Assortment Optimization under the Multinomial Logit Model with Random Choice Parameters

Paat Rusmevichientong[1], David Shmoys[2], Chaoxu Tong[3], Huseyin Topaloglu[4]

September 19, 2013

Abstract

We consider assortment optimization problems under the multinomial logit model, where the parameters of the choice model are random. The randomness in the choice model parameters is motivated by the fact that there are multiple customer segments, each with different preferences for the products, and the segment of each customer is unknown to the firm when the customer makes a purchase. This choice model is also called the mixture-of-logits model. The goal of the firm is to choose an assortment of products to offer that maximizes the expected revenue per customer, across all customer segments. We establish that the problem is NP-complete even when there are just two customer segments. Motivated by this complexity result, we focus on assortments consisting of products with the highest revenues, which we refer to as revenue-ordered assortments. We identify specially structured cases of the problem where revenue-ordered assortments are optimal. When the randomness in the choice model parameters does not follow a special structure, we derive tight approximation guarantees for revenue-ordered assortments. We extend our model to the multi-period capacity allocation problem, and prove that, when restricted to the revenue-ordered assortments, the mixture-of-logits model possesses the nesting-by-fare-order property. This result implies that revenue-ordered assortments can be incorporated into existing revenue management systems through nested protection levels. Numerical experiments show that revenue-ordered assortments perform remarkably well, generally yielding profits that are within a fraction of a percent of the optimal.

1 Introduction

A common problem faced by many firms involves choosing an assortment of products to offer to their customers with the goal of maximizing revenues. There are two sources of difficulty when dealing with such problems. First, the products may serve as substitutes and the customers may make a choice among the products that satisfy their needs. This creates a situation where the demand for each product depends on what assortment of products are offered to the customers. Second, there can be multiple customer segments served by the firm and the customers belonging to different segments may have different preferences. So, one has to consider the preferences of the different customer segments, as well as the size of each segment, when deciding which assortment of products to offer.

In this paper, we consider an assortment optimization problem to address the difficulties described above. Each customer makes a choice among the offered products according to the multinomial logit choice model. The crucial feature of our assortment optimization problem is that the parameters of the choice model are assumed to be unknown to the firm and they are modeled as random variables. This type of a situation happens when there are multiple customer segments with different preferences and

[1]Marshall School of Business, University of Southern California, Los Angeles, CA 90089, USA, Email: rusmevic@marshall.usc.edu

[2]School of Operations Research and Information Engineering and Department of Computer Science, Cornell University, Ithaca, NY 14853, USA, Email: shmoys@cs.cornell.edu

[3]School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA, Email: ct423@cornell.edu

[4]School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA, Email: topaloglu@orie.cornell.edu

the segment of a customer is not known to the firm when the customer makes a purchase. The goal of the firm is to choose an assortment of products that maximizes the expected revenue obtained from each customer.

Our work in this paper has connections to the growing literature on assortment planning models in revenue management. In their seminal work, Talluri and van Ryzin (2004) study a version of our assortment optimization problem under the assumption that the parameters of the multinomial logit model are deterministic and known. Under this assumption, the authors show that the optimal assortment includes a certain number of products with the highest revenues. Throughout the paper, we refer to such assortments as revenue-ordered assortments. In this case, the optimal assortment can be obtained efficiently by checking the expected revenue from each assortment that includes a certain number of products with the highest revenues. In contrast, if the model parameters are random, then revenue-ordered assortments are no longer optimal and the assortment optimization problem appears to be intractable. Indeed, Bront et al. (2009) study the assortment optimization problem with random choice model parameters and establish that the problem is NP-complete when the number of possible realizations of the choice model parameters is at least as large as the number of products.

Our Contributions and Main Results: The above discussion shows a strong contrast between *deterministic* versus *random* model parameters. For the case with deterministic and known model parameters, knowing the optimality of revenue-ordered assortment has crucial theoretical and practical implications. On the theoretical side, optimality of revenue-ordered assortments allows us to find an optimal assortment in linear time. On the practical side, optimality of revenue-ordered assortments are intuitively appealing as they urge the firms to shift their focus on high-contribution products, making them easy to justify. Also, optimality of revenue-ordered assortments corresponds to nested protection levels in revenue management settings, where a certain fare class remains open as long as cheaper fare classes remain open. Legacy revenue management systems are commonly tied to the assumption of nested protection levels and using revenue-ordered assortments allow compatibility with such systems. Our goal in this paper is to close the big chasm between the two cases. Building on the fact that revenue-ordered assortments are optimal when the choice model parameters are deterministic, *we investigate the question of what we can say about the performance of such assortments when the choice model parameters are random.*

In this paper, we begin by showing that the assortment optimization problem with random choice model parameters is NP-complete even when there are just two possible realizations for the model parameters (Theorem 3.2). Following the hardness result, we give two specific cases with random choice model parameters where revenue-ordered assortments remain optimal. In the first case, the choice model is such that each product has an intrinsic mean utility known to the firm, but the mean utility of a product is modified by a term that depends on its price and the price sensitivity of a customer is sampled from the uniform distribution (Theorem 4.1). In the second case, the customers attach smaller mean utilities to the products with higher prices, but they are still aware of the value provided by high-end products so that they are not extremely price sensitive in a sense we make precise (Theorem 4.3). These two results cover a wide array of potentially useful situations where the parameters of the choice model are random according to a special structure and we can still compute the optimal assortment by focusing only on revenue-ordered assortments. To our knowledge, these two cases provide first assortment optimization problems under the multinomial logit model with random parameters, where the optimal assortment remains tractable to compute.

Following these results, we focus on the performance of revenue-ordered assortments when the randomness in the choice model parameters does not follow a special structure. If there are G possible realizations for the choice model parameters and n products that we can possibly offer to the customers, then we show that revenue-ordered assortments provide an approximation guarantee of $\min\{G, \lceil n/2 \rceil\}$ (Theorem 5.1). Furthermore, this performance guarantee is tight in the sense that there are instances of the assortment optimization problem where the expected revenue from the best revenue-ordered assortment deviates from the optimal by a factor of $\min\{G, \lceil n/2 \rceil\}$ (Proposition 5.3). The tight instances involve products whose revenues differ from each other by orders of magnitude. So, in our third guarantee, we show that revenue-ordered assortments provide an approximation guarantee of $e \log(e\,\rho)$, where ρ is the ratio between the largest and smallest product revenues (Theorem 5.4). The last result intuitively suggests that unless the revenues of the products differ by orders of magnitude, revenue-ordered assortments yield a constant factor approximation guarantee. Finally, we give an approximation guarantee for revenue-ordered assortments by using information about the distribution of the utilities that a customer associates with the products (Proposition 5.5).

We extend our model to the multi-period capacity allocation setting of Talluri and van Ryzin (2004), where we have an initial capacity of airline seats that must be allocated to multiple fare classes over time. In each period, we determine the assortment of fare classes to offer to an arriving customer whose choice model is described by a mixture-of-logits model, with the goal of maximizing the total expected revenue over the selling horizon. Surprisingly, we are able to show that if we have more remaining capacity, then the optimal revenue-ordered assortment to offer becomes a larger assortment when customers choose according to the mixture-of-logits (Theorem 6.1). Thus, when we restrict our attention to the revenue-ordered assortments, the mixture-of-logits model possesses the nesting-by-fare-order property, which is to say that as the remaining capacity on the flight leg gets smaller, fare classes with lower fares stop being offered first (Talluri and van Ryzin, 2004). This result has an important managerial implication as it implies that we can use nested protection level policies. Many traditional revenue management systems are built on these policies, and thus, revenue-ordered assortments can be seamlessly incorporated into these systems. We note that nesting-by-fare-order property does not hold under the mixture-of-logits choice model when we do not limit our attention to revenue-ordered assortments. We conclude with numerical experiments, showing that revenue-ordered assortments indeed perform well in practice, yielding expected revenues within 1% of the optimal on average.

To sum up, incorporating random choice model parameters into assortment optimization problems is crucial from a practical perspective, as firms serve customers from different segments with different preferences. We establish new problem classes with random model parameters, where revenue-ordered assortments are optimal. These problem classes appear to constitute the first assortment optimization problems with random parameters that admit tractable solutions. For the general case, we give approximation guarantees for revenue-ordered assortments, one of which indicates that revenue-ordered assortments should perform well, unless product revenues differ by orders of magnitude. We extend our model to a multi-period setting and show that revenue-ordered assortments allow us to recover nesting-by-fare-order property. Our numerical experiments demonstrate that revenue-ordered assortments do perform quite well in practice.

Related Literature: We focus our literature review primarily on papers that use variants of the multinomial logit model to describe the demand process. We refer the reader to Kok et al. (2008), Farias et al. (2011) and Farias et al. (2012) for assortment problems under other choice models. Talluri and van Ryzin (2004) show that if customers choose according to the multinomial logit model and the choice model parameters are deterministic and known, then revenue-ordered assortments are optimal. Gallego et al. (2011) show that this problem can directly be formulated and solved as a linear program. Rusmevichientong et al. (2010) study the same problem with a cardinality constraint on the offered assortment and show that although revenue-ordered assortments are no longer optimal, the optimal assortment can be computed efficiently. Jagabathula et al. (2011) also focus on assortment optimization problems where there is a cardinality constraint on the offered assortment. They study the performance of an intuitive pairwise exchange algorithm and show that this algorithm finds the optimal solution when customers choose according to the multinomial logit model with known parameters. Rusmevichientong and Topaloglu (2012) show that revenue-ordered assortments are robust against the uncertainty in model parameters, in the sense that they protect against the worst-case expected revenue. This class of assortments are no longer optimal when the choice model parameters are random. Bront et al. (2009) show that the assortment problem is NP-complete in the strong sense under random model parameters, give a mixed integer programming formulation to obtain the optimal assortment and suggest a greedy heuristic. Mendez-Diaz et al. (2010) strengthen the mixed integer programming formulation through valid inequalities. There is some work on solving assortment problems under the nested logit model, which is an extension of the multinomial logit model, allowing correlations between the evaluations of different products by a particular customer. Davis et al. (2011) give a linear programming formulation of the assortment problem under the nested logit model. Li and Rusmevichientong (2012) give a greedy algorithm for the same problem. Gallego and Topaloglu (2012) show how to impose a variety of constraints on the offered assortment when customers chose according to the nested logit model. All of the current work on nested logit model is under the assumption that there is a single customer segment with known choice model parameters.

Our work is also related to revenue management models that incorporate customer choice. Talluri and van Ryzin (2004) study a revenue management problem over a single flight leg. Customers choose among the fare classes that are available for purchase and the objective is to adjust the assortment of available fare classes at each period to maximize the total expected revenue. There are a number of papers that extend this work to a flight network; see Gallego et al. (2004), Liu and van Ryzin (2008), Kunnumkal and Topaloglu (2008), Zhang and Adelman (2009), Talluri (2010), Gallego et al. (2011), Meissner and Strauss (2012), Vossen and Zhang (2012) and Meissner et al. (2012). The fundamental idea in these papers is to formulate various deterministic linear programming approximations. These linear programs have one decision variable for each subset of itinerary products, corresponding to the duration of time during which a subset of itinerary products is made available to customers. As a result, the number of decision variables can get large and it is customary to solve these linear programs by using column generation. The column generation subproblem corresponds to our assortment problem when customers choose according to the multinomial logit model with random parameters. Of particular interest in this domain is the work by Talluri (2010) and Meissner et al. (2012), where the authors focus on the case with multiple customer segments and customers from different segments choose according to different choice models. The first paper gives tractable relaxations of the problem when there are multiple segments, whereas the second paper gives valid cuts to tighten the relaxation, but as far as we are aware, it is difficult to a priori characterize the tightness of these relaxations.

The rest of the paper is organized as follows. In Section 2, we formulate our assortment optimization problem. In Section 3, we show that this assortment optimization problem is NP-complete. In Section 4, we give two specific problem classes with random choice model parameters where revenue-ordered assortments remain optimal. In Section 5, we develop approximation guarantees for revenue-ordered assortments when the randomness in the choice model parameters does not have any special structure. In Section 6, we extend our model to a multi-period setting. In Section 7, we give numerical experiments testing the performance of revenue-ordered assortments. In Section 8, we conclude.

2 Model Formulation

We have n products indexed by $1, 2, \ldots, n$, and for each i, let r_i be the revenue associated with product i. Without loss of generality, we assume that the products are indexed such that $r_1 \geq r_2 \geq \ldots \geq r_n$. The customers choose among the offered products according to random utility maximization. In particular, each customer associates the utility

$$U_i = V_i + \epsilon_i$$

with product i, where ϵ_i is a standard Gumbel random variable with mean zero and we view V_i as the mean utility of product i. We normalize the utility of the no purchase option to zero. In this case, if we offer the assortment $S \subseteq \{1, \ldots, n\}$ of products to the customers, then a customer chooses the product with the highest utility if the utility of this product is positive, but otherwise, leaves without purchasing anything. It is a standard result in discrete choice theory (see, for example, Ben-Akiva and Lerman, 1985 and Train, 2003) that if we offer the assortment S to the customers, then a customer chooses product $i \in S$ with probability

$$\pi_i(S, \boldsymbol{V}) = \frac{e^{V_i}}{1 + \sum_{j \in S} e^{V_j}},$$

where we use $\boldsymbol{V} = (V_1, \ldots, V_n)$ to denote the vector of mean utilities of the products, and make the dependence of $\pi_i(S, \boldsymbol{V})$ on S and $\boldsymbol{V}$ explicit. The choice model above is known as the multinomial logit model. If we offer the assortment S and the vector of mean utilities is $\boldsymbol{V}$, then the expected revenue obtained from a customer is

$$f(S, \boldsymbol{V}) = \sum_{i \in S} r_i \, \pi_i(S, \boldsymbol{V}).$$

When the mean utility vector $\boldsymbol{V}$ is fixed and known, we can find an assortment that maximizes the expected revenue from a customer by solving the problem $\max_{S\subseteq\{1,\ldots,n\}} f(S,\boldsymbol{V})$. The implicit assumption behind using a fixed mean utility vector is that each customer is probabilistically identical, reacting to an offered assortment in the same manner. On the other hand, if each customer has a different reaction towards an assortment, then the mean utilities that a customer attaches to the products can be modeled as random variables themselves. In that case, we can solve

$$Z^* = \max_{S\subseteq\{1,\ldots,n\}} \mathbb{E}\{f(S,\boldsymbol{V})\} \qquad \text{(Mixture Logit)}$$

to find an assortment that maximizes the expected revenue over all possible realizations of the mean utility vector. The above expectation involves the random vector $\boldsymbol{V}$ and the distribution of $\boldsymbol{V}$ naturally depends on the composition of the market and how customers in the market make a choice. The random vector $\boldsymbol{V}$ can have a discrete or continuous distribution and we assume that it is independent of $(\epsilon_1,\ldots,\epsilon_n)$. We continue referring to V_i as the mean utility of product i, although V_i is random. Throughout the paper, we focus on the Mixture Logit problem.

The mean utilities $\boldsymbol{V}$ can be either discrete or continuous random vectors, depending on the specific application. For example, if we have G customer types, with each type following a multinomial logit choice model, then $\boldsymbol{V}$ is a discrete random vector that takes G different values $\hat{\boldsymbol{V}}^1,\ldots,\hat{\boldsymbol{V}}^G$, where $\hat{\boldsymbol{V}}^g$ corresponds to the mean utilities of customer type g. In another example, we can have $\boldsymbol{V} = \boldsymbol{\mu} - B\boldsymbol{r}$, where $\boldsymbol{\mu} = (\mu_1,\ldots,\mu_n)$ is a deterministic vector, $\boldsymbol{r} = (r_1,\ldots,r_n)$ gives the prices of the products, and B is a continuous random variable. In this case, $\boldsymbol{V}$ is a continuous random vector, and B is the customer-specific price sensitivity, whose distribution reflects the price sensitivity pattern among the customers. More generally, McFadden and Train (2000) have shown that any choice model based on random utility maximization can be approximated arbitrarily well by the multinomial logit model with random parameters, although the required mixing distribution of $\boldsymbol{V}$ can be quite complicated and may be difficult to calibrate. Thus, the mixture-of-logits model, in principal, can be used as an approximation in assortment optimization problems under a general choice model.

3 Computational Complexity

If the mean utility vector $\boldsymbol{V}$ is fixed and known, then we can find the optimal assortment by solving the problem $\max_{S\subseteq\{1,\ldots,n\}} f(S,\boldsymbol{V})$. Talluri and van Ryzin (2004) show that the optimal assortment for the last optimization problem is of the form $\{1,\ldots,i^*_{\boldsymbol{V}}\}$, including $i^*_{\boldsymbol{V}}$ products with the highest revenues. We call such assortments as revenue-ordered assortments. In this case, we can find the optimal assortment in a tractable fashion simply by checking the expected revenue $f(\{1,\ldots,i\},\boldsymbol{V})$ for all $i = 1,\ldots,n$. In contrast, the following example shows that revenue-ordered assortments are no longer optimal when we have multiple possible values for the mean utility vector, suggesting that it may be difficult to obtain an optimal solution to the Mixture Logit problem.

Example 3.1 Consider an instance of the Mixture Logit problem with three products. The revenues of the products are $(r_1,r_2,r_3) = (8,4,3)$. There are two possible realizations of $\boldsymbol{V}$, denoted by $\hat{\boldsymbol{V}}^1 =$

$(\log 5, \log 20, 0)$ and $\hat{\boldsymbol{V}}^2 = (-\log 5, \log 10, \log 10)$, each equally likely. The next table shows $f(S, \hat{\boldsymbol{V}}^1)$, $f(S, \hat{\boldsymbol{V}}^2)$ and $\mathbb{E}\{f(S, \boldsymbol{V})\} = \frac{1}{2} f(S, \hat{\boldsymbol{V}}^1) + \frac{1}{2} f(S, \hat{\boldsymbol{V}}^2)$ for each possible assortment S.

S	$f(S, \hat{\boldsymbol{V}}^1)$	$f(S, \hat{\boldsymbol{V}}^2)$	$\mathbb{E}\{f(S, \boldsymbol{V})\}$
$\{1\}$	$\underline{6.67}$	1.33	4.00
$\{2\}$	3.81	3.64	3.72
$\{3\}$	1.50	2.73	2.11
$\{1,2\}$	4.62	$\underline{3.71}$	4.16
$\{1,3\}$	6.14	2.82	$\underline{4.48}$
$\{2,3\}$	3.77	3.33	3.55
$\{1,2,3\}$	4.56	3.38	3.97

If the mean utility vector is known to be $\hat{\boldsymbol{V}}^1$, then the assortment $\{1\}$ maximizes the expected revenue, whereas if the mean utility vector is known to be $\hat{\boldsymbol{V}}^2$, then the assortment $\{1, 2\}$ is optimal. On the other hand, if the mean utility vector is equally likely to take the two values, then the assortment $\{1, 3\}$ maximizes the expected revenue. Since this assortment skips over the second product, it is not revenue-ordered. Furthermore, the optimal assortment includes the third product, which does not even appear in the optimal assortments when we focus on each possible value of $\boldsymbol{V}$ separately.

In the next theorem, we formally establish that the Mixture Logit problem is difficult by showing that it is NP-complete even when we have two possible realizations of the random vector $\boldsymbol{V}$. Bront et al. (2009) show that the Mixture Logit problem is NP-complete in the strong sense and their result requires that the number of possible realizations for $\boldsymbol{V}$ is at least as large as the number of products. This result leaves open the question of whether the problem is still difficult when the random vector $\boldsymbol{V}$ does not have too many realizations, which, for example, is the case when dealing with a small number of customer segments. Theorem 3.2 affirmatively settles this question. Before proceeding to the theorem, we introduce the following decision-theoretic formulation of the assortment optimization problem with two possible realizations of $\boldsymbol{V}$. We call this problem the 2-Class Logit problem.

2-Class Logit

Inputs:

- Set of products indexed by $1, 2, \ldots, n$.
- Product revenues $r_1, r_2, \ldots, r_n$, where $r_i \in \mathbb{Z}_+$ for all $i = 1, \ldots, n$.
- Parameters of the multinomial logit model corresponding to the two realizations of the mean utilities $\hat{\boldsymbol{V}}^1 = (\hat{V}_1^1, \ldots, \hat{V}_n^1)$ and $\hat{\boldsymbol{V}}^2 = (\hat{V}_1^2, \ldots, \hat{V}_n^2)$, where for $i = 1, \ldots, n$, $g = 1, 2$, $e^{\hat{V}_i^g} \in \mathbb{Q}_+$, the set of positive rational numbers.
- Probability of observing the two realizations of the mean utilities, denoted by (α^1, α^2), where $\alpha^g \in \mathbb{Q}_+$ for $g = 1, 2$ and $\alpha^1 + \alpha^2 = 1$.
- Target revenue $K \in \mathbb{Z}_+$.

Question: Is there an assortment $S \subseteq \{1, \ldots, n\}$ whose expected revenue is at least K, that is,

$$\alpha^1 \frac{\sum_{i \in S} r_i \, e^{\hat{V}_i^1}}{1 + \sum_{i \in S} e^{\hat{V}_i^1}} + \alpha^2 \frac{\sum_{i \in S} r_i \, e^{\hat{V}_i^2}}{1 + \sum_{i \in S} e^{\hat{V}_i^2}} \;\geq\; K\ ?$$

Theorem 3.2 2-Class Logit *is NP-complete.*

The proof of Theorem 3.2 involves a reduction from the Partition problem, which is a well-known NP-complete problem (Garey and Johnson, 1979). The details are given in Appendix A in the Online Supplement. We note that our complexity result is different from that of Goyal et al. (2013), who establish the NP-hardness result for the assortment optimization problem under the mixture-of-logits model, where all products have the same revenue, with $r_i = 1$ for $i = 1, \dots, n$, but there is a cardinality constraint on the size of the assortment. Under their model, each customer segment is associated with a set of exactly two equally-preferred products, and the customers in each segment will only purchase one of these two products. This corresponds to the case where every realization of the mean utility $\boldsymbol{V}$ is a vector in $\{-\infty, 0\}^n$, where exactly two coordinates are zero and the rest are $-\infty$. They use a reduction from the Vertex Cover problem, another well-known NP-complete problem, and the number of realizations of $\boldsymbol{V}$ is equal to the number of edges in the graph.

In light of Example 3.1 and Theorem 3.2, the Mixture Logit problem does not admit a simple solution with intuitive properties. On the other hand, if the parameters of the choice model are fixed and known, then a revenue-ordered assortment becomes optimal. Computing the best revenue-ordered assortment is simple since there are only n possible revenue-ordered assortments, and one can check the expected revenue from each one of them. Furthermore, revenue-ordered assortments are intuitively appealing, urging firms to focus on high-contribution products. Therefore, there is a variety of theoretical and practical reasons to work with revenue-ordered assortments. These observations raise two natural questions. First, are there special cases of the Mixture Logit problem for which revenue-ordered assortments remain optimal? Second, can we make any statements about the performance of revenue-ordered assortments for the general instances of the Mixture Logit problem? In the next section, we begin by answering the first question. The section following the next dwells on the second question.

4 Optimality of Revenue-Ordered Assortments

In this section, we give two special cases of the Mixture Logit problem that admit revenue-ordered assortments as the optimal solution.

4.1 Product-Independent Price Sensitivity with Random No-Purchase Mean Utilities

When using the multinomial logit model in practice, it is customary to express the mean utility of a product as a function of its features, such as price, quality and durability. In this section, we follow a similar approach. In particular, we assume that the mean utility of product i is of the form

$$V_i = \mu_i + P - B\, r_i, \qquad \text{(Product-Independent Sensitivity)}$$

where μ_i is a deterministic constant, P and B are random variables and B takes positive values. We observe that this choice of mean utilities corresponds to the case where each customer associates the intrinsic mean utilities $(\mu_1, \dots, \mu_n)$ with the products. For each customer, the intrinsic mean utilities of all products are perturbed by a quantity sampled from the distribution of the random variable P. On top of this, the price sensitivity of each customer is sampled from the distribution of the random variable B and the intrinsic mean utility of each product is further perturbed by a quantity that depends on the price sensitivity of the customer and the price of each product. A customer associates the same price sensitivity B for each product. Since the utility of the no-purchase option V_0 is always set to zero, the PRODUCT-INDEPENDENT SENSITIVITY model is equivalent to

$$V_0 = -P, \qquad \text{and} \qquad V_i = \mu_i - B\, r_i, \qquad i = 1, 2, \ldots, n,$$

and thus, in this model, we allow the mean utility of the no-purchase option to vary across customers.

Surprisingly, the main result of this section shows that if the random variable B is uniformly distributed and it is independent of P, then revenue-ordered assortments are optimal for the MIXTURE LOGIT problem. We observe that this result allows the random variable P to be arbitrary, allowing for a broad class of distributions. While the uniform distribution assumption on the random variable B is limiting, it corresponds to the uninformative prior. For example, if we estimate the price sensitivity of the customers to be $\hat{B}$ from data with a margin of error $\pm\varepsilon$, then we can assume that B is uniformly distributed over the interval $[\hat{B} - \varepsilon, \hat{B} + \varepsilon]$.

Imposing a condition on the distribution of B is necessary to ensure that revenue-ordered assortments are optimal for the MIXTURE LOGIT problem, even when the mean utilities have the PRODUCT-INDEPENDENT SENSITIVITY form. To see this, consider an instance of the MIXTURE LOGIT problem with five products. The revenues of the products are $(r_1, r_2, r_3, r_4, r_5) = (92, 91, 89, 39, 13)$. The mean utility of product i is of the PRODUCT-INDEPENDENT SENSITIVITY form with $\mu_i = 0$ and $P = 0$, but there are two possible realizations of B, denoted by $\hat{B}_1 = 0.01$ and $\hat{B}_2 = 12$, occurring with probabilities 0.1 and 0.9. For this problem instance, the best revenue-ordered assortment is $\{1, 2, 3, 4, 5\}$ providing an expected revenue of about 7.67, but the optimal assortment is $\{1, 2, 3, 5\}$ with an expected revenue of about 7.72. Thus, revenue-ordered assortments are not necessarily optimal when B has an arbitrary distribution.

The main result of this section is given in the following theorem. The proof is given in Appendix B in the Online Supplement.

Theorem 4.1 *If the mean utilities of the products are of the* PRODUCT-INDEPENDENT SENSITIVITY *form, the random variables P and B are independent of each other, and B is uniformly distributed, then a revenue-ordered assortment is optimal for the* MIXTURE LOGIT *problem.*

4.2 Value Conscious Customers

In this section, we give another special case of the MIXTURE LOGIT problem where revenue-ordered assortments remain optimal. In particular, we assume that any realization of the mean utility vector $\boldsymbol{V} = (V_1, \ldots, V_n)$ satisfies

$$V_1 \le V_2 \le \ldots \le V_n \;\; \text{and} \;\; r_1\, e^{V_1} \ge r_2\, e^{V_2} \ge \ldots \ge r_n\, e^{V_n}. \qquad \text{(VALUE CONSCIOUS)}$$

Revenue-ordered assortments remain optimal when any realization of the mean utility vector satisfies the VALUE CONSCIOUS condition. Before we proceed to the proof of this result, we discuss the implications of the VALUE CONSCIOUS condition. We recall that the products are indexed such that $r_1 \geq r_2 \geq \ldots \geq r_n$. Consider a customer whose mean utility vector is given by $\boldsymbol{V} = (V_1, \ldots, V_n)$. The first requirement that $V_1 \leq V_2 \leq \ldots \leq V_n$ corresponds to a situation where the customer prefers less expensive products, which is quite reasonable in many settings. On the other hand, the second requirement that $r_1 \, e^{V_1} \geq r_2 \, e^{V_2} \geq \ldots \geq r_n \, e^{V_n}$ implies that for any $i < j$, we have

$$f(\{i\}, \boldsymbol{V}) = \frac{r_i \, e^{V_i}}{1 + e^{V_i}} \geq \frac{r_j \, e^{V_j}}{1 + e^{V_i}} \geq \frac{r_j \, e^{V_j}}{1 + e^{V_j}} = f(\{j\}, \boldsymbol{V}).$$

This means that although the customer prefers cheaper products, she is value conscious and our expected revenue from offering a premium product is still larger. We actually generalize the inequality above in the proof of Theorem 4.3, showing that $f(S \cup \{i\}, \boldsymbol{V}) \geq f(S \cup \{j\}, \boldsymbol{V})$ whenever $i < j$ and $i, j \notin S$.

In the next example, we give a concrete situation where we can have mean utilities that satisfy the VALUE CONSCIOUS condition.

Example 4.2 Consider the case where the mean utility of product i is of the form $V_i = P \, r_i^{-B}$, where P and B are (possibly dependent) arbitrary random variables and B takes values in the interval $[0, 1]$. This corresponds to a situation where the price sensitivity of each customer is sampled from the distribution of B and the customers are not too price sensitive in the sense that B takes values in the interval $[0, 1]$. It is easy to check that this choice of mean utilities satisfies the VALUE CONSCIOUS condition.

The next theorem gives the main result of this section, showing that revenue-ordered assortments are optimal under the VALUE CONSCIOUS condition.

Theorem 4.3 *If the mean utilities of the products satisfy the* VALUE CONSCIOUS *condition, then a revenue-ordered assortment is optimal for the* MIXTURE LOGIT *problem.*

Proof. We claim that an optimal solution to the problem

$$\max_{S \subseteq \{1, \ldots, n\} : |S| = k} f(S, \boldsymbol{V})$$

is always given by the assortment $\{1, \ldots, k\}$. Assume on the contrary that S^* is the optimal solution to the problem above, but there are two products i and j with $i < j$, $i \notin S^*$ and $j \in S^*$. If we let $T = S^* \setminus \{j\}$, then it follows that

$$f(T \cup \{i\}, \boldsymbol{V}) = \frac{\sum_{\ell \in T} r_\ell \, e^{V_\ell} + r_i \, e^{V_i}}{1 + \sum_{\ell \in T} e^{V_\ell} + e^{V_i}} \geq \frac{\sum_{\ell \in T} r_\ell \, e^{V_\ell} + r_j \, e^{V_j}}{1 + \sum_{\ell \in T} e^{V_\ell} + e^{V_j}} = f(T \cup \{j\}, \boldsymbol{V}) = f(S^*, \boldsymbol{V}),$$

where the inequality follows from the VALUE CONSCIOUS condition. Thus, $f(S^* \setminus \{j\} \cup \{i\}, \boldsymbol{V}) \geq f(S^*, \boldsymbol{V})$, indicating that the objective value of the optimization problem at the beginning of the proof under $S^* \setminus \{j\} \cup \{i\}$ is at least as large as the one corresponding to the solution S^*. This establishes the claim. Since an optimal solution to the problem at the beginning of the proof is always given by $\{1, \ldots, k\}$ irrespective of the value of $\boldsymbol{V}$, we observe that an optimal solution to the problem

$$\max_{S \subseteq \{1,\ldots,n\} : |S| = k} \mathbb{E}\{f(S, \boldsymbol{V})\}$$

is also given by the assortment $\{1, \ldots, k\}$. The result follows by noting that the optimal solution to the Mixture Logit problem can be obtained by solving the last problem above for all $k = 1, \ldots, n$. $\square$

5 Performance of Revenue-Ordered Assortments

In this section, we investigate the performance of revenue-ordered assortments for general instances of the Mixture Logit problem. Throughout this section, we say that revenue-ordered assortments provide an approximation guarantee of β if the expected revenue from the best revenue-ordered assortment deviates from the optimal expected revenue by no more than a factor of β. In other words, noting that the optimal objective value of the Mixture Logit problem is denoted by Z^*, if we have

$$\frac{1}{\beta} Z^* \leq \max_{i=1,\ldots,n} \mathbb{E}\{f(\{1, \ldots, i\}, \boldsymbol{V})\},$$

for some $\beta \geq 1$, then we say that revenue-ordered assortments provide an approximation guarantee of β. In the next section, we derive an approximation guarantee that increases linearly with the number of products and customer segments, and prove that it is tight. This guarantee is applicable when the number of products or customer segments is small. For larger problem instances, we derive other guarantees based on the variations in the mean utilities and product revenues in Section 5.2.

5.1 A Tight Guarantee Based on the Number of Products and Customer Segments

In the next theorem, we show that if there are G possible realizations for the vector of mean utilities and n possible products that we can offer to customers, then revenue-ordered assortments provide an approximation guarantee of $\min\{G, \lceil n/2 \rceil\}$.

Theorem 5.1 (Guarantees Based on Numbers of Mean Utility Realizations and Products) *If there are G realizations of the vector of mean utilities, then revenue-ordered assortments provide an approximation guarantee of* $\min\{G, \lceil n/2 \rceil\}$ *for the* Mixture Logit *problem.*

The proof of Theorem 5.1 makes use of the following property of the expected revenue function.

Lemma 5.2 *For every realization of the vector of mean utilities $\boldsymbol{V}$, the following two results hold.*

(i) For all assortments S and T, $f(S \cup T, \boldsymbol{V}) \leq f(S, \boldsymbol{V}) + f(T, \boldsymbol{V})$.

(ii) For all products i and j with $i < j$, $f(\{i, \ldots, j\}, \boldsymbol{V}) \leq f(\{1, \ldots, j\}, \boldsymbol{V})$.

Proof. The first result follows immediately from the fact that

$$f(S \cup T, \boldsymbol{V}) = \frac{\sum_{i \in S} r_i \, e^{V_i} + \sum_{i \in T} r_i \, e^{V_i}}{1 + \sum_{i \in S} e^{V_i} + \sum_{i \in T} e^{V_i}} \leq \frac{\sum_{i \in S} r_i \, e^{V_i}}{1 + \sum_{i \in S} e^{V_i}} + \frac{\sum_{i \in T} r_i \, e^{V_i}}{1 + \sum_{i \in T} e^{V_i}} = f(S, \boldsymbol{V}) + f(T, \boldsymbol{V}).$$

To establish the second part of the lemma, note that $f(\{i,\dots,j\},\boldsymbol{V}) = \frac{\sum_{\ell=i}^{j} r_\ell\, e^{V_\ell}}{1+\sum_{\ell=i}^{j} e^{V_\ell}}$, which implies that $f(\{i,\dots,j\},\boldsymbol{V})$ is a weighted average of $0, r_i, \dots, r_j$, where the weights associated with each one of these are $1, e^{V_i}, \dots, e^{V_j}$, respectively . By using the same reasoning, $f(\{1,\dots,j\},\boldsymbol{V})$ is a weighted average of $0, r_1, \dots, r_i, \dots, r_j$, where the weights associated with each one of these are $1, e^{V_1}, \dots, e^{V_i}, \dots, e^{V_j}$, respectively. Since $r_1 \ge \dots \ge r_i \ge \dots \ge r_j \ge 0$, it follows that $f(\{1,\dots,j\},\boldsymbol{V}) \ge f(\{i,\dots,j\},\boldsymbol{V})$. $\square$

We are now ready to show Theorem 5.1

Proof of Theorem 5.1

Proof. Let $\hat{\boldsymbol{V}}^1, \dots, \hat{\boldsymbol{V}}^G$ denote the possible realizations of the mean utility vector $\boldsymbol{V}$. The first part of the approximation guarantee follows immediately from the fact that for each $g = 1, \dots, G$, a revenue-ordered assortment solves the problem $\max_{S \subseteq \{1,\dots,n\}} f(S, \hat{\boldsymbol{V}}^g)$, giving us the approximation guarantee of G (see Talluri and van Ryzin, 2004). To establish the approximation guarantee of $\lceil n/2 \rceil$, let S^* be the optimal solution to the Mixture Logit problem. We partition the assortment S^* into k assortments $S_1^*, \dots, S_k^*$ such that each assortment S_ℓ^* contains consecutive products. That is, each assortment S_ℓ^* is of the form

$$\{i_\ell^*, i_\ell^* + 1, \dots, j_\ell^*\}$$

for some i_ℓ^* and j_ℓ^*, with $i_\ell^* \le j_\ell^*$ and

$$i_1^* \le j_1^* \ < \ i_2^* \le j_2^* \ < \ \dots \ < \ i_{k-1}^* \le j_{k-1}^* \ < \ i_k^* \le j_k^*.$$

Although each S_ℓ^* may contain a single product with $i_\ell^* = j_\ell^*$, we observe that the number of assortments k in the partition never has to be greater than $\lceil n/2 \rceil$. The desired result follows by

$$\begin{aligned} Z^* &= \mathbb{E}\{f(S^*, \boldsymbol{V})\} = \mathbb{E}\{f(\cup_{\ell=1}^k S_\ell^*, \boldsymbol{V})\} \le \sum_{\ell=1}^k \mathbb{E}\{f(S_\ell^*, \boldsymbol{V})\} = \sum_{\ell=1}^k \mathbb{E}\{f(\{i_\ell^*, \dots, j_\ell^*\}, \boldsymbol{V})\} \\ &\le \sum_{\ell=1}^k \mathbb{E}\{f(\{1, \dots, j_\ell^*, \}, \boldsymbol{V})\} \le \sum_{\ell=1}^k \max_{i=1,\dots,n} \mathbb{E}\{f(\{1,\dots,i\}, \boldsymbol{V})\} \le \lceil n/2 \rceil \max_{i=1,\dots,n} \mathbb{E}\{f(\{1,\dots,i\}, \boldsymbol{V})\}, \end{aligned}$$

where the first and second inequalities use the first and second parts of Lemma 5.2, respectively. $\square$

To see a simple application of Theorem 5.1, consider the case where we serve two customer segments, a price sensitive and a quality sensitive market segment. Price sensitive customers associate the vector of mean utilities $\hat{\boldsymbol{V}}^1 = (\hat{V}_1^1, \dots, \hat{V}_n^1)$ with the products, whereas quality sensitive customers associate the vector of mean utilities $\hat{\boldsymbol{V}}^2 = (\hat{V}_1^2, \dots, \hat{V}_n^2)$. Theorem 5.1 implies that the expected revenue from the best revenue-ordered assortment is at least half of the optimal expected revenue.

In the next proposition, we show that the approximation guarantee in Theorem 5.1 is tight, so that there are instances of the Mixture Logit where the expected revenue from the best revenue-ordered assortment deviates from the optimal by a factor arbitrarily close to $\min\{G, \lceil n/2 \rceil\}$.

Proposition 5.3 (Tight Guarantee) *There are instances of the* Mixture Logit *problem such that the expected revenue from the best revenue-ordered assortment deviates from the optimal by a factor that is arbitrarily close to* $\min\{G, \lceil n/2 \rceil\}$.

Proof. We construct a problem instance with G possible realizations of the vector of mean utilities and $n = 2G - 1$ products such that the expected revenue from the best revenue-ordered assortment deviates from the optimal by a factor arbitrarily close to $G = \lceil n/2 \rceil$. To simplify the presentation, we give a problem instance with $G = 3$ and $n = 5$ such that the expected revenue from the best revenue-ordered assortment deviates from the optimal by a factor arbitrarily close to three. Once we give this problem instance, it is easy to see how to generalize this problem instance to an arbitrary G.

We chose $\delta > 0$ and consider the following instance of the Mixture Logit problem. There are five products. There are three possible realizations of $\boldsymbol{V}$, which we denote by $\hat{\boldsymbol{V}}^1$, $\hat{\boldsymbol{V}}^2$ and $\hat{\boldsymbol{V}}^3$. The next table gives the revenues of the products and the values of $\hat{\boldsymbol{V}}^1$, $\hat{\boldsymbol{V}}^2$ and $\hat{\boldsymbol{V}}^3$. Each column in this table corresponds to a product.

Product	1	2	3	4	5
Revenues	δ	δ^2	δ^3	δ^4	δ^5
$\hat{\boldsymbol{V}}^1$	$-\log\delta$	$-2\log\delta$	$-\infty$	$-\infty$	$-\infty$
$\hat{\boldsymbol{V}}^2$	$-\infty$	$-\infty$	$-\log\delta$	$-2\log\delta$	$-\infty$
$\hat{\boldsymbol{V}}^3$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\log\delta$

The probabilities of observing the three realizations of the vector of mean utilities are $\delta^4/(\delta^4+\delta^2+1)$, $\delta^2/(\delta^4+\delta^2+1)$ and $1/(\delta^4+\delta^2+1)$.

The next table gives the expected revenue provided by each possible revenue-ordered assortment and the assortment $\{1,3,5\}$.

S	$\mathbb{E}\{f(S,\boldsymbol{V})\}$
$\{1\}$	$\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{1}{1+\delta^{-1}}$
$\{1,2\}$	$\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{2}{1+\delta^{-1}+\delta^{-2}}$
$\{1,2,3\}$	$\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{2}{1+\delta^{-1}+\delta^{-2}} + \frac{\delta^2}{\delta^4+\delta^2+1} \times \frac{\delta^2}{1+\delta^{-1}}$
$\{1,2,3,4\}$	$\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{2}{1+\delta^{-1}+\delta^{-2}} + \frac{\delta^2}{\delta^4+\delta^2+1} \times \frac{2\,\delta^2}{1+\delta^{-1}+\delta^{-2}}$
$\{1,2,3,4,5\}$	$\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{2}{1+\delta^{-1}+\delta^{-2}} + \frac{\delta^2}{\delta^4+\delta^2+1} \times \frac{2\,\delta^2}{1+\delta^{-1}+\delta^{-2}} + \frac{1}{\delta^4+\delta^2+1} \times \frac{\delta^4}{1+\delta^{-1}}$
$\{1,3,5\}$	$\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{1}{1+\delta^{-1}} + \frac{\delta^2}{\delta^4+\delta^2+1} \times \frac{\delta^2}{1+\delta^{-1}} + \frac{1}{\delta^4+\delta^2+1} \times \frac{\delta^4}{1+\delta^{-1}}$

The two terms in the expected revenue from assortment $\{1\}$ can be bounded by $\delta^4/(\delta^4+\delta^2+1) \le \delta^4$ and $1/(1+\delta^{-1}) \le 1/\delta^{-1}$. Therefore, the expected revenue from assortment $\{1\}$ is bounded by δ^5. We bound the two terms in the expected revenue from assortment $\{1,2\}$ as $\delta^4/(\delta^4+\delta^2+1) \le \delta^4$ and $2/(1+\delta^{-1}+\delta^{-2}) \le 2/\delta^{-2}$, which implies that the expected revenue from assortment $\{1,2\}$ is bounded by $2\,\delta^6$. Continuing in the same fashion, the expected revenues from assortments $\{1,2,3\}$, $\{1,2,3,4\}$ and $\{1,2,3,4,5\}$ are bounded by $2\,\delta^6+\delta^5$, $2\,\delta^6+2\,\delta^6$ and $2\,\delta^6+2\,\delta^6+\delta^5$, respectively. Therefore, the expected revenue from a revenue-ordered assortment never exceeds $4\,\delta^6+\delta^5$. Noting the expected revenue from assortment $\{1,3,5\}$, the ratio between the expected revenue from the optimal assortment and the expected revenue from the best revenue-ordered assortment is at least

$$\frac{\frac{\delta^4}{\delta^4+\delta^2+1} \times \frac{1}{1+\delta^{-1}} + \frac{\delta^2}{\delta^4+\delta^2+1} \times \frac{\delta^2}{1+\delta^{-1}} + \frac{1}{\delta^4+\delta^2+1} \times \frac{\delta^4}{1+\delta^{-1}}}{4\,\delta^6 + \delta^5}$$

$$= \frac{\frac{1}{\delta^4+\delta^2+1} \times \frac{\delta^{-1}}{1+\delta^{-1}} + \frac{1}{\delta^4+\delta^2+1} \times \frac{\delta^{-1}}{1+\delta^{-1}} + \frac{1}{\delta^4+\delta^2+1} \times \frac{\delta^{-1}}{1+\delta^{-1}}}{4\,\delta + 1},$$

which can be made arbitrarily close to three by choosing δ small enough. □

5.2 Guarantees based on the Product Revenues and Mean Utilities

The proof of Proposition 5.3 provides instances of the Mixture Logit problem where the performance of revenue-ordered assortments becomes progressively worse as the number of possible realizations of the vector of mean utilities or the number of products increases. However, in the instance that we constructed in the proof of the proposition, the revenues of the products are $\delta, \delta^2, \ldots,$ and we let δ become arbitrarily small. In this case, the revenues of the products differ from each other by orders of magnitude, which may not be realistic for many practical applications. So, a natural question is whether revenue-ordered assortments perform well when the revenues of the products are not too different from each other. In Theorem 5.4, we give an affirmative answer to this question by establishing an approximation guarantee for revenue-ordered assortments in terms of the ratio between the largest and smallest product revenues. The proof is given in Appendix C in the Online Supplement.

Theorem 5.4 (Guarantee Based on Product Revenues) *Revenue-ordered assortments provide an approximation guarantee of* $e \log(e\, r_1/r_n)$ *for the* Mixture Logit *problem.*

The important observation from the theorem above is that the approximation guarantee provided by revenue-ordered assortment scales logarithmically with the ratio of the largest and smallest product revenues. Thus, if the revenues of the products do not differ from each other by orders of magnitude, then Theorem 5.4 suggests that revenue-ordered assortments should intuitively provide a constant factor approximation guarantee.

Finally, we can also provide approximation guarantees based on the distribution of the mean utility vector. In particular, we give an approximation guarantee for revenue-ordered assortment as a function of how much the mean utility of each product differs from its expectation and how much the mean utilities of different products differ from each other. This result is stated in the next proposition, and the proof is given in Appendix D in the Online Supplement.

Proposition 5.5 (Guarantees Based on Mean Utilities) *If* $\max_{i=1,\ldots,n} |V_i - \mathbb{E}V_i| \le \delta$ *almost surely, then revenue-ordered assortments provide an approximation guarantee of* $e^{4\delta}$*. Similarly, if* $\max_{i,j=1,\ldots,n} |V_i - V_j| \le \delta$ *almost surely, then revenue-ordered assortments provide an approximation guarantee of* $e^{4\delta}$*.*

6 Multi-Period Capacity Allocation and Nesting-by-Fare-Order Property

In this section, we extend our formulation to the multi-period capacity allocation setting of Talluri and van Ryzin (2004). The model of Talluri and van Ryzin (2004) is one of the pioneering works in revenue management that demonstrates the importance of incorporating choice behavior in operational decisions. Let us briefly review the setup for this problem. We have an initial capacity of C seats on a flight leg that must be allocated over T periods. There are n products (fare classes) that can be offered to customers, indexed by $\{1, 2, \ldots, n\}$. If we sell one ticket for fare class i, then we generate a revenue of r_i. We index the fare classes such that $r_1 \geq r_2 \geq \ldots \geq r_n$. In each period, based on the remaining capacity, we must decide on the assortment of fare classes to offer to an arriving customer, who chooses a fare class from the assortment according to a mixture-of-logits choice model. The goal is to determine the revenue-maximizing policy for allocating the capacity. In this paper, for simplicity, we assume that there is exactly one customer arriving in each period; all of our results immediately extend to the case where there is a positive probability that no customer shows up in each period.

We focus on the special case where, in each period, we only offer a revenue-ordered assortment. Under this restriction, we will show that the mixture-of-logits model possesses the nesting-by-fare-order property. In other words, we will establish that as we have more remaining capacity in a period, we offer a larger revenue-ordered assortment. This result implies that as the remaining capacity on the flight leg gets smaller, fare classes with lower fares stop being offered first. An important managerial implication of this result is that revenue-ordered assortments can be implemented by using nested protection level policies, which is the standard tool in traditional revenue management systems. Therefore, our policy that offers only revenue-ordered assortments can be easily integrated with the existing revenue management controls. We emphasize that nesting-by-fare-order property does not necessarily hold under the mixture-of-logits choice model if we do not focus on revenue-ordered assortments.

To establish this result, for each $x \in \{0, 1, \ldots, C\}$, let $J_t(x)$ denote the maximum expected revenue when we have x units of inventory and t periods remaining, given that we can offer only revenue-ordered assortments. Since we offer revenue-ordered assortments in each period, $J_t(\cdot)$ satisfies the following dynamic programming equation:

$$
\begin{aligned}
J_t(x) &= \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \sum_{i=1}^{\ell} \pi_i(\{1,\ldots,\ell\}, \boldsymbol{V})\,(r_i + J_{t-1}(x-1)) + \pi_0(\{1,\ldots,\ell\}, \boldsymbol{V}) J_{t-1}(x) \right\} \\
&= \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \sum_{i=1}^{\ell} \pi_i(\{1,\ldots,\ell\}, \boldsymbol{V})\,(r_i - \Delta J_{t-1}(x)) \right\} + J_{t-1}(x) \ ,
\end{aligned}
$$

where the expectation is taken with respect to the random vector $\boldsymbol{V}$, $\pi_0(S, \boldsymbol{V})$ is the no purchase probability when we offer assortment S and the realization of the mean utilities are $\boldsymbol{V}$, and $\Delta J_{t-1}(x) = J_{t-1}(x) - J_{t-1}(x-1)$ denotes the marginal value of capacity. The boundary conditions of the dynamic program are $V_0(x) = 0$ for all x and $V_t(0) = 0$ for all t. Let $\ell_t^*(x)$ denote the optimal product index in the problem on the right side above when we have x units of remaining inventory and t periods to go, corresponding to the revenue-ordered assortment $\{1, 2, \ldots, \ell_t^*(x)\}$.

The main result of this section is stated in the following theorem.

Theorem 6.1 (Nested-by-Fare-Order Property) *For each t, $\ell_t^*(x)$ is non-decreasing in x. Moreover, for each x, $\ell_t^*(x)$ is non-increasing in t.*

Thus, we offer a larger revenue-ordered assortment in a period when we have more capacity. As discussed in Talluri and van Ryzin (2004), Theorem 6.1 implies that we can implement the resulting optimal policy using nested protection level policies, which represent a capacity allocation method traditionally used in the revenue management industry. In addition, Theorem 6.1 enables us to simplify the computation and storage of the optimal policy. In general, we need to keep the optimal product index $\ell_t^*(x)$ for each period t and for each remaining inventory level x, which, in turns, gives the optimal revenue-ordered assortment to offer for each period and for each inventory level. This requires keeping a total of $O(CT)$ numbers. However, since $\ell_t^*(x)$ is non-decreasing in x, we can simply keep track of n capacity thresholds in each period, representing the smallest amount of capacity at which we still offer a particular fare class. This requires keeping a total of $O(nT)$ numbers. In practical applications, the number of fare classes n is generally much smaller than C.

The proof of Theorem 6.1 makes use of the following lemma, which shows that when we increment the revenue of every product by a positive constant, the corresponding optimal revenue-ordered assortment becomes larger. To facilitate our discussion, let us introduce the following notation. For each $\delta \geq 0$ and each assortment S, let $f^\delta(S, \boldsymbol{V})$ denote the revenue under assortment S and mean utility $\boldsymbol{V}$, when the revenue of every product is increased by δ; that is,

$$f^\delta(S, \boldsymbol{V}) \equiv \sum_{i \in S}(r_i + \delta)\pi_i(S, \boldsymbol{V}) = \sum_{i \in S} r_i \pi_i(S, \boldsymbol{V}) + \delta \sum_{i \in S} \pi_i(S, \boldsymbol{V}) = f^0(S, \boldsymbol{V}) + \delta \sum_{i \in S} \pi_i(S, \boldsymbol{V}) \ .$$

Let $\{1, 2, \ldots, \ell^\delta\}$ denote the revenue-ordered assortment that maximizes the expected revenue under f^δ; that is,

$$\ell^\delta = \arg \max_{\ell = 1, \ldots, n} \mathbb{E}\left\{ f^\delta\left(\{1, \ldots, \ell\}, \boldsymbol{V}\right)\right\} \ .$$

Lemma 6.2 (Larger Revenues Lead to Larger Assortments) *For all $\delta \geq 0$, $\ell^0 \leq \ell^\delta$.*

Proof. We will prove this by contradiction. Suppose on the contrary that $\ell^0 > \ell^\delta$ for some $\delta \geq 0$. By definition of ℓ^δ and using the fact that $f^\delta(S, \boldsymbol{V}) = f^0(S, \boldsymbol{V}) + \delta \sum_{i \in S} \pi_i(S, \boldsymbol{V})$, we have that

$$\begin{aligned}
0 \ &\geq \ \mathbb{E}\left\{f^\delta(\{1, \ldots, \ell^0\}, \boldsymbol{V})\right\} - \mathbb{E}\left\{f^\delta(\{1, \ldots, \ell^\delta\}, \boldsymbol{V})\right\} \\
&= \ \mathbb{E}\Big\{f^0(\{1, \ldots, \ell^0\}, \boldsymbol{V}) + \delta \sum_{k=1}^{\ell^0} \pi_k(\{1, \ldots, \ell^0\}, \boldsymbol{V}) - f^0(\{1, \ldots, \ell^\delta\}, \boldsymbol{V}) - \delta \sum_{k=1}^{\ell^\delta} \pi_k(\{1, \ldots, \ell^\delta\}, \boldsymbol{V})\Big\} \\
&= \ \mathbb{E}\left\{f^0(\{1, \ldots, \ell^0\}, \boldsymbol{V})\right\} - \mathbb{E}\left\{f^0(\{1, \ldots, \ell^\delta\}, \boldsymbol{V})\right\} + \delta \mathbb{E}\left\{\pi_0(\{1, \ldots, \ell^\delta\}, \boldsymbol{V}) - \pi_0(\{1, \ldots, \ell^0\}, \boldsymbol{V})\right\} \\
&> \ \mathbb{E}\left\{f^0(\{1, \ldots, \ell^0\}, \boldsymbol{V})\right\} - \mathbb{E}\left\{f^0(\{1, \ldots, \ell^\delta\}, \boldsymbol{V})\right\} \\
&\geq \ 0 \ ,
\end{aligned}$$

where the second equality follows by noting that $\sum_{k=1}^{\ell^0} \pi_k(\{1,\ldots,\ell^0\},\boldsymbol{V}) = 1 - \pi_0(\{1,\ldots,\ell^0\},\boldsymbol{V})$ and $\sum_{k=1}^{\ell^\delta} \pi_k(\{1,\ldots,\ell^\delta\},\boldsymbol{V}) = 1 - \pi_0(\{1,\ldots,\ell^\delta\},\boldsymbol{V})$, and the second inequality follows from

$$\mathbb{E}\left\{\pi_0(\{1,\ldots,\ell^\delta\},\boldsymbol{V}) - \pi_0(\{1,\ldots,\ell^0\},\boldsymbol{V})\right\} = \mathbb{E}\left\{\frac{1}{1+\sum_{i=1}^{\ell^\delta} e^{V_i}} - \frac{1}{1+\sum_{i=1}^{\ell^0} e^{V_i}}\right\} \;>\; 0$$

because $\ell^\delta < \ell^0$. The final inequality in the first displayed chain of inequalities above follows from the definition of ℓ^0. This is a contradiction, and thus, it must be the case that $\ell^0 \le \ell^\delta$ for all $\delta \ge 0$. □

We also need the following lemma, which characterizes the structure of the value functions. In particular, this lemma shows that the value functions are concave in the remaining inventory and the first differences of the value functions decrease as we approach the end of the selling horizon. The proof of this result is by now standard within the revenue management literature (see, for example, Talluri and van Ryzin, 2004), and we omit the details.

Lemma 6.3 (Properties of the Value Functions) *For $t = 1, 2, \ldots, T$ and $x = 1, 2, \ldots, C$,*

$$\Delta J_t(x) \le \Delta J_t(x-1) \qquad \text{and} \qquad \Delta J_t(x) \ge \Delta J_{t-1}(x)\ .$$

We are now ready to show Theorem 6.1.

Proof of Theorem 6.1

Proof. By definition,

$$\begin{aligned}
\ell_t^*(x) &= \arg\max_{\ell=1,\ldots,n} \mathbb{E}\left\{\sum_{i=1}^{\ell} \pi_i(\{1,\ldots,\ell\},\boldsymbol{V})\,(r_i - \Delta J_{t-1}(x))\right\} \\
\ell_t^*(x-1) &= \arg\max_{\ell=1,\ldots,n} \mathbb{E}\left\{\sum_{i=1}^{\ell} \pi_i(\{1,\ldots,\ell\},\boldsymbol{V})\,(r_i - \Delta J_{t-1}(x-1))\right\} .
\end{aligned}$$

Since $\Delta J_{t-1}(x-1) \ge \Delta J_{t-1}(x)$ by Lemma 6.3, applying Lemma 6.2 with $\ell^0 = \ell_t^*(x-1)$ and $\delta = \Delta J_{t-1}(x-1) - \Delta J_{t-1}(x)$, it follows that $\ell_t^*(x-1) \le \ell_t^*(x)$, which proves the first part of the theorem.

To establish the second part, note that

$$\begin{aligned}
\ell_t^*(x) &= \arg\max_{\ell=1,\ldots,n} \mathbb{E}\left\{\sum_{i=1}^{\ell} \pi_i(\{1,\ldots,\ell\},\boldsymbol{V})\,(r_i - \Delta J_{t-1}(x))\right\} \\
\ell_{t-1}^*(x) &= \arg\max_{\ell=1,\ldots,n} \mathbb{E}\left\{\sum_{i=1}^{\ell} \pi_i(\{1,\ldots,\ell\},\boldsymbol{V})\,(r_i - \Delta J_{t-2}(x))\right\} .
\end{aligned}$$

Since $\Delta J_{t-1}(x) \ge \Delta J_{t-2}(x)$ by Lemma 6.3, applying Lemma 6.2 with $\ell^0 = \ell_t^*(x)$ and $\delta = \Delta J_{t-1}(x) - \Delta J_{t-2}(x)$, it follows that $\ell_t^*(x) \le \ell_{t-1}^*(x)$, which is the desired result. □

7 Numerical Experiments

In this section, our goal is to investigate the empirical performance of revenue-ordered assortments on practical instances of the Mixture Logit problem. We consider two sets of test problems in our numerical experiments. In the first set, we serve a market with a relatively small number of customer segments. Customers in different segments associate different mean utilities with the products. The objective is to maximize the expected revenue over all customer segments. In the second set, the mean utilities that a customer associates with the products have a continuous distribution. The objective is to maximize the expected revenue over all customers.

7.1 Small Number of Market Segments

Experimental Setup: We assume that we serve a market with G customer segments. A customer in segment G associates the mean utilities $\hat{\boldsymbol{V}}^g = (\hat{V}_1^g, \ldots, \hat{V}_n^g)$ with the products. The relative size of customer segment g is α^g. We view α^g as the probability of getting a customer in segment g, where we normalize $\alpha^1, \ldots, \alpha^G$ so that $\sum_{g=1}^G \alpha^g = 1$. This setup corresponds to the situation where the vector of mean utilities has G possible realizations $\hat{\boldsymbol{V}}^1, \ldots, \hat{\boldsymbol{V}}^G$ and the vector of mean utilities takes value $\hat{\boldsymbol{V}}^g$ with probability α^g.

Noting that the approximation guarantees of revenue-ordered assortments in Sections 5.1 and 5.2 depend on G, n and r_1/r_n, we vary $(G, n, r_1/r_n)$ over $\{2, 5, 10\} \times \{10, 25, 50\} \times \{10, 10^2, 10^3\}$ to obtain 27 problem classes. In each problem class, we randomly generate 10,000 problem instances. For each problem instance, we find the best revenue-ordered assortment by checking the expected revenue provided by every assortment of the form $\{1, 2, \ldots, i\}$. Let $\mathsf{Approx}(k)$ denote the expected revenue provided by the best revenue-ordered assortment for problem instance k. Bront et al. (2009) show that the Mixture Logit problem can be formulated as a mixed integer program, enabling us to compute the optimal assortment for each problem instance by using a mixed integer programming solver. Let $\mathsf{Opt}(k)$ be the expected revenue provided by the optimal assortment for problem instance k. Our goal is to compare $\mathsf{Opt}(k)$ with $\mathsf{Approx}(k)$.

We use the following strategy to generate the problem instances. To come up with the possible realizations of the vector of mean utilities, we sample σ_i from the uniform distribution over $[0, 1]$ for each product i. The parameter σ_i characterizes the variability in the mean utility of product i among the different customer segments. If σ_i is close to zero, then product i is a staple product and its mean utility does not show too much variability among the different customer segments. This corresponds a situation where the different customer segments evaluate product i in a similar fashion. If σ_i is close to one, then product i is a specialty product and its mean utility shows large variability among the different customer segments. This corresponds to a situation where the different customer segments evaluate product i in a drastically different fashion. In this case, for each customer segment g and product i, we sample ϑ_i^g from the uniform distribution over $[0, 10]$ and set $\hat{V}_i^g = \log((1 - \sigma_i)\, \vartheta_i^g / n)$ with probability $1/2$ and $\hat{V}_i^g = \log((1 + \sigma_i)\, \vartheta_i^g / n)$ with probability $1/2$. To see the motivation behind our choice of the mean utilities, we note that if σ_i takes a value close to zero, then $e^{\hat{V}_i^g}$ takes a value close to ϑ_i^g / n. On the other hand, if σ_i takes a value close to one, then $e^{\hat{V}_i^g}$ either takes a value close to zero or takes a value close to $2\vartheta_i^g/n$. Therefore the parameter σ_i indeed captures how much the mean utility of a product differs among the different customer segments. Furthermore, the expectation of $e^{\hat{V}_i^g}$ is $5/n$. If $e^{\hat{V}_i^g}$ takes a value close to its expectation and we offer all products, then the probability that a customer does not purchase anything is $1/(1+(n \times 5/n)) = 1/6$, indicating that a customer does not purchase anything with a significant probability even if we offer all products.

To come up with the revenues of the products, we always set $r_n = 1$. Depending on the problem class, we set $r_1 = 10$, 10^2 or 10^3. We sample the remaining revenues from the uniform distribution over $[r_n, r_1]$. Finally, to come up with the relative sizes of the customer segments, we sample β^g from the uniform distribution over $[0,1]$ for each customer segment g and set $\alpha^g = \beta^g / \sum_{h=1}^G \beta^h$.

Numerical Results: Table 1 gives an overview of our numerical results. Each row in this table corresponds to a problem class described by the triplet $(G, n, r_1/r_n)$. The first column lists all problem classes. Recall that we generate 10,000 problem instances in each problem class. The second column shows the percentage of problem instances out of 10,000 for which $\mathsf{Approx}(k)$ is not equal to $\mathsf{Opt}(k)$, corresponding to the frequency with which revenue-ordered assortments are not optimal. To facilitate our discussion, we denote these problem instances by defining the set $\mathsf{NonOpt} = \{k \in \{1, \ldots, 10,000\} : \mathsf{Approx}(k) < \mathsf{Opt}(k)\}$. For each problem class, we are interested in the distribution of the percent optimality gaps across all 10,000 problem instances and across the problem instances in NonOpt. In particular, these percent optimality gaps are captured by the sets of numbers

$$\mathsf{GapAll} = \left\{ 100 \times \frac{\mathsf{Opt}(k) - \mathsf{Approx}(k)}{\mathsf{Opt}(k)} : k = 1, \ldots, 10,000 \right\}$$

$$\mathsf{GapNonOpt} = \left\{ 100 \times \frac{\mathsf{Opt}(k) - \mathsf{Approx}(k)}{\mathsf{Opt}(k)} : k \in \mathsf{NonOpt} \right\}.$$

The third and fourth columns in Table 1 provide aggregate statistics for the set GapAll. In particular, the third column reports the average of the optimality gaps across all problem instances, whereas the fourth column shows the 95th percentile. The fifth and sixth columns give similar aggregate statistics, but they focus on the optimality gaps in the set $\mathsf{GapNonOpt}$.

The results in Table 1 indicate that revenue-ordered assortments perform very well across all problem instances. The average and 95th percentile of the optimality gaps are only a fraction of a percent. When we focus only on the problem instances in NonOpt, the average optimality gap is still less than 1% and the 95th percentile does not exceed 3.26%. As the number of products increases, the average and the 95th percentile slightly increase, which is in agreement with the approximation guarantees we derive in Section 5. As the number of customer segments or the ratio between the largest and smallest revenues increases, the average and 95th percentile of the optimality gaps remain stable and they even show a slightly decreasing trend.

Figure 1 shows the histograms of the optimality gaps in the set $\mathsf{GapNonOpt}$ for problem classes $(2, 10, 10)$ and $(10, 50, 10^3)$. The two histograms indicate that the optimality gaps decline sharply and the tails of the histogram tend to be quite thin. Generally, less than 2% of the problem instances have optimality gaps exceeding 1%. Although we give histograms for only two problem classes, we observed similar patterns for all of the problem classes we studied.

Prob. Class			Perc. of Suboptimal Instances	Statistics for GapAll		Statistics for GapNonOpt	
G	n	r_1/r_n		Avg.	95th Pr.	Avg.	95th Pr.
2	10	10	10.27%	0.09%	0.58%	0.92%	3.22%
2	10	10^2	7.78%	0.07%	0.32%	0.91%	3.26%
2	10	10^3	7.61%	0.07%	0.31%	0.91%	3.20%
2	25	10	30.41%	0.15%	0.90%	0.49%	1.73%
2	25	10^2	24.25%	0.12%	0.74%	0.48%	1.69%
2	25	10^3	23.96%	0.11%	0.72%	0.48%	1.75%
2	50	10	50.97%	0.18%	0.94%	0.36%	1.31%
2	50	10^2	44.05%	0.14%	0.82%	0.33%	1.24%
2	50	10^3	43.72%	0.14%	0.80%	0.33%	1.24%
5	10	10	8.84%	0.07%	0.41%	0.77%	2.33%
5	10	10^2	5.91%	0.04%	0.12%	0.75%	2.14%
5	10	10^3	5.82%	0.04%	0.12%	0.74%	2.22%
5	25	10	31.69%	0.10%	0.60%	0.33%	1.03%
5	25	10^2	23.31%	0.07%	0.45%	0.30%	1.03%
5	25	10^3	23.10%	0.07%	0.45%	0.30%	1.00%
5	50	10	58.52%	0.14%	0.62%	0.24%	0.77%
5	50	10^2	49.22%	0.10%	0.49%	0.21%	0.69%
5	50	10^3	48.82%	0.10%	0.49%	0.21%	0.69%
10	10	10	6.67%	0.04%	0.21%	0.66%	1.85%
10	10	10^2	4.42%	0.02%	0.00%	0.56%	1.75%
10	10	10^3	4.45%	0.02%	0.00%	0.54%	1.76%
10	25	10	25.10%	0.06%	0.38%	0.25%	0.76%
10	25	10^2	16.13%	0.04%	0.25%	0.23%	0.70%
10	25	10^3	15.94%	0.04%	0.24%	0.23%	0.70%
10	50	10	50.88%	0.08%	0.37%	0.16%	0.47%
10	50	10^2	39.73%	0.05%	0.26%	0.13%	0.40%
10	50	10^3	39.18%	0.05%	0.26%	0.13%	0.40%

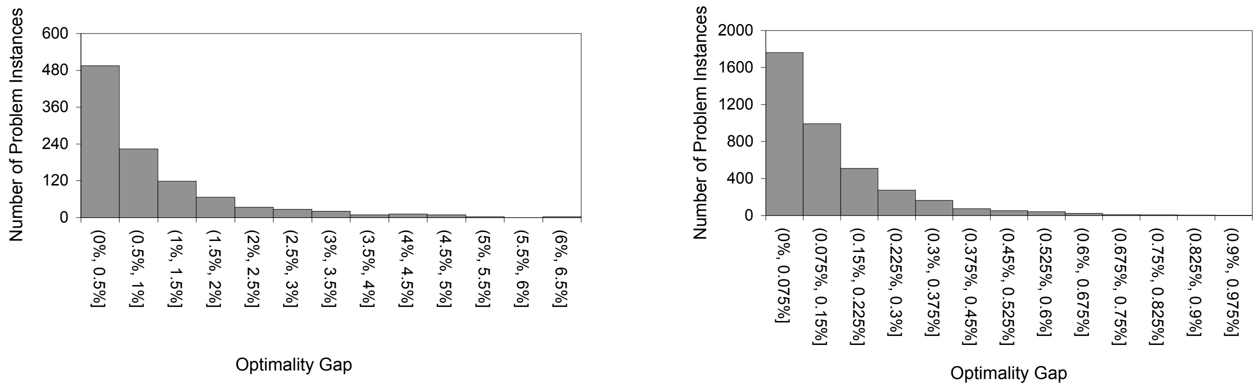
Table 1: Performance of revenue-ordered assortments when we have a small number of market segments. Each problem class includes 10,000 problem instances.

7.2 Mean Utilities Drawn from a Continuous Distribution

Experimental Setup: In the experimental setup of this section, the mean utilities of the products have a continuous distribution. In particular, we assume that a customer considers purchasing each one of the products with probability 1/2. If the customer considers purchasing product i, then the mean utility V_i associated with this product has a normal distribution. To generate problems of this nature, we proceed as follows. In each problem instance, we generate a parameter p_i for each product i. This parameter captures the inherent attractiveness of product i to a customer, with a larger value of p_i making product i more attractive to a customer. To come up with $(p_1, \ldots, p_n)$, we sample Ψ_i from the standard normal distribution and set $p_i = e^{\Psi_i} / \sum_{j=1}^n e^{\Psi_j}$. Note that we have $\sum_{i=1}^n p_i = 1$. Using the parameters $(p_1, \ldots, p_n)$, the mean utility of product i in our numerical experiments has the form

$$V_i = \log(p_i\, X_i) + \mu + \sigma\, N_i,$$

where p_i is generated as described above, X_i is a Bernoulli random variable with parameter 1/2 and N_i is a standard normal random variable. Each one of the random variables $(X_1, \ldots, X_n)$ and $(N_1, \ldots, N_n)$ are independent of others. We choose the values of the constants μ and σ such that $\mu + \sigma^2/2 = \log 10$ and $\sqrt{e^{\sigma^2/2} - 1} = \Delta$, where Δ is a parameter that we vary in our experimental setup. To see the



Problem Class $(2, 10, 10)$ Problem Class $(10, 50, 10^3)$

Figure 1: Histograms of the optimality gaps among problem instances in NonOpt for two problem classes.

reasoning behind the specific form of the mean utility V_i, we observe that if $X_i = 0$, then $e^{V_i} = 0$, which indicates that the customer does not consider purchasing product i at all. Therefore, a customer considers purchasing each product with probability 1/2. Also, if p_i is large, then the mean utility of product i tends to be large, indicating that p_i indeed captures the inherent attractiveness of product i to a customer. Furthermore, once the parameters $(p_1, \ldots, p_n)$ are fixed, we have $\mathbb{E}\{e^{V_i}\} = \frac{1}{2}\, p_i\, e^{\mu+\sigma^2/2} = 5\, p_i$. In this case, if e^{V_i} takes a value close to its expectation and we offer all products, then the probability that a customer does not purchase anything is $1/(1 + \sum_{i=1}^n 5\, p_i) = 1/6$, where we use the fact that $\sum_{i=1}^n p_i = 1$. So, the probability that a customer does not purchase anything is still significant even if we offer all products. Finally, given that $X_i = 1$, the expectation and standard deviation of e^{V_i} are respectively given by $p_i\, e^{\mu+\sigma^2/2}$ and $p_i\, e^{\mu+\sigma^2/2}\sqrt{e^{\sigma^2/2} - 1}$. Therefore, the parameter Δ corresponds to the standard deviation to mean ratio of e^{V_i} given that a customer considers purchasing product i. So, larger values of Δ yield larger variations in how much a customer is attracted to a product.

Similar to our experimental setup in Section 7.1, to come up with the revenues of the products, we always set $r_n = 1$. We sample the remaining revenues from the uniform distribution over $[r_n, r_1]$, where r_1 is a parameter that we vary in our experimental setup. In our numerical experiments, we vary $n \times r_1 \times \Delta$ over $\{10, 20\} \times \{10, 10^2, 10^3\} \times \{2, 4, 6\}$ to obtain 18 problem classes. In each problem class, we generate 1,000 problem instances

When $\boldsymbol{V} = (V_1, \ldots, V_n)$ has a continuous distribution, an important consideration is that simply computing the expected revenue $\mathbb{E}\{f(S, \boldsymbol{V})\}$ from an assortment S can be difficult, requiring the evaluation of a multi-dimensional integral. We resolve this difficulty by estimating this expectation through Monte Carlo simulation. We obtain 500 samples of $\boldsymbol{V}$ given by $\{\hat{\boldsymbol{V}}^t : t = 1, \ldots, 500\}$ and estimate $\mathbb{E}\{f(S, \boldsymbol{V})\}$ as $\frac{1}{500}\sum_{t=1}^{500} f(S, \hat{\boldsymbol{V}}^t)$. When estimating the expectation $\mathbb{E}\{f(S, \boldsymbol{V})\}$ for different assortments, we use the same set of 500 samples of $\boldsymbol{V}$. We can interpret the sample $\hat{\boldsymbol{V}}^t$ as the mean utilities that customers in market segment t associates with the products. Therefore, similar

Prob. Class			Perc. of Suboptimal	Statistics for GapAll		Statistics for GapNonOpt	
n	r_1/r_n	Δ	Instances	Avg.	95th Pr.	Avg.	95th Pr.
10	10	2	2.7%	0.01%	0.00%	0.22%	0.75%
10	10^2	4	6.6%	0.02%	0.05%	0.30%	0.89%
10	10^3	6	13.7%	0.04%	0.32%	0.30%	0.91%
10	10	2	2.3%	0.01%	0.00%	0.22%	0.85%
10	10^2	4	1.6%	0.00%	0.00%	0.19%	0.66%
10	10^3	6	1.8%	0.00%	0.00%	0.08%	0.26%
10	10	2	2.1%	0.00%	0.00%	0.16%	0.29%
10	10^2	4	1.7%	0.00%	0.00%	0.17%	0.43%
10	10^3	6	2.4%	0.00%	0.00%	0.17%	0.46%
20	10	2	4.9%	0.00%	0.00%	0.09%	0.25%
20	10^2	4	7.3%	0.01%	0.03%	0.14%	0.53%
20	10^3	6	11.2%	0.02%	0.08%	0.14%	0.49%
20	10	2	4.8%	0.00%	0.00%	0.07%	0.20%
20	10^2	4	8.2%	0.01%	0.06%	0.12%	0.35%
20	10^3	6	7.2%	0.01%	0.03%	0.08%	0.28%
20	10	2	5.2%	0.01%	0.00%	0.11%	0.42%
20	10^2	4	5.4%	0.01%	0.02%	0.11%	0.33%
20	10^3	6	7.5%	0.01%	0.04%	0.11%	0.36%

Table 2: Performance of revenue-ordered assortments when the mean utilities have a continuous distribution. Each problem class includes 1,000 problem instances.

to Section 7.1, the experimental setup in this section can be interpreted as one where we have multiple customer segments, but the number of segments in the experimental setup of this section is quite large.

Numerical Results: We present our numerical results in Table 2. The format of this table is identical to that of Table 1. The results indicate that revenue-ordered assortments continue to perform remarkably well. In the most problematic problem class, revenue-ordered assortments are not optimal in no more than 14% of the problem instances. Even if we consider problem instances where revenue-ordered assortments are not optimal, the 95th percentile of the optimality gaps is less than a percent. As Δ increases and the mean utilities associated with the products become more variable, we generally observe a slight increase in the optimality gaps, but even the largest optimality gaps are no more than a percent. Overall, our results in Sections 7.1 and 7.2 indicate that revenue-ordered assortments perform quite well over a wide range of problem parameters.

8 Conclusions

We studied the performance of revenue-ordered assortments for assortment optimization problems under the multinomial logit model with random parameters. We identified two practically useful cases with a special structure on the random nature of the choice model parameters, where revenue-ordered assortments remain optimal. When the randomness in the choice model parameters does not follow a special structure, we derived tight approximation guarantees for revenue-ordered assortments. In a multi-period model, we showed that focusing on revenue-ordered assortments allows us to use nested protection levels. Our numerical results indicated that revenue-ordered assortments perform very well in practice.

This work opens up a number of directions for future research. To begin with, it would be interesting to extend the reach of revenue-ordered assortments further by identifying other cases where such assortments remain optimal. Also, it would be useful to find a class of assortments that provide a constant factor approximation guarantee for the assortment optimization problem. One may also try to extend the analysis in this paper to more complex choice models such as the mixture-of-nested-logits. Finally, the multinomial logit model with random parameters provides a rich family of models for representing the underlying customer choice process. Finding an effective parameter estimation technique for this class of choice models is an exciting future research area.

Acknowledgement

We would like to thank the department editor Bill Cooper, the senior editor, and the referees for their helpful and insightful comments, which greatly improved the results and presentation in the paper. We are also grateful to Guillermo Gallego for encouraging us to pursue research in this area. This research is supported in part by National Science Foundation through grants CMMI-1158658, CMMI-1158659, CMMI-1157569, and CMMI-0969113.

References

Ben-Akiva, M., and S. Lerman. 1985. *Discrete choice analysis: Theory and application to travel demand.* Cambridge, MA: MIT Press.

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

Davis, J. M., G. Gallego, and H. Topaloglu. 2011. Assortment optimization under variants of the nested logit model. Working paper, Cornell University.

Farias, V. F., S. Jagabathula, and D. Shah. 2011. A Non-Parametric Approach to Modeling Choice with Limited Data. *Management Science* (to appear).

Farias, V. F., S. Jagabathula, and D. Shah. 2012. Sparse Choice Models. Technical report, MIT, Cambridge, MA.

Gallego, G., G. Iyengar, R. Phillips, and A. Dubey. 2004. Managing flexible products on a network. Working Paper, Columbia University.

Gallego, G., R. Ratliff, and S. Shebalov. 2011. A General Attraction Model and an Efficient Formulation for the Network Revenue Management Problem. Technical report, Columbia University, New York, NY.

Gallego, G., and H. Topaloglu. 2012. Constrained assortment optimization for the nested logit model. Working paper, *Cornell University.*

Garey, M. R., and D. S. Johnson. 1979. *Computers and Intractability: A Guide to The Theory of NP-Completeness*. W. H. Freeman.

Goyal, V., R. Levi, and D. Segev. 2013. Near-Optimal Algorithms for the Assortment Planning Problem under Dynamic Substitution and Stochastic Demand. Working Paper, Columbia University.

Jagabathula, S., V. Farias, and D. Shah. 2011. Assortment Optimization under General Choice. In *INFORMS Conference, Charlotte, NC.*

Kok, A. G., M. Fisher, and R. Vaidyanathan. 2008. Assortment planning: Review of literature and industry practice. In *Retail Supply Chain Management*: Springer.

Kunnumkal, S., and H. Topaloglu. 2008. A refined deterministic linear program for the network revenue management problem with customer choice behavior. *Naval Research Logistics* 55 (6): 563–580.

Li, G., and P. Rusmevichientong. 2012. Technical Note: A Simple Greedy Algorithm for Assortment Optimization in the Two-Level Nested Logit Model. Working Paper, USC.

Liu, Q., and G. J. van Ryzin. 2008. On the choice-based linear programming model for network revenue management. *Manufacturing and Service Operations Management* 10 (2): 288–310.

McFadden, D., and K. Train. 2000. Mixed MNL Models for Discrete Response. *Journal of Applied Economics* 15:447–470.

Meissner, J., A. Strauss, and K. Talluri. 2012. An Enhanced Concave Program Relaxation for Choice Network Revenue Management. *Production and Operations Management* 22 (1): 71–87.

Meissner, J., and A. K. Strauss. 2012. Network Revenue Management with Inventory-Sensitive Bid Prices and Customer Choice. *European Journal of Operational Research* 216:459–468.

Mendez-Diaz, I., J. J. M. Bront, G. Vulcano, and P. Zabala. 2010. A Branch-and-Cut Algorithm for the Latent-Class Logit Assortment Problem. *Discrete Applied Mathematics* 36:383–390.

Rusmevichientong, P., Z.-J. M. Shen, and D. B. Shmoys. 2010. Dynamic Assortment Optimization with a Multinomial Logit Choice Model and Capacity Constraint. *Operations Research* 58 (6): 1666–1680.

Rusmevichientong, P., and H. Topaloglu. 2012. Robust Assortment Optimization in Revenue Management Under the Multinomial Logit Choice Model. *Operations Research* 60 (4): 865–882.

Talluri, K. 2010. A randomized concave programming method for choice network revenue management. Working paper, Universitat Pompeu Fabra, Barcelona, Spain.

Talluri, K., and G. van Ryzin. 2004. *The theory and practice of revenue management*. New York: Springer.

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

Train, K. 2003. *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.

Vossen, T. W. M., and D. Zhang. 2012. A Dynamic Disaggregation Approach to Approximate Linear Programs for Network Revenue Management. Technical report, University of Colorado at Boulder, Boulder, CO.

Zhang, D., and D. Adelman. 2009. An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Science* 43 (3): 381–394.

Online Supplement

"Assortment Optimization under the Multinomial Logit Model with Random Choice Parameters"

Paat Rusmevichientong[5], David Shmoys[6], Chaoxu Tong[7], Huseyin Topaloglu[8]

A Proof of Theorem 3.2

Proof. It is clear that 2-Class Logit is in NP. We show that we can transform an arbitrary instance of Partition, which is a well-known NP-complete problem (see, for example, Garey and Johnson, 1979), to an equivalent 2-Class Logit problem. The Partition problem is defined as follows.

Partition
Inputs: Set of items indexed by $1, 2, \ldots, n$ and the size $c_i \in \mathbb{Z}_+$ associated with each item i.
Question: Is there a subset $S \subseteq \{1, \ldots, n\}$ such $\sum_{i \in S} c_i = \sum_{i \in \{1,\ldots,n\} \setminus S} c_i$?

Let $T = \frac{1}{2} \sum_{i=1}^n c_i$. Note that $\sum_{i \in S} c_i = \sum_{i \in \{1,\ldots,n\} \setminus S} c_i$ if and only if $\sum_{i \in S} c_i = T$. Therefore, we may assume without loss of generality that $T \in \mathbb{Z}_+$. We construct an instance of the 2-Class Logit problem as follows. We have $n+1$ products indexed by $1, \ldots, n, n+1$. We set the revenues of the products and the two realizations of the mean utilities as

$$r_i = \begin{cases} (1+8T)(3+4T) & \text{if } i = 1, \ldots, n \\ 4(1+8T)(3+4T) & \text{if } i = n+1, \end{cases}$$

$$e^{\hat{V}_i^1} = \begin{cases} 2\, c_i & \text{if } i = 1, \ldots, n \\ 1/2 & \text{if } i = n+1, \end{cases} \qquad e^{\hat{V}_i^2} = \begin{cases} 4(1+4T) c_i / 7T & \text{if } i = 1, \ldots, n \\ 1/7 & \text{if } i = n+1. \end{cases}$$

We set the probabilities of observing the two realizations of the mean utilities as $\alpha^1 = (1+4T)/(1+8T)$ and $\alpha^2 = 4T/(1+8T)$. Finally, we set the target revenue as $K = 4(1+4T)(1+2T) + 4T$. Now, we show that the Partition problem has a solution if and only if there exists an assortment $S \subseteq \{1, \ldots, , n, n+1\}$ whose expected revenue is at least K. Noting that the $r_{n+1} > \max_{i \in \{1,\ldots,n\}} r_i$, adding product $n+1$ to any assortment increases the expected revenue from the assortment. So, we have

$$\max_{S \subseteq \{1,\ldots,n,n+1\}} \left\{ \alpha^1 \frac{\sum_{i \in S} r_i\, e^{\hat{V}_i^1}}{1 + \sum_{i \in S} e^{\hat{V}_i^1}} + \alpha^2 \frac{\sum_{i \in S} r_i\, e^{\hat{V}_i^2}}{1 + \sum_{i \in S} e^{\hat{V}_i^2}} \right\}$$
$$= \max_{S \subseteq \{1,\ldots,n\}} \left\{ \alpha^1 \frac{r_{n+1}\, e^{\hat{V}_{n+1}^1} + \sum_{i \in S} r_i\, e^{\hat{V}_i^1}}{1 + v_{n+1}^1 + \sum_{i \in S} e^{\hat{V}_i^1}} + \alpha^2 \frac{r_{n+1}\, e^{\hat{V}_{n+1}^2} + \sum_{i \in S} r_i\, e^{\hat{V}_i^2}}{1 + v_{n+1}^2 + \sum_{i \in S} e^{\hat{V}_i^2}} \right\}.$$

[5]Marshall School of Business, University of Southern California, Los Angeles, CA 90089, USA, Email: rusmevic@marshall.usc.edu
[6]School of Operations Research and Information Engineering and Department of Computer Science, Cornell University, Ithaca, NY 14853, USA, Email: shmoys@cs.cornell.edu
[7]School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA, Email: ct423@cornell.edu
[8]School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA, Email: topaloglu@orie.cornell.edu

To facilitate our exposition, let $\Pi : 2^{\{1,\ldots,n\}} \to \mathbb{R}_+$ denote the objective function of the last optimization problem above. Using the definitions of $r_i, e^{\hat{V}_i^1}, e^{\hat{V}_i^2}, \alpha^1$ and α^2, we have

$$\begin{aligned}\Pi(S) = &\frac{1+4T}{1+8T}\left(\frac{2(1+8T)(3+4T)+2(1+8T)(3+4T)\sum_{i\in S}c_i}{1+1/2+2\sum_{i\in S}c_i}\right)\\ &\quad + \frac{4T}{1+8T}\left(\frac{4(1+8T)(3+4T)/7+4(1+8T)(3+4T)(1+4T)\sum_{i\in S}c_i/(7T)}{1+1/7+4(1+4T)\sum_{i\in S}c_i/(7T)}\right)\\ &= 4(1+4T)(3+4T)\left(\frac{1+\sum_{i\in S}c_i}{3+4\sum_{i\in S}c_i}\right) + 4(1+4T)(3+4T)\left(\frac{\frac{4T^2}{1+4T}+4T\sum_{i\in S}c_i}{8T+4(1+4T)\sum_{i\in S}c_i}\right)\end{aligned}$$

for any assortment S. Therefore, $\Pi(S)$ can be written as $\Pi(S) = 4(1+4T)(3+4T)F\big(\sum_{i\in S}c_i\big)$, where the function $F : \mathbb{R}_+ \to \mathbb{R}_+$ is defined for all $z \in \mathbb{R}_+$ by

$$F(z) = \frac{1+z}{3+4z} + \frac{\frac{4T^2}{1+4T}+4Tz}{8T+4(1+4T)z} = \frac{1+z}{3+4z} + \frac{\frac{T}{1+4T}+z}{2+\left(\frac{1+4T}{T}\right)z}.$$

It is straightforward to verify that the derivative of F is strictly positive over the interval $[0, T)$ and strictly negative over the interval (T, ∞). Therefore, F has a unique maximum at T, and thus, for all $z \in \mathbb{R}_+$, $F(z) \le F(T) = \left(1+2T+\frac{T}{1+4T}\right)/(3+4T)$. Hence, we obtain

$$\begin{aligned}\max_{S\subseteq\{1,\ldots,n,n+1\}}\left\{\alpha^1\frac{\sum_{i\in S}r_i\,e^{\hat{V}_i^1}}{1+\sum_{i\in S}e^{\hat{V}_i^1}} + \alpha^2\frac{\sum_{i\in S}r_i\,e^{\hat{V}_i^2}}{1+\sum_{i\in S}e^{\hat{V}_i^2}}\right\} &= \max_{S\subseteq\{1,\ldots,n\}}\Pi(S)\\ &= 4(1+4T)(3+4T)\max_{S\subseteq\{1,\ldots,n\}}F\big(\textstyle\sum_{i\in S}c_i\big) \le 4(1+4T)(3+4T)F(T)\\ &= 4(1+4T)\left(1+2T+\frac{T}{1+4T}\right) = 4(1+4T)(1+2T)+4T = K.\end{aligned}$$

So, there exists an assortment $S \subseteq \{1,\ldots,n,n+1\}$ whose expected profit is at least K if and only if the chain of inequalities above hold as equalities. For this to happen, however, we need to have $\sum_{i\in S'}c_i = T$ for some assortment $S' \subseteq \{1,\ldots,n\}$. Therefore, we reach the conclusion that there exists an assortment $S \subseteq \{1,\ldots,n,n+1\}$ whose expected profit is at least K if and only if there exists an assortment $S' \subseteq \{1,\ldots,n\}$ that satisfies $\sum_{i\in S'}c_i = T$. □

B Proof of Theorem 4.1

Proof. We assume that P has the density function h and B has uniform distribution over $[b_1, b_2]$ with $0 \le b_1 < b_2$. The same line of reasoning applies when P has a discrete distribution or when the distribution of B is degenerate with $b_1 = b_2$. By definition of $\pi_i(S, \boldsymbol{V})$,

$$\mathbb{E}\{\pi_i(S,\boldsymbol{V})\} = \frac{1}{b_2-b_1}\int_{\mathbb{R}} h(p)\int_{b_1}^{b_2}\frac{e^{\mu_i+p-b\,r_i}}{1+\sum_{j\in S}e^{\mu_j+p-b\,r_j}}\,db\,dp.$$

Using the expression above, we write the objective function of the MIXTURE LOGIT problem as

$$\begin{aligned}\mathbb{E}\{f(S,\boldsymbol{V})\} = \sum_{i\in S}r_i\,\mathbb{E}\{\pi_i(S,\boldsymbol{V})\} &= \frac{1}{b_2-b_1}\sum_{i\in S}r_i\int_{\mathbb{R}}h(p)\int_{b_1}^{b_2}\frac{e^{\mu_i+p-b\,r_i}}{1+\sum_{j\in S}e^{\mu_j+p-b\,r_j}}\,db\,dp\\ &= \frac{1}{b_2-b_1}\int_{\mathbb{R}}h(p)\int_{b_1}^{b_2}\frac{\sum_{i\in S}r_i\,e^{\mu_i+p-b\,r_i}}{1+\sum_{j\in S}e^{\mu_j+p-b\,r_j}}\,db\,dp.\end{aligned}$$

The second integral in the last expression above can be computed as

$$\int_{b_1}^{b_2} \frac{\sum_{i\in S} r_i\, e^{\mu_i+p-b\,r_i}}{1+\sum_{j\in S} e^{\mu_j+p-b\,r_j}}\, db = -\log\left(1+\sum_{i\in S} e^{\mu_i+p-b\,r_i}\right)\Bigg|_{b_1}^{b_2} = \log\left(\frac{e^{-p}+\sum_{i\in S} e^{\mu_i-b_1\,r_i}}{e^{-p}+\sum_{j\in S} e^{\mu_i-b_2\,r_i}}\right),$$

in which case, we obtain

$$\begin{aligned}\mathbb{E}\{f(S,\boldsymbol{V})\} &= \frac{1}{b_2-b_1}\int_{\mathbb{R}} h(p)\log\left(\frac{e^{-p}+\sum_{i\in S} e^{\mu_i-b_2\,r_i}e^{(b_2-b_1)\,r_i}}{e^{-p}+\sum_{i\in S} e^{\mu_i-b_2\,r_i}}\right) dp \\ &= \frac{1}{b_2-b_1}\int_0^{\infty} g(v_0)\log\left(\frac{v_0+\sum_{i\in S} e^{\mu_i-b_2\,r_i}e^{(b_2-b_1)\,r_i}}{v_0+\sum_{i\in S} e^{\mu_i-b_2\,r_i}}\right) dv_0\ ,\end{aligned}$$

where the last equality follows from the change of variable $v_0 = e^{-p}$ and $g(v_0) = \frac{h(-\log v_0)}{v_0}$. For $i = 1,\ldots,n$, let $w_i = e^{(b_2-b_1)r_i}$ and $v_i = e^{\mu_i-b_2r_i}$. Note that $w_1 \geq w_2 \geq \cdots \geq w_n > 1$ by our assumption on r_i, and $g(v_0) \geq 0$ for all v_0. Dropping the scaling constant $\frac{1}{b_2-b_1}$, the Mixture Logit problem then corresponds to the following optimization problem:

$$Z^* = \max_{S\subseteq\{1,\ldots,n\}} \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\sum_{i\in S} w_i v_i}{v_0+\sum_{i\in S} v_i}\right) dv_0 \quad\leq\quad \max_{\boldsymbol{y}\in[0,1]^n} \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\sum_{i=1}^n w_i v_i y_i}{v_0+\sum_{i=1}^n v_i y_i}\right) dv_0\ ,$$

where the inequality follows from the fact that the optimization problem on the right side is the continuous relaxation of the one on the left side. To complete the proof, we will show that revenue-ordered assortments are optimal for the continuous relaxation. To see this, note that

$$\begin{aligned}&\max_{\boldsymbol{y}\in[0,1]^n} \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\sum_{i=1}^n w_i v_i y_i}{v_0+\sum_{i=1}^n v_i y_i}\right) dv_0 \\ &= \max_{\epsilon\in[0,\sum_{i=1}^n v_i]}\ \max_{\boldsymbol{y}\in[0,1]^n:\sum_{i=1}^n v_i y_i=\epsilon}\ \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\sum_{i=1}^n w_i v_i y_i}{v_0+\sum_{i=1}^n v_i y_i}\right) dv_0 \\ &= \max_{\epsilon\in[0,\sum_{i=1}^n v_i]}\ \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\xi(\epsilon)}{v_0+\epsilon}\right) dv_0\ ,\end{aligned}$$

where $\xi(\epsilon) = \max\{\sum_{i=1}^n w_i v_i y_i \ :\ \boldsymbol{y}\in[0,1]^n,\ \ \sum_{i=1}^n v_i y_i = \epsilon\}$. Note that $\xi(\epsilon)$ is the optimal objective value of a continuous knapsack problem, and the optimal solution can be found by filling the items according to their profit-to-space ratio $w_i v_i / v_i$. Since $w_1 \geq w_2 \geq \cdots \geq w_n > 1$, it follows that $\xi(\epsilon)$ is a piecewise linear function with the following breakpoints: $\left\{\sum_{i=1}^{\ell} v_i : \ell = 1,2,\ldots,n\right\}$. Note that, for each ℓ, the breakpoint $\sum_{i=1}^{\ell} v_i$ corresponds to the total preference weight associated with the revenue-ordered assortment $\{1,\ldots,\ell\}$. The desired result then follows immediately from the following claim:

Claim: The maximum of the continuous relaxation occurs at one of the breakpoints $\left\{\sum_{i=1}^{\ell} v_i : \ell = 1,2,\ldots,n\right\}$; that is,

$$\max_{\epsilon\in[0,\sum_{i=1}^n v_i]} \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\xi(\epsilon)}{v_0+\epsilon}\right) dv_0 \;=\; \max_{\ell=1,\ldots,n} \int_{\mathbb{R}_+} g(v_0)\log\left(\frac{v_0+\xi\left(\sum_{i=1}^{\ell} v_i\right)}{v_0+\sum_{i=1}^{\ell} v_i}\right) dv_0.$$

For each $\epsilon \geq 0$, let $T(\epsilon) = \int_{\mathbb{R}_+} g(v_0) \log\left(\frac{v_0+\xi(\epsilon)}{v_0+\epsilon}\right) dv_0$, and let ϵ^* be a maximizer of T over the interval $[0, \sum_{i=1}^n v_i]$. To prove the claim, suppose on the contrary that ϵ^* occurs strictly between the breakpoints; that is, $\sum_{i=1}^{K-1} v_i < \epsilon^* < \sum_{i=1}^{K} v_i$ for some $K \geq 1$. The knapsack problem defining $\xi(\epsilon)$ can be solved by ordering the items with respect to their profit-to-space ratios. Since the profit-to-space ratio of item i in this knapsack problem is $w_i\, v_i/v_i = w_i$ and $w_1 \geq w_2 \geq \cdots \geq w_n$, it follows that, for $\sum_{i=1}^{K-1} v_i < \epsilon^* < \sum_{i=1}^{K} v_i$, we have

$$\xi(\epsilon^*) = \sum_{i=1}^{K-1} w_i v_i + w_K \left(\epsilon^* - \sum_{i=1}^{K-1} v_i\right) \qquad \text{and} \qquad \xi'(\epsilon^*) = w_K \ .$$

Let $M = \frac{\sum_{i=1}^{K-1}(w_i - w_K)v_i}{w_K - 1}$. Since ϵ^* occurs between breakpoints, we have $0 = T'(\epsilon^*)$. Note that

$$\begin{aligned} T'(\epsilon^*) &= \int_{\mathbb{R}_+} g(v_0) \frac{v_0+\epsilon^*}{v_0+\xi(\epsilon^*)} \left(\frac{(v_0+\epsilon^*)\xi'(\epsilon^*) - (v_0+\xi(\epsilon^*))}{(v_0+\epsilon^*)^2}\right) dv_0 \\ &= \int_{\mathbb{R}_+} g(v_0)\left(\frac{w_K}{v_0+\xi(\epsilon^*)} - \frac{1}{v_0+\epsilon^*}\right) dv_0 \ = \ (w_K - 1)\int_{\mathbb{R}_+} g(v_0) \frac{v_0 - M}{(v_0+\epsilon^*)(v_0+\xi(\epsilon^*))}\, dv_0 \ , \end{aligned}$$

where the last equality follows from the fact that

$$(v_0+\epsilon^*)w_K - (v_0+\xi(\epsilon^*)) = v_0(w_K - 1) - \sum_{i=1}^{K-1}(w_i - w_K)v_i = (w_K - 1)(v_0 - M) \ .$$

Since $w_K > 1$ and $T'(\epsilon^*) = 0$, it must be the case that $M > 0$. Otherwise, the last integral is strictly positive. Moreover, it is easy to verify that

$$T''(\epsilon^*) \ = \ \int_{\mathbb{R}_+} g(v_0)\left(\frac{-w_K^2}{(v_0+\xi(\epsilon^*))^2} + \frac{1}{(v_0+\epsilon^*)^2}\right) dv_0 = \int_{\mathbb{R}_+} g(v_0) \frac{(v_0+\xi(\epsilon^*))^2 - w_K^2(v_0+\epsilon^*)^2}{(v_0+\xi(\epsilon^*))^2(v_0+\epsilon^*)^2}\, dv_0.$$

Note that for any $x \in \mathbb{R}_+$ and $z \in \mathbb{R}_+$, $x^2 - z^2 = (x-z)(x+z) \geq 2(x-z)z$, which implies that

$$(v_0+\xi(\epsilon^*))^2 - w_K^2(v_0+\epsilon^*)^2 \geq 2w_K\left(v_0+\xi(\epsilon^*) - w_K(v_0+\epsilon^*)\right)(v_0+\epsilon^*) = -2w_K(w_K-1)(v_0-M)(v_0+\epsilon^*) \ ,$$

where the last equality follows from $(v_0+\epsilon^*)w_K - (v_0+\xi(\epsilon^*)) = (w_K-1)(v_0-M)$, which we have shown above. Putting everything together, we get

$$\begin{aligned} \frac{T''(\epsilon^*)}{w_K(w_K-1)} &\geq -2\int_{\mathbb{R}_+} g(v_0)\frac{(v_0-M)(v_0+\epsilon^*)}{(v_0+\xi(\epsilon^*))^2(v_0+\epsilon^*)^2}\, dv_0 = -2\int_{\mathbb{R}_+} g(v_0)\frac{(v_0-M)}{(v_0+\xi(\epsilon^*))^2(v_0+\epsilon^*)}\, dv_0 \\ &= -2\int_0^M g(v_0)\frac{(v_0-M)}{(v_0+\xi(\epsilon^*))^2(v_0+\epsilon^*)}\, dv_0 - 2\int_M^\infty g(v_0)\frac{(v_0-M)}{(v_0+\xi(\epsilon^*))^2(v_0+\epsilon^*)}\, dv_0 \\ &> -2\int_0^M g(v_0)\frac{(v_0-M)}{(M+\xi(\epsilon^*))(v_0+\xi(\epsilon^*))(v_0+\epsilon^*)}\, dv_0 - 2\int_M^\infty g(v_0)\frac{(v_0-M)}{(M+\xi(\epsilon^*))(v_0+\xi(\epsilon^*))(v_0+\epsilon^*)}\, dv_0 \\ &= -\frac{2}{M+\xi(\epsilon^*)}\int_{\mathbb{R}_+} g(v_0)\frac{(v_0-M)}{(v_0+\xi(\epsilon^*))(v_0+\epsilon^*)}\, dv_0 \ = \ -\frac{2\,T'(\epsilon^*)}{(M+\xi(\epsilon^*))(w_K-1)} \ = \ 0 \ , \end{aligned}$$

where the last strict inequality follows from the fact that $M > 0$. This means that T is strictly convex in a small neighborhood of ϵ^*, while the derivative of T vanishes at ϵ^*. This contradicts the fact that ϵ^* is a maximizer of T! Therefore, the maximum must occur at one of the breakpoints. $\square$

C Proof of Theorem 5.4

The proof of Theorem 5.4 makes use of the following lemma.

Lemma C.1 *If product i is included in an optimal assortment for the problem* $\max_{S\subseteq\{1,\ldots,n\}} f(S,\boldsymbol{V})$*, then there also exists an optimal assortment for this problem that includes any other product that has the same revenue as product i.*

Proof. Assume that S^* is an optimal solution to the problem $\max_{S\subseteq\{1,\ldots,n\}} f(S,\boldsymbol{V})$ with $i\in S^*$, $j\notin S^*$ and $r_i=r_j$. Note that $f(S^*,V)$ is a convex combination of $f(S^*\setminus\{i\},\boldsymbol{V})$ and r_i since

$$f(S^*,\boldsymbol{V})=\left(\frac{1+\sum_{\ell\in S^*\setminus\{i\}}e^{V_\ell}}{1+\sum_{\ell\in S^*}e^{V_\ell}+e^{V_i}}\right)f(S^*\setminus\{i\},\boldsymbol{V})+\left(\frac{e^{V_i}}{1+\sum_{\ell\in S^*}e^{V_\ell}+e^{V_i}}\right)r_i.$$

Noting that $f(S^*,\boldsymbol{V})\geq f(S^*\setminus\{i\},\boldsymbol{V})$ by the definition of S^* and $f(S^*,\boldsymbol{V})$ is a convex combination of $f(S^*\setminus\{i\},\boldsymbol{V})$ and r_i, we obtain $f(S^*,\boldsymbol{V})\leq r_i=r_j$. An identity similar to the one above shows that $f(S^*\cup\{j\},\boldsymbol{V})$ is also a convex combination of $f(S^*,\boldsymbol{V})$ and r_j, in which case, noting that $f(S^*,\boldsymbol{V})\leq r_j$, we obtain $f(S^*\cup\{j\},\boldsymbol{V})\geq f(S^*,\boldsymbol{V})$. Therefore we can add product j to the assortment S^* without degrading the objective value for the problem $\max_{S\subseteq\{1,\ldots,n\}} f(S,\boldsymbol{V})$. □

Here is the proof of Theorem 5.4.

Proof. We can scale the revenues of the products without changing the optimal solution to the Mixture Logit problem. Thus, we assume without loss of generality that $r_n=1$. In this case, we need to show an approximation guarantee of $e\log(e\,r_1)$. For any $\delta>0$, we let $\mathsf{Dom}_\delta=\{(1+\delta)^k : k=0,1,2,\ldots\}$ and define $\mathsf{Floor}_\delta:\mathbb{R}_+\to\mathsf{Dom}_\delta$ as $\mathsf{Floor}_\delta(x)=\max\{y\in\mathsf{Dom}_\delta : y\leq x\}$ for each $x\in\mathbb{R}_+$. Therefore, Floor_δ rounds its argument down to the closest element in Dom_δ. We let $f_\delta(S,\boldsymbol{V})$ be an approximation to $f(S,\boldsymbol{V})$ defined by

$$f_\delta(S,\boldsymbol{V})=\sum_{i\in S}\mathsf{Floor}_\delta(r_i)\,\pi_i(S,\boldsymbol{V}).$$

It is easy to verify that $f_\delta(S,\boldsymbol{V})\leq f(S,\boldsymbol{V})\leq(1+\delta)f_\delta(S,\boldsymbol{V})$ for all $S\subseteq\{1,\ldots,n\}$.

Let $K_\delta=\lfloor(\log r_1)/\log(1+\delta)\rfloor$ so that $(1+\delta)^{K_\delta}\leq r_1<(1+\delta)^{K_\delta+1}$. For $k=0,1,\ldots,K_\delta$, we define the set of products N_δ^k as $N_\delta^k=\{i=1,\ldots,n : \mathsf{Floor}_\delta(r_i)=(1+\delta)^k\}$. Therefore, the sets of products $N_\delta^1,\ldots,N_\delta^{K_\delta}$ partition the set of products $\{1,\ldots,n\}$. Furthermore, each N_δ^k includes a consecutive set of products. Finally, if $i,j\in N_\delta^k$ for some $k=0,1,\ldots,K_\delta$, then these products have the same revenue in the assortment optimization problem $\max_{S\subseteq\{1,\ldots,n\}} f_\delta(S,\boldsymbol{V})$. In this case, it holds that

$$\max_{S\subseteq\{1,\ldots,n\}} f_\delta(S,\boldsymbol{V})=\max_{i=1,\ldots,n} f_\delta(\{1,\ldots,i\},\boldsymbol{V})=\max_{k=0,\ldots,K_\delta} f_\delta(\cup_{\ell=0}^k N_\delta^\ell,\boldsymbol{V}).$$

The first equality above follows from the fact that if the vector of mean utilities are known, then revenue-ordered assortments are optimal. To see that the second equality holds, assume on the contrary that the optimal solution to the second problem above is $\{1,\ldots,i^*\}$ and it cannot be written as $\cup_{\ell=0}^k N_\delta^\ell$

for some $k = 0, \ldots, K_\delta$. We let k^* be such that $\cup_{\ell=0}^{k^*} N_\delta^\ell \subseteq \{1, \ldots, i^*\} \subseteq \cup_{\ell=0}^{k^*+1} N_\delta^\ell$. Therefore, all of the products in the set $\cup_{\ell=0}^{k^*+1} N_\delta^\ell \setminus \{1, \ldots, i^*\}$ have the same revenue in the assortment optimization problem $\max_{S \subseteq \{1,\ldots,n\}} f_\delta(S, \boldsymbol{V})$, in which case, Lemma C.1 implies that these products can be added to the assortment $\{1, \ldots, i^*\}$ and the assortment $\cup_{\ell=0}^{k^*+1} N_\delta^\ell$ is also an optimal solution to the second problem above. This establishes the second equality. To complete the proof, recall that $f_\delta(S, \boldsymbol{V}) \le f(S, \boldsymbol{V}) \le (1+\delta) f_\delta(S, \boldsymbol{V})$. Therefore, we have

$$\begin{aligned}
Z^* &= \max_{S \subseteq \{1,\ldots,n\}} \mathbb{E}\{f(S, \boldsymbol{V})\} \le (1+\delta) \max_{S \subseteq \{1,\ldots,n\}} \mathbb{E}\{f_\delta(S, \boldsymbol{V})\} \\
&\le (1+\delta)\, \mathbb{E}\left\{ \max_{S \subseteq \{1,\ldots,n\}} f_\delta(S, \boldsymbol{V}) \right\} = (1+\delta)\, \mathbb{E}\left\{ \max_{k=0,\ldots,K_\delta} f_\delta(\cup_{\ell=0}^{k} N_\delta^\ell, \boldsymbol{V}) \right\} \\
&\le (1+\delta)\, \mathbb{E}\left\{ \sum_{k=0}^{K_\delta} f_\delta(N_\delta^k, \boldsymbol{V}) \right\} = (1+\delta) \sum_{k=0}^{K_\delta} \mathbb{E}\{f_\delta(N_\delta^k, \boldsymbol{V})\},
\end{aligned}$$

where the third inequality follows by noting that $f_\delta(\cup_{\ell=0}^{k} N_\delta^\ell, \boldsymbol{V}) \le \sum_{\ell=0}^{k} f_\delta(N_\delta^\ell, \boldsymbol{V})$ by the first part of Lemma 5.2. To continue the chain of inequalities, we note that the set N_δ^ℓ includes consecutive products so that it is of the form $\{i_\delta^\ell, \ldots, j_\delta^\ell\}$. In this case, the second part of Lemma 5.2 implies that $f_\delta(N_\delta^\ell, \boldsymbol{V}) = f_\delta(\{i_\delta^\ell, \ldots, j_\delta^\ell\}, \boldsymbol{V}) \le f_\delta(\{1, \ldots, j_\delta^\ell\}, \boldsymbol{V})$. We continue to chain of inequalities above as

$$\begin{aligned}
(1+\delta) \sum_{k=0}^{K_\delta} \mathbb{E}\{f_\delta(N_\delta^k, \boldsymbol{V})\} &\le (1+\delta) \sum_{k=0}^{K_\delta} \mathbb{E}\{f_\delta(\{1, \ldots, j_\delta^k\}, \boldsymbol{V})\} \\
&\le (1+\delta) \sum_{k=0}^{K_\delta} \max_{i=1,\ldots,n} \mathbb{E}\{f_\delta(\{1, \ldots, i\}, \boldsymbol{V})\} = (1+\delta)\,(1+K_\delta) \max_{i=1,\ldots,n} \mathbb{E}\{f_\delta(\{1, \ldots, i\}, \boldsymbol{V})\} \\
&\le (1+\delta)\,(1+K_\delta) \max_{i=1,\ldots,n} \mathbb{E}\{f(\{1, \ldots, i\}, \boldsymbol{V})\} \\
&\le (1+\delta) \left(1 + \frac{\log r_1}{\log(1+\delta)}\right) \max_{i=1,\ldots,n} \mathbb{E}\{f(\{1, \ldots, i\}, \boldsymbol{V})\},
\end{aligned}$$

where the last inequality follows from the fact that $K_\delta \le (\log r_1)/\log(1+\delta)$. The result follows by choosing $\delta = e - 1$ in the last expression. □

D Proof of Proposition 5.5

Proof. To establish the first part of the proposition, it follows from our hypothesis that

$$\begin{aligned}
Z^* &= \max_{S \subseteq \{1,\ldots,n\}} \mathbb{E}\left\{ \frac{\sum_{i \in S} r_i e^{V_i}}{1 + \sum_{i \in S} e^{V_i}} \right\} \le \max_{S \subseteq \{1,\ldots,n\}} \frac{\sum_{i \in S} r_i e^{\mathbb{E}V_i + \delta}}{1 + \sum_{i \in S} e^{\mathbb{E}V_i - \delta}} = e^{2\delta} \max_{S \subseteq \{1,\ldots,n\}} \frac{\sum_{i \in S} r_i e^{\mathbb{E}V_i}}{e^\delta + \sum_{i \in S} e^{\mathbb{E}V_i}} \\
&= e^{2\delta} \max_{\ell=1,\ldots,n} \frac{\sum_{i=1}^{\ell} r_i e^{\mathbb{E}V_i}}{e^\delta + \sum_{i=1}^{\ell} e^{\mathbb{E}V_i}} \le e^{2\delta} \max_{\ell=1,\ldots,n} \frac{\sum_{i=1}^{\ell} r_i e^{\mathbb{E}V_i}}{1 + \sum_{i=1}^{\ell} e^{\mathbb{E}V_i}} \le e^{2\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i=1}^{\ell} r_i e^{V_i + \delta}}{1 + \sum_{i=1}^{\ell} e^{V_i - \delta}} \right\} \\
&= e^{4\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i=1}^{\ell} r_i e^{V_i}}{e^\delta + \sum_{i=1}^{\ell} e^{V_i}} \right\} \le e^{4\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i=1}^{\ell} r_i e^{V_i}}{1 + \sum_{i=1}^{\ell} e^{V_i}} \right\},
\end{aligned}$$

where the third equality above follows from the standard result that when the mean utility is deterministic and known, revenue-ordered assortments are optimal. This gives the desired result. For the second part of the proposition, consider an arbitrary product $\phi \in \{1, \ldots, n\}$. Since $\max_{i=1,\ldots,n} |V_i - V_\phi| \leq \delta$ almost surely,

$$
\begin{aligned}
Z^* &= \max_{S \subseteq \{1,\ldots,n\}} \mathbb{E}\left\{ \frac{\sum_{i \in S} r_i e^{V_i}}{1 + \sum_{i \in S} e^{V_i}} \right\} \leq \max_{S \subseteq \{1,\ldots,n\}} \mathbb{E}\left\{ \frac{\sum_{i \in S} r_i e^{V_\phi + \delta}}{1 + \sum_{i \in S} e^{V_\phi - \delta}} \right\} = e^{2\delta} \max_{S \subseteq \{1,\ldots,n\}} \mathbb{E}\left\{ \frac{\sum_{i \in S} r_i e^{V_\phi}}{e^{\delta} + \sum_{i \in S} e^{V_\phi}} \right\} \\
&= e^{2\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i=1}^{\ell} r_i e^{V_\phi}}{e^{\delta} + \sum_{i=1}^{\ell} e^{V_\phi}} \right\} \leq e^{2\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i \in S} r_i e^{V_i + \delta}}{1 + \sum_{i \in S} e^{V_i - \delta}} \right\} \leq e^{2\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i \in S} r_i e^{V_i + \delta}}{1 + \sum_{i \in S} e^{V_i - \delta}} \right\} \\
&= e^{4\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i=1}^{\ell} r_i e^{V_i}}{e^{\delta} + \sum_{i=1}^{\ell} e^{V_i}} \right\} \leq e^{4\delta} \max_{\ell=1,\ldots,n} \mathbb{E}\left\{ \frac{\sum_{i=1}^{\ell} r_i e^{V_i}}{1 + \sum_{i=1}^{\ell} e^{V_i}} \right\} ,
\end{aligned}
$$

where the third equality follows from the fact that the third maximization problem above satisfies the Value Conscious condition because, in this problem, the mean utility of every product is equal to V_ϕ, so we have that $r_1 e^{V_\phi} \geq \cdots \geq r_n e^{V_\phi}$, in which case, the equality follows from Theorem 4.3. This gives the desired result. □