Robust Assortment Optimization in Revenue Management Under the Multinomial Logit Choice Model

Paat Rusmevichientong[*]   Huseyin Topaloglu[†]

September 20, 2011

Abstract

We study robust formulations of assortment optimization problems under the multinomial logit choice model. The novel aspect of our formulations is that the true parameters of the logit model are assumed to be unknown, and we represent the set of likely parameter values by a compact uncertainty set. The objective is to find an assortment that maximizes the worst case expected revenue over all parameter values in the uncertainty set. We consider both static and dynamic settings. The static setting ignores inventory consideration, while in the dynamic setting, there is a limited initial inventory that must be allocated over time. We give a complete characterization of the optimal policy in both settings, show that it can be computed efficiently, and derive operational insights. We also propose a family of uncertainty sets that enables the decision maker to control the tradeoff between increasing the average revenue and protecting against the worst case scenario. Numerical experiments show that our robust approach, combined with our proposed family of uncertainty sets, is especially beneficial when there is significant uncertainty in the parameter values. When compared to other methods, our robust approach yields over 10% improvement in the worst case performance, but it can also maintain comparable average revenue if average revenue is the performance measure of interest.

1. Introduction

A traditional assumption in the revenue management literature has been that each customer arrives into the system with the intention of purchasing a fixed product. If this product is available, then the customer makes a purchase; otherwise, he leaves the system without purchasing anything. In many applications, however, the customer observes a set of available products and makes a selection among them. In this case, the firm needs to decide which assortment of products to offer, taking into account the choice process of the customers. The seminal work of Talluri and van Ryzin (2004) has demonstrated the importance of incorporating customer choice behavior into revenue management models. Since their formulation of the single leg airline seat allocation problem under a general discrete choice model, there has been a large amount of research on choice-based revenue management models. Most of this work assumes that the parameters of the underlying choice model are either known in advance or have been estimated from data. The revenue management decision is then computed based on the estimated parameter values, ignoring any uncertainty associated with the estimates. Currently, existing models do not take into account the losses that may incur when the estimated values differ significantly from the true parameter values.

---

[*]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

In this paper, we formulate assortment optimization models that explicitly incorporate the uncertainty in the choice parameters. We consider a setting where the set of products is indexed by $\mathcal{A} = \{1, \ldots, n\}$, and the revenue of each product is ordered such that $r_1 \geq r_2 \geq \cdots \geq r_n > 0$. The customer choices are driven by the multinomial logit model. However, the parameters of the logit model are unknown, and the set of likely parameter values is described by a compact uncertainty set. Our goal is to choose an assortment that maximizes the worst case expected revenue, where the worst case is taken over all likely parameter values. We consider both static and dynamic settings. The static setting ignores inventory consideration, while in the dynamic setting, we have a fixed initial inventory that has to be allocated to customers arriving over time. As described below, we make contributions in understanding the policy structure, developing efficient computational methods, and deriving important operational insights.

We consider the static problem in Section 3. When the parameters of the logit model are known in advance, Talluri and van Ryzin (2004) show that assortments of the form $\{1, 2, \ldots, i\}$ for some product i are optimal. We call such assortments as revenue-ordered assortments. Surprisingly, we are able to show that revenue-ordered assortments remain optimal, even when we are uncertain about the model parameters and we wish to protect against the worst case expected revenue (Theorem 3.2). This result shows that we can compute the robust assortment efficiently because we only need to consider n revenue-ordered assortments, instead of all 2^n subsets. In addition to generalizing the result of Talluri and van Ryzin (2004), our proof of Theorem 3.2 is novel and different from the existing techniques in the literature, contributing to a new understanding of the logit model (Lemma 3.1).

In addition to characterizing the policy structure and providing efficient computational methods, our analysis provides operational insights about how to protect against uncertainty in the model parameters. In Corollary 3.5, we show that larger uncertainty leads to a larger robust assortment. Thus, product variety serves as a buffer against worst case scenario. Moreover, if a firm is operating in a market with multiple customer segments, then to maximize the worst case expected revenue, it should target the customer segment whose corresponding preference profile gives rise to the largest optimal assortment (Theorem 3.6). We also show that firms that carry higher revenue-generating products and face uncertainty in the marketplace would benefit from offering larger assortments (Theorem 3.7). In Section 3.3, we propose a family of uncertainty sets that enables the decision maker to control the tradeoff between increasing the average revenue and protecting against the worst case scenario. The uncertainty sets we propose are characterized by a radius parameter ϵ. Large values of ϵ provide large uncertainty sets, resulting in conservative assortments that protect well against the worst case expected revenue, but lack in average performance. Small values of ϵ provide aggressive assortments with high average performance, but these assortments may not perform well in terms of the worst case expected revenue. Depending on the preferences of the decision maker, the radius parameter ϵ allows us to obtain a variety of assortments with different degrees of conservatism. Our uncertainty sets perform well in our numerical experiments, providing significant improvements in the worst case revenue or delivering high average revenue depending on the preferences of the decision maker.

Most of the structural properties in the static problem extend to the dynamic setting, which we

consider in Section 4. We show that it remains optimal to offer revenue-ordered assortments in each period, allowing us to compute the optimal robust policy efficiently (Theorem 4.2). This result shows that revenue-ordered assortments are *complete efficient sets*, and establishes that the *nesting-by-fare-order* property continues to hold in the robust setting, generalizing the result of Talluri and van Ryzin (2004) who establish these properties in the setting where the parameters of the logit model are known in advance. We show that as the remaining inventory decreases, the robust assortment shrinks. On the other hand, as we get closer to the end of the selling season, the robust assortment grows, increasing our chance of liquidating the inventory (Theorem 4.3). Thus, the robust policy satisfies intuitive properties, which should hopefully increase its practical appeal. We note that these results are under the assumption that consumers are not strategic about the timing of their purchases. Liu and van Ryzin (2008) show that strategic consumers may stockpile the products at the beginning of the selling season because of rationing risk (see, also, Su, 2010), or postpone their purchases until later in anticipation of future markdowns (Su, 2007). Incorporating strategic consumer behavior when the parameters of the underlying choice model are unknown remains an open research area.

The key takeaway message in this paper is that the cost of robustness is minimal. Computing the optimal robust policy, both in static and dynamic settings, can be done efficiently, and it is almost as easy as the non-robust case. Yet, when there is significant uncertainty in the model parameters, our numerical experiments in Section 5 show that our robust approach offers substantial benefits. The robust policy improves the worst case expected revenue by over 10%, when compared to policies which assume that the estimated parameters are accurate and do not account for uncertainty in the estimates. Furthermore, the revenues provided by the robust assortments are very stable, showing 50% less variability when compared to the revenues obtained under other approaches. The family of uncertainty sets we propose balances the competing incentives of increasing the average revenue and protecting against the worst case performance, enabling the decision maker to choose the uncertainty set that is appropriate for each decision making setting. The bottom line is that our robust approach can be quite effective in reducing the risk caused by uncertain model parameters, but it also allows for comparable revenues in average cases if average case happens to be the performance measure of interest.

2. Literature Review

There is a significant amount of research that tries to incorporate customer choice behavior into revenue management models. Talluri and van Ryzin (2004) consider a single leg airline seat allocation problem, where customers make a choice between the available fare classes. They show that if the customers make a choice according to the multinomial logit model, then it is optimal to offer only revenue-ordered assortments. Gallego et al. (2004) focus on airline networks and develop a linear program to approximate the total expected revenue. Liu and van Ryzin (2008) build on this linear program to extend bid pricing ideas to deal with customer choice. Zhang and Adelman (2009), Meissner and Strauss (2011, 2010), and Zhang (2011) use the linear programming approach to approximate dynamic programming to come up with different approximations to the value functions that appear in the dynamic programming formulation of the airline network revenue management problem with customer choice. Kunnumkal and Topaloglu (2008, 2010) provide refinements on the linear program proposed by Gallego et al. (2004)

to be able to tighten the approximations of the total expected revenue. The paper by van Ryzin and Vulcano (2008) develops effective optimization methods for computing virtual nesting controls for network revenue management problems under customer choice behavior. Bront et al. (2009) and Mendez-Diaz et al. (2010) focus on the case with multiple customer types each making a choice according to the multinomial logit model, show that the underlying problem is NP-hard, and provide effective heuristics for finding good assortments. Chaneton and Vulcano (2011) use stochastic approximation to obtain bid prices under customer choice. Talluri (2010) proposes a deterministic concave program for the network revenue management problem with customer choice behavior.

The papers above assume that the model parameters are known in advance. This paper, in contrast, considers the setting where the parameters are unknown, and we need to find an assortment that protects against the worst case expected revenue over all parameter values in the uncertainty set. To our knowledge, our paper is the first to model explicitly the uncertainty in the choice parameters and to try to minimize losses that one incurs when there are errors in the estimated parameter values. By accounting for errors that may arise from estimation under limited data, this paper complements recent work by Vulcano et al. (2011), who develop algorithms for estimating choice model parameters under censored sales transaction data. Vulcano et al. (2010) conduct an empirical study on real data by using the estimation technique suggested by Talluri and van Ryzin (2004), and their work demonstrates that such estimation techniques, when combined with capacity control policies, can significantly increase the revenue for airlines.

Farias et al. (2010) provide an alternative approach for dealing with uncertainty in the choice model. They consider the problem of estimating a revenue of an assortment, when the underlying choice model and its parameters are unknown, and the decision maker is given only partial data on customer preferences. For a fixed assortment, they formulate an optimization problem that finds the worst case revenue of the assortment, where worst case is among all choice models that are consistent with the data. However, they do not consider the combinatorial optimization problem of finding the assortment that maximizes the worst case revenue. Their approach is *non-parametric* because they do not make any a priori assumption about the structure of the underlying choice model. Our paper complements their work as it provides a parametric approach to the problem by assuming an underlying multinomial logit choice model. Our parametric framework leads to a tractable solution, and allows us to determine the assortment with the maximum worst case revenue.

If we consider the case where the uncertainty set is a singleton, then we essentially know the choice parameters in advance, and our results subsume some of the existing results that assume full knowledge of choice parameters. For example, by using different proof techniques, Talluri and van Ryzin (2004), Gallego et al. (2004), Liu and van Ryzin (2008) and Kunnumkal and Topaloglu (2008) show the optimality of revenue-ordered assortments when the choice parameters are known in advance. We generalize this result by showing that the same class of assortments are optimal when the set of likely parameter values is an arbitrary compact set. In addition, the proofs in the earlier papers resort to quasi-linearity or knapsack arguments, but our proofs directly follow from first principles.

Robust optimization has recently seen significant attention as a potential approach for dealing

with uncertainty (see, for example, Ben-Tal and Nemirovski, 1999; Bertsimas and Sim, 2004). Iyengar (2005) formalizes a robust Markov decision process framework and our dynamic model closely follows his formalism. Birbil et al. (2009) use robust optimization in single leg airline revenue management problems where the arrival probabilities for different customer types are not known and are assumed to lie in an ellipsoidal uncertainty set. We construct a similar family of uncertainty sets in Section 3.3 that enables the decision maker to control the tradeoff between increasing the average revenue and protecting against the worst case scenario. Perakis and Roels (2010) use robust optimization for network revenue management problems. The last two papers exclusively work under the assumption that the customers are interested in a single itinerary and they do not make a choice among different itineraries that may satisfy their needs, but choice is of central concern in our paper.

We use the multinomial logit model to capture the customer preferences. Although this is not the only approach to model choice, multinomial logit model is one of the most commonly used models in economics, marketing and operations management. We refer the reader to Ben-Akiva and Lerman (1985), Anderson et al. (1992), van Ryzin and Mahajan (1999), Kok et al. (2008) and the references therein for background reading on the multinomial logit model.

3.   A Static Model: Robust Assortment Optimization

In this section, we focus on a static robust assortment optimization problem under the multinomial logit model. The model is static in the sense that it ignores inventory consideration. The results in this section are relevant to environments where inventory is not a limiting issue, which is particularly the case when selling digital products such as electronic books and music. Furthermore, the insights in this section will guide our analysis of the dynamic capacity allocation problem in Section 4, where we need to allocate a fixed initial inventory over time.

In our problem setup, we have a set of products indexed by $\mathcal{A} = \{1, \ldots, n\}$. For each product i, let $r_i > 0$ denote its revenue. Without loss of generality, we assume that the products are ordered such that $r_1 \geq r_2 \geq \cdots \geq r_n > 0$. Each customer chooses a product from an assortment according to a multinomial logit model with unknown parameters described by a vector $\boldsymbol{v} = (v_0, v_1, \ldots, v_n) \in \mathbb{R}_{++}^{n+1}$, where $\mathbb{R}_{++}$ denotes the set of positive real numbers. We interchangeably refer to $\boldsymbol{v}$ as the model parameters or the customer preference weights. If we offer the assortment $S \subseteq \mathcal{A}$, then the probability $\phi_i(S, \boldsymbol{v})$ that a customer purchases product i is given by

$$\phi_i(S, \boldsymbol{v}) = \begin{cases} \dfrac{v_i}{v_0 + \sum_{\ell \in S} v_\ell} & \text{if } i \in S, \\ 0 & \text{otherwise,} \end{cases}$$

and with the remaining probability $1 - \sum_{i \in S} \phi_i(S, \boldsymbol{v}) = v_0/[v_0 + \sum_{i \in S} v_i]$, the customer simply does not purchase any product. When an assortment S is offered, the expected revenue that we obtain from a customer is given by

$$f(S, \boldsymbol{v}) = \sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) = \frac{\sum_{i \in S} r_i \, v_i}{v_0 + \sum_{i \in S} v_i} \, .$$

The uncertainty in the parameters of the multinomial logit model is represented by a compact uncertainty set $\mathcal{V} \subseteq \mathbb{R}_{++}^{n+1}$. In our model, we want to find an assortment that maximizes the worst case expected revenue over all model parameters in $\mathcal{V}$, corresponding to the optimization problem

$$Z^*(\mathcal{V}) = \max_{S \subseteq \mathcal{A}} \left\{ \min_{\boldsymbol{v} \in \mathcal{V}} \ f(S, \boldsymbol{v}) \right\} . \qquad \text{(ROBUST LOGIT)}$$

We define $S^*(\mathcal{V})$ to be an optimal assortment under the ROBUST LOGIT problem. If there are ties, then we choose $S^*(\mathcal{V})$ to be an optimal assortment with the smallest cardinality[1]. We 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 parameters 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 parameters exactly and our robust formulation reduces to the classical assortment optimization problem under the multinomial logit model. To facilitate our exposition, we use $S_{\boldsymbol{v}}^*$ to denote $S^*(\{\boldsymbol{v}\})$, which is an optimal assortment when the model parameters $\boldsymbol{v}$ are known exactly. Also, we denote $Z^*(\{\boldsymbol{v}\})$ by $Z_{\boldsymbol{v}}^*$. In other words, $S_{\boldsymbol{v}}^*$ is an optimal solution to the assortment optimization problem associated with $Z_{\boldsymbol{v}}^* = \max_{S \subseteq \mathcal{A}} f(S, \boldsymbol{v})$, and we break ties by choosing $S_{\boldsymbol{v}}^*$ as an optimal assortment with the smallest cardinality. Talluri and van Ryzin (2004), Gallego et al. (2004), Liu and van Ryzin (2008) and Kunnumkal and Topaloglu (2008) show that $S_{\boldsymbol{v}}^*$ is a revenue-ordered assortment consisting of products with the highest revenues, that is, $S_{\boldsymbol{v}}^* = \{1, 2, \ldots, i_{\boldsymbol{v}}^*\}$ for some product $i_{\boldsymbol{v}}^*$. This result allows us to consider only assortments of the form $\{1, \ldots, i\}$, instead all of 2^n possible assortments.

Surprisingly, even when we have uncertainty in the model parameters, the robust assortment $S^*(\mathcal{V})$ shares the same structural property. Before we proceed to the statement of this result, the following lemma characterizes when adding a product to an assortment increases the expected revenue.

Lemma 3.1 (When is Adding a Product Beneficial?). *For any $\boldsymbol{v} \in \mathbb{R}_{++}^{n+1}$, $A \subseteq \mathcal{A}$ and $i \notin A$, the following three statements are equivalent:*

(a) $r_i > f(A, \boldsymbol{v})$,   (b) $f(A \cup \{i\}, \boldsymbol{v}) > f(A, \boldsymbol{v})$,   *and*   (c) $r_i > f(A \cup \{i\}, \boldsymbol{v})$.

Proof. We note that $f(A \cup \{i\}, \boldsymbol{v})$ is a convex combination of r_i and $f(A, \boldsymbol{v})$ because

$$f(A \cup \{i\}, \boldsymbol{v}) = \left(\frac{v_i}{v_0 + v_i + \sum_{\ell \in \mathcal{A}} v_\ell} \cdot r_i \right) + \left(\frac{v_0 + \sum_{\ell \in A} v_\ell}{v_0 + v_i + \sum_{\ell \in \mathcal{A}} v_\ell} \cdot f(A, \boldsymbol{v}) \right) .$$

Therefore, $f(A \cup \{i\}, \boldsymbol{v})$ is between r_i and $f(A, \boldsymbol{v})$, which implies that $f(A \cup \{i\}, \boldsymbol{v}) > f(A, \boldsymbol{v})$ if and only if $r_i > f(A, \boldsymbol{v})$. This establishes the equivalence between (a) and (b). The equivalence between statements (b) and (c) follows from the same argument. $\qquad \square$

The main result of this section is stated in the following theorem, which provides an important structural property of the optimal robust assortment $S^*(\mathcal{V})$. The result shows that the optimal robust

---

assortment consists of products whose revenues exceed a particular value, which corresponds to the optimal objective value $Z^*(\mathcal{V})$ of the ROBUST LOGIT problem. Given that the product revenues are ordered such that $r_1 \geq r_2 \geq \cdots \geq r_n$, it follows from Theorem 3.2 that the robust assortment $S^*(\mathcal{V})$ is a revenue-ordered assortment of the form $\{1, 2, \ldots, i^*(\mathcal{V})\}$, where $i^*(\mathcal{V})$ is the last product whose revenue still exceeds $Z^*(\mathcal{V})$. This greatly simplifies the computation of the robust assortment since we only need to consider at most n assortments.

Theorem 3.2 (Revenue-Ordered Assortments are Robust). *For any $\mathcal{V} \subset \mathbb{R}_{++}^n$, $S^*(\mathcal{V}) = \{i : r_i > Z^*(\mathcal{V})\}$.*

Proof. We first show that $\{i : r_i > Z^*(\mathcal{V})\} \subseteq S^*(\mathcal{V})$. To prove this claim, suppose on the contrary that there exists a product i such that $r_i > Z^*(\mathcal{V})$ and $i \notin S^*(\mathcal{V})$. Consider an arbitrary model parameter vector $\boldsymbol{v} \in \mathcal{V}$. By definition of $S^*(\mathcal{V})$, we have $f(S^*(\mathcal{V}), \boldsymbol{v}) \geq Z^*(\mathcal{V})$. Suppose that $r_i > f(S^*(\mathcal{V}), \boldsymbol{v})$, in which case, it follows from statements (a) and (b) in Lemma 3.1 that $f(S^*(\mathcal{V}) \cup \{i\}, \boldsymbol{v}) > f(S^*(\mathcal{V}), \boldsymbol{v}) \geq Z^*(\mathcal{V})$. On the other hand, if $r_i \leq f(S^*(\mathcal{V}), \boldsymbol{v})$, then it follows from the contrapositives of statements (a) and (c) in Lemma 3.1 that $f(S^*(\mathcal{V}) \cup \{i\}, \boldsymbol{v}) \geq r_i > Z^*(\mathcal{V})$. So, in both cases, we have $f(S^*(\mathcal{V}) \cup \{i\}, \boldsymbol{v}) > Z^*(\mathcal{V})$. Since $\boldsymbol{v} \in \mathcal{V}$ is arbitrary and $\mathcal{V}$ is compact, it follows that $\min_{\boldsymbol{v} \in \mathcal{V}} f(S^*(\mathcal{V}) \cup \{i\}, \boldsymbol{v}) > Z^*(\mathcal{V})$. This contradicts the optimality of $S^*(\mathcal{V})$ for the ROBUST LOGIT problem. Therefore, we have $i \in S^*(\mathcal{V})$, establishing the desired claim.

To complete the proof, we show that $S^*(\mathcal{V}) \subseteq \{i : r_i > Z^*(\mathcal{V})\}$. Suppose on the contrary that there exists a product i such that $r_i \leq Z^*(\mathcal{V})$ and $i \in S^*(\mathcal{V})$. Consider an arbitrary model parameter vector $\boldsymbol{v} \in \mathcal{V}$. By definition of $S^*(\mathcal{V})$, we have $r_i \leq Z^*(\mathcal{V}) \leq f(S^*(\mathcal{V}), \boldsymbol{v})$, in which case, it follows from the contrapositives of statements (b) and (c) in Lemma 3.1 with $A = S^*(\mathcal{V}) \setminus \{i\}$ that $f(S^*(\mathcal{V}), \boldsymbol{v}) \leq f(S^*(\mathcal{V}) \setminus \{i\}, \boldsymbol{v})$. Since $\boldsymbol{v} \in \mathcal{V}$ is arbitrary, it follows that

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

which implies that $S^*(\mathcal{V}) \setminus \{i\}$ is also an optimal robust assortment. However, this contradicts our hypothesis that $S^*(\mathcal{V})$ is the optimal robust assortment with the smallest cardinality. Therefore, it must be the case that $i \notin S^*(\mathcal{V})$. This completes the proof. $\square$

We note that since Theorem 3.2 continues to hold when the uncertainty set $\mathcal{V}$ is a singleton, it provides an alternative proof of the fact that revenue-ordered assortments are optimal when the model parameters are known exactly.

To determine the optimal robust revenue $Z^*(\mathcal{V})$, it follows from Theorem 3.2 that we need to evaluate the worst case expected revenue $\min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v})$ of each revenue-ordered assortment S. When $\mathcal{V}$ is finite, we can compute this minimum by enumeration. For general uncertainty sets, however, more sophisticated optimization techniques may be required. It turns out that when $\mathcal{V}$ is a polyhedron, the worst case revenue can be computed using a linear program, as shown in the following example.

Example 3.3 (Polyhedron Uncertainty Set). Suppose that the uncertainty set $\mathcal{V}$ is a polyhedron, that is, $\mathcal{V} = \left\{\boldsymbol{v} \in \mathbb{R}_{++}^{n+1} : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{b}\right\}$ for some matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$. Then, for any assortment $S \subseteq \mathcal{A}$, the

worst case expected revenue under S is given by

$$\min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v}) \;\; = \;\; \min_{\boldsymbol{v}: \boldsymbol{A}\boldsymbol{v}=\boldsymbol{b}} \frac{\sum_{i \in S} r_i v_i}{v_0 + \sum_{i \in S} v_i} \;.$$

By making the change of variables $z = 1/(v_0 + \sum_{i \in S} v_i)$ and $y_k = v_k/(v_0 + \sum_{i \in S} v_i)$ for $k = 0, 1, \ldots, n$, the optimization problem on the right side above can be rewritten as the linear program

$$\min \left\{ \sum_{i \in S} r_i y_i : \boldsymbol{A}\boldsymbol{y} = z\boldsymbol{b}, \;\; \sum_{i=0}^n y_i = 1, \; z \geq 0, \text{ and } y_k \geq 0, \;\; k = 0, 1, \ldots, n \right\} \;.$$

The proof of equivalence between the last two optimization problems is a standard result in linear-fractional programming (see, for example, Section 4.3.2 in Boyd and Vandenberghe, 2004). Thus, it follows from Theorem 3.2 that computing the optimal robust assortment requires solving n linear programs, one for each revenue-ordered assortment.

When the uncertainty set $\mathcal{V}$ is not a polyhedron, finding the worst case revenue $\min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v})$ for each assortment S remains tractable as long as $\mathcal{V}$ is convex. In particular, the convexity of $\mathcal{V}$ allows us to approximate it by a polyhedron that is constructed iteratively by using the supporting hyperplanes of $\mathcal{V}$. Iteratively approximating a convex set by using its supporting hyperplanes is well-studied in nonlinear programming literature and we refer the reader to Section 7.6.2 in Ruszczynski (2006) for further details.

In the special case of a rectangular uncertainty set, it turns out that we do not need to solve any linear program. The robust assortment, in this case, is simply the optimal assortment associated with specific parameter values. This result is given in the following example.

Example 3.4 (Rectangular Uncertainty Set). Consider $\mathcal{V} = \prod_{i=0}^n [l_i, u_i] \subset \mathbb{R}_{++}^{n+1}$, where for each i, the interval $[l_i, u_i] \subset \mathbb{R}_{++}$ represents the range of the likely values of the preference weight associated with product i. We will show that the robust assortment $S^*(\mathcal{V})$ corresponds to the optimal assortment under the parameter vector $\boldsymbol{h} := (u_0, l_1, l_2, \ldots, l_n)$, that is,

$$Z^*(\mathcal{V}) = \max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v}) = \max_{S \subseteq \mathcal{A}} f(S, (u_0, l_1, l_2, \ldots, l_n)) = f(S_{\boldsymbol{h}}^*, \boldsymbol{h}) \;.$$

Since $\boldsymbol{h} \in \mathcal{V}$, it follows immediately that $Z^*(\mathcal{V}) \leq \max_{S \subseteq \mathcal{A}} f(S, \boldsymbol{h}) = f(S_{\boldsymbol{h}}^*, \boldsymbol{h})$. Suppose that $S_{\boldsymbol{h}}^* = \{1, 2, \ldots, k\}$ for some product k. Let $q : \mathcal{V} \to \mathbb{R}_{++}$ be defined by: for all $\boldsymbol{v} \in \mathcal{V}$,

$$q(\boldsymbol{v}) := f(S_{\boldsymbol{h}}^*, \boldsymbol{v}) = \frac{\sum_{i=1}^k r_i v_i}{v_0 + \sum_{i=1}^k v_i} \;.$$

To complete the proof, it suffices to show that $\min_{\boldsymbol{v} \in \mathcal{V}} q(\boldsymbol{v}) = \min_{\boldsymbol{v} \in \mathcal{V}} f(S_{\boldsymbol{h}}^*, \boldsymbol{v}) = f(S_{\boldsymbol{h}}^*, \boldsymbol{h})$, which implies $Z^*(\mathcal{V}) \geq \min_{\boldsymbol{v} \in \mathcal{V}} f(S_{\boldsymbol{h}}^*, \boldsymbol{v}) = f(S_{\boldsymbol{h}}^*, \boldsymbol{h})$. It is easy to verify that $q(\cdot)$ is continuously differentiable and

$$\frac{\partial q}{\partial v_i}(\boldsymbol{v}) = \begin{cases} -q(\boldsymbol{v})/(v_0 + \sum_{s=1}^k v_s), & \text{if } i = 0, \\ (r_i - q(\boldsymbol{v}))/(v_0 + \sum_{s=1}^k v_s), & \text{if } i = 1, 2, \ldots, k, \end{cases}$$

8

Let $\boldsymbol{v}^* = (v_0^*, v_1^*, \ldots, v_n^*) \in \mathcal{V}$ denote a minimizer of $q(\cdot)$, that is, $q(\boldsymbol{v}^*) = \min_{\boldsymbol{v} \in \mathcal{V}} q(\boldsymbol{v}) = \min_{\boldsymbol{v} \in \mathcal{V}} f(S_{\boldsymbol{h}}^*, \boldsymbol{v})$. It follows from the expression for the partial derivatives that $\frac{\partial q}{\partial v_0}(\boldsymbol{v}^*) = -q(\boldsymbol{v}^*)/(v_0^* + \sum_{s=1}^k v_s^*) < 0$, and for $i = 1, 2, \ldots, k$,

$$\frac{\partial q}{\partial v_i}(\boldsymbol{v}^*) = \frac{r_i - q(\boldsymbol{v}^*)}{v_0^* + \sum_{s=1}^k v_s^*} \geq \frac{r_i - q(\boldsymbol{h})}{v_0^* + \sum_{s=1}^k v_s^*} = \frac{r_i - f(S_{\boldsymbol{h}}^*, \boldsymbol{h})}{v_0^* + \sum_{s=1}^k v_s^*} > 0 \ ,$$

where the first inequality follows from the fact that $q(\boldsymbol{h}) \geq q(\boldsymbol{v}^*)$, and the last inequality follows from Theorem 3.2 because $i \in S_{\boldsymbol{h}}^*$, and thus, $r_i > f(S_{\boldsymbol{h}}^*, \boldsymbol{h})$. Since the gradient of $q(\cdot)$ at the minimizer $\boldsymbol{v}^*$ is nonzero, $\boldsymbol{v}^*$ must be at the boundary of $\mathcal{V}$. From the sign of the partial derivatives, we must have $v_0^* = u_0$ and $v_i^* = l_i$ for $i = 1, 2, \ldots, k$. So, $\min_{\boldsymbol{v} \in \mathcal{V}} q(\boldsymbol{v}) = q(\boldsymbol{h}) = f(S_{\boldsymbol{h}}^*, \boldsymbol{h})$, which is the desired result.

Intuitively, Example 3.4 shows that, to protect against the worst case revenue, we should target customers who are least likely to purchase our products, corresponding to those customers with the lowest product preference weights (that is, $v_i^* = l_i$ for $i = 1, \ldots n$), and highest no purchase weight ($v_0^* = u_0$). Of course, when $\mathcal{V}$ is a general set, it is not possible to tell a priori which model parameters to target so as to protect against the worst case revenue.

3.1   Comparative Statistics and Operational Insights

In this section, we analyze how the robust assortment changes with the uncertainty set and the revenue of each product. It turns out that to protect against larger uncertainty in the model parameters, we should offer a larger assortment. This result is stated in the next corollary, whose proof follows immediately from the structural property of the robust assortment in Theorem 3.2.

Corollary 3.5 (Larger Uncertainty Implies Larger Robust Assortment). *For any $\mathcal{V} \subseteq \mathcal{V}' \subseteq \mathbb{R}_{++}^n$,*

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

Proof. Since $\mathcal{V} \subseteq \mathcal{V}'$, $\min_{\boldsymbol{v} \in \mathcal{V}'} f(S, \boldsymbol{v}) \leq \min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v})$ for any assortment S, which implies that $Z^*(\mathcal{V}') \leq Z^*(\mathcal{V})$ by definition. Since $Z^*(\mathcal{V}') \leq Z^*(\mathcal{V})$, we also obtain $S^*(\mathcal{V}) = \{i : r_i > Z^*(\mathcal{V})\} \subseteq \{i : r_i > Z^*(\mathcal{V}')\} = S^*(\mathcal{V}')$ by Theorem 3.2. $\square$

The above result tells us that as the uncertainty about the model parameters increases, the robust assortment gets larger. Thus, a large product variety serves as a safety buffer against large uncertainty in the model parameters. Besides maximizing revenue, some firms are interested in maximizing their sales volume, and this motivates firms to offer large assortments to reduce the likelihood that the customers leave without purchasing anything. For example, Phillips (2011) describes a model for an airline to balance the potentially conflicting objectives of maximizing revenue and maximizing sales volume. Corollary 3.5 points out that large uncertainty in the choice model parameters forms another motivation for firms to offer large assortments.

It turns out that the robust assortment corresponds to the largest optimal assortment among $\{S_{\boldsymbol{v}}^* : \boldsymbol{v} \in \mathcal{V}\}$, as shown in the following theorem.

Theorem 3.6 (Largest Optimal Assortment is Robust). *For any $\mathcal{V} \subseteq \mathbb{R}^n_{++}$, $S^*(\mathcal{V}) = \cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}}$.*

Proof. It follows from Corollary 3.5 that $S^*_{\boldsymbol{v}} = S^*(\{\boldsymbol{v}\}) \subseteq S^*(\mathcal{V})$ for all $\boldsymbol{v} \in \mathcal{V}$ because $\{\boldsymbol{v}\} \subseteq \mathcal{V}$. Therefore, $\cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}} \subseteq S^*(\mathcal{V})$. To complete the proof, we will now show that $S^*(\mathcal{V}) = \cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}}$. Suppose on the contrary that $\cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}} \subset S^*(\mathcal{V})$. Let $\bar{\boldsymbol{v}}$ denote the optimal solution to the problem $\min_{\boldsymbol{v} \in \mathcal{V}} f(S^*(\mathcal{V}), \boldsymbol{v})$, which exists by compactness of $\mathcal{V}$. Since $\cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}} \subset S^*(\mathcal{V})$, we also have $S^*_{\bar{\boldsymbol{v}}} \subset S^*(\mathcal{V})$, which implies that there exist products $\{i_1, i_2, \ldots, i_K\}$ with $K \geq 1$ such that $S^*(\mathcal{V}) = S^*_{\bar{\boldsymbol{v}}} \cup \{i_1, i_2 \ldots, i_K\}$ and $S^*_{\bar{\boldsymbol{v}}} \cap \{i_1, i_2 \ldots, i_K\} = \varnothing$. Without loss of generality, we assume that the products $\{i_1, i_2, \ldots, i_K\}$ are ordered such that $r_{i_1} \geq r_{i_2} \geq \cdots \geq r_{i_K}$.

We show by induction on k that $r_{i_k} > f(S^*_{\bar{\boldsymbol{v}}} \cup \{i_1, \ldots, i_k\}, \bar{\boldsymbol{v}})$ for all $k = 1, \ldots, K$. The result is true for $k = K$ because $i_K \in S^*(\mathcal{V})$, and Theorem 3.2 implies that $r_{i_K} > Z^*(\mathcal{V}) = f(S^*(\mathcal{V}), \bar{\boldsymbol{v}}) = f(S^*_{\bar{\boldsymbol{v}}} \cup \{i_1, i_2 \ldots, i_K\}, \bar{\boldsymbol{v}})$. Assuming that the result holds for $k + 1$, we have $f(S^*_{\bar{\boldsymbol{v}}} \cup \{i_1, \ldots, i_k, i_{k+1}\}, \bar{\boldsymbol{v}}) < r_{i_{k+1}}$. Applying statements (a) and (c) in Lemma 3.1 with $A = S^*_{\bar{\boldsymbol{v}}} \cup \{i_1, i_2 \ldots, i_k\}$, we have $f(S^*_{\bar{\boldsymbol{v}}} \cup \{i_1, \ldots, i_k\}, \bar{\boldsymbol{v}}) < r_{i_{k+1}} \leq r_{i_k}$, and the result also holds for k. This completes the induction.

Using the above result with $k = 1$, we obtain $r_{i_1} > f(S^*_{\bar{\boldsymbol{v}}} \cup \{i_1\}, \bar{\boldsymbol{v}})$, and it follows from the statements (b) and (c) in Lemma 3.1 that $f(S^*_{\bar{\boldsymbol{v}}} \cup \{i_1\}, \bar{\boldsymbol{v}}) > f(S^*_{\bar{\boldsymbol{v}}}, \bar{\boldsymbol{v}})$. This contradicts the optimality of $S^*_{\bar{\boldsymbol{v}}}$ for the problem $\max_{S \subseteq \mathcal{A}} f(S, \bar{\boldsymbol{v}})$. Therefore, it must be the case that $S^*(\mathcal{V}) = \cup_{\boldsymbol{v} \in \mathcal{V}} S^*_{\boldsymbol{v}}$. $\square$

The next theorem shows that as the revenue of each product increases by the same additive increment, the robust optimal assortment becomes larger. To facilitate our exposition, for any $\delta \geq 0$ and $\mathcal{V} \subset \mathbb{R}^{n+1}_{++}$, let $Z^*_\delta(\mathcal{V})$ denote the objective value of the robust assortment where the revenue of each product is increased by δ; that is

$$Z^*_\delta(\mathcal{V}) = \max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}} \sum_{i \in S} (r_i + \delta) \, \phi_i(S, \boldsymbol{v}) \ ,$$

and let $S^*_\delta(\mathcal{V})$ denote the corresponding optimal assortment, where we break ties by choosing the assortment with the smallest cardinality. Note that $Z^*_0(\mathcal{V}) = Z^*(\mathcal{V})$ and $S^*_0(\mathcal{V}) = S^*(\mathcal{V})$, where $Z^*(\mathcal{V})$ and $S^*(\mathcal{V})$ are defined earlier.

Theorem 3.7 (Additive Incremental Revenues Lead to Larger Robust Assortment). *For any $\delta \geq 0$, $S^*(\mathcal{V}) \subseteq S^*_\delta(\mathcal{V})$.*

Proof. For any assortment $S \subseteq \mathcal{A}$ and $\boldsymbol{v} \in \mathcal{V}$, we have that

$$\sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) \leq \sum_{i \in S} (r_i + \delta) \, \phi_i(S, \boldsymbol{v}) \leq \sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) + \delta \ ,$$

where we use the fact that $\sum_{i \in S} \phi_i(S, \boldsymbol{v}) \leq 1$. Since the chain of inequalities above holds for any $\boldsymbol{v} \in \mathcal{V}$, taking minimum over all $\boldsymbol{v} \in \mathcal{V}$, we obtain

$$\min_{\boldsymbol{v} \in \mathcal{V}} \left\{ \sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) \right\} \leq \min_{\boldsymbol{v} \in \mathcal{V}} \left\{ \sum_{i \in S} (r_i + \delta) \, \phi_i(S, \boldsymbol{v}) \right\} \leq \min_{\boldsymbol{v} \in \mathcal{V}} \left\{ \sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) \right\} + \delta \ .$$

Similarly, since the chain of inequalities above holds for any $S \subseteq \mathcal{A}$, taking maximum over all $S \subseteq \mathcal{A}$, it follows that

$$\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}} \left\{ \sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) \right\} \leq \max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}} \left\{ \sum_{i \in S} (r_i + \delta) \, \phi_i(S, \boldsymbol{v}) \right\} \leq \max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}} \left\{ \sum_{i \in S} r_i \, \phi_i(S, \boldsymbol{v}) \right\} + \delta \ .$$

Therefore, we have $Z^*(\mathcal{V}) \leq Z_\delta^*(\mathcal{V}) \leq \delta + Z^*(\mathcal{V})$, which implies that

$$S^*(\mathcal{V}) = \{i : r_i > Z^*(\mathcal{V})\} = \{i : r_i + \delta > Z^*(\mathcal{V}) + \delta\} \subseteq \{i : r_i + \delta > Z_\delta^*(\mathcal{V})\} = S_\delta^*(\mathcal{V}) \ ,$$

where the first and last equalities follow from Theorem 3.2. $\qquad\square$

Intuitively, Theorem 3.7 suggests that as the opportunity cost of not making a sale gets larger, the robust assortment also gets larger. This theorem can be of independent interest, but we also use this result in Section 4 to compare the assortments that are offered by our dynamic model at different time periods and at different remaining inventory levels.

Note that the revenue of each product in Theorem 3.7 is increased by the same amount δ. This result is adequate to compare the assortments that are offered by our dynamic model at different time periods and at different inventory levels. However, one can make slight generalizations in Theorem 3.7 to allow different increases in the revenues of different products. For example, if the revenue of product i is increased by δ_i and the increases in the product revenues satisfy $\sum_{\ell \in S} \delta_\ell \, \phi_\ell(S, \boldsymbol{v}) \leq \delta_i$ for all $i \in \mathcal{A}$, for any assortment $S \subseteq \mathcal{A}$ and $\boldsymbol{v} \in \mathcal{V}$, then we can follow the same argument in the proof of Theorem 3.7 to show that the optimal robust assortment is larger when the revenue of each product i is increased by δ_i. The assumption that $\sum_{\ell \in S} \delta_\ell \, \phi_\ell(S, \boldsymbol{v}) \leq \delta_i$ for all $i \in \mathcal{A}$ holds when the preference weight associated with the no purchase option is large enough for all $\boldsymbol{v} \in \mathcal{V}$.

3.2 Constraining the Size of Assortments: A Rectangular Uncertainty Set

In many cases, it may not be possible to include an arbitrarily large number of products in the assortment offered to the customers. This is particularly the case when the offered assortment is limited by the space available in the display window or the size of a web page. In this section, we extend our robust logit formulation to allow for a constraint on the size of assortments that we offer, corresponding to the following optimization problem:

$$Y^*(\mathcal{V}) = \max_{S \subseteq \mathcal{A}: |S| \leq K} \left\{ \min_{\boldsymbol{v} \in \mathcal{V}} \ f(S, \boldsymbol{v}) \right\} \ , \qquad \text{(SIZE-CONSTRAINED ROBUST)}$$

where K denotes the largest allowable size for an assortment. In general, the SIZE-CONSTRAINED ROBUST problem appears to be intractable. However, the next theorem shows that if the uncertainty set $\mathcal{V}$ has the rectangular form $\prod_{i=0}^{n}[l_i, u_i] \subseteq \mathbb{R}_{++}^{n+1}$ considered in Example 3.4, then the SIZE-CONSTRAINED ROBUST problem reduces to an assortment optimization problem with known preference weights, which can be solved efficiently.

Theorem 3.8 (Robust Size-Constrained Assortment under Rectangular Uncertainty Sets). *For any* $\mathcal{V} = \prod_{i=0}^{n}[l_i, u_i] \subset \mathbb{R}_{++}^{n+1}$,

$$Y^*(\mathcal{V}) = \max_{S \subseteq \mathcal{A}: |S| \leq K} \frac{\sum_{i \in S} l_i \, r_i}{u_0 + \sum_{i \in S} l_i} = \max_{S \subseteq \mathcal{A}: |S| \leq K} f(S, (u_0, l_1, l_2, \ldots, l_n)) \ .$$

Proof. By definition of $Y^*(\mathcal{V})$, we have

$$
\begin{aligned}
Y^*(\mathcal{V}) &= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \min_{\boldsymbol{v} \in \mathcal{V}} f(S, \boldsymbol{v}) \geq \lambda\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \forall \boldsymbol{v} \in \mathcal{V}, \ f(S, \boldsymbol{v}) \geq \lambda\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \forall \boldsymbol{v} \in \mathcal{V}, \ \sum_{i \in S} v_i \, r_i \geq \lambda\left(v_0 + \sum_{i \in S} v_i\right)\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \forall \boldsymbol{v} \in \mathcal{V}, \ \frac{1}{v_0} \sum_{i \in S} v_i \, (r_i - \lambda) \geq \lambda\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \min_{\boldsymbol{v} \in \mathcal{V}} \frac{1}{v_0} \sum_{i \in S} v_i \, (r_i - \lambda) \geq \lambda\right\} \\
&= \max\left\{\lambda : \max_{S:|S| \leq K} \min_{\boldsymbol{v} \in \mathcal{V}} \frac{1}{v_0} \sum_{i \in S} v_i \, (r_i - \lambda) \geq \lambda\right\} \ ,
\end{aligned}
$$

where the third equality follows from the definition of $f(S, \boldsymbol{v})$. For any $x \in \mathbb{R}$, let $x^+ = \max\{x, 0\}$ denote the positive part of x. For all λ, we have

$$
\begin{aligned}
\max_{S:|S| \leq K} \min_{\boldsymbol{v} \in \mathcal{V}} \frac{1}{v_0} \sum_{i \in S} v_i(r_i - \lambda) &= \max_{S:|S| \leq K} \min_{\boldsymbol{v} \in \mathcal{V}} \frac{1}{v_0} \sum_{i \in S} v_i(r_i - \lambda)^+ \\
&= \max_{S:|S| \leq K} \frac{1}{u_0} \sum_{i \in S} l_i(r_i - \lambda)^+ = \max_{S:|S| \leq K} \frac{1}{u_0} \sum_{i \in S} l_i(r_i - \lambda) \ ,
\end{aligned}
$$

where the first equality follows from the fact that the optimal solution to the optimization problem $\max_{S:|S| \leq K} \min_{\boldsymbol{v} \in \mathcal{V}} \frac{1}{v_0} \sum_{i \in S} v_i \, (r_i - \lambda)$ never includes any product for which $r_i - \lambda$ is negative, the second equality follows from the fact that $\mathcal{V} = \prod_{i=0}^{n}[l_i, u_i]$ and the final equality follows from the fact that the optimal solution to the optimization problem $\max_{S:|S| \leq K} \frac{1}{u_0} \sum_{i \in S} l_i \, (r_i - \lambda)$ never includes any product for which $r_i - \lambda$ is negative. Putting everything together, we have

$$
\begin{aligned}
Y^*(\mathcal{V}) &= \max\left\{\lambda : \max_{S:|S| \leq K} \min_{\boldsymbol{v} \in \mathcal{V}} \frac{1}{v_0} \sum_{i \in S} v_i \, (r_i - \lambda) \geq \lambda\right\} = \max\left\{\lambda : \max_{S:|S| \leq K} \frac{1}{u_0} \sum_{i \in S} l_i \, (r_i - \lambda) \geq \lambda\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \sum_{i \in S} l_i \, (r_i - \lambda) \geq u_0 \, \lambda\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } \sum_{i \in S} l_i \, r_i \geq \lambda\left(u_0 + \sum_{i \in S} l_i\right)\right\} \\
&= \max\left\{\lambda : \exists S \text{ s.t. } |S| \leq K \text{ and } f(S, (u_0, l_1, l_2, \dots, l_n)) \geq \lambda\right\} \\
&= \max_{S \subseteq \mathcal{A}:|S| \leq K} f(S, (u_0, l_1, l_2, \dots, l_n)) \ ,
\end{aligned}
$$

where the chain of equalities above follow from an argument similar to the one that we use at the beginning of the proof and we have the desired result. $\qquad\square$

Theorem 3.8 shows that if the uncertainty set has a rectangular form, then we can convert the SIZE-CONSTRAINED ROBUST problem into a size-constrained assortment optimization problem with *known*

preference weights. Building upon the work of Megiddo (1979), Rusmevichientong et al. (2010) show that the optimal solution for the size-constrained assortment optimization problem with known preference weights can be computed in $O(n^2)$ time. Therefore, the SIZE-CONSTRAINED ROBUST problem under rectangular uncertainty sets can also be solved in $O(n^2)$ time.

3.3   Constructing an Uncertainty Set: Balancing Average Revenue and Variability

Previously, we assume that the uncertainty set $\mathcal{V}$ is given in advance, and we focus on establishing properties of the corresponding optimal robust assortment $S^*(\mathcal{V})$. In this section, we discuss the construction of an uncertainty set, and the inherent tradeoffs in such a construction between worst case revenue and average performance. A common approach in the robust optimization literature is to describe an uncertainty set as an ellipsoid in $\mathbb{R}_{++}^{n+1}$, that is,

$$\mathcal{V}_{\bar{\boldsymbol{v}},\epsilon} = \left\{ \boldsymbol{v} \in \mathbb{R}_{++}^{n+1} : \|\boldsymbol{v} - \bar{\boldsymbol{v}}\| \leq \epsilon \right\} ,$$

where $\|\cdot\|$ denotes an appropriately chosen norm, and $\bar{\boldsymbol{v}} \in \mathbb{R}_{++}^{n+1}$ and $\epsilon \in \mathbb{R}_+$ denote the parameters associated with the uncertainty set (see, for example, Bertsimas and Sim, 2003; Bertsimas et al., 2004). The vector $\bar{\boldsymbol{v}}$ corresponds to the center of the ellipsoid, and can be interpreted as the most likely parameter vector associated with the underlying multinomial logit choice model.

The parameter $\epsilon \in \mathbb{R}_+$ reflects our degree of uncertainty in the model parameters. Intuitively, larger values of ϵ provide larger uncertainty sets, resulting in conservative assortments that protect well against the worst case expected revenue, but may lack in average performance. On the other hand, smaller values of ϵ provide smaller uncertainty sets, which are likely to yield aggressive assortments that may have high average performance, but these assortments may not do well in terms of the worst case expected revenue. We can thus view ϵ as a knob that controls the tradeoff between the worst case expected revenue and average performance. Depending on the preferences of the decision maker, we can adjust ϵ to achieve a desired tradeoff. These intuitive expectations are indeed verified in our numerical experiments in Section 5. By using larger values of ϵ, we obtain assortments that do quite well in terms of the worst case expected revenues and the variability of revenues. In particular, our robust assortments can provide 10% improvement in the worst case expected revenue when compared with methods that assume precise knowledge of the model parameters. On the other hand, if average performance is the performance measure of interest, then we are able to use our robust approach with a smaller value of ϵ, and the robust assortments that we obtain in this fashion perform well in terms of average performance. Additional details are given in Section 5.

Let $\mathsf{Ext}\,(\mathcal{V}_{\bar{\boldsymbol{v}},\varepsilon})$ denote the set of extreme points of the uncertainty set $\mathcal{V}_{\bar{\boldsymbol{v}},\epsilon}$. The following lemma will be useful momentarily, and it shows that it suffices to consider only the extreme points of the uncertainty set when computing the optimal robust assortment.

Lemma 3.9. *For any $\bar{\boldsymbol{v}} \in \mathbb{R}_{++}^{n+1}$ and $\epsilon \in \mathbb{R}_+$, $Z^*(\mathcal{V}_{\bar{\boldsymbol{v}},\varepsilon}) = Z^*(\mathsf{Ext}\,(\mathcal{V}_{\bar{\boldsymbol{v}},\varepsilon}))$ and $S^*(\mathcal{V}_{\bar{\boldsymbol{v}},\varepsilon}) = S^*(\mathsf{Ext}\,(\mathcal{V}_{\bar{\boldsymbol{v}},\varepsilon}))$.*

Proof. For any assortment $S \subseteq \mathcal{A}$, the revenue function $f(S, \boldsymbol{v})$ is a quasi-linear function in $\boldsymbol{v}$ because it is the ratio between two linear functions in $\boldsymbol{v}$. It follows from standard results in convex optimization

13

(see, for example, Boyd and Vandenberghe, 2004) that

$$\min_{\boldsymbol{v} \in \mathcal{V}_{\bar{\boldsymbol{v}}, \varepsilon}} f(S, \boldsymbol{v}) = \min_{\boldsymbol{v} \in \mathsf{Ext}(\mathcal{V}_{\bar{\boldsymbol{v}}, \varepsilon})} f(S, \boldsymbol{v}) \,.$$

Since S is arbitrary, taking the maximum of both sides of the equality above over $S \subseteq \mathcal{A}$, we obtain $Z^*(\mathcal{V}_{\bar{\boldsymbol{v}}, \varepsilon}) = Z^*(\mathsf{Ext}(\mathcal{V}_{\bar{\boldsymbol{v}}, \varepsilon}))$. The equality between optimal assortments follows from Theorem 3.2. □

According to Lemma 3.9, it suffices to focus on the extreme points of the uncertainty set when computing the optimal robust assortment. This result can simplify the computation of the optimal robust revenue, as shown in the following example.

Example 3.10 (Uncertainty Sets for a Heterogeneous Market)**.** Suppose the true parameter vector for the underlying choice model can be one of G vectors, denoted by $\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G$. This may correspond to a setting where a firm operates in a market with G heterogeneous customer types, and for each $g = 1, \ldots, G$, the preference weights associated customer type g is defined by a vector $\boldsymbol{v}^g \in \mathbb{R}_{++}^{n+1}$. Although the preference profile of each customer type is known by the firm in advance, the firm may not know the type of each arriving customer. Let $\theta^1, \ldots, \theta^G$ denote the proportion of each customer type, with $\theta^g \geq 0$ for all g and $\sum_{g=1}^{G} \theta^g = 1$.

Finding the assortment that maximizes the average revenue across G customer types corresponds to the optimization problem $\max_{S \subseteq \mathcal{A}} \sum_{g=1}^{G} \theta^g f(S, \boldsymbol{v}^g)$. Bront et al. (2009) and Rusmevichientong et al. (2010) show that this optimization problem is NP-hard, even in the simplest setting where $G = 2$.

An alternative to maximizing the average revenue is to find the assortment that maximizes the worst case revenue, where we construct the uncertainty set by taking an estimate of the proportion of each customer type into consideration. This approach would not be equivalent to maximizing the average revenue $\sum_{g=1}^{G} \theta^g f(S, \boldsymbol{v}^g)$ across G customer types, but it provides an effective way to find an assortment to offer, without having to solve an NP-hard optimization problem. Assume that the vectors $\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G$ are linearly independent, and let $\mathsf{Span}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$ denote the subspace spanned by these vectors. For any $\boldsymbol{v} \in \mathsf{Span}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$ such that $\boldsymbol{v} = \sum_{g=1}^{G} \lambda^g \boldsymbol{v}^g$, let $\|\boldsymbol{v}\| = \max_{g=1,\ldots,G} |\lambda^g|$. It is easy to verify that $\|\cdot\|$ defines a norm on this subspace.[2]

To construct the uncertainty set, let $\mathsf{Conv}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$ denote the convex hull of $\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G$. Let $\bar{\boldsymbol{v}} = \sum_{g=1}^{G} \theta^g \boldsymbol{v}^g \in \mathsf{Conv}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$ be the proportionally weighted average of the preference vectors for all G customer types, and consider the ellipsoidal uncertainty set

$$\begin{aligned}
\mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon} &= \left\{ \boldsymbol{v} \in \mathsf{Conv}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right) : \|\boldsymbol{v} - \bar{\boldsymbol{v}}\| \leq \epsilon \right\} \\
&= \left\{ \sum_{g=1}^{G} \lambda^g \boldsymbol{v}^g : \lambda^g \geq 0 \text{ for all } g \text{ and } \sum_{g=1}^{G} \lambda^g = 1 \text{ and } \max_{g=1,\ldots,G} |\lambda^g - \theta^g| \leq \epsilon \right\} . \quad (\text{ELLIPSOIDAL})
\end{aligned}$$

For $\epsilon = 1$, we have $\mathcal{V}_{\bar{\boldsymbol{v}}, 1} = \mathsf{Conv}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$ and $\mathsf{Ext}\left(\mathcal{V}_{\bar{\boldsymbol{v}}, 1}\right) = \left\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right\}$. It follows from Lemma 3.9 that this uncertainty set is equivalent to the finite set $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\}$. The choice of

---

[2]The assumption of linear independence among $\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G$ is not essential. If the vectors are linearly dependent, then we choose any basis $\boldsymbol{b}^1, \ldots, \boldsymbol{b}^d$ with $d \leq G$ for the subspace $\mathsf{Span}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$. Then, for any $\boldsymbol{v} \in \mathsf{Span}\left(\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^G\right)$ such that $\boldsymbol{v} = \sum_{g=1}^{d} \lambda^g \boldsymbol{b}^g$, we simply define $\|\boldsymbol{v}\| = \max_{g=1,\ldots,d} |\lambda^g|$.

$\epsilon = 1$ essentially assumes that we do not have a good estimate of how likely each customer type is, leading to an uncertainty set that ignores the relative weights of each customer type. As our numerical experiments in Section 5 indicates, this choice of the uncertainty set leads to an assortment with a high worst case expected revenue and low variability in revenues, but the average performance can be somewhat unsatisfactory. As we decrease the value of ϵ, we can control the tradeoff between the worst case revenue and average performance. In fact, for appropriately chosen ϵ, the robust assortment $S^*(\mathcal{V}_{\bar{\boldsymbol{v}},\epsilon})$ yields quite high average performance. Therefore, our uncertainty sets allow us to tune the degree of conservatism we get from our robust approach.

4. A Dynamic Model: Robust Capacity Allocation

In this section, we extend the structural properties and insights from the previous section to a dynamic setting, where we now consider the problem of allocating C units of initial resources to arriving customers over time periods $1, 2, \ldots, T$. As before, our products are indexed by $\mathcal{A} = \{1, 2, \ldots, n\}$, and the revenues are ordered such that $r_1 \geq r_2 \geq \cdots \geq r_n > 0$. At each time period t, we need to choose an assortment of products to offer to an arriving customer, who chooses a product from the assortment according to the multinomial logit model with unknown parameters $\boldsymbol{v}_t \in \mathbb{R}_{++}^{n+1}$. If there is a purchase, then one unit of resource capacity is consumed. For notational brevity, we assume that there is exactly one customer arrival in each period, but it is straightforward to extend to the case where there may be no customer arrivals in each period. Our goal is to find a policy for choosing an assortment in each period that protects against the worst case total expected revenue over the selling horizon, where the worst case is taken over all model parameters that could have been chosen in response to the assortments that we have offered to the customers. This setting is the robust analogue of the single leg airline seat allocation problem considered by Talluri and van Ryzin (2004), where the resource corresponds to the seat availability on the flight leg and the products correspond to different fare classes.

We capture the uncertainty in the model parameters $\boldsymbol{v}_t$, at time period t, through a compact uncertainty set $\mathcal{V}_t \subseteq \mathbb{R}_{++}^{n+1}$, which represents the set of likely parameter values that can occur. We allow the uncertainty set $\mathcal{V}_t$ to change over time to model situations where the set of possible parameter values may shift due to a change in the customer profile as the selling season progresses. In our robust formulation, it is conceptually helpful to view the choice parameters in each period as being chosen by an "adversary", who wishes to minimize the total expected future revenue. Thus, once we choose an assortment in period t, the model parameter $\boldsymbol{v}_t$ is chosen by the adversary to generate the smallest total expected revenue in the remaining portion of the selling horizon. When we move to the next time period $t + 1$, we choose a new assortment and the adversary then chooses a new model parameter in response to this assortment. Thus, our objective is to find a policy π that solves the problem

$$\max_{\pi} \ \min_{\psi} \ \mathbb{E}_{\pi,\psi}\left[\sum_{t=1}^{T} f(S_t^{\pi,\psi}, \boldsymbol{v}_t^{\psi}(S_t^{\pi,\psi})) \right],$$

where π is the policy that chooses the assortments, and ψ is the policy of the adversary that chooses the model parameters. The random variables $S_1^{\pi,\psi}, S_2^{\pi,\psi}, \ldots, S_T^{\pi,\psi}$ are the assortments generated under the policies π and ψ, and $\boldsymbol{v}_t^{\psi}(S_t^{\pi,\psi})$ is the model parameters chosen by the policy ψ at time period t in

15

response to the assortment $S_t^{\pi,\psi}$.

For each time period t and $x \geq 0$, let $J_t(x)$ denote the maximum worst case total expected revenue that can be obtained over the time periods $t, t+1, \ldots, T$ given that we have x units of capacity remaining at the beginning of time period t. Iyengar (2005) shows that the value function $J_t(\cdot)$ satisfies the Bellman equation

$$J_t(x) = \max_{S_t \subseteq \mathcal{A}} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, \boldsymbol{v}_t)(r_i + J_{t+1}(x-1)) + \left(1 - \sum_{i \in S_t} \phi_i(S_t, \boldsymbol{v}_t)\right) J_{t+1}(x) \right\}$$

$$= \max_{S_t \subseteq \mathcal{A}} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x)) \right\} + J_{t+1}(x), \qquad \text{(DYNAMIC ROBUST)}$$

where $\Delta J_{t+1}(x) = J_{t+1}(x) - J_{t+1}(x-1)$ denotes the marginal value of capacity. The boundary conditions are $J_t(0) = 0$ for all $t = 1, 2, \ldots, T$ and $J_{T+1}(x) = 0$ for all $x = 0, 1, 2, \ldots, C$. Let $S_t^*(x)$ denote the assortment that maximizes the right hand side of the Bellman equation above, corresponding to the optimal assortment to offer when the remaining capacity at the beginning of time period t is x.

4.1 Structural Properties

The following theorem shows that the value functions are concave in the remaining capacity and the marginal value of capacity is decreasing over time. The proof technique for this result is quite standard (see, for example, Talluri and van Ryzin, 2004), and the details are given in Appendix A.

Theorem 4.1 (Monotonicity). *For any $x \in \{0, 1, \ldots, C\}$ and $t = 1, \ldots, T$,*

$$\Delta J_t(x+1) \leq \Delta J_t(x) \quad and \quad \Delta J_{t+1}(x) \leq \Delta J_t(x) .$$

The next theorem provides a complete characterization of the optimal assortment at each time period, showing that it suffices to consider only revenue-ordered assortments. This result enables us to compute the value functions and the optimal policy efficiently. Talluri and van Ryzin (2004) previously proved a similar result when there is no uncertainty in the parameters of the underlying multinomial logit model. Surprisingly, Theorem 4.2 shows that their result can be extended even when we have uncertainty in the model parameters. Moreover, this result has an important managerial implication by showing that the optimal assortment in each period consists of the set of products whose revenues are larger than the opportunity cost, as measured by $J_t(x) - J_{t+1}(x-1)$.

Theorem 4.2 (Revenue-Ordered Assortments are Optimal). *For any x and t,*

$$S_t^*(x) = \{i : r_i > J_t(x) - J_{t+1}(x-1)\} .$$

Proof. By definition of $J_t(x)$, we have that

$$J_t(x) - J_{t+1}(x) = \max_{S_t \subseteq \mathcal{A}} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x)) \right\} ,$$

16

and $S_t^*(x)$ is the optimal solution associated with the above optimization problem, where we break ties by choosing the assortment with the smallest cardinality. Note that it is not optimal to use a product i in the solution to the above problem for which $r_i < \Delta J_{t+1}(x)$. Thus, the optimization problem associated with $S_t^*(x)$ is the same as the robust assortment optimization problem considered in Section 3, where the revenue of each product i is now given by $r_i - \Delta J_{t+1}(x)$. In this case, it follows from Theorem 3.2 that

$$S_t^*(x) = \{i : r_i - \Delta J_{t+1}(x) > J_t(x) - J_{t+1}(x)\} = \{i : r_i > J_t(x) - J_{t+1}(x-1)\} \ ,$$

where the last inequality follows from the fact that $J_t(x) - J_{t+1}(x) + \Delta J_{t+1}(x) = J_t(x) - J_{t+1}(x-1)$. $\square$

We note that Theorem 4.2 also shows that revenue-ordered assortments are *complete efficient sets*, and the *nesting-by-fare-order* property continues to hold for the robust setting. This generalizes the result by Talluri and van Ryzin (2004), who establish these properties in the setting where the choice model parameters are fixed and known in advance. In addition, as noted by Talluri and van Ryzin (2004), the nesting-by-fare-order property enables us to implement the optimal policy using the classical nested protection levels or bid price control, broadening the potential application of our approach.

4.2 Comparative Statistics and Operational Insights

Intuitively, we expect that as we have fewer units of remaining inventory, all else being equal, the opportunity cost of a unit of resource would increase and we would offer smaller assortments to the customers, focusing on products with higher revenues. Similarly, as we approach the end of the selling horizon, we have fewer opportunities to sell the remaining capacity, and we should offer larger assortments to increase the probability of purchase and liquidate the inventory. The next theorem verifies these intuitions.

Theorem 4.3 (Larger Remaining Capacity Leads to Larger Assortment). *For any x and t,*

$$S_t^*(x) \subseteq S_t^*(x+1) \ ,$$

and if, in addition $\mathcal{V}_t \subseteq \mathcal{V}_{t+1}$, then

$$S_t^*(x) \subseteq S_{t+1}^*(x) \ .$$

Proof. It follows from the definition of $S_t^*(x+1)$ that

$$S_t^*(x+1) = \arg\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S} \phi_i(S, \boldsymbol{v}_t) \left(r_i - \Delta J_{t+1}(x+1) \right) \right\}$$

$$= \arg\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S} \phi_i(S, \boldsymbol{v}_t) \left(r_i - \Delta J_{t+1}(x) + [\Delta J_{t+1}(x) - \Delta J_{t+1}(x+1)] \right) \right\} \ .$$

By Theorem 4.1, the expression $\Delta J_{t+1}(x) - \Delta J_{t+1}(x+1)$ in the square brackets above is nonnegative. On the other hand, noting the definition of $S_t^*(x)$, we can compute $S_t^*(x)$ simply by dropping the expression in the square brackets above and solving the resulting optimization problem. In this case, it follows from Theorem 3.7, with $\delta = \Delta J_{t+1}(x) - \Delta J_{t+1}(x+1)$, that $S_t^*(x) \subseteq S_t^*(x+1)$.

17

Now, suppose that $\mathcal{V}_t \subseteq \mathcal{V}_{t+1}$. By definition of $S_{t+1}^*(x)$, we have that

$$S_{t+1}^*(x) = \arg\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}_{t+1}} \left\{ \sum_{i \in S} \phi_i(S, \boldsymbol{v}) \left(r_i - \Delta J_{t+2}(x) \right) \right\} .$$

Let $Q \subseteq \{1, 2, \ldots n\}$ be defined by

$$Q = \arg\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}_t} \left\{ \sum_{i \in S} \phi_i(S, \boldsymbol{v}) \left(r_i - \Delta J_{t+2}(x) \right) \right\} .$$

Since $\mathcal{V}_t \subseteq \mathcal{V}_{t+1}$, it follows from Theorem 3.6 that $Q \subseteq S_{t+1}^*(x)$. On the other hand, it follows from the definition of Q that

$$Q = \arg\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}_t} \left\{ \sum_{i \in S} \phi_i(S, \boldsymbol{v}) \left(r_i - \Delta J_{t+1}(x) + [\Delta J_{t+1}(x) - \Delta J_{t+2}(x)] \right) \right\} .$$

By Theorem 4.1, we have that $\Delta J_{t+1}(x) - \Delta J_{t+2}(x) \geq 0$. Therefore, it follows from Theorem 3.7 and the definition of $S_t^*(x)$ that $S_t^*(x) \subseteq Q$, which completes the proof. $\qquad \square$

We note that the result of Theorem 4.3 assumes that consumers are not strategic about the timing of their purchases. When consumers are strategic, stockpiling or postponing of purchases may occur. We refer the reader to Liu and van Ryzin (2008) and Su (2007, 2010) for more details on models with strategic consumers.

5. Numerical Experiments

In this section, we provide numerical experiments demonstrating the benefits from the robust assortment optimization and robust capacity allocation models described in Sections 3 and 4.

5.1 Robust Assortment Optimization Problems

In this section, we work with the robust assortment optimization model described in Section 3. We begin by describing our numerical setup, followed by our numerical results.

5.1.1 Numerical Setup

We design our numerical experiments to understand how estimation errors in model parameters impact the expected revenue and to investigate when our robust approach can protect us against the uncertainty in parameter estimation. We consider Example 3.10 in Section 3.3 where a firm operates in a market with G customer types, and for $g = 1, \ldots, G$, the choices of a customer of type g are governed by the multinomial logit model with preference weights $\boldsymbol{v}^g \in \mathbb{R}_{++}^{n+1}$. We assume that the true proportion of each customer type is *unknown*, and we model the unknown proportions by a G-dimensional Dirichlet random variable $\boldsymbol{\Theta} = (\Theta^1, \ldots, \Theta^G)$ where, $\sum_{g=1}^G \Theta^g = 1$. We denote the mean vector of $\boldsymbol{\Theta}$ by $\boldsymbol{\theta} = (\theta^1, \ldots, \theta^G)$, and let the coefficients of variation for each component of $\boldsymbol{\Theta}$ be given by $(\rho^1, \ldots, \rho^G)$.

In our numerical experiments, we view the mean vector $\boldsymbol{\theta} = (\theta^1, \ldots, \theta^G)$ as our *estimate* of the proportion of each customer type, and we try to compute assortments by using only the knowledge of $\boldsymbol{\theta}$. We then evaluate the performance of the assortments under the *true unknown* proportions $\boldsymbol{\Theta}$. In this case, if we offer an assortment S to the customers, then the actual expected revenue we obtain is given by the random variable

$$\mathsf{Rev}(S, \boldsymbol{\Theta}) = \sum_{g=1}^{G} \Theta^g f(S, \boldsymbol{v}^g).$$

We observe that $\mathsf{Rev}(S, \boldsymbol{\Theta})$ is a random variable because $\boldsymbol{\Theta}$ is the Dirichlet random variable. Our numerical setup is intended to capture a situation where the unknown true proportions $\boldsymbol{\Theta}$ do not exactly match the estimated proportions $\boldsymbol{\theta}$, but they match in expectation. By varying the coefficients of variation $(\rho^1, \ldots, \rho^G)$, we control how much the unknown proportions $\boldsymbol{\Theta}$ tend to deviate from the estimated proportions $\boldsymbol{\theta}$.

Given the estimated proportions $\boldsymbol{\theta} = (\theta^1, \ldots, \theta^G)$, we evaluate two classes of assortments, MIXTURE and ROBUST. Under the mixture model, we attempt to find an assortment that maximizes the expected revenue per customer by solving the problem $\max_{S \subseteq \mathcal{A}} \sum_{g=1}^{G} \theta^g f(S, \boldsymbol{v}^g)$. Although this assortment optimization problem is NP-hard even when $G = 2$ (see Bront et al., 2009; Rusmevichientong et al., 2010), Bront et al. (2009) show that this problem can be formulated as a mixed integer linear program. Let S^{MIX} denote the assortment that we obtain by using this mixed integer linear programming formulation, where the superscript MIX emphasizes the fact that the expected revenue in this approach is obtained by mixing the expected revenue under G parameter vectors.

In the ROBUST assortment, we use the ELLIPSOIDAL uncertainty set $\mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon}$ given in Section 3.3, where

$$\mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon} = \left\{ \sum_{g=1}^{G} \lambda^g \boldsymbol{v}^g : \lambda^g \geq 0 \text{ for all } g \text{ and } \sum_{g=1}^{G} \lambda^g = 1 \text{ and } \max_{g=1,\ldots,G} |\lambda^g - \theta^g| \leq \epsilon \right\}.$$

We then solve the assortment optimization problem $\max_{S \subseteq \mathcal{A}} \min_{\boldsymbol{v} \in \mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon}} f(S, \boldsymbol{v})$ to find a robust assortment $S_{\epsilon}^{\text{ROB}}$, where the superscript ROB emphasizes that this assortment is obtained under a robust formulation. We experiment with different values for ϵ in our numerical experiments. Larger values for ϵ yield larger uncertainty sets and provide more conservative assortments that protect against larger uncertainty in proportions. Noting that the uncertainty set $\mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon}$ is a polyhedron, we can compute $\min_{\boldsymbol{v} \in \mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon}} f(S, \boldsymbol{v})$ for any $S \subseteq \mathcal{A}$ by solving a linear program as in Example 3.3. Since Theorem 3.2 shows that the robust assortment is revenue-ordered, we obtain $S_{\epsilon}^{\text{ROB}}$ by computing $\min_{\boldsymbol{v} \in \mathcal{V}_{\bar{\boldsymbol{v}}, \epsilon}} f(S, \boldsymbol{v})$ for each revenue-ordered assortment and picking the one that provides the largest objective function value. Therefore, while the assortment optimization model that provides S^{MIX} is NP-hard, solving the assortment optimization model that provides $S_{\epsilon}^{\text{ROB}}$ involves solving n linear programs.

Our goal is to compare the *distributions* of the random variables $\mathsf{Rev}(S^{\text{MIX}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(S_{\epsilon}^{\text{ROB}}, \boldsymbol{\Theta})$, which correspond to the (random) revenues obtained by the assortments S^{MIX} and $S_{\epsilon}^{\text{ROB}}$, respectively. On average, the expected revenue under the assortment S^{MIX} is higher than the expected revenue under

the assortment S_ϵ^{ROB} because

$$\mathbb{E}\left[\text{Rev}(S^{\text{MIX}},\boldsymbol{\Theta})\right] = \sum_{g=1}^{G}\mathbb{E}\left[\Theta^g\right]f(S^{\text{MIX}},\boldsymbol{v}^g) = \sum_{g=1}^{G}\theta^g f(S^{\text{MIX}},\boldsymbol{v}^g) = \max_{S\subseteq\mathcal{A}}\sum_{g=1}^{G}\theta^g f(S,\boldsymbol{v}^g)$$

$$\geq \sum_{g=1}^{G}\theta^g f(S_\epsilon^{\text{ROB}},\boldsymbol{v}^g) = \mathbb{E}\left[\text{Rev}(S_\epsilon^{\text{ROB}},\boldsymbol{\Theta})\right] \ .$$

Although our robust formulation yields a smaller average performance, we show that the random variable $\text{Rev}(S_\epsilon^{\text{ROB}},\boldsymbol{\Theta})$ exhibits significantly smaller variability and higher left tails than $\text{Rev}(S^{\text{MIX}},\boldsymbol{\Theta})$. Moreover, by adjusting the parameter ϵ of the uncertainty set, we can reach a desired balance between the worst case and average performance of the assortment S_ϵ^{ROB}.

Problem Classes: To test the effectiveness of both MIXTURE and ROBUST assortments, we consider multiple problem classes. Each problem class is characterized by the number of customer types G, the coefficient of variation vector $(\rho^1,\ldots,\rho^G)$ and the estimated proportions $\boldsymbol{\theta}$, which defines the mean vector for the Dirichlet random variable $\boldsymbol{\Theta}$. We use two different forms for $\boldsymbol{\theta}$. In the first form for $\boldsymbol{\theta}$, we have $\boldsymbol{\theta} \propto (1,1,\ldots,1)$, whereas in the second form, we have $\boldsymbol{\theta} \propto (1,4,9,\ldots,G^2)$. Both of these vectors are normalized so that their components add up to one. The first form for $\boldsymbol{\theta}$ represents a case where each customer type is estimated to be equally likely, whereas the second form for $\boldsymbol{\theta}$ corresponds to a case where some customer types are estimated to be much more likely than others. We generate 250 individual test problems for each problem class. Each test problem in a problem class is determined by randomly generating G vectors $\boldsymbol{v}^1,\ldots,\boldsymbol{v}^G$ and randomly generating the revenues $r_1,\ldots,r_n$ of the products. We then compare the histograms of the random variables $\text{Rev}(S_\epsilon^{\text{ROB}},\boldsymbol{\Theta})$ and $\text{Rev}(S^{\text{MIX}},\boldsymbol{\Theta})$ by generating 100,000 samples from the random vector $\boldsymbol{\Theta}$. In our test problems, we have 20 products. In the appendix, we give the complete details on how we randomly generate $\boldsymbol{v}^1,\ldots,\boldsymbol{v}^G$ and $r_1,\ldots,r_n$.

5.1.2 Numerical Results for Static Problems

Figure 1 shows the histograms of $\text{Rev}(S_\epsilon^{\text{ROB}},\boldsymbol{\Theta})$ for $\epsilon \in \{0.01,0.1,0.25,0.5\}$ and the histogram of $\text{Rev}(S^{\text{MIX}},\boldsymbol{\Theta})$ for two test problems. Each chart in Figure 1 corresponds to one test problem. Both test problems have 3 customer types and 20 products and we have $\boldsymbol{\theta} = (1/3,1/3,1/3)$. The solid data series plot the histograms of $\text{Rev}(S_\epsilon^{\text{ROB}},\boldsymbol{\Theta})$ for different values of ϵ, whereas the dashed data series plot the histogram of $\text{Rev}(S^{\text{MIX}},\boldsymbol{\Theta})$.

We observe that the histograms for $\text{Rev}(S_\epsilon^{\text{ROB}},\boldsymbol{\Theta})$ get tighter as ϵ increases. In particular, since the size of the uncertainty set grows as ϵ gets larger, we obtain more conservative assortments and the variability in the revenue decreases. For example, in Figure 1(a), the standard deviation of the histogram for $\text{Rev}(S_{0.01}^{\text{ROB}},\boldsymbol{\Theta})$ is 363, whereas the standard deviation of the histogram for $\text{Rev}(S_{0.5}^{\text{ROB}},\boldsymbol{\Theta})$ is 181. Although the histograms get tighter as ϵ increases, the average of the histograms get smaller. In Figure 1(a), the average of the histogram for $\text{Rev}(S_{0.01}^{\text{ROB}},\boldsymbol{\Theta})$ is 1337, whereas the average of the histogram for $\text{Rev}(S_{0.5}^{\text{ROB}},\boldsymbol{\Theta})$ is 1090. Therefore, ϵ serves as a knob to adjust the tradeoff between the variability and average performance of the robust assortment. As ϵ gets larger the variability in performance gets smaller, but so does the average performance. As expected, the average of the histogram for
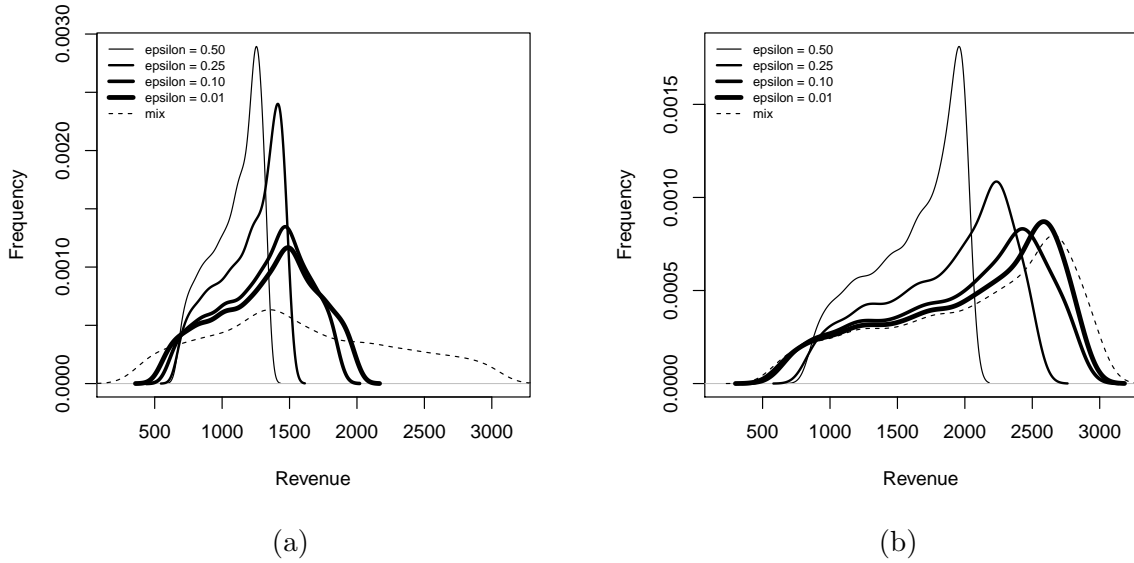
Figure 1: Comparison between the distribution of revenues provided by $S_\epsilon^{\mathrm{ROB}}$ for $\epsilon \in \{0.01, 0.1, 0.25, 0.5\}$ and S^{MIX} for two test problems.

$\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ is larger than the average of the histograms for $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$, but the width of the histogram for $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ is also larger than the width of the histograms for $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$. The average and the standard deviation of the histograms for $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ are respectively 1585 and 678 in Figure 1(a). We also note that the tails of the histograms for $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ can be much shorter than those of the histograms for $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$. In Figure 1(a), the first percentiles of the histograms for $\mathsf{Rev}(S_{0.5}^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(S_{0.01}^{\mathrm{ROB}}, \boldsymbol{\Theta})$ are respectively 690 and 574, whereas the first percentile of the histogram for $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ is 393. In other words, for this test problem, if we offer the robust assortment $S_\epsilon^{\mathrm{ROB}}$ with $\epsilon = 0.5$, then the expected revenue per customer exceeds 690 with 99% probability, whereas if we offer the assortment S^{MIX}, then the expected revenue per customer exceeds 393 with the same probability. Therefore, the robust assortment provides significantly better protection against the worst case expected revenues. Similar observations apply to Figure 1(b).

In Table 1, we compare the assortments $S_\epsilon^{\mathrm{ROB}}$ and S^{MIX} for different values of ϵ over 12 problem classes. We use the three attributes (T, G, CV) to represent each problem class. The attribute T takes value I if we use the form $\boldsymbol{\theta} \propto (1, 1 \ldots, 1)$, and takes value II if we use the form $\boldsymbol{\theta} \propto (1, 4, 9, \ldots, G^2)$. The attribute G corresponds to the number of customer types. The attribute CV is the coefficient of variation of $\boldsymbol{\Theta}$. If we use the form $\boldsymbol{\theta} \propto (1, 1 \ldots, 1)$ for $\boldsymbol{\theta}$, then the coefficient of variation for each component $\boldsymbol{\Theta}$ is the same and CV corresponds to the common coefficient of variation. If we use the form $\boldsymbol{\theta} \propto (1, 4, 9, \ldots, G^2)$ for $\boldsymbol{\theta}$, then CV corresponds to the smallest coefficient of variation for the components of $\boldsymbol{\Theta}$. In each problem class, we generate 250 individual test problems. 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 $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$. In other words, let the random variables $\mathsf{Rev}^k(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}^k(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ denote the

Prob. class (T, G, CV)	ϵ	Ratios between			Prob. class (T, G, CV)	ϵ	Ratios between		
		1 Per.	Sdv.	Avg.			1 Per.	Sdv.	Avg.
(I, 3, 1.0)	0.50	1.06	0.79	0.95	(II, 3, 0.5)	0.50	1.04	0.84	0.97
(I, 3, 1.0)	0.25	1.05	0.85	0.97	(II, 3, 0.5)	0.25	1.04	0.87	0.98
(I, 3, 1.0)	0.10	1.03	0.92	0.98	(II, 3, 0.5)	0.10	1.03	0.90	0.99
(I, 3, 1.0)	0.01	1.01	0.96	0.99	(II, 3, 0.5)	0.01	1.02	0.94	0.99
(I, 3, 1.2)	0.50	1.08	0.79	0.95	(II, 3, 0.7)	0.50	1.05	0.84	0.97
(I, 3, 1.2)	0.25	1.07	0.85	0.97	(II, 3, 0.7)	0.25	1.04	0.87	0.98
(I, 3, 1.2)	0.10	1.04	0.92	0.98	(II, 3, 0.7)	0.10	1.03	0.90	0.99
(I, 3, 1.2)	0.01	1.02	0.96	0.99	(II, 3, 0.7)	0.01	1.03	0.94	0.99
(I, 6, 1.2)	0.50	1.08	0.74	0.94	(II, 6, 0.7)	0.50	1.06	0.80	0.96
(I, 6, 1.2)	0.25	1.08	0.77	0.95	(II, 6, 0.7)	0.25	1.06	0.82	0.97
(I, 6, 1.2)	0.10	1.06	0.86	0.98	(II, 6, 0.7)	0.10	1.06	0.86	0.98
(I, 6, 1.2)	0.01	1.02	0.94	0.99	(II, 6, 0.7)	0.01	1.04	0.93	0.99
(I, 6, 2.0)	0.50	1.18	0.74	0.94	(II, 6, 1.0)	0.50	1.08	0.80	0.96
(I, 6, 2.0)	0.25	1.17	0.77	0.95	(II, 6, 1.0)	0.25	1.08	0.82	0.97
(I, 6, 2.0)	0.10	1.12	0.86	0.98	(II, 6, 1.0)	0.10	1.07	0.86	0.98
(I, 6, 2.0)	0.01	1.06	0.94	0.99	(II, 6, 1.0)	0.01	1.04	0.93	0.99
(I, 12, 2.0)	0.50	1.10	0.69	0.92	(II, 12, 1.0)	0.50	1.06	0.75	0.95
(I, 12, 2.0)	0.25	1.10	0.72	0.93	(II, 12, 1.0)	0.25	1.06	0.78	0.96
(I, 12, 2.0)	0.10	1.09	0.78	0.96	(II, 12, 1.0)	0.10	1.06	0.82	0.97
(I, 12, 2.0)	0.01	1.02	0.94	0.99	(II, 12, 1.0)	0.01	1.04	0.91	0.99
(I, 12, 3.0)	0.50	1.19	0.69	0.92	(II, 12, 1.5)	0.50	1.11	0.75	0.95
(I, 12, 3.0)	0.25	1.19	0.72	0.93	(II, 12, 1.5)	0.25	1.11	0.78	0.96
(I, 12, 3.0)	0.10	1.17	0.78	0.96	(II, 12, 1.5)	0.10	1.10	0.82	0.97
(I, 12, 3.0)	0.01	1.04	0.94	0.99	(II, 12, 1.5)	0.01	1.06	0.91	0.99

Table 1: Comparison between the revenues provided by $S_\epsilon^{\mathrm{ROB}}$ and S^{MIX} over 12 problem classes.

revenues under the assortments $S_\epsilon^{\mathrm{ROB}}$ and S^{MIX}, respectively, in the kth test problem. We compute

$$\frac{\text{1st Per Rev}^k(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})}{\text{1st Per Rev}^k(S^{\mathrm{MIX}}, \boldsymbol{\Theta})}, \quad \frac{\text{Std Rev}^k(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})}{\text{Std Rev}^k(S^{\mathrm{MIX}}, \boldsymbol{\Theta})}, \quad \frac{\text{Avg Rev}^k(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})}{\text{Avg Rev}^k(S^{\mathrm{MIX}}, \boldsymbol{\Theta})}.$$

In Table 1, we report the averages of these ratios over 250 test problems in each problem class. The first column in Table 1 shows the problem class by using the attributes (T, G, CV). The second column shows the value of ϵ that we use to compute $S_\epsilon^{\mathrm{ROB}}$. The third, fourth and fifth 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 $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ tend to be larger than those of $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$, indicating that the left tails of $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ tend to be shorter than the left tails of $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$. The assortments provided by the robust model with large values of ϵ are especially effective in protecting against the worse cast expected revenues, and the first percentiles of $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ with $\epsilon = 0.5$ can exceed those of $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ by up to 19%. Furthermore, as the coefficient of variation CV increases, the gap in the first percentiles of $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$ increases. Therefore, the risk associated with the assortment S^{MIX} becomes more apparent when $\boldsymbol{\Theta}$ can deviate significantly from its mean. The standard deviations of $\mathsf{Rev}(S_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ are significantly smaller than those of $\mathsf{Rev}(S^{\mathrm{MIX}}, \boldsymbol{\Theta})$. The gap between the standard deviations can be as large as 31% for the test problems with large number of customer classes. Finally, the average performance of the robust assortment $S_\epsilon^{\mathrm{ROB}}$ lags behind the average performance of the assortment S^{MIX}, but if we use a small value for ϵ, then the gaps in the average performance can shrink to merely a fraction of a percent. Therefore,

depending on the desired tradeoff between protecting against the worst case revenue and obtaining good average performance, we can find a value of ϵ that balances the worst case revenue and average performance provided by S_ϵ^{ROB}.

5.2 Robust Capacity Allocation Problems

In this section, we work with the robust capacity allocation model described in Section 4.

5.2.1 Numerical Setup

Our numerical setup is similar to the one that we use for robust assortment optimization problems. In particular, our goal is to understand how parameter estimation errors impact the expected revenue and how our robust approach can alleviate estimation errors. We consider a firm that dynamically makes assortment offer decisions in a market with G customer types. The choices of a customer of type g are governed by the multinomial logit model with preference weights $\boldsymbol{v}^g$. At each time period, we need to decide which assortment of products to offer to an arriving customer, whose type is unknown to us. This situation arises in brick-and-mortar retail or e-commerce settings where it is difficult to know the type of each customer.

In the absence of information on the exact customer type, let θ^g be our estimate of the probability that an arriving customer at each time period is of type g. Then, we can consider the problem of maximizing the total expected revenue based on our estimate $\boldsymbol{\theta} = (\theta^1, \ldots, \theta^G)$. The optimal policy for this problem can be found by solving the Bellman equation

$$V_t(x) = \max_{S_t \subseteq \mathcal{A}} \left\{ \sum_{i \in S_t} \left[\sum_{g=1}^{G} \theta^g \phi_i(S_t, \boldsymbol{v}^g) \right] (r_i + V_{t+1}(x-1)) + \left(1 - \sum_{i \in S_t} \left[\sum_{g=1}^{G} \theta^g \phi_i(S_t, \boldsymbol{v}^g) \right] \right) V_{t+1}(x) \right\},$$

where the expression in the square brackets above corresponds to the probability that there is a purchase for product i given that we offer the assortment S_t at time period t. The optimality equation above naturally defines a policy that allows us to decide which assortment to offer at each time period as a function of the remaining inventory. We use π^{MIX} to denote this policy, where the superscript MIX emphasizes the fact that the purchase probabilities in the square brackets above are obtained by mixing G multinomial logit models.

We focus on the situation where there are errors in the estimated probabilities $\boldsymbol{\theta}$. In this case, another approach for choosing an assortment to offer at each time period is to solve the Bellman equation that characterizes the DYNAMIC ROBUST problem in Section 4 with the ELLIPSOIDAL uncertainty set $\mathcal{V}_{\bar{\boldsymbol{v}},\epsilon}$ given in Section 3.3. This optimality equation also defines a policy that allows us to decide which assortment to offer at each time period as a function of the remaining inventory. We use $\pi_\epsilon^{\text{ROB}}$ to denote the policy that we obtain in this way, where the superscript ROB stands for robust.

Similar to our earlier development, to model in the error in the estimates of the probabilities, we let Θ^g be the true probability that a customer is of type g, and assume that $\boldsymbol{\Theta} = (\Theta^1, \ldots, \Theta^G)$ is a G-dimensional Dirichlet random variable. We set the mean vector of $\boldsymbol{\Theta}$ to be $\boldsymbol{\theta}$. We assume that the

coefficient of variation of Θ^g is the same for all g, and we denote its value by ρ. Larger values of ρ imply that the true probabilities $\boldsymbol{\Theta}$ have higher propensity to differ from the estimated values $\boldsymbol{\theta}$. Let $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ denote the total expected revenue obtained by the policy $\pi_\epsilon^{\mathrm{ROB}}$ when, in each period, the customer arrivals occur according to the probabilities $\boldsymbol{\Theta}$. Similarly, let $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$ denote the total expected revenue under π^{MIX}. Both $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$ are random variables because $\boldsymbol{\Theta}$ is a random variable. Our goal is to compare the distributions of these two random variables. Note that for a specific sample $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^G)$ of $\boldsymbol{\Theta}$, we can compute the total expected revenue under any policy π using the recursion

$$V_t^\pi(x, \boldsymbol{\alpha}) = \sum_{i \in \pi_t(x)} \left[\sum_{g=1}^G \alpha^g \, \phi_i(S_t, \boldsymbol{v}^g) \right] \left(r_i + V_{t+1}^\pi(x - 1, \boldsymbol{\alpha}) \right) + \left(1 - \sum_{i \in \pi_t(x)} \left[\sum_{g=1}^G \alpha^g \phi_i(S_t, \boldsymbol{v}^g) \right] \right) V_{t+1}^\pi(x, \boldsymbol{\alpha}) \,,$$

where $\pi_t(x) \subseteq \mathcal{A}$ denotes the assortment under the policy π when there are x units of capacity remaining in period t. Then, we can let $\mathsf{Rev}(\pi, \boldsymbol{\alpha}) = V_1^\pi(C, \boldsymbol{\alpha})$.

As shown in the next section, the total expected revenue obtained by the policy $\pi_\epsilon^{\mathrm{ROB}}$ has substantially smaller variability and shorter left tails when compared with the total expected revenue obtained by the policy π^{MIX}. Therefore, the policy $\pi_\epsilon^{\mathrm{ROB}}$ does a better job of protecting against undesirable outcomes when there is uncertainty in the model parameters.

5.2.2 Numerical Results for Dynamic Problems

In this section, we compare the total expected revenues obtained by the policies $\pi_\epsilon^{\mathrm{ROB}}$ and π^{MIX} on test problems that are based on the experimental setup in Talluri and van Ryzin (2004). In these test problems, the number of products n is 10 and the number of customer types G is 2. Table 2 shows the product revenues and the product preference weights for each customer type. We observe that the first customer type is somewhat price insensitive because it associates relatively high preference weights with the expensive products. On the other hand, the second customer type associates relatively low preference weights with the expensive products, and it is unlikely to purchase expensive products even if these products are in the offered assortment. Both customer types associate the preference weight of one with the no purchase option, that is, $v_0^1 = v_0^2 = 1$. We assume that $\boldsymbol{\theta} = (0.5, 0.5)$ so that each customer type is estimated to be equally likely to arrive in each period. We sample the true probabilities $\boldsymbol{\Theta} = (\Theta^1, \Theta^2)$ from the two-dimensional Dirichlet distribution with mean $\boldsymbol{\theta}$. The coefficient of variation for Θ^1 and Θ^2 is ρ. The number of time periods in the selling horizon is 100 and we have C units of initial inventory. We vary C and ρ over $(C, \rho) \in \{30, 50, 70\} \times \{0.5, 0.9\}$ and this provides 6 test problems. Varying ϵ over $[0.1, 0.5]$ provides the most interesting tradeoffs for these test problems. For all of our test problems, the results we get with values of ϵ beyond 0.5 are identical to those we get with a value of 0.5. Thus, we use $\epsilon \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ in our numerical experiments.

In Table 3, we compare the distributions of the random variables $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$, which correspond to the total expected revenues obtained by the policies $\pi_\epsilon^{\mathrm{ROB}}$ and π^{MIX}, when the true fraction of each customer type is given by $\boldsymbol{\Theta}$. The first column in this table shows the test problems by using the pair (C, ρ). The second column shows the value of ϵ that we use to compute $\pi_\epsilon^{\mathrm{ROB}}$. The third, fourth and fifth columns respectively show the first percentile, standard deviation and mean of the

i	1	2	3	4	5	6	7	8	9	10
r_i	600	550	475	400	300	280	240	200	185	175
v_i^1	0.407	0.438	0.490	0.549	0.638	0.657	0.698	0.741	0.758	0.769
v_i^2	0.050	0.064	0.093	0.135	0.223	0.247	0.301	0.368	0.397	0.417

Table 2: Product revenues and preference weights associated with the products in the experimental setup in Talluri and van Ryzin (2004).

random variable $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$. We estimate these summary statistics by drawing 100,000 realizations from the distribution of $\boldsymbol{\Theta}$. In particular, for each realization $\boldsymbol{\alpha}$ of the random variable $\boldsymbol{\Theta}$, we evaluate $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\alpha})$ using the recursion from the previous section. The interpretations of the sixth, seventh and eighth columns are similar to those of the third, fourth and fifth, but these columns provide summary statistics for the random variable $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$. The last three columns compare the random variables $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ and $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$ by providing the ratios between the corresponding statistics.

The results in Table 3 indicate that the first percentiles of $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ are noticeably larger than those of $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$, showing that the left tails of $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ are shorter than the left tails of $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$. For larger values of ϵ, the robust model provides assortments that are especially effective in protecting against the worst cases, and the first percentiles of the policies $\pi_\epsilon^{\mathrm{ROB}}$ and π^{MIX} can differ by up to 19%. Furthermore, the gaps in the first percentiles are especially large when the initial capacity is small. Therefore, the risk associated with the policy π^{MIX} becomes more apparent when inventory is scarce. The standard deviations of $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ are substantially smaller than those of $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$, which indicates that $\mathsf{Rev}(\pi_\epsilon^{\mathrm{ROB}}, \boldsymbol{\Theta})$ tends to fluctuate less than $\mathsf{Rev}(\pi^{\mathrm{MIX}}, \boldsymbol{\Theta})$. There are test problems where the gap between the standard deviations is as large as 52%. These observations suggest that there is noticeably higher risk associated with the policy π^{MIX} when compared with the policy $\pi_\epsilon^{\mathrm{ROB}}$. Overall, we observe test problems where the first percentile and standard deviation provided by $\pi_\epsilon^{\mathrm{ROB}}$ are respectively 19% and 50% better than those provided by π^{MIX}, but the average performance of $\pi_\epsilon^{\mathrm{ROB}}$ is only 3% worse than that of π^{MIX}. The average performance of the policy $\pi_\epsilon^{\mathrm{ROB}}$ generally lags behind the average performance of the policy π^{MIX}. However, if average performance is of concern, then by using a small value for ϵ, we can sacrifice the first percentile of the performance of the policy $\pi_\epsilon^{\mathrm{ROB}}$, but get an average performance from this policy that is comparable to that of the policy π^{MIX}.

6. Conclusions

Motivated by applications where there is uncertainty about the parameters of the underlying choice model, we formulated novel robust assortment optimization problems under the multinomial logit choice model. Both static and dynamic variants admit efficient solution methods. The solutions under our robust approach have intuitive practical appeal, and offer important operational insights. In addition, we propose a family of uncertainty sets that allows the decision maker to control the tradeoff between improving the average performance and protecting against the worst case revenues. Our numerical experiments demonstrate that our robust approach is effective in reducing the risk when the estimated model parameters can significantly differ from the true values, while delivering similar average performance when compared to other heuristics if average performance happens to be the

Prob. Class (C,ρ)	ϵ	Stats. for $\mathsf{Rev}(\pi_\epsilon^{\mathrm{Rob}},\boldsymbol{\Theta})$			Stats. for $\mathsf{Rev}(\pi^{\mathrm{Mix}},\boldsymbol{\Theta})$			Ratios between stats.		
		1 Per.	Sdv.	Avg.	1 Per.	Sdv.	Avg.	1 Per.	Sdv.	Avg.
$(30,0.5)$	0.5	12,295	809	15,180	10,705	1,695	15,958	1.15	0.48	0.95
$(30,0.5)$	0.4	12,125	1,010	15,624	10,705	1,695	15,958	1.13	0.60	0.98
$(30,0.5)$	0.3	11,535	1,290	15,889	10,705	1,695	15,958	1.08	0.76	1.00
$(30,0.5)$	0.2	11,010	1,522	15,960	10,705	1,695	15,958	1.03	0.90	1.00
$(30,0.5)$	0.1	10,675	1,682	15,949	10,705	1,695	15,958	1.00	0.99	1.00
$(30,0.9)$	0.5	11,425	1,192	14,884	9,635	2,366	15,361	1.19	0.50	0.97
$(30,0.9)$	0.4	11,055	1,489	15,240	9,635	2,366	15,361	1.15	0.63	0.99
$(30,0.9)$	0.3	10,435	1,874	15,374	9,635	2,366	15,361	1.08	0.79	1.00
$(30,0.9)$	0.2	9,855	2,192	15,380	9,635	2,366	15,361	1.02	0.93	1.00
$(30,0.9)$	0.1	9,340	2,393	15,333	9,635	2,366	15,361	0.97	1.01	1.00
$(50,0.5)$	0.5	15,830	1,826	20,687	14,475	3,261	22,150	1.09	0.56	0.93
$(50,0.5)$	0.4	15,485	2,263	21,407	14,475	3,261	22,150	1.07	0.69	0.97
$(50,0.5)$	0.3	15,215	2,703	21,909	14,475	3,261	22,150	1.05	0.83	0.99
$(50,0.5)$	0.2	14,585	3,159	22,139	14,475	3,261	22,150	1.01	0.97	1.00
$(50,0.5)$	0.1	13,955	3,367	22,089	14,475	3,261	22,150	0.96	1.03	1.00
$(50,0.9)$	0.5	14,855	2,425	20,343	13,295	4,256	21,639	1.12	0.57	0.94
$(50,0.9)$	0.4	14,445	2,983	20,987	13,295	4,256	21,639	1.09	0.70	0.97
$(50,0.9)$	0.3	13,935	3,600	21,422	13,295	4,256	21,639	1.05	0.85	0.99
$(50,0.9)$	0.2	13,485	4,154	21,608	13,295	4,256	21,639	1.01	0.98	1.00
$(50,0.9)$	0.1	12,795	4,382	21,539	13,295	4,256	21,639	0.96	1.03	1.00
$(70,0.5)$	0.5	16,570	2,730	22,816	15,640	3,911	24,026	1.06	0.70	0.95
$(70,0.5)$	0.4	16,040	3,345	23,698	15,640	3,911	24,026	1.03	0.86	0.99
$(70,0.5)$	0.3	15,990	3,506	23,836	15,640	3,911	24,026	1.02	0.90	0.99
$(70,0.5)$	0.2	15,905	3,837	24,014	15,640	3,911	24,026	1.02	0.98	1.00
$(70,0.5)$	0.1	14,890	4,113	23,914	15,640	3,911	24,026	0.95	1.05	1.00
$(70,0.9)$	0.5	15,740	3,558	22,590	14,515	5,183	23,913	1.08	0.69	0.94
$(70,0.9)$	0.4	14,910	4,373	23,392	14,515	5,183	23,913	1.03	0.84	0.98
$(70,0.9)$	0.3	14,780	4,635	23,611	14,515	5,183	23,913	1.02	0.89	0.99
$(70,0.9)$	0.2	14,505	5,112	23,918	14,515	5,183	23,913	1.00	0.99	1.00
$(70,0.9)$	0.1	13,715	5,492	23,711	14,515	5,183	23,913	0.94	1.06	0.99
$(90,0.5)$	0.5	16,595	2,950	22,917	16,000	3,716	23,999	1.04	0.79	0.95
$(90,0.5)$	0.4	16,000	3,714	24,000	16,000	3,716	23,999	1.00	1.00	1.00
$(90,0.5)$	0.3	15,985	3,716	23,999	16,000	3,716	23,999	1.00	1.00	1.00
$(90,0.5)$	0.2	16,000	3,716	23,999	16,000	3,716	23,999	1.00	1.00	1.00
$(90,0.5)$	0.1	15,025	4,239	24,001	16,000	3,716	23,999	0.94	1.14	1.00
$(90,0.9)$	0.5	15,575	3,944	22,870	14,865	5,049	23,949	1.05	0.78	0.95
$(90,0.9)$	0.4	14,865	5,048	23,948	14,865	5,049	23,949	1.00	1.00	1.00
$(90,0.9)$	0.3	14,865	5,048	23,948	14,865	5,049	23,949	1.00	1.00	1.00
$(90,0.9)$	0.2	14,865	5,049	23,948	14,865	5,049	23,949	1.00	1.00	1.00
$(90,0.9)$	0.1	13,650	5,817	23,942	14,865	5,049	23,949	0.92	1.15	1.00

Table 3: Comparison between the total expected revenues obtained by the policies $\pi_\epsilon^{\mathrm{Rob}}$ and π^{Mix}.

performance measure of interest.

In this paper, we assumed that the model parameters lie in an uncertainty set, but we do not refine the uncertainty set as we observe the purchases of the customers. This is a reasonable assumption when it is difficult to learn the model parameters either because there are a few sale opportunities or the different customers are so distinct from each other that the purchase behavior of one customer tells us little about the purchase behavior of others. Another option is to resolve the problem as more information becomes available and the uncertainty sets are refined. Nevertheless, there are undoubtedly settings where one could learn from the purchase behavior of the past customers to predict the preferences of the future customers. In this case, adaptively refining the offered assortment and the uncertainty set by embedding a learning mechanism into the assortment optimization model would be beneficial.

It would also be worthwhile to compare the solution under our robust formulation to the one obtained by putting a prior distribution on the likely values of the model parameters. Understanding the tradeoffs between these two approaches might lead to a new framework for thinking about assortment optimization problems. Finally, it would be interesting to extend our analysis to more complex choice models.

Acknowledgement

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.

Ben-Tal, A., and A. Nemirovski. 1999. Robust solutions of uncertain linear programs. *Operations Research Letters* 25 (1): 1–13.

Bertsimas, D., D. Pachamanovab, and M. Sim. 2004. Robust linear optimization under general norms. *Operations Research Letters* 32 (6): 510–516.

Bertsimas, D., and M. Sim. 2003. Robust discrete optimization and network flows. *Mathematical Programming* 98 (1): 49–71.

Bertsimas, D., and M. Sim. 2004. The price of robustness. *Operations Research* 52 (1): 35–53.

Birbil, S. I., J. B. G. Frank, J. A. S. Gromicho, and S. Zhang. 2009. The role of robust optimization in single-leg airline revenue management. *Management Science* 55:148–163.

Boyd, S., and L. Vandenberghe. 2004. *Convex optimization*. Cambridge, UK: Cambridge University 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.

Chaneton, J., and G. Vulcano. 2011. Computing bid-prices for revenue management under customer choice behavior. To appear in *Manufacturing and Service Operations Management*.

Farias, V. F., S. Jagabathula, and D. Shah. 2010. A non-parametric approach to modeling choice with limited data. Working paper, MIT.

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

Iyengar, G. N. 2005. Robust dynamic programming. *Mathematics of Operations Research* 30 (2): 257–280.

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.

Kunnumkal, S., and H. Topaloglu. 2010. A New Dynamic Programming Decomposition Method for the Network Revenue Management Problem with Customer Choice Behavior. *Production and Operations Management* 19 (5): 575–590.

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.

Liu, Q., and G. J. van Ryzin. 2008. Strategic capacity rationing to induce early purchases. *Management Science* 54 (6): 1115–1131.

Megiddo, N. 1979. Combinatorial optimization with rational objective functions. *Mathematics of Operations Research* 4 (4): 414–424.

Meissner, J., and A. K. Strauss. 2010. Improved bid prices for choice-based network revenue management. Working paper, Department of Management Science, Lancaster University.

Meissner, J., and A. K. Strauss. 2011. Network revenue management with inventory-sensitive bid prices and customer choice. To appear in *European Journal of Operational Research*.

Mendez-Diaz, I., J. M. Bront, G. Vulcano, and P. Zabala. 2010. A Branch-and-Cut Algorithm for the Latent-Class Logit Assortment Problem. Working Paper, New York University.

Perakis, G., and G. Roels. 2010. Robust controls for network revenue management. *Manufacturing & Service Operations Management* 12:56–76.

Phillips, R. 2011. Efficient frontiers in revenue management. INFORMS Revenue Management and Pricing Section Conference, New York, NY.

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., D. B. Shmoys, and H. Topaloglu. 2010. Assortment optimization with mixtures of logits. Working paper, Cornell University.

Ruszczynski, A. 2006. *Nonlinear optimization*. Princeton, New Jersey: Princeton University Press.

Su, X. 2007. Intertemporal pricing with strategic customer behavior. *Management Science* 53 (5): 726–741.

Su, X. 2010. Intertemporal pricing and consumer stockpiling. *Operations Research* 58 (4): 1133–1147.

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

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. J., and S. Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science* 45:1496–1509.

van Ryzin, G. J., and G. Vulcano. 2008. Computing virtual nesting controls for network revenue management under customer choice behavior. *Manufacturing and Service Operations Management* 10 (3): 448–467.

Vulcano, G., G. J. van Ryzin, and W. Chaar. 2010. Choice-based revenue management: An empirical study of estimation and optimization. *Manufacturing and Service Operations Management* 12 (3): 371–392.

Vulcano, G., G. J. van Ryzin, and R. Ratliff. 2011. Estimating primary demand for substitutable products from sales transaction data. To appear in *Operations Research*.

Zhang, D. 2011. An improved dynamic programming decomposition approach for network revenue management. *Manufacturing and Service Operations Management* 13 (1): 35–52.

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

A.   Appendix: Proof of Theorem 4.1

We first establish, by induction, that the value functions are concave in the remaining capacity; that is, $\Delta J_t(x)$ is decreasing in x for all t. The result is trivially true for $\Delta J_{T+1}(\cdot)$. So, assume that it holds for $\Delta J_{t+1}(\cdot)$. We will now prove that it also holds for $\Delta J_t(\cdot)$. It follows from the Bellman equation in the DYNAMIC ROBUST problem and the definition of $S_t^*(x)$ that

$$\Delta J_t(x+2) - \Delta J_t(x+1) = J_t(x+2) - J_t(x+1) - J_t(x+1) + J_t(x)$$

$$= \max_{S_t \subseteq \mathcal{A}} \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, v_t)(r_i - \Delta J_{t+1}(x+2)) \right\} + J_{t+1}(x+2)$$

$$- \max_{S_t \subseteq \mathcal{A}} \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, v_t)(r_i - \Delta J_{t+1}(x+1)) \right\} - J_{t+1}(x+1)$$

$$- \max_{S_t \subseteq \mathcal{A}} \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, v_t)(r_i - \Delta J_{t+1}(x+1)) \right\} - J_{t+1}(x+1)$$

$$+ \max_{S_t \subseteq \mathcal{A}} \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, v_t)(r_i - \Delta J_{t+1}(x)) \right\} + J_{t+1}(x)$$

$$= \Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1) + \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), v_t)(r_i - \Delta J_{t+1}(x+2)) \right\}$$

$$- \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), v_t)(r_i - \Delta J_{t+1}(x+1)) \right\}$$

$$+ \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), v_t)(r_i - \Delta J_{t+1}(x)) \right\}$$

$$- \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), v_t)(r_i - \Delta J_{t+1}(x+1)) \right\} .$$

We will now bound the terms in the right hand side of the above equality. By definition, $S_t^*(x+1)$ is the optimal solution to the problem $\max_{S_t \subseteq \{1,2,\dots,n\}} \min_{v_t \in \mathcal{V}_t} \sum_{i \in S_t} \phi_i(S_t, v_t)(r_i - \Delta J_{t+1}(x+1))$, which implies that

$$\min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), v_t)(r_i - \Delta J_{t+1}(x+1)) \right\} \geq \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), v_t)(r_i - \Delta J_{t+1}(x+1)) \right\}$$

$$= \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), v_t)(r_i - \Delta J_{t+1}(x)) + \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), v_t)(\Delta J_{t+1}(x) - \Delta J_{t+1}(x+1)) \right\}$$

$$\geq \min_{v_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), v_t)(r_i - \Delta J_{t+1}(x)) \right\}$$

$$+ \left\{ \min_{v_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), v_t) \right\} (\Delta J_{t+1}(x) - \Delta J_{t+1}(x+1)) ,$$

where the last inequality follows from our hypothesis that $\Delta J_{t+1}(x) - \Delta J_{t+1}(x+1) \geq 0$. Therefore,

$$\min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x)) \right\} - \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+1)) \right\}$$

$$\leq \left\{ \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+1) - \Delta J_{t+1}(x)) \ .$$

Using a similar argument, we have that

$$\min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+1)) \right\}$$

$$\geq \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+1)) \right\}$$

$$= \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+2)) \right.$$

$$\left. + \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t)\, (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1)) \right\}$$

$$\geq \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+2)) \right\}$$

$$+ \left\{ \max_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1)) \ ,$$

where the last inequality follows from the fact that $\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1) \leq 0$, which means that

$$\min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t)\, (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1)) \right\}$$

$$= \left\{ \max_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1)) \ .$$

Therefore, we have

$$\min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+2)) \right\}$$

$$- \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), \boldsymbol{v}_t)\, (r_i - \Delta J_{t+1}(x+1)) \right\}$$

$$\leq - \left\{ \max_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1)) \ .$$

Putting everything together, we bound $\Delta J_t(x+2) - \Delta J_t(x+1)$ by

$$\Delta J_t(x+2) - \Delta J_t(x+1)$$

$$\leq \Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1) - \left\{ \max_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1))$$

$$+ \left\{ \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+1) - \Delta J_{t+1}(x))$$

$$= \left\{ 1 - \max_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x+2)} \phi_i(S_t^*(x+2), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1))$$

$$+ \left\{ \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), \boldsymbol{v}_t) \right\} (\Delta J_{t+1}(x+1) - \Delta J_{t+1}(x))$$

$$\leq 0 \,,$$

where the last inequality follows from the fact that $\max_{\boldsymbol{v}_t \in \mathcal{V}_t} \sum_{i \in S_t} \phi_i(S_t^*(x+2), \boldsymbol{v}_t) \leq 1$ and from our inductive hypothesis that $\Delta J_{t+1}(x+2) - \Delta J_{t+1}(x+1) \leq 0$ and $\Delta J_{t+1}(x+1) - \Delta J_{t+1}(x) \leq 0$. This completes the induction.

We will now use the concavity of the value functions to show that marginal value of capacity is decreasing with time; that is, $\Delta J_{t+1}(x+1) \leq \Delta J_t(x+1)$. It follows from the Bellman equation in the DYNAMIC ROBUST problem that

$$\Delta J_t(x+1) - \Delta J_{t+1}(x+1)$$
$$= J_t(x+1) - J_t(x) - \Delta J_{t+1}(x+1)$$
$$= \max_{S_t} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x+1)) \right\} - \max_{S_t} \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t} \phi_i(S_t, \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x)) \right\}$$
$$= \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x+1)) \right\} - \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x)) \right\}$$
$$\geq \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x+1)} \phi_i(S_t^*(x+1), \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x+1)) \right\} - \min_{\boldsymbol{v}_t \in \mathcal{V}_t} \left\{ \sum_{i \in S_t^*(x)} \phi_i(S_t^*(x), \boldsymbol{v}_t)(r_i - \Delta J_{t+1}(x+1)) \right\}$$
$$\geq 0 \,,$$

where the first inequality follows from the fact that $\Delta J_{t+1}(x) \geq \Delta J_{t+1}(x+1)$ and the second inequality follows from the definition $S_t^*(x+1)$.

B. Appendix: Generation of Test Problems

To come up with the preference weights for each customer type, 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 types. 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 types. In this case, we sample the parameter ϑ_i^g from the uniform distribution over $[0, 10]$ for each product i and customer type 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 types. For the no purchase weights, we set $v_0^g = 1$ for each customer type g. 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 $r_i = [n + 1 - i]\,\kappa_i$. In this way, we hope to 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 $r_1 \geq r_2 \geq \cdots \geq r_n$.