Technical Note: Capacity Constraints Across Nests in Assortment Optimization Under the Nested Logit Model

Jacob B. Feldman, Huseyin Topaloglu
School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14853, USA
jbf232@cornell.edu, topaloglu@orie.cornell.edu

March 16, 2015

Abstract

We consider assortment optimization problems when customers choose according to the nested logit model and there is a capacity constraint limiting the total capacity consumption of all products offered in all nests. When each product consumes one unit of capacity, our capacity constraint limits the cardinality of the offered assortment. For the cardinality constrained case, we develop an efficient algorithm to compute the optimal assortment. When the capacity consumption of each product is arbitrary, we give an algorithm to obtain a 4-approximate solution. We show that we can compute an upper bound on the optimal expected revenue for an individual problem instance by solving a linear program. In our numerical experiments, we consider problem instances involving products with arbitrary capacity consumptions. Comparing the expected revenues from the assortments obtained by our 4-approximation algorithm with the upper bounds on the optimal expected revenues, our numerical results indicate that the 4-approximation algorithm performs quite well, yielding less than 2% optimality gap on average.

1 Introduction

A conventional approach to modeling demand in revenue management is to assume that each customer arrives into the system with the intention of purchasing a fixed product. If this product is available for sale, then the customer purchases it. Otherwise, the customer leaves the system without making a purchase. In reality, however, there may be multiple products that can potentially serve the needs of a customer, in which case, customers may make a choice between the products and may substitute a product for another one when their favorite product is not available. This kind of a choice process creates interactions between the demand for the different products, inflating the demand for an available product when some other product is not available so that customers satisfy their needs by substituting for the available product. A common question that arises in this setting is what products to make available to customers so as to maximize the expected revenue, given that customers choose and substitute according to a particular choice model.

In this paper, we consider assortment optimization problems when customers choose according to the nested logit model and there is limited capacity for the products in the offered assortment. We consider a setting where we need to decide which assortment of products to offer. Each arriving customer chooses among the offered products according to the nested logit model. Under the nested logit model, the products are organized in nests. Each customer, after viewing the offered assortment, decides either to make a purchase within one of the nests or to leave the system without purchasing anything. If a nest is chosen, then the customer purchases one of the products within the chosen nest. There is a capacity constraint limiting the total capacity consumption of the products in the offered assortment. The goal is to choose an assortment of products to offer so as to maximize the expected revenue obtained from each customer. We consider two types of capacity constraints. In the first type of constraints, each product occupies one unit of space, in which case, the capacity constraint limits the total number of products in the offered assortment. We refer to this type of a capacity constraint as a *cardinality constraint.* In the second type of constraints, the capacity consumption of a product is arbitrary, possibly reflecting the space or capital requirement of a product. We refer to this type of a capacity constraint as a *space constraint.*

Under a cardinality constraint, we show that we can obtain an optimal assortment by solving a linear program with $O(m^2n)$ decision variables and $O(m^2n^4)$ constraints, where m in the number of nests and n is the number of products in each nest. As far as we are aware, the assortment problem was not known to be tractable when customers choose according to the nested logit model and there is a cardinality constraint limiting the total number of products in the offered assortment. This paper gives the first exact solution method for this problem. On the other hand, under a space constraint, we show that we can obtain a 4-approximate solution by solving a linear program with $O(m)$ decision variables and $O(mn^4)$ constraints. The running time of this algorithm scales polynomially with the number of products and the number of nests. To our knowledge, this paper gives the first algorithm for the assortment problem that scales polynomially with the number of nests, when there is a capacity constraint on the space consumption of all offered products and customers choose according to the nested logit model. In addition to giving algorithms to solve the assortment problem, we give a tractable linear program that computes an upper bound on the optimal expected revenue. By comparing the expected revenues of the assortments obtained by our 4-approximation algorithm with the upper bounds on the optimal expected revenues, we demonstrate that our 4-approximation algorithm performs quite well in practice.

An attractive approach for modeling the customer choice process is to use the utility maximization principle, where a customer associates a random utility with each product and chooses the product with the largest utility. Multinomial logit model is one of the popular choice models that are based on the utility maximization principle where the utilities of the products are independent of each other; see Luce (1959) and McFadden (1974). Due to the independence of the utilities of the products, the multinomial logit model implicitly assumes that how highly a customer evaluates a certain product has nothing to do with how highly the same customer evaluates another product. The nested logit model remedies this shortcoming by organizing the products in nests such that the utilities of the products in the same nest can be dependent on each other; see Train (2003). This feature allows the modeler to capture situations where the products in the same nest are alike and how highly a customer evaluates a certain product can be a good indicator of how highly the same customer evaluates another product.

Talluri and van Ryzin (2004) and Gallego et al. (2004) consider revenue management problems under customer choice behavior. During the course of their analyses, they give efficient approaches to solve the assortment problem under the multinomial logit model without any constraints. Rusmevichientong et al. (2010), Wang (2012) and Wang (2013) consider assortment problems under variants of the multinomial logit model with a cardinality constraint on the offered assortment and show that the problem can be solved efficiently. Bront et al. (2009), Mendez-Diaz et al. (2010) and Rusmevichientong et al. (2013) consider assortment problems where there are multiple customer types and customers of different types choose according to multinomial logit models with different parameters. The authors show that the problem is NP-complete, study heuristics and investigate valid cuts for integer programming formulations.

Davis et al. (2013) show how to solve the assortment problem under the nested logit model without any constraints. Li and Rusmevichientong (2014) give a greedy algorithm for the same problem. Gallego and Topaloglu (2014) study constrained assortment problems under the nested logit model, but they impose capacity constraints separately on the assortment offered in each nest. Similar to us, Rusmevichientong et al. (2009) and Desir and Goyal (2013) consider assortment problems under the nested logit model, where there is a constraint on the total capacity consumption of the products offered in all nests. They give approximation schemes that tradeoff running time with solution quality, but the running time for their approaches grows exponentially with the number of nests. For example, to obtain a 4-approximate solution, Rusmevichientong et al. (2009) need $O(m(m^6 n^6 \log(mn))^m)$ operations, which gets prohibitive when m exceeds two or three but there are practical applications where the number of nests easily exceeds two or three; see Train et al. (1987), Slade (2009) and Grigolon and Verboren (2013).

As mentioned above, Gallego and Topaloglu (2014) study assortment problems under the nested logit model, but they impose capacity constraints separately on the assortment offered in each nest. To understand their constraint structure, consider a retailer that is interested in finding an assortment of cars to offer to its customers. We model the demand for cars through a nested logit model, where each car category, such as compact, mid-size and sedan, corresponds to a different nest. The products within a nest correspond to cars of different models within the car category corresponding to the nest. The constraint structure in Gallego and Topaloglu (2014) assumes that there is a separate fixed amount of space reserved for the cars offered in each car category and the retailer is interested in finding a set of cars to offer such that the set of cars offered in each car category does not violate the space reserved for that car category. In contrast, we impose a capacity constraint on the assortment offered over all nests. Thus, our constraint structure assumes that there is a fixed amount of space available for all car categories and the retailer is interested in finding a set of cars to offer such that the total amount of space consumed by all offered cars in all car categories does not violate the space availability.

One line of attack for assortment optimization under the nested logit model has been to identify a collection of good candidate subsets to offer in each nest. Once these collections are identified, it is possible to solve a separate linear program to pick a subset to offer in each nest so that the combined subsets over all nests provide a good assortment. This is the strategy followed by Davis et al. (2013) and Gallego and Topaloglu (2014). We follow an approach similar to theirs in identifying the collections of candidate subsets in each nest, but due to the fact that our capacity constraint limits the total capacity consumption of the subsets of products offered in all nests, different nests interact with each other, making the assortment problem substantially more difficult. The linear programs used by Davis et al. (2013) and Gallego and Topaloglu (2014) become ineffective and we cannot build on the earlier work to figure out how to pick a subset to offer in each nest so that the combined subsets over all nests provide the highest possible expected revenue.

We get around this difficulty by using the following general methodology. The expected revenue function under the nested logit model is a fraction. We convert the problem of finding an assortment that maximizes the expected revenue into the problem of finding the fixed point of a function. Computing the value of this function at any point requires solving an optimization problem, which involves finding a subset of products to offer in each nest so as to maximize a nonlinear function. Under a cardinality constraint, we show that we can consider a small number of candidate subsets in each nest without incurring any loss. Furthermore, we can maximize the nonlinear function by using dynamic programming. Under a space constraint, we show that we can consider a small number of candidate subsets in each nest while incurring a constant factor loss. Furthermore, we can maximize the nonlinear function with a constant factor loss by solving the linear programming relaxation of a multiple choice knapsack problem.

The discussion in the previous paragraph indicates that our approach is related to maximizing a fraction. Meggido (1979) shows how to build on a tractable algorithm for a combinatorial optimization problem with a linear objective function to develop a tractable algorithm for the same combinatorial optimization problem with a fractional objective function, where the numerator and the denominator of the fraction are linear. Hashizume et al. (1987) extend this result by showing how to build on an approximation algorithm for the former problem to give an approximation algorithm for the latter, but this extension requires the existence of an approximation algorithm when the linear objective function has negative coefficients. Correa et al. (2010) show how to drop this requirement. Mittal and Schulz (2013) make generalizations to sums of fractions. One of the important distinguishing features of our assortment problem is that the objective function is a fraction where the numerator and the denominator involve nonlinearities. Thus, the general methods in the papers outlined in this paragraph do not immediately apply.

2 Problem Formulation

In this section, we formulate the assortment optimization problem that we want to solve. There are m nests indexed by $M = \{1, \ldots, m\}$. In each nest, there are n products that we can offer to customers and we index the products by $N = \{1, \ldots, n\}$. Although we assume that each nest has the same number of products, this assumption is only for notational brevity and it is straightforward to extend our results to the case where different nests have different numbers of products. Under the nested logit model, a customer decides either to make a purchase within one of the nests or to leave without purchasing anything. If the customer decides to make a purchase within one of the nests, then the customer chooses one of the products offered in this nest. We let v_{ij} be the preference weight associated with product j in nest i. Given that we offer the assortment $S_i \subset N$ of products in nest i, we use $V_i(S_i) = \sum_{j \in S_i} v_{ij}$ to denote the total preference weight of the products in the offered assortment. Under the nested logit model, if we offer the assortment S_i in nest i and a customer has already decided to make a purchase in this nest, then this customer chooses product $j \in S_i$ in nest i with probability $v_{ij}/V_i(S_i)$. We let r_{ij} be the revenue associated with product j in nest i. In this case, given that we offer the assortment S_i in nest i and a customer has already decided to make a purchase in this nest, the expected revenue that we obtain from this customer can be written as

$$R_i(S_i) = \sum_{j \in S_i} \frac{v_{ij}}{V_i(S_i)} \, r_{ij} = \frac{\sum_{j \in S_i} v_{ij} \, r_{ij}}{V_i(S_i)}.$$

Associated with each nest, there is a parameter $\gamma_i \in [0, 1]$ capturing the degree of dissimilarity between the products in nest i. The preference weight of the no purchase option is v_0. Under the nested logit model, if we offer the assortment $(S_1, \ldots, S_m)$ over all nests with $S_i \subset N$ for all $i \in M$, then a customer chooses nest i with probability $Q_i(S_1, \ldots, S_m) = V_i(S_i)^{\gamma_i} / (v_0 + \sum_{l \in M} V_l(S_l)^{\gamma_l})$, which corresponds to the probability that a customer is attracted to nest i as a function of the assortment $(S_1, \ldots, S_m)$ offered over all nests. So, if we offer the assortment $(S_1, \ldots, S_m)$ over all nests, then we obtain an expected revenue of

$$\Pi(S_1, \ldots, S_m) = \sum_{i \in M} Q_i(S_1, \ldots, S_m) \, R_i(S_i) = \frac{\sum_{i \in M} V_i(S_i)^{\gamma_i} R_i(S_i)}{v_0 + \sum_{i \in M} V_i(S_i)^{\gamma_i}}$$

from each customer. Our goal is to find an assortment of products so as to maximize the expected revenue from each customer, subject to a capacity constraint on the offered assortment.

We consider two types of capacity constraints. In the first type of constraint, we limit the total number of products offered over all nests to c. Thus, the set of feasible assortments can be written as $\{(S_1, \ldots, S_m) : \sum_{i \in M} |S_i| \leq c, \; S_i \subset N \; \forall \, i \in M\}$. We refer to this constraint as a cardinality constraint. In the second type of constraint, we let w_{ij} be the space requirement of product j in nest i and limit the total space requirement of the products offered over all nests to c. In this case, the set of feasible assortments is $\{(S_1, \ldots, S_m) : \sum_{i \in M} \sum_{j \in S_i} w_{ij} \leq c, \; S_i \subset N \; \forall \, i \in M\}$. We refer to this constraint as a space constraint. For uniformity, we use $C_i(S_i)$ to denote the capacity consumption of the assortment S_i offered in nest i. We have $C_i(S_i) = |S_i|$ under a cardinality constraint and $C_i(S_i) = \sum_{j \in S_i} w_{ij}$ under a space constraint. In this case, we can write the set of feasible assortments as $\{(S_1, \ldots, S_m) : \sum_{i \in M} C_i(S_i) \leq c, \; S_i \subset N \; \forall \, i \in M\}$ under capacity or space constraints. We want to find an assortment that maximizes the expected revenue from each customer subject to a capacity constraint, yielding the problem

$$z^* \;\; = \max_{\substack{(S_1, \ldots, S_m) : \\ \sum_{i \in M} C_i(S_i) \leq c, \\ S_i \subset N \; \forall \, i \in M}} \Big\{ \Pi(S_1, \ldots, S_m) \Big\}, \tag{1}$$

where $C_i(S_i)$ may correspond to a cardinality or space constraint. If $C_i(S_i)$ corresponds to a cardinality constraint, then we can assume without loss of generality that c is an integer. In this paper, we show that if we have a cardinality constraint on the offered assortment, then we can obtain an optimal solution to problem (1) by solving a tractable linear program. On the other hand, if we have a space constraint, then Lemma 2.1 in Rusmevichientong et al. (2009) shows that problem (1) is NP-hard even when there is a single nest with a dissimilarity parameter of one. Therefore, obtaining an optimal solution to problem (1) under a space constraint is likely to be intractable. In this paper, we show that if we have a space constraint, then we can obtain a 4-approximate solution to problem (1) by solving a tractable linear program.

3 Fixed Point Representation

In this section, we describe the connection of problem (1) to the problem of computing the fixed point of a function. This connection plays an important role throughout the paper and it becomes critical for constructing an efficient solution approach for problem (1) under a cardinality or space constraint. To connect problem (1) to the problem of computing the fixed point of a function, we define the function $f(\cdot) : \Re_+ \to \Re_+$ as

$$f(z) \quad = \max_{\substack{(S_1,\ldots,S_m):\\ \sum_{i\in M} C_i(S_i)\le c,\\ S_i\subset N\ \forall\, i\in M}} \left\{ \sum_{i\in M} V_i(S_i)^{\gamma_i}(R_i(S_i)-z)\right\}. \tag{2}$$

Offering the empty assortment over all nests is a feasible solution to the problem above providing the objective value of zero, so that $f(z) \ge 0$ for all $z \in \Re_+$. Consider a value of $\hat{z}$ that satisfies $f(\hat{z}) = v_0\,\hat{z}$, corresponding to the fixed point of the function $f(\cdot)/v_0$. Such a value of $\hat{z} \ge 0$ always exists since $f(\cdot)$ is a decreasing function and $f(0) \ge 0$. The next theorem shows that the value of $\hat{z}$ that satisfies $f(\hat{z}) = v_0\,\hat{z}$ is useful in identifying an optimal solution to problem (1).

Theorem 1 *Let $\hat{z}$ be such that $f(\hat{z}) = v_0\,\hat{z}$. If the assortment $(\hat{S}_1,\ldots,\hat{S}_m)$ satisfies*

$$\sum_{i\in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i)-\hat{z}) \ge f(\hat{z}),$$

then we have $\Pi(\hat{S}_1,\ldots,\hat{S}_m) \ge z^*$, where z^* is the optimal objective value of problem (1).

Proof. We claim that $z^* = \hat{z}$. First, we show that $z^* \ge \hat{z}$. We let $(\tilde{S}_1,\ldots,\tilde{S}_m)$ be an optimal solution to problem (2) when we solve this problem with $z = \hat{z}$. Thus, we have $v_0\,\hat{z} = f(\hat{z}) = \sum_{i\in M} V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i)-\hat{z})$. Focusing on the first and last expressions in this chain of equalities and solving for $\hat{z}$ yields $\hat{z} = \sum_{i\in M} V_i(\tilde{S}_i)^{\gamma_i} R_i(\tilde{S}_i)/(v_0 + \sum_{i\in M} V_i(\tilde{S}_i)^{\gamma_i}) = \Pi(\tilde{S}_1,\ldots,\tilde{S}_m)$. Noting that $(\tilde{S}_1,\ldots,\tilde{S}_m)$ is a feasible solution to problem (1), we have $\Pi(\tilde{S}_1,\ldots,\tilde{S}_m) \le z^*$. Using this inequality with the last chain of equalities, we obtain $z^* \ge \hat{z}$. Second, we show that $z^* \le \hat{z}$. Using $(S_1^*,\ldots,S_m^*)$ to denote an optimal solution to problem (1), we have $z^* = \Pi(S_1^*,\ldots,S_m^*) =$

$\sum_{i \in M} V_i(S_i^*)^{\gamma_i} R_i(S_i^*)/(v_0 + \sum_{i \in M} V_i(S_i^*)^{\gamma_i})$. Focusing on the first and last expressions in this chain of equalities and solving for z^*, we obtain $v_0\, z^* = \sum_{i \in M} V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - z^*)$. In this case, we obtain $v_0\, \hat{z} = f(\hat{z}) \geq \sum_{i \in M} V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - \hat{z}) \geq \sum_{i \in M} V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - z^*) = v_0\, z^*$, where the first inequality follows by the fact that $(S_1^*, \ldots, S_m^*)$ is a feasible solution to problem (2) when we solve this problem with $z = \hat{z}$ and the second inequality uses the fact that $z^* \geq \hat{z}$, which is shown above. The last chain of inequalities indicate that $z^* \leq \hat{z}$, establishing the claim. Since $f(\hat{z}) = v_0\, \hat{z}$, we write the inequality in the theorem as $\sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - \hat{z}) \geq v_0\, \hat{z}$. Solving for $\hat{z}$ in this inequality yields $\hat{z} \leq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i} R_i(\hat{S}_i)/(v_0 + \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}) = \Pi(\hat{S}_1, \ldots, \hat{S}_m)$, in which case, the desired result follows by noting that $\hat{z} = z^*$. $\square$

Theorem 1 suggests the following procedure to obtain an optimal solution to problem (1). We find $\hat{z}$ such that $f(\hat{z}) = v_0\, \hat{z}$ and solve problem (2) with $z = \hat{z}$ to obtain an optimal solution $(\hat{S}_1, \ldots, \hat{S}_m)$. In this case, it is possible to show that $(\hat{S}_1, \ldots, \hat{S}_m)$ is an optimal solution to problem (1). To see this result, we have $f(\hat{z}) = \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - \hat{z})$ by the definition of $(\hat{S}_1, \ldots, \hat{S}_m)$, in which case, $(\hat{S}_1, \ldots, \hat{S}_m)$ satisfies the inequality in Theorem 1 and we obtain $\Pi(\hat{S}_1, \ldots, \hat{S}_m) \geq z^*$. Since the solution $(\hat{S}_1, \ldots, \hat{S}_m)$ is feasible to problem (2), it is feasible to problem (1) as well and we have $\Pi(\hat{S}_1, \ldots, \hat{S}_m) \leq z^*$. Therefore, we have $\Pi(\hat{S}_1, \ldots, \hat{S}_m) = z^*$, establishing that $\Pi(\hat{S}_1, \ldots, \hat{S}_m)$ is an optimal solution to problem (1), as desired. Later in the paper, we show that we can efficiently find $\hat{z}$ that satisfies $f(\hat{z}) = v_0\, \hat{z}$ when we have a cardinality constraint. However, finding such $\hat{z}$ may be difficult when we have a space constraint. The next corollary gives an approximate version of Theorem 1 that does not require finding $\hat{z}$ such that $f(\hat{z}) = v_0\, \hat{z}$. To state this corollary, we let $f^R(\cdot)$ be an approximation to $f(\cdot)$ that satisfies $\alpha\, f^R(z) \geq f(z)$ for all $z \in \Re_+$ for some $\alpha \geq 1$. We do not yet specify how to construct this approximation. We only assume that $f^R(\cdot)$ is a decreasing function similar to $f(\cdot)$ and $f^R(0) \geq 0$, in which case, we can always find $\hat{z} \geq 0$ satisfying $f^R(\hat{z}) = v_0\, \hat{z}$. The next corollary shows that we can use this value of $\hat{z}$ to get an approximation guarantee for problem (1). Its proof is similar to that of Theorem 1 and deferred to Online Appendix A.

Corollary 2 *Let $f^R(\cdot)$ be an approximation to $f(\cdot)$ that satisfies $\alpha\, f^R(z) \geq f(z)$ for all $z \in \Re_+$ for some $\alpha \geq 1$ and $\hat{z}$ be such that $f^R(\hat{z}) = v_0\, \hat{z}$. If the assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ satisfies*

$$\beta \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - \hat{z}) \geq f^R(\hat{z})$$

for some $\beta \geq 1$, then we have $\alpha\, \beta\, \Pi(\hat{S}_1, \ldots, \hat{S}_m) \geq z^*$, where z^* is the optimal objective value of problem (1).

4 Cardinality Constraint

In this section, we consider problem (1) under a cardinality constraint. Thus, we have $C_i(S_i) = |S_i|$ throughout this section. First, we show how to solve problem (2), which allows us to compute $f(z)$ at a particular value of z. Second, we show how to find a value of $\hat{z}$ that satisfies $f(\hat{z}) = v_0\, \hat{z}$. In this case, noting the discussion that follows Theorem 1, we can find an optimal solution to problem (1) by finding a value of $\hat{z}$ that satisfies $f(\hat{z}) = v_0\,\hat{z}$ and solving problem (2) with $z = \hat{z}$.

4.1 Computation at a Particular Point

We consider solving problem (2), which allows us to compute $f(z)$ at a particular value of z. The starting point for our discussion is a result due to Gallego and Topaloglu (2014), who study the problem of maximizing the expected revenue obtained from each customer subject to a separate cardinality constraint on the assortment offered in each nest. In particular, for any $z \in \Re_+$ and $b_i \in \mathbb{Z}_+$, Gallego and Topaloglu (2014) focus on the problem

$$\max_{\substack{S_i \,:\, C_i(S_i) \le b_i, \\ S_i \subset N}} \Big\{ V_i(S_i)^{\gamma_i} \, (R_i(S_i) - z) \Big\} \quad = \quad \max_{\substack{S_i \,:\, |S_i| \le b_i, \\ S_i \subset N}} \Big\{ V_i(S_i)^{\gamma_i} \, (R_i(S_i) - z) \Big\}. \tag{3}$$

The authors construct $O(n^2)$ different orderings of the products such that for any $z \in \Re_+$ and $b_i \in \mathbb{Z}_+$, an optimal solution to problem (3) can be obtained by sorting the products according to one of these orderings and using an assortment that includes some number of earliest products in this ordering. In other words, letting $\{\sigma^g : g \in \mathcal{G}_i\}$ with $|\mathcal{G}_i| = O(n^2)$ be the orderings constructed by Gallego and Topaloglu (2014) and $S_i(\sigma^g, k)$ be the assortment that includes the first k products when the products in nest i are sorted according to the ordering σ^g, for any $z \in \Re_+$ and $b_i \in \mathbb{Z}_+$, an optimal solution to problem (3) can always be found in the collection of assortments $\{S_i(\sigma^g, k) : g \in \mathcal{G}_i,\ k = 0, \ldots, n\}$. Since $|\mathcal{G}_i| = O(n^2)$, there are $O(n^3)$ assortments in this collection. For notational brevity, we use $\mathcal{A}_i = \{S_{it} : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n^3)$ to denote the collection of assortments $\{S_i(\sigma^g, k) : g \in \mathcal{G}_i,\ k = 0, \ldots, n\}$ and the next lemma follows.

Lemma 3 *There exists a collection of assortments $\mathcal{A}_i = \{S_{it} : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n^3)$ such that for any $z \in \Re_+$ and $b_i \in \mathbb{Z}_+$, an optimal solution to problem* (3) *can be found in $\mathcal{A}_i$.*

Lemma 3 allows us to focus only on the assortments in the collections $\mathcal{A}_1, \ldots, \mathcal{A}_m$ in problem (2). In particular, we can write problem (2) equivalently as

$$f(z) \quad = \max_{\substack{(S_1, \ldots, S_m) \,: \\ \sum_{i \in M} C_i(S_i) \le c, \\ S_i \in \mathcal{A}_i \ \forall\, i \in M}} \left\{ \sum_{i \in M} V_i(S_i)^{\gamma_i} (R_i(S_i) - z) \right\}. \tag{4}$$

To see that problems (2) and (4) have the same optimal objective values, we note that problem (2) allows using assortments of the form $(S_1, \ldots, S_m)$ with $S_i \subset N$ for all i, whereas problem (4) allows using assortments of the form $(S_1, \ldots, S_m)$ with $S_i \in \mathcal{A}_i$ for all $i \in M$. Therefore, the optimal objective value of problem (2) is at least as large as the optimal objective value of problem (4). On the other hand, letting $(\hat{S}_1, \ldots, \hat{S}_m)$ be an optimal solution to problem (2) and $C_i(\hat{S}_i) = \hat{b}_i$, since $(\hat{S}_1, \ldots, \hat{S}_m)$ is a feasible solution to problem (2), we have $\sum_{i \in M} \hat{b}_i = \sum_{i \in M} C_i(\hat{S}_i) \le c$. By

Lemma 3, the collection of assortments $\mathcal{A}_i$ includes an optimal solution to problem (3) for any $z \in \Re_+$ and $b_i \in \mathbb{Z}_+$. Using this result with $b_i = \hat{b}_i$ and noting that $\hat{S}_i$ is a feasible solution to problem (3) when we solve this problem with $b_i = \hat{b}_i$, it follows that there exists $\tilde{S}_i \in \mathcal{A}_i$ such that $V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i) - z) \geq V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - z)$ and $C_i(\tilde{S}_i) \leq \hat{b}_i$. Adding the last two inequalities over all $i \in M$, we have $\sum_{i \in M} V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i) - z) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - z)$ and $\sum_{i \in M} C_i(\tilde{S}_i) \leq \sum_{i \in M} \hat{b}_i \leq c$. Since $\tilde{S}_i \in \mathcal{A}_i$ for all $i \in M$, the last two chains of inequalities show that $(\tilde{S}_1, \ldots, \tilde{S}_m)$ is a feasible solution to problem (4) and the objective value provided by this solution for problem (4) is at least as large as the optimal objective value of problem (2). Therefore, the optimal objective value of problem (4) is at least as large as the optimal objective value of problem (2), establishing that problems (2) and (4) have the same optimal objective values.

The discussion above indicates that we can compute $f(z)$ at a particular value of z by solving problem (4), instead of problem (2). However, solving problem (4) in a brute force fashion is still difficult since there are $|\mathcal{A}_1| \times \ldots \times |\mathcal{A}_m|$ different combinations of assortments that we can choose from different nests and the number of such possible combinations grows exponentially fast with the number of nests. To solve problem (4) in a tractable fashion, the critical observation is that the objective function of this problem is separable by the nests. If we offer the assortment S_i in nest i, then we obtain a contribution of $V_i(S_i)^{\gamma_i}(R_i(S_i) - z)$. Problem (4) finds one assortment to offer in each nest to maximize the total contribution subject to the constraint that the total cardinality of the assortments offered over all nests does not exceed c. Therefore, we can solve problem (4) by using a dynamic program. In this dynamic program, the decision epochs correspond to the nests. The state variable in each decision epoch is the remaining capacity left from the earlier nests just before choosing the assortment offered in a particular nest. Finally, the action variable in each decision epoch is the assortment offered in a particular nest. Thus, we can compute $f(z)$ at a particular value of z by solving the dynamic program

$$J_i(b \,|\, z) = \max_{\substack{S_i \,:\, C_i(S_i) \leq b \\ S_i \in \mathcal{A}_i}} \left\{ V_i(S_i)^{\gamma_i}(R_i(S_i) - z) + J_{i+1}(b - C_i(S_i) \,|\, z) \right\}, \tag{5}$$

with the boundary condition $J_{m+1}(\cdot \,|\, z) = 0$. Under a cardinality constraint, we can assume that c is an integer that does not exceed mn, which is the total number of products in all of the nests. Thus, the state space in the dynamic program above is $0, \ldots, mn$. Computing the value functions $\{J_i(b \,|\, z) : b = 0, \ldots, mn, \; i \in M\}$, the value of $J_1(c \,|\, z)$ corresponds to $f(z)$.

The dynamic program in (5) provides an efficient approach for computing $f(z)$ at a particular value of z. Since there are m decision epochs, the state space is $0, \ldots, mn$ and $|\mathcal{A}_i| = O(n^3)$, this dynamic program can be solved in $O(m^2 \, n^4)$ operations. In the next section, we build on the dynamic program to find $\hat{z}$ that satisfies $f(\hat{z}) = v_0 \, \hat{z}$.

4.2 Finding the Fixed Point

We consider the problem of finding $\hat{z}$ that satisfies $f(\hat{z}) = v_0\,\hat{z}$. For this purpose we use the linear programming representation of the dynamic program in (5). A dynamic program with finite states and actions has a linear programming representation. In this linear program, there is one decision variable for each state and decision epoch corresponding to the value function at each state and decision epoch. Inspired by this linear program, we propose solving

$$\min \quad \Theta_1(c) \tag{6}$$

$$\text{st} \quad \Theta_i(b) \geq V_i(S_i)^{\gamma_i}(R_i(S_i) - z) + \Theta_{i+1}(b - C_i(S_i)) \quad \forall\, i \in M,\ b = 0, \ldots, mn,\ S_i \in \mathcal{F}_i(b) \tag{7}$$

$$\Theta_1(c) = v_0\, z, \tag{8}$$

to find $\hat{z}$ satisfying $f(\hat{z}) = v_0\,\hat{z}$. The decision variables are $\Theta = \{\Theta_i(b) : i \in M,\ b = 0, \ldots, mn\}$ and z in the linear program above. We use the convention that $\Theta_{m+1}(b) = 0$ for all $b = 0, \ldots, mn$. The set $\mathcal{F}_i(b)$ is given by $\mathcal{F}_i(b) = \{S_i : C_i(S_i) \leq b,\ S_i \in \mathcal{A}_i\}$, capturing the set of feasible actions at decision epoch i and state b. If we drop the second constraint in problem (6)-(8) and minimize the objective function subject to the first set of constraints for a fixed value of z, then it is well known that the optimal value of the decision variable $\Theta_1(c)$ gives the value function $J_1(b \,|\, z)$ computed through the dynamic program in (5); see Puterman (1994). Interestingly, if we solve problem (6)-(8) as formulated, then the optimal value of the decision variable z gives the value of $\hat{z}$ satisfying $f(\hat{z}) = v_0\,\hat{z}$. The next theorem shows this result.

Theorem 4 *Letting* $(\hat{\Theta}, \hat{z})$ *be an optimal solution to problem* (6)-(8)*,* $\hat{z}$ *satisfies* $f(\hat{z}) = v_0\,\hat{z}$.

Proof. We let $(\hat{S}_1, \ldots, \hat{S}_m)$ be an optimal solution to problem (2) when we solve this problem with $z = \hat{z}$. We define $\{\hat{b}_i : i \in M\}$ as $\hat{b}_1 = c$ and $\hat{b}_{i+1} = \hat{b}_i - C_i(\hat{S}_i)$ so that $\hat{b}_i$ corresponds to the total capacity consumption of the assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ in nests $1, \ldots, i-1$. Since $(\hat{\Theta}, \hat{z})$ is a feasible solution to problem (6)-(8), it satisfies the first set of constraints for state and action $(\hat{b}_i, \hat{S}_i)$ for all $i \in M$. So, we have $\hat{\Theta}_i(\hat{b}_i) \geq V_i(\hat{S}_i)^{\gamma_i}\,(R_i(\hat{S}_i) - \hat{z}) + \hat{\Theta}_{i+1}(\hat{b}_i - C_i(\hat{S}_i))$ for all $i \in M$. Noting that $\hat{b}_{i+1} = \hat{b}_i - C_i(\hat{S}_i)$, adding these inequalities gives $v_0\,\hat{z} = \hat{\Theta}_1(c) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}\,(R_i(\hat{S}_i) - \hat{z}) = f(\hat{z})$, where the first equality uses the fact that $(\hat{\Theta}, \hat{z})$ satisfies the second constraint in problem (6)-(8) and the second equality is by the definition of $(\hat{S}_1, \ldots, \hat{S}_m)$. So, we have $v_0\,\hat{z} \geq f(\hat{z})$. To get a contradiction, assume that $v_0\,\hat{z} > f(\hat{z})$ in the rest of the proof and let $\tilde{z}$ be such that $f(\tilde{z}) = v_0\,\tilde{z}$.

Compute the value functions $J(\tilde{z}) = \{J_i(b \,|\, \tilde{z}) : i \in M,\ b = 0, \ldots, mn\}$ through the dynamic program in (5) with $z = \tilde{z}$. Noting the way the value functions are computed in (5), we have $J_i(b \,|\, \tilde{z}) \geq V_i(S_i)^{\gamma_i}(R_i(S_i) - \tilde{z}) + J_{i+1}(b - C_i(S_i) \,|\, \tilde{z})$ for all $i \in M$, $b = 0, \ldots, mn$ and $S_i \in \mathcal{A}_i$ such that $C_i(S_i) \leq b$, which indicates that $(J(\tilde{z}), \tilde{z})$ satisfies the first set of constraints in problem (6)-(8). Furthermore, we know that $J_1(c \,|\, \tilde{z})$ provides the optimal objective value of problem (4) when this problem is solved with $z = \tilde{z}$, so that $J_1(c \,|\, \tilde{z}) = f(\tilde{z}) = v_0\,\tilde{z}$. Thus, the solution $(J(\tilde{z}), \tilde{z})$ satisfies the second constraint in problem (6)-(8) as well. The optimal objective value $\hat{\Theta}_1(c)$ of problem (6)-(8) must be no larger than the objective value $J_1(c \,|\, \tilde{z})$ at the feasible solution $(J(\tilde{z}), \tilde{z})$, implying $v_0\, \hat{z} = \hat{\Theta}_1(c) \leq J_1(c \,|\, \tilde{z}) = v_0\, \tilde{z}$. So, we obtain $f(\hat{z}) < v_0\, \hat{z} \leq v_0\, \tilde{z} = f(\tilde{z})$, but since $f(\cdot)$ is decreasing, we cannot have $v_0\, \hat{z} \leq v_0\, \tilde{z}$ and $f(\hat{z}) < f(\tilde{z})$, yielding a contradiction. □

To sum up, we can solve the linear program in (6)-(8) to obtain $\hat{z}$ satisfying $f(\hat{z}) = v_0\, \hat{z}$. Since $|\mathcal{F}_i(b)| \leq |\mathcal{A}_i| = O(n^3)$, there are $O(m^2 n)$ decision variables and $\sum_{i \in M} O(mn|\mathcal{A}_i|) = O(m^2 n^4)$ constraints in this linear program. Once we have $\hat{z}$, noting the discussion that follows Theorem 1, we can solve problem (4) with $z = \hat{z}$ to obtain an optimal solution to problem (1). To solve problem (4) with $z = \hat{z}$, we can use the dynamic program in (5). The dynamic program in (5) can be solved in $O(m^2\, n^4)$ operations and the computational effort for solving this dynamic program is negligible when compared with that for solving the linear program in (6)-(8).

5 Space Constraint

In this section, we consider problem (1) under a space constraint. Thus, we have $C_i(S_i) = \sum_{j \in S_i} w_{ij}$ throughout this section. First, we show how to construct an approximation $f^R(\cdot)$ to $f(\cdot)$ such that $2\, f^R(z) \geq f(z)$ for all $z \in \Re_+$. Second, we show how to find $\hat{z}$ satisfying $f^R(\hat{z}) = v_0\, \hat{z}$. Third, we show how to find an assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ such that $2 \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i} (R_i(\hat{S}_i) - \hat{z}) \geq f^R(\hat{z})$ and $\sum_{i \in M} C_i(\hat{S}_i) \leq c$. In this case, we obtain $4\, \Pi(\hat{S}_1, \ldots, \hat{S}_m) \geq z^*$ by Corollary 2 and $(\hat{S}_1, \ldots, \hat{S}_m)$ is a feasible solution to problem (1). Therefore, it follows that $(\hat{S}_1, \ldots, \hat{S}_m)$ is a 4-approximate solution to problem (1) under a space constraint.

5.1 Approximation at a Particular Point

We consider constructing an approximation $f^R(\cdot)$ to $f(\cdot)$ that satisfies $2\, f^R(z) \geq f(z)$ for all $z \in \Re_+$. Similar to our development under the cardinality constraint, the construction of our approximation to $f(\cdot)$ builds on a result that is due to Gallego and Topaloglu (2014). In addition to a separate cardinality constraint on the assortment offered in each nest, the authors study the problem of maximizing the expected revenue obtained from each customer subject to a separate space constraint on the assortment offered in each nest. Within this setting, for any $z \in \Re_+$ and $b_i \in \Re_+$, Gallego and Topaloglu (2014) focus on the problem

$$\max_{\substack{S_i \,:\, C_i(S_i) \leq b_i, \\ S_i \subset N}} \Big\{ V_i(S_i)^{\gamma_i} \, (R_i(S_i) - z) \Big\} \quad = \quad \max_{\substack{S_i \,:\, \sum_{j \in S_i} w_{ij} \leq b_i, \\ S_i \subset N}} \Big\{ V_i(S_i)^{\gamma_i} \, (R_i(S_i) - z) \Big\}. \qquad (9)$$

The authors construct $O(n^2)$ different orderings between the products such that for any $z \in \Re_+$ and $b_i \in \Re_+$, a 2-approximate solution to problem (9) can be obtained by sorting the products according to one of these orderings, dropping the products whose space consumption exceeds b_i from consideration and using an assortment that includes some number of earliest products in this ordering. In other words, we use $\{\sigma^g : g \in \mathcal{G}_i\}$ with $|\mathcal{G}_i| = O(n^2)$ to denote the orderings constructed by Gallego and Topaloglu (2014) and $S_i(\sigma^g, k, b_i)$ to denote the assortment that includes the first k products when the products in nest i are sorted according to the ordering σ^g and the products whose space consumption exceeds b_i are dropped from consideration. In this case, Gallego and Topaloglu (2014) show that for any $z \in \Re_+$ and $b_i \in \Re_+$, a 2-approximate solution to problem (9) can always be found in the collection of assortments $\{S_i(\sigma^g, k, b_i) : \sigma^g \in \mathcal{G}_i,\ k = 0, \ldots, n,\ b_i \in \Re_+\}$. A critical observation is that we can consider only $b_i \in \{w_{i1}, \ldots, w_{in}\}$, rather than $b_i \in \Re_+$, without changing the collection of assortments $\{S_i(\sigma^g, k, b_i) : \sigma^g \in \mathcal{G}_i,\ k = 0, \ldots, n,\ b_i \in \Re_+\}$, since if b_i takes a value other than $\{w_{i1}, \ldots, w_{in}\}$, then we can decrease the value of b_i to the closest element in $\{w_{i1}, \ldots, w_{in}\}$ without changing the set of products whose space consumptions exceed b_i. Therefore, noting that $|\mathcal{G}_i| = O(n^2)$ and we can consider only $b_i \in \{w_{i1}, \ldots, w_{in}\}$, there are $O(n^4)$ assortments in the collection $\{S_i(\sigma^g, k, b_i) : \sigma^g \in \mathcal{G}_i,\ k = 0, \ldots, n,\ b_i \in \Re_+\}$. For notational brevity, we use $\mathcal{A}_i = \{S_{it} : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n^4)$ to denote the collection of assortments $\{S_i(\sigma^g, k, b_i) : g \in \mathcal{G}_i,\ k = 0, \ldots, n,\ b_i \in \Re_+\}$ and obtain the next lemma. This lemma becomes useful when constructing our approximation $f^R(\cdot)$ to $f(\cdot)$.

Lemma 5 *There exists a collection of assortments $\mathcal{A}_i = \{S_{it} : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n^4)$ such that for any $z \in \Re_+$ and $b_i \in \Re_+$, a 2-approximate solution to problem (9) can be found in $\mathcal{A}_i$.*

We construct our approximation to $f(\cdot)$ by focusing only on the assortments in the collections $\mathcal{A}_1, \ldots, \mathcal{A}_m$. Using the decision variables $x = \{x_i(S_i) : i \in M,\ S_i \in \mathcal{A}_i\}$, we define $f^R(\cdot)$ as

$$f^R(z) \quad = \quad \max \quad \sum_{i \in M} \sum_{S_i \in \mathcal{A}_i} V_i(S_i)^{\gamma_i} (R_i(S_i) - z)\, x_i(S_i) \tag{10}$$

$$\text{st} \quad \sum_{i \in M} \sum_{S_i \in \mathcal{A}_i} C_i(S_i)\, x_i(S_i) \le c \tag{11}$$

$$\sum_{S_i \in \mathcal{A}_i} x_i(S_i) = 1 \qquad \forall\, i \in M \tag{12}$$

$$x_i(S_i) \ge 0 \qquad \forall\, i \in M,\ S_i \in \mathcal{A}_i, \tag{13}$$

which corresponds to the optimal objective value of a linear program with $\sum_{i \in M} O(|\mathcal{A}_i|) = O(mn^4)$ decision variables and $O(m)$ constraints. In problem (2), each assortment S_i offered in nest i provides a contribution of $V_i(S_i)^{\gamma_i}(R_i(S_i) - z)$. This problem finds one assortment $S_i \subset N$ to offer in each nest i such that the total contribution over all nests is maximized and the total capacity consumption over all nests does not exceed c. Similarly, each assortment S_i offered in nest i provides a contribution of $V_i(S_i)^{\gamma_i}(R_i(S_i) - z)$ in problem (10)-(13). If we impose integrality constraints on the decision variables in problem (10)-(13), then noting the second set of constraints, this problem finds one assortment $S_i \in \mathcal{A}_i$ to offer in each nest i such that the total contribution over all nests in maximized and the total capacity consumption over all nests does not exceed c. We note that $f^R(z)$ is decreasing in z. Also, we assume that $C_i(S_i) \le c$ for all $i \in M$ and $S_i \in \mathcal{A}_i$. If $C_i(S_i) > c$ for some $i \in M$ and $S_i \in \mathcal{A}_i$, then we can drop this assortment from $\mathcal{A}_i$ since using this assortment in problem (1) would yield an infeasible solution.

It is possible to use Lemma 5 to show that our approximation $f^R(\cdot)$ to $f(\cdot)$ satisfies $2\, f^R(z) \ge f(z)$ for all $z \in \Re_+$. To see this result, we let $(\hat{S}_1, \ldots, \hat{S}_m)$ be an optimal solution to problem

(2) and $\hat{b}_i = C_i(\hat{S}_i)$. Since $(\hat{S}_1, \ldots, \hat{S}_m)$ is a feasible solution to problem (2), we have $\sum_{i \in M} \hat{b}_i = \sum_{i \in M} C_i(\hat{S}_i) \leq c$. Lemma 5 indicates that the collection of assortments $\mathcal{A}_i$ includes a 2-approximate solution to problem (9) for any $z \in \Re_+$ and $b_i \in \Re_+$. Using this result with $b_i = \hat{b}_i$ and noting the fact that $\hat{S}_i$ is a feasible solution to problem (9) when this problem is solved with $b_i = \hat{b}_i$, it follows that there exists $\tilde{S}_i \in \mathcal{A}_i$ such that $2\, V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i) - z) \geq V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - z)$ and $C_i(\tilde{S}_i) \leq \hat{b}_i$. Adding the last two inequalities over all $i \in M$, we have $2 \sum_{i \in M} V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i) - z) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - z)$ and $\sum_{i \in M} C_i(\tilde{S}_i) \leq \sum_{i \in M} \hat{b}_i \leq c$.

To obtain the desired result, we define the solution $\tilde{x}$ to problem (10)-(13) as $\tilde{x}_i(\tilde{S}_i) = 1$ for all $i \in M$ and $\tilde{x}_i(S_i) = 0$ for all $i \in M$ and $S_i \in \mathcal{A}_i \setminus \{\tilde{S}_i\}$. The solution $\tilde{x}$ is feasible to problem (10)-(13) since the definition of $\tilde{x}$ implies that $\sum_{i \in M} \sum_{S_i \in \mathcal{A}_i} C_i(S_i)\, \tilde{x}_i(S_i) = \sum_{i \in M} C_i(\tilde{S}_i) \leq c$ and $\sum_{S_i \in \mathcal{A}_i} \tilde{x}_i(S_i) = \tilde{x}_i(\tilde{S}_i) = 1$. Furthermore, the objective value provided by the solution $\tilde{x}$ for problem (10)-(13) satisfies $\sum_{i \in M} \sum_{S_i \in \mathcal{A}_i} V_i(S_i)^{\gamma_i}(R_i(S_i) - z)\, \tilde{x}_i(S_i) = \sum_{i \in M} V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i) - z) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - z)/2 = f(z)/2$, where the first equality follows from the definition of $\tilde{x}$, the inequality uses the fact that $2 \sum_{i \in M} V_i(\tilde{S}_i)^{\gamma_i}(R_i(\tilde{S}_i) - z) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - z)$ shown in the previous paragraph and the second equality is by the fact that $(\hat{S}_1, \ldots, \hat{S}_m)$ is an optimal solution to problem (2). Thus, there exists a feasible solution to problem (10)-(13) providing an objective value for this problem that is at least $f(z)/2$, which implies that the optimal objective value $f^R(z)$ of problem (10)-(13) satisfies $f^R(z) \geq f(z)/2$, establishing the desired result.

5.2 Finding the Fixed Point

We consider the problem of finding the value of $\hat{z}$ that satisfies $f^R(\hat{z}) = v_0\, \hat{z}$. Noting that $f^R(z)$ is given by the optimal objective value of the linear program in (10)-(13), we use the dual of this problem to find the value of $\hat{z}$ that satisfies $f^R(\hat{z}) = v_0\, \hat{z}$. In particular, associating the dual variables Δ and $y = \{y_i : i \in M\}$ respectively with the two sets of constraints in problem (10)-(13), we propose solving the linear program

$$\min \quad c\,\Delta + \sum_{i \in M} y_i \tag{14}$$

$$\text{st} \quad C_i(S_i)\,\Delta + y_i \geq V_i(S_i)^{\gamma_i}(R_i(S_i) - z) \qquad \forall\, i \in M,\ S_i \in \mathcal{A}_i \tag{15}$$

$$c\,\Delta + \sum_{i \in M} y_i = v_0\, z \tag{16}$$

$$\Delta \geq 0,\ y_i \text{ is free},\ z \text{ is free} \qquad \forall\, i \in M \tag{17}$$

to find $\hat{z}$ satisfying $f^R(\hat{z}) = v_0\, \hat{z}$. The decision variables are Δ, y and z in the problem above. If we drop the second constraint in problem (14)-(17) and minimize the objective function subject to the first set of constraints for a fixed value of z, then this problem corresponds to the dual of problem (10)-(13). With the second constraint added, problem (14)-(17) allows us to find $\hat{z}$ that satisfies $f^R(\hat{z}) = v_0\, \hat{z}$, as shown in the next theorem.

Theorem 6 *Letting $(\hat{\Delta}, \hat{y}, \hat{z})$ be an optimal solution to problem (14)-(17), $\hat{z}$ satisfies $f^R(\hat{z}) = v_0\, \hat{z}$.*

Proof. Associating the dual variables Δ and $y = \{y_i : i \in M\}$ with the two sets of constraints in problem (10)-(13), the dual of this problem is

$$f^R(z) \;=\; \min \quad c\,\Delta + \sum_{i \in M} y_i \tag{18}$$

$$\text{st} \quad C_i(S_i)\,\Delta + y_i \geq V_i(S_i)^{\gamma_i}(R_i(S_i) - z) \qquad \forall\, i \in M,\ S_i \in \mathcal{A}_i \tag{19}$$

$$\Delta \geq 0,\ y_i \text{ is free} \qquad \forall\, i \in M. \tag{20}$$

Therefore, the solution $(\hat{\Delta}, \hat{y})$ is feasible to problem (18)-(20) when we solve this problem with $z = \hat{z}$, which implies that $f^R(\hat{z}) \leq c\,\hat{\Delta} + \sum_{i \in M} \hat{y}_i = v_0\,\hat{z}$, where the equality follows from the fact that $(\hat{\Delta}, \hat{y}, \hat{z})$ is a feasible solution to problem (14)-(17). To get a contradiction, assume that the last inequality is strict so that $f^R(\hat{z}) < v_0\,\hat{z}$. We let $\tilde{z}$ be such that $f^R(\tilde{z}) = v_0\,\tilde{z}$ and $(\tilde{\Delta}, \tilde{y})$ be an optimal solution to problem (18)-(20) when we solve this problem with $z = \tilde{z}$. Thus, we get $v_0\,\tilde{z} = f^R(\tilde{z}) = c\,\tilde{\Delta} + \sum_{i \in M} \tilde{y}_i$, which indicates that $(\tilde{\Delta}, \tilde{y}, \tilde{z})$ is a feasible solution to problem (14)-(17). In this case, it follows that $v_0\,\tilde{z} = f^R(\tilde{z}) = c\,\tilde{\Delta} + \sum_{i \in M} \tilde{y}_i \geq c\,\hat{\Delta} + \sum_{i \in M} \hat{y}_i = v_0\,\hat{z} > f^R(\hat{z})$, where the first inequality is by the fact that $(\tilde{\Delta}, \tilde{y}, \tilde{z})$ is a feasible, but not necessarily an optimal solution to problem (14)-(17). The last chain of inequalities yields $v_0\,\tilde{z} \geq v_0\,\hat{z}$ and $f^R(\tilde{z}) > f^R(\hat{z})$, which contradict the fact that $f^R(\cdot)$ is a decreasing function. □

5.3 Construction of an Approximate Assortment

By the earlier discussion in this section, our approximation $f^R(\cdot)$ to $f(\cdot)$ satisfies $2\,f^R(z) \geq f(z)$ for all $z \in \Re_+$. Furthermore, we can find the value of $\hat{z}$ that satisfies $f^R(\hat{z}) = v_0\,\hat{z}$ by solving the linear program in (14)-(17). In the remainder of this section, we consider the problem of finding an assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ such that $2\,\sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(R_i(\hat{S}_i) - \hat{z}) \geq f^R(\hat{z})$ and $\sum_{i \in M} C_i(\hat{S}_i) \leq c$. In this case, Corollary 2 implies that we have $4\,\Pi(\hat{S}_1, \ldots, \hat{S}_m) \geq z^*$ and $(\hat{S}_1, \ldots, \hat{S}_m)$ is a feasible solution to problem (1). So, $(\hat{S}_1, \ldots, \hat{S}_m)$ is a 4-approximate solution to problem (1).

It is a simple exercise in linear programming duality to show that any basic optimal solution to problem (10)-(13) includes at most two fractional components; see Sinha and Zoltners (1979). We let $\hat{x}$ be a basic optimal solution to problem (10)-(13) when we solve this problem with $z = \hat{z}$. We make two observations. First, if $\hat{x}_{i'}(P_{i'}) \in (0, 1]$ for some nest $i' \in M$ and assortment $P_{i'} \in \mathcal{A}_{i'}$, then noting the second set of constraints in problem (10)-(13), there must be some other assortment $Q_{i'} \in \mathcal{A}_{i'}$ such that $\hat{x}_{i'}(Q_{i'}) \in [0, 1)$ as well. Second, since $\hat{x}$ has as most two fractional components, there can be no other fractional component of $\hat{x}$. In this case, noting the second set of constraints in problem (10)-(13) once more, it follows that for each nest $i \in M \setminus \{i'\}$, there exists a single assortment $\tilde{S}_i$ such that $\hat{x}_i(\tilde{S}_i) = 1$. Therefore $\{\hat{x}_i(\tilde{S}_i) : i \in M \setminus \{i'\}\} \cup \{\hat{x}_{i'}(P_{i'})\} \cup \{\hat{x}_{i'}(Q_{i'})\}$ includes all components of $\hat{x}$ taking strictly positive values.

In the rest of the discussion, we assume that the basic optimal solution $\hat{x}$ has two fractional components. In particular, noting the second set of constraints in problem (10)-(13), $\hat{x}$ cannot have one fractional component and the result holds in a straightforward fashion when $\hat{x}$ has no fractional components. As described in the previous paragraph, if the basic optimal solution $\hat{x}$ to problem (10)-(13) has two fractional components, then there exist some nest $i' \in M$ and assortments $P_{i'}, Q_{i'} \in \mathcal{A}_{i'}$ such that $\hat{x}_{i'}(P_{i'}), \hat{x}_{i'}(Q_{i'}) \in (0,1)$ and there is no other fractional component of $\hat{x}$. Without loss of generality, we assume that $C_{i'}(P_{i'}) \leq C_{i'}(Q_{i'})$. Furthermore, for each nest $i \in M \setminus \{i'\}$, there exists a single assortment $\tilde{S}_i$ such that $\hat{x}_i(\tilde{S}_i) = 1$. Using the solution $\hat{x}$, we construct two assortments $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$ as follows. The first one of these assortments is constructed as $(\hat{S}_1^1, \ldots, \hat{S}_m^1) = (\tilde{S}_1, \ldots \tilde{S}_{i'-1}, P_{i'}, \tilde{S}_{i'+1}, \ldots, \tilde{S}_m)$. In other words, the assortment $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ uses the components of the solution $\hat{x}$ that take value one, along with the fractional component of the solution $\hat{x}$ with the smaller capacity consumption. We construct the second assortment as $(\hat{S}_1^2, \ldots, \hat{S}_m^2) = (\emptyset, \ldots, \emptyset, Q_{i'}, \emptyset, \ldots, \emptyset)$, offering the subset $Q_{i'}$ in nest i', but offering empty subsets in all of the other nests. A crucial observation is that the two assortments $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$ as defined above collectively include all components of the solution $\hat{x} = \{\hat{x}_i(S_i) : i \in M,\ S_i \in \mathcal{A}_i\}$ taking a strictly positive value. In this case, we get

$$\begin{aligned} f^R(\hat{z}) &= \sum_{i \in M} \sum_{S_i \in \mathcal{A}_i} V_i(S_i)^{\gamma_i} (R_i(S_i) - \hat{z})\, \hat{x}_i(S_i) \leq \sum_{i \in M} V_i(\hat{S}_i^1)^{\gamma_i} (R_i(\hat{S}_i^1) - \hat{z}) + \sum_{i \in M} V_i(\hat{S}_i^2)^{\gamma_i} (R_i(\hat{S}_i^2) - \hat{z}) \\ &\leq 2 \max\left\{ \sum_{i \in M} V_i(\hat{S}_i^1)^{\gamma_i} (R_i(\hat{S}_i^1) - \hat{z}), \sum_{i \in M} V_i(\hat{S}_i^2)^{\gamma_i} (R_i(\hat{S}_i^2) - \hat{z}) \right\}, \end{aligned}$$

where the first inequality is by the fact that if $\hat{x}_i(S_i) > 0$ for some $i \in M$ and $S_i \in \mathcal{A}_i$, then we have $S_i = \hat{S}_i^1$ or $\hat{S}_i = \hat{S}_i^2$. Therefore, the chain of inequalities above shows that if we choose $(\hat{S}_1, \ldots, \hat{S}_m)$ as one of the assortments $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$, then $(\hat{S}_1, \ldots, \hat{S}_m)$ satisfies $2 \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i} (R_i(\hat{S}_i) - \hat{z}) \geq f^R(\hat{z})$. Furthermore, we note that both of the solutions $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$ are feasible to problem (1). The solution $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$ is feasible since this solution only offers $Q_{i'}$ in nest i' and we have $\sum_{i \in M} C_i(\hat{S}_i^2) = C_{i'}(Q_{i'}) \leq c$, where the last inequality uses the assumption that $C_i(S_i) \leq c$ for all $i \in M$ and $S_i \in \mathcal{A}_i$. To see the feasibility of the solution $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ to problem (1), we observe that $\sum_{i \in M} C_i(\hat{S}_i^1) = \sum_{i \in M \setminus \{i'\}} C_i(\tilde{S}_i) + C_{i'}(P_{i'}) \leq \sum_{i \in M \setminus \{i'\}} C_i(\tilde{S}_i) + C_{i'}(P_{i'})\, \hat{x}_{i'}(P_{i'}) + C_{i'}(Q_{i'})\, \hat{x}_{i'}(Q_{i'}) = \sum_{i \in M} \sum_{S_i \in \mathcal{A}_i} C_i(S_i)\, \hat{x}_i(S_i) \leq c$, where the first inequality uses the fact that $C_{i'}(P_{i'}) \leq C_{i'}(Q_{i'})$ and $\hat{x}_{i'}(P_{i'}) + \hat{x}_{i'}(Q_{i'}) = 1$ by the second set of constraints in problem (10)-(13) and the second equality uses the fact that $\{\hat{x}_i(\tilde{S}_i) : i \in M \setminus \{i'\}\} \cup \{\hat{x}_{i'}(P_{i'})\} \cup \{\hat{x}_{i'}(Q_{i'})\}$ correspond to all components of $\hat{x}$ taking strictly positive values.

To sum up, we can solve the linear program in (14)-(17) to find $\hat{z}$ satisfying $f^R(\hat{z}) = v_0\, \hat{z}$. This linear program has $O(m)$ decision variables and $\sum_{i \in M} O(|\mathcal{A}_i|) = O(mn^4)$ constraints. Once we have $\hat{z}$, we can solve problem (10)-(13) with $z = \hat{z}$ to obtain an optimal solution $\hat{x}$ and construct the assortments $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$ as in the previous paragraph. If we choose $(\hat{S}_1, \ldots, \hat{S}_m)$ as one of the assortments $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$, then $(\hat{S}_1, \ldots, \hat{S}_m)$ satisfies $2 \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i} (R_i(\hat{S}_i) - \hat{z}) \geq f^R(\hat{z})$ and $(\hat{S}_1, \ldots, \hat{S}_m)$ is a feasible solution to problem (1). Since our approximation $f^R(\cdot)$ to $f(\cdot)$ satisfies $2\, f^R(z) \geq f(z)$ for all $z \in \Re_+$, by Corollary 2, we obtain $4\, \Pi(\hat{S}_1, \ldots, \hat{S}_m) \geq z^*$, so that $(\hat{S}_1, \ldots, \hat{S}_m)$ is a 4-approximate solution to problem (1). Thus, if we check the expected revenue provided by each one of the assortments $(\hat{S}_1^1, \ldots, \hat{S}_m^1)$ and $(\hat{S}_1^2, \ldots, \hat{S}_m^2)$ for problem (1) and pick the best one, then we obtain a 4-approximate solution to problem (1).

6 Numerical Experiments

In this section, our goal is to numerically test the performance of the 4-approximation algorithm described in Section 5. Since we can obtain the optimal solution to problem (1) when we have a cardinality constraint, we do not provide numerical experiments under a cardinality constraint.

6.1 Numerical Setup

In our numerical experiments, we randomly generate a large number of problem instances. We generate each problem instance by using the following procedure. We set the number of nests as $m = 3$ or $m = 5$ and the number of products in each nest as $n = 15$ or $n = 30$. To come up with the revenues and preference weights of the products, we sample C_{ij} from the uniform distribution over $[0, 1]$. Similarly, we sample X_{ij} and Y_{ij} from the uniform distribution over $[0.75, 1.25]$. We set the revenue and preference weight of product j in nest i respectively as $r_{ij} = 10 \times C_{ij}^2 \times X_{ij}$ and $v_{ij} = 10 \times (1 - C_{ij}) \times Y_{ij}$. Through C_{ij}, we ensure that the products having larger revenues generally tend to have smaller preference weights, indicating more expensive products tend to be less attractive. Squaring C_{ij} in the expression for r_{ij} skews the distribution of the revenues so that we have a small number of products with large revenues. Through X_{ij} and Y_{ij}, we incorporate idiosyncratic noise into the revenues and preference weights so that not all products with large revenues have small preference weights. We sample the dissimilarity parameter γ_i for each nest i from the uniform distribution over $[0.25, 0.75]$. We set the preference weight v_0 of the no purchase option such that the probability of no purchase is 0.4 even when we offer all products in all nests. We sample the space requirement w_{ij} of product j in nest i from the uniform distribution over $[1, 10]$. We set the capacity availability as $c = \kappa \sum_{i \in M} \sum_{j \in N} w_{ij}$, corresponding to a κ fraction of the total space consumption of all products in all nests. We use $\kappa = 0.1$, $\kappa = 0.15$ or $\kappa = 0.2$.

We vary (m, n, κ) over $\{3, 5\} \times \{15, 30\} \times \{0.1, 0.15, 0.2\}$ to get 12 parameter combinations. In each parameter combination, we generate 5,000 individual problem instances by using the approach in the previous paragraph. For each problem instance, we use the approach in Section 5 to obtain a 4-approximate solution. In Online Appendix B, we also give a linear program that provides an upper bound on the optimal expected revenue for a particular problem instance. To assess the quality of the 4-approximate solution, we check the gap between the expected revenue from the 4-approximate solution and the upper bound on the optimal expected revenue.

6.2 Numerical Results

Our numerical results are given in Table 1. The first column in this table shows the parameter combinations by using (m, n, κ). We recall that we generate 5,000 individual problem instances in each parameter combination. For each problem instance, we use the approach in Section 5 to obtain a 4-approximate solution to problem (1). We use Rev^p to denote the expected revenue obtained by the 4-approximate solution for problem instance p. Using Bnd^p to denote the upper bound on the optimal total expected revenue for problem instance p, the second column in Table 1 gives the average percent gap between Rev^p and Bnd^p, where average is taken over all 5,000 problem instances in a parameter combination. The third and fourth columns respectively give the 95th percentile and maximum of the percent gaps between Rev^p and Bnd^p over all 5,000 problem instances in a particular parameter combination. Thus, the second, third and fourth columns respectively give the average, 95th percentile and maximum of the data $\{100 \times (\mathsf{Bnd}^p - \mathsf{Rev}^p)/\mathsf{Bnd}^p : p = 1, \ldots, 5{,}000\}$. In this way, these columns give an indication of the optimality gaps of the 4-approximate solutions. The fifth column gives the number of products in the 4-approximate solution, averaged over all 5,000 problem instances. The sixth column gives the number of problem instances for which the percent gap between Rev^p and Bnd^p is less than 1%. Similarly, the seventh, eighth and ninth columns give the number of problem instances where the percent gap between Rev^p and Bnd^p is respectively less than 2.5%, 5% and 10%. The tenth and eleventh columns attempt to give a feel for the tightness of the space constraint. The tenth column shows the average number of products in the optimal assortment when there are no space constraints. The eleventh column shows the number of problem instances for which this unconstrained solution violates the space constraint. For comparison, we also use a greedy algorithm to obtain a heuristic solution to problem (1). The greedy algorithm starts with the empty assortment, finds the product that provides the largest improvement in the expected revenue per unit of space consumption and adds this product to the assortment, until there is no available space or no improvement in the expected revenue. The twelfth, thirteenth and fourteenth columns respectively show the average, 95th percentile and maximum of the percent gaps between the expected revenue obtained by the greedy algorithm and the upper bound on the optimal expected revenue. So, using Gre^p to denote the expected revenue obtained by the greedy algorithm for problem instance p, these three columns show the average, 95th percentile and maximum of the data $\{100 \times (\mathsf{Bnd}^p - \mathsf{Gre}^p)/\mathsf{Bnd}^p : p = 1, \ldots, 5{,}000\}$.

Our results indicate that our 4-approximation algorithm performs quite well. Over all problem instances, the average optimality gap of this algorithm is no larger than 1.56%. In 58,034 out of all 60,000 problem instances, the optimality gaps of the 4-approximate solutions are less than 5%. As a general trend, the optimality gaps tend to get smaller as κ gets larger. As κ gets larger, the capacity availability gets larger and each product occupies a smaller fraction of the available capacity. So, problem instances where each product occupies a smaller fraction of the available capacity appear to be easier to approximate. This observation is aligned with the intuition that the linear programming relaxation of a knapsack problem becomes tighter as each item occupies a smaller fraction of the knapsack capacity. In particular, it is known that if each item occupies no larger than a fraction ϵ of the knapsack capacity, then the optimal objective value of the linear programming relaxation exceeds the optimal objective value of a knapsack problem by at most a factor of $1/(1-\epsilon)$. The most problematic parameter combination in Table 1 is $(3, 15, 0.1)$, corresponding to a small value of κ with $\kappa = 0.1$. Even for this parameter combination, in more than 75% of the problem instances, the optimality gap of the 4-approximate solution is no larger than 5%. Also, we note that the reported optimality gaps are pessimistic estimates, since these optimality gaps are obtained by

Param. Combin. (m, n, κ)	% Gap btw. Rev^p, Bnd^p Avg.	95th	Max.	Avg. Assr. Size	No. Prob. with Certain % Gap btw. Rev^p, Bnd^p 1%	2.5%	5%	10%	Uncn. Assr. Size	No. Capac. Prbs.	% Gap btw. Gre^p, Bnd^p Avg.	95th	Max.
$(3, 15, 0.20)$	1.48	3.86	8.68	9.89	2,196	4,090	4,939	5,000	17.02	4,899	2.78	8.16	28.11
$(3, 15, 0.15)$	2.21	5.62	14.36	8.06	1,423	3,258	4,610	4,996	17.02	4,985	3.46	9.61	31.87
$(3, 15, 0.10)$	3.35	8.38	21.53	6.13	861	2,322	3,875	4,898	17.02	4,999	4.51	11.57	36.78
$(3, 30, 0.20)$	0.82	1.95	4.95	20.09	3,430	4,924	5,000	5,000	33.45	4,941	1.43	4.35	13.25
$(3, 30, 0.15)$	1.19	2.75	7.78	16.36	2,464	4,622	4,996	5,000	33.45	4,996	1.82	5.35	15.07
$(3, 30, 0.10)$	1.79	4.06	7.36	12.45	1,551	3,740	4,926	5,000	33.45	5,000	2.39	6.55	17.50
$(5, 15, 0.20)$	1.05	2.34	5.18	16.88	2,776	4,828	4,999	5,000	28.47	4,982	2.31	6.03	13.70
$(5, 15, 0.15)$	1.54	3.41	6.05	13.78	1,798	4,163	4,987	5,000	28.47	4,999	2.83	7.00	15.10
$(5, 15, 0.10)$	2.26	5.16	10.33	10.53	1,086	3,196	4,704	4,999	28.47	5,000	3.44	8.13	17.30
$(5, 30, 0.20)$	0.68	1.37	4.52	33.82	4,097	4,992	5,000	5,000	56.06	4,993	1.30	3.27	7.75
$(5, 30, 0.15)$	0.96	1.88	3.96	27.57	3,010	4,972	5,000	5,000	56.06	5,000	1.63	3.99	9.70
$(5, 30, 0.10)$	1.35	2.73	5.09	21.04	1,907	4,606	4,998	5,000	56.06	5,000	2.03	4.87	13.53
Avg./Total	1.56	3.63	8.32		26,599	49,713	58,034	59,893			2.49	6.57	18.30

Table 1: Performance of the 4-approximate solutions and the greedy algorithm on 5,000 randomly generated problem instances.

comparing the expected revenue from an assortment with an upper bound on the optimal expected revenue, rather than the optimal expected revenue itself. The greedy algorithm performs noticeably worse than the 4-approximation. There are parameter combinations such as $(3, 15, 0.20)$, where the 95th percentile of the optimality gaps from the 4-approximation algorithm is 3.86%, but the 95th percentile of the optimality gaps from the greedy algorithm is 8.16%. The running times for the 4-approximation algorithm are reasonable. We use Java 1.6.033 on an Intel Xeon 2.00 GHz CPU and Gurobi 5.1.0 as the linear programming solver. For the largest problem instances with $m = 5$ and $n = 30$, the average running time for the 4-approximation algorithm is 3.56 seconds.

7 Conclusions

We gave tractable methods to solve assortment problems under the nested logit model when there is a cardinality or space constraint on the assortment offered over all nests. As a direction for future research, the 4-approximation algorithm does not provide any guidance as to how we can obtain better solutions if we are willing to increase the computational effort. Gallego and Topaloglu (2014) show how to generate candidate assortments that tradeoff running time with solution quality. Furthermore, Frieze and Clarke (1984) develop approximations to multiple choice knapsack problems that tradeoff running time with solution quality. It is interesting to see whether we can join these two approaches to develop an approximation algorithm for the assortment problem under a space constraint that tradeoff running time with solution quality.

Acknowledgements. We thank the area editor, the associate editor and the two anonymous referees for their valuable comments that improved the paper in many ways.

References

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

Correa, J. R., Fernandes, C. G. and Wakabayashi, Y. (2010), ‘Approximating a class of combinatorial problems with rational objective function', *Mathematical Programming* **124**(1-2), 255–269.

Davis, J., Gallego, G. and Topaloglu, H. (2013), 'Assortment optimization under variants of the nested logit model', *Operations Research* (to appear).

Desir, A. and Goyal, V. (2013), An FPTAS for capacity constrained assortment optimization, Technical report, Columbia University, School of Industrial Engineering and Operations Research.

Frieze, A. M. and Clarke, M. R. B. (1984), 'Approximation algorithms for the m-dimensional 0-1 knapsack problem: Worst-case and probabilistic analyses', *European Journal of Operational Research* **15**, 100–109.

Gallego, G., Iyengar, G., Phillips, R. and Dubey, A. (2004), Managing flexible products on a network, Technical report, Columbia University.

Gallego, G. and Topaloglu, H. (2014), 'Constrained assortment optimization for the nested logit model', *Management Science* (to appear).

Grigolon, L. and Verboren, F. (2013), 'Nested logit or random coefficients logit? A comparison of alternative discrete choice models of product differentiation', *The Review of Economics and Statistics* (to appear).

Hashizume, S., Fukushima, M., Katoh, N. and Ibaraki, T. (1987), 'Approxiation algorithms for combinatorial fractional programming problems', *Mathematical Programming* **37**(3), 255–267.

Li, G. and Rusmevichientong, P. (2014), 'A greedy algorithm for assortment optimization in the two-level nested logit model', *Operations Research Letters Letters* **42**(5), 319–324.

Luce, R. D. (1959), *Individual Choice Behavior: A Theoretical Analysis*, Wiley, New York, NY.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, *in* P. Zarembka, ed., '*Frontiers in Economics*', Academic Press, pp. 105–142.

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

Mendez-Diaz, I., Bront, J. J. M., Vulcano, G. and Zabala, P. (2010), 'A branch-and-cut algorithm for the latent-class logit assortment problem', *Discrete Applied Mathematics* **36**, 383–390.

Mittal, S. and Schulz, A. S. (2013), 'A general framework for designing approximation schemes for combinatorial optimization problems with many objectives combined into one', *Operations Research* **61**(2), 389–397.

Puterman, M. L. (1994), *Markov Decision Processes*, John Wiley and Sons, Inc., New York.

Rusmevichientong, P., Shen, Z.-J. M. and Shmoys, D. B. (2009), 'A PTAS for capacitated sum-of-ratios optimization', *Operations Research Letters* **37**(4), 230–238.

Rusmevichientong, P., Shen, Z.-J. M. and Shmoys, D. B. (2010), 'Dynamic assortment optimization with a multinomial logit choice model and capacity constraint', *Operations Research* **58**(6), 1666–1680.

Rusmevichientong, P., Shmoys, D. B., Tong, C. and Topaloglu, H. (2013), 'Assortment optimization under the multinomial logit model with random choice parameters', *Production and Operations Management* (to appear).

Sinha, P. and Zoltners, A. A. (1979), 'The multiple-choice knapsack problem', *Operations Research* **27**(3), 503–515.

Slade, M. E. (2009), Merge-simulations of unilateral effects: What can we learn from the UK brewing industry?, *in* B. Lyons, ed., '*Cases in European Competition Policy: The Economic Analysis*', Cambridge University Press, Cambridge.

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

Train, K. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, UK.

Train, K. E., McFadden, D. L. and Ben-Akiva, M. (1987), 'The demand for local telephone service: A fully discrete model of residential calling patterns and service choices', *The RAND Journal of Economics* **18**(1), 109–123.

Wang, R. (2012), 'Capacitated assortment and price optimization under the multinomial logit model', *Operations Research Letters* **40**, 492–497.

Wang, R. (2013), 'Assortment management under the generalized attraction model with a capacity constraint', *Journal of Revenue and Pricing Management* **12**(3), 254–270.

A Online Appendix: Proof of Corollary 2

We let $(S_1^*, \ldots, S_m^*)$ be an optimal solution to problem (1) so that $z^* = \Pi(S_1^*, \ldots, S_m^*) = \sum_{i \in M} V_i(S_i^*)^{\gamma_i} R_i(S_i^*)/(v_0 + \sum_{i \in M} V_i(S_i^*)^{\gamma_i})$. Focusing on the first and last terms in this chain of equalities and solving for z^*, we obtain $v_0\, z^* = \sum_{i \in M} V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - z^*)$. Since $(S_1^*, \ldots, S_m^*)$ is a feasible solution to problem (2) when we solve this problem with $z = z^*$, we obtain $f(z^*) \geq \sum_{i \in M} V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - z^*)$, in which case, using the last equality, we have $f(z^*) \geq v_0\, z^*$. We claim that $\alpha\, \hat{z} \geq z^*$. To get a contradiction, assume that $\alpha\, \hat{z} < z^*$. In this case, we obtain $f(z^*) \geq v_0\, z^* > \alpha\, v_0\, \hat{z} = \alpha\, f^R(\hat{z}) \geq f(\hat{z})$, where the equality follows from the definition of $\hat{z}$. Since $f(\cdot)$ is decreasing, having $f(z^*) \geq f(\hat{z})$ implies that $z^* \leq \hat{z} \leq \alpha\, \hat{z}$, which contradicts the assumption that $\alpha\, \hat{z} < z^*$ and the claim follows. To obtain the desired result, we observe that $\sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(\alpha\, \beta\, R_i(\hat{S}_i) - z^*) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(\alpha\, \beta\, R_i(\hat{S}_i) - \beta\, z^*) \geq \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}(\alpha\, \beta\, R_i(\hat{S}_i) - \alpha\, \beta\, \hat{z}) \geq \alpha\, f^R(\hat{z}) = \alpha\, v_0\, \hat{z} \geq v_0\, z^*$, where the first inequality follows from the fact that $\beta \geq 1$, the second inequality is by the fact that $\alpha\, \hat{z} \geq z^*$ and the third inequality follows from the inequality given in the corollary. Focusing on the first and last expressions in the last chain of inequalities and solving for z^*, we obtain $z^* \leq \alpha\, \beta \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i} R_i(\hat{S}_i)/(v_0 + \sum_{i \in M} V_i(\hat{S}_i)^{\gamma_i}) = \alpha\, \beta\, \Pi(\hat{S}_1, \ldots, \hat{S}_m)$.

B Online Appendix: An Upper Bound

The approach in Section 5 obtains a 4-approximate solution under a space constraint, indicating that this approach never performs arbitrarily badly. However, knowing that a solution provides at least a quarter of the optimal expected revenue may not be thoroughly satisfying from a practical perspective. In this section, we develop a tractable approach for obtaining an upper bound on the optimal expected revenue for an individual instance of problem (1) under a space constraint. By comparing this upper bound on the optimal expected revenue with the expected revenue obtained by a particular assortment, we can get a feel for the optimality gap of the assortment on hand.

To construct an upper bound on the optimal expected revenue in problem (1), for each nest i, we partition the interval $[0, c]$ into K intervals $\{[b_i^{k-1}, b_i^k] : k = 1, \ldots, K\}$, where we have $0 = b_i^0 \leq b_i^1 \leq \ldots \leq b_i^{K-1} \leq b_i^K = c$. Noting that the total preference weight of the products offered in nest i can at most be $\sum_{j \in N} v_{ij}$, we let $\bar{v}_i = \sum_{j \in N} v_{ij}$ and partition the interval $[0, \bar{v}_i]$ into L intervals $\{[\nu_i^{q-1}, \nu_i^q] : q = 1, \ldots, L\}$ with $0 = \nu_i^0 \leq \nu_i^1 \leq \ldots \leq \nu_i^{L-1} \leq \nu_i^L = \bar{v}_i$. Using the decision variables $x_i = \{x_{ij} : j \in N\} \in [0, 1]^n$, we define $\phi_i^{kq}(z)$ as

$$\phi_i^{kq}(z) \quad = \quad \max \quad (\nu_i^{q-1})^{\gamma_i} \left\{ \frac{\sum_{j \in N} v_{ij}\, r_{ij}\, x_{ij}}{\nu_i^{q-1}} - z \right\} \tag{21}$$

$$\text{st} \quad \sum_{j \in N} w_{ij}\, x_{ij} \leq b_i^k \tag{22}$$

$$\sum_{j \in N} v_{ij}\, x_{ij} \leq \nu_i^q \tag{23}$$

$$0 \leq x_{ij} \leq \mathbf{1}(w_{ij} \leq b_i^k) \qquad \forall\, j \in N, \tag{24}$$

which is a continuous knapsack problem with two dimensions. The selection of the intervals $\{[b_i^{k-1}, b_i^k] : k = 1, \ldots, K\}$ and $\{[\nu_i^{q-1}, \nu_i^q] : q = 1, \ldots, L\}$ can be completely arbitrary, as long as these intervals respectively cover $[0, c]$ and $[0, \bar{v}_i]$. We observe that $\phi_i^{kq}(z)$ is a linear function of z. If $q = 1$, then $\nu_i^{q-1} = 0$, in which case, we have a zero in the denominator of the fraction above. To deal with this case, if $q = 1$, then we follow the convention that $\phi_i^{kq}(z) = 0$ for all $k = 1, \ldots, K$ and $z \in \Re_+$. Roughly speaking, we can interpret problem (21)-(24) as a continuous version of problem (9). In the objective function of problem (21)-(24), the term ν_i^{q-1} corresponds to $V_i(S_i)$ in the objective function of problem (9). Noting that $R_i(S_i) = \sum_{j \in S_i} r_{ij}\, v_{ij} / V_i(S_i)$, the term $\sum_{j \in N} v_{ij}\, r_{ij}\, x_{ij} / \nu_i^{q-1}$ in the objective function of problem (21)-(24) corresponds to $R_i(S_i)$ in the objective function of problem (9). The first constraint in problem (21)-(24) imposes the capacity constraint, whereas the second constraint ensures that the total preference weight of the offered products are computed correctly. We use ν_i^{q-1} in the objective function, but ν_i^q in the constraint to ultimately ensure that we can use $\phi_i^{kq}(z)$ to obtain an upper bound on the optimal expected revenue. Using the decision variables Δ, $y = \{y_i : i \in M\}$ and z, to obtain an upper bound on the optimal expected revenue, we propose solving the problem

$$\min \quad c\,\Delta + \sum_{i \in M} y_i \tag{25}$$

$$\text{st} \quad b_i^{k-1}\,\Delta + y_i \geq \phi_i^{kq}(z) \qquad \forall\, i \in M,\ k = 1, \ldots, K,\ q = 1, \ldots, L \tag{26}$$

$$c\,\Delta + \sum_{i \in M} y_i = v_0\, z \tag{27}$$

$$\Delta \geq 0,\ y_i \text{ is free},\ z \text{ is free} \quad \forall\, i \in M. \tag{28}$$

Since $\phi_i^{kq}(\cdot)$ is linear, the problem above is a linear program. The next theorem shows that we can use this problem to obtain an upper bound on the optimal expected revenue z^* in problem (1).

Theorem 7 *Letting $(\hat{\Delta}, \hat{y}, \hat{z})$ be an optimal solution to problem* (25)-(28)*, we have $\hat{z} \geq z^*$.*

Proof. We let $(S_1^*, \ldots, S_m^*)$ be an optimal solution to problem (1), k_i' be such that $C_i(S_i^*) \in [b_i^{k_i'-1}, b_i^{k_i'}]$ and q_i' be such that $V_i(S_i^*) \in [\nu_i^{q_i'-1}, \nu_i^{q_i'}]$. Since $(\hat{\Delta}, \hat{y}, \hat{z})$ is a feasible solution to problem (25)-(28), we have $b_i^{k_i'-1}\,\hat{\Delta} + \hat{y}_i \geq \phi_i^{k_i' q_i'}(\hat{z})$ for all $i \in M$. Adding this inequality over all $i \in M$, we obtain $\sum_{i \in M} \phi_i^{k_i' q_i'}(\hat{z}) \leq \sum_{i \in M} b_i^{k_i'-1}\,\hat{\Delta} + \sum_{i \in M} \hat{y}_i \leq \sum_{i \in M} C_i(S_i^*)\,\hat{\Delta} + \sum_{i \in M} \hat{y}_i \leq c\,\hat{\Delta} + \sum_{i \in M} \hat{y}_i = v_0\,\hat{z}$, where the second inequality uses the fact that $C_i(S_i^*) \geq b_i^{k_i'-1}$, the third inequality uses the fact that $(S_1^*, \ldots, S_m^*)$ is a feasible solution to problem (1) and the equality is by the fact that $(\hat{\Delta}, \hat{y}, \hat{z})$ is a feasible solution to problem (25)-(28). Thus, the last chain of inequalities implies that $\sum_{i \in M} \phi_i^{k_i' q_i'}(\hat{z}) \leq v_0\,\hat{z}$. On the other hand, consider a solution x_i^* to problem (21)-(24) obtained by letting $x_{ij}^* = 1$ if $j \in S_i^*$ and $x_{ij}^* = 0$ otherwise. Since $\sum_{j \in N} w_{ij}\, x_{ij}^* = C_i(S_i^*) \leq b_i^{k_i'}$ and $\sum_{j \in N} v_{ij}\, x_{ij}^* = V_i(S_i^*) \leq \nu_i^{q_i'}$, the solution x_i^* is feasible to problem (21)-(24) when we solve this problem with $k = k_i'$, $q = q_i'$ and $z = \hat{z}$. So, the optimal objective value of problem (21)-(24) is at least as large as the objective value provided by the feasible solution x_i^* and we obtain

$$\begin{aligned}\phi_i^{k_i' q_i'}(\hat{z}) &\geq (\nu_i^{q_i'-1})^{\gamma_i} \left\{ \frac{\sum_{j \in N} v_{ij}\, r_{ij}\, x_{ij}^*}{\nu_i^{q_i'-1}} - \hat{z} \right\} = \frac{\sum_{j \in N} v_{ij}\, r_{ij}\, x_{ij}^*}{(\nu_i^{q_i'-1})^{1-\gamma_i}} - (\nu_i^{q_i'-1})^{\gamma_i}\, \hat{z} \\ &\geq \frac{\sum_{j \in N} v_{ij}\, r_{ij}\, x_{ij}^*}{V_i(S_i^*)^{1-\gamma_i}} - V_i(S_i^*)^{\gamma_i}\, \hat{z} = V_i(S_i^*)^{\gamma_i} \left\{ \frac{\sum_{j \in S_i^*} v_{ij}\, r_{ij}}{V_i(S_i^*)} - \hat{z} \right\} = V_i(S_i^*)^{\gamma_i} (R_i(S_i^*) - \hat{z}),\end{aligned}$$

where the second inequality uses the fact that $V_i(S_i^*) \geq \nu_i^{q_i'-1}$ and the second equality uses the definition of x_i^*. Since we have $\sum_{i \in M} \phi_i^{k_i' q_i'}(\hat{z}) \leq v_0 \hat{z}$ as shown at the beginning of the proof, the chain of inequalities above implies that $v_0\, \hat{z} \geq \sum_{i \in M} \phi_i^{k_i' q_i'}(\hat{z}) \geq \sum_{i \in M} V_i(S_i^*)^{\gamma_i}\, (R_i(S_i^*) - \hat{z})$. If we focus on the first and last expressions in this chain of inequalities and solve for $\hat{z}$, then we obtain $\hat{z} \geq \sum_{i \in M} V_i(S_i^*)^{\gamma_i}\, R_i(S_i^*) / (v_0 + \sum_{i \in M} V_i(S_i^*)^{\gamma_i}) = \Pi(S_1^*, \ldots, S_m^*) = z^*$. $\square$