Technical Note: Robust Logit Assortments

Paat Rusmevichientong* Huseyin Topaloglu†

July 7, 2009

Abstract

In this paper, we analyze a robust assortment optimization problem under the multinomial logit choice model. The novel aspect of our analysis is that the model parameter is assumed to be unknown and we use an uncertainty set to capture the possible values for the unknown model parameter. The goal is to choose an assortment that maximizes the worst-case revenue over all parameter values in the uncertainty set. We show that the robust assortment can be found efficiently and it includes more products as the uncertainty set grows. Therefore, larger product variety serves as a safety buffer against larger parameter uncertainty. Numerical experiments indicate that the robust assortments are especially beneficial when the uncertainty in the model parameter is high.

1. Introduction and Problem Formulation

Many firms face the problem of choosing an assortment of substitutable products to offer to their customers, who in turn, observe the assortment that is available for purchase and possibly make a purchase within the assortment. The goal of the firm is to maximize the expected revenue generated from each customer. In this paper, we focus on the setting where the choices of the customers are governed by the multinomial logit choice model, which is one of the most commonly used choice models in economics, marketing and operations management; see Ben-Akiva and Lerman (1985), Anderson et al. (1992), van Ryzin and Mahajan (1999) and the references therein.

Assortment optimization is an active area of research with many applications and the reader is referred to Kok et al. (2008) for an excellent review of the literature in this area. Here, we primarily focus our attention to papers that use the multinomial logit choice model to describe the customer choice process. The paper by van Ryzin and Mahajan (1999) uses the multinomial logit choice model to study a classical retail problem that involves balancing the revenue benefit of offering product variety with the inventory cost of carrying a large number of products. This paper has been extended by Mahajan and van Ryzin (2001) to include the possibility that one product can substitute for another in case of a stock out. Cachon et al. (2005) also build on the work of van Ryzin and Mahajan (1999), and they incorporate the search costs by considering the possibility that a customer may find an acceptable product in one store, but still purchase nothing hoping that another store may carry a more desirable product. In addition to applications in retail, the multinomial logit choice model has applications in revenue management. Talluri and van Ryzin (2004) consider a revenue management problem over a single flight leg, where customers choose among the fare classes that are available for purchase. The

*School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA. E-mail: paatrus@cornell.edu

†School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA. E-mail: topaloglu@orie.cornell.edu

goal is to dynamically adjust the assortment of available fare classes to maximize the total expected revenue generated until the departure time of the flight leg. Gallego et al. (2004), Zhang and Adelman (2006), Liu and van Ryzin (2008) and Kunnumkal and Topaloglu (2008) extend this model to multiple flight legs and itineraries with connections.

All of the work thus far assumes that the parameter of the multinomial logit choice model is known in advance or has been estimated from data. In many cases, however, we are uncertain about the true value of the model parameter and we might want to incorporate this uncertainty directly into our problem formulation. In this paper, we analyze a robust assortment optimization problem that incorporates the uncertainty in the model parameter through an uncertainty set. The goal is to choose an assortment that maximizes the worst-case revenue over all parameter values in the uncertainty set. We show that the robust assortment can be found efficiently, establish its structural properties and demonstrate its benefits through extensive numerical experiments.

Before we proceed to our main results, let us briefly describe the problem setup. We have a set of products indexed by $1, 2, \ldots, n$, and for each i, $w_i > 0$ denotes the revenue of product i. Without loss of generality, we assume that the products are ordered such that $w_1 \geq w_2 \geq \ldots \geq w_n$. Each customer chooses a product from an assortment according to a multinomial logit choice model with an *unknown* parameter $\boldsymbol{v} = (v_1, \ldots, v_n) \in \Re^n_+$. Following the terminology of Vulcano et al. (2008), we interchangeably refer to $\boldsymbol{v}$ as the model parameter or the customer preference weights. If we offer the assortment $S \subseteq \{1, \ldots, n\}$, then the customer purchases product i with probability

$$\phi_i(S, \boldsymbol{v}) = \begin{cases} \dfrac{v_i}{1 + \sum_{\ell \in A} v_\ell} & \text{if } i \in S, \\ 0 & \text{otherwise,} \end{cases}$$

and with the remaining probability $1 - \sum_{i \in S} \theta_i(S, \boldsymbol{v}) = 1/[1 + \sum_{i \in S} v_i]$, the customer does not purchase any product. The expected revenue per customer under the assortment S is given by

$$f(S, \boldsymbol{v}) = \sum_{i \in S} w_i \, \phi_i(S, \boldsymbol{v}) = \frac{\sum_{i \in S} w_i \, v_i}{1 + \sum_{i \in S} v_i} .$$

We model the uncertainty in the model parameter $\boldsymbol{v}$ through a compact uncertainty set $\mathcal{V} \subseteq \Re^n_+$. In our robust formulation, we are interested in choosing an assortment that maximizes the worst-case expected revenue over all model parameters in $\mathcal{V}$, corresponding to the optimization problem

$$\text{(Robust Logit)} \qquad Z^*(\mathcal{V}) \equiv \max_{S \subseteq \{1, \ldots, n\}} \left\{ \min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v}) \right\} .$$

We use $S^*(\mathcal{V})$ to denote an optimal assortment under Robust Logit problem. If there are ties, then we choose $S^*(\mathcal{V})$ to be an optimal assortment with the largest cardinality. Note that $Z^*(\mathcal{V})$ is well defined because $\mathcal{V}$ is compact.

The uncertainty set $\mathcal{V}$ represents the set of likely values of the model parameter and the size of $\mathcal{V}$ reflects our degree of uncertainty. If $\mathcal{V} = \{\boldsymbol{v}\}$ is a singleton, then we know the underlying model parameter exactly and our robust formulation reduces to the classical assortment optimization problem under multinomial logit choice model. To facilitate our exposition, we use $S^*_{\boldsymbol{v}}$ to denote $S^*(\{\boldsymbol{v}\})$, which is an optimal assortment when the model parameter $\boldsymbol{v}$ is known in advance. In other words, $S^*_{\boldsymbol{v}}$ is an optimal solution to the problem $\max_{S \subseteq \{1,\ldots,n\}} f(S, \boldsymbol{v})$ and we break ties by choosing $S^*_{\boldsymbol{v}}$ as an optimal assortment with the largest cardinality. Gallego et al. (2004), Liu and van Ryzin (2008), and Kunnumkal and Topaloglu (2008) characterize the assortment $S^*_{\boldsymbol{v}}$. In particular, they show that $S^*_{\boldsymbol{v}}$ consists of products with the highest revenues; that is, $S^*_{\boldsymbol{v}} = \{1, 2, \ldots, i^*_{\boldsymbol{v}}\}$ for some product $i^*_{\boldsymbol{v}}$. This is a powerful result since it allows us to consider only assortments of the form $\{1, \ldots, i\}$, instead of searching through all of 2^n possible assortments.

Surprisingly, even when we have uncertainty in the model parameter, the robust assortment $S^*(\mathcal{V})$ has the same structural property. This is the main result of our paper and it is stated in the following theorem, whose proof is given in the next section.

Theorem 1.1 (Robust Assortments)**.** *For any $\mathcal{V} \subset \Re^n_+$, the assortment $S^*(\mathcal{V})$ is the largest assortment among $\{S^*_{\boldsymbol{v}} : \boldsymbol{v} \in \mathcal{V}\}$; that is,* $S^*(\mathcal{V}) = \cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}}$.

Theorem 1.1 offers both computational and operational insights. The result shows that it suffices to consider assortments of the form $\{1, 2, \ldots, i\}$. This greatly simplifies the computation of the robust assortment since we only need to consider at most n possible assortments. Moreover, the theorem tells us that to protect against the uncertainty in the preference weights, we should offer the largest possible assortment among $\{S^*_{\boldsymbol{v}} : \boldsymbol{v} \in \mathcal{V}\}$. Thus, a large product variety serves as a safety buffer against uncertainty in the model parameter. In fact, as our uncertainty increases, the robust assortment gets larger, while of course, the worst-case expected revenue decreases. This result is stated in the following corollary whose proof follows immediately from Theorem 1.1.

Corollary 1.2 (Larger Uncertainty Implies Larger Robust Assortment)**.** *For any $\mathcal{V} \subseteq \mathcal{V}' \subseteq \Re^n_+$,*

$$Z^*(\mathcal{V}') \leq Z^*(\mathcal{V}) \quad \textit{and} \quad S^*(\mathcal{V}) \subseteq S^*(\mathcal{V}') \ .$$

A common criticism for a robust formulation is that such a model leads to an overly conservative solution. An alternative formulation, which we refer to as Mixture Logit problem, is to put a prior distribution $\theta : \mathcal{V} \to \Re_+$ over the set of likely parameter values, where $\sum_{\boldsymbol{v} \in \mathcal{V}} \theta(\boldsymbol{v}) = 1$. We can then consider the problem of choosing an assortment that maximizes the average expected revenue, where the average is taken over the set of likely parameter values. This yields the optimization problem

$$(\text{Mixture Logit}) \qquad \max_{S \subseteq \{1,\ldots,n\}} \sum_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v})\, \theta(\boldsymbol{v}) \ .$$

Mixture Logit problem has the advantage of producing a solution that is less conservative since it uses the probability that each model parameter value occurs based on the prior distribution. Unfortunately, Bront et al. (2009) show that Mixture Logit problem is NP-complete. Therefore, it is unlikely that we can find an efficient method for solving this problem. In contrast, as shown in Theorem 1.1, Robust Logit problem admits an efficient solution. Our numerical experiments in Section 3 will show that the assortments produced by Robust Logit problem are especially beneficial when the uncertainty in the model parameter is high and we do not have good estimates for the prior distribution.

2. Proof of Theorem 1.1

The proof of Theorem 1.1 makes use of the following lemmas. The first lemma characterizes the ordering between the convex combination of two numbers and its constituents. The proof follows from simple algebra and we omit the details.

Lemma 2.1. *For all $x \in \Re_+$, $y \in \Re_+$, and $\gamma \in (0,1)$, we have $\gamma x + (1-\gamma) y \geq y$ if and only if $x \geq y$, and $\gamma x + (1-\gamma) y \leq y$ if and only if $x \leq y$.*

The next lemma establishes conditions under which adding products to an assortment increases the expected revenue.

Lemma 2.2 (Positive Increments)**.** *For all $\boldsymbol{v} \in \Re^n_+$, $A \subseteq \{1,\ldots,n\}$, $B \subseteq \{1,\ldots,n\}$, and $A \cap B = \varnothing$, the following statements are equivalent:*
(a) $\sum_{i\in A} w_i\, v_i / \sum_{i\in A} v_i \geq f(B,\boldsymbol{v})$,
(b) $f(A \cup B, \boldsymbol{v}) \geq f(B,\boldsymbol{v})$, *and*
(c) $\sum_{i\in A} w_i\, v_i / \sum_{i\in A} v_i \geq f(A \cup B,\boldsymbol{v})$.

Proof. Note that $f(A\cup B,\boldsymbol{v})$ is a convex combination of $\sum_{i\in A} w_i\, v_i / \sum_{i\in A} v_i$ and $f(B,\boldsymbol{v})$ because

$$f(A\cup B,\boldsymbol{v}) = \left(\frac{\sum_{i\in A} v_i}{1+\sum_{i\in A} v_i + \sum_{i\in B} v_i} \cdot \frac{\sum_{i\in A} w_i\, v_i}{\sum_{i\in A} v_i} \right) + \left(\frac{1+\sum_{i\in B} v_i}{1+\sum_{i\in A} v_i + \sum_{i\in B} v_i} \cdot f(B,\boldsymbol{v}) \right)$$

and the equivalence between (a) and (b) follows immediately from the first part of Lemma 2.1 with $x = \sum_{i\in A} w_i\, v_i / \sum_{i\in A} v_i$ and $y = f(B,\boldsymbol{v})$. The equivalence between (a) and (c) follows from a similar argument using the second part of Lemma 2.1 with $x = f(B,\boldsymbol{v})$ and $y = \sum_{i\in A} w_i\, v_i / \sum_{i\in A} v_i$. □

Here is the proof of Theorem 1.1.

Proof. We show the result by induction on the number of products. When we have only one product, we trivially have $S^*(\mathcal{V}) = \{1\} = \cup_{\boldsymbol{v}\in\mathcal{V}} S^*_{\boldsymbol{v}}$ and the result holds. Assume that the result holds when we have $n-1$ products and consider the case with n products where the revenues of the products satisfy $w_1 \geq \ldots \geq w_n > 0$. We separately consider the cases $Z^*(\mathcal{V}) \leq w_n$ and $Z^*(\mathcal{V}) > w_n$.

Case 1. Suppose that $Z^*(\mathcal{V}) \leq w_n$. First, we show that $S^*(\mathcal{V}) = \{1,\ldots,n\}$. Suppose on the contrary that there exists $i \in \{1,\ldots,n\}$ such that $i \notin S^*(\mathcal{V})$. We partition $\mathcal{V}$ into subsets $\mathcal{V}_0 = \{\boldsymbol{v} \in \mathcal{V} : w_i < f(S^*(\mathcal{V}),\boldsymbol{v})\}$ and $\mathcal{V}_1 = \{\boldsymbol{v} \in \mathcal{V} : w_i \geq f(S^*(\mathcal{V}),\boldsymbol{v})\}$ so that $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$ and $\mathcal{V}_0 \cap \mathcal{V}_1 = \varnothing$. For all $\boldsymbol{v}_0 \in \mathcal{V}_0$, we have $w_i < f(S^*(\mathcal{V}),\boldsymbol{v}_0)$, in which case, the contrapositives of statements (a) and (c) in Lemma 2.2 with $A = \{i\}$ and $B = S^*(\mathcal{V})$ imply that

$$f(S^*(\mathcal{V}) \cup \{i\}, \boldsymbol{v}_0) \;>\; w_i \;\geq\; w_{i+1} \;\geq\; \ldots \;\geq\; w_n \;\geq\; Z^*(\mathcal{V}) \;.$$

On the other hand, for all $\boldsymbol{v}_1 \in \mathcal{V}_1$, we have $w_i \geq f(S^*(\mathcal{V}),\boldsymbol{v}_1)$, in which case, statements (a) and (b) in Lemma 2.2 with $A = \{i\}$ and $B = S^*(\mathcal{V})$ imply that

$$f(S^*(\mathcal{V}) \cup \{i\}, \boldsymbol{v}_1) \;\geq\; f(S^*(\mathcal{V}),\boldsymbol{v}_1) \;\geq\; \min_{\boldsymbol{v}\in\mathcal{V}} f(S^*(\mathcal{V}),\boldsymbol{v}) \;=\; Z^*(\mathcal{V}) \;.$$

Putting it all together, we obtain $\min_{\boldsymbol{v}\in\mathcal{V}} f(S^*(\mathcal{V})\cup\{i\},\boldsymbol{v}) = \min_{\boldsymbol{v}\in\mathcal{V}_0\cup\mathcal{V}_1} f(S^*(\mathcal{V})\cup\{i\},\boldsymbol{v}) \geq Z^*(\mathcal{V})$ so that $S^*(\mathcal{V})\cup\{i\}$ is an optimal assortment for Robust Logit problem. This contradicts the fact that $S^*(\mathcal{V})$ is an optimal assortment with the largest cardinality! So, we must have $S^*(\mathcal{V}) = \{1,\ldots,n\}$.

Second, we show that $\cup_{\boldsymbol{v}\in\mathcal{V}} S^*_{\boldsymbol{v}} = \{1,\ldots,n\}$. Suppose on the contrary that we have $\cup_{\boldsymbol{v}\in\mathcal{V}} S^*_{\boldsymbol{v}} \subset \{1,\ldots,n\}$. As we mention in Section 1, for all $\boldsymbol{v}\in\mathcal{V}$, $S^*_{\boldsymbol{v}}$ is an assortment of the form $\{1,\ldots,i^*_{\boldsymbol{v}}\}$ for some product $i^*_{\boldsymbol{v}}$. Therefore, having $\cup_{\boldsymbol{v}\in\mathcal{V}} S^*_{\boldsymbol{v}} \subset \{1,\ldots,n\}$ implies that $S^*_{\boldsymbol{v}} \subseteq \{1,\ldots,n-1\}$ for all $\boldsymbol{v}\in\mathcal{V}$. Furthermore, since we have $S^*_{\boldsymbol{v}} \subseteq \{1,\ldots,n-1\}$ for all $\boldsymbol{v}\in\mathcal{V}$, the definition of $S^*_{\boldsymbol{v}}$ implies that $f(\{1,\ldots,n\},\boldsymbol{v}) < f(S^*_{\boldsymbol{v}},\boldsymbol{v})$ for all $\boldsymbol{v}\in\mathcal{V}$. In this case, the contrapositives of statements (b) and (c) in Lemma 2.2 with $A = \{1,\ldots,n\}\setminus S^*_{\boldsymbol{v}}$ and $B = S^*_{\boldsymbol{v}}$ yield

$$f(\{1,\ldots,n\},\boldsymbol{v}) > \frac{\sum_{i\in\{1,\ldots,n\}\setminus S^*_{\boldsymbol{v}}} w_i\, v_i}{\sum_{i\in\{1,\ldots,n\}\setminus S^*_{\boldsymbol{v}}} v_i} \geq w_n$$

for all $\boldsymbol{v}\in\mathcal{V}$, where the second inequality follows from the fact that $w_1 \geq \ldots \geq w_n$. Using the inequality above, the contrapositives of statements (b) and (c) in Lemma 2.2 with $A = \{n\}$ and $B = \{1,\ldots,n-1\}$ tell us that $f(\{1,\ldots,n-1\},\boldsymbol{v}) > f(\{1,\ldots,n\},\boldsymbol{v}) = f(S^*(\mathcal{V}),\boldsymbol{v})$ for all $\boldsymbol{v}\in\mathcal{V}$, where the equality follows from the fact that $S^*(\mathcal{V}) = \{1,\ldots,n\}$, which we have shown in the paragraph above. Therefore, we have $\min_{\boldsymbol{v}\in\mathcal{V}} f(\{1,\ldots,n-1\},\boldsymbol{v}) > \min_{\boldsymbol{v}\in\mathcal{V}} f(S^*(\mathcal{V}),\boldsymbol{v})$ and the assortment $\{1,\ldots,n-1\}$ provides a strictly larger objective function value for Robust Logit problem than does $S^*(\mathcal{V})$. This contradicts the optimality of $S^*(\mathcal{V})$! So, we must have $\cup_{\boldsymbol{v}\in\mathcal{V}} S^*_{\boldsymbol{v}} = \{1,\ldots,n\}$.

Collecting the two results that we have shown under Case 1, we have $S^*(\mathcal{V}) = \{1,\ldots,n\} = \cup_{\boldsymbol{v}\in\mathcal{V}} S^*_{\boldsymbol{v}}$ and the desired result holds under Case 1.

Case 2. Suppose that $Z^*(\mathcal{V}) > w_n$. First, we show that $n \notin S^*(\mathcal{V})$. Suppose on the contrary that we have $n \in S^*(\mathcal{V})$. Since $Z^*(\mathcal{V}) = \min_{\boldsymbol{v}\in\mathcal{V}} f\,(S^*(\mathcal{V}),\boldsymbol{v})$, we have $w_n < f(S^*(\mathcal{V}),\boldsymbol{v})$ for all $\boldsymbol{v}\in\mathcal{V}$. In this case, the contrapositives of the statements (b) and (c) in Lemma 2.2 with $A = \{n\}$ and $B = S^*(\mathcal{V})\setminus\{n\}$ imply that $f(S^*(\mathcal{V}),\boldsymbol{v}) < f(S^*(\mathcal{V})\setminus\{n\},\boldsymbol{v})$ for all $\boldsymbol{v}\in\mathcal{V}$. Therefore, we obtain

$$\min_{\boldsymbol{v}\in\mathcal{V}} f(S^*(\mathcal{V}),\boldsymbol{v}) < \min_{\boldsymbol{v}\in\mathcal{V}} f(S^*(\mathcal{V})\setminus\{n\},\boldsymbol{v})$$

so that the assortment $S^*(\mathcal{V})\setminus\{n\}$ provides a strictly larger objective function value for Robust Logit problem than does $S^*(\mathcal{V})$. This contradicts the fact that $S^*(\mathcal{V})$ is an optimal assortment for Robust Logit problem! So, we must have $n \notin S^*(\mathcal{V})$.

Second, we show that $n \notin S^*_{\boldsymbol{v}}$ for all $\boldsymbol{v}\in\mathcal{V}$. Suppose on the contrary that we have $n \in S^*_{\boldsymbol{v}}$ for some $\boldsymbol{v}\in\mathcal{V}$. Since $w_n < Z^*(\mathcal{V}) = \min_{\boldsymbol{u}\in\mathcal{V}} f\,(S^*(\mathcal{V}),\boldsymbol{u})$, we have $w_n < f(S^*(\mathcal{V}),\boldsymbol{v}) \leq f(S^*_{\boldsymbol{v}},\boldsymbol{v})$, where the second inequality follows from the definition of $S^*_{\boldsymbol{v}}$. In this case, the contrapositives of the statements (b) and (c) in Lemma 2.2 with $A = \{n\}$ and $B = S^*_{\boldsymbol{v}}\setminus\{n\}$ imply that $f(S^*_{\boldsymbol{v}},\boldsymbol{v}) < f(S^*_{\boldsymbol{v}}\setminus\{n\},\boldsymbol{v})$. The last inequality contradicts the definition of $S^*_{\boldsymbol{v}}$! So, we must have $n \notin S^*_{\boldsymbol{v}}$ for all $\boldsymbol{v}\in\mathcal{V}$.

Collecting the two results that we have shown under Case 2, we have $n \notin S^*(\mathcal{V})$ and $n \notin S^*_{\boldsymbol{v}}$ for all $\boldsymbol{v}\in\mathcal{V}$. Therefore, we can drop product n from consideration without changing $S^*(\mathcal{V})$, $\{S^*_{\boldsymbol{v}} : \boldsymbol{v}\in\mathcal{V}\}$ and $Z^*(\mathcal{V})$, in which case, we have an assortment optimization problem with $n-1$ products and the desired result follows from the induction hypothesis. □

3. Numerical Experiments

In this section, our goal is to investigate the performance of the assortment that is obtained under Robust Logit problem and contrast this assortment with the one that is obtained under Mixture Logit problem described in Section 1.

3.1 Experimental Setup

We consider a finite uncertainty set $\mathcal{V} = \left\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^G\right\}$ consisting of G vectors, where $\boldsymbol{v}^g \in \Re^n_+$ for $g = 1, \ldots, G$. The vector $\boldsymbol{v}^g = (v^g_1, \ldots, v^g_n)$ can be interpreted as the preference weights of a customer class g and we have a total of G customer classes. If the proportion of customers that belong to customer class g is given by Θ^g, then we can choose an assortment by solving Mixture Logit problem after replacing $\theta(\boldsymbol{v}^g)$ with Θ^g. In our numerical experiments, we assume that we do not know the proportions $\boldsymbol{\Theta} = (\Theta^1, \ldots, \Theta^G)$, but we estimate them to be $\bar{\boldsymbol{\theta}} = (\bar{\theta}^1, \ldots, \bar{\theta}^G)$. In this case, a reasonable approach for choosing an assortment is to solve Mixture Logit problem after replacing $\theta(\boldsymbol{v}^g)$ with the estimated proportion $\bar{\theta}^g$. We use M^* to denote the assortment that we choose in this fashion. Bront et al. (2009) show that Mixture Logit problem can be transformed into a mixed integer program and we use the mixed integer programming routine in CPLEX to compute M^*. If the proportions $\boldsymbol{\Theta} = (\Theta^1, \ldots, \Theta^G)$ are not known, then another reasonable approach for choosing an assortment is to solve Robust Logit problem corresponding to the uncertainty set $\mathcal{V} = \left\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^G\right\}$. We use R^* to denote the assortment that we choose by solving Robust Logit problem.

To model the unknown proportions, we assume that $\boldsymbol{\Theta}$ is given by a G-dimensional Dirichlet random variable. The mean vector of $\boldsymbol{\Theta}$ is given by $\bar{\boldsymbol{\theta}}$ and the coefficient of variation of Θ^g is given by ρ for all $g = 1, \ldots, G$. We vary ρ in our numerical experiments. Our experimental setup is intended to capture a situation where the unknown proportions $\boldsymbol{\Theta}$ do not exactly match the estimated proportions $\bar{\boldsymbol{\theta}}$, but the two sets of proportions match in expectation. By varying the coefficient of variation ρ, we control how much the unknown proportions $\boldsymbol{\Theta}$ tend to deviate from the estimated proportions $\bar{\boldsymbol{\theta}}$. Our goal is to compare the distributions of the random variables

$$\mathsf{MixtureRev} \equiv \sum_{g=1}^{G} \Theta^g f\left(M^*, \boldsymbol{v}^g\right) \qquad \text{and} \qquad \mathsf{RobustRev} \equiv \sum_{g=1}^{G} \Theta^g f\left(R^*, \boldsymbol{v}^g\right) \ ,$$

which are the actual expected revenues under the assortments M^* and R^*, respectively. On average, the expected revenue under the assortment M^* is higher than the expected revenue under the assortment R^* because $\mathbb{E}\left[\mathsf{MixtureRev}\right] = \sum_{g=1}^{G} \mathbb{E}\left[\Theta^g\right] f\left(M^*, \boldsymbol{v}^g\right) = \sum_{g=1}^{G} \bar{\theta}^g f\left(M^*, \boldsymbol{v}^g\right) = \max_{S \subseteq \{1,\ldots,n\}} \sum_{g=1}^{G} \bar{\theta}^g f\left(S, \boldsymbol{v}^g\right)$ $\geq \sum_{g=1}^{G} \bar{\theta}^g f\left(R^*, \boldsymbol{v}^g\right) = \mathbb{E}\left[\mathsf{RobustRev}\right]$. However, as we shortly show in our numerical experiments, the expected revenue under the assortment R^* has a *significantly smaller* variability when compared with the expected revenue under the assortment M^*.

To test the effectiveness of our assortments, we consider multiple problem classes. Each problem class is characterized by the number of customer classes G, the number of products n and the coefficient of variation ρ. We generate 1,000 test problems for each problem class. Each test problem in a problem class (G, n, ρ) is determined by randomly generating G vectors $\left\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^G\right\}$ to form an uncertainty set

$\mathcal{V}$, randomly generating the revenues $(w_1, \ldots, w_n)$ of the products and randomly choosing the estimated distribution $\bar{\boldsymbol{\theta}}$. We then compare the histograms of the random variables $\mathsf{MixtureRev}$ and $\mathsf{RobustRev}$ by generating 1,000,000 samples from the random vector $\boldsymbol{\Theta}$. In the appendix, we give the complete details on how we generate $\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^G\}$, $(w_1, \ldots, w_n)$ and $\bar{\boldsymbol{\theta}}$.

3.2 Numerical Results

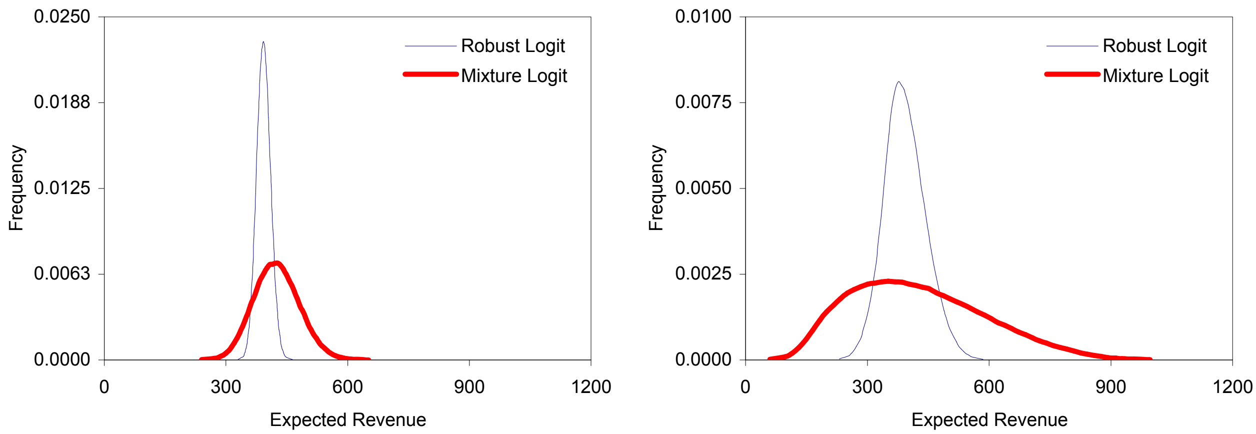
Figure 1 shows the histograms of $\mathsf{MixtureRev}$ and $\mathsf{RobustRev}$ for two test problems. Each chart in Figure 1 corresponds to one test problem. Both test problems have 3 customer classes and 20 products. All of the problem parameters for the two test problems are the same with the exception of the coefficient of variation ρ. The coefficient of variations are 0.28 and 0.85 for the test problems on the left and right side, respectively. The thick data series plot the histogram for $\mathsf{MixtureRev}$, whereas the thin data series plot the histogram for $\mathsf{RobustRev}$.

We observe that the widths of the histograms for $\mathsf{RobustRev}$ tend to be significantly tighter than those of the histograms for $\mathsf{MixtureRev}$. For example, in Figure 1(b), the standard deviation of the histogram for $\mathsf{RobustRev}$ is 72, which is smaller than that of the histogram for $\mathsf{MixtureRev}$, which is 165. As expected, the average revenues under the robust assortment R^* are smaller than those under the assortment M^*. Going back to Figure 1(b) again, the average of the histogram for $\mathsf{RobustRev}$ is 393, whereas the average of the histogram for $\mathsf{MixtureRev}$ is 424. Finally, the tails of the histograms for $\mathsf{RobustRev}$ tend to be shorter than those of the histograms for $\mathsf{MixtureRev}$. In Figure 1(b), the first percentiles of the histograms for $\mathsf{RobustRev}$ and $\mathsf{MixtureRev}$ are 282 and 135, respectively. In other words, for this test problem, if we offer the robust assortment R^* to our customers, then the expected revenue per customer exceeds 282 with 99% probability, whereas if we offer the assortment M^*, then the expected revenue per customer exceeds 135 with the same probability. Comparing the charts on the left and right sides of Figure 1 with each other, we observe that as the coefficient of variation increases, the tail of the histogram for $\mathsf{MixtureRev}$ gets longer when compared with the tail of the histogram for $\mathsf{RobustRev}$. Therefore, the risk associated with the assortment M^* becomes more apparent as the coefficient of variation increases. This observation confirms our intuition that it is a good idea use the Mixture Logit formulation when the distribution of $\boldsymbol{\Theta}$ is tightly concentrated around its mean and we can estimate $\boldsymbol{\Theta}$ with high accuracy. On the other hand, it is beneficial to switch to the Robust Logit formulation when we cannot estimate $\boldsymbol{\Theta}$ with high accuracy.

In Table 1, we compare the assortments chosen by Robust Logit and Mixture Logit problems across 27 problem classes, each problem class corresponding to a different combination of (G, n, ρ). For the kth test problem in a problem class, we carry out an analysis similar to the one in Figure 1 and compute the ratios between the first percentiles, standard deviations and means of the random variables $\mathsf{MixtureRev}$ and $\mathsf{RobustRev}$. In other words, letting $\mathsf{MixtureRev}(k)$ and $\mathsf{RobustRev}(k)$, respectively, be the expected revenues under the assortments M^* and R^* in the kth test problem, we compute

$$\frac{\text{1st Per } \mathsf{RobustRev}(k)}{\text{1st Per } \mathsf{MixtureRev}(k)}, \quad \frac{\text{Std } \mathsf{RobustRev}(k)}{\text{Std } \mathsf{MixtureRev}(k)}, \quad \frac{\text{Avg } \mathsf{RobustRev}(k)}{\text{Avg } \mathsf{MixtureRev}(k)}.$$

In Table 1, we report the averages of these ratios over 1,000 test problems in each problem class. The



(a) $\rho = 0.28$ (b) $\rho = 0.85$

Figure 1: Comparison between the revenues under the assortments chosen by ROBUST LOGIT and MIXTURE LOGIT problems for two test problems. Both test problems have 3 customer classes and 20 products. The test problem corresponding to the chart on the right side has a higher coefficient of variation.

first column in Table 1 shows G, n and ρ for the problem class. The second, third and fourth columns focus on the ratios between the first percentiles, standard deviations and means, respectively.

The results in Table 1 agree with those in Figure 1. The first percentiles of RobustRev tend to be larger than those of MixtureRev, indicating that the tails of RobustRev tend to be shorter than the tails of MixtureRev. Furthermore, as the coefficient of variation ρ increases, the gap in the first percentiles of RobustRev and MixtureRev generally increases. Therefore, the risk associated with the assortment M^* becomes more apparent when $\boldsymbol{\Theta}$ can deviate significantly from its mean. The standard deviations of RobustRev are significantly smaller than those of MixtureRev. The gap between the standard deviations can be as large as 36% for the test problems with a large number of customer classes. Finally, the average performance of the robust assortment R^* is not as good as the average performance of the assortment M^*, and the gap can be as high as 10%. Therefore, although the average performance of the robust assortment R^* is worse than the average performance of the assortment M^*, the risk associated with the robust assortment is significantly smaller than the risk associated with the assortment M^*, especially when we have high coefficient of variations and the unknown proportions $\boldsymbol{\Theta}$ can deviate significantly from the estimated proportions $\bar{\boldsymbol{\theta}}$.

4. Conclusions

We formulated a robust assortment optimization problem under the multinomial logit choice model with an unknown parameter. The formulation admits an efficient solution and the optimal solution indicates that a large product variety serves as a safety buffer against uncertainty in the model parameter. It would be interesting to see whether our analysis can be extended to more complex choice models, such as the nested multinomial logit choice model.

Prob class (G,n,ρ)	Ratios between 1st Per	Std	Avg
$(3,20,1.0)$	1.17	0.75	0.94
$(3,20,1.2)$	1.12	0.74	0.94
$(3,20,1.4)$	1.13	0.74	0.94
$(3,40,1.0)$	1.10	0.75	0.94
$(3,40,1.2)$	1.10	0.75	0.94
$(3,40,1.4)$	1.11	0.75	0.94
$(3,60,1.0)$	1.09	0.75	0.94
$(3,60,1.2)$	1.09	0.75	0.94
$(3,60,1.4)$	1.09	0.75	0.94
$(6,20,1.0)$	1.04	0.69	0.92
$(6,20,1.4)$	1.11	0.69	0.92
$(6,20,2.0)$	1.16	0.70	0.92
$(6,40,1.0)$	1.03	0.68	0.92
$(6,40,1.4)$	1.09	0.68	0.92
$(6,40,2.0)$	1.14	0.67	0.92
$(6,60,1.0)$	1.04	0.68	0.92
$(6,60,1.4)$	1.09	0.68	0.92
$(6,60,2.0)$	1.12	0.67	0.92
$(12,20,1.5)$	1.04	0.66	0.91
$(12,20,2.0)$	1.10	0.66	0.90
$(12,20,3.0)$	1.22	0.65	0.90
$(12,40,1.5)$	1.03	0.64	0.90
$(12,40,2.0)$	1.09	0.64	0.90
$(12,40,3.0)$	1.17	0.65	0.91
$(12,60,1.5)$	1.03	0.64	0.91
$(12,60,2.0)$	1.08	0.64	0.90
$(12,60,3.0)$	1.16	0.64	0.90

Table 1: Comparison between the revenues under the assortments chosen by Robust Logit and Mixture Logit problems for 27 problem classes.

References

Anderson, S., A. de Palma, and J. F. Thisse. 1992. *Discrete choice theory of product differentiation*. Cambridge, MA: MIT Press.

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.

Cachon, G., C. Terwiesch, and Y. Xu. 2005. Assortment Planning in the Presence of Consumer Search. *Manufacturing and Service Operations Management* 7:330–346.

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

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.

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.

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

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.

van Ryzin, G., and S. Mahajan. 1999. On the Relationship Between Inventory Costs and Variety Benefits in Retail Assortments. *Management Science* 45:1496–1509.

Vulcano, G., G. J. van Ryzin, and R. Ratliff. 2008. Estimating primary demand for substitutable products from sales transaction data. Working paper, Stern Business School.

Zhang, D., and D. Adelman. 2006. An Approximate Dynamic Programming Approach to Network Revenue Management with Customer Choice. Technical report, University of Chicago, Graduate School of Business.

A. Appendix: Description of Test Problems

In all of our test problems, we assume that the estimated proportions are given by $\bar{\boldsymbol{\theta}} = (1/G, \ldots, 1/G)$ so that we estimate each customer class to be equally likely. To come up with the preference weights for each customer class, we sample the parameter σ_i from the uniform distribution over $[0, 1]$ for each product i. A large value for σ_i indicates that product i is a specialty product and the preference weights associated with product i show large variability among the different customer classes. On the other hand, a small value for σ_i indicates that product i is a staple product and the preference weights associated with product i does not show large variability among the different customer classes. In this case, we sample the parameter ϑ_i^g from the uniform distribution over $[0, 10]$ for each product i and customer class g, and set $v_i^g = (1-\sigma_i)\,\vartheta_i^g/n$ with probability 1/2 and $v_i^g = (1+\sigma_i)\,\vartheta_i^g/n$ with probability 1/2. To see the motivation behind our choice of the preference weights, we observe that if σ_i is close to zero, then v_i^g is close to ϑ_i^g/n. On the other hand, if σ_i is close to one, then v_i^g is either close to zero or close to $2\,\vartheta_i^g$. Therefore, as we mention above, the parameter σ_i is a measure of the variability of the preference weights associated with product i among the different customer classes. Furthermore, we observe that the expectation of v_i^g is $5/n$. If v_i^g takes a value close to its expectation and we offer all products to our customers, then the probability that a customer leaves without purchasing anything is given by $1/[1+\sum_{i=1}^n 5/n] = 1/6$. Therefore, we expect to have a significant probability that a customer leaves without purchasing anything even if we offer all products to our customers.

To come up with the revenues of the products, we sample κ_i from the uniform distribution over $[0, 200]$ and set $w_i = [n+1-i]\,\kappa_i$. In this way, we hope the generate a set of products with a large variety in revenues. After generating the revenues of all of the products, we renumber the products to make sure that $w_1 \geq w_2 \geq \ldots \geq w_n$.