Online Appendix

Cross-Category Retailing Management: Substitution and Complementarity

In this Online Appendix, we provide technical proofs for the paper titled "Cross-Category Retailing Management: Substitution and Complementarity."

Proof of Theorem 1. Let $\Pi(S,\mathbf{T},\mathbf{p},\mathbf{r})=z$. Then, the pricing problem (6) can be rewritten by

$$\max_{\mathbf{p},\mathbf{r}}\left\{z:\ z=\sum_{i\in S^+}\frac{\exp(u_i-p_i)}{1+\sum_{j\in S}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T_i}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{k\in T_i}\exp(v_{ik}-r_k)}\right)\right\}$$

$$=\max_{\mathbf{p},\mathbf{r}}\left\{z:\ z=\sum_{i\in S}\exp(u_i-p_i)\cdot\left((p_i-c_i-z)+\sum_{l\in T_i}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{k\in T_i}\exp(v_{ik}-r_k)}\right)+\sum_{l\in T_0}\frac{(r_l-s_l)\cdot\exp(v_{0l}-r_l)}{1+\sum_{k\in T_0}\exp(v_{0k}-r_k)}\right\}.$$

Note that the right-hand side (RHS) of the equation inside of the $\max(\cdot)$ function is strictly decreasing in z for any given $(\mathbf{p},\mathbf{r})$. Then, the above equation continues

$$z=\max_{\mathbf{p},\mathbf{r}}\left\{\sum_{i\in S}\exp(u_i-p_i)\cdot\left((p_i-c_i-z)+\sum_{l\in T_i}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{k\in T_i}\exp(v_{ik}-r_k)}\right)+\sum_{l\in T_0}\frac{(r_l-s_l)\cdot\exp(v_{0l}-r_l)}{1+\sum_{k\in T_0}\exp(v_{0k}-r_k)}\right\}$$

$$\Longleftrightarrow z=\max_{\mathbf{r}}\left\{\sum_{i\in S}\max_{p_i}\left\{\exp(u_i-p_i)\cdot\left((p_i-c_i-z)+\sum_{l\in T_i}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{k\in T_i}\exp(v_{ik}-r_k)}\right)\right\}\right.$$
$$\left.+\sum_{l\in T_0}\frac{(r_l-s_l)\cdot\exp(v_{0l}-r_l)}{1+\sum_{k\in T_0}\exp(v_{0k}-r_k)}\right\}.$$

Given $\mathbf{r}$ and z, the optimal solution to the optimization problem with respect to (w.r.t.) p_i is $p_i=c_i+z+1-\sum_{l\in T_i}(r_l-s_l)\cdot\exp(v_{il}-r_l)/\big(1+\sum_{k\in T_i}\exp(v_{ik}-r_k)\big)$. We have also derived a by-product result: $p_i-c_i+t_i=z+1$, where $t_i=\sum_{l\in T_i}(r_l-s_l)\cdot\exp(v_{il}-r_l)/\big(1+\sum_{k\in T_i}\exp(v_{ik}-r_k)\big)$. It suggests the aggregate markup $p_i-c_i+t_i$ is the same across all main products at optimality.

Then, the above chain of equations continues as follows: $z=H(z)$, where $H(z):=\max_{\mathbf{r}}H(\mathbf{r};z)$ and

$$H(\mathbf{r};z)=\sum_{i\in S}\exp\left((u_i-c_i-1-z)+\sum_{l\in T_i}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{k\in T_i}\exp(v_{ik}-r_k)}\right)+\sum_{l\in T_0}\frac{(r_l-s_l)\cdot\exp(v_{0l}-r_l)}{1+\sum_{k\in T_0}\exp(v_{0k}-r_k)}.$$

We observe that $H(z)$ is strictly decreasing in z. Denote z^* as the unique fixed point of $H(z)$, i.e., $z^*=H(z^*)$. Immediately, the aggregate profit is equal to z^*, and the optimal price of each main product i is

$$p_i^*=c_i+z^*+1-\sum_{l\in T_i}(r_l^*-s_l)\cdot\exp(v_{il}-r_l^*)/\big(1+\sum_{k\in T_i}\exp(v_{ik}-r_k^*)\big),$$

where $\mathbf{r}^*$ is the solution to the pricing problem for the secondary products $\max_{\mathbf{r}}H(\mathbf{r};z^*)$.

To derive the optimal prices $(\mathbf{p}^*,\mathbf{r}^*)$, we can combine with the bisection method and repeatedly solve the problem $\max_{\mathbf{r}}H(\mathbf{r};z)$ for different z. A sketch of the algorithm is provided as follows.

Step 0: Choose lower and upper bounds for z^*: $\underline{z}\le z^*\le\bar{z}$ and select the stopping criterion $\delta>0$.

Step 1: Solve the optimization problem $H(z'):=\max_{\mathbf{r}}H(\mathbf{r};z')$ regarding the middle point $z'=(\underline{z}+\bar{z})/2$.

Step 2: If $H(z')\ge z'$, set $\underline{z}=z'$; otherwise set $\bar{z}=z'$. If $|\bar{z}-\underline{z}|<\delta$, stop; otherwise, go to Step 1. □

Proof of Proposition 1. For $T_0 = \emptyset$, the function $H(\mathbf{r}; z)$ defined in the proof of Theorem 1 becomes

$$H(\mathbf{r}; z) = \exp(-z) \cdot \sum_{i \in S} \exp\left((u_i - c_i - 1) + \sum_{l \in T_i} \frac{(r_l - s_l) \cdot \exp(v_{il} - r_l)}{1 + \sum_{k \in T_i} \exp(v_{ik} - r_k)} \right).$$

Immediately, by Theorem 1 the optimal total expected profit z^* is the unique solution to

$$z \cdot \exp(z) = \max_{\mathbf{r}} \sum_{i \in S} \exp\left((u_i - c_i - 1) + \sum_{l \in T_i} \frac{(r_l - s_l) \cdot \exp(v_{il} - r_l)}{1 + \sum_{k \in T_i} \exp(v_{ik} - r_k)} \right).$$

One can observe that we need to solve the above optimization problem w.r.t. $\mathbf{r}$ only once. □

Proof of Proposition 2. By the proof of Theorem 1, the problem $\max_{\mathbf{p}} \Pi(S, \mathbf{T}, \mathbf{p}, \mathbf{r}^*)$ is equivalent to the problem $z = \sum_{i \in S} \exp(u_i - (1 + z + c_i - t_i^*)) + t_0^*$, where $t_i^* = \sum_{l \in T_i} (r_l^* - s_l) \cdot \exp(v_{il} - r_l^*) / \big(1 + \sum_{k \in T_i} \exp(v_{ik} - r_k^*)\big)$ for any $i \in S$. The unique solution w.r.t. z is denoted by z^*. Taking full derivatives for $z^* = \sum_{i \in S} \exp(u_i - (1 + z^* + c_i - t_i^*)) + t_0^*$ w.r.t. t_i^*, and rearranging terms yields the derivative of z^* w.r.t. t_i^* as follows: $\partial z^* / \partial t_i^* = \exp(u_i - (1 + z^* + c_i - t_i^*)) / \big(1 + \sum_{j \in S} \exp(u_j - (1 + z^* + c_j - t_j^*))\big)$. Obviously, $0 < \partial z^* / \partial t_i^* < 1$. Recall that the optimal price can be expressed as $p_i^* = 1 + z^* + c_i - t_i^*$. Then the optimal price p_i^* is decreasing in t_i^* although z^* is increasing, because $\partial p_i^* / \partial t_i^* = \partial(1 + z^* + c_i - t_i^*) / \partial t_i^* = \partial z^* / \partial t_i^* - 1 < 0$. Moreover, the optimal price p_i^* is the largest at $t_i^* = 0$, which indicates the optimal price could be lower when secondary products are provided. □

Proof of Theorem 2. With the *disjoint secondary sets*, the multi-secondary-product pricing problem can be decoupled into multiple independent pricing problems of smaller size corresponding to each main product. Let $\sum_{l \in T_i} (r_l - s_l) q_{l|i} = t_i$ for the secondary set T_i, and we have $r_l - s_l = 1 + t_i$ by using the seminal result in Li and Huh (2011) and Gallego and Wang (2014). Replace r_l by $1 + s_l + t_i$. After using some algebra and rearranging terms, problem (8) can be rewritten as follows: $t_i = \sum_{l \in T_i} \exp(v_{il} - 1 - s_l - t_i)$. Obviously, the above equation has a unique solution, denoted by $t_i^\dagger$, and immediately $r_l^\dagger = 1 + s_l + t_i^\dagger$.

Recall $p_i = 1 + z + c_i - t_i$ for any main product i at optimality by Theorem 1. Combining it with $r_l - s_l = 1 + t_i$ results in

$$p_i - c_i + r_l - s_l = 2 + z, \ \forall i \in S, \ l \in T_i, \tag{EC.1}$$

where z represents the firm's total expected profit from all main and secondary products. The equation (EC.1) shows the aggregate markup of any pair of main product and its associated secondary product is the same across all pairs of main and secondary products at optimality. □

Proof of Proposition 3. By the proof of Theorem 2, the optimal price for each secondary product $l \in T_i$ is equal to $1 + t_i^\dagger + s_l$. Considering the derivative of $r_l^\dagger$ w.r.t. β_{il} yields

$$\frac{\partial r_l^\dagger}{\partial \beta_{il}} = \frac{\partial(1 + t_i^\dagger + s_l)}{\partial \beta_{il}} = \frac{\partial t_i^\dagger}{\partial \beta_{il}} = \exp(\beta_{il} v_l - 1 - s_l - t_i^\dagger)\left(v_l - \frac{\partial t_i^\dagger}{\partial \beta_{il}}\right) - \frac{\partial t_i^\dagger}{\partial \beta_{il}} \sum_{k \in T_i, k \neq l} \exp(\beta_{ik} v_k - 1 - s_k - t_i^\dagger).$$

After using some algebra and rearranging terms, we have $\partial r_l^\dagger / \partial \beta_{il} = v_l \exp(\beta_{il} v_l - 1 - s_l - t_i^\dagger) / \big(1 + \sum_{k \in T_i} \exp(\beta_{ik} v_k - 1 - s_k - t_i^\dagger)\big) > 0$. Thus, $r_l^\dagger$ is increasing in β_{il}.

Similarly, for any other secondary product $k \in T_i$, $k \neq l$, the derivative of $r_k^\dagger$ w.r.t. β_{il} is $\partial r_k^\dagger / \partial \beta_{il} = \partial(1 + t_i^\dagger + s_k) / \partial \beta_{il} = \partial t_i^\dagger / \partial \beta_{il} > 0$. Thus, $r_k^\dagger$ is also increasing in β_{il} for $k \in T_i$ and $k \neq l$. Therefore, if β_{il} increases, both the optimal profit $t_i^\dagger$ and the prices for all secondary products in T_i increase. Moreover, $\partial r_l^\dagger / \partial \beta_{il} = \partial r_k^\dagger / \partial \beta_{il} = \partial t_i^\dagger / \partial \beta_{il}$. □

Proof of Proposition 4. By Theorem 1, we have the optimal price $p_i^\dagger = 1 + z^\dagger + c_i - t_i^\dagger$ for each main product $i \in S$ with disjoint secondary sets. Because a one-to-one relationship exists between $t_i^\dagger$ and β_{il}, we take the derivative of $p_i^\dagger$ w.r.t. β_{il} as follows by using the chain rules: $\partial p_i^\dagger/\partial \beta_{il} = \partial(1 + z^\dagger + c_i - t_i^\dagger)/\partial \beta_{il} = \partial(1 + z^\dagger + c_i - t_i^\dagger)/\partial t_i^\dagger \cdot \partial t_i^\dagger/\partial \beta_{il} = (\partial z^\dagger/\partial t_i^\dagger - 1) \cdot \partial t_i^\dagger/\partial \beta_{il}$. By Proposition 3, we have $\partial t_i^\dagger/\partial \beta_{il} > 0$, so we next consider the derivative of $z^\dagger$ w.r.t. $t_i^\dagger$: $\partial z^\dagger/\partial t_i^\dagger = \exp\left(u_i - (1 + z^\dagger + c_i - t_i^\dagger)\right)(1 - \partial z^\dagger/\partial t_i^\dagger) - \partial z^\dagger/\partial t_i^\dagger \cdot \sum_{j\in S, j\neq i} \exp\left(u_j - (1 + z^\dagger + c_j - t_j^\dagger)\right)$. After using some algebra and rearranging terms, we have $\partial z^\dagger/\partial t_i^\dagger = \exp\left(u_i - (1 + z^\dagger + c_i - t_i^\dagger)\right)/\left(1 + \sum_{j\in S} \exp\left(u_j - (1 + z^\dagger + c_j - t_j^\dagger)\right)\right) < 1$. Therefore, $p_i^\dagger$ is decreasing in β_{il}.

In a similar fashion, for any other main product $j \in S$, $i \neq j$, the derivative of $p_j^\dagger$ w.r.t β_{il} is $\partial p_j^\dagger/\partial \beta_{il} = \partial(1 + z^\dagger + c_j - t_j^\dagger)/\partial \beta_{il} = \partial(1 + z^\dagger + c_j - t_j^\dagger)/\partial t_i^\dagger \cdot \partial t_i^\dagger/\partial \beta_{il} = \partial z^\dagger/\partial t_i^\dagger \cdot \partial t_i^\dagger/\partial \beta_{il} > 0$. Therefore, $p_j^\dagger$ is increasing in β_{il}. Because all secondary sets are disjoint, $t_j^\dagger$ remains unchanged w.r.t. β_{il}, which is obvious by the expression of $t_j^\dagger$. Moreover, the total optimal profit $z^\dagger$ increases in β_{il} indirectly because $\partial p_j^\dagger/\partial \beta_{il} = \partial z^\dagger/\partial t_i^\dagger \cdot \partial t_i^\dagger/\partial \beta_{il} = \partial z^\dagger/\partial \beta_{il} > 0$. Thus, except for the optimal price for the main product i, the optimal prices for other main products increase in β_{il}.

The market share for main product i is denoted by d_i. Taking the derivative of d_i w.r.t. β_{il}, we have

$$\frac{\partial d_i}{\partial \beta_{il}} = \exp(u_i - p_i^\dagger) \cdot \frac{-\frac{\partial p_i^\dagger}{\partial \beta_{il}}\left(1 + \sum_{j\neq i, j\in S} \exp(u_j - p_j^\dagger)\right) + \frac{\partial p_j^\dagger}{\partial \beta_{il}} \sum_{j\neq i, j\in S} \exp(u_j - p_j^\dagger)}{\left(1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)\right)^2} > 0,$$

because $\partial p_i^\dagger/\partial \beta_{il} < 0$ and $\partial p_j^\dagger/\partial \beta_{il} > 0$. Therefore, the market share for product i increases in β_{il}.

Similarly, for any other main product $j \in S$, $j \neq i$, the derivative of d_j w.r.t. β_{il} is

$$\frac{\partial d_j}{\partial \beta_{il}} = \exp(u_j - p_j^\dagger) \cdot \frac{-\frac{\partial p_j^\dagger}{\partial \beta_{il}}\left(1 + \exp(u_i - p_i^\dagger)\right) + \frac{\partial p_i^\dagger}{\partial \beta_{il}} \exp(u_i - p_i^\dagger)}{\left(1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)\right)^2} < 0.$$

Therefore, market shares for other products in S decrease in β_{il}. Moreover, the choice probability of the no-purchase option 0 also decreases in β_{il}, because

$$\begin{aligned}
\frac{\partial d_0}{\partial \beta_{il}} &= \frac{\frac{\partial p_j^\dagger}{\partial \beta_{il}} \sum_{j\in S, j\neq i} \exp(u_j - p_j^\dagger) + \frac{\partial p_i^\dagger}{\partial \beta_{il}} \exp(u_i - p_i^\dagger)}{\left(1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)\right)^2} = \frac{\frac{\partial z^\dagger}{\partial t_i^\dagger} \cdot \frac{\partial t_i^\dagger}{\partial \beta_{il}} \sum_{j\in S} \exp(u_j - p_j^\dagger) - \frac{\partial t_i^\dagger}{\partial \beta_{il}} \exp(u_i - p_i^\dagger)}{\left(1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)\right)^2} \\
&= \frac{\frac{\partial t_i^\dagger}{\partial \beta_{il}} \cdot \left(\frac{\exp(u_i - p_i^\dagger)}{1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)} \sum_{j\in S} \exp(u_j - p_j^\dagger) - \exp(u_i - p_i^\dagger)\right)}{\left(1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)\right)^2} < 0.
\end{aligned}$$

The second equality holds because $\partial p_i^\dagger/\partial \beta_{il} = (\partial z^\dagger/\partial t_i^\dagger - 1) \cdot (\partial t_i^\dagger/\partial \beta_{il})$ and $\partial p_j^\dagger/\partial \beta_{il} = (\partial z^\dagger/\partial t_i^\dagger) \cdot (\partial t_i^\dagger/\partial \beta_{il})$. Thus, if β_{il} increases, the market share for main product i increases, whereas the market shares for any other main product in S and the no-purchase option decrease.

Recall that the market share for the secondary product $l \in T_i$ is $q_l = d_i q_{l|i}$ in this special case. Then considering the derivative of q_l w.r.t. β_{il} yields

$$\frac{\partial q_l}{\partial \beta_{il}} = \frac{\partial d_i}{\partial \beta_{il}} \cdot q_{l|i} + d_i \cdot \exp(\beta_{il} v_l - r_l^\dagger) \cdot \frac{v_l\left(1 + \sum_{k\neq l, k\in T_i} \exp(\beta_{ik} v_k - r_k^\dagger)\right) - \frac{v_l \exp(\beta_{il} v_l - r_l^\dagger)}{1 + \sum_{k\in T_i} \exp(\beta_{ik} v_k - r_k^\dagger)}}{\left(1 + \sum_{k\in T_i} \exp(\beta_{ik} v_k - r_k^\dagger)\right)^2} > 0.$$

Thus, the market share for secondary product $l \in T_i$ increases in β_{il}.

For any other secondary product $k \in T_i$, $k \neq l$, the derivative of q_k w.r.t. β_{il} is

$$\frac{\partial q_k}{\partial \beta_{il}} = \exp(u_i - p_i^\dagger) \cdot \frac{-\frac{\partial p_i^\dagger}{\partial \beta_{il}}\left(1 + \sum_{j\neq i, j\in S} \exp(u_j - p_j^\dagger)\right) + \frac{\partial p_j^\dagger}{\partial \beta_{il}} \sum_{j\neq i, j\in S} \exp(u_j - p_j^\dagger)}{\left(1 + \sum_{j\in S} \exp(u_j - p_j^\dagger)\right)^2} \cdot \frac{\exp(\beta_{ik} v_k - r_k^\dagger)}{1 + \sum_{l\in T_i} \exp(\beta_{il} v_l - r_l^\dagger)}$$

$$+\frac{\exp(u_i-p_i^\dagger)}{1+\sum_{j\in S}\exp(u_j-p_j^\dagger)}\cdot\exp(\beta_{ik}v_k-r_k^\dagger)\cdot\frac{-v_l\exp(\beta_{il}v_l-r_l^\dagger)-\frac{\partial r_k^\dagger}{\partial\beta_{il}}}{\left(1+\sum_{k\in T_i}\exp(\beta_{ik}v_k-r_k^\dagger)\right)^2}$$
$$=d_i\cdot q_{k|i}\cdot\left(\frac{-\frac{\partial p_i^\dagger}{\partial\beta_{il}}+\frac{\partial t_i^\dagger}{\partial\beta_{il}}\sum_{j\neq i,j\in S}\exp(u_j-p_j^\dagger)}{1+\sum_{j\in S}\exp(u_j-p_j^\dagger)}-\frac{v_l\exp(\beta_{il}v_l-r_l^\dagger)+\frac{\partial t_i^\dagger}{\partial\beta_{il}}}{1+\sum_{k\in T_i}\exp(\beta_{ik}v_k-r_k^\dagger)}\right)$$
$$=d_i\cdot q_{k|i}\cdot\frac{\partial t_i^\dagger}{\partial\beta_{il}}\cdot\left(\frac{1-d_i+\sum_{j\neq i,j\in S}\exp(u_j-p_j^\dagger)}{1+\sum_{j\in S}\exp(u_j-p_j^\dagger)}-\frac{1}{1+\sum_{k\in T_i}\exp(\beta_{ik}v_k-r_k^\dagger)}-1\right)<0.$$

The last equality holds because $\partial p_i^\dagger/\partial\beta_{il}=(\partial z^\dagger/\partial t_i^\dagger-1)\cdot(\partial t_i^\dagger/\partial\beta_{il})$ and $\partial t_i^\dagger/\partial\beta_{il}=v_l\exp(\beta_{il}v_l-r_l^\dagger)/\big(1+\sum_{k\in T_i}\exp(\beta_{ik}v_k-r_k^\dagger)\big)>0$. Thus, market shares for other secondary items in T_i decrease in β_{il}. □

Proof of Theorem 3. Let $(\mathbf{p}^*,\mathbf{r}^*)$ be the optimal prices for problem (9). For any $i\in S^+$ and $l\in T$, we have $w_l-\delta\le v_{il}\le w_l+\delta$. Then, for each $i\in S^+$ we have

$$\Pi_i(T,\mathbf{r}^*)=\sum_{l\in T}\frac{(r_l^*-s_l)\cdot\exp(v_{il}-r_l^*)}{1+\sum_{k\in T}\exp(v_{ik}-r_k^*)}\le\sum_{l\in T}\frac{(r_l^*-s_l)\cdot\exp(w_l+\delta-r_l^*)}{1+\sum_{k\in T}\exp(w_k-\delta-r_k^*)}$$
$$\le\exp(2\delta)\cdot\sum_{l\in T}\frac{(r_l^*-s_l)\cdot\exp(w_l-r_l^*)}{1+\sum_{k\in T}\exp(w_k-r_k^*)}\le\exp(2\delta)\cdot\sum_{l\in T}\frac{(r_l^o-s_l)\cdot\exp(w_l-r_l^o)}{1+\sum_{k\in T}\exp(w_k-r_k^o)}=\exp(2\delta)\cdot\Pi(T,\mathbf{r}^o).$$

The last inequality holds because $\mathbf{r}^o$ is the optimal solution to the problem (11). Then,

$$\Pi(S,T,\mathbf{p}^*,\mathbf{r}^*)=\sum_{i\in S^+}\frac{\exp(u_i-p_i^*)}{1+\sum_{j\in S}\exp(u_j-p_j^*)}\cdot\left(p_i^*-c_i+\Pi_i(T,\mathbf{r}^*)\right)$$
$$\le\sum_{i\in S^+}\frac{\exp(u_i-p_i^*)}{1+\sum_{j\in S}\exp(u_j-p_j^*)}\cdot\Big(p_i^*-c_i+\exp(2\delta)\cdot\Pi(T,\mathbf{r}^o)\Big)$$
$$\le\exp(2\delta)\cdot\sum_{i\in S^+}\frac{\exp(u_i-p_i^*)}{1+\sum_{j\in S}\exp(u_j-p_j^*)}\cdot\Big(p_i^*-c_i+\Pi(T,\mathbf{r}^o)\Big)$$
$$\le\exp(2\delta)\cdot\sum_{i\in S^+}\frac{\exp(u_i-p_i^o)}{1+\sum_{j\in S}\exp(u_j-p_j^o)}\cdot\Big(p_i^o-c_i+\Pi(T,\mathbf{r}^o)\Big)=\exp(2\delta)\cdot\Pi(S,T,\mathbf{p}^o,\mathbf{r}^o).$$

The last inequality holds because $\mathbf{p}^o$ is the optimal solution to the problem (12). Therefore, $\Pi(S,T,\mathbf{p}^o,\mathbf{r}^o)\ge\exp(-2\delta)\cdot\Pi(S,T,\mathbf{p}^*,\mathbf{r}^*)$. □

Proof of Corollary 1. (a) If consumers who did not buy any main product do not consider secondary products, i.e., $T_0=\emptyset$, problem (10) can be updated as follows:

$$\max_{\mathbf{p}}\Pi(S,T,\mathbf{p},\mathbf{r}):=\sum_{i\in S}\frac{\exp(u_i-p_i)}{1+\sum_{j\in S}\exp(u_j-p_j)}\cdot\big((p_i-c_i)+t^*\big),\tag{EC.2}$$

where $t^*=\sum_{l\in T}(r_l^*-s_l)\cdot\exp(w_l-r_l^*)/\big(1+\sum_{k\in T}\exp(w_k-r_k^*)\big)$ represents the optimal total profit of the identical set of secondary products, and it is the same across different main products because $v_{il}=w_l$ for any $i\in S$ and $l\in T$.

Let $\Pi(S,T,\mathbf{p},\mathbf{r})=z$, the problem (EC.2) is equivalent to the problem $z=\sum_{i\in S}\exp(u_i-(1+z+c_i-t^*))$ because $p_i=1+z+c_i-t^*$ for any main product $i\in S$ by Theorem 1, and the unique solution is denoted by z^*. Immediately, $p_i^*=1+z^*+c_i-t^*$. We first consider the derivative of the optimal total expect profit z^* w.r.t. the optimal profit from all secondary products t^*: $\partial z^*/\partial t^*=\partial\big\{\sum_{i\in S}\exp(u_i-(1+z^*+c_i-t^*))\big\}/\partial t^*=\big(1-\partial z^*/\partial t^*\big)\cdot\sum_{j\in S}\exp(u_j-(1+z^*+c_j-t^*))$. After rearranging terms, we obtain $\partial z^*/\partial t^*=\sum_{j\in S}\exp(u_j-(1+z^*+c_j-t^*))/\big(1+\sum_{j\in S}\exp(u_j-(1+z^*+c_j-t^*))\big)<1$. Then the optimal price p_i^* decreases in t^*

because $\partial p_i^*/\partial t^* = \partial(1+z^*+c_i-t^*)/\partial t^* = \partial z^*/\partial t^* - 1 < 0$, and it achieves its maximum when $t^*=0$, which indicates the main product prices with secondary products are lower.

(b) If consumers who did not buy any main product also consider secondary products, i.e., $T_0 = T$, problem (10) can be rewritten as follows under the assumption of $v_{il} = w_l$ for any $i \in S^+$ and $l \in T$:

$$\max_{\mathbf{p},\mathbf{r}} \Pi(S,T,\mathbf{p},\mathbf{r}) := \sum_{i\in S} \frac{(p_i - c_i)\cdot \exp(u_i - p_i)}{1+\sum_{j\in S}\exp(u_j - p_j)} + \sum_{l\in T} \frac{(r_l - s_l)\cdot \exp(w_l - r_l)}{1+\sum_{k\in T}\exp(w_k - r_k)}. \tag{EC.3}$$

Obviously, the above optimization problem (EC.3) can be decoupled into two pricing problems under the standard MNL model, and therefore the multi-main-product pricing problem can be solved independently of the secondary products. □

Proof of Proposition 5. We will construct an example to show that in the assortment planning problem, ignoring the complementarity effects may result in a solution (i.e., offer sets for the main and secondary products), whose profit is arbitrarily close to zero while the optimal profit is strictly positive.

Consider a simple example with one main product and two secondary products: $\exp(u_1 - p_1) = 1$ for the main product 1; $\exp(v_1 - r_1) = \exp(v_2 - r_2) = 1$ for these two secondary products. Set $r_2 - s_2 = (r_1 - s_1)/3 > 0$. Assume the firm ignores the complementarity effects, so it considers $\exp(v_1 - r_1)$ and $\exp(v_2 - r_2)$ for the secondary products 1 and 2 in consumers' choice process. It is straightforward to verify that the firm will offer the main product 1 and provide the secondary product 1, i.e., its offer sets are $S^o := \{1\}$ and $T^o := \{1\}$. The firm's perceived profit is $\Pi^o(S^o,T^o) = (p_1 - c_1)/2 + (r_1 - s_1)/4$.

However, there exist complementarity effects between the main product 1 and the secondary products, and the effects are measured by the coefficients β_{11} and β_{12}. Therefore, consumers in fact consider $\exp(\beta_{11}v_1 - r_1)$ and $\exp(\beta_{12}v_2 - r_2)$ for the secondary products 1 and 2 in their choice processes. We consider an extreme case $\beta_{11} \to -\infty$ while $\beta_{12} = 1$. Obviously, it is optimal to offer the main product 1 and provide the secondary product 2 as $\beta_{11} \to -\infty$, and the optimal offer sets are $S^* := \{1\}$ and $T^* := \{2\}$. The total expected profit of the offer sets (S^*, T^*) is $\Pi(S^*,T^*) = (p_1 - c_1)/2 + (r_2 - s_2)/4$, while the actual total expected profit of the offer sets (S^o, T^o) is $\Pi(S^o,T^o) = (p_1 - c_1)/2$, which is different from the firm's perceived profit $\Pi^o(S^o,T^o)$ discussed above. We compare the profits as follows:

$$\lim_{(p_1-c_1)\to 0} \frac{\Pi(S^o,T^o)}{\Pi(S^*,T^*)} = \lim_{(p_1-c_1)\to 0} \frac{(p_1-c_1)/2}{(p_1-c_1)/2 + (r_2-s_2)/4} = 0.$$

The limit holds because $(r_2 - s_2) > 0$. Till now, we have constructed an example that shows ignoring the complementarity effects may yield an arbitrarily bad solution. □

Proof of Theorem 4. Denote the optimal secondary sets by $\mathbf{T}^* := \left(T_i^*\right)_{i\in\mathcal{M}^+}$. For given price vectors $\mathbf{p}$ and $\mathbf{r}$, denote $z := \max_{S\subseteq\mathcal{M}} \Pi(S,\mathbf{T}^*,\mathbf{p},\mathbf{r})$. Rearranging terms and using some algebra yields

$$z = \max_{S\subseteq\mathcal{M}} \sum_{i\in S} (p_i - c_i + t_i^* - z)\exp(u_i - p_i) + t_0^*, \tag{EC.4}$$

where t_i^* (i.e., $t_i(T_i^*)$), $i \in \mathcal{M}^+$ represents the optimal profit collected from the secondary set T_i^*. Note that the RHS of the above equation is a series of strictly decreasing functions w.r.t. z while the LHS is strictly increasing in z, thus, there exists a unique solution, denoted by z^*. Recall the main products are labeled in the order of $p_1 - c_1 + t_1^* \geq p_2 - c_2 + t_2^* \geq \cdots \geq p_M - c_M + t_M^*$, so main product $i \in \mathcal{M}$ will be offered if and only if $p_i - c_i + t_i^* \geq z^*$. If we denote $p_{i^*} - c_{i^*} + t_{i^*}^* \geq z^* > p_{i^*+1} - c_{i^*+1} + t_{i^*+1}^*$, then the main product offer set $\{1, 2, \cdots, i^*\}$, which is an aggregate-markup-ordered assortment, is the optimal solution to the assortment problem (13). □

Proof of Corollary 2. After obtaining the optimal markup-ordered assortment T_i^* for each main product $i \in S^+$, we relabel main products such that they are in the decreasing order of aggregate markup (which is $p_i - c_i + t_i^*$ for $i \in S$). Immediately, an aggregate-markup-ordered assortment of main products is optimal for the assortment optimization problem (14) by Theorem 4. □

Proof of Proposition 6. (a) The assortment optimization problem for secondary products can be completely separated according to each main product $i \in S^+$. In particular, let $\sum_{l \in T_i}(r_l - s_l) q_{l|i}(T_i, \mathbf{r}) = t_i$ for main product i at the pre-determined price vector $\mathbf{r}$. Rearranging terms and using some algebra, the assortment problem $\max_{T_i \subseteq \mathcal{N}_i} \sum_{l \in T_i}(r_l - s_l) q_{l|i}(T_i, \mathbf{r})$ is equivalent to the problem below:

$$t_i = \max_{T_i \subseteq \mathcal{N}_i} \sum_{l \in T_i} h_l(t_i, \beta_{il}), \tag{EC.5}$$

where $h_l(t_i, \beta_{il}) = (r_l - s_l - t_i)\exp(\beta_{il} v_l - r_l)$ and it strictly decreases in t_i. Denote the optimal profit of the above problem by t_i^*, and the optimal secondary set for main product i is T_i^*, which is a markup-ordered assortment.

Now, we replace the complementarity coefficient β_{il} by β'_{il} for any $l \in T_i^*$, and $\beta_{il} > \beta'_{il}$. Then we have $h_l(t_i, \beta_{il}) > h_l(t_i, \beta'_{il})$ for any $t_i \geq 0$. Immediately, the unique profit solution to $t_i = \max_{T_i}\{\sum_{k \in T_i/\{l\}} h_k(t_i, \beta_{ik}) + h_l(t_i, \beta'_{il})\}$ is denoted by t_i^o, and $t_i^o < t_i^*$. The corresponding optimal secondary set is denoted by T_i^o. Since any secondary product k will be included in the optimal secondary set T_i^o if and only if $r_k - s_k \geq t_i^o$, we immediately have $T_i^* \subseteq T_i^o$.

Therefore, if β_{il} increases, the optimal profit t_i^* increases and the seller offers fewer (we use "fewer" in a weak sense in Section 4, i.e., fewer or equal) secondary products in the secondary set T_i^*. By the function for any $j \in S^+, j \neq i$ (including the no-purchase option in the first stage if necessary), $t_j = \max_{T_j} \sum_{l \in T_j}(r_l - s_l - t_j)\exp(\beta_{jl} v_l - r_l)$, we find that the change in β_{il} does not affect the best value of t_j and the assortment of the secondary set T_j. Therefore, the secondary sets for other main products remain unchanged as β_{il} increases.

(b) Following the proof of Theorem 4, let $g_i(t_i^*, z) = (p_i - c_i + t_i^* - z)\exp(u_i - p_i)$ for any $i \in S$, and it strictly increases in t_i^*. Recalling that $t_i^o < t_i^*$ and t_j^* is unchanged for any $j \in S^+, j \neq i$ by part (a), we have $g_i(t_i^o, z) < g_i(t_i^*, z)$ and $g_j(t_j^*, z)$ remains unchanged in t_i^* for any $z \geq 0$. Thus, the unique profit solution to $z = \max_{S \subseteq \mathcal{M}}\{\sum_{j \in S/\{i\}} g_j(t_j^*, z) + g_i(t_i^o, z)\} + t_0^*$ is z^o, and $z^o < z^*$. The corresponding optimal offer set of main products is S^o.

For main product i, it will be included in the offer set S^* if and only if $p_i - c_i \geq z^* - t_i^*$. As t_i^* increases due to larger β_{il}, the value of $z^* - t_i^*$ decreases because $\partial(z^* - t_i^*)/\partial t_i^* = (-1 - \sum_{j \neq i, j \in S}\exp(u_j - p_j))/(1 + \sum_{j \in S}\exp(u_j - p_j)) < 0$. Therefore, product i is still included in the offer set as t_i increases.

Any other main product j will be included if and only if $p_j - c_j \geq z^* - t_j^*$. Because z^* is larger but t_j^* remains unchanged as β_{il} increases, we derive $S^* \subseteq S^o$. Thus, if β_{il} is larger, the firm offers fewer main products at optimality, but main product i remains in the offer set, becoming more profitable to the firm and preferred by consumers. □

Proof of Proposition 7. The assortment problem with cardinality constraints for main and secondary products can be solved by linear programming. To have a concise paper and limit its length, the details of the technical proof are provided in a separate supplemental file, which is also available upon request. □

Proof of Proposition 8. To establish the NP-completeness for the Feasibility question regarding the assortment problem (15), we show that any Partition problem, which is a well-known NP-complete problem (see, e.g., Garey and Johnson 1979), can be transformed to an equivalent instance of Feasibility question for problem (15). We will show the NP-completeness via a reduction from an instance of the Partition problem: Partition Problem: given a sequence of positive rational numbers $\{a_1, a_2, \ldots, a_n\}$ with $\sum_{l=1}^n a_l = 2/3$, can we find a subset $T \subset \{1, 2, \ldots, n\}$ such that $\sum_{l\in T} a_l = 1/3 = \frac{1}{2} \cdot \sum_{l=1}^n a_l$?

For any sequence of positive integers, the condition $\sum_{l=1}^n a_l = 2/3$ can be satisfied by scaling. The Assortment Feasibility question is defined as follows:
Assortment Feasibility: given a profit threshold K, are there an offer set S and an identical secondary set T that generate a total profit of K or more in problem (15)?

We will construct a feasibility problem below. Suppose that there are two main products in consideration, i.e., $\mathcal{M} = \{1, 2\}$; there are $n+1$ products for possible secondary sets, i.e., $\mathcal{N} = \{1, 2, \ldots, n, n+1\}$. The parameters are: for main products $p_i - c_i = 300$ and $a_i = 1$ ($a_i = \exp(u_i - p_i)$) for $i = 1, 2$. For secondary products $r_l - s_l = 1$ and $a_{1,l} = a_{2,l} = a_l$ ($a_{i,l} = \exp(v_{il} - r_l)$) for $l = 1, 2, \ldots, n$; $r_{n+1} - s_{n+1} = 11/8$ and $a_{1,n+1} = 16/3$, $a_{2,n+1} = 2$. For the outside option, we have $a_0 = 1$ in consumers' choice for main products and $a_{i,0'} = 1$ for $i = 1, 2$ in consumers' choice for secondary products. Suppose that no secondary products are provided for the outside option in the first stage, i.e., $T_0 = \emptyset$. The threshold for the Assortment Feasibility question is $K = 200 + 83/120$.

We first claim that both main products should be offered; otherwise, the total profit cannot be equal to or greater than K. Second, we claim the secondary product $n+1$ should be included in the secondary set; otherwise, the total profit cannot be equal to or greater than the threshold, i.e.,

$$\Pi(\mathcal{M}, T_n) = \sum_{i=1}^{2} \frac{a_i}{a_0 + a_1 + a_2} \cdot \left(p_i - c_i + \sum_{l=1}^{n} \frac{(r_l - s_l) a_{i,l}}{a_{i,0'} + \sum_{l=1}^n a_{i,l}} \right) = 200 + 4/15 < K.$$

Therefore, product $n+1$ should be included in the secondary set. Denote the identical secondary set for both main products by $T \cup \{n+1\}$. Then, the total profit including the main and secondary products can be expressed as follows:

$$\begin{aligned} \Pi(\mathcal{M}, T \cup \{n+1\}) &= \frac{1}{3} \cdot \left(p_1 - c_1 + \frac{(r_{n+1} - s_{n+1}) a_{1,n+1} + \sum_{l\in T} (r_l - s_l) a_{1,l}}{a_{1,0'} + a_{1,n+1} + \sum_{l\in T} a_{1,l}} \right) \\ &+ \frac{1}{3} \cdot \left(p_2 - c_2 + \frac{(r_{n+1} - s_{n+1}) a_{2,n+1} + \sum_{l\in T} (r_l - s_l) a_{2,l}}{a_{2,0'} + a_{2,n+1} + \sum_{l\in T} a_{2,l}} \right) = 200 + \frac{1}{3} \cdot \left[\frac{22/3 + x}{19/3 + x} + \frac{11/4 + x}{3 + x} \right], \end{aligned}$$

where $x = \sum_{l\in T} a_{i,l}$, for $i = 1, 2$. Denote the terms inside the square bracket by $F(x)$ and consider its derivative as follows:

$$\frac{\partial F(x)}{\partial x} := \frac{\partial}{\partial x} \left[\frac{22/3 + x}{19/3 + x} + \frac{11/4 + x}{3 + x} \right] = \frac{(37/6 + 3x/2)(1/6 - x/2)}{(19/3 + x)^2 \cdot (3 + x)^2}.$$

Then $F(x)$ is increasing in x for $x \le 1/3$ and is decreasing thereafter, so $\max_{0 \le x \le 2/3} F(x) = F(1/3)$. Immediately, we have $\Pi(\mathcal{M}, T \cup \{n+1\}) \le 200 + \frac{1}{3} \cdot F(1/3) = 200 + 83/120$. The equality holds if and only if there exists a subset T such that $\sum_{l\in T} a_l = 1/3 = \frac{1}{2} \cdot \sum_{l=1}^n a_l$, which is exactly a solution to the Partition Problem. Thus, we have shown that the Feasibility question to the assortment problem (15) under the two-stage MNL model is NP-complete. □

Proof of Proposition 9. Denote the optimal offer sets of main and secondary products by S^* and T^*. Then, we have

$$\Pi(S^*,T^*) = \sum_{i\in S^*+} \frac{\exp(u_i-p_i)}{1+\sum_{j\in S^*}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T^*}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T^*}\exp(v_{il}-r_l)}\right).$$

Note that $p_0-c_0=0$ corresponds to the no-purchase option in the first stage. Consider the assortment problem of secondary products for the no-purchase option in the first stage:

$$\max_{T_0\subseteq\widetilde{\mathcal{N}}}\sum\nolimits_{l\in T_0}(r_l-s_l)\cdot q_{l|0}(T_0,\mathbf{r}). \tag{EC.6}$$

This is an assortment optimization problem under the standard MNL model, so by the seminal result in Talluri and van Ryzin (2004a) a margin-ordered assortment is optimal, denoted by $T_{n^\dagger}$.

We first show $\Pi(S^*,T^*)\le(1+\kappa)\cdot\Pi(S^*,T_{n^\dagger})$ as follows.

$$\begin{aligned}
\Pi(S^*,T^*) &= \sum_{i\in S^*+} \frac{\exp(u_i-p_i)}{1+\sum_{j\in S^*}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T^*}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T^*}\exp(v_{il}-r_l)}\right)\\
&\le \sum_{i\in S^*} \frac{\exp(u_i-p_i)}{1+\sum_{j\in S^*}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T^*}\frac{\kappa(p_i-c_i)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T^*}\exp(v_{il}-r_l)}\right)\\
&\quad+\frac{1}{1+\sum_{j\in S^*}\exp(u_j-p_j)}\cdot\left(\sum_{l\in T^*}\frac{(r_l-s_l)\cdot\exp(v_{0l}-r_l)}{1+\sum_{l\in T^*}\exp(v_{0l}-r_l)}\right)\\
&\le \sum_{i\in S^*} \frac{\exp(u_i-p_i)}{1+\sum_{j\in S^*}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T^*}\frac{\kappa(p_i-c_i)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T^*}\exp(v_{il}-r_l)}\right)\\
&\quad+\frac{1}{1+\sum_{j\in S^*}\exp(u_j-p_j)}\cdot\left(\sum_{l\in T_{n^\dagger}}\frac{(r_l-s_l)\cdot\exp(v_{0l}-r_l)}{1+\sum_{l\in T_{n^\dagger}}\exp(v_{0l}-r_l)}\right)\\
&\le (1+\kappa)\cdot\sum_{i\in S^*+} \frac{\exp(u_i-p_i)}{1+\sum_{j\in S}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T_{n^\dagger}}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T_{n^\dagger}}\exp(v_{il}-r_l)}\right)=(1+\kappa)\cdot\Pi(S^*,T_{n^\dagger}).
\end{aligned}$$

The first inequality holds because $(r_l-s_l)\le\kappa(p_i-c_i)$ for any $l\in\mathcal{N}$ and $i\in\mathcal{M}$; the second inequality holds because $T_{n^\dagger}$ is the optimal solution to the assortment problem (EC.6) for the no-purchase option in the first stage; the third inequality holds strictly if $\sum_{l\in T_{n^\dagger}}(r_l-s_l)\cdot\exp(v_{0l}-r_l)/\big(1+\sum_{l\in T_{n^\dagger}}\exp(v_{0l}-r_l)\big)\ge 0$.

By the proposed heuristic, for $n=n^\dagger$, i.e., the secondary set is $T_{n^\dagger}$, we derive the best offer set for main products, denoted by $S_{m^\dagger}$. Obviously, we have $\Pi(S_{m^\dagger},T_{n^\dagger})\ge\Pi(S^*,T_{n^\dagger})$. Therefore,

$$\Pi(S_{m^o},T_{n^o})\ge\Pi(S_{m^\dagger},T_{n^\dagger})\ge\Pi(S^*,T_{n^\dagger})\ge 1/(1+\kappa)\cdot\Pi(S^*,T^*).$$

Thus, the proposed heuristic yields a $1/(1+\kappa)-$optimal solution. □

Proof of Theorem 5. We will first show that for any given main product offer set S, a markup-ordered assortment is optimal for the secondary set. Assume by contrast that for a main product offer set S, a secondary product t' is included in the secondary set T but another secondary product t is not included, where $t<t'$. Note that $(r_t-s_t)\cdot\exp(v_{it}-r_t)\ge(r_{t'}-s_{t'})\cdot\exp(v_{it'}-r_{t'})$ and $\exp(v_{it}-r_t)\le\exp(v_{it'}-r_{t'})$. We will show that replacing the secondary product t' with t yields higher profit.

To show the improvement, we compare the total expected profits of (S,T) and $\big(S,T/\{t'\}\cup\{t\}\big)$ as follows. For any main product $i\in S^+$, we have

$$\Pi_i(T) = \frac{\sum_{l\in T}(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T}\exp(v_{il}-r_l)} = \frac{\sum_{l\in T/\{t'\}}(r_l-s_l)\cdot\exp(v_{il}-r_l)+(r_{t'}-s_{t'})\cdot\exp(v_{it'}-r_{t'})}{1+\sum_{l\in T/\{t'\}}\exp(v_{il}-r_l)+\exp(v_{it'}-r_{t'})}$$

$$\leq \frac{\sum_{l\in T/\{t'\}}(r_l-s_l)\cdot\exp(v_{il}-r_l)+(r_t-s_t)\cdot\exp(v_{it}-r_t)}{1+\sum_{l\in T/\{t'\}}\exp(v_{il}-r_l)+\exp(v_{it}-r_t)}$$
$$=\sum_{l\in T/\{t'\}\cup\{t\}}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T/\{t'\}\cup\{t\}}\exp(v_{il}-r_l)}.$$

The inequality holds because $(r_t-s_t)\cdot\exp(v_{it}-r_t)\geq(r_{t'}-s_{t'})\cdot\exp(v_{it'}-r_{t'})$ and $\exp(v_{it}-r_t)\leq\exp(v_{it'}-r_{t'})$.

Therefore, for any offer set S of main products, the following comparison holds:

$$\begin{aligned}\Pi(S,T)&=\left[\sum_{i\in S^+}\frac{\exp(u_i-p_i)}{1+\sum_{j\in S}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T}\exp(v_{il}-r_l)}\right)\right]\\&\leq\left[\sum_{i\in S^+}\frac{\exp(u_i-p_i)}{1+\sum_{j\in S}\exp(u_j-p_j)}\cdot\left((p_i-c_i)+\sum_{l\in T/\{t'\}\cup\{t\}}\frac{(r_l-s_l)\cdot\exp(v_{il}-r_l)}{1+\sum_{l\in T/\{t'\}\cup\{t\}}\exp(v_{il}-r_l)}\right)\right]\\&=\Pi\big(S,T/\{t'\}\cup\{t\}\big).\end{aligned}$$

In other words, for any given main product offer set S, a markup-ordered set of secondary products is optimal. Thus, we only need to consider the offer sets of secondary products: $T_1,T_2,\cdots,T_N$.

For any markup-ordered secondary set T_n, we consider the main product offering problem:

$$\max_{S\subseteq\mathcal{M}}\sum_{i\in S^+}\frac{\exp(u_i-p_i)}{1+\sum_{j\in S}\exp(u_j-p_j)}\cdot\big((p_i-c_i)+\Pi_i(T_n)\big),$$

where for notational convenience $\Pi_i(T_n)=\sum_{l\in T_n}(r_l-s_l)\cdot\exp(v_{il}-r_l)/\big(1+\sum_{l\in T_n}\exp(v_{il}-r_l)\big)$. The value $\big[(p_i-c_i)+\Pi_i(T_n)\big]$ can be viewed as the aggregate markup for the main product i, so one of the aggregate markup-ordered assortments is optimal. In particular, we only need to consider the assortments $S_m:=\{\sigma(1,n),\ldots,\sigma(m,n)\}$, where the main products are relabelled in the aggregate markup decreasing order such that $p_{\sigma(1,n)}-c_{\sigma(1,n)}+t_{\sigma(1,n)}(T_n)\geq\cdots\geq p_{\sigma(i,n)}-c_{\sigma(i,n)}+t_{\sigma(i,n)}(T_n)\geq\cdots\geq p_{\sigma(M,n)}-c_{\sigma(M,n)}+t_{\sigma(M,n)}(T_n)$. Thus, we conclude that the proposed heuristic returns an optimal solution. □

Proof of Proposition 10. By $q_{l|ki}(\boldsymbol{\vartheta},\boldsymbol{\beta}_i)=\exp\big(\beta_{il}(\boldsymbol{\vartheta}^T\cdot\mathbf{y}_l)-r_{kl}\big)/\sum_{j\in T_{ki}^+}\exp\big(\beta_{ij}(\boldsymbol{\vartheta}^T\cdot\mathbf{y}_j)-r_{kj}\big)$, the log-likelihood function $\mathcal{LL}_2(\boldsymbol{\vartheta},\boldsymbol{\beta})$ can be expressed as follows:

$$\begin{aligned}\mathcal{LL}_2(\boldsymbol{\vartheta},\boldsymbol{\beta})&=\sum_{k=1}^{K}\sum_{i\in S_k^+}\sum_{l\in T_{ki}^+}n_{kil}\cdot\log\big(q_{l|ki}(\boldsymbol{\vartheta},\boldsymbol{\beta}_i)\big)\\&=\sum_{k=1}^{K}\sum_{i\in S_k^+}\left[\Big(\sum_{l\in T_{ki}^+}n_{kil}\cdot\big(\beta_{il}(\boldsymbol{\vartheta}^T\cdot\mathbf{y}_l)-r_{kl}\big)\Big)-n_{ki}\log\Big(\sum_{j\in T_{ki}^+}\exp\big(\beta_{ij}(\boldsymbol{\vartheta}^T\cdot\mathbf{y}_j)-r_{kj}\big)\Big)\right].\end{aligned}$$

(a) First, we will show $\mathcal{LL}_2(\boldsymbol{\vartheta},\boldsymbol{\beta})$ is jointly concave in the vector $\boldsymbol{\vartheta}$ for any given $\boldsymbol{\beta}$. Its first-order partial derivative w.r.t. $\boldsymbol{\vartheta}^T$ can be derived as follows:

$$\frac{\partial\mathcal{LL}_2(\boldsymbol{\vartheta},\boldsymbol{\beta})}{\partial\boldsymbol{\vartheta}^T}=\sum_{k=1}^{K}\sum_{i\in S_k^+}\left[\sum_{l\in T_{ki}^+}\big(n_{kil}-n_{ki}\cdot q_{l|ki}\big)\cdot\beta_{il}\mathbf{y}_l\right],$$

where $q_{l|ki}=\exp\big(\beta_{il}(\boldsymbol{\vartheta}^T\cdot\mathbf{y}_l)-r_{kl}\big)/\sum_{j\in T_{ki}^+}\exp\big(\beta_{ij}(\boldsymbol{\vartheta}^T\cdot\mathbf{y}_j)-r_{kj}\big)$ for notational brevity. By considering the second-order partial derivative of $\mathcal{LL}_2(\boldsymbol{\vartheta},\boldsymbol{\beta})$ w.r.t. $\boldsymbol{\vartheta}$, we derive the Hessian matrix as follows:

$$\frac{\partial^2\mathcal{LL}_2(\boldsymbol{\vartheta},\boldsymbol{\beta})}{\partial\boldsymbol{\vartheta}^T\partial\boldsymbol{\vartheta}}=\sum_{k=1}^{K}\sum_{i\in S_k^+}-n_{ki}\cdot\left[\sum_{l\in T_{ki}^+}\frac{\partial q_{l|ki}}{\partial\boldsymbol{\vartheta}}\cdot\beta_{il}\mathbf{y}_l\right]=\sum_{k=1}^{K}\sum_{i\in S_k^+}-n_{ki}\cdot\left[\sum_{l\in T_{ki}^+}q_{l|ki}\cdot\Big((\beta_{il}\mathbf{y}_l)^T-\sum_{l\in T_{ki}^+}q_{l|ki}(\beta_{il}\mathbf{y}_l)^T\Big)\cdot\beta_{il}\mathbf{y}_l\right]$$

$$= \sum_{k=1}^{K} \sum_{i \in S_k^+} -n_{ki} \cdot \left[\sum_{l \in T_{ki}^+} q_{l|ki} (\beta_{il} \mathbf{y}_l)(\beta_{il} \mathbf{y}_l)^T - \sum_{l \in T_{ki}^+} q_{l|ki} (\beta_{il} \mathbf{y}_l) \cdot \sum_{l \in T_{ki}^+} q_{l|ki} (\beta_{il} \mathbf{y}_l)^T \right]$$
$$= -\sum_{k=1}^{K} \sum_{i \in S_k^+} n_{ki} \cdot \left[\sum_{l \in T_{ki}^+} q_{l|ki} (\beta_{il} \mathbf{y}_l - \overline{\mathbf{y}}_{ki})^T \cdot (\beta_{il} \mathbf{y}_l - \overline{\mathbf{y}}_{ki}) \right],$$

where $\overline{\mathbf{y}}_{ki} = \sum_{l \in T_{ki}^+} q_{l|ki} \cdot \beta_{il} \mathbf{y}_l$.

Because the moment matrix $(\beta_{il} \mathbf{y}_l - \overline{\mathbf{y}}_{ki})^T \cdot (\beta_{il} \mathbf{y}_l - \overline{\mathbf{y}}_{ki})$ is positive semidefinite, the Hessian matrix is negative semidefinite. Therefore, the log-likelihood function $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta})$ is jointly concave in the vector $\boldsymbol{\vartheta}$ for any given $\boldsymbol{\beta}$. We would like to point out that $\partial^2 \mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta}) / \partial \boldsymbol{\vartheta}^T \partial \boldsymbol{\vartheta}$ is negative definite under some very mild conditions, for example $K \geq |\boldsymbol{\beta}_i|$, where $|\boldsymbol{\beta}_i|$ represents the dimensions of $\boldsymbol{\beta}_i$. The readers of interest are referred to page 116 of McFadden (1974) for detailed discussion on the uniqueness of the maximizer of $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta})$.

(b) For any given $\boldsymbol{\vartheta}$, the log-likelihood function $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta})$ can be decoupled according to each main product $i \in \mathcal{M}^+$ to $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta}) = \sum_{i \in \mathcal{M}^+} \mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)$, where $\mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)$ is defined as follows: for any main product $i \in \mathcal{M}^+$,

$$\mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i) := \sum_{k=1}^{K} \sum_{l \in T_{ki}^+} n_{kil} \cdot \log \left(q_{l|ki}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i) \right),$$

where $\boldsymbol{\beta}_i := (\beta_{il})_{l \in \mathcal{N}^+}$. Recall that we assume $\beta_{0l} = 1$ for any $l \in \mathcal{N}^+$. For the concavity of $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta})$ w.r.t. $\boldsymbol{\beta}$ given any $\boldsymbol{\vartheta}$, it is sufficient to show the concavity of $\mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)$ w.r.t. $\boldsymbol{\beta}_i$.

Let $\mathbf{z}_l$ represent a vector: each element is zero except the l-th element is $\boldsymbol{\vartheta}^T \cdot \mathbf{y}_l$. Then, the choice probability $q_{l|ki}$ can be rewritten by $q_{l|ki} = \exp\left(\boldsymbol{\beta}_i^T \cdot \mathbf{z}_l - r_{kl} \right) / \sum_{j \in T_{ki}^+} \exp\left(\boldsymbol{\beta}_i^T \cdot \mathbf{z}_j - r_{kj} \right)$. Then, the first-order and second-order partial derivatives of $\mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)$ w.r.t. $\boldsymbol{\beta}_i$ can be derived as follows:

$$\frac{\partial \mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i^T} = \sum_{k=1}^{K} \left[\sum_{l \in T_{ki}^+} \left(n_{kil} - n_{ki} \cdot q_{l|ki} \right) \cdot \mathbf{z}_l \right],$$

$$\frac{\partial^2 \mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)}{\partial \boldsymbol{\beta}_i^T \partial \boldsymbol{\beta}_i} = -\sum_{k=1}^{K} n_{ki} \cdot \left[\sum_{l \in T_{ki}^+} \frac{\partial q_{l|ki}}{\partial \boldsymbol{\beta}_i} \cdot \mathbf{z}_l \right] = -\sum_{k=1}^{K} n_{ki} \cdot \left[\sum_{l \in T_{ki}^+} q_{l|ki} \cdot \left(\mathbf{z}_l^T - \sum_{l \in T_{ki}^+} q_{l|ki} \mathbf{z}_l^T \right) \cdot \mathbf{z}_l \right]$$
$$= -\sum_{k=1}^{K} n_{ki} \cdot \left[\sum_{l \in T_{ki}^+} q_{l|ki} \cdot \left(\mathbf{z}_l - \overline{\mathbf{z}}_{ki} \right)^T \cdot \left(\mathbf{z}_l - \overline{\mathbf{z}}_{ki} \right) \right],$$

where $\overline{\mathbf{z}}_{ki} = \sum_{l \in T_{ki}^+} q_{l|ki} \cdot \mathbf{z}_l$.

Axiom 5 in McFadden (1974) shows $\partial^2 \mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i) / \partial \boldsymbol{\beta}_i^T \partial \boldsymbol{\beta}_i$ is negative definite if the $\sum_{k=1}^{K} |S_k| \times (N+1)$ matrix whose rows are $[\mathbf{z}_l - \overline{\mathbf{z}}_{ki}]$ for any k and l is of rank $(N+1)$. Note that $\mathbf{z}_l$ is an $(N+1)$-dimensional vector with only one non-zero element. Moreover, we fix $\beta_{i0'} = 1$ for the no-purchase option in the second stage associated with each main product i. As discussed above, the log-likelihood function $\mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)$ is jointly concave in the vector $\boldsymbol{\beta}_i$. Immediately, $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta}) = \sum_{i \in \mathcal{M}^+} \mathcal{LL}_{2i}(\boldsymbol{\vartheta}, \boldsymbol{\beta}_i)$ is also jointly concave in the vector (matrix) $\boldsymbol{\beta}$ for any given $\boldsymbol{\vartheta}$.

(c) In the AO method, the log-likelihood function $\mathcal{LL}_2(\widehat{\boldsymbol{\vartheta}}^{(t)}, \widehat{\boldsymbol{\beta}}^{(t)})$ is monotonically increasing in each iteration l, i.e., $\mathcal{LL}_2(\widehat{\boldsymbol{\vartheta}}^{(t)}, \widehat{\boldsymbol{\beta}}^{(t)}) \geq \mathcal{LL}_2(\widehat{\boldsymbol{\vartheta}}^{(t-1)}, \widehat{\boldsymbol{\beta}}^{(t-1)})$. Moreover, $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta})$ is bounded above, e.g., $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta}) \leq 0$, so the AO method converges to a stationary point of $\mathcal{LL}_2(\boldsymbol{\vartheta}, \boldsymbol{\beta})$. The readers of interest are referred to Csiszar and Tusnady (1984) and Bezdek and Hathaway (2002) for detailed discussions on the convergence of the AO method. □