holds with the overlapping assortment model. In other words, both the independent and overlapping assortment model violate the IIA property with respect to the no-purchase variant. Hence, while the iterative no-search model correctly evaluates the relative preference between any two product variants, it does not evaluate correctly the relative preference of the no-purchase variant: even though the true no-purchase preference is fixed, given that the preference for variant 1 is fixed, $\hat{v}_0(x)$ depends on x. The next Theorem describes the bias caused by this result.

Theorem 7 *Take any assortment x and the preferences estimated from that assortment, $\hat{v}(x) = (\hat{v}_0(x), ..., \hat{v}_x(x), \hat{v}_{x+1}(n), ..., \hat{v}_n(n))$. With any narrower assortment x', i.e. $x' < x$, the no-search model under-estimates the no-purchase demand, $d_0(x') \geq q_0^m(x'|\hat{v}(x))$. With any deeper assortment x'', i.e. $x'' > x$, the no-search model over-estimates the no-purchase demand, $d_0(x'') \leq q_0^m(x''|\hat{v}(x))$.*

Proof. We first show that $d_0(x') \geq q_0^m(x'|\hat{v}(x))$ for the independent assortment. Hereafter, we abbreviate the notation $H(\bar{U}, x)$ to $H(x)$. Let $\widehat{v}_i(x')$ denote the estimation of v_i under assortment x', which are obtained from the following equations:

$$d_0(x') = \frac{\widehat{v}_0(x')}{\widehat{v}_0(x') + \sum_{j \leq x'} \widehat{v}_j(x')} \tag{15}$$

$$d_i(x') = \frac{\widehat{v}_i(x')}{\widehat{v}_0(x') + \sum_{j \leq x'} \widehat{v}_j(x')}, \text{ for } 0 < i \leq x'. \tag{16}$$

To establish that $d_0(x') \geq q_0^m(x'|\hat{v}(x))$, it is sufficient to show that $\frac{\widehat{v}_0(x')}{\widehat{v}_i(x')} \geq \frac{\widehat{v}_0(x)}{\widehat{v}_i(x)}$. Taking ratios of the left hand sides and right hand sides of (15) and (16) respectively, we have $\frac{\widehat{v}_0(x')}{\widehat{v}_i(x')} = \frac{d_0(x')}{d_i(x')} = \frac{q_0^m(x') + \sum_{j \leq x'} q_j^m(x')H(x')}{q_i^m(x')(1-H(x'))}$. Similarly, we have $\frac{\widehat{v}_0(x)}{\widehat{v}_i(x)} = \frac{q_0^m(x) + \sum_{j \leq x} q_j^m(x)H(x)}{q_i^m(x)(1-H(x))}$. Then, $\frac{\widehat{v}_0(x')}{\widehat{v}_i(x')} \geq \frac{\widehat{v}_0(x)}{\widehat{v}_i(x)}$ implies the following,

$$\frac{q_0^m(x') + \sum_{j \leq x'} q_j^m(x')H(x')}{q_i^m(x')(1 - H(x'))} \geq \frac{q_0^m(x) + \sum_{j \leq x} q_j^m(x)H(x)}{q_i^m(x)(1 - H(x))}.$$

Substituting (2) into $q_i^m(x')$ and $q_i^m(x)$ and the expressions for $H(.)$, and collecting terms, we obtain,

$$\frac{\left(\sum_{j=0}^{x'} v_j\right) \exp\left(-\lambda \sum_{j=0}^{x'} v_j\right)}{1 - \exp\left(-\lambda \sum_{j=0}^{x'} v_j\right)} \geq \frac{\left(\sum_{j=0}^{x} v_j\right) \exp\left(-\lambda \sum_{j=0}^{x} v_j\right)}{1 - \exp\left(-\lambda \sum_{j=0}^{x} v_j\right)}.$$

Define function $Z(\delta) = \frac{\delta \exp(-\lambda\delta)}{1-\exp(-\lambda\delta)}$. The above inequality implies that $Z(\sum_{j=0}^{x'} v_j) \geq Z(\sum_{j=0}^{x} v_j)$. Therefore, all we need to show is that $Z(\delta)$ is a decreasing function in δ. Notice that the first derivative of $Z(\delta)$ is:

$$Z'(\delta) = \frac{[1 - \lambda\delta - \exp(-\lambda\delta)] \exp(-\lambda\delta)}{[1 - \exp(-\lambda\delta)]^2}.$$

By Taylor expansion of $\exp(-\lambda\delta)$, we can see that $1 - \lambda\delta - \exp(-\lambda\delta) \leq 0$ for any $\lambda\delta \geq 0$. Therefore, we have established that $Z'(\delta) \leq 0$, and thus the fact that $Z(\delta)$ is a decreasing function in δ. This proves $d_0(x') \geq q_0^m(x'|\hat{v}(x))$. The above process can be applied to show $d_0(x'') \leq q_0^m(x''|\widehat{v}(x))$. The similar logic can also be applied to show the results for the overlapping assortment.■

From Theorem 7, the no-search model overestimates the retailer's demand with a narrower assortment and underestimates the retailer's demand with a deeper assortment. The retailer overestimates demand with a narrower assortment because the retailer does not account for consumer search behavior: consumer search is more likely as the assortment becomes narrower but that behavior is not reflected in the estimates of the preferences. Given Theorem 7, the result in the next Theorem is expected: the iterative application of the no-search model results in an assortment that is never deeper than the optimal assortment.

Theorem 8 *The heuristic equilibrium resulting from the iterative application of the no-search assortment planning model never contains more variants than the optimal assortment, i.e.,* $x^* \leq x^o$.

Proof. We prove the result for the independent assortment first. From the definition of the optimal assortment, we have $\pi(x^o) = \max_{0 \leq x \leq n}\{\pi(x)\}$. For an "heuristic equilibrium" x^*, x^* is believed as the "optimal" assortment under $\widehat{v}(x^*)$, therefore, we have $\pi(x^*) = \pi(x^*|\widehat{v}(x^*)) = \max_{0 \leq x \leq n}\{\pi(x|\widehat{v}(x^*))\}$.

As we show in the previous section, the assortment x^* will not contain any variant with negative expected profit, i.e., $\pi_i(x^*) \geq 0$, for all $1 \leq i \leq x^*$. It implies that $\pi_i(x|\widehat{v}(x^*)) \geq 0$ at any narrower assortment $1 \leq x < x^*$ for all $1 \leq i \leq x$. In other words, based on $\widehat{v}(x^*)$, the expected profit of every variant contained in a narrower assortment than x^* is estimated to be positive, because the expected profit of a variant decreases with the depth of the assortment.

By the above Theorem, we know $d_0(x) \geq q_0^m(x|\widehat{v}(x^*))$, for $1 \leq x < x^*$. Since $\pi_i(x|\widehat{v}(x^*)) \geq$

0, for all $1 \leq i \leq x$, $d_0(x) \geq q_0^m(x|\widehat{v}(x^*))$ implies that the retailer over-estimates the expected profit of every variant contained in an assortment x that is narrower than x^*, i.e., $\pi_i(x) \leq \pi_i(x|\widehat{v}(x^*))$. Consequently, at x^*, the retailer over-estimates the expected profit of the assortment x that is narrower than x^*, i.e., $\pi(x) \leq \pi(x|\widehat{v}(x^*))$, for $1 \leq x < x^*$.

Suppose we have $x^* > x^o$. Based on the above discussion, at x^*, the retailer will over-estimate the expected profit of the assortment x^o, i.e., $\pi(x^o|\widehat{v}(x^*)) \geq \pi(x^o)$. Therefore, according to the definition of x^*, we have $\pi(x^*) = \pi(x^*|\widehat{v}(x^*)) = \max_{0 \leq x \leq n}\{\pi(x|\widehat{v}(x^*))\} \geq \pi(x^o|\widehat{v}(x^*)) \geq \pi(x^o)$, which is a contradiction to x^o being the optimal assortment. Hence, we have $x^* \leq x^o$. The similar process can be applied to show the result for the overlapping assortment model.■

If the retailer conducts assortment planning with the traditional model that does not explicitly incorporate consumer search behavior, then the retailer runs the risk of choosing an assortment that is narrower than optimal. Alternatively, the no-search model might suggest that x^o is too deep an assortment when in fact it is optimal.

5.2 A Numerical Study

We conducted a numerical study to explore the performance of the no-search assortment planning model in an environment with consumer search characterized either by the independent assortment model or the overlapping assortment model. We are interested in whether the heuristic equilibrium with the no-search model (assuming it exists and is unique) corresponds to the optimal assortment, and if not, what is the magnitude of the profit loss.

In all of the constructed scenarios there are eight potential variants, $n = 8$. Three functions are used to assign utilities: $u_i = 8 - \phi_1(1 - e^{-\phi_2(i-1)})$ for $\{\phi_1, \phi_2\} \in \{\{1, 1/2\}, \{3, 1/2\}, \{6, 1/2\}\}$. The rate of decline in the utilities is least with the first function and greatest with the last one, but with each function $u_1 = 8$. Three levels are used for the utility of the no-purchase variant: $u_0 = u_1 - 0.25(u_1 - u_8)$, $u_0 = u_1 - 0.75(u_1 - u_8)$, and $u_0 = 0.75u_8$. The first case has a relatively attractive no-purchase utility while in the latter case the no-purchase utility is even lower than the least attractive product variant. Nine different cost functions are constructed: $c(q) = \nu q^\beta$, with $\nu \in \{0.2, 0.5, 0.8\}$ and $\beta \in \{0.2, 0.5, 0.7\}$. These functions include newsvendor like cost functions as well as the EOQ cost function. The remaining parameters are the margins, $m \in \{3, 5, 7\}$, and the nine search costs $b \in \{0.1, 0.4, ..., 2.5\}$. The variants have identical prices, normalized to an arbitrary constant.

Even though we assume the no-search assortment planning model is implemented iteratively by the retailer, consumer choice is determined either as in the independent or overlapping assortment models. With the independent assortment model we choose three utilities for the outside alternative, $u_r \in \{u_1, u_1 - 0.5(u_1 - u_8), u_8\}$. Therefore, the combination of nine utility patterns for the product variants, the nine cost functions, 27 margin-search cost pairs and three outside alternative utilities yields a total of 6561 scenarios. There are no additional parameters needed with the overlapping assortment, so with that search model there are 2187 scenarios.

With each scenario we evaluated the optimal assortment according to the correct search model. With the independent assortment model we need only search among the set of popular assortments, P, to find the optimal assortment. However, with the overlapping assortment model we enumerated all possible assortments to search for the optimal one. In all scenarios the optimal assortment was nevertheless among the set of popular assortments, P.

With each scenario we identified the set of heuristic equilibria when the no-search assortment planning model is applied iteratively. With both independent and overlapping search there are some scenarios that do not have any heuristic equilibrium (198 and 230 respectively). Without a heuristic equilibrium the retailer would constantly adjust his assortment. With overlapping search we found 18 scenarios with two heuristic equilibria and none with more than two equilibria. In all cases the two equilibria differ only by one variant. Therefore, in the majority of scenarios there exists a unique heuristic equilibrium. Only those scenarios are considered in the following analysis.

Figures 1 and 2 display the heuristic equilibrium assortments with overlapping and independent search respectively. As predicted by Theorem 8, the heuristic equilibrium with independent search never chooses an assortment with more variants than optimal. Given that the optimal assortment with overlapping search is always among the set of popular assortments, the same results also holds with that type of search.

Excluding scenarios in which the optimal assortment has one or fewer variants, with independent search the heuristic equilibrium is optimal in about 75% of the scenarios, and with overlapping search that figure is 50%. Figures 3 and 4 display the percentage profit loss relative to the optimal profit of the heuristic equilibrium assortment with the overlapping and independent search models respectively. From these figures we conclude that the no-search assortment planning model performs well in essentially all of our scenarios with independent search: 75% of scenarios have no profit loss, the average profit loss is 0.29% and the maximum profit loss is 10.68%. However, ignoring search in the overlapping search model can be costly: the average profit loss is 10.26% and the maximum profit loss is 100%. These results are consistent with our analytical results which show that search has a greater influence on consumer choice in the overlapping model than in the independent model (Theorem 3). As a result, the gap between the estimate of the no-purchase utility, $\hat{v}_0(x)$, and the true no purchase utility, v_0, is reasonably small for any assortment depth in the independent search model but can be quite large with the overlapping search model, especially for assortments with only a few variants.

Figures 3 and 4 also illustrate that the profit loss of the heuristic equilibrium increases as search costs are reduced. While the direction of this effect is intuitive (as search becomes cheaper it has a great influence on consumer choice), the magnitude of the effect is surprisingly strong, in particular with overlapping search. In fact, with a sufficiently low search cost we observe several scenarios with a 100% profit loss: the heuristic equilibrium leads the retailer to close down the category (i.e., an assortment with no variants) even though there exists an assortment which could generate a positive profit for the retailer. This is consistent with Theorem 6: when search costs are sufficiently low the full assortment is optimal with overlapping search.

Figure 5 further illustrates how search can influence assortment planning. For the parameters listed with overlapping search, the optimal and the heuristic equilibria assortments are displayed for different search costs. As the search cost decreases the optimal assortment becomes deeper in order to prevent consumer search. But the heuristic equilibrium has completely the opposite response, it carries fewer variants. The heuristic equilibrium cuts back on the assortment because lower search costs manifest as a higher no-purchase utility, which makes the category look less attractive. But as the assortment is reduced further, the estimate of the no-purchase utility becomes even worse because the narrower assortment makes consumers search even more. With sufficiently low search costs the heuristic equilibrium closes the category even though the full assortment is optimal.

While we demonstrated that the cannibalization effect of a deeper assortment implies that every variant must make a strictly positive profit with independent search, we conjectured that the optimal assortment with overlapping search might include some money losing variants. In fact, we found 54 scenarios in which the optimal assortment with overlapping search included variants with a negative expected profit. Search costs were low in all these scenarios, so the retailer is willing to carry some money losing variants in order to prevent consumer search.

6 Estimation in the Presence of Search

Given the importance of consumer search and the resulting impact on the assortment planning problem established in the previous sections, a retailer needs to be able to estimate consumer search behavior. In this section, we first provide a warning indicator that retailers can use to determine to what extent they are affected by consumer search. We then present a procedure for estimating the search parameters.

In presence of consumer search, the ratio between the number of no purchase cases and any product variant i in the assortment, $\frac{d_0(x)}{d_i(x)}$, depends on the assortment depth x instead of being a constant as predicted by the IIA. Specifically, we can use Theorem 7 to see that the ratio $\frac{d_0(x)}{d_i(x)}$ is decreasing in x. Since the ratio $\frac{d_0(x)}{d_i(x)}$ can easily be estimated empirically, it provides an indicator for the retailer that his business is affected by consumer search. I.e., a retailer who observes that the "no purchase" demand under a narrower (deeper) assortment is higher (lower) than what was predicted from the assortment depth the sales data was taken from (e.g. a depth test) should treat the recommendations from an MNL based assortment model with suspicion.

Our analytical results can also be used by a retailer to estimate the parameters associated with consumer search. For both cases, independent and overlapping, the IIA holds for all variants, except the no-purchase option. Hence, fixing one preference v_i, say v_1, allows the retailer to estimate the remaining v_i, for $x \geq i > 1$ from sales data collected based on an assortment depth x. This leaves two parameters to be estimated. In both, the independent and the overlapping case, the retailer needs to estimate the no-purchase utility, v_0. For the independent case, the retailer also needs to estimate the search threshold, $\bar{U}$, while she needs to estimate the search cost, b, for the overlapping case.

Consider the overlapping case first. If the retailer offers the full assortment, there will be no consumer search. Thus, if the retailer conducts a depth test based on one full assortment, she is able to estimate all v_i for $i = 0, 1, ..., n$ directly. To estimate the consumer search cost b, the retailer needs to set up one store experiment based on an assortment $x < n$ and then collect the resulting consumer choice data. On average, there will be $d_0(x) = q_0^m(x) + \sum_{j \leq x} q_j^m(x) H(\bar{U}(x), x)$ percent of consumers who do not purchase. Since the v_is are already known, the retailer can directly compute the unknown variable $\bar{U}(x)$. Once $\bar{U}(x)$ is known, the retailer can use (4) to infer the consumer search cost b.

In the independent assortment model, the retailer needs to set up two store experiments based on assortments x_1 and x_2. Theoretically, there will be $d_0(x_1) = q_0^m(x_1) + \sum_{j \leq x_1} q_j^m(x_1) H(\bar{U}, x_1)$ and $d_0(x_2) = q_0^m(x_2) + \sum_{j \leq x_2} q_j^m(x_2) H(\bar{U}, x_2)$ percent of consumers who do purchase from these two respective assortments. Based on these two equations, the retailer can then infer the two unknown variables $\bar{U}$ and v_0.

In this paper, we have not considered the costs associated with conducting experiments in retail settings. Specifically, while being a common practice in retailing, setting up an experimental full-assortment store is costly. Moreover, we have assumed consumer preferences to remain constant and the estimators to be free of sampling error. We do point out that the presence of limited data availability as well as sampling errors will only amplify the importance of our analytical findings. In presence of sampling errors, a retailer is likely to interpret deviations between actual and predicted no-purchase incidents as noise, opposed to understanding the systemic bias outlined in Theorem 7.

7 Conclusions and Future Research

This paper develops three related assortment planning models. The traditional no-search model utilizes the multinomial logit model of consumer choice without explicitly considering consumer search. The other two models incorporate consumer search into the traditional model. The first search model, which we call the independent search model, presumes that there is an essentially unlimited pool of variants, which implies a low probability that the same product variant is carried in the assortment of two retailers. The second model, called the overlapping search model, presumes a limited pool of variants so that overlap between retailers is likely. The value of search to a consumer in the independent model is limited by the possibility of not finding an even better product if search is chosen. However, search is quite valuable to a consumer in the overlapping model because in that situation the consumer is never worse off with search. Hence, consumer choice is most sensitive to search in the overlapping model and least sensitive to search in the independent model.

With each of the three assortment models we derive the optimal assortment and present a procedure for estimating the necessary parameters to implement the models in practice. Nevertheless, a retailer may choose to use the no-search assortment planning model even though search influences consumer choice. We presume a retailer would apply the model iteratively: an assortment is chosen, then updated parameters are estimated using the sales data with that assortment, an updated assortment is chosen, etc. A potential issue with this method is that the no-search model does not distinguish between consumers who abstain from purchase because they do not like the variants in the current assortment in absolute terms (i.e., no variant exceeds their no-purchase utility) and consumers who do not purchase because they engage in search (i.e., no variant exceeds their no-search threshold).

Because the no-search model does not distinguish between the two reasons for a consumer to choose to not purchase any variant, we demonstrate that the estimated parameters with the no-search model are not independent of the assortment chosen and the iterative application of that model can lead to a heuristic equilibrium: an assortment that is optimal given the parameters estimated from the data observed with that assortment. We show that the heuristic equilibrium assortment never contains more variants than the optimal assortment and may very well contain fewer variants.

While the no-search assortment planning model yields biased parameter estimates, if the true model is independent search, we find that the iterative application of the no-search assortment planning model generally leads to good assortments: out of 2868 scenarios, the average profit loss was only 0.29% and the maximum profit loss relative to the optimal profit was only 10.26%. However, the same does not hold if the no-search assortment planning model is applied when the overlapping search model is appropriate: out of 910 scenarios, the average profit loss was 10.26% and the maximum profit loss was 100%. Furthermore, the average profit loss across the tested scenarios increases as consumer search cost decreases. We conclude that retailers should explicitly consider consumer search in their assortment planning process in markets with a limited pool of variants (overlapping search) and low consumer search costs. But in markets that are better characterized by the independent search model, the traditional no-search assortment planning model yields good results.

We believe our model addresses an important issue in the retailing industry, yet several interesting questions remain unexplored. In this paper, we consider a static setting that ignores the competitive interactions associated with a retailer's assortment. Thus, an extension of our model could explore assortment planning with competing retailers in the presence of consumer search. Using retail data to empirically compare estimates from our model with the traditional no-search estimates also seems to be an interesting avenue for future research.

Reference

Anderson, S., A. de Palma and J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation*, The MIT Press, Cambridge, MA.

Anderson, S. and R. Renault. 1999. Pricing, product diversity, and search costs: a Bertrand-Chamberlin-Diamond model. *Rand Journal of Economics.* **30**(4) 719-735.

Armory, M. and E. Plambeck. 2002. The impact of duplicate orders on demand estimation and capacity investment. Working paper, New York University.

Aydin, G. and W. H. Hausman. 2002. Supply chain coordination and assortment planning. presented at INFORMS Conference, San Jose.

Cachon, G. P. and G. Kok. 2002. Heuristic equilibrium in the newsvendor model with clearance pricing. Working paper, The Wharton School, University of Pennsylvania.

Chong, J.K., T. H. Ho, and C. S. Tang. 2001. A modeling framework for category assortment planning. *Manufacturing & Service Operations Management.* **3**(3) 191-210.

Fisher, M. and K. Rajaram. 2000. Accurate retail testing of fashion merchandise: methodology and application. *Marketing Science.* **19**(3) 266-278.

Guadagni, P. M. and J. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Science.* **2**(3) 203-238.

Hoch, S. J., E. T. Bradlow, and B. Wansink. 1999. The variety of an assortment. *Marketing Science.* **18**(4) 527-546.

Huffman, C. and B. E. Kahn. 1998. Variety for sale: mass customization or mass confusion?. *Journal of Retailing.* **74**(4) 491-513.

Mahajan, S. and G. van Ryzin. 2001. Stocking retail assortments under dynamic consumer substitution. *Operations Research.* **49**(3) 334-351.

Morgan, P. B. 1983. Search and optimal sample sizes. *Review of Economic Studies.* **50** 659-675.

Salop, S. and J. Stiglitz. Bargains and ripoffs: a model of monopolistically competitive price dispersion. *Review of Economic Studies.* **44** 493-510.

Simonson, I. 1999. The effect of product assortment on buyer preferences. *Journal of Retailing*, **75**(3) 347-370.

Smith, S. 2002. Choosing optimal retail assortments using conjoint preference data. presented at INFORMS Conference, San Jose.

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

Stahl, K. 1982. Differentiated products, consumer search, and locational oligopoly. *Journal of Industrial Economics.* **31** 97-114.

Stahl, D. O. 1989. Oligopolistic pricing with sequential consumer search. *American Economic Review.* **79**(4) 700-712.

Stigler, G. J. 1961. The economics of information. *The Journal of Political Economy.* **69**(3) 213-225.

Mahajan, S. and G. van Ryzin. 1998. Retail inventories and consumer choice. *Quantitative Models for Supply Chain Management*, Edited by Tayur, S., Ganeshan, R., and Magazine, M., Kluwer Academic Publishers, Boston, MA, 491-551.

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

Weitzman, M.L. 1979. Optimal search for the best alternative. *Econometrica.* **47**(3) 641-654.

Wolinsky, A. 1983. Retail trade concentration due to consumers' imperfect information. *Bell Journal of Economics.* 275-282.

Wolinsky, A. 1984. Product differentiation with imperfect information. *Review of Economic Studies.* **51** 53-61.

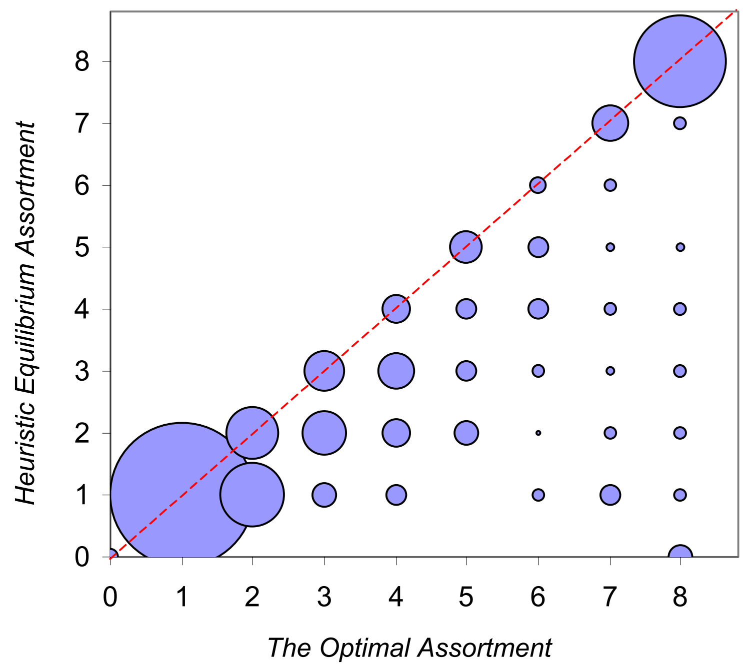

Figure 1: Heuristic equilibrium assortment when the no-search assortment planning model is implemented with overlapping search. The diameter of a circle represents the frequency of that assortment in the numerical study.

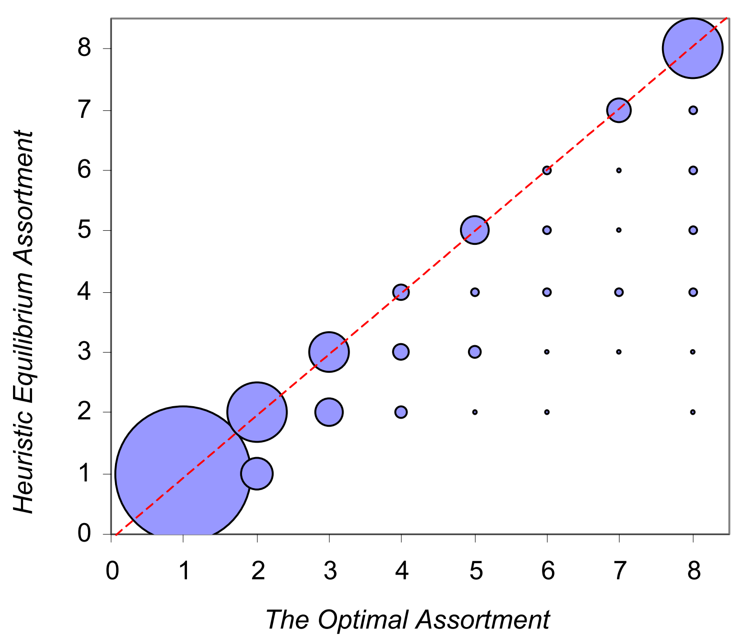

Figure 2: Heuristic equilibrium assortment when the no-search assortment planning model is implemented with independent search. The diameter of a circle represents the frequency of that assortment in the numerical study.

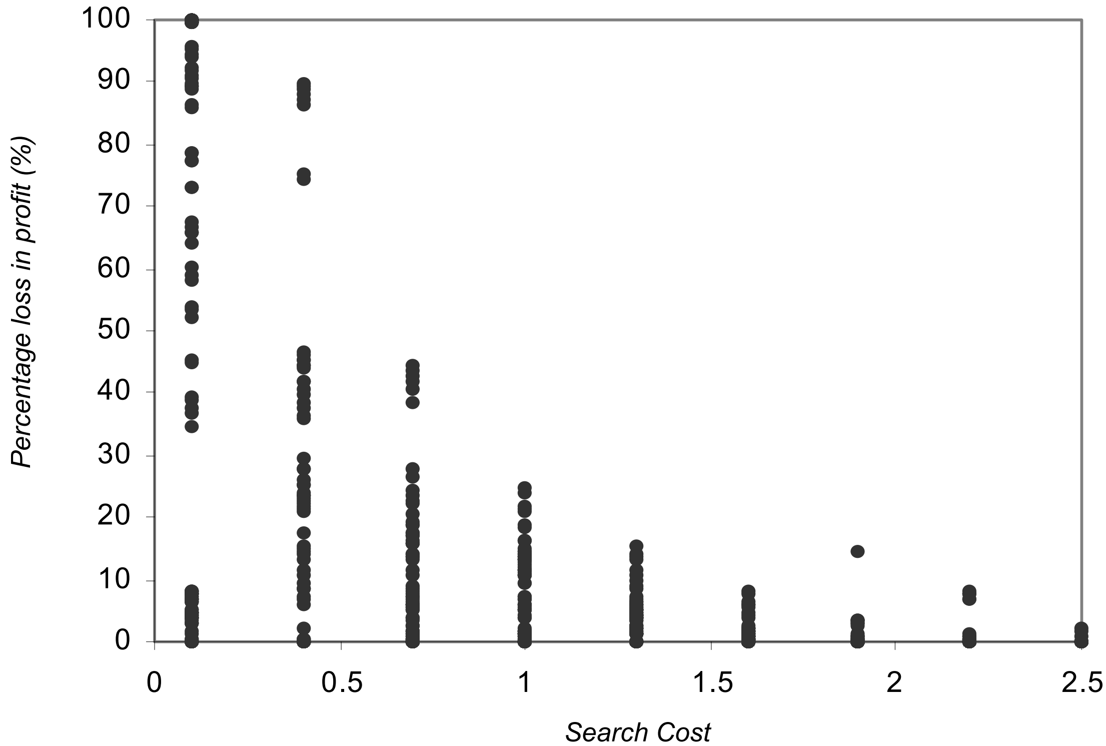

Figure 3: Percentage profit loss relative to the optimal profit of the heuristic equilibrium assortment when the no-search assortment model is implemented with overlapping search.

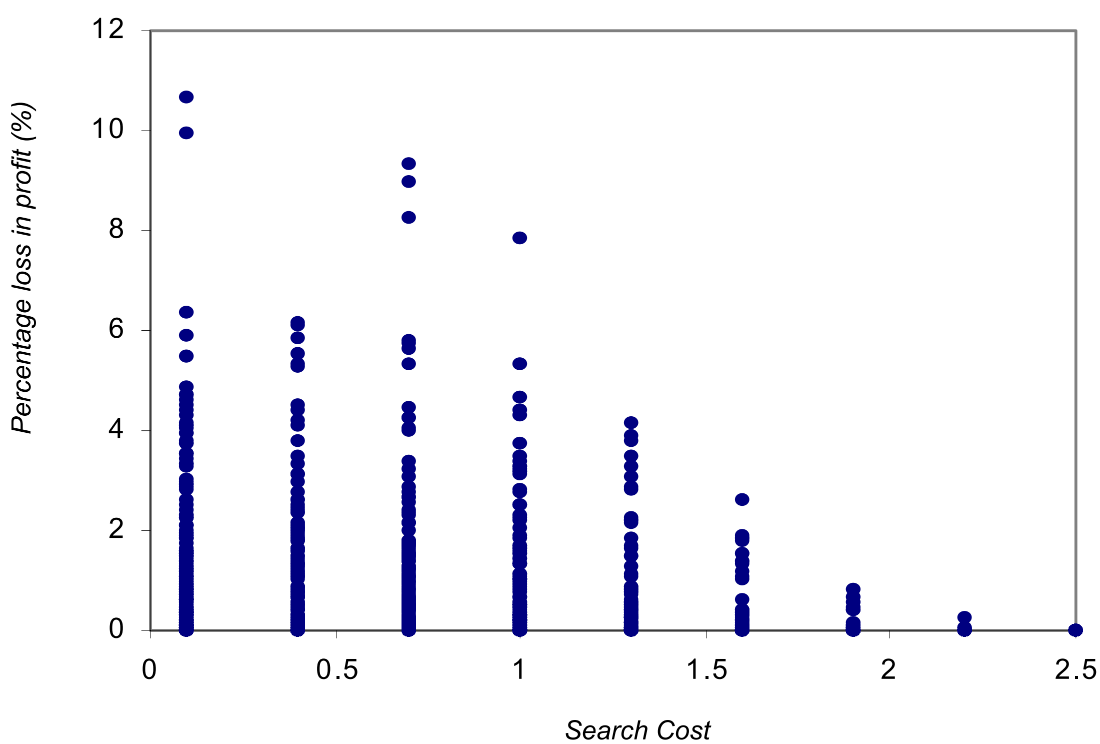

Figure 4: Percentage profit loss relative to the optimal profit the heuristic equilibrium assortment when the no-search assortment model is implemented with independent search.

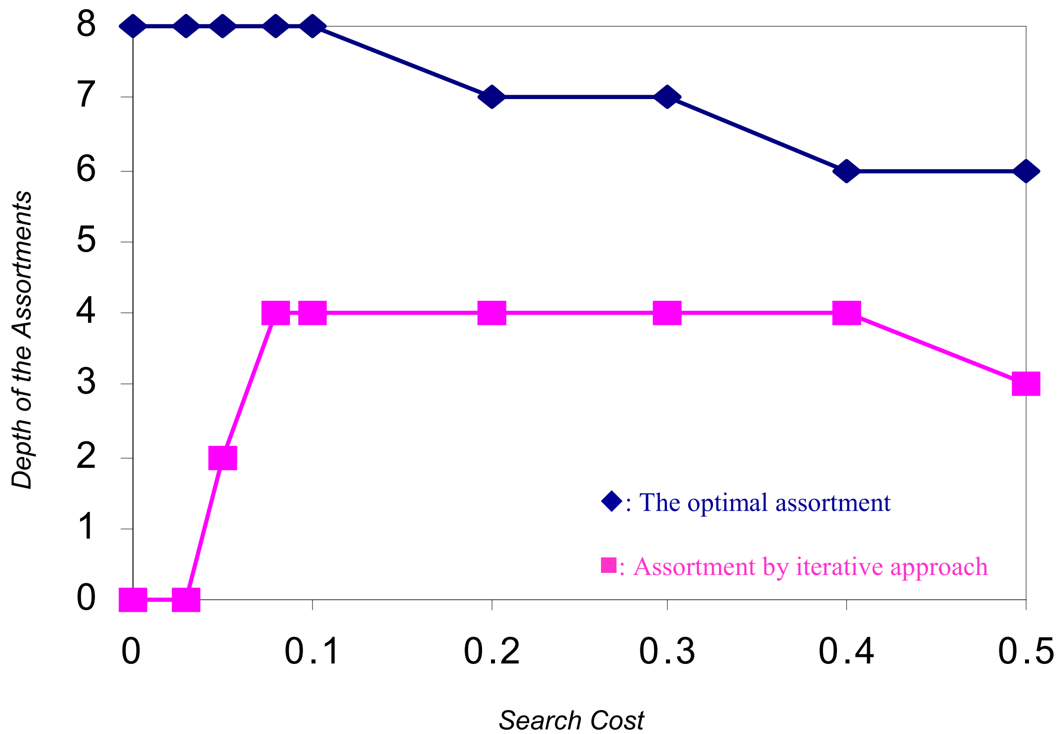

Figure 5: The trajectories of the optimal assortment and the heuristic equilibrium assortment with parameters $\phi_1 = 3$, $\phi_2 = 0.5$, $u_0 = u_1 - 0.75(u_1 - u_8)$, $m = 5$, $\nu = 0.5$, and $\beta = 0.5$.