Online Supplement

Constrained Assortment Optimization for the Nested Logit Model

Guillermo Gallego
Department of Industrial Engineering and Operations Research
Columbia University, New York, New York 10027, USA
gmg2@columbia.edu

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

February 11, 2014

A Online Supplement: Proposition 7

Proposition 7 is used in Section 7 and it shows that the optimal objective value of problem (13) is an upper bound on the optimal expected revenue when we have space constraints.

Proposition 7 *If we use $\hat{z}$ to denote the optimal objective value of the linear program in* (13)*, then we have $\hat{z} \geq Z^*$.*

Proof. We let $(\hat{z}, \hat{y})$ be an optimal solution to problem (13) and $(S_1^*, \ldots, S_m^*)$ be an optimal solution to problem (1). We claim that $\hat{y}_i \geq V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - \hat{z})$ for all $i \in M$. To establish this claim, we note that if u is large enough, then the optimal solution to the linear programming relaxation of problem (10) is zero. Thus, letting $\mathcal{I}_i^g$ and x_i^g be as defined in Section 5, there exists an interval $\mathcal{I}_i^g$ such that the solution x_i^g to the linear programming relaxation of problem (10) is zero when u takes values in this interval. In this case, the second set of constraints in problem (13) implies that $\hat{y}_i \geq 0$ for all $i \in M$. Therefore, if $S_i^* = \bar{0}$ or $R_i(S_i^*) \leq \hat{z}$, then we have $V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - \hat{z}) \leq 0 \leq \hat{y}_i$ and our claim trivially holds when $S_i^* = \bar{0}$ or $R_i(S_i^*) \leq \hat{z}$. So, it is enough to establish our claim with $S_i^* \neq \bar{0}$ and $R_i(S_i^*) > \hat{z}$.

We let $\hat{u} = \gamma_i \, \hat{z} + (1 - \gamma_i) \, R_i(S_i^*)$. Letting g be the index of the interval $\mathcal{I}_i^g$ that includes $\hat{u}$, by definition, x_i^g is the optimal solution to the linear programming relaxation of problem (10) when we solve this problem with $u = \hat{u}$. Therefore, we have $\sum_{j \in N} v_{ij} \, (r_{ij} - \hat{u}) \, x_{ij}^g \geq \sum_{j \in N} v_{ij} \, (r_{ij} - \hat{u}) \, S_{ij}^*$, where we use the fact that x_i^g is the optimal solution to the linear programming relaxation of problem (10) when we solve this problem with $u = \hat{u}$ but offering the products in S_i^* provides a feasible, but not necessarily an optimal, solution to this problem. First, we assume that $x_i^g \neq \bar{0}$. In this case, slightly abusing the notation to let $V_i(x_i^g) = \sum_{j \in N} v_{ij} \, x_{ij}^g$ and $R_i(x_i^g) = \sum_{j \in N} v_{ij} \, r_{ij} \, x_{ij}^g / V_i(x_i^g)$, the last inequality can equivalently be written as $V_i(x_i^g) \, (R_i(x_i^g) - \hat{u}) \geq V_i(S_i^*) \, (R_i(S_i^*) - \hat{u})$.

Noting that $(\hat{z}, \hat{y})$ is a feasible solution to problem (13), this solution satisfies the second set of constraints in problem (13) for the index g defined at the beginning of the previous paragraph, in which case, we can write this constraint in a compact fashion as $\hat{y}_i \geq V_i(x_i^g)^{\gamma_i}(R_i(x_i^g) - \hat{z})$. Also, using the fact that u^{γ_i} is a concave function of u, the subgradient inequality yields $V_i(S_i^*)^{\gamma_i} \leq V_i(x_i^g)^{\gamma_i} + \gamma_i \, V_i(x_i^g)^{\gamma_i - 1} \, (V_i(S_i^*) - V_i(x_i^g)) = \gamma_i \, V_i(x_i^g)^{\gamma_i - 1} \, V_i(S_i^*) + (1 - \gamma_i) \, V_i(x_i^g)^{\gamma_i}$. Using these observations, we have the chain of inequalities

$$
\begin{aligned}
\hat{y}_i \geq V_i(x_i^g)^{\gamma_i}(R_i(x_i^g) - \hat{z}) &= V_i(x_i^g)^{\gamma_i}(R_i(x_i^g) - \hat{u}) + (1 - \gamma_i) \, V_i(x_i^g)^{\gamma_i}(R_i(S_i^*) - \hat{z}) \\
&\geq V_i(x_i^g)^{\gamma_i - 1} \, V_i(S_i^*) \, (R_i(S_i^*) - \hat{u}) + (1 - \gamma_i) \, V_i(x_i^g)^{\gamma_i}(R_i(S_i^*) - \hat{z}) \\
&= \gamma_i \, V_i(x_i^g)^{\gamma_i - 1} \, V_i(S_i^*) \, (R_i(S_i^*) - \hat{z}) + (1 - \gamma_i) \, V_i(x_i^g)^{\gamma_i}(R_i(S_i^*) - \hat{z}) \\
&\geq V_i(S_i^*)^{\gamma_i} \, (R_i(S_i^*) - \hat{z}),
\end{aligned}
$$

where the first inequality follows from the inequality we establish at the beginning of this paragraph, the first equality follows by using the definition of $\hat{u}$ and arranging the terms, the second inequality follows from the fact that $V_i(x_i^g)\,(R_i(x_i^g) - \hat{u}) \geq V_i(S_i^*)\,(R_i(S_i^*) - \hat{u})$, which is established in the previous paragraph, the second equality follows by using the definition of $\hat{u}$ and the third inequality follows by noting that $V_i(S_i^*)^{\gamma_i} \leq \gamma_i\, V_i(x_i^g)^{\gamma_i - 1}\, V_i(S_i^*) + (1-\gamma_i)\, V_i(x_i^g)^{\gamma_i}$, which is established above by using the subgradient inequality. So, our claim holds when $x_i^g \neq \bar{0}$.

Second, we assume that $x_i^g = \bar{0}$ so that the optimal solution to the linear programming relaxation of problem (10) is zero when this problem is solved with $u = \hat{u}$. Thus, the utility of each product in this problem should be negative, yielding $r_{ij} \leq \hat{u} = \gamma_i\,\hat{z} + (1-\gamma_i)\,R_i(S_i^*)$ for all $j \in N$. Noting $R_i(S_i^*) > \hat{z}$, the last inequality gives $r_{ij} < R_i(S_i^*)$ for all $j \in N$. However, since $R_i(S_i^*) = \sum_{j \in N} v_{ij}\, r_{ij}\, S_{ij}^* / \sum_{j \in N} v_{ij}\, S_{ij}^*$ by definition, $R_i(S_i^*)$ is a weighted average of the product revenues $\{r_{ij} : j \in N\}$ in the assortment S_i^*. So, we cannot have $r_{ij} < R_i(S_i^*)$ for all $j \in N$, indicating that the case $x_i^g = \bar{0}$ cannot occur. Thus, our claim is established and we have $\hat{y}_i \geq V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - \hat{z})$ for all $i \in M$. Adding these inequalities over all $i \in M$ and noting that $v_0\,\hat{z} \geq \sum_{i \in M} \hat{y}_i$ by the first constraint in problem (13), we obtain $v_0\,\hat{z} \geq \sum_{i \in M} V_i(S_i^*)^{\gamma_i}(R_i(S_i^*) - \hat{z})$, in which case, solving for $\hat{z}$ in this inequality and noting the definition of $\Pi(S_1, \ldots, S_m)$, we get $\hat{z} \geq \Pi(S_1^*, \ldots, S_m^*)$. So, $\hat{z}$ is an upper bound on the expected revenue from the optimal assortment, as desired. $\Box$

B Online Supplement: Extensions to Other Constraints

In this section, we show how to extend our approach to additional types of constraints.

B.1 Parent Product Constraints

We consider the case where certain products in a nest are designated as parent products. Each parent product has a set of child products associated with it and if a parent product is not offered, then none of its child products can be offered. We refer to this type of constraints as *parent product constraints*. Parent product constraints arise when company policy or law requires offering certain products before offering others. For example, a company may be required to offer the generic version of a drug before it can offer the brand name versions, in which case, the generic version acts as the parent product, the brand name versions are the child products of the generic version and if the generic version is not offered, then none of the brand name versions can be offered. To capture parent product constraints, we use P_i to denote the set of parent products in nest i. A parent product j in nest i has the set of child products C_{ij}. If product j is neither a parent product nor a child product, then we assume that product j is a parent product with an empty set of child products. Also, we assume that $C_{ij} \cap C_{ik} = \emptyset$ for all distinct $j, k \in P_i$ so that two parent products have different sets of child products. By the last two assumptions, the sets of products P_i and $\{C_{ij} : j \in P_i\}$ collectively partition N. So, the feasible assortments in nest i under parent constraints are given by $\mathcal{C}_i = \{S_i \in \{0,1\}^n : S_{ik} \leq S_{ij}\ \forall\, j \in P_i,\ k \in C_{ij}\}$, ensuring that if we do not offer a parent product, then none of its child products can be offered.

Under parent product constraints, we show that we can come up with a collection of assortments

$\{A_i^t : t \in \mathcal{T}_i\}$ that includes an optimal solution to problem (7) for any $u \in \Re_+$. Furthermore, this collection of assortments includes $O(n)$ assortments. These results, together with Theorems 2 and 4, imply that we can solve a linear program with $1 + m$ decision variables and $O(mn)$ constraints to obtain the optimal assortment under parent product constraints. To characterize the optimal solution to problem (7) for any $u \in \Re_+$, we use the decision variables $x_i = (x_{i1}, \ldots, x_{in}) \in \{0,1\}^n$ to write problem (7) under parent product constraints as

$$\max\left\{ \sum_{j\in N} v_{ij}\,(r_{ij} - u)\,x_{ij} : x_{ik} \le x_{ij}\ \forall\, j \in P_i,\ k \in C_{ij},\ x_{ij} \in \{0,1\}\ \forall\, j \in N \right\}. \tag{14}$$

Assume that parent product j is offered in the optimal solution to problem (14). In this case, we are free to offer any of the child products of parent product j. Since it is optimal to offer one of these child products when their objective function coefficient is positive, if parent product j is offered, then the total contribution of parent product j and all of its child products to the objective function of the problem above is given by $f_{ij}(u) = v_{ij}\,(r_{ij} - u) + \sum_{k\in C_{ij}} v_{ik}\,[r_{ik} - u]^+$, where we use $[\cdot]^+ = \max\{\cdot, 0\}$. On the other hand, if parent product j is not offered, then none of its child products can be offered, in which case, parent product j and all of its child products make a contribution of zero to the objective function of the problem above. Therefore, it is optimal to offer parent product j as long as $f_{ij}(u) > 0$. The function $f_{ij}(\cdot)$ is decreasing and piecewise linear with points of nondifferentiability occurring at $\{r_{ik} : k \in C_{ij}\}$. Thus, we can find a value of $\bar{u}_{ij}$ such that $f_{ij}(u) > 0$ for any $u < \bar{u}_{ij}$ and $f_{ij}(u) \le 0$ for any $u \ge \bar{u}_{ij}$, in which case, we offer parent product j in the optimal solution to the problem above when $u < \bar{u}_{ij}$ and we do not offer parent product j when $u \ge \bar{u}_{ij}$. If it is optimal not to offer parent product j, then none of its child products are offered, whereas if it is optimal to offer parent product j, then its child product k is offered when $u < r_{ik}$. Therefore, by comparing the value of u with $\bar{u}_{ij}$ and $\{r_{ik} : k \in C_{ij}\}$, we can decide whether it is optimal to offer parent product j and any of its child products in the optimal solution to problem (14). Furthermore, it is straightforward to obtain the point $\bar{u}_{ij}$. Repeating the same reasoning for all of the parent products, we obtain the collections of points $\{\bar{u}_{ij} : j \in P_i\}$ and $\{r_{ik} : j \in P_i,\ k \in C_{ij}\}$. Since P_i and $\{C_{ij} : j \in P_i\}$ partition N, there are a total of n points in the collections $\{\bar{u}_{ij} : j \in P_i\}$ and $\{r_{ik} : j \in P_i,\ k \in C_{ij}\}$. These points completely characterize the optimal solution to problem (14) since we can compare u with $\bar{u}_{ij}$ to decide whether it is optimal to offer parent product j. If this is the case, then we can decide whether it is optimal to offer its child product k by comparing u with r_{ik}.

Since there are a total of n points in the collections $\{\bar{u}_{ij} : j \in P_i\}$ and $\{r_{ik} : j \in P_i,\ k \in C_{ij}\}$, these points partition the positive real line into $O(n)$ intervals and we denote these intervals by $\{\mathcal{I}_i^t : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n)$. We observe that as long as u takes values in one of the intervals $\{\mathcal{I}_i^t : t \in \mathcal{T}_i\}$, the ordering between u and any of the points in the collections $\{\bar{u}_{ij} : j \in P_i\}$ and $\{r_{ik} : j \in P_i,\ k \in C_{ij}\}$ does not change. This observation, in view of the discussion in the paragraph above, implies that the optimal solution to problem (14) does not change as long as u takes values in one of the intervals $\{\mathcal{I}_i^t : t \in \mathcal{T}_i\}$. Therefore, by comparing the value of u with $\{\bar{u}_{ij} : j \in P_i\}$ and $\{r_{ik} : j \in P_i,\ k \in C_{ij}\}$ in each one of the intervals $\{\mathcal{I}_i^t : t \in \mathcal{T}_i\}$, we can come up with a collection of assortments $\{A_i^t : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n)$ such that this collection always includes an optimal solution to problem (7) for any $u \in \Re_+$, as desired. In this case, Theorem 4 with $\alpha = 1$ implies that the best assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ with $\hat{S}_i \in \{A_i^t : t \in \mathcal{T}_i\}$ is the optimal solution to problem (1). By Theorem 2, we can find this best assortment by solving a linear program with $1+m$ decision variables and $O(mn)$ constraints.

B.2 Cardinality and Space Constraints

In this section, we consider the case where we have both cardinality and space constraints on the assortment offered in each nest. In particular, if we use b_i to denote the limit on the cardinality of the assortment offered in nest i, c_i to denote the space availability in nest i and w_{ij} to denote the space requirement of product j in nest i, then the feasible assortments in nest i are given by $\mathcal{C}_i = \{S_i \in \{0,1\}^n : \sum_{j \in N} S_{ij} \le b_i,\ \sum_{j \in N} w_{ij}\, S_{ij} \le c_i\}$. For this case, we begin by showing that we can come up with a collection of assortments $\{A_i^t : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(n^2)$ such that this collection includes a 3-approximate solution to problem (7) for any $u \in \Re_+$. Thus, we can solve a linear program with $1 + m$ decision variables and $O(mn^2)$ constraints to obtain a solution to the assortment optimization problem whose expected revenue deviates from the optimal expected revenue by no more than a factor of three. Later in this section, we refine our analysis to show that we can improve this performance guarantee from three to two.

The discussion in this section follows the development in Sections 4 and 5 closely. So, we mostly focus on the main points. Using the decision variables $x_i = (x_{i1}, \ldots, x_{in}) \in \{0,1\}^n$, we write problem (7) under cardinality and space constraints as

$$\max \left\{ \sum_{j \in N} v_{ij}\, (r_{ij} - u)\, x_{ij} : \sum_{j \in N} x_{ij} \le b_i,\ \sum_{j \in N} w_{ij}\, x_{ij} \le c_i,\ x_{ij} \in \{0,1\}\ \forall\, j \in N \right\}, \tag{15}$$

which is a cardinality constrained knapsack problem. A basic exercise in duality theory shows that there are at most n^2 possible optimal bases to the linear programming relaxation of problem (15). Naturally, for any $u \in \Re_+$, the optimal solution to the linear programming relaxation of problem (15) must correspond to one of these optimal bases. In other words, for any $u \in \Re_+$, the optimal solution to the linear programming relaxation of problem (15) is one of n^2 solutions. By using the parametric simplex method over $u \in \Re_+$, we can generate all of these n^2 solutions. We use $\{x_i^g : g \in \mathcal{G}_i\}$ with $|\mathcal{G}_i| = O(n^2)$ to denote all possible solutions to the linear programming relaxation of problem (15). Thus, for any $u \in \Re_+$, there exists some x_i^g with $g \in \mathcal{G}_i$ such that x_i^g is the optimal solution to the linear programming relaxation of problem (15).

Implicitly treating the upper bounds $0 \le x_{ij} \le 1$ for all $j \in N$ in the linear programming relaxation of problem (15), we observe that there must be two basic decision variables in any basic optimal solution and the other decision variables have integer values. Thus, the solution x_i^g has at most two fractional components. Using the solution x_i^g, we define the assortment $S_i^g = (S_{i1}^g, \ldots, S_{in}^g) \in \{0,1\}^n$ such that $S_{ij}^g = \lfloor x_{ij}^g \rfloor$ for all $j \in N$. In other words, the assortment S_i^g

includes the products that take value one in the solution x_i^g. In this case, augmenting the collection of assortments $\{S_i^g : g \in \mathcal{G}_i\}$ with the collection of singleton assortments $\{\{j\} : j \in N\}$, it is possible to show that the collection of assortments $\{S_i^g : g \in \mathcal{G}_i\} \cup \{\{j\} : j \in N\}$ always includes a 3-approximate solution to problem (15) for any $u \in \Re_+$.

To see this result, assume that we solve problem (15) for some $u \in \Re_+$ and let g be such that x_i^g is the optimal solution to the linear programming relaxation of problem (15) when we solve this problem with the value of u in consideration. By the discussion at the beginning of the paragraph above, the solution x_i^g has at most two fractional components. We use j_1^g to denote the first fractional component of x_i^g when there is one. Similarly, we use j_2^g to denote the second fractional component of x_i^g when there is one. In this case, if we let $z^*(u)$ be the optimal objective value of problem (15), then noting that the optimal objective value of the linear programming relaxation provides an upper bound on $z^*(u)$, we obtain the chain of inequalities

$$\begin{aligned} z^*(u) \leq \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, x_{ij}^g &\leq \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, S_{ij}^g + v_{ij_1^g}(r_{ij_1^g} - u) + v_{ij_2^g}(r_{ij_2^g} - u) \\ &\leq 3 \max\left\{ \sum_{j \in N} v_{ij^g}\,(r_{ij^g} - u)\, S_{ij}^g, v_{ij_1^g}(r_{ij_1^g} - u), v_{ij_2^g}(r_{ij_2^g} - u) \right\}, \end{aligned} \tag{16}$$

where the second inequality follows from the fact that the assortment S_i^g, together with j_1^g and j_2^g, includes all components of the solution x_i^g that take strictly positive values. If the solution x_i^g has fewer than two fractional components, then the inequalities above continue to hold if we ignore the terms that involve j_1^g or j_2^g. From (16), we observe that either one of the assortments S_i^g, $\{j_1^g\}$ and $\{j_2^g\}$ is a 3-approximate solution to problem (15). Therefore, the collection of assortments $\{S_i^g : g \in \mathcal{G}_i\} \cup \{\{j\} : j \in N\}$ includes a 3-approximate solution to problem (15) for any $u \in \Re_+$ and there are $O(n^2)$ assortments in this collection, establishing the desired result.

We can tighten the approximation guarantee from three to two by using a somewhat more involved definition of the assortment S_i^g. If the solution x_i^g has zero or one fractional component, then we continue defining $S_i^g = (S_{i1}^g, \ldots, S_{in}^g) \in \{0,1\}^n$ such that $S_{ij}^g = \lfloor x_{ij}^g \rfloor$ for all $j \in N$. However, if the solution x_i^g has two fractional components, then using j_1^g and j_2^g to denote these fractional components with the convention that $w_{ij_1^g} \leq w_{ij_2^g}$, we define the assortment $S_i^g = (S_{i1}^g, \ldots, S_{in}^g) \in \{0,1\}^n$ as $S_{ij}^g = 1$ when $x_{ij}^g = 1$ or $j = j_1^g$, otherwise $S_{ij}^g = 0$. So, the assortment S_i^g includes all products that take value one and the product with the smaller space requirement that takes a fractional value in the solution x_i^g. In this case, we can show that the collection of assortments $\{S_i^g : g \in \mathcal{G}_i\} \cup \{\{j\} : j \in N\}$ includes a 2-approximate solution to problem (15) for any $u \in \Re_+$. To see this result, we can follow another basic exercise in duality theory to show that if there are two fractional components in a basic optimal solution to the linear programming relaxation of problem (15), then both constraints must be satisfied as equality. Thus, if x_i^g has two fractional components j_1^g and j_2^g, then $x_{ij_1^g}^g + x_{ij_2^g}^g = 1$. The last expression, together with the fact that $w_{ij_1^g} \leq w_{ij_2^g}$ and x_i^g is a feasible solution to the linear programming relaxation of problem (15), yields $c_i \geq \sum_{j \in N} w_{ij}\, x_{ij}^g = \sum_{j \in N} w_{ij}\, \mathbf{1}(x_{ij}^g = 1) + w_{ij_1^g}\, x_{ij_1^g} + w_{ij_2^g}\, x_{ij_2^g} \geq \sum_{j \in N} w_{ij}\, \mathbf{1}(x_{ij}^g = 1) + w_{ij_1^g}$. So, the assortment S_i^g, which includes the products $\{j \in N : x_{ij}^g = 1\} \cup \{j_1^g\}$, is feasible to problem (15). In this case, we can use the same line of reasoning in (16) to get

$$z^*(u) \leq \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, x_{ij}^g \leq \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, S_{ij}^g + v_{ij_2^g}(r_{ij_2^g} - u)$$

$$\leq 2 \max \left\{ \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, S_{ij}^g,\, v_{ij_2^g}(r_{ij_2^g} - u) \right\},$$

where the second inequality holds since the assortment S_i^g includes all strictly positive components of x_i^g except for j_2^g. The chain of inequalities above shows that either S_i^g or $\{j_2^g\}$ is a 2-approximate solution to problem (15), as desired.

C Online Supplement: Better Performance Guarantees under Space Constraints

In Section 5, we describe an approach to obtain a $\min\{2, 1/(1-\epsilon)\}$-approximate solution to problem (1) under space constraints. The smallest possible value of ϵ that we can use in this performance guarantee is $\bar{\epsilon} = \max\{w_{ij}/c_i : i \in M,\ j \in N\}$, indicating that the best performance guarantee from the approach described in Section 5 is given by $\min\{2, 1/(1-\bar{\epsilon})\}$. In particular, even if we are willing to increase the computational effort, the approach described in Section 5 does not provide any guidance as to how we can improve this performance guarantee. In this section, our goal is to show how we can obtain better performance guarantees under space constraints as long as we are willing to increase the computational effort.

C.1 Improving the Performance Guarantee

The starting point for our discussion is problem (10), which is equivalent to problem (7) under space constraints. We recall that if we can come up with a collection of assortments $\{A_i^t : t \in \mathcal{T}_i\}$ such that this collection includes an α-approximate solution to problem (10) for any $u \in \Re_+$, then Theorem 4 implies that the best assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ with $\hat{S}_i \in \{A_i^t : t \in \mathcal{T}_i\}$ is an α-approximate solution to problem (1). Furthermore, by Theorem 2, we can find this best assortment by solving a linear program with $1 + m$ decision variables and $1 + \sum_{i \in M} |\mathcal{T}_i|$ constraints. In this section, we show that if we are given any $\alpha > 1$, then we can come up with a collection of assortments $\{A_i^t : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(\lceil \alpha/(\alpha-1) \rceil\, n^{\lceil \alpha/(\alpha-1) \rceil + 2})$ such that this collection always includes an α-approximate solution to problem (10) for any $u \in \Re_+$. In this case, by Theorem 4, the best assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ with $\hat{S}_i \in \{A_i^t : t \in \mathcal{T}_i\}$ is an α-approximate solution to problem (1). By Theorem 2, we can find this best assortment by solving a linear program with $1 + m$ decision variables and $O(m \lceil \alpha/(\alpha-1) \rceil\, n^{\lceil \alpha/(\alpha-1) \rceil + 2})$ constraints. Thus, by choosing α closer to one, we can obtain a performance guarantee that is closer to one as long as we are willing to increase the number of constraints in the linear program.

To characterize approximate solutions to problem (10), we use a special linear programming relaxation to this problem. Using the decision variables $x_i = (x_{i1}, \ldots, x_{in}) \in [0,1]^n$, for any given

$J \subset N$, we consider the problem

$$\max \left\{ \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, x_{ij} \;:\; \sum_{j \in N} w_{ij}\, x_{ij} \le c_i, \right.$$

$$\left. x_{ij} = 1 \;\forall\, j \in J, \;\; 0 \le x_{ik} \le \mathbf{1}\Big(v_{ik}\,(r_{ik} - u) \le \min_{j \in J}\{ v_{ij}\,(r_{ij} - u)\}\Big) \;\forall\, k \in N \setminus J \right\}. \quad (17)$$

We can interpret the problem above as the linear programming relaxation of a knapsack problem after fixing the values of some of the decision variables at zero or one. In particular, we fix the values of the decision variables corresponding to the products in J at one. For the remaining decision variables, if the utility of a product corresponding to one of these decision variables exceeds the minimum of the utilities of the products in J, then we fix the value of this decision variable at zero. The role of the indicator function in the third set of constraints is to drop the products with utilities exceeding the minimum of the utilities of the products in J from consideration. Similar constraints appear in Frieze and Clarke (1984). The problem above may be infeasible for a certain J, in which case, we set the values of all decision variables to zero by convention. Problem (17) is the linear programming relaxation of a knapsack problem, where the utility of product j is $v_{ij}\,(r_{ij} - u)$ and the capacity consumption of product j is w_{ij}. So, we can solve this problem by using the following procedure. We put all of the products in J into the knapsack and drop these products from consideration. We order the other products with respect to their utilities. If there are any products whose utilities exceed the smallest of the utilities of the products in J, then we drop these products from consideration as well. Considering the remaining products, we fill the knapsack starting from the product with the largest utility to space consumption ratio, as long as the utility of the product exceeds zero. This procedure implies that the optimal solution to problem (17) does not change as long as the ordering of the utilities, ordering of the utility to space consumption ratios and signs of the utilities of the products do not change. Also, there is at most one fractional decision variable in the optimal solution to problem (17) obtained by using this procedure.

To exploit the fact that the optimal solution to problem (17) does not change as long as the ordering of the utilities, ordering of the utility to space consumption ratios and signs of the utilities of the products do not change, we define the linear functions $h_{ij}(u) = v_{ij}\,(r_{ij} - u)$ and $f_{ij}(u) = v_{ij}\,(r_{ij} - u)/w_{ij}$ for $j \in N$ and $h_{i0}(u) = 0$, $f_{i0}(u) = 0$. In this case, $h_{ij}(u)$ and $f_{ij}(u)$ respectively capture the utility and utility to space consumption ratio of product j in problem (17). We use $\{\bar{w}_i^g : g \in \mathcal{H}_i\}$ to denote the set of intersection points of the $n+1$ linear functions $\{h_{ij}(\cdot) : j \in N \cup \{0\}\}$ and $\{\bar{u}_i^g : g \in \mathcal{H}_i\}$ to denote the set of intersection points of the $n+1$ linear functions $\{f_{ij}(\cdot) : j \in N \cup \{0\}\}$. Thus, there are $|\mathcal{H}_i| = O(n^2)$ points in each one of these two sets of intersection points. Collecting the points in the two sets $\{\bar{w}_i^g : g \in \mathcal{H}_i\}$ and $\{\bar{u}_i^g : g \in \mathcal{H}_i\}$ together, we observe that the points $\{\bar{w}_i^g : g \in \mathcal{H}_i\} \cup \{\bar{u}_i^g : g \in \mathcal{H}_i\}$ partition the positive real line into $O(2\,|\mathcal{H}_i|) = O(n^2)$ intervals. We use $\{\mathcal{I}_i^g : g \in \mathcal{G}_i\}$ with $\mathcal{G}_i = O(n^2)$ to denote these intervals, in which case, the ordering of the utilities, ordering of the utility to space consumption ratios and signs of the utilities of the products in problem (17) do not change as long as u takes values in one of these intervals. Since the optimal solution to problem (17) depends only on the ordering of the utilities, ordering of the utility to space consumption ratios and signs of the utilities of the products, the optimal solution to problem (17) does not change either when u takes values in one of these intervals. We use $x_i^g(J)$ to denote the optimal solution to problem (17) when u takes values in the interval $\mathcal{I}_i^g$. Our notation for $x_i^g(J)$ reflects the fact that the optimal solution to problem (17) depends on the choice of J.

Using the solution $x_i^g(J)$, we define the assortment $S_i^g(J) = (S_{i1}^g(J), \ldots, S_{in}^g(J)) \in \{0,1\}^n$ as $S_{ij}^g(J) = \lfloor x_{ij}^g(J) \rfloor$ for all $j \in N$. Therefore, the assortment $S_i^g(J)$ includes the products taking value one in the solution $x_i^g(J)$. In this case, using $\wp_q$ to denote the set of subsets of N with cardinality not exceeding q, we propose using the collection of assortments $\{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$ as a collection of possibly good solutions to problem (10). Noting that $|\wp_q| = O(qn^q)$ and $|\mathcal{G}_i| = O(n^2)$, there are $O(qn^{q+2})$ assortments in the collection $\{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$, which can be manageable when q is not too large. The next lemma shows that this collection always includes a $q/(q-1)$-approximate solution to problem (10) for any $u \in \Re_+$. To keep the focus on our main result, we defer the proof of this lemma to Section C.2.

Lemma 8 *Letting $S_i^g(J)$ be as given above, the collection of assortments $\{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$ includes a $q/(q-1)$-approximate solution to problem* (10) *for any $u \in \Re_+$.*

For any desired performance guarantee $\alpha > 1$, setting $\alpha = q/(q-1)$ and solving for q, we obtain $q = \alpha/(\alpha-1)$. Thus, if we choose $q = \lceil \alpha/(\alpha-1) \rceil$ in the lemma above, then the collection of assortments $\{S_i^g(J) : J \in \wp_{\lceil \alpha/(\alpha-1) \rceil},\ g \in \mathcal{G}_i\}$ includes an α-approximate solution to problem (10) for any $u \in \Re_+$. To come up with this collection of assortments, we compute the intervals $\{\mathcal{I}_i^g : g \in \mathcal{G}_i\}$ by finding the intersection points of the linear functions $\{h_{ij}(\cdot) : j \in N \cup \{0\}\}$ and $\{f_{ij}(\cdot) : j \in N \cup \{0\}\}$. In this case, the ordering of the utilities, ordering of the utility to space consumption ratios and signs of the utilities of the products in problem (17) do not change when u takes values in one of the intervals $\{\mathcal{I}_i^g : g \in \mathcal{G}_i\}$. Once these intervals are computed, we focus on each one of them one by one. For each interval $\mathcal{I}_i^g$ and for each $J \in \wp_{\lceil \alpha/(\alpha-1) \rceil}$, we solve problem (17) to get the optimal solution $x_i^g(J)$ and define the assortment $S_i^g(J)$ as above. Since $|\wp_{\lceil \alpha/(\alpha-1) \rceil}| = O(\lceil \alpha/(\alpha-1) \rceil\, n^{\lceil \alpha/(\alpha-1) \rceil})$ and $|\mathcal{G}_i| = O(n^2)$, there are $O(\lceil \alpha/(\alpha-1) \rceil\, n^{\lceil \alpha/(\alpha-1) \rceil + 2})$ assortments in the collection $\{S_i^g(J) : J \in \wp_{\lceil \alpha/(\alpha-1) \rceil},\ g \in \mathcal{G}_i\}$. The next theorem collects our observations.

Theorem 9 *Under space constraints, for any $\alpha > 1$, there exists a collection of assortments $\{A_i^t : t \in \mathcal{T}_i\}$ with $|\mathcal{T}_i| = O(\lceil \alpha/(\alpha-1) \rceil\, n^{\lceil \alpha/(\alpha-1) \rceil + 2})$ such that this collection includes an α-approximate solution to problem* (7) *for any $u \in \Re_+$.*

Thus, Theorem 4 implies that the best assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ with $\hat{S}_i \in \{A_i^t : t \in \mathcal{T}_i\}$ provides a performance guarantee of α for problem (1) under space constraints. Noting Theorem 2, this best assortment can be obtained by solving a linear program with $1 + m$ decision variables and

$O(m\lceil\alpha/(\alpha-1)\rceil\, n^{\lceil\alpha/(\alpha-1)\rceil+2})$ constraints. So, for any desired performance guarantee $\alpha > 1$, finding an assortment that provides this performance guarantee amounts to solving a linear program with $1 + m$ decision variables and $O(m\lceil\alpha/(\alpha-1)\rceil\, n^{\lceil\alpha/(\alpha-1)\rceil+2})$ constraints. This result demonstrates how we can improve the performance guarantee by increasing the number of constraints in the linear program. This approach naturally becomes computationally intractable when α gets too close to one, but if, for example, we want a performance guarantee of $\alpha = 3/2$, then the number of constraints we need comes out to be $O(mn^5)$.

The development in this section builds on Frieze and Clarke (1984), where the authors develop polynomial time approximation schemes for multi-dimensional knapsack problems. However, the focus of Frieze and Clarke (1984) is on solving a single instance of a multi-dimensional knapsack problem, but we are interested in finding good solutions to problem (10) for all $u \in \Re_+$.

C.2 Proof of Lemma 8

In this section, we give a proof for Lemma 8. Throughout this section, we denote an assortment offered in nest i by using a subset $S_i \subset N$. This is a slight deviation from our earlier notation where we use a vector $S_i \in \{0,1\}^n$ to denote an assortment offered in nest i, but given a vector $S_i \in \{0,1\}^n$, we can define the corresponding subset as $\{j \in N : S_{ij} = 1\}$. Using a subset $S_i \subset N$ to denote an assortment offered in nest i considerably simplifies our notation in this section.

Fixing u at an arbitrary $\hat{u} \in \Re_+$, we use x_i^* to denote the optimal solution to problem (10) when solved with $u = \hat{u}$. We have the assortment $S_i^* = \{j \in N : x_{ij}^* = 1\} \subset N$ corresponding to this optimal solution. Throughout the proof, we let $\hat{g}$ be such that $\hat{u}$ takes a value in the interval $\mathcal{I}_i^{\hat{g}}$, where the intervals $\{\mathcal{I}_i^g : g \in \mathcal{G}_i\}$ are as defined in Section C.1 of Online Supplement C. We begin by considering the case $|S_i^*| \le q$ so that there are q or fewer products in the assortment S_i^*. If $|S_i^*| \le q$, then we have $S_i^* \in \wp_q$. In this case, noting that $J \subset S_i^g(J)$ by the definitions of $x_i^g(J)$ and $S_i^g(J)$, we have $S_i^* \subset S_i^{\hat{g}}(S_i^*)$, which implies that $\sum_{j \in S_i^*} v_{ij}\,(r_{ij} - \hat{u}) \le \sum_{j \in S_i^{\hat{g}}(S_i^*)} v_{ij}\,(r_{ij} - \hat{u})$. Thus, the assortment $S_i^{\hat{g}}(S_i^*)$ provides a better objective value for problem (10) than the assortment S_i^* when this problem is solved with $u = \hat{u}$. So, the assortment $S_i^{\hat{g}}(S_i^*)$ is also optimal to problem (10) when solved with $u = \hat{u}$. Also, since $S_i^* \in \wp_q$, we have $S_i^{\hat{g}}(S_i^*) \in \{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$, showing that the collection of assortments $\{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$ includes an optimal solution to problem (10) when this problem is solved with $u = \hat{u}$. Since $\hat{u}$ is arbitrary, we are done.

In the rest of the proof, we assume that $|S_i^*| > q$. We let J_i^* be the subset of S_i^* that includes the q elements of S_i^* with the largest utilities in problem (10) when this problem is solved with $u = \hat{u}$. Since $|S_i^*| > q$, J_i^* is well defined. Consider the optimal solution $x_i^{\hat{g}}(J_i^*)$ to problem (17) when this problem is solved with $u = \hat{u}$ and $J = J_i^*$. If this solution has a fractional component j', then the third set of constraints in problem (17) implies that product j' satisfies $v_{ij'}\,(r_{ij'} - \hat{u}) \le v_{ij}\,(r_{ij} - \hat{u})$ for all $j \in J_i^*$. Thus, using $z^*(u)$ to denote the optimal objective value of problem (10), we have $z^*(\hat{u}) = \sum_{j \in S_i^*} v_{ij}\,(r_{ij} - \hat{u}) = \sum_{j \in J_i^*} v_{ij}\,(r_{ij} - \hat{u}) + \sum_{j \in S_i^* \setminus J_i^*} v_{ij}\,(r_{ij} - \hat{u}) \ge q\, v_{ij'}\,(r_{ij'} - \hat{u})$, where the inequality follows by the fact that $|J_i^*| = q$ and $v_{ij'}\,(r_{ij'} - \hat{u}) \leq v_{ij}\,(r_{ij} - \hat{u})$ for all $j \in J_i^*$. Thus, the last chain of inequalities yields $v_{ij'}\,(r_{ij'} - \hat{u}) \leq z^*(\hat{u})/q$.

To finish the proof, let $z^*(u)$ and $\zeta^*(u, J)$ respectively be the optimal objective values of problems (10) and (17). We claim that $z^*(\hat{u}) \leq \zeta^*(\hat{u}, J_i^*)$. To see this claim, we note that J_i^* includes the q products with the largest utilities among the products taking value one in the optimal solution to problem (10) when this problem is solved with $u = \hat{u}$. The products in $S_i^* \setminus J_i^*$ also take value one in the optimal solution to problem (10), but by the definition of J_i^*, these products satisfy $v_{ik}\,(r_{ik} - \hat{u}) \leq \min_{j \in J_i^*}\{v_{ij}\,(r_{ij} - \hat{u})\}$ for all $k \in S_i^* \setminus J_i^*$. This inequality, together with the third set of constraints in problem (17), implies that if we solve problem (17) with $u = \hat{u}$ and $J = J_i^*$, then we fix the values of the decision variables corresponding to the products in J_i^* at one, but the values of the decision variables corresponding to the products in $S_i^* \setminus J_i^*$ are free between zero and one. Furthermore, problem (17) does not require the decision variables to take binary values. Thus, the solution S_i^* is optimal to problem (10) and feasible to problem (17) when we solve the first problem with $u = \hat{u}$ and the second problem with $u = \hat{u}$ and $J = J_i^*$. So, our claim holds and we have $z^*(\hat{u}) \leq \zeta^*(\hat{u}, J_i^*)$. In this case, if we use, as above, $x_i^{\hat{g}}(J_i^*)$ to denote the optimal solution to problem (17) when this problem is solved with $u = \hat{u}$ and $J = J_i^*$, then letting j' be the fractional component of $x_i^{\hat{g}}(J_i^*)$ when there is one, we obtain

$$
\begin{aligned}
z^*(\hat{u}) \leq \zeta^*(\hat{u}, J_i^*) &= \sum_{j \in N} v_{ij}\,(r_{ij} - \hat{u})\, x_{ij}^{\hat{g}}(J_i^*) \\
&\leq \sum_{j \in S_i^{\hat{g}}(J_i^*)} v_{ij}\,(r_{ij} - \hat{u}) + v_{ij'}\,(r_{ij'} - \hat{u}) \leq \sum_{j \in S_i^{\hat{g}}(J_i^*)} v_{ij}\,(r_{ij} - \hat{u}) + z^*(\hat{u})/q,
\end{aligned}
$$

where the second inequality is by noting that $S_i^{\hat{g}}(J_i^*)$ includes all strictly positive and integer valued components of $x_i^{\hat{g}}(J_i^*)$ and the only possibly fractional component is j' and the third inequality is by the fact that $v_{ij'}\,(r_{ij'} - \hat{u}) \leq z^*(\hat{u})/q$, which is shown in the paragraph above. Focusing on the first and last terms in the chain of inequalities above, we have $z^*(\hat{u}) \leq (q/(q-1)) \sum_{j \in S_i^{\hat{g}}(J_i^*)} v_{ij}\,(r_{ij} - \hat{u})$, showing that the assortment $S_i^{\hat{g}}(J_i^*)$ is a $q/(q-1)$-approximate solution to problem (10) when solved with $u = \hat{u}$. Since $|J_i^*| = q$, we have $S_i^{\hat{g}}(J_i^*) \in \{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$. Noting that the choice of $\hat{u}$ is arbitrary, we conclude that the collection of assortments $\{S_i^g(J) : J \in \wp_q,\ g \in \mathcal{G}_i\}$ includes a $q/(q-1)$-approximate solution to problem (10) for any $u \in \Re_+$, as desired.

D Online Supplement: Maximizing the Expected Revenue from a Single Nest

Problem (1) finds an assortment $(S_1, \ldots, S_m)$ that maximizes the expected revenue over all nests. In this section, we elaborate on the connections of this problem to the problem of maximizing the expected revenue only from a single nest. To that end, as a function of $u \in \Re_+$, we let $\hat{S}_i(u)$ be an optimal solution to the problem

$$
\max_{S_i \in \mathcal{C}_i} \Big\{ V_i(S_i)\,(R_i(S_i) - u) \Big\} = \max_{S_i \in \mathcal{C}_i} \Bigg\{ \sum_{j \in N} v_{ij}\,(r_{ij} - u)\, S_{ij} \Bigg\}, \tag{18}
$$

where the equality follows from (8). Applying Theorem 4 with $\alpha = 1$ implies that if we obtain an optimal solution to the problem above for all $u \in \Re_+$ and use $\{\hat{S}_i(u) : u \in \Re_+\}$ as a collection of candidate assortments for nest i, then the best assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ with $\hat{S}_i \in \{\hat{S}_i(u) : u \in \Re_+\}$ is an optimal solution to problem (1). Therefore, if we use $\{\hat{S}_i(u) : u \in \Re_+\}$ for all $i \in M$ as collections of candidate assortments for the different nests, then we can stitch together the optimal solution to problem (1) by using an assortment from each one of these collections. We proceed to showing that the collection $\{\hat{S}_i(u) : u \in \Re_+\}$ does not only allow us to stitch together an assortment that maximizes the expected revenue over all nests, but it also includes an assortment solving the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$. In other words, the collection of assortments $\{\hat{S}_i(u) : u \in \Re_+\}$ also includes an assortment that maximizes the expected revenue only from nest i.

To see that the collection of assortments $\{\hat{S}_i(u) : u \in \Re_+\}$ includes an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$, we let $z^* = \max_{S_i \in \mathcal{C}_i} R_i(S_i)$. So, we have $R_i(S_i) \leq z^*$ for all $S_i \in \mathcal{C}_i$ and the inequality holds as equality at the assortment that maximizes the expected revenue from nest i. Noting that $R_i(S_i) = \sum_{j \in N} v_{ij}\, r_{ij}\, S_{ij} / \sum_{j \in N} v_{ij}\, S_{ij}$, the inequality $R_i(S_i) \leq z^*$ can be written as $\sum_{j \in N} v_{ij}\,(r_{ij} - z^*)\, S_{ij} \leq 0$. In this case, we have $\sum_{j \in N} v_{ij}\,(r_{ij} - z^*)\, S_{ij} \leq 0$ for all $S_i \in \mathcal{C}_i$ and the inequality holds as equality at the assortment that maximizes the expected revenue from nest i. Therefore, an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} \sum_{j \in N} v_{ij}\,(r_{ij} - z^*)\, S_{ij}$ is also an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$. Since an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} \sum_{j \in N} v_{ij}\,(r_{ij} - z^*)\, S_{ij}$ is given by $\hat{S}_i(z^*)$, it follows that $\hat{S}_i(z^*)$ is an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$. Noting that $\hat{S}_i(z^*) \in \{\hat{S}_i(u) : u \in \Re_+\}$, the collection of assortments $\{\hat{S}_i(u) : u \in \Re_+\}$ indeed includes an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$.

We can construct the collection of assortments $\{\hat{S}_i(u) : u \in \Re_+\}$ by obtaining an optimal solution to problem (18) for all $u \in \Re_+$. In this case, if we use $\{\hat{S}_i(u) : u \in \Re_+\}$ for all $i \in M$ as collections of candidate assortments for the different nests, then the best assortment $(\hat{S}_1, \ldots, \hat{S}_m)$ with $\hat{S}_i \in \{\hat{S}_i(u) : u \in \Re_+\}$ is an optimal solution to problem (1). Furthermore, the collection of assortments $\{\hat{S}_i(u) : u \in \Re_+\}$ also includes an optimal solution to the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$ and if we want to maximize the expected revenue only from nest i, then it is enough to focus on the assortments in the collection $\{\hat{S}_i(u) : u \in \Re_+\}$. Rusmevichientong et al. (2010) consider the problem of maximizing the expected revenue under cardinality constraints when customers choose according to the multinomial logit model. Since the multinomial logit model can be viewed as the nested logit model with a single nest, their problem is similar to the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$ with $\mathcal{C}_i$ corresponding to a cardinality constraint. As a result, they construct their collection of candidate assortments by solving a problem similar to problem (18).

The discussion in this section shows that we can use the collection $\{\hat{S}_i(u) : u \in \Re_+\}$ as candidate assortments for nest i both to maximize the expected revenue over all nests and to maximize the expected revenue only from nest i. However, the particular one of the assortments in this collection that we end up using to maximize the expected revenue over all nests can be entirely different from the particular one that we end up using to maximize the expected revenue only from nest

i. Indeed, it is possible to generate examples such that if we solve the problem $\max_{S_i \in \mathcal{C}_i} R_i(S_i)$ to find an assortment $\tilde{S}_i$ that maximizes the expected revenue only from nest i and use the assortment $(\tilde{S}_1, \ldots, \tilde{S}_m)$ as a possible solution to problem (1), then the assortment $(\tilde{S}_1, \ldots, \tilde{S}_m)$ performs arbitrarily poorly for problem (1).

References

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.

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.