Online Supplement

Efficient Frontier and Applications in Product Offering and Pricing

This note serves the online supplement for the paper titled "Efficient Frontier and Applications in Product Offering and Pricing." In this note, we first provide the additional results and examples not included in the main context. Then, we present all the technical proofs for the findings presented in the paper. At last, we extend the analysis to the case under the nested logit model.

A. Additional Results and Examples

A.1. Unconstrained Assortment and Pricing

Proposition A-1. (Unconstrained Pricing under a Given Assortment) *For any given offer set S, without price bounds for each product, we have the following:*

(a) For any $\rho \in (0,1)$, the optimal pricing to $R(S,\rho)$ in (5) *is uniquely determined by a constant adjusted markup for all products in S, i.e.,*

$$\theta(S,\rho) = p_i - c_i - 1/\beta_i, \quad \forall i \in S, \text{ determined by } \sum_{i\in S} d_i(S, \theta(S,\rho)) = \rho. \tag{A-1}$$

(b) For any proper subset of S, i.e., $S' \subset S$, we have $R(S,\rho) > R(S',\rho)$ for any $\rho \in (0,1)$. Besides, any optimal assortment for problem (5) *is fully capacitated.*

(c) The function $R(S,\rho)$ is concave with respect to the total choice probability $\rho \in (0,1)$, and the optimal solution is denoted by ρ^*, which resolves the optimization problem (2) *with a given S.*

Proof of Proposition A-1. First, we consider the first-order condition (FOC) of the total expected profit with respect to (w.r.t.) price p_j for any $j \in S$, i.e., $\frac{\partial}{\partial p_j}\sum_{i\in S}(p_i - c_i)d_i(S,\mathbf{p}_S) = 0$. Rearranging the FOC equation results in $d_j(S,\mathbf{p}_S)\cdot\left[1-\beta_j(p_j-c_j)+\beta_j\sum_{i\in S}(p_i-c_i)d_i(S,\mathbf{p}_S)\right]=0$. This equation can be satisfied by either setting $d_j(S,\mathbf{p}_S)=0$ that requires $p_j=+\infty$ (which contradicts the case in real business) or letting the inner term of the square bracket equal zero, which is equivalent to

$$p_j - c_j - \frac{1}{\beta_j} = \sum_{i\in S'}(p_i - c_i)d_i(S',\mathbf{p}_{S'}), \tag{A-2}$$

where $d_0(S',\mathbf{p}_{S'}) = 1-\sum_{i\in S'} d_i(S',\mathbf{p}_{S'})$ and S' is the subset of S including all the products with finite prices. The right-hand side (RHS) of equation (A-2) is independent of product index j, so $p_j - c_j - 1/\beta_j$ is constant for all $j \in S'$ at optimality. Denote $\theta = p_j - c_j - 1/\beta_j$, then $p_j = \theta + c_j + 1/\beta_j$. We immediately formulate the problem as $R(S',\rho) = \sum_{i\in S'}(\theta + \frac{1}{\beta_i})d_i(S',\theta)$, *s.t.*, $\sum_{i\in S'} d_i(S',\theta) = \rho$, where $d_i(S',\theta) = \frac{e^{\widetilde{\alpha}_i - \beta_i\theta}}{1+\sum_{s\in S'} e^{\widetilde{\alpha}_s-\beta_s\theta}}$ and $\widetilde{\alpha}_s = \alpha_s - \beta_s c_s - 1$ for all $s \in S'$.

From $\sum_{i\in S'} d_i(S',\theta) = \rho$, we have $\sum_{s\in S'} e^{\widetilde{\alpha}_s-\beta_s\theta} = \frac{\rho}{1-\rho}$. Define $\theta(S',\rho)$ such that total choice probability within the offer set S' equals to ρ. In other words, $\theta(S',\rho)$ is the unique solution to $\sum_{s\in S'} e^{\widetilde{\alpha}_s-\beta_s\theta} = \frac{\rho}{1-\rho}$. Then, we have the following total expected profit formulation: $R(S',\rho) = \rho\theta(S',\rho) + (1-\rho)\sum_{i\in S'} e^{\widetilde{\alpha}_i-\beta_i\theta(S',\rho)}/\beta_i$.

Denote $R_2(S,\rho,\theta)=\rho\theta+(1-\rho)\sum_{i\in S}\frac{e^{\widetilde{\alpha}_i-\beta_i\theta}}{\beta_i}$, where θ is a free variable. Apparently, $R_2(S,\rho,\theta)$ is strictly convex in θ for any $0<\rho<1$ and $S\subseteq\mathcal{M}$. Then, we find

$$\frac{\partial R_2(S,\rho,\theta)}{\partial\theta}\bigg|_{\theta=\theta(S,\rho)}=\rho-(1-\rho)\sum_{i\in S}e^{\widetilde{\alpha}_i-\beta_i\theta}\bigg|_{\theta=\theta(S,\rho)}=0.$$

Thus, for any $S'\subseteq S\subseteq\mathcal{M}$ we can establish the result that $R(S',\rho)$ is the minimum of $R_2(S',\rho,\theta)$ with respect to θ, i.e., $R(S',\rho)=\min_\theta R_2(S',\rho,\theta)$, and $\theta(S',\rho)=\arg\min_\theta R_2(S',\rho,\theta)$.

Then, we will show that all the products should be charged finite prices such that the adjusted markup is constant for all $i\in S$. Consider that another product, say product k, is brought to set S', i.e., product k charges a finite price such that equation (A-2) is satisfied. Denote the new set by $S'^{+}=S'\cup\{k\}$. Similarly, define $R_2(S'^{+},\rho,\theta)=\rho\theta+(1-\rho)\sum_{i\in S'^{+}}\frac{e^{\widetilde{\alpha}_i-\beta_i\theta}}{\beta_i}$, and $R(S'^{+},\rho)$ is the minimum of $R_2(S'^{+},\rho,\theta)$, so we have the following comparison

$$R(S'^{+},\rho)=R_2(S'^{+},\rho,\theta)|_{\theta=\theta(S'^{+},\rho)}>R_2(S',\rho,\theta)|_{\theta=\theta(S'^{+},\rho)}\geq R_2(S',\rho,\theta)|_{\theta=\theta(S',\rho)}=R(S',\rho).$$

The first inequality holds because $R_2(S'^{+},\rho,\theta)>R_2(S',\rho,\theta)$ for any θ.

Therefore, all products should be priced finite such that p_j-c_j-1/β_j is constant for each $j\in S$. It is straightforward that an optimal assortment to problem (5) is fully capacitated.

Consider the first- and second-order derivatives of $R(S,\rho)$ with respect to ρ,

$$\frac{\partial R(S,\rho)}{\partial\rho}=\theta(S,\rho)+\rho\frac{\partial\theta(S,\rho)}{\partial\rho}-\sum_{i\in S}\frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S,\rho)}}{\beta_i}-(1-\rho)\sum_{i\in S}e^{\widetilde{\alpha}_i-\beta_i\theta(S,\rho)}\frac{\partial\theta(S,\rho)}{\partial\rho}=\theta(S,\rho)-\sum_{i\in S}\frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S,\rho)}}{\beta_i},$$
$$\frac{\partial R^2(S,\rho)}{\partial\rho^2}=\frac{\partial\theta(S,\rho)}{\partial\rho}+\sum_{i\in S'}e^{\widetilde{\alpha}_i-\beta_i\theta(S,\rho)}\frac{\partial\theta(S,\rho)}{\partial\rho}=\frac{\partial\theta(S,\rho)}{\partial\rho}(1+\sum_{i\in S}e^{\widetilde{\alpha}_i-\beta_i\theta(S,\rho)})<0,$$

because $\partial\theta(S,\rho)/\partial\rho=-1/(d_0\sum_{i\in S}\beta_i d_i)<0$ by considering the full differentiation of equation $\sum_{i\in S}d_i(S,\theta(S,\rho))=\rho$ with respect to ρ. Therefore, the function $R(S,\rho)$ is concave in ρ for each assortment $S\subseteq\mathcal{M}$.

Moreover, to fully resolve the joint optimization problem on assortment and pricing, we consider the FOC, yielding $\theta(S,\rho)=\sum_{i\in S}\frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S,\rho)}}{\beta_i}$. Since the left-hand side (LHS) of this equation is strictly increasing while the RHS is decreasing, we can immediately find the unique fixed point, denoted by θ^*. Then, the unique optimal price for each product i can be derived as $p_i^*=\theta^*+c_i+1/\beta_i$. Thus, we have completed the proof for Proposition A-1. □

Corollary A-1. (Unconstrained Assortment and Pricing) *Without any constraint on price and capacity, the joint optimal assortment and pricing are fully characterized as:*

(a) All products in consideration should be offered, i.e., $S^*=\mathcal{M}$.

(b) The adjusted markup, i.e., $p_i^*-c_i-1/\beta_i$ *is constant for all products in* $\mathcal{M}$.

(c) The problem in equation (5) *is given by* $r(\rho)=R(\mathcal{M},\rho)$ *that is concave in* $\rho\in(0,1)$.

Proof of Corollary A-1. The results can be derived directly from Proposition A-1, so we omit the details. □

A.2. Examples

Example A-1. We revisit Example 1 and consider the scenario where the prices for each product $i = 1, 2, 3$ are confined within adjusted-markup bounds of $\{\underline{\theta}_i, \overline{\theta}_i\} = \{5.5, 6.5\}$. Then, we can derive the minimum and maximum total choice probabilities for each feasible offer set, e.g., $\{\underline{\rho}(\{1,2\}), \overline{\rho}(\{1,2\}\} = \{0.4390, 0.5351\}$ for offer set $\{1,2\}$, and $\{\underline{\rho}(\{2,3\}), \overline{\rho}(\{2,3\}\} = \{0.3157, 0.3842\}$ for offer set $\{2,3\}$. Clearly, under price constraints, some values of ρ may be unattainable with any single feasible offer set. For instance, a total choice probability of $\rho = 0.4$ cannot be achieved by any single (deterministic) feasible offer set, but can be realized through the strategic combination of two feasible sets, namely $\{1,2\}$ and $\{2,3\}$. This underscores the advantage of employing randomized assortments, particularly when dealing with limited values of ρ, as it offers greater flexibility in meeting specific choice probability targets. ☐

Example A-2. Continuing with Example 1, we consider the dynamic problem. Assume that customers arrive following a nonhomogeneous Poisson process with rates $\lambda_1 = 0.003$, $\lambda_2 = 0.004$, $\lambda_3 = 0.003$, $\lambda_4 = 0.004, \cdots$. The value function $J(x,t)$ is obtained by solving (19) backwardly for each (x,t).

Table A-1 Time Thresholds $t^*(x)$

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$t^*(x)$	0	340	969	1653	2358	3072	3792	4515	5239	5964	6690	7417	8143	8870	9597

The optimal policy for the dynamic product selection and price optimization has a time-threshold structure with thresholds listed in Table A-1. As discussed above, there are two efficient sets: $\{1,2\}$ and $\{2,3\}$; and only one efficient intersection $\mathcal{A}(1,3) = 6.20$, so we omit l in this example for the sake of brevity. For example, $t^*(5) = 2358$ states that the optimal product selection policy for the remaining inventory level $x = 5$ is the following: it is optimal to offer products 2 and 3 if the time-to-go t is greater than 2358; otherwise, it is optimal to offer products 1 and 2. The associated optimal prices can be found by solving equation (21). Table A-1 demonstrates that $t^*(x)$ is increasing in x: the higher level the remaining resource, the more likely to offer products 1 and 2. If the resource is at a higher level, the optimal aggregate resource consumption rate is higher, so it is more profitable to offer products 1 and 2 as shown in Figure 4. ☐

Table A-2 Comparison on Profits under the Static Policy (J^s) and the Optimal Policy (J^*) in a Dynamic Setting

x	J^s	J^*	J^s/J^*	x	J^s	J^*	J^s/J^*	x	J^s	J^*	J^s/J^*	x	J^s	J^*	J^s/J^*
1	15.13	25.33	59.7%	11	107.52	115.26	93.3%	21	148.13	148.74	99.6%	31	156.50	156.51	100.0%
2	28.15	41.85	67.2%	12	113.55	119.92	94.7%	22	150.01	150.47	99.7%	32	156.61	156.62	100.0%
3	39.86	54.83	72.7%	13	119.12	124.20	95.9%	23	151.58	151.92	99.8%	33	156.68	156.69	100.0%
4	50.60	65.80	76.9%	14	124.24	128.19	96.9%	24	152.86	153.10	99.8%	34	156.73	156.73	100.0%
5	60.55	75.38	80.3%	15	128.91	131.96	97.7%	25	153.88	154.06	99.9%	35	156.76	156.76	100.0%
6	69.80	83.88	83.2%	16	133.14	135.49	98.3%	26	154.69	154.81	99.9%	36	156.77	156.77	100.0%
7	78.43	91.50	85.7%	17	136.94	138.75	98.7%	27	155.30	155.38	99.9%	37	156.78	156.78	100.0%
8	86.48	98.36	87.9%	18	140.32	141.71	99.0%	28	155.76	155.82	100.0%	38	156.79	156.79	100.0%
9	94.00	104.55	89.9%	19	143.31	144.37	99.3%	29	156.10	156.13	100.0%	39	156.79	156.79	100.0%
10	101.00	110.16	91.7%	20	145.90	146.71	99.4%	30	156.33	156.36	100.0%	40	156.80	156.80	100.0%

B. Technical Proofs

Proof of Proposition 1. In the proof of Proposition A-1, it has been shown that $R_2(S,\rho,\theta)$ is strictly convex in θ for any $0<\rho<1$ and arbitrary $S\subseteq\mathcal{M}$. Moreover, $R(S,\rho)$ is equal to the minimum of $R_2(S,\rho,\theta)$ with $\theta(S,\rho)=\arg\min_\theta R_2(S,\rho,\theta)$. ☐

Proof of Proposition 2. Suppose that for $0<\rho<1$ an optimizer to $\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)$ is θ^* and $S(\theta^*,\rho)$. Apparently, we have $\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)=\max\limits_{|S|=C}R_2(S,\rho,\theta^*)\geq\max\limits_{|S|=C}\min\limits_{\theta}R_2(S,\rho,\theta)$. The inequality holds because $R_2(S,\rho,\theta^*)\geq\min\limits_{\theta}R_2(S,\rho,\theta)$ for each $S\subseteq\mathcal{M}$. For completeness, we next establish the existence of the optimizer θ^*. Recall that $R_2(S,\rho,\theta)=\rho\theta+(1-\rho)\sum_{i\in S}\frac{e^{\tilde{\alpha}_i-\beta_i\theta}}{\beta_i}$ is strictly convex in θ for any $0<\rho<1$ and $S\subseteq\mathcal{M}$ in the proof of Proposition 1. Thus, $\max\limits_{|S|=C}R_2(S,\rho,\theta)$ is convex in θ by the preservation of convexity (see, e.g., Boyd and Vandenberghe 2009). Then, there exists a solution to $\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)$ for ρ falling within a specific range, which will be explicitly explained in the proof of Lemma B-2.

If there exists a set S^* (with $|S^*|=C$) such that $\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)=\min\limits_{\theta}R_2(S^*,\rho,\theta)$, we consider

$$\max_{|S|=C}\min_{\theta}R_2(S,\rho,\theta)\geq\min_{\theta}R_2(S^*,\rho,\theta)=\min_{\theta}\max_{|S|=C}R_2(S,\rho,\theta).$$

Thus, $\max\limits_{|S|=C}\min\limits_{\theta}R_2(S,\rho,\theta)=\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)$. For the existence and the solution of S^*, we will also present the detailed arguments in the proof of Lemma B-2. ☐

Proof of Proposition 3. Before presenting the proof, we will need following few Lemmas.

LEMMA B-1. *Function $R_3(\rho,\theta)$ is strictly convex in θ for any $\rho\in(0,1)$.*

Proof of Lemma B-1. First, the exponential function $e^{\tilde{\alpha}_i-\beta_i\theta}$ is convex in θ for each $i\in\mathcal{M}$, so $\max\limits_{|S|=C}\sum_{i\in S}\frac{e^{\tilde{\alpha}_i-\beta_i\theta}}{\beta_i}$ is convex in θ by the preservation of convexity (see, e.g., Boyd and Vandenberghe 2009). Then, it is straightforward that $R_3(\rho,\theta)$ is strictly convex in θ for any $\rho\in(0,1)$. ☐

LEMMA B-2. *If $\theta(\rho):=\arg\min_\theta R_3(\rho,\theta)$ is not an efficient intersection, i.e., $\theta(\rho)\notin\mathcal{E}$, then $r(\rho)=\max\limits_{|S|=C}\min\limits_{\theta}R_2(S,\rho,\theta)=\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)=\min\limits_{\theta}R_3(\rho,\theta)$.*

Proof of Lemma B-2. From Lemma B-1, $R_3(\rho,\theta)$ is strictly convex in θ for any $0<\rho<1$, therefore $\theta(\rho)$ is well-defined. Note that $R_3(\rho,\theta)$ is differentiable in θ when θ is not an efficient intersection. If $\theta(\rho)$ is not equal to any efficient intersection, it is the unique solution to $\partial R_3(\rho,\theta)/\partial\theta=0$, i.e., $\theta(\rho)$ satisfies $\rho-(1-\rho)\sum_{i\in S_C^{l(\rho)}}e^{\tilde{\alpha}_i-\beta_i\theta}=0$ for ρ in a certain range. In other words, there must be a set, denoted by $S_C^{l(\rho)}$, such that $\min\limits_{\theta}R_3(\rho,\theta)=R_3(\rho,\theta(\rho))=R_2(S_C^{l(\rho)},\rho,\theta(\rho))=\min\limits_{\theta}\max\limits_{|S|=C}R_2(S,\rho,\theta)$.

Specifically, the set $S_C^{l(\rho)}$ can be identified as follows. Noting that for θ varying between any two successive efficient intersections, e.g., for any $\theta\in(\Theta_l,\Theta_{l+1})$, the candidate assortment S_C^l resolves the problem (11), which is then given by

$$R_3(\rho,\theta)=\rho\theta+(1-\rho)\mathcal{H}(\theta),\qquad\text{where}\qquad\mathcal{H}(\theta)=\sum_{i\in S_C^l}h_i(\theta),\quad\forall\,\theta\in\big(\Theta_l,\Theta_{l+1}\big).\tag{B-1}$$

For any $\theta \in (\Theta_l, \Theta_{l+1})$, the corresponding optimal offer set S_C^l, and we can calculate the corresponding interval of ρ. Then, the unit region $(0,1)$ can be divided into multiple segments such that when ρ falls within a specific segmented interval, an offer set $S_C^{l(\rho)}$ exists and $S_C^{l(\rho)} = S_C^l$ for ρ varying in the same interval. Such a candidate assortment serves as the optimal S^* presented in Proposition 2, thus, the strong duality holds, implying $\max\limits_{|S|=C} \min\limits_{\theta} R_2(S,\rho,\theta) = \min\limits_{\theta} \max\limits_{|S|=C} R_2(S,\rho,\theta)$. ☐

Lemma B-3. *For any $l = 1, 2, \ldots, L$, $S_C^l = S_C^{l-1} \setminus \{i_l\} \cup \{j_l\}$ where i_l, j_l are associated with the two $\{h_i(\theta)\}_{i\in\mathcal{M}}$ curves whose intersection is Θ_l, i.e., $\Theta_l = \mathcal{A}(i_l, j_l)$.*

Proof of Lemma B-3. First, we focus on the case that only two curves intersect at any intersection. Recall that $\mathcal{A}(i_l, j_l)$ is an intersection between the two curves $\{h_{i_l}(\theta), h_{j_l}(\theta)\}$. If i_l, j_l are both inside or outside S_C^l, the candidate assortment does not change when θ crosses point $\mathcal{A}(i_l, j_l)$, i.e., $S_C^l = S_C^{l-1}$, this violates the definition of efficient intersection of Θ_l. Thus, one of $\{i_l, j_l\}$ is inside S_C^l and the other is outside S_C^l. Since $\beta_{i_l} > \beta_{j_l}$ by definition of $\mathcal{A}(i_l, j_l)$, $h_{i_l}(\theta) > h_{j_l}(\theta)$ for any $\theta < \mathcal{A}(i_l, j_l)$ and $h_{i_l}(\theta) < h_{j_l}(\theta)$ for any $\theta > \mathcal{A}(i_l, j_l)$, thus, $i_l \in S_C^{l-1}$ and $j_l \in S_C^l$. Therefore, candidate assortment $S_C^l = S_C^{l-1} \setminus \{i_l\} \cup \{j_l\}$.

In addition, we consider the possibility that more than two curves intersect at the same point $\Theta_l = \mathcal{A}(i_l, j_l)$. For this case, we can group the curves inside S_C^{l-1} by $\{i_l\}'s$, and the curves outside S_C^{l-1} by $\{j_l\}'s$. Note that both groups exclude the curve(s) inside (or outside) both two successive candidate assortments S_C^l and S_C^{l-1}. Thus, both groups $\{i_l\}'s$ and $\{j_l\}'s$ have the same number of curves. Then, we can express the candidate assortments as $S_C^l = S_C^{l-1} \setminus \{i_l\}'s \cup \{j_l\}'s$. For simplicity of presentation, we will focus on the case with only two curves intersecting at an intersection. We claim that this is only made for notation simplicity, the same arguments can be used to establish results for the case with more than two curves intersect at the same intersection. Furthermore, focusing on the case where only two curves intersect at each intersection is important because it represents the worst-case scenario with the largest number of intersections. ☐

Now, we are ready to present the proof of Proposition 3. First, we will show $\rho_1^- > \rho_1^+ > \ldots > \rho_L^- > \rho_L^+$. Because $e^{\widehat{\alpha}_{i_l} - \beta_{i_l}\Theta_l} = e^{\widehat{\alpha}_{j_l} - \beta_{j_l}\Theta_l}$ and $\beta_{i_l} > \beta_{j_l}$ by the definition of $\mathcal{A}(i_l, j_l)$, then $e^{\widetilde{\alpha}_{i_l} - \beta_{i_l}\Theta_l} > e^{\widetilde{\alpha}_{j_l} - \beta_{j_l}\Theta_l}$. Since $S_C^l = S_C^{l-1} \setminus \{i_l\} \cup \{j_l\}$ by Lemma B-3, then $\rho_l^+ < \rho_l^-$. By definition $\left.\frac{\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i\theta}}{1+\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i\theta}}\right|_{\theta=\Theta_{l-1}} = \rho_{l-1}^+$, then $\rho_{l-1}^+ > \rho_l^-$ because $\Theta_{l-1} < \Theta_l$, thus, we have the sequence $\rho_1^- > \rho_1^+ > \ldots > \rho_L^- > \rho_L^+$ for all efficient intersections.

Next, we show S_C^{l-1} is the optimal assortment for any $\rho_l^- < \rho < \rho_{l-1}^+$. First, we focus on problem $\max\limits_{|S|=C} R_2(S,\rho,\theta)$. One can see that once the range of θ is given, the optimal offer set(s) to equation (11) is determined, and the corresponding region of ρ can be calculated. In particular, for $\theta \in (\Theta_{l-1}, \Theta_l)$, the optimal assortment is S_C^{l-1} and the corresponding range for ρ is (ρ_l^-, ρ_{l-1}^+), where $\left.\frac{\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i\theta}}{1+\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i\theta}}\right|_{\theta=\Theta_l} = \rho_l^-$ and $\left.\frac{\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i\theta}}{1+\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i\theta}}\right|_{\theta=\Theta_{l-1}} = \rho_{l-1}^+$ based on the definitions in (13). It is clear that $\theta(\rho) = \arg\min_\theta R_3(\rho,\theta)$ is not equal to any efficient intersection for any $\rho_l^- < \rho < \rho_{l-1}^+$ based on the above analysis, and we have $\max\limits_{|S|=C} \min\limits_{\theta} R_2(S,\rho,\theta) = \min\limits_{\theta} \max\limits_{|S|=C} R_2(S,\rho,\theta)$ from Lemma B-2 for the case of $\theta(\rho)$ not being an efficient intersection. Therefore, the optimal assortment for problem (5) is set S_C^{l-1} for any $\rho_l^- < \rho < \rho_{l-1}^+$.

By Proposition 2, $r(\rho) = \max\limits_{|S|=C} \min\limits_{\theta} R_2(S,\rho,\theta) = \min\limits_{\theta} \max\limits_{|S|=C} R_2(S,\rho,\theta) = \min_\theta \Big(\rho\theta + (1-\rho)\mathcal{H}(\theta)\Big)$, $\forall \rho_l^- < \rho < \rho_{l-1}^+$. This can also be expressed more explicitly as follows: $r(\rho) = \rho\theta(S_C^{l-1},\rho) + (1-\rho)\sum_{i\in S_C^{l-1}} h_i(\theta(S_C^{l-1},\rho))$, $\forall \rho_l^- < \rho < \rho_{l-1}^+$, where $\theta(S_C^{l-1},\rho)$ is uniquely determined by $\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i-\beta_i\theta} = \frac{\rho}{1-\rho}$. Therefore, $r(\rho) = \max\limits_{|S|=C} R(S,\rho) = R(S_C^{l-1},\rho), \quad \forall \rho \in (\rho_l^-, \rho_{l-1}^+)$. □

Proof of Theorem 1. Since $R(S,\rho)$ equals $R_2(S,\rho,\theta(S,\rho))$ defined in equation (8), $R(S_C^{l-1},\rho) = R(S_C^l,\rho)$ can be rewritten as follows:

$$\begin{aligned}
&\textstyle\sum_{i\in S_C^l} e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^l,\rho)} = \frac{\rho}{1-\rho}, \quad \sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^{l-1},\rho)} = \frac{\rho}{1-\rho},\\
&\textstyle\rho\theta(S_C^l,\rho) + (1-\rho)\sum_{i\in S_C^l} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^l,\rho)}}{\beta_i} = \rho\theta(S_C^{l-1},\rho) + (1-\rho)\sum_{i\in S_C^{l-1}} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^{l-1},\rho)}}{\beta_i}.
\end{aligned} \tag{B-2}$$

We will show that there exists a unique solution in $[\rho_l^+, \rho_l^-]$ to equations (B-2), denoted by ρ_l. Consider the derivative of $R(S_C^{l-1},\rho) - R(S_C^l,\rho)$ as follows:

$$\frac{\partial}{\partial\rho}\left(R(S_C^{l-1},\rho) - R(S_C^l,\rho)\right) = \left(\theta(S_C^{l-1},\rho) - \sum_{i\in S_C^{l-1}} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^{l-1},\rho)}}{\beta_i}\right) - \left(\theta(S_C^l,\rho) - \sum_{i\in S_C^l} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^l,\rho)}}{\beta_i}\right),$$

where $\theta(S_C^l,\rho)$ and $\theta(S_C^{l-1},\rho)$ are uniquely determined by equations $\sum_{i\in S_C^l} e^{\widetilde{\alpha}_i-\beta_i\theta} = \frac{\rho}{1-\rho}$ and $\sum_{i\in S_C^{l-1}} e^{\widetilde{\alpha}_i-\beta_i\theta} = \frac{\rho}{1-\rho}$, respectively.

For any $\rho_l^+ < \rho \le \rho_l^-$, $\theta(S_C^l,\rho) < \Theta_l \le \theta(S_C^{l-1},\rho)$ because $\theta(S,\rho)$ is decreasing in ρ for each given assortment S. Then, $\sum_{i\in S_C^{l-1}} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^{l-1},\rho)}}{\beta_i} \le \mathcal{H}(\Theta_l) < \sum_{i\in S_C^l} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^l,\rho)}}{\beta_i}$. Thus, $\frac{\partial R(S_C^{l-1},\rho)}{\partial\rho} > \frac{\partial R(S_C^l,\rho)}{\partial\rho}$, $\forall \rho_l^+ < \rho \le \rho_l^-$.

Compare $R(S_C^{l-1},\rho_l^+)$ and $R(S_C^{l-1},\rho_l^-)$ to $R(S_C^l,\rho_l^+)$ and $R(S_C^l,\rho_l^-)$, respectively, yielding

$$R(S_C^l,\rho_l^+) = \min_\theta R_2(S_C^l,\rho_l^+,\theta) = R_2(S_C^l,\rho_l^+,\Theta_l), \text{ and } R(S_C^{l-1},\rho_l^+) = \min_\theta R_2(S_C^{l-1},\rho_l^+,\theta) < R_2(S_C^{l-1},\rho_l^+,\Theta_l).$$

Thus, $R(S_C^l,\rho_l^+) > R(S_C^{l-1},\rho_l^+)$ as $R_2(S_C^l,\rho_l^+,\Theta_l) = R_2(S_C^{l-1},\rho_l^+,\Theta_l)$. Similarly, $R(S_C^{l-1},\rho_l^-) > R(S_C^l,\rho_l^-)$.

Recalling that $R(S_C^{l-1},\rho) - R(S_C^l,\rho)$ is strictly increasing in ρ, and each $R(S_C^l,\rho)$ for any l is continuous in ρ according to its specific expression, i.e., $R(S_C^l,\rho) = \rho\theta(S_C^l,\rho) + (1-\rho)\sum_{i\in S_C^l} \frac{e^{\widetilde{\alpha}_i-\beta_i\theta(S_C^l,\rho)}}{\beta_i}$, thus there exists a unique solution to $R(S_C^{l-1},\rho) = R(S_C^l,\rho)$ for $\rho_l^+ < \rho \le \rho_l^-$, denoted by ρ_l. Therefore, $R(S_C^l,\rho) < R(S_C^{l-1},\rho)$ for any $\rho_l < \rho \le \rho_l^-$ and $R(S_C^l,\rho) > R(S_C^{l-1},\rho)$ for any $\rho_l^+ < \rho < \rho_l$.

Note that the key idea to prove the results in Theorem 1 is by showing that the equation $R(S_C^{l-1},\rho) = R(S_C^l,\rho)$ has a unique solution ρ_l in interval $[\rho_l^+, \rho_l^-]$. Then we can fully characterize the joint optimization problem on assortment and pricing. Here we only compare the revenues or profits generated from offer sets S_C^l and S_C^{l-1}. This is because for any $\rho \in [\rho_l^+, \rho_l^-]$, we have

$$R(S_C^l,\rho) = R_2(S_C^l,\rho,\theta(S_C^l,\rho)) > R_2(S_C^k,\rho,\theta(S_C^l,\rho)) > \min_\theta R_2(S_C^k,\rho,\theta) = R(S_C^k,\rho), \forall k > l, \tag{B-3}$$

$$R(S_C^{l-1},\rho) = R_2(S_C^{l-1},\rho,\theta(S_C^{l-1},\rho)) > R_2(S_C^k,\rho,\theta(S_C^{l-1},\rho)) > \min_\theta R_2(S_C^k,\rho,\theta) = R(S_C^k,\rho), \forall k < l-1. \tag{B-4}$$

In equation (B-3), the first inequality holds because $\theta(S_C^l,\rho) \le \Theta_l$ for any $\rho \in [\rho_l^+, \rho_l^-]$, thus any candidate assortment with larger index generates less profit. By similar arguments, the first inequality in (B-4) holds. Therefore, other offer sets are dominated by S_C^l and S_C^{l-1} on profitability for $\rho \in [\rho_l^+, \rho_l^-]$. Consequently, we only need to consider the revenues of offer sets S_C^l and S_C^{l-1} in this region to characterize $r(\rho)$.

Similarly, we can construct a sequence $\rho_{L+1} < \rho_L < \ldots < \rho_1 < \rho_0$ with each $\rho_l^+ \leq \rho_l \leq \rho_l^-$ for $l = 1, 2, \ldots, L$. Note that ρ_{L+1} and ρ_0 are added for describing the optimal policy, and they are actually two inaccessible boundary points of ρ. Thus, the optimal assortment is S_C^{l-1} for any $\rho \in (\rho_l, \rho_{l-1})$, and assortments S_C^{l-1} and S_C^l are indifferent at $\rho = \rho_l$. The ρ_l's can be efficiently calculated by Algorithm 1. $\square$

Algorithm 1 Finding ϵ-optimal ρ_l

```
step 0 (initialization): Set $\overline{\rho}_l^{(0)} = \rho_l^-$, $\underline{\rho}_l^{(0)} = \rho_l^+$, and the stop criterion $\epsilon > 0$.
step 1 (iteration $k$, part a): Denote the middle point as $\sigma_l^{(k)} = \big(\overline{\rho}_l^{(k-1)} + \underline{\rho}_l^{(k-1)}\big)/2$. Solve $\sigma_l^{(k)} - (1 - \sigma_l^{(k)}) \sum_{i \in S_C^l} e^{\widetilde{\alpha}_i - \beta_i \theta_l} = 0$ to obtain $\theta_l^{(k)}$; solve $\sigma_l^{(k)} - (1 - \sigma_l^{(k)}) \sum_{i \in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i \theta_{l-1}} = 0$ to obtain $\theta_{l-1}^{(k)}$.
step 2 (iteration $k$, part b): Compute slope $\gamma_l^{(k)} := \Big( \sum_{i \in S_C^l} e^{\widetilde{\alpha}_i - \beta_i \theta_l^{(k)}} / \beta_i - \sum_{i \in S_C^{l-1}} e^{\widetilde{\alpha}_i - \beta_i \theta_{l-1}^{(k)}} / \beta_i \Big) \Big/ \big(\theta_l^{(k)} - \theta_{l-1}^{(k)}\big)$.
    (a) If $\gamma_l^{(k)} > -\frac{\sigma_l^{(k)}}{1-\sigma_l^{(k)}}$, set $\overline{\rho}_l^{(k)} = \sigma_l^{(k)}$ and $\underline{\rho}_l^{(k)} = \underline{\rho}_l^{(k-1)}$;
    (b) If $\gamma_l^{(k)} < -\frac{\sigma_l^{(k)}}{1-\sigma_l^{(k)}}$, set $\overline{\rho}_l^{(k)} = \overline{\rho}_l^{(k-1)}$ and $\underline{\rho}_l^{(k)} = \sigma_l^{(k)}$;
    (c) Otherwise, stop.
step 3 (iteration $k$, part c): If $|\overline{\rho}_l^{(k)} - \underline{\rho}_l^{(k)}| < \epsilon$, stop; otherwise, then go to step 1.
```

Proof of Corollary 1. The result follows from Theorem 1 and Proposition A-1, so we omit the details. $\square$

Proof of Proposition 4. According to Definition 1 and formation (4), an efficient frontier function is increasing and concave. Note that the concavity for Formation (4) follows from the standard parametric linear programming results (see, e.g., Bertsimas and Tsitsikis 1997, Theorem 5.1, p.213).

For the explicit expression of the efficient frontier function $\overline{r}(\rho)$ under cardinality constraints, consider the following equations:

$$\frac{\partial R(S_C^{l-1}, \rho)}{\partial \rho}\bigg|_{\rho=\rho'} = \frac{\partial R(S_C^l, \rho)}{\partial \rho}\bigg|_{\rho=\rho''} = \frac{R(S_C^l, \rho') - R(S_C^{l-1}, \rho'')}{\rho' - \rho''}. \tag{B-5}$$

We will next show that the portfolio (ρ_l^-, ρ_l^+) is the unique solution to (B-5) in the region $[\rho_l^+, \rho_l^-]$.

$$\frac{\partial R(S_C^{l-1}, \rho)}{\partial \rho}\bigg|_{\rho=\rho_l^-} = \theta - \sum_{s \in S_C^{l-1}} \frac{e^{\widetilde{\alpha}_s - \beta_s \theta}}{\beta_s}\bigg|_{\Theta_l} = \Theta_l - \sum_{s \in S_C^{l-1}} e^{\widehat{\alpha}_s - \beta_s \Theta_l}.$$

Similarly, $\frac{\partial R(S_C^l, \rho)}{\partial \rho}\big|_{\rho=\rho_l^+} = \Theta_l - \sum_{s \in S_C^l} e^{\widehat{\alpha}_s - \beta_s \Theta_l}$. Because $\sum_{s \in S_C^{l-1}} e^{\widehat{\alpha}_s - \beta_s \Theta_l} = \sum_{s \in S_C^l} e^{\widehat{\alpha}_s - \beta_s \Theta_l}$, then we can observe $\frac{\partial R(S_C^{l-1}, \rho)}{\partial \rho}\big|_{\rho=\rho_l^-} = \frac{\partial R(S_C^l, \rho)}{\partial \rho}\big|_{\rho=\rho_l^+}$. Consider $r(\rho)$ at points ρ_l^- and ρ_l^+, $r(\rho_l^-) = R(S_C^{l-1}, \rho_l^-) = \rho_l^- \Theta_l + (1 - \rho_l^-) \sum_{s \in S_C^{l-1}} e^{\widehat{\alpha}_s - \beta_s \Theta_l}$ and $r(\rho_l^+) = R(S_C^l, \rho_l^+) = \rho_l^+ \Theta_l + (1 - \rho_l^+) \sum_{s \in S_C^l} e^{\widehat{\alpha}_s - \beta_s \Theta_l}$. Then,

$$\frac{R(S_C^{l-1}, \rho_l^-) - R(S_C^l, \rho_l^+)}{\rho_l^- - \rho_l^+} = \Theta_l - \sum_{s \in S_C^l} e^{\widehat{\alpha}_s - \beta_s \Theta_l} = \frac{\partial R(S_C^{l-1}, \rho_l^-)}{\partial \rho^-} = \frac{\partial R(S_C^l, \rho_l^+)}{\partial \rho_l^+}. \tag{B-6}$$

Thus, we have proven that (ρ_l^-, ρ_l^+) is a solution to equations (B-5). Moreover, we have $\frac{\partial R(S_C^{l-1}, \rho)}{\partial \rho} > \frac{\partial R(S_C^l, \rho)}{\partial \rho}$, $\forall \rho_l^+ < \rho < \rho_l^-$. The above inequality holds because both $\partial R(S_C^{l-1}, \rho)/\partial \rho$ and $\partial R(S_C^l, \rho)/\partial \rho$ are decreasing in ρ, implied by the concavity of $R(S_C^{l-1}, \rho)$ and $R(S_C^l, \rho)$.

For $\rho_l^- < \rho < \rho_{l-1}^+ < \rho_{l-1}^-$, $r(\rho) = R(S_C^{l-1}, \rho)$ is concave in ρ and $\overline{r}(\rho) = r(\rho)$. Since $R(S_C^l, \rho)$ and $R(S_C^{l-1}, \rho)$ are both strictly concave in ρ, for any $\rho_l^+ \leq \rho \leq \rho_l^-$,

$$R(S_C^l, \rho) \leq R(S_C^l, \rho_l^+) + \tfrac{\partial R(S_C^l, \rho_l^+)}{\partial \rho_l^+} \cdot (\rho - \rho_l^+) = \tfrac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+} \cdot R(S_C^l, \rho_l^+) + \tfrac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+} \cdot R(S_C^{l-1}, \rho_l^-),$$

$$R(S_C^{l-1}, \rho) \leq R(S_C^{l-1}, \rho_l^-) - \tfrac{\partial R(S_C^{l-1}, \rho_l^-)}{\partial \rho^-} \cdot (\rho^- - \rho) = \tfrac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+} \cdot R(S_C^l, \rho_l^+) + \tfrac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+} \cdot R(S_C^{l-1}, \rho_l^-),$$

where $\frac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+} + \frac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+} = 1$ and $\frac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+} \cdot \rho_l^+ + \frac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+} \cdot \rho_l^- = \rho$. Therefore, for $\rho_l^+ \leq \rho \leq \rho_l^-$,

$$\overline{r}(\rho) = \tfrac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+} \cdot R(S_C^l, \rho_l^+) + \tfrac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+} \cdot R(S_C^{l-1}, \rho_l^-) = \tfrac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+} \cdot r(\rho_l^+) + \tfrac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+} \cdot r(\rho_l^-).$$

For $\rho \geq \rho^*$, it is obvious that $\overline{r}(\rho) = r(\rho^*)$. ☐

Proof of Proposition 5. To show $\big(p_i(\theta)\big)_{i \in S}$ is a feasible solution to problem $R(S, \rho(\theta))$ in (15) under the price-bound constraints, we consider the equivalent problem of $R(S, \rho(\theta))$ in (15) as follows:

$$\begin{aligned} \max_{\mathbf{p}_S} \ & \textstyle\sum_{i \in S} (p_i - c_i) d_i(S, \mathbf{p}_S) \\ s.t., \ & l_i \leq p_i \leq u_i \ \forall i \in S, \ \textstyle\sum_{i \in S} e^{\alpha_i - \beta_i p_i} = \tfrac{\rho}{1-\rho}. \end{aligned} \tag{B-7}$$

Note that the condition $d_i(S, \mathbf{p}_S) \geq 0$ for any $i \in S$ in (15) is always satisfied under the MNL model.

Let $w_i = e^{\alpha_i - \beta_i p_i}$. Then, $p_i = (\alpha_i - \log(w_i))/\beta_i$. The above problem (B-7) can be rewritten as follows

$$\begin{aligned} \max_{\mathbf{w}_S} \ & (1-\rho) \cdot \textstyle\sum_{i \in S} \left(\tfrac{\alpha_i - \log(w_i)}{\beta_i} - c_i \right) w_i \\ s.t., \ & e^{\alpha_i - \beta_i u_i} \leq w_i \leq e^{\alpha_i - \beta_i l_i} \ \forall i \in S, \ \textstyle\sum_{i \in S} w_i = \tfrac{\rho}{1-\rho}. \end{aligned} \tag{B-8}$$

The concavity of the objective function can easily be established because of the concavity of each $\left(\frac{\alpha_i - \log(w_i)}{\beta_i} - c_i \right) w_i$ in $\mathbf{w}_S$ and the preservation of concavity in Boyd and Vandenberghe (2009). In addition, the constraints are all linear in $\mathbf{w}_S$. Then, we consider the Karush-Kuhn-Tucker (KKT) conditions for problem (B-8):

$$\begin{gathered} \frac{(\alpha_i - \beta_i c_i - 1) - \log(w_i)}{\beta_i} - (\mu_i - \nu_i) + \eta = 0; \\ e^{\alpha_i - \beta_i u_i} \leq w_i \leq e^{\alpha_i - \beta_i l_i}, \forall i \in S, \sum_{i \in S} w_i = \frac{\rho}{1-\rho}, \\ \mu_i (w_i - e^{\alpha_i - \beta_i l_i}) = 0, \ \nu_i (e^{\alpha_i - \beta_i u_i} - w_i) = 0, \ \mu_i, \nu_i \geq 0, \forall i \in S. \end{gathered} \tag{B-9}$$

We will next construct a solution to the KKT conditions. Let $w_i^*(\theta) = e^{\alpha_i - \beta_i p_i(\theta)}$; let $\mu_i^* = 0$ if $l_i < p_i(\theta)$; let $\nu_i^* = 0$ if $p_i(\theta) < u_i$.

If there exists an $i \in S$ such that $l_i < p_i(\theta) < u_i$, the equation (B-9) becomes $p_i(\theta) - c_i - \frac{1}{\beta_i} + \eta = 0$. Combing it with the efficient price defined in (6) results in $\eta^* = -\theta$.

If $p_i(\theta) = l_i$, the equation (B-9) becomes $l_i - c_i - \frac{1}{\beta_i} - \mu_i + \eta^* = 0$, then $\mu_i^* = l_i - c_i - \frac{1}{\beta_i} - \theta$; if $p_i(\theta) = u_i$, the equation (B-9) becomes $u_i - c_i - \frac{1}{\beta_i} + \nu_i + \eta^* = 0$, then $-\nu_i^* = u_i - c_i - \frac{1}{\beta_i} - \theta$. So far, we have constructed a unique solution $(w_i^*(\theta), \mu_i^*, \nu_i^*, \eta^*)$ to the KKT conditions if there exists an $i \in S$ such that $l_i < p_i(\theta) < u_i$.

If there does not exist an $i \in S$ such that $l_i < p_i(\theta) < u_i$, the solution to the KKT conditions is not unique and $(w_i^*(\theta), \mu_i^*, \nu_i^*, \eta^*)$ is one of them. However, $(w_i^*(\theta))_{i \in S}$ is the unique solution to problem (B-8). It can be shown that the objective function in problem (B-8) is strictly jointly concave in the vector $\mathbf{w}_S$ because $\partial^2 (w_i \log(w_i)) / \partial w_i^2 = 1/w_i > 0$. The constraints in problem (B-8) are affine, so the optimal solution is unique. Thus, the efficient price vector $\big(p_i(\theta)\big)_{i \in S}$ is the unique solution to $R(S, \rho(\theta))$ defined in problem (15). ☐

LEMMA B-4. *The function $h_i(\theta)$ is continuous, decreasing, differentiable and convex in θ.*

Proof of Lemma B-4. Substituting efficient prices into $h_i(\theta)$, we have

$$h_i(\theta) := (p_i(\theta) - c_i - \theta) \cdot e^{\alpha_i - \beta_i p_i(\theta)} = \begin{cases} (l_i - c_i - \theta) e^{\alpha_i - \beta_i l_i}, & \theta < \underline{\theta}_i, \\ e^{\alpha_i - \beta_i c_i - 1 - \log(\beta_i) - \beta_i \theta}, & \underline{\theta}_i \leq \theta \leq \overline{\theta}_i, \\ (u_i - c_i - \theta) e^{\alpha_i - \beta_i u_i}, & \theta > \overline{\theta}_i. \end{cases} \tag{B-10}$$

To show the continuity of $h_i(\theta)$, we observe

$$\lim_{\theta \nearrow l_i - c_i - 1/\beta_i} (l_i - c_i - \theta) e^{\alpha_i - \beta_i l_i} = 1/\beta_i \cdot e^{\alpha_i - \beta_i c_i - 1 - \beta_i \theta}\big|_{\theta = l_i - c_i - 1/\beta_i} = 1/\beta_i \cdot e^{\alpha_i - \beta_i l_i},$$
$$\lim_{\theta \searrow u_i - c_i - 1/\beta_i} (u_i - c_i - \theta) e^{\alpha_i - \beta_i u_i} = 1/\beta_i \cdot e^{\alpha_i - \beta_i c_i - 1 - \beta_i \theta}\big|_{\theta = u_i - c_i - 1/\beta_i} = 1/\beta_i \cdot e^{\alpha_i - \beta_i u_i},$$

where $\nearrow$ represents the left-sided limit and $\searrow$ represents the right-sided limit. To show $h_i(\theta)$ is differentiable in any θ, we consider

$$\lim_{\theta \nearrow l_i - c_i - 1/\beta_i} \frac{\partial\big((l_i - c_i - \theta) e^{\alpha_i - \beta_i l_i}\big)}{\partial \theta} = \frac{\partial\big(1/\beta_i \cdot e^{\alpha_i - \beta_i c_i - 1 - \beta_i \theta}\big)}{\partial \theta}\Big|_{\theta = l_i - c_i - 1/\beta_i} = -e^{\alpha_i - \beta_i l_i},$$
$$\lim_{\theta \searrow u_i - c_i - 1/\beta_i} \frac{\partial\big((u_i - c_i - \theta) e^{\alpha_i - \beta_i u_i}\big)}{\partial \theta} = \frac{\partial\big(1/\beta_i \cdot e^{\alpha_i - \beta_i c_i - 1 - \beta_i \theta}\big)}{\partial \theta}\Big|_{\theta = u_i - c_i - 1/\beta_i} = -e^{\alpha_i - \beta_i u_i}.$$

The decreasing monotonicity is obvious. Furthermore, $h_i(\theta)$ is a smooth function and the linear functions at both sides are the tangents of the exponential function in between, so $h_i(\theta)$ is decreasing convex with respect to θ. Therefore, the desired structural properties for the function $h_i(\theta)$ hold as claimed. □

Proof of Proposition 6. Under price-bound constraints, the piece-wise function $h_i(\theta)$ in (B-10) may not always be positive. For certain values of θ, the product i whose corresponding $h_i(\theta)$ is negative will be removed from the offer set in order to maximize $\sum_{i \in S} h_i(\theta)$, although there are no cardinality constraints. Therefore, we only need to compare each $h_i(\theta), i \in \mathcal{M}$ with $h_0(\theta) := 0$ to calculate the intersection point. Clearly, all intersections between the right-end linear line of any $h_i(\theta)$ and $h_0(\theta)$ are efficient intersections, so there are m efficient intersections. Recall that products are relabeled in the decreasing order according to their *max markups*, i.e., $u_1 - c_1 \geq u_2 - c_2 \geq \cdots \geq u_m - c_m$. Let Θ_i denote the intersection point between $h_{m+1-i}(\theta)$ and $h_0(\theta)$. For example, Θ_1 represents the intersection point between $h_m(\theta)$ and $h_0(\theta)$. Immediately, the ordering of the intersections is given by $0 = \Theta_0 < \Theta_1 \leq \cdots \leq \Theta_m < \Theta_{m+1} = +\infty$. Again, two end points Θ_0 and Θ_{m+1} are added for describing the algorithm. For a given $\theta \in (\Theta_l, \Theta_{l+1}]$, we select the first $m - l$ products whose corresponding $h_i(\theta)$'s are positive values for the offer set, leading to the candidate assortment $S^l := \{1, 2, \ldots, m - l\}$ for any $l = 0, 1, \cdots, m - 1$. Candidate assortments change across the efficient intersections, thus, there are m candidate assortments following the max markup ordering. □

Proof of Proposition 7 First, we rewrite the function $h_i(\theta)$ with price bounds $l_i \leq p_i \leq u_i$ as follows:

$$h_i(\theta) := \begin{cases} (l_i - c_i - \theta) e^{\alpha_i - \beta_i l_i}, & \theta < \underline{\theta}_i, \\ e^{\widehat{\alpha}_i - \beta_i \theta}, & \underline{\theta}_i \leq \theta \leq \overline{\theta}_i, \\ (u_i - c_i - \theta) e^{\alpha_i - \beta_i u_i}, & \theta > \overline{\theta}_i, \end{cases}$$

where $\underline{\theta}_i = l_i - c_i - 1/\beta_i$ and $\overline{\theta}_i = u_i - c_i - 1/\beta_i$. Lemma B-4 has shown that the $h_i(\theta)$ function is differentiable in θ for each i. Moreover, this smooth function is characterized by having linear functions on both sides that act as tangents to the exponential function in between. As a result, each $h_i(\theta)$ is decreasing convex with respect to θ. Now, we consider the number of intersections between $h_i(\theta)$ and $h_j(\theta)$, for any $j \neq i$.

Note that there exists at most one intersection between any two exponential functions, i.e., there is at most one solution regarding θ to equation $e^{\widehat{\alpha}_i - \beta_i \theta} = e^{\widehat{\alpha}_j - \beta_j \theta}$ for any $j \neq i$. Similarly, there is also at most one intersection between any two linear functions, i.e., at most one θ satisfies the equation $(l_i - c_i - \theta)e^{\alpha_i - \beta_i l_i} = (l_j - c_j - \theta)e^{\alpha_j - \beta_j l_j}$. Regarding the number of intersections between a linear function and an exponential function, we will show that there can be at most two. Consider a function $g(\theta) = (l_i - c_i - \theta)e^{\alpha_i - \beta_i l_i} - e^{\widehat{\alpha}_j - \beta_j \theta}$. Obviously, $g(\theta)$ is concave because $g''(\theta) < 0$, thus there exist at most two solutions to $g(\theta) = 0$.

Next, we will consider the number of intersections between $h_i(\theta)$ and $h_j(\theta)$ for any $i \neq j$ by considering following two cases based on whether or not an intersection exists between $e^{\widehat{\alpha}_i - \beta_i \theta}$ and $e^{\widehat{\alpha}_j - \beta_j \theta}$.

Case I: there is no intersection between $e^{\widehat{\alpha}_i - \beta_i \theta}$ and $e^{\widehat{\alpha}_j - \beta_j \theta}$. Without loss of generality, suppose that $e^{\widehat{\alpha}_i - \beta_i \theta} > e^{\widehat{\alpha}_j - \beta_j \theta}$ for any θ. That is, the curve $e^{\widehat{\alpha}_i - \beta_i \theta}$ lies above $e^{\widehat{\alpha}_j - \beta_j \theta}$ for all values of θ. Recall that each $h_i(\theta)$ is linear when $\theta \leq \underline{\theta}_i$ and $\theta \geq \overline{\theta}_i$, and a exponential function when $\theta \in [\underline{\theta}_i, \overline{\theta}_i]$. Thus, $h_i(\theta)$ may be passed once by $h_j(\theta)$ for $\theta \leq \underline{\theta}_i$, i.e., there exists at most one intersection between $h_i(\theta)$ and $h_j(\theta)$ within this range. In particular, the intersection could occur between the left linear piece of $h_i(\theta)$ and the left linear piece or the exponential piece of $h_j(\theta)$. Similarly, there exists at most one intersection between $h_i(\theta)$ and $h_j(\theta)$ for $\theta \geq \overline{\theta}_i$, and they could intersect at the right linear piece of $h_i(\theta)$ with the right linear piece or the exponential piece of $h_j(\theta)$. Therefore, there are at most two intersections in total between $h_i(\theta)$ and $h_j(\theta)$ for this case.

Case II: there exists an intersection between $e^{\widehat{\alpha}_i - \beta_i \theta}$ and $e^{\widehat{\alpha}_j - \beta_j \theta}$, denoted by $\tilde{\theta}$. Without loss of generality, suppose that $e^{\widehat{\alpha}_i - \beta_i \theta} > e^{\widehat{\alpha}_j - \beta_j \theta}$ for any $\theta < \tilde{\theta}$ and $e^{\widehat{\alpha}_i - \beta_i \theta} < e^{\widehat{\alpha}_j - \beta_j \theta}$ for any $\theta > \tilde{\theta}$. We consider the following three subcases based on the location of $\tilde{\theta}$.

<u>*Case II-a)*</u>: $\tilde{\theta} \leq \underline{\theta}_j$. ① If $\tilde{\theta} > \underline{\theta}_i$, there is no intersection between $h_i(\theta)$ and $h_j(\theta)$ for $\theta \leq \underline{\theta}_i$; and there are at most two intersections between them for $\theta > \underline{\theta}_i$. More precisely, the intersections could occur once between the exponential piece of $h_i(\theta)$ with the left linear piece of $h_j(\theta)$ and the other between the exponential piece or the right linear piece of $h_i(\theta)$ with the right linear piece of $h_j(\theta)$. Therefore, there are at most two intersections in total.

② If $\tilde{\theta} \leq \underline{\theta}_i$, There are at most two intersections between $h_i(\theta)$ and $h_j(\theta)$. More Specifically, if the value of $\underline{\theta}_i$ is still smaller than the (smaller) intersection between $e^{\widehat{\alpha}_i - \beta_i \theta}$ and $(l_j - c_j - \theta)e^{\alpha_j - \beta_j l_j}$, this case is similar to ①, i.e., no intersection between $h_i(\theta)$ and $h_j(\theta)$ for $\theta \leq \underline{\theta}_i$; and at most two intersections between them for $\theta > \underline{\theta}_i$. Otherwise, there exists respectively at most one intersection between $h_i(\theta)$ and $h_j(\theta)$ for $\theta \leq \underline{\theta}_i$ and $\theta > \underline{\theta}_i$. In particular, for $\theta \leq \underline{\theta}_i$, the intersection could occur between the left linear section of $h_i(\theta)$ with the left linear piece of $h_j(\theta)$. For $\theta > \underline{\theta}_i$, the intersection could occur between the exponential piece or the right linear piece of $h_i(\theta)$ with the right linear piece of $h_j(\theta)$.

Case II-b): $\tilde{\theta} \geq \overline{\theta}_j$, this is similar to *Case II-a* , so the detail is omitted.

Case II-c): $\underline{\theta}_j < \tilde{\theta} < \overline{\theta}_j$. ① If $\tilde{\theta} > \underline{\theta}_i$, there exists at most one intersection between $h_i(\theta)$ and $h_j(\theta)$ for $\theta \leq \underline{\theta}_i$, and this intersection could occur between the left linear section of $h_i(\theta)$ with the left linear piece or the exponential piece of $h_j(\theta)$. And, there exist at most two intersections between them for $\theta > \underline{\theta}_i$, one is the intersection $\tilde{\theta}$ and the other could occur between the right linear section of $h_j(\theta)$ with the right linear piece or the exponential piece of $h_i(\theta)$. Therefore, there are at most three intersections.

② If $\tilde{\theta} \leq \underline{\theta}_i$, there exist at most two intersections between $h_i(\theta)$ and $h_j(\theta)$ for $\theta \leq \underline{\theta}_i$. The intersections could occur twice between the left linear piece of $h_i(\theta)$ with the exponential piece of $h_j(\theta)$; or occur between the left linear piece of $h_i(\theta)$ with the exponential piece of $h_j(\theta)$ once and with the left linear piece of $h_j(\theta)$ once. For $\theta > \underline{\theta}_i$, there is at most one intersection between them, and the intersection could occur between the right linear section of $h_j(\theta)$ with the right linear piece or the exponential piece of $h_i(\theta)$. Therefore, there are at most three intersections.

We have thus shown that there are at most three intersections between functions $h_i(\theta)$ and $h_j(\theta)$. ☐

Proof of Theorem 2. Recall the results in Proposition 3, for any $\rho \in (\rho_l^-, \rho_{l-1}^+)$, the adjusted markup $\theta(\rho)$ (which is the solution to $r(\rho) = \min_\theta R_3(\rho, \theta)$) falls between two successive efficient intersections Θ_{l-1} and Θ_l, so the optimal offer set is S_C^{l-1} and thus $\max\limits_{|S|=C} R(S, \rho) = R(S_C^{l-1}, \rho)$. That is, it is best for the firm to offer S_C^{l-1} with one hundred percent probability, and the choice probability for S_C^{l-1} is equal to ρ for randomized assortment and pricing. The operational strategy is same to that for problem (5).

For any $\rho \in [\rho_l^+, \rho_l^-]$, we have the result that either S_C^l or S_C^{l-1} serves as the best offer set for problem (5) in this region in the proof of Theorem 1. In the proof of Proposition 4, we have shown that the combination of S_C^l and S_C^{l-1} generates more profit than either of them for ρ in this region. Naturally, we can offer a combination of these two sets to achieve profit maximization for problem (16). At the intersection point Θ_l, we have the choice probabilities ρ_l^+ and ρ_l^- associated with offer sets S_C^l and S_C^{l-1} respectively. Considering the constraints $\sum \pi_S = 1$ and $\sum \pi_S \rho_S = \rho$ in (16), we construct the equation system with $\rho_{S_C^{l-1}} = \rho_l^-$ and $\rho_{S_C^l} = \rho_l^+$: $\pi_{S_C^l} + \pi_{S_C^{l-1}} = 1$ and $\pi_{S_C^{l-1}} \rho_l^- + \pi_{S_C^l} \rho_l^+ = \rho$, leading to $\pi_{S_C^l} = \frac{\rho_l^- - \rho}{\rho_l^- - \rho_l^+}$ and $\pi_{S_C^{l-1}} = \frac{\rho - \rho_l^+}{\rho_l^- - \rho_l^+}$. Thus, the randomized optimal policy can be characterized by the efficient frontier specified in Proposition 4. ☐

Proof of Proposition 8. The proof is similar to that of Theorem 2, so we omit it here. ☐

Proof of Proposition 9. It is equivalent to show that for any $\Delta J(x, t-1)$, the maximum of $r(\rho) - \rho \Delta J(x, t-1)$ subject to $0 < \rho < 1$ is equal to the maximum of $\overline{r}(\rho) - \rho \Delta J(x, t-1)$.

Suppose that ρ^* is an optimal solution to problem $\max_{0<\rho<1} \{r(\rho) - \rho \Delta J(x, t-1)\}$. Note that any solution to problem $\max_{0<\rho<1} (r(\rho) - \Delta J(x, t-1)\rho)$ satisfies the first order condition: $r'(\rho) = \Delta J(x, t-1)$, noting that $r'(\rho)$ refers to its left-hand or right-hand derivative when $r(\rho)$ is not differentiable at $\rho = \rho_l$, $l = 1, 2, \ldots, L$. If there are multiple solutions to $r'(\rho) = \Delta J(x, t-1)$, denote $\rho^*(\Delta J(x, t-1))$ as the minimum one that maximizes $r(\rho)-\Delta J(x,t-1)\rho$. Suppose that ρ^o is an optimal solution to problem $\max_{0<\rho<1}\big\{\overline{r}(\rho)-\rho\Delta J(x,t-1)\big\}$ and (ρ'^o,ρ''^o,μ^o) corresponds to the definition of efficient frontier at point ρ^o. Then, we have

$$\begin{aligned}&r(\rho^*)-\rho^*\Delta J(x,t-1)\le\overline{r}(\rho^o)-\rho^o\Delta J(x,t-1)=\mu^o r(\rho'^o)+(1-\mu^o)r(\rho''^o)-\rho^o\Delta J(x,t-1)\\&\le\mu^o\Big(r(\rho'^o)-\rho'^o\Delta J(x,t-1)\Big)+(1-\mu^o)\Big(r(\rho''^o)-\rho''^o\cdot\Delta J(x,t-1)\Big).\end{aligned}$$

The second inequality holds because of the transformation (4). Recall that $r(\rho^*)-\rho^*\Delta J(x,t-1)\ge r(\rho'^o)-\rho'^o\Delta J(x,t-1)$ and $r(\rho^*)-\rho^*\Delta J(x,t-1)\ge r(\rho''^o)-\rho''^o\Delta J(x,t-1)$. Then, we obtain

$$r(\rho^*)-\rho^*\Delta J(x,t-1)\ge\mu^o\Big(r(\rho'^o)-\rho'^o\Delta J(x,t-1)\Big)+(1-\mu^o)\Big(r(\rho''^o)-\rho''^o\cdot\Delta J(x,t-1)\Big).$$

Therefore, we get

$$r(\rho^*)-\rho^*\Delta J(x,t-1)=r(\rho'^o)-\rho'^o\Delta J(x,t-1)=r(\rho''^o)-\rho''^o\cdot\Delta J(x,t-1)=\overline{r}(\rho^o)-\rho^o\Delta J(x,t-1).$$

Here ρ^* and ρ^o are optimal solutions to $\max_{0<\rho<1}\big\{r(\rho)-\rho\Delta J(x,t-1)\big\}$ and $\max_{0<\rho<1}\big\{\overline{r}(\rho)-\rho\Delta J(x,t-1)\big\}$, respectively. Thus, profit function $r(\rho)$ can be replaced by its corresponding efficient frontier in the Bellman function (18). Actually, we will show in the proof of Proposition 10 that both $r(\rho)-\rho\Delta J(x,t-1)$ and $\overline{r}(\rho)-\rho\Delta J(x,t-1)$ have the same optimizer. □

Proof of Proposition 10. Any solution to problem $\max_{0<\rho<1}\big(r(\rho)-\mu\rho\big)$ satisfies the first order condition: $r'(\rho)=\mu$, noting that $r'(\rho)$ refers to its left-hand or right-hand derivative when $r(\rho)$ is not differentiable at $\rho=\rho_l,\ l=1,2,\ldots,L$. If there are multiple solutions to $r'(\rho)=\mu$, denote $\rho^*(\mu)$ as the minimum one that maximizes $r(\rho)-\mu\rho$, i.e., $\rho^*(\mu):=\min\{\arg\max_{0<\rho<1}\{r(\rho)-\mu\rho\}\}$. We will next show that point $(\rho^*(\mu),r(\rho^*(\mu)))$ is on the efficient frontier $\overline{r}(\cdot)$.

Denote the tangent line of function $r(\rho)$ at point $(\rho^*(\mu),r(\rho^*(\mu)))$ as $\mathcal{L}:\ \big(\rho,r(\rho^*(\mu))+\mu(\rho-\rho^*(\mu))\big)$. If function $r(\rho)$ is below the tangent line $\mathcal{L}$, i.e., $r(\rho)\le r(\rho^*(\mu))+\mu(\rho-\rho^*(\mu))$ for any ρ, then $(\rho^*(\mu),r(\rho^*(\mu)))$ is on the frontier. We find that this inequality $r(\rho)\le r(\rho^*(\mu))+\mu(\rho-\rho^*(\mu))$ is equivalent to $r(\rho)-\mu\rho\le r(\rho^*(\mu))-\mu\rho^*(\mu)$, that is the definition of $\rho^*(\mu)$, thus the inequality is always true, which implies that function $r(\rho)$ is always below the tangent line $\mathcal{L}$ at $(\rho^*(\mu),r(\rho^*(\mu))$ for the whole interval of ρ.

Therefore, the optimal solution to problem $\max_{0<\rho<1}\{r(\rho)-\mu\rho\}$ for any $\mu\ge0$ must be on the efficient frontier of function $r(\rho)$. Moreover, $\big(\rho^*(\mu),r(\rho^*(\mu))\big)$ is an extreme point on the efficient frontier, i.e., equivalently $\rho^*(\mu)=\max_{0<\rho<1}\{\overline{r}(\rho)-\mu\rho\}$.

The optimal solution to problem $\max_{0<\rho<1}\{r(\rho)-\mu\rho\}$ (or equivalently $\max_{0<\rho<1}\{\overline{r}(\rho)-\mu\rho\}$) for any $\mu\ge0$ must be less than or equal to ρ^*. Actually, any $\rho\in(\rho^*,1)$ can not be an optimal solution for any $\mu\ge0$ because it consumes more resource but produces less profit. By Theorem 1 and Proposition 10, the only efficient sets are $S_C^{l^*},S_C^{l^*+1},\ldots,S_C^L$, where $l^*=\min\{l:\ \rho_{l+1}^-\le\rho^*=\rho^*(0)\}$. □

Proof of Theorem 3. Define $\tau_l = \Theta_l - \mathcal{H}(\Theta_l)$, since $\mathcal{H}(\cdot)$ (defined in (B-1)) is decreasing, it is obvious that $\tau_{l^*} < \tau_{l^*+1} < \ldots < \tau_L$. Based on (14) and the proof of Proposition 4, e.g., (B-6), the derivative of $\overline{r}(\rho)$ is given by:

$$\frac{\partial \overline{r}(\rho)}{\partial \rho} = \begin{cases} \theta(S_C^{l-1}, \rho) - \sum_{i \in S_C^{l-1}} \frac{e^{\widetilde{\alpha}_s - \beta_s \theta(S_C^{l-1}, \rho)}}{\beta_s}, & \rho_l^- \leq \rho \leq \rho_{l-1}^+, \\ \Theta_l - \mathcal{H}(\Theta_l), & \rho_l^+ < \rho < \rho_l^-, \\ 0, & \rho \geq \rho^*. \end{cases}$$

The optimization problem in (18) at each state (x,t) is equivalent to (19) or

$$\max_{0<\rho<1} \Big\{ \overline{r}(\rho) - \rho \Delta J(x, t-1) \Big\}, \tag{B-11}$$

which is clearly the problem $\max_{0<\rho<1}\{\overline{r}(\rho) - \mu\rho\}$ with $\mu = \Delta J(x, t-1)$. Because $\overline{r}(\rho)$ is strictly concave in ρ, for any $\Delta J(x, t-1) \in (\tau_{l-1}, \tau_l)$, the optimal solution to problem (B-11) satisfies the first-order condition $\partial \overline{r}(\rho)/\partial\rho = \Delta J(x, t-1)$. Moreover, since $\overline{r}(\rho)$ is concave in ρ, then $\Theta_{l-1} - \mathcal{H}(\Theta_{l-1}) \leq \theta(S_C^{l-1}, \rho) - \sum_{i \in S_C^{l-1}} \frac{e^{\widetilde{\alpha}_s - \beta_s \theta(S_C^{l-1}, \rho)}}{\beta_s} \leq \Theta_l - \mathcal{H}(\Theta_l)$. Because the definition of $\tau_{l-1} = \Theta_{l-1} - \mathcal{H}(\Theta_{l-1})$ and $\tau_l = \Theta_l - \mathcal{H}(\Theta_l)$, therefore, for any $\Delta J(x, t-1) \in (\tau_{l-1}, \tau_l)$, the optimal assortment is S_C^{l-1}; for $\Delta J(x, t-1) = \tau_l$, assortments S_C^{l-1} and S_C^l are both optimal. The associated optimal aggregate resource consumption rate and prices can be obtained by solving $\partial \overline{r}(\rho)/\partial\rho = \Delta J(x, t-1)$. Note that $\tau_{l^*} = \Theta_{l^*} - \mathcal{H}(\Theta_{l^*}) \leq 0$, and $\tau_{L+1} = \Theta_{L+1} - \mathcal{H}(\Theta_{L+1}) = +\infty$. Therefore, any $\Delta J(x, t-1) \in (\tau_l^*, \tau_{L+1})$ can be guaranteed.

In particular, the aggregate resource consumption rate corresponding to S_C^l and $\Delta J(x, t-1)$, denoted by $\rho(S_C^l, \Delta J(x, t-1))$, is given by $\rho(S_C^l, \Delta J(x, t-1)) = \frac{\sum_{s \in S_C^l} e^{\widetilde{\alpha}_s - \beta_s \theta(S_C^l, \Delta J(x, t-1))}}{1 + \sum_{s \in S_C^l} e^{\widetilde{\alpha}_s - \beta_s \theta(S_C^l, \Delta J(x, t-1))}}$, where $\theta(S_C^l, \Delta J(x, t-1))$ is the unique solution to $\Delta J(x, t-1) = \theta - \sum_{i \in S_C^l} \frac{e^{\widetilde{\alpha}_i - \beta_i \theta}}{\beta_i}$.

The optimal price corresponding to assortment S_C^l and ρ is $p_i(S_C^l, \rho) = 1/\beta_i + c_i + \theta(S_C^l, \rho)$ for any $i \in S_C^l$, where $\theta(S_C^l, \rho)$ is given by $\sum_{s \in S_C^l} e^{\widetilde{\alpha}_s - \beta_s \theta(S_C^l, \rho)} = \frac{\rho}{1-\rho}$. □

Proof of Theorem 4. Recall that $\tau_{l^*} \leq 0$, and $\tau_{L+1} = +\infty$. We can derive that $t^*(x, l^*) \leq 0$ and $t^*(x, L+1) = +\infty$ for any x. Therefore, any t belonging to the interval $(t^*(x, l^*), t^*(x, L+1))$ can be guaranteed. The proof is similar to Gallego and van Ryzin (1994) and Zhao and Zheng (2000), thus is omitted here. □

C. Nested Logit Model

The nested logit (NL) model is a popular generalization of the standard MNL model. Under the two-stage NL model, the customer makes the product selection decision sequentially: at the first stage, a branch is selected, referred to as a "nest" that includes multiple similar products; at the second stage, the customer chooses a product within the nest that is chosen at the first stage. Moreover, the *independence of irrelevant alternatives* (IIA) property no longer holds when the two alternatives do not belong to the same nest.

Suppose that the substitutable products constitute n nests, denoted by $\mathcal{N} := \{1, 2, \cdots, n\}$. There are m_i products that can be chosen to offer in nest i, denoted by $\mathcal{M}_i := \{1, 2, \cdots, m_i\}$. Denote the product set to offer by S_i and the corresponding price vector $\mathbf{p}_i$ for each $i \in \mathcal{N}$, where $S_i \subseteq \mathcal{M}_i$ and $\mathbf{p}_i := (p_{ij})_{j \in S_i}$. Some constraints can also be involved with product sets to offer and their corresponding prices under the NL model. In particular, the cardinality constraints are included in $\mathfrak{S}_i = \{S_i \subseteq \mathcal{M}_i : |S_i| \leq C_i\}$, where $|S_i|$ represents the cardinality of set S_i and C_i is the upper bound for the number of products to offer in nest i. The constraints on prices are expressed as : $p_{ij} \in \mathfrak{P}_{ij} := [l_{i,j}, u_{i,j}]$. For notational convenience, let $\mathbf{p}_i \in \mathfrak{P}_i$

represent $p_{ij} \in \mathfrak{P}_{ij}$ for each $j \in S_i$. The joint assortment and pricing under the NL model can be formulated as follows:

$$\begin{aligned} r(\rho) \overset{\text{def}}{=} \max_{S_i \in \mathfrak{S}_i, \mathbf{p}_i \in \mathfrak{P}_i, i \in \mathcal{N}} \quad & \textstyle\sum_{i \in \mathcal{N}} \sum_{j \in S_i} (p_{ij} - c_{ij}) d_{ij}\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) \\ s.t., \quad & \textstyle\sum_{i \in \mathcal{N}} \sum_{j \in S_i} d_{ij}\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) = \rho, \\ & d_{ij}\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) \geq 0, \ \forall i \in \mathcal{N}, j \in S_i, \end{aligned} \tag{C-1}$$

where $d_{ij}\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) = Q_i\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) \cdot q_{j|i}(S_i, \mathbf{p}_i)$ is the probability that the customer selects product j in nest i under the NL model, given the product sets and their prices $(S_i, \mathbf{p}_i)_{i \in \mathcal{N}}$. In particular, $Q_i\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) = \frac{\left(I_i(S_i, \mathbf{p}_i)\right)^{\gamma_i}}{1 + \sum_{k \in \mathcal{N}} \left(I_k(S_k, \mathbf{p}_k)\right)^{\gamma_k}}$ represents the choice probability that a customer selects nest i at the first stage; and $q_{j|i}(S_i, \mathbf{p}_i) = \frac{e^{\alpha_{ij} - \beta_i p_{ij}}}{\sum_{s \in S_i} e^{\alpha_{is} - \beta_i p_{is}}}$ denotes the probability that product j of nest i is selected at the second stage, given that she selected nest i at the first stage, by following the literature Greene (2007). Note that $\sum_{j \in S_i} q_{j|i}(S_i, \mathbf{p}_i) = 1$ and $\sum_{j \in S_i} d_{ij}\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big) = Q_i\big((S_i, \mathbf{p}_i)_{i \in \mathcal{N}}\big)$. Moreover, α_{ij} can be interpreted as the feature utility value of product j in nest i to the customer, and $I_i(S_i, \mathbf{p}_i) = \sum_{s \in S_i} e^{\alpha_{is} - \beta_i p_{is}}$ represents the attractiveness of nest i. The parameter γ_i can be viewed as the degree of heterogeneity among products in nest i and we assume that $0 < \gamma_i \leq 1$ to be compatible with the random utility maximization theory. In particular, if $\gamma_i = 1$, the two-stage NL degenerates to the standard MNL model; if $\gamma_i < 1$, the products within each nest i are more similar than the products across nests. Because we consider the products within the same nest are more similar (i.e., $0 < \gamma_i < 1$ for each $i \in \mathcal{N}$), assuming each nest products associated with the same level of price sensitivity is reasonable, i.e., $\beta_{ij} = \beta_i$ for any $j \in S_i$ (see, e.g., Li and Huh 2011).

Similar to the analysis under the MNL model, we solve the problem sequentially by first considering price optimization ($R(\mathbf{S}, \rho)$) then assortment optimization $r(\rho) = \max_{S_i \in \mathfrak{S}_i, i \in \mathcal{N}} R(\mathbf{S}, \rho)$. The following Corollary C-1 shows that the counterparts of Proposition A-1 hold under the NL model.

Corollary C-1. (Unconstrained Pricing under the NL Model) *Given assortment S_i for each nest $i \in \mathcal{N}$, we have the following:*

(a) At optimality, the product-level markup, defined as $p_{ij} - c_{ij} = \theta_i + 1/\beta_i$, is product-invariant for all offered products in the same nest i. Moreover, the adjusted nest-level markup, defined as $\phi = \theta_i + 1/\beta_i - 1/(\gamma_i \beta_i)$ is nest invariant for all nests.

(b) Any optimal assortment S_i^* for each nest i is fully capacitated.

(c) $R(\mathbf{S}, \rho)$ is concave with respect to the total choice probability ρ under the NL model.

Proof. The proof is similar to Gallego and Wang (2014), so it is omitted here for the sake of brevity. □

We focus on the joint optimization problem with cardinality constraint in the following analysis to derive new and deep managerial insights. By Corollary C-1, the average profit generated from each nest i, denoted by $R_i(S_i, \mathbf{p}_i) = \sum_{j \in S_i} (p_{ij} - c_{ij}) q_{j|i}(S_i, \mathbf{p}_i)$, is equal to $\theta_i + 1/\beta_i$, where θ_i is the constant markup for all products in S_i, $i \in \mathcal{N}$. The definition $\theta(S, \rho)$ in (7) can be extended and derived by

$$\rho = \sum_{i \in \mathcal{N}} Q_i(\boldsymbol{\theta}) \Longleftrightarrow \sum_{i \in \mathcal{N}} \Big(\sum_{s \in S_i} e^{\widehat{\alpha}_{is} - \beta_i \phi(\theta_i)} \Big)^{\gamma_i} = \frac{\rho}{1 - \rho}, \tag{C-2}$$

where $\widehat{\alpha}_{is} = \alpha_{is} - \beta_i c_{is} - 1/\gamma_i$, and $\phi(\theta_i) = \theta_i + 1/\beta_i - 1/(\gamma_i \beta_i)$. There exists a one-to-one relationship between $\phi(\theta_i)$ and ρ, and each θ_i is uniquely determined by a common ϕ (i.e., $\phi(\theta_i) = \phi$ for all i), which is crucial for simplifying the optimization problem. For any given ρ and assortments $\mathbf{S}$, the total expected profit can be expressed by

$$R(\mathbf{S}, \rho) = \rho\phi + (1-\rho) \sum_{i \in \mathcal{N}} \frac{\left(\sum_{s \in S_i} e^{\widehat{\alpha}_{is} - \beta_i \phi}\right)^{\gamma_i}}{\gamma_i \beta_i} = \rho\phi + (1-\rho) \sum_{i \in \mathcal{N}} \Big(\sum_{s \in S_i} e^{\widehat{\alpha}_{is} - \eta_i - \beta_i \phi} \Big)^{\gamma_i}, \tag{C-3}$$

where $\eta_i = 1/\gamma_i \log(\gamma_i \beta_i)$, and ϕ is $\phi(\mathbf{S}, \rho)$ with arguments omitted for brevity. Then $r(\rho)$ in (C-1) can be efficiently solved under cardinality constraints via introducing an auxiliary function $R_3(\rho, \phi)$, analogous to (11), for any $\rho \in (0,1)$ and a free variable $\phi > 0$,

$$R_3(\rho, \phi) := \rho\phi + (1-\rho)\mathcal{H}(\phi) = \rho\phi + (1-\rho) \sum_{i \in \mathcal{N}} \max_{|S_i| = C_i} \Big(\sum_{s \in S_i} h_{i,s}(\phi) \Big)^{\gamma_i} \text{ where } h_{i,s}(\phi) = e^{\hat{\alpha}_{is} - \eta_i - \beta_i \phi}. \tag{C-4}$$

Note that the strong duality presented in Proposition 2 continues to hold under the NL model for ρ varying in certain range, which can be demonstrated by the similar arguments in the proof of Lemma B-2. Then, we follow a similar idea under the capacitated MNL model to select up-to C_i top strictly decreasing and positive curves (e.g., $h_{i,s}(\phi)$) for each nest i as ϕ varies in a certain interval. More specifically, for each nest i, we suppose there are T_i efficient intersections among $\{h_{i,s}(\theta_i) : s \in \mathcal{M}_i\}$, and we rank the intersections from the smallest to the largest, i.e., $0 = \mathcal{A}(i,0) < \mathcal{A}(i,1) \leq \cdots \leq \mathcal{A}(i,T_i) < \mathcal{A}(i,T_i+1) = +\infty$. Thus, for any $\theta_i \in [\mathcal{A}(i,t_i), \mathcal{A}(i,t_i+1))$, $t_i \in \mathcal{T}_i := \{1, 2, \cdots T_i\}$, the candidate assortment is formulated as $S_i^{t_i} = \{\sigma_1^{t_i}, \cdots, \sigma_{C_i}^{t_i}\}$. Note that all θ_i's are linked to the common ϕ at optimality. Let $\mathcal{B}(i,t_i)$ be the maximum of ϕ with respect to $\theta_i \in [\mathcal{A}(i,t_i), \mathcal{A}(i,t_i+1)]$, i.e., $\mathcal{B}(i,t_i) := \max_{\theta_i \in [\mathcal{A}(i,t_i), \mathcal{A}(i,t_i+1)]} \phi(\theta_i) = \mathcal{A}(i,t_i+1) + 1/\beta_i - 1/(\gamma_i \beta_i)$.

Suppose the sequence $\{\mathcal{B}(i,t_i) : t_i \in \mathcal{T}_i\}$ is increasing in t_i for each nest i; otherwise, we can construct an increasing sequence based on $\{\mathcal{B}(i,t_i) : t_i \in \mathcal{T}_i\}$, which is monotonously increasing in k as follows: $\mathcal{B}\big(i, t_i(1)\big) < \mathcal{B}\big(i, t_i(2)\big) < \cdots < \mathcal{B}\big(i, t_i(k)\big) < \cdots$. Note that $t_i(1) = 1$. Then, denote the size of sequence $\{\mathcal{B}(i,t_i)_{t_i \in \mathcal{T}_i, i \in \mathcal{N}}\}$ by $\widehat{T}$ and let $\mathcal{T} := \{1, 2, \cdots, \widehat{T}\}$. Rank all the elements from the smallest to the largest, yielding the sequence: $\mathcal{B}\big(u(1), v(1)\big) \leq \mathcal{B}\big(u(2), v(2)\big) \leq \cdots \leq \mathcal{B}\big(u(\widehat{T}), v(\widehat{T})\big)$. We can then characterize the joint assortment and pricing problem under the capacitated NL model in Theorem C-1.

Theorem C-1. *For any* $\phi \in \big[\mathcal{B}(u(k-1), v(k-1)), \mathcal{B}(u(k), v(k))\big]$, $k \in \mathcal{T}$, *there exists the unique interval index, denoted by* $t_i(k)$ *for each nest* $i \in \mathcal{N}$ *such that* $\big[\mathcal{B}(u(k-1), v(k-1)), \mathcal{B}(u(k), v(k))\big] \subseteq \big[\mathcal{B}(i, t_i(k) - 1), \mathcal{B}(i, t_i(k))\big]$, *and the optimal assortment for each nest* i *is* $S_i^{t_i(k)}$, *i.e.,*

$$\mathcal{H}(\phi) = \sum_{i \in \mathcal{N}} \Big(\sum_{s \in S_i^{t_i(k)}} h_{i,s}(\phi) \Big)^{\gamma_i}, \quad \phi \in \big[\mathcal{B}(u(k-1), v(k-1)), \mathcal{B}(u(k), v(k))\big].$$

Proof. We follow a similar idea under the capacitated MNL model to select up-to-C_i top strictly decreasing and positive curves (e.g., $h_{i,s}(\phi)$) for each nest i as ϕ varies in a certain interval. Based on the above analysis, we have the sequence: $\mathcal{B}\big(u(1), v(1)\big) \leq \mathcal{B}\big(u(2), v(2)\big) \leq \cdots \leq \mathcal{B}\big(u(\widehat{T}), v(\widehat{T})\big)$. For any interval $[\mathcal{B}\big(u(k-1), v(k-1)\big), \mathcal{B}\big(u(k), v(k)\big)]$, $k \in \mathcal{T}$, we can find the unique interval index, denoted by $t_i(k)$ for each nest $i \in \mathcal{N}$ such that $[\mathcal{B}\big(u(k-1),v(k-1)\big),\mathcal{B}\big(u(k),v(k)\big)]\subseteq[\mathcal{B}\big(i,t_i(k)-1\big),\mathcal{B}\big(i,t_i(k)\big)]$, thus, the optimal assortment for each nest i is $S_i^{t_i(k)}$. Immediately, we can derive the following result on the constrained assortment optimization under the NL model: $\mathcal{H}(\phi)=\sum_{i\in\mathcal{N}}\left(\sum_{s\in S_i^{t_i(k)}}h_{i,s}(\phi)\right)^{\gamma_i}$, $\phi\in[\mathcal{B}\big(u(k-1),v(k-1)\big),\mathcal{B}\big(u(k),v(k)\big)]$. ☐

It is clear that the algorithm on how to identify a candidate assortment under the NL model is more complicated because multiple nests coexist. Despite this, it is important to note that the developed methodology (e.g., dimension and search space reduction) works efficiently for tackling optimization problems under more general consumer choice models. Moreover, C-1 provides a detailed analysis on the computational complexity of our method for the assortment problem under the constrained NL model.

Proposition C-1. *Let* $m=\max\{m_i:i\in\mathcal{N}\}$ *and* $C=\max\{C_i:i\in\mathcal{N}\}$. *It needs to run at most* $O\big(nC(m-C)\log(n)\big)$ *operations to obtain the sequence* $\mathcal{B}\big(u(1),v(1)\big)\le\mathcal{B}\big(u(2),v(2)\big)\le\cdots\le\mathcal{B}\big(u(\widehat{T}),v(\widehat{T})\big)$.

Proof. Let $m=\max\{m_i:i\in\mathcal{N}\}$ and $C=\max\{C_i:i\in\mathcal{N}\}$. The size of sequence of "$\mathcal{B}\big(i,t_i(1)\big)<\mathcal{B}\big(i,t_i(2)\big)<\cdots<\mathcal{B}\big(i,t_i(k)\big)<\cdots$" is at most $C(m-C)$. The complexity mainly comes from obtaining sequence "$\mathcal{B}\big(u(1),v(1)\big)\le\mathcal{B}\big(u(2),v(2)\big)\le\cdots\le\mathcal{B}\big(u(\widehat{T}),v(\widehat{T})\big)$", which is equal to merging n sorted arrays. The running time of merging n sorted arrays of size m is $O\big(nm\log(n)\big)$, which can be achieved by some well known algorithms, e.g., the multi-array merge algorithm. Immediately, the running time of merging sequence $\{\mathcal{B}(i,t_i)_{t_i\in\mathcal{T}_i,i\in\mathcal{N}}\}$ is $O\big(nC(m-C)\log(n)\big)$. ☐