Online Appendix: Assortment Optimization for the Multinomial Logit Model with Repeated Customer Interactions

Chenhao Wang,[a,b] Ningyuan Chen,[c] Pin Gao,[d] Yao Wang,[e,*] Su Zhao[f]

[a] School of Economics and Management, Tongji University, Shanghai 200092, China

[b] Laboratory of High Quality Urban Development and Strategic Decision, Tongji University, Shanghai 200092, China

[c] Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada

[d] School of Data Science, School of Management and Economics, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China

[e] School of Management, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

[f] Meituan, Beijing, 100102, China

Appendix A: Design FPTAS for the Joint Assortment Planning Problem (12)

In this section, we design an FPTAS based on dynamic programming. The main idea is to find a set of assortments $\{S_1, S_{2,0}, S_{2,1}\}$ that approximate the optimal ones $\{S_1^*, S_{2,0}^*, S_{2,1}^*\}$ in terms of $V(S)$ and $W(S)$. Specifically, given positive constant $\epsilon \in (0,1)$, we aim at

$$V(S^*) \le V(S) \le (1+\epsilon)V(S^*),\ W(S) \ge (1-\epsilon)W(S^*), \tag{A.1}$$

for the three assortments. To achieve the objective, for each period $j \in \{1,2\}$, let $\gamma^j_{\max} = \max_{i\in\mathcal{N}_j} r_i v_i$, $\gamma^j_{\min} = \min_{i\in\mathcal{N}_j} r_i v_i$, $\alpha^j_{\max} = \max_{i\in\mathcal{N}_j} v_i$, and $\alpha^j_{\min} = \min_{i\in\mathcal{N}_j} v_i$. As a result, we can express the extreme values of the assortments: for example, $V(S_1) \in [\alpha^1_{\min}, |\mathcal{N}_1|\alpha^1_{\max}]$. We then discretize the intervals geometrically with the grid $\Delta^\epsilon_j = \{\gamma^j_{\min}\cdot(1+\epsilon)^l, l = 0,\dots,L^{\max}_j\}$ and $\Gamma^\epsilon_j = \{\alpha^j_{\min}\cdot(1+\epsilon)^l, l=0,\dots,G^{\max}_j\}$, where $L^{\max}_j = \lceil \log(|\mathcal{N}_j|\gamma^j_{\max}/\gamma^j_{\min})/\log(1+\epsilon)\rceil$ and $G^{\max}_j = \lceil \log(|\mathcal{N}_j|\alpha^j_{\max}/\alpha^j_{\min})/\log(1+\epsilon)\rceil$ for $j\in\{1,2\}$. The construction of the grid achieves two goals simultaneously. First, for the optimal assortment $\{S_1^*, S_{2,0}^*, S_{2,1}^*\}$, the nearby grid satisfies (A.1) because of the controlled step size. We will elaborate on how to find assortments whose $V(\cdot)$ and $W(\cdot)$ correspond to that grid point shortly. Second, searching the grid is relatively cheap, because of the geometric step size.

For a point on the grid $(g_j, h_j) \in \Gamma^\epsilon_j \times \Delta^\epsilon_j$, the goal is to find an assortment S solving the following problem:

$$\max_{S\subseteq\mathcal{N}_j} V(S) \quad s.t.\ V(S) \le g_j, W(S) \ge h_j. \tag{A.2}$$

To understand why such an optimization leads us to a solution to (A.1), observe that for the grid point (g^*, h^*) satisfying $g^* < V(S^*) < (1+\epsilon)g^*$ and $(1+\epsilon)^{-1}h^* < W(S^*) < h^*$, the optimal assortment in (A.1) S^* must be a feasible solution to (A.2) for the grid point $((1+\epsilon)g^*, (1+\epsilon)^{-1}h^*)$. Thus, the optimal solution S to Problem (A.2) satisfies $W(S) \ge (1+\epsilon)^{-1}h^* \ge (1-\epsilon)h^* \ge (1-\epsilon)W(S^*)$ and $V(S^*) \le V(S) \le (1+\epsilon)V(S^*)$, where in the last inequality we have used the optimality of S in (A.2) as $V(S^*) \le V(S)$. This guarantees that the optimal solution S to (A.2) for the grid point (g^*, h^*) satisfies (A.1).

We develop dynamic programming to solve (A.2) approximately. More precisely, let

$$\gamma_{j,i} = \left\lfloor \frac{|\mathcal{N}_j|}{\epsilon \cdot h_j} \cdot r_i v_i \right\rfloor,\ v_{j,i} = \left\lceil \frac{|\mathcal{N}_j|}{\epsilon \cdot g_j} \cdot v_i \right\rceil.$$

Given $I_1^j = \lfloor |\mathcal{N}_j|/\epsilon \rfloor - |\mathcal{N}_j|$ and $I_2^j = \lceil |\mathcal{N}_j|/\epsilon \rceil + |\mathcal{N}_j|$, we consider the optimization problem (A.3).

$$\max_{S_j \subseteq \mathcal{N}_j} V(S_j) \quad s.t. \sum_{i \in S_j} v_{j,i} \le I_2^j, \sum_{i \in S_j} \gamma_{j,i} \ge I_1^j. \tag{A.3}$$

The connection between (A.3) and (A.2) can be established: for any feasible solution S_j of Problem (A.3), we have $V(S_j) \le (1+2\epsilon) g_j$ and $W(S_j) \ge (1-2\epsilon) h_j$; Moreover, any feasible solution to (A.2) is feasible for (A.3). Therefore, solving (A.3) provides an approximately optimal solution to (A.2).

Problem (A.3) turns out to be more accessible to analysis, as a dynamic program can be developed: let $F^j(i_1^j, i_2^j, k)$ be the optimal value of the following optimization problem:

$$\max_{S_j \subseteq \{1,\dots,k\}} \sum_{i \in S_j} v_i \qquad s.t. \sum_{i \in S_j} \gamma_{j,i} \ge i_1^j, \text{ and } \sum_{i \in S_j} v_{j,i} \le i_2^j,$$

where $(i_1^j, i_2^j, k) \in [I_1^j] \times [I_2^j] \times [|\mathcal{N}_j|]$. By definition, the optimal solution to the state $(I_1^j, I_2^j, |\mathcal{N}_j|)$ in dynamic programming is the optimal solution to (A.3). The Bellman equation for the value function $F^j(i_1^j, i_2^j, k)$ is given by

$$F^j(i_1^j, i_2^j, k) = \max\{v_k + F^j(i_1^j - \gamma_{j,k}, i_2^j - v_{j,k}, k-1), F^j(i_1^j, i_2^j, k-1)\}. \tag{A.4}$$

For $k=1$, we have

$$F^j(i_1^j, i_2^j, 1) = \begin{cases} v_1 & \text{if } 0 \le i_1^j \le \gamma_{j,1}, i_2^j \ge v_{j,1}, \\ 0 & \text{if } i_1^j \le 0, i_2^j \ge 0, \\ -\infty & \text{otherwise.} \end{cases}$$

After solving the dynamic program, we collect the optimal solution to the state $(I_1^j, I_2^j, |\mathcal{N}_j|)$ for each pair of (g_j, h_j) and denote the collection by $\mathcal{S}_\epsilon^j$. We calculate the expected revenue for $(S_1, S_{2,1}, S_{2,0}) \in \mathcal{S}_\epsilon^1 \times \mathcal{S}_\epsilon^2 \times \mathcal{S}_\epsilon^2$ and return the combination with the highest revenue. The details are described in Algorithm 1. Formalizing the intuition provided above, we derive a performance guarantee for its expected revenue and running time in Proposition 2.

Appendix B: Numerical Performance of the Fixed Policy

In this section, we test the performance of the fixed policy in the disjoint product setting. When there are m periods and n products in each period, the computational complexity of finding the optimal adaptive policy by enumeration is $O(2^{n \cdot 2^m})$. For tractability, we consider a special setting. In the first $m-1$ stages, the seller only has one unique product with zero revenue. Then based on the interactions in the first $m-1$ stage, the seller offers a personalized assortment. In this case, only 2^{m-1} assortments need to be determined in the optimal adaptive policy, and the computational complexity reduces to $O(2^{n+m})$. In this setting, the seller learns the preference of customer through the repeated interactions in the first $m-1$ periods.

We use the following experimental setup: the preference weights of all products are sampled uniformly from the range $[0,5]$ and the revenue of products in period m is sampled uniformly from the range $[0,10]$. We consider $n \in \{2,5,7,9,11,13\}$ and $m \in \{2,3,4,5,6\}$. We calculate the adaptive ratio, which is the ratio of the

Algorithm 1 Approximation Algorithm for Problem (12).

Input: r_i, v_i for $i \in \mathcal{N}_1 \cup \mathcal{N}_2$, ϵ

Output: $S_1^f, S_{2,0}^f, S_{2,1}^f$

for $j \in \{1,2\}$ **do**

  for $(g_j, h_j) \in \Gamma_j^\epsilon \times \Delta_j^\epsilon$ **do**

    Calculate $F^j(i_1^j, i_2^j, k)$ for $i_1^j \le I_1^j, i_2^j \le I_2^j, k \le |\mathcal{N}_j|$ based on (A.4)

    Add the optimal solution S_j to the state $(I_1^j, I_2^j, |\mathcal{N}_j|)$ to $\mathcal{S}_\epsilon^j$

  end for

end for

$(S_1^f, S_{2,0}^f, S_{2,1}^f) = \arg\max_{(S_1, S_{2,0}, S_{2,1}) \in \mathcal{S}_\epsilon^1 \times \mathcal{S}_\epsilon^2 \times \mathcal{S}_\epsilon^2} \mathscr{R}_A(S_1, S_{2,0}, S_{2,1})$

expected revenue of the optimal adaptive policy over the optimal fixed policy. For each setting, we repeat 1000 trials and report the average and maximum adaptive ratio.

	$m=2$		$m=3$		$m=4$		$m=5$		$m=6$	
n	mean	max	mean	max	mean	max	mean	max	mean	max
2	1.0048	1.0950	1.0077	1.1256	1.0100	1.1274	1.0115	1.1353	1.0125	1.1421
5	1.0117	1.1211	1.0181	1.1440	1.0224	1.1572	1.0255	1.1670	1.0278	1.1782
7	1.0131	1.1129	1.0203	1.1503	1.0251	1.1526	1.0285	1.1596	1.0310	1.1703
9	1.0131	1.0641	1.0211	1.1009	1.0262	1.1196	1.0299	1.1281	1.0326	1.1384
11	1.0133	1.0722	1.0215	1.0937	1.0268	1.0961	1.0307	1.1145	1.0335	1.1266
13	1.0135	1.0766	1.0216	1.1023	1.0269	1.1154	1.0308	1.1227	1.0339	1.1290

Table EC.1 The adaptive ratio for the experiment in Appendix B

In Table EC.1, we summarize the performance of the fixed policy in different settings. The result shows that the adaptive ratio increases as the total number of stages m increases, in which the seller has more interactions with the customer. The results highlight the value of adaptive policy, the maximum revenue improvement could be 9.5% with two products and one past interaction. The average adaptive ratio also increases as more products are available. The maximal average adaptive ratio is around 1.034, which is achieved when $m = 6$ and $n = 13$.

Appendix C: Multiple Purchases

In this section, we extend our model to a scenario where a customer may purchase multiple products in a single period. We adopt the (single-period) multi-purchase choice model from Bai et al. (2024). In this framework, a customer will purchase at most M products, where M can potentially be a random variable. In our analysis, we assume M is deterministic and known for a customer for tractability. Because M is an unobserved latent variable, a random M can cause identifiability issues for estimation. Therefore, this simplifying assumption is also imposed in the numerical studies in Bai et al. (2024).

If the customer decides to purchase m products, they choose the m products with the highest utilities, provided they exceed the no-purchase option (U_0). The unconditional probability of selecting product $i \in S_1$ in the first period is thus:

$$\mathbb{P}_m(i, S_1) = \Pr(\{\text{rank}(i, S_1) \leq m\} \cap \{U_i > U_0\}),$$

where $\text{rank}(i, S_1)$ is the ranking of product i's utility in the assortment S_1. Given the customer's purchase decision $A \subseteq S_1$ with $|A| \leq m$, we define $\underline{U}_A = \min_{i \in A} U_i$ and $\overline{U}_B = \max_{i \in B} U_i$ for $B = S_1 \setminus A$. Clearly, we have $\underline{U}_A > \max(U_0, \overline{U}_B)$.

To optimize the assortment S in the second period, we must characterize the conditional choice probability given the purchase history in the first period. We derive a recursive formulation for the joint probability of this event

$$\mathbb{P}_m(i, S, A, B) = \Pr\left(\{\text{rank}(i, S) \leq m\} \cap \{U_i > U_0\} \cap \{\underline{U}_A > \max(U_0, \overline{U}_B)\}\right). \tag{C.1}$$

That is, the customer purchases a subset A of products in period one and moves on to purchase product i (in addition to other products) in the offered assortment S of period two.

Proposition C.1. *Let $T = S \cup S_1 \cup \{0\}$. The joint choice probability $\mathbb{P}_m(i, S, A, B)$ is given by the recursive formulation:*

$$\begin{aligned}\mathbb{P}_m(i, S, A, B) = &\frac{v_i}{V(T)} \cdot \mathcal{Q}(A \setminus \{i\}, B) \\ &+ \sum_{j \in (S \setminus B) \setminus \{i\}} \frac{v_j}{V(T)} \cdot \mathbb{P}_{m-1}(i, S \setminus \{j\}, A \setminus \{j\}, B) \\ &+ \sum_{j \in A \setminus S} \frac{v_j}{V(T)} \cdot \mathbb{P}_m(i, S, A \setminus \{j\}, B).\end{aligned}$$

Where $\mathcal{Q}(A', B)$ represents the probability that all items in A' are preferred to the items in B plus the outside option. This quantity can be calculated recursively as:

$$\mathcal{Q}(A', B) = \begin{cases} \sum_{k \in A'} \frac{v_k}{1 + V(A' \cup B)} \cdot \mathcal{Q}(A' \setminus \{k\}, B), & \textit{if } A' \neq \emptyset, \\ 1, & \textit{if } A' = \emptyset. \end{cases}$$

The recursive formulation in Proposition C.1 decomposes the joint event into a linear combination of simpler sub-events. Similar to the unconditional choice probability analyzed in Bai et al. (2024), the computational complexity of the joint probability depends on $|S|$ exponentially. In the special case where $m = 1$, it is straightforward to verify that the corresponding conditional choice probability reduces to the formulation established in Theorem 1.

While Proposition C.1 provides the theoretical foundation for calculating conditional probabilities, their complexity motivates a numerical study to extract practical insights. To this end, we design an experiment to quantify the value of conditioning on a customer's multi-purchase history when optimizing the conditional assortment.

C.1. Numerical Experiments

We conduct a numerical experiment to quantify the value of conditioning on the historical purchase in the multi-purchase setting. For the experiment, we assume (a) a customer would purchase at most $m=2$ items in a period; (2) a set A of $|A|=2$ products are purchased in period one, and (3) the available products of the two periods are disjoint ($\mathcal{N}_1 \cap \mathcal{N}_2 = \emptyset$). In addition, we assume the purchased products in A have the same preference weights (i.e., $v_i = V(A)/|A|$ for $i \in A$). Under this assumption, the pair $(V(A), V(B))$ serves as a sufficient statistic to record the historical event $\underline{U}_A > \max\{U_0, \overline{U}_B\}$ and thus assess the probability $\mathbb{P}_2(i, S, A, B)$. To evaluate the impact of the first-period outcome, we generate 400 scenarios of $(V(A), V(B))$ by sampling 20 values each for $V(A)$ and $V(B)$ from a uniform grid on $[10^{-3}, 2]$. For each of these 400 scenarios, we run 1000 replications. In each replication, we generate revenues $\boldsymbol{r}$ and weights $\boldsymbol{v}$ for the 10 available products in the second period (i.e., $|\mathcal{N}_2| = 10$). The revenues $\boldsymbol{r}$ are sampled uniformly from $[0, 10]$ and the weights $\boldsymbol{v}$ are sampled uniformly from $[0, 2]$. We then compare two strategies: (1) the benchmark Unconditional Optimal Policy (S^*_{uncond}), which ignores the history, and (2) our proposed Conditional Optimal Policy (S^*_{cond}), which maximizes the conditional expected revenue given (A, B). Our metric is the percentage revenue improvement of the conditional policy over the unconditional benchmark, averaged across the 1000 replications for each scenario.

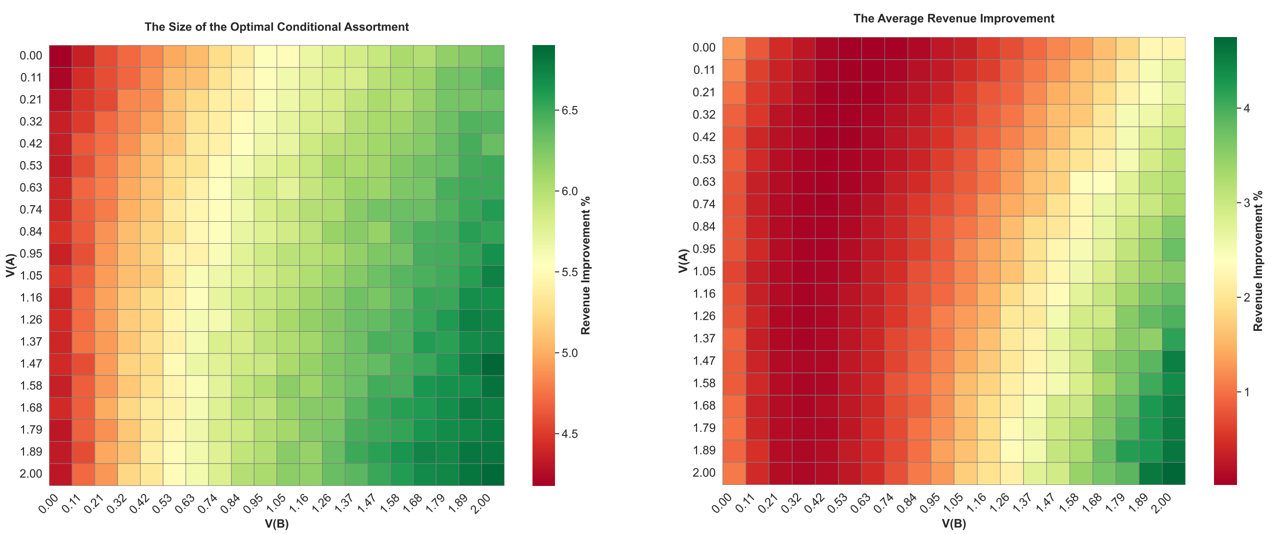
Our numerical results reveal a complex interaction effect between the purchased set value ($V(A)$) and the non-purchased set value ($V(B)$). As visualized in Figures EC.1 to EC.3, the value of the optimal conditional policy is driven not by any single factor, but by what has been inferred from the customer's first-period choice.

First, from Panel (a) of Figure EC.1, we observe that the optimal assortment size generally increases with the preference weights of the purchased set ($V(A)$) and the non-purchased set ($V(B)$). When $V(A)$ is small, we can infer that the utility of the no-purchase option U_0 is likely low, implying a high purchase probability in the second period. In this scenario (an "eager" buyer), the seller should limit the assortment size to mitigate the cannibalization effect, which is more prevalent in larger assortments. Conversely, a larger $V(B)$ implies a stronger competitive threshold $\overline{U}_B$. Conditional on the event $\underline{U}_A > \max\{U_0, \overline{U}_B\}$, a high $V(B)$ necessitates a higher realization of $\underline{U}_A$. This effectively lifts the upper bound for U_0, stochastically increasing its conditional distribution (i.e., making higher realizations of U_0 more probable). Consequently, the seller must offer a larger assortment to absorb this potential no-purchase demand in the second period.

Second, the mean revenue improvement exhibits distinct behavior, which are particularly evident in the right panel of Figure EC.3. When $V(B)$ is relatively high (approximately > 0.4), the mean revenue improvement increases as both $V(A)$ and $V(B)$ increase. This combination (high $V(A)$ and high $V(B)$) signals a "picky" customer, i.e., one with a high no-purchase utility (U_0). The unconditional policy, blind to this information, typically offers a standard assortment and suffers from a high rate of no-purchase outcomes. In contrast, when $V(B)$ is low, revenue improvement is negatively correlated with $V(A)$ and $V(B)$. A combination of low $V(A)$ and low $V(B)$ provides a clear, alternative signal: the customer purchased from a low-value set, identifying them as an "eager" buyer with a low U_0. Our analysis of the optimal assortment cardinality confirms that for this segment, the conditional policy prescribes a smaller assortment. Thus, S^*_{cond} creates value by pruning low-revenue items, mitigating the substitution effect and maximizing the probability that the eager customer selects a high-margin product.

Finally, the heatmap in Figure EC.2, which plots the *maximum observed revenue improvement* for each $(V(A),\ V(B))$ scenario, underscores the value of personalization. The maximum improvement can reach as high as 29%, which occurs when the customer is identified as picky. Furthermore, even when identifying the eager buyer (low $V(A)$ and low $V(B)$), the maximum improvement in specific instances remains significant, reaching as high as 15%.

In summary, our numerical experiments distinguish between two primary customer types determined jointly by the values of $V(A)$ and $V(B)$. We provide actionable insights for managers regarding when to expand or shrink the assortment size based on historical purchase behavior. These results confirm the substantial value of repeated interactions in the multi-purchase setting, demonstrating that the revenue gains from history-dependent optimization can be significant.



(a) Average size of S^*_{cond} (b) Revenue improvement (%)

Figure EC.1 Heatmaps illustrating the conditional policy's (a) average assortment size and (b) revenue improvement (%) as a function of the first-period purchased set value $V(A)$ and non-purchased set value $V(B)$.

Appendix D: Estimation Procedure for Section 7

In this section, we explain how to estimate the parameters of MNL and RCI-MNL models for the numerical experiments in Section 7. To simplify the notation, we consider customers having at most two interactions with the seller in the training data. The estimation procedure is similar when the customers have more interactions. We split the training data $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$, where $\mathcal{T}_1 = \{(S_t, i_t)\}_{t=1}^{|\mathcal{T}_1|}$ and $\mathcal{T}_2 = \{(S_t^1, i_t^1, S_t^2, i_t^2)\}_{t=1}^{|\mathcal{T}_2|}$. Specifically, $\mathcal{T}_1$ contains the data from the customers who have a single interaction with the seller and $\mathcal{T}_2$ contains the data from the customers who have two interactions with the seller. Under the MNL model, the likelihood of the events (S_t, i_t) and $(S_t^1, i_t^1, S_t^2, i_t^2)$ can be represented as follows:

$$\mathbb{P}^{\mathrm{MNL}}(S_t, i_t) = \frac{v_{i_t}}{1 + \sum_{i \in S_t} v_i},$$

$$\mathbb{P}^{\mathrm{MNL}}(S_t^1, i_t^1, S_t^2, i_t^2) = \frac{v_{i_t^1}}{1 + \sum_{i \in S_t^1} v_i} \cdot \frac{v_{i_t^2}}{1 + \sum_{i \in S_t^2} v_i}.$$

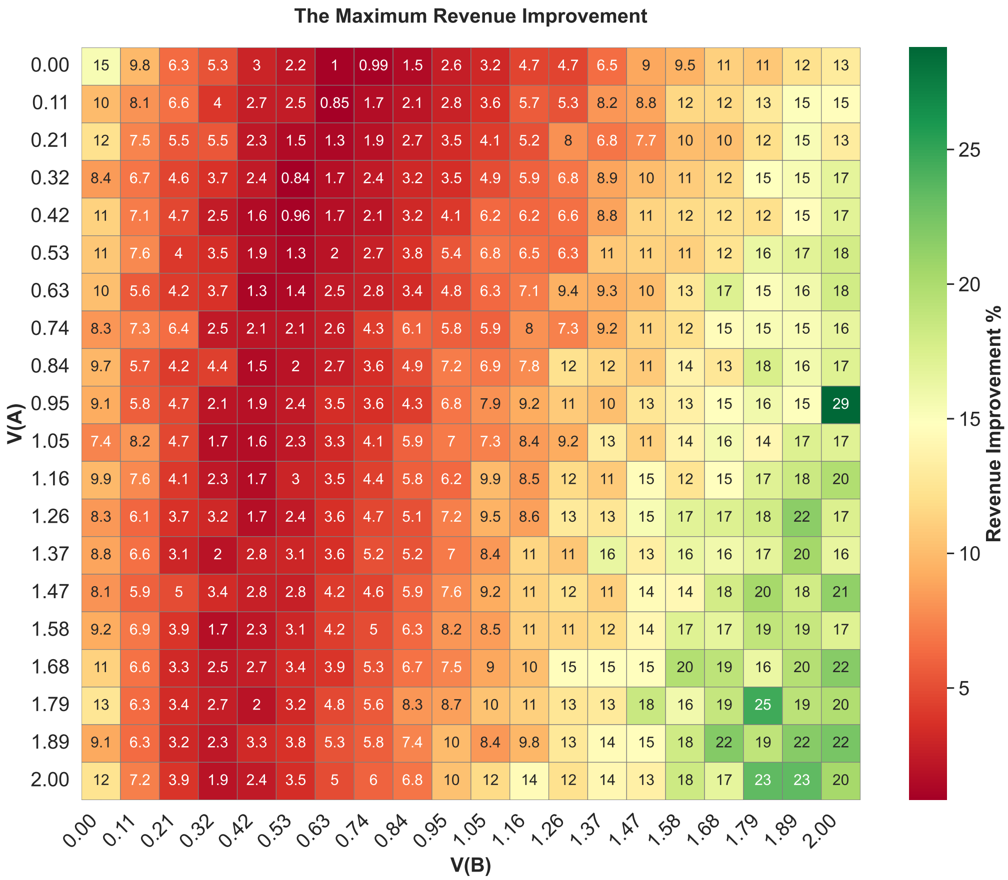


Figure EC.2 **The maximum revenue improvement of the conditional policy (S^*_{cond}) as a function of the first-period purchased set value $V(A)$ and non-purchased set value $V(B)$.**

For the same events, the probability expressions under the RCI-MNL model are

$$\mathbb{P}^{\text{RCI-MNL}}(S_t, i_t) = \frac{v_{i_t}}{1 + \sum_{i \in S_t} v_i},$$
$$\mathbb{P}^{\text{RCI-MNL}}(S^1_t, i^1_t, S^2_t, i^2_t) = \frac{v_{i^1_t}}{1 + \sum_{i \in S^1_t} v_i} \cdot \mathbb{P}^{\text{RCI-MNL}}(i^2_t | i^1_t, S^1_t, S^2_t),$$

where $\mathbb{P}^{\text{RCI-MNL}}(i^2_t | i^1_t, S^1_t, S^2_t)$ is the conditional probability of purchasing product i^2_t in the assortment S^2_t, which can be calculated based on Theorem 1. Note that we can convert v_i to e^{μ_i}. We then estimate the parameters in both models by maximizing the log-likelihood. That is, we solve the following optimization problems:

$$\max_{\boldsymbol{\mu}} \sum_{t=1}^{|\mathcal{T}_1|} \log \mathbb{P}^{\text{MNL}}(S_t, i_t) + \sum_{t=1}^{|\mathcal{T}_2|} \log \mathbb{P}^{\text{MNL}}(S^1_t, i^1_t, S^2_t, i^2_t),$$
$$\max_{\boldsymbol{\mu}} \sum_{t=1}^{|\mathcal{T}_1|} \log \mathbb{P}^{\text{RCI-MNL}}(S_t, i_t) + \sum_{t=1}^{|\mathcal{T}_2|} \log \mathbb{P}^{\text{RCI-MNL}}(S^1_t, i^1_t, S^2_t, i^2_t).$$

We solve both problems by the nonlinear solver *fmincon* in the MATLAB Optimization Toolbox.

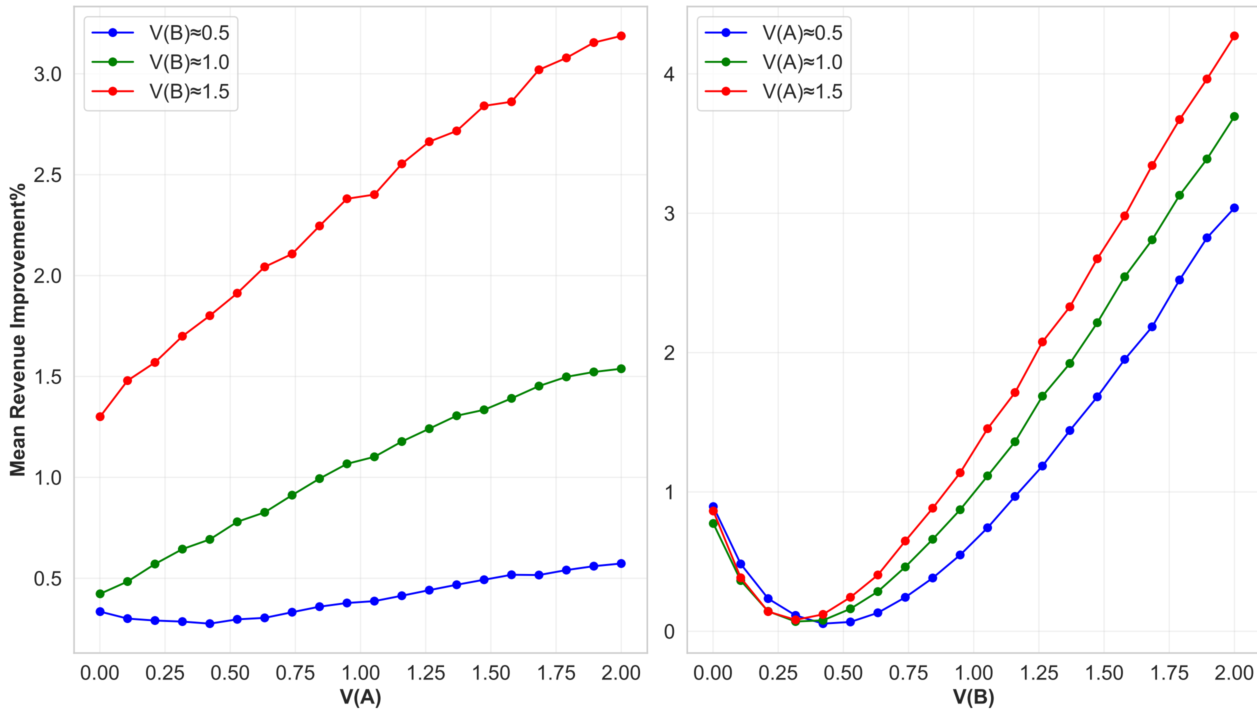


Figure EC.3 Mean revenue improvement (%) sliced by $V(A)$ (left) and $V(B)$ (right).

Appendix E: Hypothesis Testing for MNL and RCI-MNL Models

In this section, we provide a statistical framework to compare our proposed model with the MNL model. We formulate a hypothesis test to determine which model better explains the observed data, which consists of T customers, each observed over two choice periods. The two competing hypotheses are:

H_0: the customer behavior follows the MNL model, where the utilities are resampled.

H_1: the customer behavior follows the RCI-MNL model, where the utilities are not resampled.

As H_0 (independent utilities) and H_1 (invariant utilities) represent fundamentally different, non-nested data-generating processes, a standard likelihood-ratio test is not applicable. We therefore employ Vuong's Test Vuong (1989), which is specifically designed for comparing non-nested hypotheses.

First, we define the log-likelihood functions for both models. Let the observed data for customer i be $(S_{i,1}, c_{i,1})$ and $(S_{i,2}, c_{i,2})$, representing the assortments and choices in periods $t = 1, 2$. Let $\boldsymbol{\mu}$ be the vector of mean utility parameters (the logarithm of the preference weights). Under H_0, the log-likelihood ℓ_0 is the sum of all individual choice events:

$$\ell_0(\boldsymbol{\mu}) = \sum_{i=1}^{T} [\log \mathbb{P}(c_{i,1}|S_{i,1}; \boldsymbol{\mu}) + \log \mathbb{P}(c_{i,2}|S_{i,2}; \boldsymbol{\mu})]$$

where the probability for any single choice is given by the standard MNL formula $\mathbb{P}(c|S; \boldsymbol{\mu}) = \frac{\exp(\mu_c)}{1+\sum_{j\in S}\exp(\mu_j)}$. Under H_1, the choice in the second period is conditional on the outcome of the first. The log-likelihood ℓ_1 is based on the law of conditional probability:

$$\ell_1(\boldsymbol{\mu}) = \sum_{i=1}^{T} [\log \mathbb{P}(c_{i,1}|S_{i,1}; \boldsymbol{\mu}) + \log \mathbb{P}(c_{i,2}|S_{i,2}, c_{i,1}, S_{i,1}; \boldsymbol{\mu})] \,.$$

The conditional probability $\mathbb{P}(c_{i,2}|\cdot)$ depends on the history $(c_{i,1}, S_{i,1})$. Its form is derived in the main body and varies based on whether a purchase has been made in the first period.

To compare the models, we use the Maximum Likelihood Estimation (MLE) to obtain the maximal per-customer log-likelihood values, $\hat{\ell}_{0,i}$ and $\hat{\ell}_{1,i}$. We then compute the log-likelihood ratio for each customer i: $m_i = \hat{\ell}_{1,i} - \hat{\ell}_{0,i}$. The Vuong statistic V is constructed as the t-statistic for the mean of m_i, $V = \frac{\sqrt{T}\cdot\bar{m}}{s_m}$, where $\bar{m} = \frac{1}{T}\sum_{i=1}^{T} m_i$ is the sample mean of the ratios, and $s_m = \sqrt{\frac{1}{T-1}\sum_{i=1}^{T}(m_i - \bar{m})^2}$ is the sample standard deviation. Under the null hypothesis that the two models are indistinguishable, V asymptotically follows a standard normal distribution, $\mathcal{N}(0,1)$. Using a significance level of $p = 0.05$:

- If $V > 1.96$, we reject the null hypothesis. The test favors the H_1 (RCI-MNL) model as providing a significantly better fit;
- If $V < -1.96$, we reject the null hypothesis. The test favors the H_0 (standard MNL) model as providing a significantly better fit;
- If $|V| \le 1.96$, the test is inconclusive, and we cannot statistically distinguish between the two models' fit to the data.

E.1. Numerical Experiment

We conduct a simulation study to validate the statistical power of the Vuong test in distinguishing between H_0 (the standard MNL) and H_1 (RCI-MNL). The goal is to confirm that the test can reliably identify the correct data-generating process (DGP) as the sample size increases.

We consider 10 products and the preference weights are uniformly sampled from $[0,5]$. For each customer i, we generate two distinct assortments $S_{i,1}$ and $S_{i,2}$. We test a range of sample sizes $T \in \{100, 250, 500, 1000, 2000\}$. For each sample size T, we run $R = 1000$ independent simulation replications. We generate data under two ground-truth scenarios:

- H_0 (MNL): For each customer i, we generate two independent choices, following the MNL model;
- H_1 (RCI-MNL): For each customer i, we generate two dependent choices, following RCI-MNL choice model.

In each simulation, we fit both the H_0 and H_1 models to the generated dataset (regardless of the ground truth) and calculate the Vuong statistic V.

Figure EC.4 confirms the efficacy of the Vuong test in discriminating between the two models. The results demonstrate that the test reliably identifies the true data-generating process, with statistical power converging rapidly; notably, power exceeds 95% for both scenarios at a sample size of only $T = 500$. This validation establishes that the standard MNL and the proposed RCI-MNL are statistically distinguishable even with moderate data.

Appendix F: Proofs

We first show several technical lemmas that are useful in proving the main theorems.

LEMMA F.1. *For a Gumbel random variable $U \sim Gumbel(\mu, 1)$, the probability density function of U is $f_\mu(u) = \exp\left(-\left(u - \mu + e^{-(u-\mu)}\right)\right)$ and the cumulative distribution function $F_\mu(u) = \exp\left(-e^{-(u-\mu)}\right)$.*

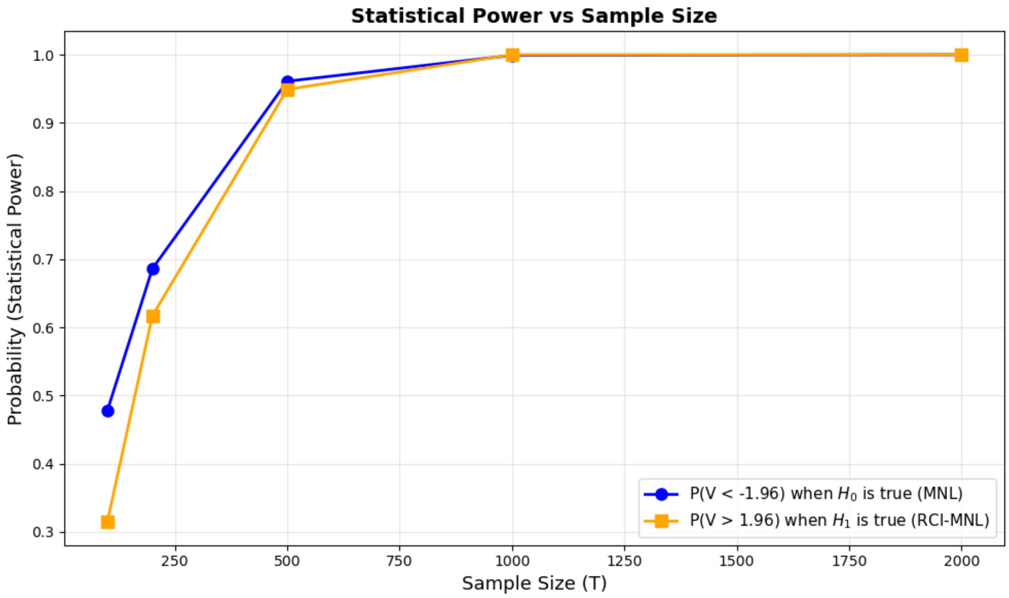


Figure EC.4 Statistical power of the Vuong test as a function of sample size (T).

LEMMA F.2. *Let $U_i \sim$ Gumbel$(\mu_i, 1), i = 1, \ldots, k$, $X \sim$ Gumbel$(\mu_x, 1)$ and assume they are all independent. We have:*

$$\mathbb{P}(U_1 \leq X, \ldots, U_k \leq X, a \leq X \leq b) = \frac{e^{-e^{-(b-\mu_x)}(\sum_{i=1}^k e^{\mu_i - \mu_x}+1)} - e^{-e^{-(a-\mu_x)}(\sum_{i=1}^k e^{\mu_i - \mu_x}+1)}}{\sum_{i=1}^k e^{\mu_i - \mu_x} + 1}.$$

Proof:

$$\begin{aligned}\mathbb{P}(U_1 \leq X, \ldots, U_k \leq X, a \leq X \leq b) &= \int_a^b \mathbb{P}(U_1 \leq X, \ldots, U_k \leq X | X = x) f_{\mu_x}(x) dx \\ &= \int_a^b \prod_{i=1}^k \mathbb{P}(U_i \leq x) f_{\mu_x}(x) dx \\ &= \int_a^b \prod_{i=1}^k e^{-e^{-(x-\mu_i)}} \cdot e^{-(x-\mu_x)} \cdot e^{-e^{-(x-\mu_x)}} dx.\end{aligned}$$

Let $t = e^{-(x-\mu_x)}$, then $dt = -e^{-(x-\mu_x)} dx$. Denote $a_x = e^{-(a-\mu_x)}$, $b_x = e^{-(b-\mu_x)}$, and $\nu_i = \mu_i - \mu_x$ $i = 1, \ldots, k$. We rewrite the probability as:

$$\int_{a_x}^{b_x} -e^{-t(\sum_{i=1}^k e^{\nu_i}+1)} dt = \frac{e^{-b_x(\sum_{i=1}^k e^{\nu_i}+1)} - e^{-a_x(\sum_{i=1}^k e^{\nu_i}+1)}}{\sum_{i=1}^k e^{\nu_i} + 1}.$$

After replacing the value of $\{\nu_i\}_{i=1}^k$ and $\{a_x, b_x\}$, we obtain the final result. $\square$

Based on Lemma F.2, we show Lemma F.3 as follows:

LEMMA F.3. *Let X and Y be independent random variables, the probability distributions are Gumbel$(\mu_x, 1)$, Gumbel$(\mu_y, 1)$, and d is a given constant, then we have :*

$$\mathbb{P}(X \geq \max\{Y, c\}) = \frac{e^{\mu_x}}{e^{\mu_x} + e^{\mu_y}} \cdot \left(1 - e^{-e^{-(c-\mu_x)}} \cdot e^{-e^{-(c-\mu_y)}}\right).$$

Proof:

$$
\begin{aligned}
\mathbb{P}(X \geq \max\{Y,c\}) &= \int_{-\infty}^{\infty} \mathbb{P}(X \geq \max\{y,c\}|Y=y) f_{\mu_y}(y)dy \\
&= \int_{-\infty}^{c} \mathbb{P}(X \geq c|Y=y) f_{\mu_y}(y)dy + \int_{c}^{\infty} \mathbb{P}(X \geq y|Y=y) f_{\mu_y}(y)dy \\
&= \int_{-\infty}^{c} \mathbb{P}(X \geq c) f_{\mu_y}(y)dy + \int_{c}^{\infty} \mathbb{P}(X \geq y) f_{\mu_y}(y)dy \\
&= \left(1-e^{-e^{-(c-\mu_x)}}\right)\cdot e^{-e^{-(c-\mu_y)}} + \mathbb{P}(Y\geq c) - \mathbb{P}(X<Y, Y\geq c) \\
&\overset{(a)}{=} \left(1-e^{-e^{-(c-\mu_x)}}\right)\cdot e^{-e^{-(c-\mu_y)}} + 1 - e^{-e^{-(c-\mu_y)}} - \frac{1-e^{-e^{-(c-\mu_y)}\cdot(1+e^{\mu_x-\mu_y})}}{1+e^{\mu_x-\mu_y}} \\
&= -e^{-e^{-(c-\mu_x)}}\cdot e^{-e^{-(c-\mu_y)}} + \frac{e^{-d\cdot e^{-(c-\mu_y)}}}{d} + 1 - \frac{1}{d} \quad (\text{where } d = 1+e^{\mu_x-\mu_y}) \\
&= (1-\frac{1}{d})\cdot\left(1-e^{-e^{-(c-\mu_x)}}\cdot e^{-e^{-(c-\mu_y)}}\right) \\
&= \frac{e^{\mu_x}}{e^{\mu_x}+e^{\mu_y}}\cdot\left(1-e^{-e^{-(c-\mu_x)}}\cdot e^{-e^{-(c-\mu_y)}}\right),
\end{aligned}
$$

where (a) follows from Lemma F.2. □

LEMMA F.4. *Let $U_i \sim$ Gumbel$(\mu_i,1)$, $i=1,\ldots,5$ and they are all independent. We have*

$$
\mathbb{P}(U_1 \geq \max\{U_2,U_3\}, U_4 \geq \max\{U_3,U_5\}) = \\
\frac{v_1v_4v_3}{(v_1+v_2)(v_4+v_5)}\left[\frac{1}{v_3} - \frac{1}{v_1+v_2+v_3} - \frac{1}{v_4+v_5+v_3} + \frac{1}{v_1+v_2+v_3+v_4+v_5}\right] \tag{F.1}
$$

Proof:

$$
\begin{aligned}
\mathbb{P}(U_1 \geq \max\{U_2,U_3\}, U_4 \geq \max\{U_3,U_5\}) &= \int_{-\infty}^{\infty} \mathbb{P}(U_1 \geq \max\{U_2,U_3\}, U_4 \geq \max\{U_3,U_5\}|U_3=c) f_{\mu_3}(c)dc \\
&= \int_{-\infty}^{\infty} \mathbb{P}(U_1 \geq \max\{U_2,c\})\cdot\mathbb{P}(U_4\geq\max\{c,U_5\}) f_{\mu_3}(c)dc
\end{aligned} \tag{F.2}
$$

Notice that from Lemma F.3, we have :

$$
\begin{aligned}
\mathbb{P}(U_1 \geq \max\{U_2,c\}) &= \frac{e^{\mu_1}}{e^{\mu_1}+e^{\mu_2}}\cdot\left(1-e^{-e^{-(c-\mu_1)}}\cdot e^{-e^{-(c-\mu_2)}}\right), \\
\mathbb{P}(U_4 \geq \max\{U_5,c\}) &= \frac{e^{\mu_4}}{e^{\mu_4}+e^{\mu_5}}\cdot\left(1-e^{-e^{-(c-\mu_4)}}\cdot e^{-e^{-(c-\mu_5)}}\right).
\end{aligned}
$$

Then (F.2) equals

$$
\begin{aligned}
&\frac{e^{\mu_1}e^{\mu_4}}{(e^{\mu_1}+e^{\mu_2})(e^{\mu_4}+e^{\mu_5})}\int_{-\infty}^{\infty}\left(1-e^{-e^{-(c-\mu_1)}}\cdot e^{-e^{-(c-\mu_2)}}\right)\cdot\left(1-e^{-e^{-(c-\mu_4)}}\cdot e^{-e^{-(c-\mu_5)}}\right)e^{-e^{-(c-\mu_3)}}\cdot e^{-(c-\mu_3)}dc \\
&= \frac{v_1v_4}{(v_1+v_2)(v_4+v_5)}\left[\int_{-\infty}^{\infty} e^{-e^{-(c-\mu_3)}}\cdot e^{-(c-\mu_3)}dc - \int_{-\infty}^{\infty}\prod_{i=1}^{3} e^{-e^{-(c-\mu_i)}}\cdot e^{-(c-\mu_3)}dc - \int_{-\infty}^{\infty}\prod_{i=3}^{5} e^{-e^{-(c-\mu_i)}}\cdot e^{-(c-\mu_3)}dc\right. \\
&\left.+\int_{-\infty}^{\infty}\prod_{i=1}^{5} e^{-e^{-(c-\mu_i)}} e^{-(c-\mu_3)}dc\right] \\
&= \frac{v_1v_4}{(v_1+v_2)(v_4+v_5)}\cdot\left[1-\frac{v_3}{v_1+v_2+v_3} - \frac{v_3}{v_4+v_5+v_3} + \frac{v_3}{v_1+v_2+v_4+v_5+v_3}\right] \\
&= \frac{v_1v_4v_3}{(v_1+v_2)(v_4+v_5)}\cdot\left[\frac{1}{v_3}-\frac{1}{v_1+v_2+v_3} - \frac{1}{v_4+v_5+v_3} + \frac{1}{v_1+v_2+v_4+v_5+v_3}\right].
\end{aligned}
$$

□

Lemma F.5. *Let $U_i \sim Gumbel(\mu_i, 1)$, $i=1,\dots,4$ and they are all independent. We have*

$$\mathbb{P}(U_1 \geq U_3 - d, U_3 \geq U_1, U_3 \geq U_4 | U_1 \geq U_2) = \frac{v_3}{v_1+v_2+v_3+v_4} - \frac{v_3}{e^d(v_1+v_2)+v_3+v_4}.$$

Proof: The probability has the following equivalent expression

$$\begin{aligned}
&\mathbb{P}(U_3 \geq U_1, U_3 \geq U_4 | U_1 \geq U_2) - \mathbb{P}(U_3 \geq U_1 + d, U_3 \geq U_4 | U_1 \geq U_2) \\
=&\mathbb{P}(U_3 \geq \max\{U_1,U_2\}, U_3 \geq U_4 | U_1 \geq U_2) - \mathbb{P}(U_3 \geq \max\{U_1,U_2\} + d, U_3 \geq U_4 | U_1 \geq U_2) \\
=&\mathbb{P}(U_3 \geq U_4 | U_3 \geq \max\{U_1,U_2\}, U_1 \geq U_2)\mathbb{P}(U_3 \geq \max\{U_1,U_2\} | U_1 \geq U_2) \\
&-\mathbb{P}(U_3 \geq U_4 | U_3 \geq \max\{U_1,U_2\} + d, U_1 \geq U_2)\mathbb{P}(U_3 \geq \max\{U_1,U_2\} + d | U_1 \geq U_2) \\
\overset{(a)}{=}&\mathbb{P}(U_3 \geq U_4 | U_3 \geq \max\{U_1,U_2\})\mathbb{P}(U_3 \geq \max\{U_1,U_2\}) - \mathbb{P}(U_3 \geq U_4 | U_3 \geq \max\{U_1,U_2\} + d)\mathbb{P}(U_3 \geq \max\{U_1,U_2\} + d) \\
=&\mathbb{P}(U_3 \geq \max\{U_1,U_2,U_4\}) - \mathbb{P}(U_3 \geq \max\{U_1+d,U_2+d,U_4\}) \\
=&\frac{e^{\mu_3}}{e^{\mu_1}+e^{\mu_2}+e^{\mu_4}+e^{\mu_3}} - \frac{e^{\mu_3}}{e^{\mu_1+d}+e^{\mu_2+d}+e^{\mu_4}+e^{\mu_3}} \\
=&\frac{v_3}{v_1+v_2+v_3+v_4} - \frac{v_3}{e^d(v_1+v_2)+v_3+v_4},
\end{aligned}$$

where (a) is from the property of the Gumbel random variable that the distribution of $\max\{U_1,U_2\}$ is independent of the event $U_1 \geq U_2$, the proof can be found in Appendix A in Gao et al. (2021). □

Lemma F.6. *Let $U_i \sim Gumbel(\mu_i, 1)$, $i=1,\dots,5$ and they are all independent. We have*

$$\begin{aligned}
&\mathbb{P}(U_4 \geq \max\{U_1,U_2,U_5-d\}, U_5 \geq \max\{U_1,U_3\}) = \\
&\frac{v_5}{v_3+v_5}\cdot\frac{v_4}{v_1+v_2+v_4+e^{-d}(v_3+v_5)} - \frac{v_1}{(v_1+v_3+v_5)(v_3+v_5)}\cdot\frac{v_4v_5}{v_1+v_2+v_3+v_4+v_5}
\end{aligned}$$

Proof:

$$\begin{aligned}
&\mathbb{P}(U_4 \geq \max\{U_1,U_2,U_5-d\}, U_5 \geq \max\{U_1,U_3\}) \\
=&\int\int \mathbb{P}(x \geq U_1, x \geq U_2, x \geq y-d, y \geq U_1, y \geq U_3) f_{\mu_4}(x) f_{\mu_5}(y) dy dx \\
=&\int\int_{-\infty}^{x+d} \mathbb{P}(\min\{x,y\} \geq U_1)\mathbb{P}(x \geq U_2)\mathbb{P}(y \geq U_3) f_{\mu_4}(x) f_{\mu_5}(y) dy dx \\
=&\int\int_{-\infty}^{x} \mathbb{P}(y \geq U_1)\mathbb{P}(x \geq U_2)\mathbb{P}(y \geq U_3) f_{\mu_4}(x) f_{\mu_5}(y) dy dx \\
+&\int\int_{x}^{x+d} \mathbb{P}(x \geq U_1)\mathbb{P}(x \geq U_2)\mathbb{P}(y \geq U_3) f_{\mu_4}(x) f_{\mu_5}(y) dy dx \\
\overset{(a)}{=}&\frac{e^{\mu_5}}{e^{\mu_5}+e^{\mu_1}+e^{\mu_3}}\cdot\int e^{-e^{-(x-\mu_5)}(1+e^{\mu_1-\mu_5}+e^{\mu_3-\mu_5})}\mathbb{P}(x \geq U_2) f_{\mu_4}(x) dx \\
+&\frac{e^{\mu_5}}{e^{\mu_5}+e^{\mu_3}}\int \mathbb{P}(x \geq U_1)\mathbb{P}(x \geq U_2)\left(e^{-e^{-(x+d-\mu_5)}(1+e^{\mu_3-\mu_5})} - e^{-e^{-(x-\mu_5)}(1+e^{\mu_3-\mu_5})}\right) f_{\mu_4}(x) dx \\
=&\frac{e^{\mu_5}}{e^{\mu_5}+e^{\mu_1}+e^{\mu_3}}\cdot\frac{e^{\mu_4}}{e^{\mu_2}+e^{\mu_5}+e^{\mu_1}+e^{\mu_3}+e^{\mu_4}} + \frac{e^{\mu_5}}{e^{\mu_5}+e^{\mu_3}}\left(\frac{e^{\mu_4}}{e^{\mu_2}+e^{\mu_5-d}+e^{\mu_1}+e^{\mu_3-d}+e^{\mu_4}} - \frac{e^{\mu_4}}{e^{\mu_2}+e^{\mu_5}+e^{\mu_1}+e^{\mu_3}+e^{\mu_4}}\right) \\
=&\frac{v_5}{v_3+v_5}\cdot\frac{v_4}{v_1+v_2+v_4+e^{-d}(v_3+v_5)} - \frac{v_1}{(v_1+v_3+v_5)(v_3+v_5)}\cdot\frac{v_4v_5}{v_1+v_2+v_3+v_4+v_5},
\end{aligned}$$

where (a) is based on the results in Lemma F.2. □

Proof of Theorem 1. We first introduce the following notation.

- $U^c = \max_{j\in S_1\cap S_2} U_j$

- $U_1^t = \max_{j\in S_1\setminus S_2} U_j$
- $U_2^t = \max_{j\in S_2\setminus S_1} U_j$
- $U^c(i) = \max_{j\in (S_1\cap S_2)\setminus\{i\}} U_j$
- $U_1^t(i) = \max_{j\in S_1\setminus (S_2\cup\{i\})} U_j$
- $U_2^t(i) = \max_{j\in S_2\setminus (S_1\cup\{i\})} U_j$

Case 1: $i=0$ In this case, the customer doesn't purchase anything in the first period, thus he wouldn't consider $i\in S_1\cap S_2$ in the second period again, the probability that he would purchase product $j\in S_2\setminus S_1$ is

$$\mathbb{P}(U_j\geq \max\{U_2^t(j),U^c,U_0\}|U_0\geq \max\{U_1^t,U^c\}) = \mathbb{P}(U_j\geq \max\{U_2^t(j),U_0\}|U_0\geq \max\{U_1^t,U^c\})$$

The probability can be calculated from Theorem 2.1 in Gao et al. (2021).

Lemma F.7 **(Adapted from Theorem 2.1 Gao et al. (2021))**. *Let $U_i\sim$ Gumbel$(\mu_i,1)$, $i=1,\ldots,4$, and they are all independent, then we have :*

$$\mathbb{P}(U_1\geq \max\{U_2,U_3\}|U_3\geq U_4) = \mathbb{P}(U_1\geq \max\{U_2,U_3,U_4\}) = \frac{e^{\mu_1}}{\sum_i e^{\mu_i}} = \frac{v_1}{\sum_i v_i}.$$

Based on Lemma F.7, we conclude that

$$\mathbb{P}^1(j|i,S_1,S_2) = \begin{cases} 0 & j\in S_2\cap S_1, \\ \frac{v_j}{1+V(S_1)+V(S_2\setminus S_1)} & j\in S_2\setminus S_1. \end{cases}$$

And the conditional probability that the customer doesn't purchase anything is :

$$\begin{aligned}\mathbb{P}(U_0\geq U_2^t|U_0\geq \max\{U_1^t,U^c\}) &= \frac{\mathbb{P}(U_0\geq U_2^t, U_0\geq \max\{U_1^t,U^c\})}{\mathbb{P}(U_0\geq \max\{U_1^t,U^c\})} \\ &= \frac{\mathbb{P}(U_0\geq \max\{U_2^t,U_1^t,U^c\})}{\mathbb{P}(U_0\geq \max\{U_1^t,U^c\})} \\ &= \frac{1+V(S_1)}{1+V(S_1)+V(S_2\setminus S_1)}.\end{aligned}$$

Case 2 $i\in S_1\cap S_2$

Case 2.1 $j\in S_1\cap S_2$ In this case, if the customer still selects a product $j\in S_1\cap S_2$ in the second period, based on the utility maximization principle, we conclude that $j=i$. Then the conditional choice probability is

$$\begin{aligned}\mathbb{P}(U_i\geq U_2^t|U_i\geq \max\{U_1^t,U_0,U^c(i)\}) &= \frac{\mathbb{P}(U_i\geq U_2^t, U_i\geq \max\{U_1^t,U_0,U^c(i)\})}{\mathbb{P}(U_i\geq \max\{U_1^t,U_0,U^c(i)\})} \\ &= \frac{v_i}{1+V(S_1)+V(S_2\setminus S_1)}\cdot\frac{1+V(S_1)}{v_i} \\ &= \frac{1+V(S_1)}{1+V(S_1)+V(S_2\setminus S_1)}.\end{aligned}$$

Case 2.2 $j\in S_2\setminus S_1$ The conditional choice probability in this case is

$$\mathbb{P}(U_j\geq \max\{U_2^t(j),U_0,U^c(i),U_i\}|U_i\geq \max\{U_1^t,U_0,U^c(i)\}) = \mathbb{P}(U_j\geq \max\{U_2^t(j),U_i\}|U_i\geq \max\{U_1^t,U_0,U^c(i)\}).$$

By Lemma F.7, this probability equals to

$$\mathbb{P}(U_j\geq \max\{U_2^t(j),U_i,U_1^t,U_0,U^c(i)\}) = \frac{v_j}{1+V(S_1)+V(S_2\setminus S_1)}.$$

Case 3 $i\in S_1\setminus S_2$

Case 3.1 $j \in S_2 \setminus S_1$ In this case, the conditional choice probability is

$$\mathbb{P}(U_j \geq \max\{U_2^t(j), U^c, U_0\} | U_i \geq \max\{U^c, U_0, U_1^t(i)\}).$$

Since this conditional probability is hard to solve directly, we consider the joint probability:

$$\mathbb{P}(U_j \geq \max\{U_2^t(j), U^c, U_0\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\}).$$

Denote $U_0^c = \max\{U_0, U^c\}$, then we have

$$\mathbb{P}(U_j \geq \max\{U_2^t(j), U^c, U_0\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\}) = \mathbb{P}(U_j \geq \max\{U_2^t(j), U_0^c\}, U_i \geq \max\{U_0^c, U_1^t(i)\}).$$

Then based on Lemma F.4, we could write the joint probability as:

$$\frac{v_i v_j (1 + V(S_1 \cap S_2))}{V(S_1 \setminus S_2) \cdot V(S_2 \setminus S_1)} \cdot \left[\frac{1}{1 + V(S_1 \cap S_2)} - \frac{1}{1 + V(S_1)} - \frac{1}{1 + V(S_2)} + \frac{1}{1 + V(S_1) + V(S_2 \setminus S_1)} \right]$$

and the conditional probability is :

$$\frac{v_j \cdot (1 + V(S_1)) \cdot (1 + V(S_1 \cap S_2))}{V(S_1 \setminus S_2) \cdot V(S_2 \setminus S_1)} \cdot \left[\frac{1}{1 + V(S_1 \cap S_2)} - \frac{1}{1 + V(S_1)} - \frac{1}{1 + V(S_2)} + \frac{1}{1 + V(S_1) + V(S_2 \setminus S_1)} \right]$$

which can be simplified as

$$v_j \cdot \left[\frac{1}{1 + V(S_1) + V(S_2 \setminus S_1)} + \frac{1 + V(S_1)}{(1 + V(S_2)) \cdot (1 + V(S_1) + V(S_2 \setminus S_1))} \right].$$

Case 3.2 : $j \in S_2 \cap S_1$ In this case, the joint probability can be written as :

$$\begin{aligned}
\mathbb{P}(U_j \geq \max\{U_2^t, U^c(j), U_0\}, U_i \geq \max\{U_1^t(i), U_j\}) &= \mathbb{P}(U_i \geq \max\{U_1^t(i), U_j\} | U_j \geq \max\{U_2^t, U^c(j), U_0\}) \cdot \mathbb{P}(U_j \geq \max\{U_2^t, U^c(j), U_0\}) \\
&= \mathbb{P}(U_i \geq \max\{U_1^t(i), U_j, U_2^t, U^c(j), U_0\}) \cdot \mathbb{P}(U_j \geq \max\{U_2^t, U^c(j), U_0\}) \\
&= \frac{v_i}{1 + V(S_1) + V(S_2 \setminus S_1)} \cdot \frac{v_j}{1 + V(S_2)}
\end{aligned}$$

Then the conditional probability is

$$\frac{1 + V(S_1)}{1 + V(S_1) + V(S_2 \setminus S_1)} \cdot \frac{v_j}{1 + V(S_2)}.$$

Case 3.3 : $j = 0$ In this case, the joint probability can be written as :

$$\begin{aligned}
\mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\}) &= \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\} | U_0 \geq U^c) \mathbb{P}(U_0 \geq U^c) + \\
&\quad \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\} | U_0 \leq U^c) \mathbb{P}(U_0 \leq U^c) \\
&= \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\} | U_0 \geq U^c) \mathbb{P}(U_0 \geq U^c) \\
&= \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}, U_i \geq \max\{U^c, U_0, U_1^t(i)\}, U_0 \geq U^c) \\
&= \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}, U_i \geq \max\{U_0, U_1^t(i)\}) \\
&= \mathbb{P}(U_i \geq \max\{U_0, U_1^t(i)\} | U_0 \geq \max\{U_2^t, U^c\}) \cdot \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}) \\
&\overset{(a)}{=} \mathbb{P}(U_i \geq \max\{U_0, U_1^t(i), U_2^t, U^c\}) \cdot \mathbb{P}(U_0 \geq \max\{U_2^t, U^c\}) \\
&= \frac{v_i}{1 + V(S_1) + V(S_2 \setminus S_1)} \cdot \frac{1}{1 + V(S_2)}
\end{aligned}$$

Where (a) is from Lemma F.7. Then the conditional probability is

$$\frac{1 + V(S_1)}{1 + V(S_1) + V(S_2 \setminus S_1)} \cdot \frac{1}{1 + V(S_2)}.$$

□

Proof of Lemma 1.

For a given assortment $S_2 \setminus \{i\} \subseteq \mathcal{N}_2 \setminus S_1$ and $j \in \mathcal{N}_2 \setminus \{S_1 \cup S_2\}$, we have:

$$
\begin{aligned}
R^2(S_2 \cup \{j\}|i, S_1) - R^2(S_2|i, S_1) &= \frac{W(S_2 \setminus S_1) + r_j v_j + r_i(1+V(S_1))}{1+V(S_1)+V(S_2 \setminus S_1)+v_j} - \frac{W(S_2 \setminus S_1) + r_i(1+V(S_1))}{1+V(S_1)+V(S_2 \setminus S_1)} \\
&= \frac{v_j}{1+V(S_1)+V(S_2 \setminus S_1)+v_j} \cdot \left(r_j - \frac{W(S_2 \setminus S_1)}{1+V(S_1)+V(S_2 \setminus S_1)} \right) \\
&\quad - r_i(1+V(S_1)) \cdot \left(\frac{1}{1+V(S_1)+V(S_2 \setminus S_1)} - \frac{1}{1+V(S_1)+V(S_2 \setminus S_1)+v_j} \right) \\
&= \frac{v_j}{1+V(S_1)+V(S_2 \setminus S_1)+v_j} \cdot \left(r_j - \frac{W(S_2 \setminus S_1) + r_i(1+V(S_1))}{1+V(S_1)+V(S_2 \setminus S_1)} \right) \\
&= \frac{v_j}{1+V(S_1)+V(S_2 \setminus S_1)+v_j} \cdot (r_j - R^2(S_2|i, S_1)).
\end{aligned}
$$

Thus, adding a product would increase the expected revenue as long as its revenue is larger than a threshold.

□

Proof of Lemma 2. We first simplify the expression of $R^3(S_2|i, S_1)$ as follows.

$$
\begin{aligned}
R^3(S_2|i, S_1) = &\frac{1+V(S_1)}{1+V(S_2)} \cdot \frac{W(S_1 \cap S_2)}{1+V(S_1)+V(S_2 \setminus S_1)} + \\
&W(S_2 \setminus S_1) \cdot \left(\frac{1}{1+V(S_1)+V(S_2 \setminus S_1)} + \frac{1+V(S_1)}{(1+V(S_2)) \cdot (1+V(S_1)+V(S_2 \setminus S_1))} \right) \\
= &\frac{1+V(S_1)}{1+V(S_2)} \cdot \frac{W(S_2)}{1+V(S_1)+V(S_2 \setminus S_1)} + \frac{W(S_2 \setminus S_1)}{1+V(S_1)+V(S_2 \setminus S_1)}.
\end{aligned}
$$

Then given an assortment S_2 and $k \in S_1 \cap \mathcal{N}_2, k \neq i$, the difference in the expected revenue is $R^3(S_2 \cup \{k\}|i, S_1) - R^3(S_2|i, S_1)$

$$
\begin{aligned}
&= \frac{1+V(S_1)}{1+V(S_2)+v_k} \cdot \frac{W(S_2) + r_k v_k}{1+V(S_1)+V(S_2 \setminus S_1)} - \frac{1+V(S_1)}{1+V(S_2)} \cdot \frac{W(S_2)}{1+V(S_1)+V(S_2 \setminus S_1)} \\
&= \frac{1+V(S_1)}{1+V(S_2)+v_k} \cdot \frac{r_k v_k}{1+V(S_1)+V(S_2 \setminus S_1)} - \frac{(1+V(S_1)) \cdot W(S_2)}{1+V(S_1)+V(S_2 \setminus S_1)} \cdot \frac{v_k}{(1+V(S_2)) \cdot (1+V(S_2)+v_k)} \\
&= \frac{1+V(S_1)}{1+V(S_2)+v_k} \cdot \frac{v_k}{1+V(S_1)+V(S_2 \setminus S_1)} \cdot \left(r_k - \frac{W(S_2)}{1+V(S_2)} \right) \\
&= \frac{1+V(S_1)}{1+V(S_2)+v_k} \cdot \frac{v_k}{1+V(S_1)+V(S_2 \setminus S_1)} \cdot (r_k - R(S_2)).
\end{aligned}
$$

It shows that the revenue would increase if and only if $r_k \geq R(S_2)$. Assume the optimal conditional assortment in this case is S_2^*, then $k \notin S_2^*$ is equivalent to $R^3(S_2^* \cup \{k\}|i, S_1) < R^3(S_2^*|i, S_1)$, which is equivalent to $r_k < R(S_2^*)$. Thus the intersection with the assortment in the first period S_1 contains all products in $S_1 \cap \mathcal{N}_2$ with revenue larger than $R(S_2^*)$. □

Proof of Lemma 3. Let $L = V(S_1)$. For any $S \subseteq \mathcal{N}_2$, the objective of Problem (9) can be written as

$$
\Phi_L(S) = W(S) \left(\frac{1}{1+V(S)} - \frac{1}{(1+L)(1+L+V(S))} \right) = R(S) - H_L(S),
$$

where $R(S) = W(S)/(1+V(S))$ is the standard MNL revenue and $H_L(S) = W(S)/[(1+L)(1+L+V(S))]$.

Let $\mathcal{R}$ be the family of revenue-ordered assortments. By the standard revenue-ordering property of the MNL assortment problem, no $S \notin \mathcal{R}$ can maximize $R(S)$. Indeed, if $i \notin S$, $j \in S$, and $r_i > r_j$, then either adding i or deleting j strictly improves $R(S)$, according as $r_i > R(S)$ or $r_i \leq R(S)$.

Since there are finitely many assortments, the optimality gap

$$\Delta := R^* - \max_{S\subseteq\mathcal{N}_2,\ S\notin\mathcal{R}} R(S)$$

is strictly positive, where $R^* = \max_{S\subseteq\mathcal{N}_2} R(S)$. Let $S^M \in \arg\max_{S\subseteq\mathcal{N}_2} R(S)$. Then $S^M \in \mathcal{R}$. Moreover, for every $S\subseteq\mathcal{N}_2$,

$$0 \le H_L(S) \le \frac{W_{\max}}{(1+L)^2}, \qquad W_{\max} := \sum_{i\in\mathcal{N}_2} r_i v_i.$$

Now take any $S\notin\mathcal{R}$. Since $R(S^M) - R(S) \ge \Delta$ and $H_L(S)\ge 0$,

$$\Phi_L(S^M) - \Phi_L(S) = R(S^M) - R(S) - H_L(S^M) + H_L(S) \ge \Delta - \frac{W_{\max}}{(1+L)^2}.$$

Choose

$$C = \max\left\{0, \sqrt{\frac{2W_{\max}}{\Delta}} - 1\right\}.$$

Then, whenever $L = V(S_1) \ge C$,

$$\Phi_L(S^M) - \Phi_L(S) \ge \frac{\Delta}{2} > 0$$

for every $S\notin\mathcal{R}$. Therefore no non-revenue-ordered assortment can be optimal for Problem (9). Hence every optimal solution is revenue-ordered when $V(S_1)\ge C$. □

Proof of the Proposition 1. We use a reduction to a well-known NP-complete problem, the partition problem. An instance of the partition problem can be defined as follows: given n products indexed by $[n] = \{1,\dots,n\}$. The weight of product i is $c_i \in \mathbb{N}_+$, and we have $\sum_{i\in[n]} c_i = 2T$. The aim of partition problem is to find a subset $S\subseteq[n]$, such that $\sum_{i\in S} c_i = \sum_{i\in[n]\setminus S} c_i = T$. Now we construct an instance of (9) from an instance of the partition problem as follows.

Consider $n+1$ products indexed by $1,\dots,n,n+1$, the revenue and preference weights are as follows:

$$r_i = \begin{cases} \frac{69}{37} & \text{if } i = 1,\dots,n \\ 3 & \text{if } i = n+1, \end{cases}, \quad v_i = \begin{cases} \frac{c_i}{T} & \text{if } i = 1,\dots,n \\ 1 & \text{if } i = n+1, \end{cases}$$

Let $V(S_1)=1$. Then the objective function (9) becomes $\frac{W(S)}{1+V(S)} - \frac{W(S)}{2(2+V(S))}$. We first show the optimal assortment must contain the last product. For any assortment $S\subseteq[n]$, the objective value (9) for the assortment $S\cup\{n+1\}$ is

$$\begin{aligned}
&W(S\cup\{n+1\})\cdot\left[\frac{1}{1+V(S\cup\{n+1\})} - \frac{1}{2(2+V(S\cup\{n+1\}))}\right]\\
=&\left(V(S)\cdot\frac{69}{37}+3\right)\cdot\left[\frac{1}{1+V(S)+1} - \frac{1}{2(2+V(S)+1)}\right]\\
\ge&\frac{69}{37}\cdot(V(S)+1)\cdot\left[\frac{1}{1+V(S)+1} - \frac{1}{2(2+V(S)+1)}\right]\\
\ge&\frac{69}{37}\cdot V(S)\cdot\left[\frac{1}{1+V(S)} - \frac{1}{2(2+V(S))}\right].
\end{aligned}$$

The last inequality is due to the monotone-increasingness of function $x\cdot(\frac{1}{1+x} - \frac{1}{2(2+x)}) = \frac{1}{2} - \frac{1}{(1+x)(2+x)}$ for $x\ge 0$. This implies that including product $n+1$ strictly improves the objective value. Then we have:

$$\begin{aligned}
\max_{S\subseteq[n+1]} \frac{W(S)}{1+V(S)} - \frac{W(S)}{2(2+V(S))} &= \max_{S\subseteq[n]} \frac{W(S)+r_{n+1}v_{n+1}}{1+V(S)+v_{n+1}} - \frac{W(S)+r_{n+1}v_{n+1}}{2(2+V(S)+v_{n+1})}\\
&= \max_{S\subseteq[n]}\left[\frac{1}{2+V(S)} - \frac{1}{2(3+V(S))}\right](V(S)\cdot\frac{69}{37}+3).
\end{aligned}$$

Define $F(x) = \left[\frac{1}{2+x} - \frac{1}{2(3+x)}\right](\frac{69}{37}\cdot x+3)$, then the derivative of F is $\frac{(21x+51)(1-x)}{37(x+2)^2(x+3)^2}$, which is strictly positive over the interval $[0,1]$ and strictly negative over the interval $(1,\infty)$. Therefore, F achieves the maximum at 1, and $F(x)\le F(1)$. This implies that

$$\max_{S\subseteq[n]} \frac{W(S)}{1+V(S)} - \frac{W(S)}{2(3+V(S))} = \max_{S\subseteq[n]} F(\sum_{i\in S} c_i\cdot\frac{1}{T}) \le F(1).$$

An assortment achieves the expected revenue $F(1)$ if and only if the equality above holds. In this case, we have $\sum_{i\in S} c_i = T$ for some assortment $S\subseteq[n]$. This implies that one can solve the partition problem. Therefore, an instance of the partition problem can be translated to (9) and thus the latter is also NP-hard. □

Proof of Lemma 4. Denote the optimal conditional assortment in two cases are $S^*_{2,2,i}$ and $S^*_{2,3,i}$. Then the conditional expected revenue is as follows:

$$R^2(S^*_{2,2,i}|i,S_1) = \frac{W(S^*_{2,2,i}\setminus S_1)}{1+V(S_1)+V(S^*_{2,2,i}\setminus S_1)} + \frac{(1+V(S_1))}{1+V(S_1)+V(S^*_{2,2,i}\setminus S_1)}\cdot r_i,$$
$$R^3(S^*_{2,3,i}|i,S_1) = \frac{W(S^*_{2,3,i}\setminus S_1)}{1+V(S_1)+V(S^*_{2,3,i}\setminus S_1)} + \frac{1+V(S_1)}{1+V(S_1)+V(S^*_{2,3,i}\setminus S_1)}\cdot\frac{W(S^*_{2,3,i})}{1+V(S^*_{2,3,i})}.$$

When $r_i \ge \frac{W(S^*_{2,3,i})}{1+V(S^*_{2,3,i})}$, then we have:

$$\begin{aligned} R^3(S^*_{2,3,i}|i,S_1) &\le \frac{W(S^*_{2,3,i}\setminus S_1)}{1+V(S_1)+V(S^*_{2,3,i}\setminus S_1)} + \frac{(1+V(S_1))}{1+V(S_1)+V(S^*_{2,3,i}\setminus S_1)}\cdot r_i \\ &\le \frac{W(S^*_{2,2,i}\setminus S_1)}{1+V(S_1)+V(S^*_{2,2,i}\setminus S_1)} + \frac{(1+V(S_1))}{1+V(S_1)+V(S^*_{2,2,i}\setminus S_1)}\cdot r_i \\ &= R^2(S^*_{2,2,i}|i,S_1). \end{aligned}$$

□

Proof of Theorem 2. Denote the optimal conditional assortment as $S^*_{2,i} = S^*\cup S^*_2$, where $S^*\subseteq \mathcal{N}_1\setminus\{i\}\cap\mathcal{N}_2$ is a revenue-ordered assortment and $S^*_2\subseteq\mathcal{N}_2\setminus S_1$. Given $\alpha\in(0,1)$, we consider the following two cases.

Case 1: $\frac{1+V(S_1)}{1+V(S^*)+V(S^*_2)}\cdot\frac{W(S^*)}{1+V(S_1)+V(S^*_2)} \ge \alpha R(S^*_{2,i}|i,S_1)$. In this case, if we only offer the assortment S^*, the expected revenue can be lower bounded as follows,

$$\frac{1+V(S_1)}{1+V(S^*)}\cdot\frac{W(S^*)}{1+V(S_1)} \ge \frac{1+V(S_1)}{1+V(S^*)+V(S^*_2)}\cdot\frac{W(S^*)}{1+V(S_1)+V(S^*_2)} \ge \alpha R(S^*_{2,i}|i,S_1).$$

Case 2: $\frac{1+V(S_1)}{1+V(S^*)+V(S^*_2)}\cdot\frac{W(S^*)}{1+V(S_1)+V(S^*_2)} < \alpha R(S^*_{2,i}|i,S_1)$. Notice the expression for the conditional assortment $R(S^*_{2,i}|i,S_1)$ is as follows:

$$\frac{1+V(S_1)}{1+V(S^*)+V(S^*_2)}\cdot\frac{W(S^*)}{1+V(S_1)+V(S^*_2)} + \frac{2+V(S_1)+V(S^*)+V(S^*_2)}{(1+V(S_1)+V(S^*_2))(1+V(S^*)+V(S^*_2))}\cdot W(S^*_2).$$

Based on our condition, we have

$$\frac{1+V(S_1)}{1+V(S^*)+V(S^*_2)}\cdot\frac{W(S^*)}{1+V(S_1)+V(S^*_2)} \le \frac{\alpha}{1-\alpha}\cdot\frac{2+V(S_1)+V(S^*)+V(S^*_2)}{(1+V(S_1)+V(S^*_2))(1+V(S^*)+V(S^*_2))}\cdot W(S^*_2). \quad \text{(F.3)}$$

Then an upper bound for the optimal expected revenue is $\frac{1}{1-\alpha}\cdot\frac{2+V(S_1)+V(S^*)+V(S^*_2)}{(1+V(S_1)+V(S^*_2))(1+V(S^*)+V(S^*_2))}\cdot W(S^*_2)$.

Moreover, we can upper bound the right part in (F.3) by the following analysis:

$$\begin{aligned}\frac{2+V(S_1)+V(S^*)+V(S_2^*)}{(1+V(S_1)+V(S_2^*))(1+V(S^*)+V(S_2^*))}\cdot W(S_2^*) &= \frac{W(S_2^*)}{1+V(S^*)+V(S_2^*)}+\frac{1+V(S^*)}{1+V(S^*)+V(S_2^*)}\cdot\frac{W(S_2^*)}{1+V(S_1)+V(S_2^*)}\\ &\le \frac{W(S_2^*)}{1+V(S^*)+V(S_2^*)}+\frac{W(S_2^*)}{1+V(S_1)+V(S_2^*)}\\ &\le 2\cdot\frac{W(S_2^*)}{1+V(S_2^*)}\\ &\le 2\max_{S_2\subseteq\mathcal{N}_2\setminus S_1}\frac{W(S_2)}{1+V(S_2)}.\end{aligned}$$

Denote $S_2^{ro}=\arg\max_{S_2\subseteq\mathcal{N}_2\setminus S_1}\frac{W(S_2)}{1+V(S_2)}$, which is a revenue-ordered assortment. Then consider the policy that only provides S_2^{ro}, the expected revenue is

$$\frac{W(S_2^{ro})}{1+V(S_2^{ro})}+\frac{1}{1+V(S_2^{ro})}\cdot\frac{W(S_2^{ro})}{1+V(S_1)+V(S_2^{ro})}\ge\frac{W(S_2^{ro})}{1+V(S_2^{ro})}\ge\frac{1}{2}\cdot\frac{2+V(S_1)+V(S^*)+V(S_2^*)}{(1+V(S_1)+V(S_2^*))(1+V(S^*)+V(S_2^*))}\cdot W(S_2^*).$$

Then the ratio of the expected revenue can be obtained as follows:

$$\begin{aligned}\frac{\frac{W(S_2^{ro})}{1+V(S_2^{ro})}+\frac{1}{1+V(S_2^{ro})}\cdot\frac{W(S_2^{ro})}{1+V(S_1)+V(S_2^{ro})}}{R(S_{2,i}^*|i,S_1)} &\ge \frac{\frac{1}{2}\cdot\frac{2+V(S_1)+V(S^*)+V(S_2^*)}{(1+V(S_1)+V(S_2^*))(1+V(S^*)+V(S_2^*))}\cdot W(S_2^*)}{\frac{1}{1-\alpha}\cdot\frac{2+V(S_1)+V(S^*)+V(S_2^*)}{(1+V(S_1)+V(S_2^*))(1+V(S^*)+V(S_2^*))}\cdot W(S_2^*)}\\ &= \frac{1-\alpha}{2}.\end{aligned}$$

Finally, let $\alpha=\frac{1-\alpha}{2}$, we have $\alpha=\frac{1}{3}$. In summary, the assortments S^* and S_2^{ro} in two cases separately could guarantee an approximation ratio $\frac{1}{3}$. Notice that S^* and S_2^{ro} are the feasible revenue-ordered assortments, then the optimal union of two revenue-ordered assortments under our policy provides at least the approximation ratio $\frac{1}{3}$.

Then we consider the special case that $S_1\cap\mathcal{N}_2=\emptyset$. The conditional assortment optimization problem is as follows:

$$\max_{S_2} W(S_2)\cdot\left(\frac{1}{1+V(S_2)}-\frac{1}{(1+V(S_1))\cdot(1+V(S_1)+V(S_2))}\right).$$

Let $\alpha=1+V(S_1)$, we denote the objective function as $R(S,\alpha)$ and simplify it as follows:

$$\begin{aligned}\left(\frac{1}{1+V(S)}-\frac{1}{\alpha(\alpha+V(S))}\right)\cdot W(S) &= \frac{\alpha(\alpha+V(S))-(1+V(S))}{\alpha(1+V(S))(\alpha+V(S))}\cdot W(S)\\ &= \frac{\alpha^2-1+(\alpha-1)V(S)}{\alpha(1+V(S))(\alpha+V(S))}\cdot W(S)\\ &= \frac{\alpha-1}{\alpha}\cdot\frac{1+\alpha+V(S)}{(1+V(S))(\alpha+V(S))}\cdot W(S)\\ &= \frac{\alpha-1}{\alpha}\cdot\left(1+\frac{1}{\alpha+V(S)}\right)\cdot\frac{W(S)}{1+V(S)}\end{aligned}$$

Let $S_2^{ro}=\arg\max_{S\subseteq\mathcal{N}_2}\frac{W(S)}{1+V(S)}$ and S^* be the optimal conditional assortment, then we have

$$\begin{aligned}R(S^*,\alpha) &= \frac{\alpha-1}{\alpha}\cdot\left(1+\frac{1}{\alpha+V(S^*)}\right)\cdot\frac{W(S^*)}{1+V(S^*)}\\ &\le \frac{\alpha-1}{\alpha}\cdot\left(1+\frac{1}{\alpha}\right)\cdot\frac{W(S^*)}{1+V(S^*)}\\ &\le \frac{\alpha-1}{\alpha}\cdot\left(1+\frac{1}{\alpha}\right)\frac{W(S_2^{ro})}{1+V(S_2^{ro})}\\ &= \frac{\alpha^2-1}{\alpha^2}\frac{W(S_2^{ro})}{1+V(S_2^{ro})}.\end{aligned}$$

Where the first inequality is due to the fact that $V(S^*) \geq 0$ and the second inequality uses the definition of S_2^{ro}. Finally, we have

$$\begin{aligned}R(S_2^{ro},\alpha) &= \frac{\alpha-1}{\alpha}\cdot\left(1+\frac{1}{\alpha+V(S_2^{ro})}\right)\cdot\frac{W(S_2^{ro})}{1+V(S_2^{ro})}\\ &\geq \frac{\alpha-1}{\alpha}\cdot\frac{W(S_2^{ro})}{1+V(S_2^{ro})}\\ &\geq \frac{\alpha-1}{\alpha}\cdot\frac{\alpha^2}{\alpha^2-1}\cdot R(S^*,\alpha)\\ &= \frac{\alpha}{1+\alpha}\cdot R(S^*,\alpha).\end{aligned}$$

The result shows providing the revenue-ordered assortment S_2^{ro} guarantees an approximation ratio $\frac{\alpha}{1+\alpha}$, which is $\frac{1+V(S_1)}{2+V(S_1)}$.

□

Proof of Proposition 2.

For any (g_j,h_j), the dynamic programming (A.4) returns an assortment S_j satisfies:

$$\begin{aligned}\sum_{i\in S_j}\gamma_{j,i} &= \sum_{i\in S_j}\left\lfloor\frac{|\mathcal{N}_j|}{\epsilon\cdot h_j}\cdot r_i v_i\right\rfloor \geq \lfloor|\mathcal{N}_j|/\epsilon\rfloor - |\mathcal{N}_j|,\\ \sum_{i\in S_j} v_{j,i} &= \sum_{i\in S_j}\left\lceil\frac{|\mathcal{N}_j|}{\epsilon\cdot g_j}\cdot v_i\right\rceil \leq \lceil|\mathcal{N}_j|/\epsilon\rceil + |\mathcal{N}_j|.\end{aligned}$$

We have the following inequalities:

$$\begin{aligned}\sum_{i\in S_j}\frac{|\mathcal{N}_j|}{\epsilon\cdot h_j}\cdot r_i v_i &\geq \lfloor|\mathcal{N}_j|/\epsilon\rfloor - |\mathcal{N}_j|,\\ \sum_{i\in S_j}\frac{|\mathcal{N}_j|}{\epsilon\cdot g_j}\cdot v_i &\leq \lceil|\mathcal{N}_j|/\epsilon\rceil + |\mathcal{N}_j|.\end{aligned}$$

Which means

$$\sum_{i\in S_j} r_i v_i \geq \left(\frac{\lfloor|\mathcal{N}_j|/\epsilon\rfloor}{|\mathcal{N}_j|}-1\right)\cdot\epsilon h_j \geq (1-2\epsilon)h_j, \tag{F.4}$$

$$\sum_{i\in S_j} v_i \leq \left(\frac{\lceil|\mathcal{N}_j|/\epsilon\rceil}{|\mathcal{N}_j|}+1\right)\cdot\epsilon g_j \leq (1+2\epsilon)g_j \leq (1+\epsilon)^2 g_j. \tag{F.5}$$

The second inequality in (F.4) and (F.5) holds since $\frac{\lfloor|\mathcal{N}_j|/\epsilon\rfloor}{|\mathcal{N}_j|} \geq \frac{1}{\epsilon}-1$ and $\frac{\lceil|\mathcal{N}_j|/\epsilon\rceil}{|\mathcal{N}_j|} \leq \frac{1}{\epsilon}+1$.

Moreover, for a given assortment S, assume that $\sum_{i\in S} r_i v_i \geq h_j, \sum_{i\in S} v_i \leq g_j$, we have

$$\begin{aligned}\sum_{i\in S}\left\lfloor\frac{|\mathcal{N}_j|}{\epsilon\cdot h_j}\cdot r_i v_i\right\rfloor &\geq \sum_{i\in S}\left(\frac{|\mathcal{N}_j|}{\epsilon\cdot h_j}\cdot r_i v_i - 1\right) \geq \frac{|\mathcal{N}_j|}{\epsilon} - |S| \geq \lfloor|\mathcal{N}_j|/\epsilon\rfloor - |\mathcal{N}_j|,\\ \sum_{i\in S}\left\lceil\frac{|\mathcal{N}_j|}{\epsilon\cdot g_j}\cdot v_i\right\rceil &\leq \sum_{i\in S}\left(\frac{|\mathcal{N}_j|}{\epsilon\cdot g_j}\cdot v_i + 1\right) \leq \frac{|\mathcal{N}_j|}{\epsilon} + |S| \leq \lceil|\mathcal{N}_j|/\epsilon\rceil + |\mathcal{N}_j|.\end{aligned}$$

Assume the optimal solution satisfies that:

$$\begin{aligned}g_1^*\cdot(1+\epsilon)^{-1} &\leq \sum_{i\in S_1^*} v_i \leq g_1^*, h_1^*\cdot(1+\epsilon)^{-1} \leq \sum_{i\in S_1^*} r_i v_i \leq h_1^*,\\ g_2^*\cdot(1+\epsilon)^{-1} &\leq \sum_{i\in S_{2,0}^*} v_i \leq g_2^*, h_2^*\cdot(1+\epsilon)^{-1} \leq \sum_{i\in S_{2,0}^*} r_i v_i \leq h_2^*,\\ g_3^*\cdot(1+\epsilon)^{-1} &\leq \sum_{i\in S_{2,1}^*} v_i \leq g_3^*, h_3^*\cdot(1+\epsilon)^{-1} \leq \sum_{i\in S_{2,1}^*} r_i v_i \leq h_3^*.\end{aligned}$$

Then we can find a set of assortment $(S_1, S_{2,0}, S_{2,1})$ from the dynamic program in the guesses $\{(g_i^*, h_i^*(1+\epsilon)^{-1})\}_{i=1}^3$, such that:

$$V(S_1^*) \le V(S_1) = \sum_{i\in S_1} v_i \le g_1^* \cdot (1+\epsilon)^2, W(S_1) = \sum_{i\in S_1} r_i v_i \ge (1-2\epsilon) h_1^*(1+\epsilon)^{-1},$$
$$V(S_{2,0}^*) \le V(S_{2,0}) = \sum_{i\in S_{2,0}} v_i \le g_2^* \cdot (1+\epsilon)^2, W(S_{2,0}) = \sum_{i\in S_{2,0}} r_i v_i \ge (1-2\epsilon) h_2^*(1+\epsilon)^{-1},$$
$$V(S_{2,1}^*) \le V(S_{2,1}) = \sum_{i\in S_{2,1}} v_i \le g_3^* \cdot (1+2\epsilon), W(S_{2,1}) = \sum_{i\in S_{2,1}} r_i v_i \ge (1-2\epsilon) h_3^*(1+\epsilon)^{-1}.$$

We bound the expected revenue generated by $S_1, S_{2,0}, S_{2,1}$, respectively. For the expected revenue in the first period, we have

$$\frac{W(S_1)}{1+V(S_1)} \ge \frac{(1-2\epsilon)h_1^*(1+\epsilon)^{-1}}{1+g_1^*\cdot(1+\epsilon)^2} \ge \frac{(1-2\epsilon)W(S_1^*)(1+\epsilon)^{-1}}{1+V(S_1^*)\cdot(1+\epsilon)^3} \ge \frac{1-2\epsilon}{(1+\epsilon)^4}\cdot\frac{W(S_1^*)}{1+V(S_1^*)} \ge (1-6\epsilon)R(S_1^*).$$

Then we consider the expected revenue when the customer makes no purchase in the first period.

$$\begin{aligned}\frac{1}{1+V(S_1)}\cdot\frac{W(S_{2,0})}{1+V(S_1)+V(S_{2,0})} &\ge \frac{1}{1+g_1^*\cdot(1+\epsilon)^2}\cdot\frac{(1-2\epsilon)h_2^*(1+\epsilon)^{-1}}{1+g_1^*\cdot(1+\epsilon)^2+g_2^*\cdot(1+\epsilon)^2}\\
&\ge \frac{1-2\epsilon}{(1+\epsilon)^7}\cdot\frac{1}{1+V(S_1^*)}\cdot\frac{W(S_{2,0}^*)}{1+V(S_1^*)+V(S_{2,0}^*)}\\
&\ge (1-9\epsilon)\cdot\frac{1}{1+V(S_1^*)}\cdot\frac{W(S_{2,0}^*)}{1+V(S_1^*)+V(S_{2,0}^*)}.\end{aligned}$$

For the last part $W(S_{2,1})\left(\frac{1}{1+V(S_{2,1})} - \frac{1}{(1+V(S_1))\cdot(1+V(S_1)+V(S_{2,1}))}\right)$, the lower bound is as follows:

$$\begin{aligned}&\frac{(1-2\epsilon)h_3^*(1+\epsilon)^{-1}}{1+g_3^*\cdot(1+2\epsilon)} - \frac{(1-2\epsilon)h_3^*(1+\epsilon)^{-1}}{(1+V(S_1))\cdot(1+V(S_1)+V(S_{2,1}))}\\
\ge&\frac{(1-2\epsilon)W(S_{2,1}^*)(1+\epsilon)^{-1}}{1+(1+\epsilon)V(S_{2,1}^*)\cdot(1+2\epsilon)} - \frac{(1-2\epsilon)W(S_{2,1}^*)}{(1+V(S_1^*))\cdot(1+V(S_1^*)+V(S_{2,1}^*))}\\
\ge&\frac{(1-2\epsilon)W(S_{2,1}^*)}{1+V(S_{2,1}^*)} - \frac{(1-2\epsilon)W(S_{2,1}^*)}{(1+V(S_1^*))\cdot(1+V(S_1^*)+V(S_{2,1}^*))} -\\
&\left(\frac{(1-2\epsilon)W(S_{2,1}^*)}{1+V(S_{2,1}^*)} - \frac{(1-2\epsilon)W(S_{2,1}^*)(1+\epsilon)^{-1}}{1+(1+\epsilon)V(S_{2,1}^*)\cdot(1+2\epsilon)}\right)\\
=&\frac{(1-2\epsilon)W(S_{2,1}^*)}{1+V(S_{2,1}^*)} - \frac{(1-2\epsilon)W(S_{2,1}^*)}{(1+V(S_1^*))\cdot(1+V(S_1^*)+V(S_{2,1}^*))} -\\
&\frac{\epsilon W(S_{2,1}^*)}{1+V(S_{2,1}^*)}\cdot\frac{(1-2\epsilon)(2\epsilon^2V(S_{2,1}^*)+5\epsilon V(S_{2,1}^*)+4V(S_{2,1}^*)+1)}{(1+\epsilon)(1+(1+\epsilon)(1+2\epsilon)V(S_{2,1}^*)}.\end{aligned}$$

Since we have

$$\begin{aligned}\frac{(1-2\epsilon)(2\epsilon^2V(S_{2,1}^*)+5\epsilon V(S_{2,1}^*)+4V(S_{2,1}^*)+1)}{(1+\epsilon)(1+(1+\epsilon)(1+2\epsilon)V(S_{2,1}^*))} &\le \frac{(2\epsilon^2V(S_{2,1}^*)+5\epsilon V(S_{2,1}^*)+4V(S_{2,1}^*)+1)}{1+(1+\epsilon)(1+2\epsilon)V(S_{2,1}^*)}\\
&= \frac{1+(1+\epsilon)(1+2\epsilon)V(S_{2,1}^*)+(3+2\epsilon)V(S_{2,1}^*)}{1+(1+\epsilon)(1+2\epsilon)V(S_{2,1}^*)}\\
&\le 1+\frac{(3+2\epsilon)V(S_{2,1}^*)}{1+(1+\epsilon)(1+2\epsilon)V(S_{2,1}^*)}\\
&\le 4.\end{aligned}$$

Then the lower bound is

$$\frac{(1-2\epsilon)W(S_{2,1}^*)}{1+V(S_{2,1}^*)} - \frac{(1-2\epsilon)W(S_{2,1}^*)}{(1+V(S_1^*))\cdot(1+V(S_1^*)+V(S_{2,1}^*)))} - \frac{4\epsilon W(S_{2,1}^*)}{1+V(S_{2,1}^*)}.$$

Combining the analysis in the three parts, we conclude that

$$\begin{aligned}\mathscr{R}_A(S_1,S_{2,0},S_{2,1}) &\geq (1-6\epsilon)R(S_1^*)+(1-9\epsilon)\cdot\frac{1}{1+V(S_1^*)}\cdot\frac{W(S_{2,0}^*)}{1+V(S_1^*)+V(S_{2,0}^*)}+\\ &\frac{(1-2\epsilon)W(S_{2,1}^*)}{1+V(S_{2,1}^*)}-\frac{(1-2\epsilon)W(S_{2,1}^*)}{(1+V(S_1^*))\cdot(1+V(S_1^*)+V(S_{2,1}^*)))}-\frac{4\epsilon W(S_{2,1}^*)}{1+V(S_{2,1}^*)}\\ &\geq (1-9\epsilon)\mathscr{R}_A(S_1^*,S_{2,0}^*,S_{2,1}^*)-\frac{4\epsilon W(S_{2,1}^*)}{1+V(S_{2,1}^*)}\\ &\geq (1-13\epsilon)\mathscr{R}_A(S_1^*,S_{2,0}^*,S_{2,1}^*).\end{aligned}$$

In the last inequality, we use the fact that

$$\mathscr{R}_A(S_1^*,S_{2,0}^*,S_{2,1}^*)\geq\mathscr{R}_A(\emptyset,S_{2,1}^*,\emptyset)=\frac{W(S_{2,1}^*)}{1+V(S_{2,1}^*)}.$$

For the running time, notice that $|\Delta_j^\epsilon|=O(\log(|\mathcal{N}_j|\gamma_{\max}^j/\gamma_{\min}^j)/\epsilon)),|\Gamma_j^\epsilon|=O(\log(|\mathcal{N}_j|\alpha_{\max}^j/\alpha_{\min}^j)/\epsilon)$, and for a given $(h_j,g_j)\in\Delta_j^\epsilon\times\Gamma_j^\epsilon$, the total states in the dynamic programming F^j is $O(|\mathcal{N}_j|^3/\epsilon^2)$. Thus the time complexity for calculating the candidates in $\mathcal{S}_\epsilon^j$ is $O(\log(|\mathcal{N}_j|\gamma_{\max}^j/\gamma_{\min}^j)/\epsilon\times\log(|\mathcal{N}_j|\alpha_{\max}^j/\alpha_{\min}^j)/\epsilon\times|\mathcal{N}_j|^3/\epsilon^2)=O(\log(|\mathcal{N}_j|\gamma_{\max}/\gamma_{\min})\times\log(|\mathcal{N}_j|\alpha_{\max}/\alpha_{\min})\times|\mathcal{N}_j|^3/\epsilon^4)$. Finally, we need to enumerate all assortments in $\mathcal{S}_\epsilon^1\times\mathcal{S}_\epsilon^2\times\mathcal{S}_\epsilon^2$, the time complexity is $O(|\mathcal{S}_\epsilon^1|\times|\mathcal{S}_\epsilon^2|^2)$. Since $|\mathcal{S}_\epsilon^j|=|\Delta_j^\epsilon|\times|\Gamma_j^\epsilon|=O(\log(|\mathcal{N}_j|\gamma_{\max}/\gamma_{\min})/\epsilon)\times\log(|\mathcal{N}_j|\alpha_{\max}/\alpha_{\min})/\epsilon)$, the time complexity in the enumeration procedure is $O\left(\frac{\prod_{j=1}^2(\log^j(|\mathcal{N}_j|\gamma_{\max}/\gamma_{\min})\log^j(|\mathcal{N}_j|\alpha_{\max}/\alpha_{\min}))}{\epsilon^6}\right)$. Let $\mathcal{N}=\mathcal{N}_1\cup\mathcal{N}_2$, $\gamma=\frac{\max_{i\in\mathcal{N}}r_iv_i}{\min_{j\in\mathcal{N}}r_jv_j},\alpha=\frac{\max_{i\in\mathcal{N}}v_i}{\min_{j\in\mathcal{N}}v_j}$, we have $\log(|\mathcal{N}_j|\gamma_{\max}/\gamma_{\min}))\log(|\mathcal{N}_j|\alpha_{\max}/\alpha_{\min})$ is upper bounded by $\log(|\mathcal{N}|\gamma)\log(|\mathcal{N}|\alpha)$, then the total time complexity is $O\left(\max\left\{\log(|\mathcal{N}|\gamma)\log(|\mathcal{N}|\alpha)\frac{|\mathcal{N}|^3}{\epsilon^4},\frac{\log^3(|\mathcal{N}|\gamma)\log^3(|\mathcal{N}|\alpha)}{\epsilon^6}\right\}\right)$, which is $O\left(\log^3(|\mathcal{N}|\gamma)\log^3(|\mathcal{N}|\alpha)\frac{|\mathcal{N}|^3}{\epsilon^6}\right)$. $\square$

Proof of Theorem 3. To bound the second term in the expression of $\mathscr{R}_A$ in (11), notice that $\frac{a+V(S)}{a+b+V(S)}\geq\frac{a}{a+b}$ for any $a,b>0$, and thus we have

$$\begin{aligned}\left[\frac{1}{a(a+V(S))}-\frac{1}{(a+b)(a+b+V(S))}\right]W(S)&=\frac{W(S)}{a+V(S)}\cdot\left[\frac{1}{a}-\frac{1}{a+b}\cdot\frac{a+V(S)}{a+b+V(S)}\right]\\ &\leq\frac{W(S)}{a+V(S)}\cdot\left[\frac{1}{a}-\frac{1}{a+b}\cdot\frac{a}{a+b}\right]\end{aligned}\tag{F.6}$$

Using this result, we can upper bound $\mathscr{R}_A$ in (11) as follows:

$$\begin{aligned}\mathscr{R}_A&\leq R(S_1^{ro})+\frac{1}{1+V(S_1^*)}\cdot\frac{W(S_{2,0}^*)}{1+V(S_1^*)+V(S_{2,0}^*)}+\frac{W(S_{2,1}^*)}{1+V(S_{2,1}^*)}\cdot\left(1-\frac{1+V(S_{2,1}^*)}{(1+V(S_1^*))\cdot(1+V(S_1^*)+V(S_{2,1}^*))}\right)\\ &\leq R(S_1^{ro})+\frac{1}{1+V(S_1^*)}\cdot\frac{W(S_{2,0}^*)}{1+V(S_1^*)+V(S_{2,0}^*)}+\left(1-\frac{1}{(1+V(S_1^*))^2}\right)R(S_2^{ro}) &&\text{(F.7)}\\ &\leq R(S_1^{ro})+\left(1+\frac{1}{1+V(S_1^*)}-\frac{1}{(1+V(S_1^*))^2}\right)R(S_2^{ro})\\ &\leq R(S_1^{ro})+\frac{5}{4}R(S_2^{ro}) &&\text{(F.8)}\\ &\leq\frac{5}{4}\cdot\mathscr{R}_F(S_1^{ro},S_2^{ro}).\end{aligned}$$

In the first inequality, we use the fact that S_1^{ro} maximizes $R(S)$ for $S\subseteq\mathcal{N}_1$. Inequality (F.7) utilizes (F.6), where $a=1$ and $b=V(S_1^*)$, as well as the fact that $R(S_{2,1}^*)\leq R(S_2^{ro})$. In the third inequality, we again use the fact that $R(S_{2,0}^*)\leq R(S_2^{ro})$ and $1+V(S_1^*)+V(S_{2,0}^*)>1+V(S_{2,0}^*)$. Inequality (F.8) is due to that the maximum of $g(x)=1+\frac{1}{1+x}-\frac{1}{(1+x)^2}$ is achieved at $x=1$ when $x>0$. Hence, we conclude that $\mathscr{R}_F(S_1^{ro},S_2^{ro})\geq\frac{4}{5}\mathscr{R}_A$. $\square$

Proof of Proposition 3: The difficulty of finding the optimal adaptive policy is to identify the optimal conditional assortment in the second stage. When the consumer purchases a product i from the set S_1, Corollary 1 implies that the case is straightforward if $i \in S^{ro}$. On the other hand, if $i \in S_1 \setminus S^{ro}$, then we need to decide whether to provide product i in period two again, which requires us to find the maximum of $R^2(S^*_{2,2,i}|i,S_1)$ and $R^3(S^*_{2,3,i}|i,S_1)$, the latter of which has been shown to be NP-hard. To avoid this comparison, we construct a unified upper bound of the conditional expected revenue for product $i \in S_1 \setminus S^{ro}$. Then the total expected revenue could be upper bounded by the value of the following function $J(S_1,\{S_{2,i}\}_{i\in S_1\cup\{0\}})$.

$$\frac{W(S_1)}{1+V(S_1)}+\sum_{i\in S_1\cap S^{ro}}\frac{v_i}{1+V(S_1)}R^2(S_{2,i}|i,S_1)+\sum_{i\in S_1\setminus S^{ro}}\frac{v_i}{1+V(S_1)}UP(S_{2,i}|S_1)+\frac{1}{1+V(S_1)}\frac{W(S_{2,0}\setminus S_1)}{1+V(S_1)+V(S_{2,0}\setminus S_1)}, \tag{F.9}$$

where $UP(S_2|S_1) \triangleq \frac{W(S_2\setminus S_1)}{1+V(S_1)+V(S_2\setminus S_1)} + \frac{1+V(S_1)}{1+V(S_1)+V(S_2\setminus S_1)}R^*$ and $R^* \triangleq \max_{S\subseteq\mathcal{N}} R(S) = R(S^{ro})$. Here S^{ro} denotes the revenue-ordered assortment that maximizes the standard MNL revenue for the product set $\mathcal{N}$. Essentially we replace the hard-to-analyze term in (2), $\max\{R^2(S^*_{2,i}|i,S^*_1), R^3(S^*_{2,i}|i,S^*_1)\}$, with $UP(S^*_{2,i}|S^*_1)$. The following lemma shows (F.9) can serve as an upper bound for the original problem (2). We defer the proofs of auxiliary lemmas to the end of the proof of Proposition 3.

Lemma F.8. *The optimal value of the following problem*

$$\max_{S_1,\{S_{2,i}\}_{i\in S_1\cup\{0\}}} J(S_1,\{S_{2,i}\}_{i\in S_1\cup\{0\}}), \tag{F.10}$$

is an upper bound for the joint assortment optimization problem (2).

Denote the optimal solution for the Problem (F.10) as $\{S^*_1,\{S^*_{2,i}\}_{i\in S^*_1\cup\{0\}}\}$, it turns out that analyzing the optimal solution of the function $J(S_1,\{S_{2,i}\}_{i\in S_1\cup\{0\}})$ is easier. We first characterize the optimal solution to the sub-problem that maximizing $UP(S_2|S_1)$ in the following auxiliary lemma.

Lemma F.9. *Let $S^*_2 = \arg\max_{S_2\subseteq\mathcal{N}\setminus S_1} UP(S_2|S_1)$, then S^*_2 is a revenue-ordered assortment. Moreover, $S^*_2 = \emptyset$ when $S^{ro}\subseteq S_1$.*

Based on the above lemma, we could further reveal the structural property of S^*_1.

Lemma F.10. *The optimal solution for the upper bound $S^*_1 \subseteq S^{ro}$ and it is revenue-ordered.*

Combining the auxiliary lemmas, we then give a formal proof of Proposition 3 as follows. Denote the optimal solution for the Problem (2) and (F.10) as $\{S^*_1,\{S^*_{2,i}\}_{i\in S^*_1\cup\{0\}}\}$ and $\{S^*_3,\{S_{4,i}\}_{i\in S^*_3\cup\{0\}}\}$. In Lemma F.8, we conclude that

$$J(S^*_3,\{S^*_{4,i}\}_{i\in S^*_3\cup\{0\}}) \geq \mathscr{R}_A(S^*_1,\{S^*_{2,i}\}_{i\in S^*_1\cup\{0\}}). \tag{F.11}$$

Then in Lemma F.10, we show S^*_3 is revenue-ordered and it is a subset of S^{ro}. In this case, we have $S^*_{4,i}=\{i\}$ for $i\in S^*_3$, according to Corollary 1. Then

$$J(S^*_3,\{S^*_{4,i}\}_{i\in S^*_1\cup\{0\}}) = \frac{2W(S^*_3)}{1+V(S^*_3)}+\frac{1}{1+V(S^*_3)}\cdot\frac{W(S^*_{4,0}\setminus S^*_3)}{1+V(S^*_3)+V(S^*_{4,0}\setminus S^*_3)} = \mathscr{R}_A(S^*_3,\{S^*_{4,i}\}_{i\in S^*_3\cup\{0\}}). \tag{F.12}$$

Combined with (F.11) and (F.12), we conclude that $\{S^*_3,\{S^*_{4,i}\}_{i\in S^*_3\cup\{0\}}\}$ is the optimal solution to problem (2). $\square$

Next we provide proofs for Lemmas F.8 to F.10.

Proof of Lemma F.8: Denote the optimal solution to the problem (2) as $\{S_1^*, \{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}}\}$. Then the difference between $\mathscr{R}_A(S_1^*, \{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})$ and $J(S_1^*, \{S_{2,i}^*\}_{i\in S_1\cup\{0\}})$ can be represented as follows:

$$\sum_{i\in S_1^*\setminus S^{ro}} \max\{R^2(S_{2,i}^*|i,S_1^*), R^3(S_{2,i}^*|i,S_1^*)\} - UP(S_{2,i}^*|S_1^*).$$

Then we show for $i\in S_1^*\setminus S^{ro}$, the conditional expected revenue is upper bounded by $UP(S_{2,i}^*|S_1^*)$ through the following analysis:

$$\begin{aligned}
R^3(S_{2,i}^*|i,S_1^*) &= \frac{W(S_{2,i}^*\setminus S_1^*)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)} + \frac{1+V(S_1^*)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\cdot\frac{W(S_{2,i}^*)}{1+V(S_{2,i}^*)}\\
&\le \frac{W(S_{2,i}^*\setminus S_1^*)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)} + \frac{1+V(S_1^*)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\cdot R^* = UP(S_{2,i}^*|S_1^*),
\end{aligned}$$

and

$$\begin{aligned}
R^2(S_{2,i}^*|i,S_1^*) &= \frac{W(S_{2,i}^*\setminus S_1^*)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)} + \frac{(1+V(S_1^*))}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\cdot r_i\\
&\le \frac{W(S_{2,i}^*\setminus S_1^*)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)} + \frac{(1+V(S_1^*))}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\cdot R^* = UP(S_{2,i}^*|S_1^*),
\end{aligned}$$

In the first inequality, we use the result that $\frac{W(S_{2,i}^*)}{1+V(S_{2,i}^*)} \le R^*$ and the second inequality is due to the fact that $r_i \le R^*$ for $i \notin S^{ro}$ according to Lemma 5. Then the analysis shows that $\mathscr{R}_A(S_1^*, \{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}}) \le J(S_1, \{S_{2,i}^*\}_{i\in S_1\cup\{0\}})$. Since $\{S_1^*, \{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}}\}$ is a feasible solution for Problem (F.10), the optimal value of the Problem (F.10) is at least $\mathscr{R}_A(S_1^*, \{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})$ and we conclude the proof. $\square$

Next, we give a proof for Lemma F.9.

Proof of Lemma F.9: The first part of the proof is similar to the result in Lemma 1. For a given assortment $S_2 \subseteq \mathcal{N}\setminus S_1$ and $j\in\mathcal{N}\setminus(S_1\cup S_2)$, we have:

$$\begin{aligned}
UP(S_2\cup\{j\}|S_1) - UP(S_2|S_1) &= \frac{W(S_2\setminus S_1)+r_jv_j+(1+V(S_1))R^*}{1+V(S_1)+V(S_2\setminus S_1)+v_j} - \frac{W(S_2\setminus S_1)+(1+V(S_1))R^*}{1+V(S_1)+V(S_2\setminus S_1)}\\
&= \frac{v_j}{1+V(S_1)+V(S_2\setminus S_1)+v_j}\cdot\left(r_j - \frac{W(S_2\setminus S_1)}{1+V(S_1)+V(S_2\setminus S_1)}\right)\\
&\quad - R^*(1+V(S_1))\cdot\left(\frac{1}{1+V(S_1)+V(S_2\setminus S_1)} - \frac{1}{1+V(S_1)+V(S_2\setminus S_1)+v_j}\right)\\
&= \frac{v_j}{1+V(S_1)+V(S_2\setminus S_1)+v_j}\cdot\left(r_j - \frac{W(S_2\setminus S_1)+R^*(1+V(S_1))}{1+V(S_1)+V(S_2\setminus S_1)}\right)\\
&= \frac{v_j}{1+V(S_1)+V(S_2\setminus S_1)+v_j}\cdot(r_j - UP(S_2|S_1)).
\end{aligned}$$

Thus, adding a product would increase the expected revenue as long as its revenue is larger than a threshold. For the second part, notice that $UP(S_2|S_1)$ can be rewritten as follows

$$UP(S_2|S_1) = \frac{V(S_2\setminus S_1)}{1+V(S_1)+V(S_2\setminus S_1)}\cdot\frac{W(S_2\setminus S_1)}{V(S_2\setminus S_1)} + \frac{1+V(S_1)}{1+V(S_1)+V(S_2\setminus S_1)}\cdot R^*,$$

which is a convex combination of $\frac{W(S_2\setminus S_1)}{V(S_2\setminus S_1)}$ and R^*. Then S_2^* should only contain product with revenue larger than R^*. When $S^{ro}\subseteq S_1$, we have $r_i < R^*$ for $i\in\mathcal{N}\setminus S_1$. Thus, the maximum is achieved by setting $S_2^*=\emptyset$ and $UP(S_2^*|S_1)=R^*$. $\square$

Finally, we give a proof for Lemma F.10.

Proof of Lemma F.10: We introduce $\{\gamma_i\}_{i\in\mathcal{N}\cup\{0\}}$ for better illustration, where $\gamma_i = r_i$ for $i \in S^{ro}\cup\{0\}$ and $\gamma_i = R^*$ for $i\in\mathcal{N}\setminus S^{ro}$. And we assume $r_1 \ge r_2 \ge \ldots \ge r_{|\mathcal{N}|} > r_0 = 0$. We consider two cases in the following analysis, that is $S^{ro}\cap S_1^* = S^{ro}$ and $S^{ro}\cap S_1^* \ne S^{ro}$.

Case 1: $S^{ro}\cap S_1^* = S^{ro}$. According to Corollary 1, we have $R^2(S_{2,i}^*|i,S_1^*) = r_i$ for $i\in S^{ro}\cap S_1^*$. Based on Lemma F.9, we have $UP(S_{2,i}^*|S_1^*) = R^*$ for $i\in S_1^*\setminus S^{ro}$. Then the expected revenue can be expressed as follows:

$$J(S_1^*,\{S_{2,i}^*\}_{i\in S_1\cup\{0\}}) = \frac{W(S_1^*)}{1+V(S_1^*)} + \frac{W(S^{ro})}{1+V(S_1^*)} + \frac{V(S_1^*\setminus S^{ro})}{1+V(S_1^*)}R^* + \frac{1}{1+V(S_1^*)}\frac{W(S_{2,0}^*\setminus S_1^*)}{1+V(S_1^*)+V(S_{2,0}^*\setminus S_1^*)}.$$

In this case, we can show the objective value will increase by replacing the assortment S_1^* with S^{ro} and $S_{2,0}^*$ with $S_{2,0}^*\setminus S_1^*$, while maintaining the other assortments. The difference between $J(S^{ro},\{S_{2,i}^*\}_{i\in S^{ro}}, S_{2,0}^*\setminus S_1^*)$ and $J(S_1^*,\{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})$ is as follows.

$$\begin{aligned}
&J(S^{ro},\{S_{2,i}^*\}_{i\in S^{ro}\cup\{0\}}, S_{2,0}^*\setminus S_1^*) - J(S_1^*,\{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})\\
&= 2R^* - R(S_1^*) - \frac{W(S^{ro})}{1+V(S_1^*)} - \frac{V(S_1^*\setminus S^{ro})}{1+V(S_1^*)}R^*\\
&\quad + \frac{1}{1+V(S^{ro})}\frac{W(S_{2,0}^*\setminus S_1^*)}{1+V(S^{ro})+V(S_{2,0}^*\setminus S_1^*)} - \frac{1}{1+V(S_1^*)}\frac{W(S_{2,0}^*\setminus S_1^*)}{1+V(S_1^*)+V(S_{2,0}^*\setminus S_1^*)}\\
&\ge 2R^* - R(S_1^*) - \frac{W(S^{ro})}{1+V(S_1^*)} - \frac{V(S_1^*\setminus S^{ro})}{1+V(S_1^*)}R^*\\
&= R^* - R(S_1^*) + \frac{V(S_1^*\setminus S^{ro})W(S^{ro})}{(1+V(S_1^*))(1+V(S^{ro}))} - \frac{V(S_1^*\setminus S^{ro})}{1+V(S_1^*)}R^*\\
&= R^* - R(S_1^*) + \frac{V(S_1^*\setminus S^{ro})}{1+V(S_1^*)}\left(\frac{W(S^{ro})}{1+V(S^{ro})} - R^*\right)\\
&= R^* - R(S_1^*)\\
&\ge 0.
\end{aligned}$$

In the first inequality, we use the fact that S^{ro} is a subset of S_1^*, which implies $V(S^{ro})\le V(S_1^*)$. Thus in this case, we conclude that if $S^{ro}\subseteq S_1^*$, then $S_1^* = S^{ro}$, which is a revenue-ordered assortment.

Case 2: $S^{ro}\cap S_1^* \ne S^{ro}$. Assume S_1^* is not revenue-ordered. Then we can always represent $S_1^* = S^{k-1}\cup\{k+l,\ldots,\}$, where $S^{k-1} = \{1,\ldots,k-1\}\subseteq S_1^*$ and $k\in S^{ro}$. For $i\in S_1^*\cap S^{ro}$ and $j\in S_1^*\setminus S^{ro}$, we know $S_{2,i}^*\setminus S_1^*$ and $S_{2,j}^*\setminus S_1^*$ is a revenue-ordered assortment in $\mathcal{N}\setminus S_1^*$ based on Lemma 1 and Lemma F.9 separately. Since $r_k = \max_{i\in\mathcal{N}\setminus S_1^*} r_i$, we have $k\in S_{2,i}^*$ for $i\in S_1^*\setminus S^{k-1}\cup\{0\}$. And we know $S_{2,i}^* = \{i\}$ for $i\in S^{k-1}$ based on the results in Corollary 1. Then a new policy $\{S_1^*\cup\{k\},\{S_{2,i}^*\}_{i\in S^k},\{S_{2,i}^*\setminus\{k\}\}_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\}$ is suboptimal, the objective value of this new policy is

$$\begin{aligned}
&\frac{W(S_1^*)+r_kv_k}{1+V(S_1^*)+v_k} + \sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{v_i}{1+V(S_1^*)+v_k}\left(\frac{W(S_{2,i}^*\setminus S_1^*)-r_kv_k}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)} + \frac{(1+V(S_1^*)+v_k)}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\cdot\gamma_i\right)\\
&+\frac{W(S^k)}{1+V(S_1^*)+v_k},
\end{aligned}\tag{F.13}$$

then the difference between $J(S_1^*,\{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})$ and (F.13) is

$$\frac{v_k}{1+V(S_1^*)+v_k}\cdot\left(R(S_1^*) - r_k + \frac{W(S^{k-1})}{1+V(S_1^*)} - r_k + \sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{v_i(r_k+W(S_{2,i}^*\setminus S_1^*)/(1+V(S_1^*)))}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\right)\ge 0,$$

which indicates that

$$\left(2-\frac{\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}v_i}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\right)r_k\le R(S_1^*)+\frac{W(S^{k-1})}{1+V(S_1^*)}+\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{v_iW(S_{2,i}^*\setminus S_1^*)/(1+V(S_1^*))}{(1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*))}. \quad \text{(F.14)}$$

Denote the right hand in (F.14) as B and $2-\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{v_i}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}$ as A. From (F.14), we also conclude that $1\in S_1^*$, otherwise, we have $S^{k-1}=\emptyset$ and

$$\begin{aligned}
B&\le R(S_1^*)+\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{v_i}{1+V(S_1^*)}\cdot\frac{r_1V(S_{2,i}^*\setminus S_1^*)}{(1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*))}\\
&=R(S_1^*)+\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{r_1v_i}{1+V(S_1^*)}\cdot\left(1-\frac{1+V(S_1^*)}{(1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*))}\right)\\
&=R(S_1^*)+r_1-\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}\frac{r_1v_i}{(1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*))}\\
&\le\left(2-\frac{\sum_{i\in S_1^*\setminus S^{k-1}\cup\{0\}}v_i}{1+V(S_1^*)+V(S_{2,i}^*\setminus S_1^*)}\right)r_1.
\end{aligned}$$

In the first inequality, we use the fact that $r_1\ge\frac{W(S_{2,i}^*\setminus S_1^*)}{V(S_{2,i}^*\setminus S_1^*)}$ and $R(S_1^*)\le R^*\le r_1$ is used in the last inequality. Since $B\le Ar_1$, we know the difference between $J(S_1^*,\{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})$ and (F.13) is non-positive, thus adding the product 1 into S_1^* is always profitable, we conclude S^{k-1} is not empty.

Let $p=k+l$, $D=1+V(S_1^*)$, and $T_i=S_{2,i}^*\setminus S_1^*$ for $i\in S_1^*\cup\{0\}$. Define $I=(S_1^*\setminus S^{k-1})\cup\{0\}$ and $I^-=I\setminus\{p\}$. Recall from (F.14) that $Ar_k\le B$, where

$$A=2-\sum_{i\in I}\frac{v_i}{D+V(T_i)},\qquad B=R(S_1^*)+\frac{W(S^{k-1})}{D}+\sum_{i\in I}\frac{v_iW(T_i)}{D(D+V(T_i))}.$$

Now consider the policy obtained by removing product $p=k+l$ from the first-period assortment and adding it to the second-period assortments corresponding to all $i\in I^-$. Denote the objective value of this new policy by $\widetilde{J}_p$. A direct calculation gives

$$\widetilde{J}_p-J(S_1^*,\{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})=\frac{v_p}{D-v_p}\left[B-Ar_p-\frac{W(T_p)-r_pV(T_p)+(D-v_p)(\gamma_p-r_p)}{D+V(T_p)}\right].$$

Thus it suffices to show that

$$\frac{W(T_p)-r_pV(T_p)+(D-v_p)(\gamma_p-r_p)}{D+V(T_p)}<B-Ar_p.$$

Since $T_p\subseteq\mathcal{N}\setminus S_1^*$ and $r_k=\max_{i\in\mathcal{N}\setminus S_1^*}r_i$, we have $W(T_p)\le r_kV(T_p)$. Moreover, $\gamma_p\le r_k$: if $p\in S^{ro}$, then $\gamma_p=r_p<r_k$; if $p\notin S^{ro}$, then $\gamma_p=R^*\le r_k$ because $k\in S^{ro}$. Therefore,

$$\frac{W(T_p)-r_pV(T_p)+(D-v_p)(\gamma_p-r_p)}{D+V(T_p)}\le\frac{(r_k-r_p)V(T_p)+(D-v_p)(r_k-r_p)}{D+V(T_p)}\le r_k-r_p.$$

On the other hand, since $Ar_k\le B$, we have $B-Ar_p\ge A(r_k-r_p)$. Furthermore,

$$A=2-\sum_{i\in I}\frac{v_i}{D+V(T_i)}\ge2-\frac{\sum_{i\in I}v_i}{D}=2-\frac{1+V(S_1^*\setminus S^{k-1})}{D}=1+\frac{V(S^{k-1})}{D}>1,$$

where the last inequality follows from $S^{k-1}\ne\emptyset$ and $v_i>0$. Because $r_k>r_p$, we obtain $B-Ar_p\ge A(r_k-r_p)>r_k-r_p$. Combining the two inequalities yields

$$\frac{W(T_p)-r_pV(T_p)+(D-v_p)(\gamma_p-r_p)}{D+V(T_p)}<B-Ar_p.$$

Hence $\widetilde{J}_p > J(S_1^*, \{S_{2,i}^*\}_{i\in S_1^*\cup\{0\}})$. Therefore, removing product $p=k+l$ from the first-period assortment and adding it to the relevant second-period assortments strictly improves the objective value, contradicting the optimality of S_1^*. Hence no such non-revenue-ordered S_1^* can be optimal.

Through the analysis in two cases, we conclude that $S_1^* \subseteq S^{ro}$ and S_1^* is also a revenue-ordered assortment. $\square$

Proof of Theorem 4: From the expression of $\mathscr{R}_A$ in (13), we can see that the optimal revenue from the adaptive policy is upper bounded by $3R(S^{ro})$:

$$\frac{2W(S_1^*)}{1+V(S_1^*)} + \frac{1}{1+V(S_1^*)} \cdot \frac{W(S_{2,0}^*)}{1+V(S_1^*)+V(S_{2,0}^*)} \leq 2R(S^{ro}) + \frac{W(S_{2,0}^*)}{1+V(S_{2,0}^*)} \leq 3R(S^{ro}).$$

On the other hand, the expected revenue from the optimal fixed policy is $2R(S^{ro})$. We conclude the ratio is at least $\frac{2}{3}$. $\square$

Proof of Proposition 4. We introduce some notations for better illustrations. For simplicity, we index the periods of product purchases $\mathcal{A}_m = \{i\in[m] | c_i \neq 0\}$ as $\{p_1,\dots,p_a\}$. For $c_i\neq 0$, let $U_{-c_i} = \max_{j\in S_i\setminus\{c_i\}} U_j$ and $\mu_{-c_i} = \log(\sum_{j\in S_i\setminus\{c_i\}} e^{\mu_j}) = \log(V(S_i\setminus c_i))$. Based on the property of Gumbel random variable, we know $U_{-c_i}\sim$ Gumbel$(\mu_{-c_i},1)$. Similarly, let $U^{\mathcal{B}_m} = \max_{i\in\cup_{j\in\mathcal{B}_m} S_j}\{U_i\}$ and $\mu^{\mathcal{B}_m} = \log(\sum_{i\in\cup_{j\in\mathcal{B}_m} S_j} e^{\mu_i}) = \log(\sum_{j\in\mathcal{B}_m} V(S_j))$, we have $U^{\mathcal{B}_m}\sim$ Gumbel$(\mu^{\mathcal{B}_m},1)$.

We first develop the following auxiliary lemmas.

LEMMA F.11. *For the joint event $(\mathcal{A}_m,\mathcal{C}_m,\mathcal{B}_m,\mathcal{S}_{[m]})$, we have*

$$\begin{aligned}\mathbb{P}(\mathcal{A}_m,\mathcal{C}_m,\mathcal{B}_m,\mathcal{S}_{[m]}) = \frac{v_{c_{p_a}}}{V(S_{p_a})}\Big(&\mathbb{P}(\mathcal{A}_m\setminus\{p_a\},\mathcal{C}_m\setminus\{c_{p_a}\},\mathcal{B}_m,\mathcal{S}_{[m]}\setminus S_{p_a}) \\ &-\mathbb{P}(\mathcal{A}_m\setminus\{p_a\},\mathcal{C}_m\setminus\{c_{p_a}\},\mathcal{B}_m\cup\{p_a\},\mathcal{S}_{[m]})\Big).\end{aligned} \tag{F.15}$$

Proof of Lemma F.11:

$$\begin{aligned}\mathbb{P}(\mathcal{A}_m,\mathcal{C}_m,\mathcal{B}_m,\mathcal{S}_{[m]}) &= \mathbb{P}(U_{c_{p_a}}\geq\max\{U_{-c_{p_a}},U_0\},\dots,U_{c_{p_1}}\geq\max\{U_{-c_{p_1}},U_0\},U_0\geq U^{\mathcal{B}_m}) \\ &= \int_{-\infty}^{\infty}\mathbb{P}(U_{c_{p_a}}\geq\max\{U_{-c_{p_a}},c\},\dots,U_{c_{p_1}}\geq\max\{U_{-c_{p_1}},c\},c\geq U^{\mathcal{B}_m}|U_0=c)f_{\mu_0}(c)dc \\ &= \int_{-\infty}^{\infty}\prod_{i=1}^{a}\mathbb{P}(U_{c_{p_i}}\geq\max\{U_{-c_{p_i}},c\})\cdot\mathbb{P}(U^{\mathcal{B}_m}\leq c)f_{\mu_0}(c)dc\end{aligned} \tag{F.16}$$

From Lemma F.3, the term $\mathbb{P}(U_{c_{p_a}}\geq\max\{U_{-c_{p_a}},c\})$ can be simplified as follows:

$$\begin{aligned}\mathbb{P}(U_{c_{p_a}}\geq\max\{U_{-c_{p_a}},c\}) &= \frac{e^{\mu_{c_{p_a}}}}{e^{\mu_{c_{p_a}}}+e^{\mu_{-c_{p_a}}}}\cdot\left[1-e^{-e^{-(c-\mu_{c_{p_a}})}}\cdot e^{-e^{-(c-\mu_{-c_{p_a}})}}\right] \\ &= \frac{v_{c_{p_a}}}{V(S_{p_a})}\cdot\left[1-e^{-e^{-(c-\mu_{c_{p_a}})}}\cdot e^{-e^{-(c-\mu_{-c_{p_a}})}}\right].\end{aligned} \tag{F.17}$$

Thus we have:

$$\begin{aligned}(\text{F.16}) &= \frac{v_{c_{p_a}}}{V(S_{p_a})}\cdot\int_{-\infty}^{\infty}\prod_{i=1}^{a-1}\mathbb{P}(U_{c_{p_i}}\geq\max\{U_{-c_{p_i}},c\})\cdot\mathbb{P}(U^{\mathcal{B}_m}\leq c)f_{\mu_0}(c)\left[1-e^{-e^{-(c-\mu_{c_{p_a}})}}\cdot e^{-e^{-(c-\mu_{-c_{p_a}})}}\right]dc \\ &= \frac{v_{c_{p_a}}}{V(S_{p_a})}\cdot\Big[\mathbb{P}(\mathcal{A}\setminus\{p_a\},\mathcal{P}\setminus\{c_{p_a}\},\mathcal{B}_m,\mathcal{S}_{[m]}\setminus S_{p_a})- \\ &\int_{-\infty}^{\infty}\left[\prod_{i=1}^{a-1}\mathbb{P}(U_{c_{p_i}}\geq\max\{U_{-c_{p_i}},c\})\cdot\mathbb{P}(U^{\mathcal{B}_m}\leq c)\right]e^{-e^{-(c-\mu_{c_{p_a}})}}\cdot e^{-e^{-(c-\mu_{-c_{p_a}})}}f_{\mu_0}(c)dc\Big].\end{aligned} \tag{F.18}$$

Since $e^{-e^{-(c-\mu_{c_{p_a}})}} \cdot e^{-e^{-(c-\mu_{-c_{p_a}})}} = e^{-e^{-(c-\ln(e^{\mu_{c_{p_a}}}+e^{\mu_{-c_{p_a}}}))}}$, we can explain this term as $\mathbb{P}(\nu_{c_{p_a}} \le c)$ based on Lemma F.1, where $\nu_{c_{p_a}} \sim \text{Gumbel}(\log(V(S_{p_a})),1)$. Then the integration in (F.18) is equivalent to

$$\begin{aligned}
&\int_{-\infty}^{\infty}\left[\prod_{i=1}^{a-1}\mathbb{P}(U_{c_{p_i}} \ge \max\{U_{-c_{p_i}},c\})\cdot \mathbb{P}(U^{\mathcal{B}_m}\le c)\right]\mathbb{P}(\nu_{c_{p_a}}\le c)f_{\mu_0}(c)dc \\
&=\int_{-\infty}^{\infty}\left[\prod_{i=1}^{a-1}\mathbb{P}(U_{c_{p_i}} \ge \max\{U_{-c_{p_i}},c\})\right]\cdot \mathbb{P}(\max\{U^{\mathcal{B}_m},\nu_{c_{p_a}}\}\le c)f_{\mu_0}(c)dc \qquad \text{(F.19)}
\end{aligned}$$

Since $\max\{U^{\mathcal{B}_m},\nu_{c_{p_a}}\} \sim \text{Gumbel}(\log\left(\sum_{j\in\mathcal{B}_m}V(S_j)+V(S_{p_a})\right),1)$, we have $\mathbb{P}(\max\{U^{\mathcal{B}_m},\nu_{c_{p_a}}\} \le c) = \mathbb{P}(U^{\mathcal{B}_m\cup\{p_a\}}\le c)$, thus we have

$$\begin{aligned}
\text{(F.19)} &= \int_{-\infty}^{\infty}\left[\prod_{i=1}^{a-1}\mathbb{P}(U_{c_{p_i}} \ge \max\{U_{-c_{p_i}},c\})\right]\cdot \mathbb{P}(U^{\mathcal{B}_m\cup\{p_a\}}\le c)f_{\mu_0}(c)dc \qquad \text{(F.20)}\\
&= \mathbb{P}(\mathcal{A}_m\setminus\{p_a\},\mathcal{C}_m\setminus\{c_{p_a}\},\mathcal{B}_m\cup\{p_a\},\mathcal{S}_{[m]}). \qquad \text{(F.21)}
\end{aligned}$$

Combined with (F.21) and (F.18), we finish the proof. □

Now we show the proof of Proposition 4.

Proof of Proposition 4: We prove (4) by induction on the number of purchased periods $a=|\mathcal{A}_m|$. Recall that for any set of periods $\mathcal{B}$, we use the notation $V(\mathcal{B})=\sum_{q\in\mathcal{B}}V(S_q)$.

When $a=1$, write $\mathcal{A}_m=\{p_1\}$. By Theorem 2.1 in Gao et al. (2021), we have

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_m,\mathcal{C}_m,\mathcal{B}_m,\mathcal{S}_{[m]}) &= \frac{v_{c_{p_1}}}{(1+V(\mathcal{B}_m)+V(S_{p_1}))(1+V(\mathcal{B}_m))}\\
&= \frac{v_{c_{p_1}}}{V(S_{p_1})}\left[\frac{1}{1+V(\mathcal{B}_m)}-\frac{1}{1+V(\mathcal{B}_m)+V(S_{p_1})}\right]. \qquad \text{(F.22)}
\end{aligned}$$

This is exactly (4), since the only subsets of $\mathcal{A}_m=\{p_1\}$ are $\emptyset$ and $\{p_1\}$.

Now suppose (4) holds for all histories with $a-1$ purchased periods. Consider a history with a purchased periods. Fix one purchased period $p_a\in\mathcal{A}_m$, and define

$$\mathcal{A}'=\mathcal{A}_m\setminus\{p_a\},\qquad \mathcal{C}'=\{c_p : p\in\mathcal{A}'\}.$$

Let $\mathcal{S}'$ denote the assortment history obtained from $\mathcal{S}_{[m]}$ by removing period p_a. By Lemma F.11,

$$\mathbb{P}(\mathcal{A}_m,\mathcal{C}_m,\mathcal{B}_m,\mathcal{S}_{[m]}) = \frac{v_{c_{p_a}}}{V(S_{p_a})}\left[\mathbb{P}(\mathcal{A}',\mathcal{C}',\mathcal{B}_m,\mathcal{S}')-\mathbb{P}(\mathcal{A}',\mathcal{C}',\mathcal{B}_m\cup\{p_a\},\mathcal{S}_{[m]})\right]. \qquad \text{(F.23)}$$

By the induction hypothesis,

$$\mathbb{P}(\mathcal{A}',\mathcal{C}',\mathcal{B}_m,\mathcal{S}') = \frac{\prod_{p\in\mathcal{A}'}v_{c_p}}{\prod_{p\in\mathcal{A}'}V(S_p)}\sum_{T\subseteq\mathcal{A}'}(-1)^{|T|}\frac{1}{1+V(\mathcal{B}_m)+\sum_{q\in T}V(S_q)}, \qquad \text{(F.24)}$$

and

$$\mathbb{P}(\mathcal{A}',\mathcal{C}',\mathcal{B}_m\cup\{p_a\},\mathcal{S}_{[m]}) = \frac{\prod_{p\in\mathcal{A}'}v_{c_p}}{\prod_{p\in\mathcal{A}'}V(S_p)}\sum_{T\subseteq\mathcal{A}'}(-1)^{|T|}\frac{1}{1+V(\mathcal{B}_m\cup\{p_a\})+\sum_{q\in T}V(S_q)}. \qquad \text{(F.25)}$$

Since $V(\mathcal{B}_m \cup \{p_a\}) = V(\mathcal{B}_m) + V(S_{p_a})$, substituting (F.24) and (F.25) into (F.23) gives

$$\begin{aligned}
\mathbb{P}(\mathcal{A}_m, \mathcal{C}_m, \mathcal{B}_m, \mathcal{S}_{[m]}) &= \frac{\prod_{p\in\mathcal{A}_m} v_{c_p}}{\prod_{p\in\mathcal{A}_m} V(S_p)} \Bigg[\sum_{T\subseteq\mathcal{A}'} (-1)^{|T|} \frac{1}{1+V(\mathcal{B}_m)+\sum_{q\in T} V(S_q)} \\
&\qquad - \sum_{T\subseteq\mathcal{A}'} (-1)^{|T|} \frac{1}{1+V(\mathcal{B}_m)+V(S_{p_a})+\sum_{q\in T} V(S_q)} \Bigg] \\
&= \frac{\prod_{p\in\mathcal{A}_m} v_{c_p}}{\prod_{p\in\mathcal{A}_m} V(S_p)} \Bigg[\sum_{T\subseteq\mathcal{A}'} (-1)^{|T|} \frac{1}{1+V(\mathcal{B}_m)+\sum_{q\in T} V(S_q)} \\
&\qquad + \sum_{T\subseteq\mathcal{A}'} (-1)^{|T|+1} \frac{1}{1+V(\mathcal{B}_m)+V(S_{p_a})+\sum_{q\in T} V(S_q)} \Bigg]. \qquad \text{(F.26)}
\end{aligned}$$

The first sum in (F.26) collects all subsets of $\mathcal{A}_m$ that do not contain p_a. The second sum collects all subsets of $\mathcal{A}_m$ that contain p_a: for each $T\subseteq\mathcal{A}'$, define $U = T\cup\{p_a\}$. Then

$$|U| = |T|+1, \qquad \sum_{q\in U} V(S_q) = V(S_{p_a}) + \sum_{q\in T} V(S_q).$$

Therefore,

$$\sum_{T\subseteq\mathcal{A}'} (-1)^{|T|} \frac{1}{1+V(\mathcal{B}_m)+\sum_{q\in T} V(S_q)} + \sum_{T\subseteq\mathcal{A}'} (-1)^{|T|+1} \frac{1}{1+V(\mathcal{B}_m)+V(S_{p_a})+\sum_{q\in T} V(S_q)}$$

is exactly

$$\sum_{U\subseteq\mathcal{A}_m} (-1)^{|U|} \frac{1}{1+V(\mathcal{B}_m)+\sum_{q\in U} V(S_q)}.$$

Thus,

$$\mathbb{P}(\mathcal{A}_m, \mathcal{C}_m, \mathcal{B}_m, \mathcal{S}_{[m]}) = \frac{\prod_{p\in\mathcal{A}_m} v_{c_p}}{\prod_{p\in\mathcal{A}_m} V(S_p)} \sum_{U\subseteq\mathcal{A}_m} (-1)^{|U|} \frac{1}{1+V(\mathcal{B}_m)+\sum_{q\in U} V(S_q)}.$$

This proves (4). The induction is complete. $\square$

Proof of Theorem 5. Notice that given the historical event in the first m periods, the seller obtains partial information about the random variable U_0 and makes recommendation decision. However, if the seller knows the realization of U_0, the optimal assortment under such information will lead to higher expected revenue. Formally, finding the optimal assortment given $U_0 = u_0$ in period m requires solving the following problem according to Wang (2022):

$$\max_{S\subseteq\mathcal{N}_m} D(S,u_0) \triangleq \max_{S\subseteq\mathcal{N}_m} \frac{W(S)}{V(S)} \left(1 - e^{-e^{-u_0} V(S)}\right). \qquad \text{(F.27)}$$

We hope to use $R(S_m^{ro})$ to bound $\max_{S\subseteq\mathcal{N}_m} D(S,u_0)$ and thus the optimal adaptive policy. We take the expectation over U_0 and obtain an upper bound for the expected revenue of the adaptive policy in period m: $\mathbb{E}[\max_{S\subset\mathcal{N}_m} D(S,U_0)]$. To find an upper bound, we aim to find $\alpha(u_0)$ satisfying the following inequality:

$$\frac{W(S)}{V(S)} \left(1 - e^{-e^{-u_0} V(S)}\right) \le \frac{W(S)}{1+V(S)} \cdot \alpha(u_0) \le R(S_m^{ro}) \cdot \alpha(u_0). \qquad \text{(F.28)}$$

If (F.28) holds, we have $\mathbb{E}[D(S,U_0)] \le \mathbb{E}[\alpha(U_0)] \cdot R(S_m^{ro})$ for all S.

To guarantee (F.28) to hold for any $V(S)$, we calculate $\alpha(u_0)$ using the following optimization problem:

$$\max_{x\ge 0} f(x,u) = \frac{1+x}{x} \left(1 - e^{-ux}\right), \qquad \text{(F.29)}$$

where $u = e^{-u_0}$. Let $z = ux$, we transform $f(x,u)$ to $f(z,u) = (1+\frac{u}{z})\cdot(1-e^{-z})$ and $\frac{\partial f}{\partial z} = \frac{\left(z^2+zu-ue^z+u\right)e^{-z}}{z^2}$. Let $f_{\mu_0}(\cdot)$ be the probability measure of U_0. We consider two cases.

Case 1: $u \geq 2$. In this case, we have $\frac{\partial f}{\partial z} \leq 0$ for $z \geq 0$. Thus $\max_{z\geq 0} f(z,u) = \lim_{z\to 0} f(z,u) = u$. Then we have

$$\mathbb{E}[\alpha(u_0)\mathbb{1}(e^{-u_0}\geq 2)] = \int_{-\infty}^{-\log 2} f_{\mu_0}(u_0)e^{-u_0}du_0 = \int_{-\infty}^{-\log 2} e^{-e^{-u_0}}e^{-u_0}e^{-u_0}du_0 = \int_2^{\infty} te^{-t}dt = 3e^{-2},$$

where $\mathbb{1}(\cdot)$ is the indicator function.

Case 2: $u < 2$. In this case, the optimal solution satisfies $\frac{\partial f}{\partial z} = 0$. we have $z^2 = u\cdot(e^z - z - 1)$, that is $u = \frac{z^2}{e^z-z-1}$. Denote this function as $h(z)$. Since $h(z)$ is a monotonically decreasing function and the range is $[0,2]$, we know the inverse function is well-defined. Let $z(u)$ satisfy $u = h(z(u))$, then

$$\max_{z\geq 0} f(z,u) = f(z(u),h(z(u))) = \left(1+\frac{z(u)}{e^{z(u)}-z(u)-1}\right)\cdot(1-e^{-z(u)}) = \frac{e^{z(u)}(1-e^{-z(u)})^2}{e^{z(u)}-z(u)-1} \triangleq K(z(u)).$$

And we have

$$h'(z) = \frac{z^2\cdot(1-e^z)}{(-z+e^z-1)^2} + \frac{2z}{-z+e^z-1}$$

$$\begin{aligned}
\mathbb{E}[\alpha(u_0)\mathbb{1}(e^{-u_0}<2)] &= \int_{-\log 2}^{\infty} f_{\mu_0}(u_0)K(z(e^{-u_0}))du_0 \\
&= \int_{-\log 2}^{\infty} e^{-e^{-u_0}}e^{-u_0}K(z(e^{-u_0}))du_0 \qquad \text{(Let } u = e^{-u_0}\text{)} \\
&= \int_0^2 e^{-u}K(z(u))du \\
&= \int_0^2 e^{-u}K(z(u))dh(z(u)) \\
&= \int_{\infty}^0 e^{-\frac{z(u)^2}{e^{z(u)}-z(u)-1}}\left(1+\frac{z(u)}{e^{z(u)}-z(u)-1}\right)\cdot(1-e^{-z(u)})\cdot \\
&\quad \left(\frac{z(u)^2\cdot\left(1-e^{z(u)}\right)}{\left(-z(u)+e^{z(u)}-1\right)^2} + \frac{2z(u)}{-z(u)+e^{z(u)}-1}\right)dz(u) \\
&\approx 1.05421.
\end{aligned}$$

Thus we have

$$\mathbb{E}[\alpha(u_0)] = \mathbb{E}[\alpha(u_0)\mathbb{1}(e^{-u_0}\geq 2)] + \mathbb{E}[\alpha(u_0)\mathbb{1}(e^{-u_0}<2)] \approx 3e^{-2} + 1.05421 \approx 1.4603.$$

We conclude that the expected revenue in each period m is upper bounded by $1.4603\cdot R(S_m^{ro})$.

Next we construct an example such that the ratio $\frac{\mathscr{R}_A}{\mathscr{R}_F} \geq 1+\frac{1}{e}$. We assume the seller only has one product in the first $M-1$ periods, and the preference weight for these products is 1 and the revenue is 0. Since the value of U_0 has been realized before the selling horizon and is unchanged, then in each period $i \leq M-1$, the customer will make purchase with probability $\mathbb{P}(U_i \geq u_0)$. Assume the customer makes purchase in $M_1(M)$ periods among the first $M-1$ periods, then we have

$$\lim_{M\to\infty} \frac{M_1(M)}{M-1} = \mathbb{P}(U \geq u_0).$$

By inverting the CDF of the Gumbel distribution, the seller will eventually know the value of u_0. In the last period, suppose there are two products. Their utilities and revenues are $v_1 = \epsilon,\ r_1 = \frac{1}{\epsilon},\ v_2 = \frac{1}{\epsilon},\ r_2 = 1-\epsilon$. For the optimal fixed policy, only product 1 will be provided and the expected revenue is $r_1v_1/(1+v_1) = \frac{1}{1+\epsilon}$. We consider the following adaptive policy:

- If $u_0 \le 0$, we provide $S_1 = \{1\}$ and the conditional expected revenue is $\frac{1}{\epsilon}(1-e^{-e^{-u_0}\epsilon})$.
- If $u_0 > 0$, we provide $S_2 = \{1,2\}$ and the conditional expected revenue is $D(S_2,u_0) = \frac{1}{1+\epsilon^2}\left(1-e^{-e^{-u_0}(\epsilon+\frac{1}{\epsilon})}\right)$.

When ϵ goes to zero, we have

$$\lim_{\epsilon\to 0} D(S_1,u_0) = e^{-u_0},\ \lim_{\epsilon\to 0} D(S_2,u_0) = 1.$$

In this case, the expected revenue obtained from such an adaptive policy is

$$\int_{-\infty}^{0} f_{\mu_0}(u_0)e^{-u_0}du_0 + \int_0^{\infty} f_{\mu_0}(u_0)du_0 = \int_{-\infty}^{0} e^{-2u_0}e^{-e^{-u_0}}du_0 + 1 - \frac{1}{e} = 1 + \frac{1}{e}.$$

Since the expected revenue from the fixed policy is 1, this completes the proof. □

Proof of Proposition 5. We first analyze the conditional expected revenue in period $m+1$. Assume under k-th trajectory, the purchase behavior in the first m periods is denoted by $(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k)$. Then the conditional probability of purchasing product $j \in S_k(m+1)$ is $\mathbb{P}(j, S_k(m+1)|\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]})$.

Under the assumption $v_i \ge \frac{1}{s-1}$, we show the conditional choice probability is less than $s \cdot \frac{v_j}{1+V(S_k(m+1))}$. The difference between $\mathbb{P}(j, S_k(m+1)|\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]})$ and $\frac{sv_j}{1+V(S_k(m+1))}$ are as follows:

$$\begin{aligned}
&\frac{\mathbb{P}(\mathcal{A}_m^k \cup \{m+1\}, \mathcal{C}_m^k \cup \{j\}, \mathcal{B}_m^k, \mathcal{S}_{[m]} \cup \{S_k(m+1)\})}{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]})} - \frac{sv_j}{1+V(S_k(m+1))}\\
&= \frac{v_j}{V(S_k(m+1))}\frac{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]}) - \mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k \cup \{m+1\}, \mathcal{S}_{[m]} \cup \{S_k(m+1)\})}{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]})} - \frac{sv_j}{1+V(S_k(m+1))}\\
&= \frac{v_j}{V(S_k(m+1))} \cdot \left[1 - \frac{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k \cup \{m+1\}, \mathcal{S}_{[m]} \cup \{S_k(m+1)\})}{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]})} - \frac{sV(S_k(m+1))}{1+V(S_k(m+1))}\right]\\
&= \frac{v_j}{V(S_k(m+1))} \cdot \left[\frac{1-(s-1)V(S_k(m+1))}{1+V(S_k(m+1))} - \frac{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k \cup \{m+1\}, \mathcal{S}_{[m]} \cup \{S_k(m+1)\})}{\mathbb{P}(\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]})}\right]\\
&\le 0,
\end{aligned}$$

where the first equality is due to Lemma F.11 and the last inequality is from our assumption. We conclude that $\mathbb{P}(j, S_k(m+1)|\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]}) \le \frac{sv_j}{1+V(S_k(m+1))}$. Then the conditional expected revenue is bounded as follows:

$$R_{m+1}^k(S_k(m+1)) = \sum_{j\in S_k(m+1)} r_j \mathbb{P}(j, S_k(m+1)|\mathcal{A}_m^k, \mathcal{C}_m^k, \mathcal{B}_m^k, \mathcal{S}_{[m]}) \le \sum_{j\in S_k(m+1)} \frac{sr_jv_j}{1+V(S_k(m+1))} = sR(S_k(m+1)) \le sR(S_{m+1}^{ro}).$$

And we establish that

$$\begin{aligned}
\mathscr{R}_A &= R(S_0^*(1)) + \sum_{m=2}^{M}\sum_{k=0}^{2^{m-1}-1} P_m^k R_m^k(S_k^*(m))\\
&\le R(S_1^{ro}) + \sum_{m=2}^{M}\sum_{k=0}^{2^{m-1}-1} P_m^k sR(S_m^{ro})\\
&\le R(S_1^{ro}) + \sum_{m=2}^{M} sR(S_m^{ro})\\
&\le s \cdot \mathscr{R}_F.
\end{aligned}$$

□

Proof of Theorem 6. For simplicity, we assume the products in period m are $\{1,2,\dots,n_m\}$ and the revenue is non-increasing, that is $r_1 \ge r_2 \ge \dots \ge r_{n_m}$. Notice that the customer purchases at most one product, if the seller knows the utility of all products, then the optimal assortment will only contain one product, that is the product with the largest revenue from the product set $\{i \in \mathcal{N}_m | U_i > U_0\}$. Moreover, the expected revenue will be maximized. For such a clairvoyant seller, the expected revenue can be expressed as follows:

$$\begin{aligned}
&r_1 \cdot \mathbb{P}(U_1 \ge U_0) + \sum_{i=2}^{n_m} r_i \cdot \mathbb{P}(U_i \ge U_0, \max_{j<i} U_j \le U_0) \\
&= \frac{r_1 v_1}{1+v_1} + \sum_{i=2}^{n_m} \frac{r_i v_i}{(1+\sum_{j=1}^{i-1} v_j)(1+\sum_{j=1}^{i} v_j)}. \qquad \text{(F.30)}
\end{aligned}$$

The equation can be derived from Lemma F.7. Let $S^i = \{1,2,\dots,i\}$ and $V(S^0)=0$, we now give an upper bound for (F.30).

$$\begin{aligned}
\text{(F.30)} &= \frac{W(S^1)}{1+V(S^1)} + \sum_{i=2}^{n_m} \frac{W(S^i)-W(S^{i-1})}{(1+V(S^{i-1}))(1+V(S^i))} \\
&= \frac{W(S^1)}{1+V(S^1)} + \sum_{i=2}^{n_m} \frac{W(S^i)}{(1+V(S^{i-1}))(1+V(S^i))} - \sum_{i=2}^{n_m} \frac{W(S^{i-1})}{(1+V(S^{i-1}))(1+V(S^i))} \\
&= \frac{W(S^1)}{1+V(S^1)} + \sum_{i=2}^{n_m} \frac{W(S^i)}{(1+V(S^{i-1}))(1+V(S^i))} - \sum_{i=1}^{n_m-1} \frac{W(S^i)}{(1+V(S^i))(1+V(S^{i+1}))} \\
&= \frac{W(S^1)}{1+V(S^1)} \cdot \left(1 - \frac{1}{1+V(S^2)}\right) + \sum_{i=2}^{n_m-1} \frac{W(S^i)}{1+V(S^i)} \cdot \left(\frac{1}{1+V(S^{i-1})} - \frac{1}{1+V(S^{i+1})}\right) + \frac{W(S^{n_m})}{(1+V(S^{n_m-1}))(1+V(S^{n_m}))} \\
&= \sum_{i=1}^{n_m-1} \frac{W(S^i)}{1+V(S^i)} \cdot \left(\frac{1}{1+V(S^{i-1})} - \frac{1}{1+V(S^{i+1})}\right) + \frac{W(S^{n_m})}{(1+V(S^{n_m-1}))(1+V(S^{n_m}))} \\
&\le \max_{S \subseteq \mathcal{N}_m} \frac{W(S)}{1+V(S)} \cdot \left(1 + \sum_{i=2}^{n_m} \frac{1}{1+V(S^{i-1})} - \frac{1}{1+V(S^i)}\right) \\
&= R(S_m^{ro}) \cdot \left(1 + \frac{1}{1+V(S^1)} - \frac{1}{1+V(S^{n_m})}\right) \\
&\le 2R(S_m^{ro})
\end{aligned}$$

We conclude that the expected revenue obtained from any adaptive policy is upper bounded $2R(S_m^{ro})$. Then we establish that

$$\mathscr{R}_A \le R(S_1^{ro}) + \sum_{m=2}^{M} 2R(S_m^{ro}) \le 2\mathscr{R}_F.$$

□

Proof of Lemma 6. We use the same notations as in proving Theorem 1 and denote $U^c(i,j) = \max_{\ell \in (S_1 \cap S_2) \setminus \{i,j\}} U_\ell$. The conditional probability are calculated in four cases separately.

Case 1: $j = i$. In this case, the conditional probability can be rewritten as follows.

$$\begin{aligned}
\mathbb{P}(U_i - d \ge \max\{U_2^t, U_0, U^c(i)\} | U_i \ge \max\{U_1^t, U_0, U^c(i)\}) &= \frac{\mathbb{P}(U_i \ge \max\{U_1^t, U_0+d, U^c(i)+d, U_2^t+d\})}{\mathbb{P}(U_i \ge \max\{U_1^t, U_0, U^c(i)\})} \\
&= \frac{v_i}{e^d(1+V(S_2 \setminus \{i\})) + v_i + V(S_1 \setminus S_2)} \cdot \frac{1+V(S_1)}{v_i} \\
&= \frac{1+V(S_1)}{e^d(1+V(S_2 \setminus \{i\})) + v_i + V(S_1 \setminus S_2)} \\
&= \frac{e^{-d}(1+V(S_1))}{1+V(S_2 \setminus \{i\}) + e^{-d}(v_i + V(S_1 \setminus S_2))}.
\end{aligned}$$

Case 2: $j \in S_2 \setminus S_1$. In this case, the conditional probability can be rewritten as follows.

$$\begin{aligned}&\mathbb{P}(U_j \geq \max\{U_2^t(j), U_0, U^c(i), U_i - d\} | U_i \geq \max\{U_1^t, U_0, U^c(i)\}) \\ =& \frac{\mathbb{P}(U_j \geq \max\{U_2^t(j), U_0, U^c(i), U_i - d\}, U_i \geq \max\{U_1^t, U_0, U^c(i)\})}{\mathbb{P}(U_i \geq \max\{U_1^t, U_0, U^c(i)\})}.\end{aligned}$$

According to Lemma F.6, the joint probability $\mathbb{P}(U_j \geq \max\{U_2^t(j), U_0, U^c(i), U_i - d\}, U_i \geq \max\{U_1^t, U_0, U^c(i)\})$ has the following expression.

$$\frac{v_i}{v_i + V(S_1 \setminus S_2)} \cdot \frac{v_j}{1 + V(S_2 \setminus \{i\}) + e^{-d}(V(S_1 \setminus S_2) + v_i)} - \frac{1 + V(S_1 \cap S_2) - v_i}{(1 + V(S_1))(v_i + V(S_1 \setminus S_2))} \cdot \frac{v_i v_j}{1 + V(S_2) + V(S_1 \setminus S_2)}.$$

Then the conditional probability can be expressed as follows:

$$\begin{aligned}&\frac{v_j}{v_i + V(S_1 \setminus S_2)} \cdot \left(\frac{1 + V(S_1)}{1 + V(S_2 \setminus \{i\}) + e^{-d}(V(S_1 \setminus S_2) + v_i)} - \frac{1 + V(S_1 \cap S_2) - v_i}{1 + V(S_2) + V(S_1 \setminus S_2)} \right) \\ =& \frac{v_j}{v_i + V(S_1 \setminus S_2)} \cdot \frac{(v_i + V(S_1 \setminus S_2)) \cdot [1 + V(S_2) + V(S_1 \setminus S_2) + (1 - e^{-d})(1 + V(S_1 \cap S_2) - v_i)]}{[1 + V(S_2 \setminus \{i\}) + e^{-d}(V(S_1 \setminus S_2) + v_i)] \cdot [1 + V(S_2) + V(S_1 \setminus S_2)]} \\ =& \frac{(1 - e^{-d})(1 + V(S_1 \cap S_2) - v_i) + 1 + V(S_2) + V(S_1 \setminus S_2)}{[1 + V(S_2 \setminus \{i\}) + e^{-d}(V(S_1 \setminus S_2) + v_i)] \cdot [1 + V(S_2) + V(S_1 \setminus S_2)]} \cdot v_j.\end{aligned}$$

Case 3: $j \in S_1 \cap S_2$, $j \neq i$. In this case, the conditional probability can be rewritten as follows.

$$\begin{aligned}&\mathbb{P}(U_j \geq \max\{U_2^t, U_0, U^c(i,j), U_i - d\} | U_i \geq \max\{U_1^t, U_0, U^c(i,j), U_j\}) \\ =& \frac{\mathbb{P}(U_j \geq \max\{U_2^t, U_0, U^c(i,j), U_i - d\}, U_i \geq \max\{U_1^t, U_0, U^c(i,j), U_j\})}{\mathbb{P}(U_i \geq \max\{U_1^t, U_0, U^c(i,j), U_j\})}.\end{aligned}$$

According to Lemma F.5, the joint probability $\mathbb{P}(U_j \geq \max\{U_2^t, U_0, U^c(i,j), U_i - d\}, U_i \geq \max\{U_1^t, U_0, U^c(i,j), U_j\})$ has the following expression.

$$\frac{v_j}{1 + V(S_2) - v_i} \cdot \left[\frac{v_i}{1 + V(S_2) + V(S_1 \setminus S_2)} - \frac{v_i}{e^d(1 + V(S_2) - v_i) + v_i + V(S_1 \setminus S_2)} \right].$$

Then the conditional probability can be expressed as follows:

$$\begin{aligned}&\frac{v_j(1 + V(S_1))}{1 + V(S_2 \setminus \{i\})} \cdot \left[\frac{1}{1 + V(S_2) + V(S_1 \setminus S_2)} - \frac{1}{e^d(1 + V(S_2 \setminus \{i\})) + v_i + V(S_1 \setminus S_2)} \right] \\ =& \frac{(1 - e^{-d})(1 + V(S_1))}{[1 + V(S_2 \setminus \{i\}) + e^{-d}(v_i + V(S_1 \setminus S_2))] \cdot [1 + V(S_2) + V(S_1 \setminus S_2)]} \cdot v_j.\end{aligned}$$

Case 4: $j = 0$. In this case, the conditional probability can be rewritten as follows.

$$\begin{aligned}&\mathbb{P}(U_0 \geq \max\{U_2^t, U^c(i,j), U_i - d\} | U_i \geq \max\{U_1^t, U_0, U^c(i,j)\}) \\ =& \frac{\mathbb{P}((U_0 \geq \max\{U_2^t, U^c(i,j), U_i - d\}, U_i \geq \max\{U_1^t, U_0, U^c(i,j)\})}{\mathbb{P}(U_i \geq \max\{U_1^t, U_0, U^c(i,j)\})}.\end{aligned}$$

According to Lemma F.5, the joint probability $\mathbb{P}((U_0 \geq \max\{U_2^t, U^c(i,j), U_i - d\}, U_i \geq \max\{U_1^t, U_0, U^c(i,j)\})$ has the following expression.

$$\frac{1}{1 + V(S_2) - v_i} \cdot \left[\frac{v_i}{1 + V(S_2) + V(S_1 \setminus S_2)} - \frac{v_i}{e^d(1 + V(S_2) - v_i) + v_i + V(S_1 \setminus S_2)} \right].$$

Then the conditional probability can be expressed as follows:

$$\begin{aligned}&\frac{(1 + V(S_1))}{1 + V(S_2 \setminus \{i\})} \cdot \left[\frac{1}{1 + V(S_2) + V(S_1 \setminus S_2)} - \frac{1}{e^d(1 + V(S_2 \setminus \{i\})) + v_i + V(S_1 \setminus S_2)} \right] \\ =& \frac{(1 - e^{-d})(1 + V(S_1))}{[1 + V(S_2 \setminus \{i\}) + e^{-d}(v_i + V(S_1 \setminus S_2))] \cdot [1 + V(S_2) + V(S_1 \setminus S_2)]}.\end{aligned}$$

$\square$

Proof of Theorem 7. Denote the optimal conditional assortment in the second period as S_2^* and the corresponding expected revenue as $R(S_2^*|i,S_1)$. We first reveal a structural property for S_2^*, that is the subset of $S_2^* \cap S_1$ is a revenue-ordered assortment. Consider adding a new product $k \in S_1 \setminus S_2$ into S_2 and denote the new assortment as S_2^k, we then analyze the revenue difference between S_2 and S_2^k.

For ease of exposition, we introduce the following notations

- $D(S_2) = 1 + V(S_2 \setminus \{i\}) + e^{-d}(v_i + V(S_1 \setminus S_2))$
- $Z(S_2) = 1 + V(S_2) + V(S_1 \setminus S_2)$
- $N(j|S_2) = \mathbb{P}_d^2(j|i,S_1,S_2) \cdot D(S_2)$

Then we have

$$
\begin{aligned}
D(S_2^k) &= 1 + V(S_2^k \setminus \{i\}) + e^{-d}(v_i + V(S_1 \setminus S_2^k)) \\
&= 1 + V((S_2 \setminus \{i\}) \cup \{k\}) + e^{-d}(v_i + V((S_1 \setminus S_2) \setminus \{k\})) \\
&= 1 + (V(S_2 \setminus \{i\}) + v_k) + e^{-d}(v_i + V(S_1 \setminus S_2) - v_k) \\
&= \left[1 + V(S_2 \setminus \{i\}) + e^{-d}(v_i + V(S_1 \setminus S_2))\right] + v_k - e^{-d} v_k \\
&= D(S_2) + v_k(1 - e^{-d}).
\end{aligned}
$$

The term $N(j|S_2)$ for the choice probabilities change as follows:

- For $j = i$: $N(i|S_2^k) = N(i|S_2)$
- For $j \in S_2 \setminus S_1$: $N(j|S_2^k) = N(j|S_2) + \frac{v_k(1-e^{-d})}{Z(S_2)} v_j$
- For $j \in (S_1 \cap S_2) \setminus \{i\}$: $N(j|S_2^k) = N(j|S_2)$
- For the new product k: $N(k|S_2) = 0$, while $N(k|S_2^k) = \frac{(1-e^{-d})(1+V(S_1))v_k}{Z(S_2)}$.

Let $\mathcal{N}(S_2) = \sum_{j \in S_2} r_j N(j|S_2)$, then the revenue difference can be expressed as

$$
\begin{aligned}
\Delta R &= R(S_2^k|i,S_1) - R(S_2|i,S_1) \\
&= \frac{\mathcal{N}(S_2^k)}{D(S_2^k)} - \frac{\mathcal{N}(S_2)}{D(S_2)} \\
&= \frac{D(S_2)\Delta\mathcal{N} - \Delta D \mathcal{N}(S_2)}{D(S_2) D(S_2^k)} \\
&= \frac{1}{D(S_2^k)} \cdot (\Delta\mathcal{N} - \Delta D \cdot R(S_2|i,S_1)).
\end{aligned}
$$

Notice that $\Delta\mathcal{N}$ is

$$
\begin{aligned}
\Delta\mathcal{N} &= \sum_{j \in S_2^k} r_j (N(j|S_2^k) - N(j|S_2)) \\
&= r_k(N(k|S_2^k) - 0) + \sum_{j \in S_2 \setminus S_1} r_j (N(j|S_2^k) - N(j|S_2)) \\
&= r_k \cdot \frac{(1-e^{-d})(1+V(S_1))v_k}{Z(S_2)} + \sum_{j \in S_2 \setminus S_1} r_j \cdot \frac{v_k(1-e^{-d})}{Z(S_2)} v_j \\
&= \frac{v_k(1-e^{-d})}{Z(S_2)} \left(r_k(1+V(S_1)) + \sum_{j \in S_2 \setminus S_1} r_j v_j \right) \\
&= \frac{v_k(1-e^{-d})(1+V(S_1))}{Z(S_2)} \left(r_k + \frac{W(S_2 \setminus S_1)}{1+V(S_1)} \right).
\end{aligned}
$$

Then we can further simplify the expression of ΔR as follows:

$$\Delta R = \frac{v_k(1-e^{-d})}{D(S_2^k)} \cdot \frac{1+V(S_1)}{Z(S_2)} \cdot \left(r_k + \frac{W(S_2 \setminus S_1)}{1+V(S_1)} - \frac{Z(S_2)}{1+V(S_1)} \cdot R(S_2|i,S_1) \right).$$

Thus as long as the revenue of product k is larger than $\frac{Z(S_2)}{1+V(S_1)} \cdot R(S_2|i,S_1) - \frac{W(S_2\setminus S_1)}{1+V(S_1)}$, adding product k increases the expected revenue. We conclude that $S_2^* \cap S_1$ is a revenue-ordered assortment.

Now we consider the conditional assortment optimization problem by restricting the common product set $S_u \subseteq \mathcal{N}_2 \cap S_1 \setminus \{i\}$, then we need to find another unique product set $S \subseteq \mathcal{N}_2 \setminus S_1$ by solving the following problem:

$$\max_{S\subseteq \mathcal{N}_2\setminus S_1} \frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)+V(S)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{(\bar{D}(S_u)+V(S))\cdot(\bar{Z}+V(S))} + \frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S)}{(\bar{D}(S_u)+V(S))\cdot(\bar{Z}+V(S))} \cdot W(S), \tag{F.31}$$

where $\bar{D}(S_u) = 1+V(S_u)+e^{-d}\big(v_i + V(S_1 \setminus (S_u \cup \{i\}))\big)$ and $\bar{Z} = 1+V(S_1)$. Denote the optimal assortment for Problem (F.31) as S^* and the optimal expected revenue as R^*. Given $\alpha \in (0,1)$, we consider the following two cases.

Case 1: $\frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)+V(S^*)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \geq \alpha R^*$. In this case, if we choose S as the empty set, the expected revenue can be lower bounded as follows,

$$\frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{\bar{D}(S_u)\cdot\bar{Z}} \geq \frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)+V(S^*)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \geq \alpha R^*.$$

Case 2: $\frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)+V(S^*)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} < \alpha R^*$. In this case, we have

$$\frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)+V(S^*)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \leq \frac{\alpha}{1-\alpha} \cdot \frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S^*)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \cdot W(S^*).$$

We conclude that $R^* \leq \frac{1}{1-\alpha} \cdot \frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S^*)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \cdot W(S^*)$. Moreover, we can upper bound the last part in the expression (F.31) as follows:

$$\begin{aligned}
\frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S^*)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \cdot W(S^*) &= \left(1 + (1-e^{-d}) \cdot \frac{1+V(S_u\setminus\{i\})}{\bar{Z}+V(S^*)}\right) \cdot \frac{W(S^*)}{\bar{D}(S_u)+V(S^*)} \\
&\leq \left(1 + (1-e^{-d}) \cdot \frac{1+V(S_u\setminus\{i\})}{\bar{Z}}\right) \cdot \frac{W(S^*)}{\bar{D}(S_u)+V(S^*)} \\
&= \left(1 + (1-e^{-d}) \cdot \frac{1+V(S_u\setminus\{i\})}{1+V(S_1)}\right) \cdot \frac{W(S^*)}{\bar{D}(S_u)+V(S^*)} \\
&\leq \left(1 + (1-e^{-d}) \cdot (1 - \frac{v_i}{1+V(S_1)})\right) \cdot \frac{W(S^*)}{\bar{D}(S_u)+V(S^*)} \\
&\leq \left(2-e^{-d}\right) \cdot \max_{S\subseteq \mathcal{N}_2\setminus S_1} \frac{W(S)}{\bar{D}(S_u)+V(S)}.
\end{aligned}$$

Denote $S_u^{ro} = \arg\max_{S\subseteq \mathcal{N}_2\setminus S_1} \frac{W(S)}{\bar{D}(S_u)+V(S)}$, then the revenue guarantee of S_u^{ro} is lower bounded by

$$\begin{aligned}
&\frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S_u^{ro})}{(\bar{D}(S_u)+V(S_u^{ro}))\cdot(\bar{Z}+V(S_u^{ro}))} \cdot W(S_u^{ro}) \\
&\geq \frac{W(S_u^{ro})}{\bar{D}(S_u)+V(S_u^{ro})} \\
&\geq \frac{1}{2-e^{-d}} \cdot \frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S^*)}{(\bar{D}(S_u)+V(S^*))\cdot(\bar{Z}+V(S^*))} \cdot W(S^*) \\
&\geq \frac{1-\alpha}{2-e^{-d}} R^*.
\end{aligned}$$

Finally, let $\alpha = \frac{1-\alpha}{2-e^{-d}}$, we have $\alpha = \frac{1}{3-e^{-d}}$. Then for the Problem (F.31), providing an empty set and S_u^{ro} in two cases separately could guarantee an approximation ratio $\frac{1}{3-e^{-d}}$.

Lastly, for the generally conditional assortment optimization problem, it can be rewritten as follows:

$$\max_{S_u\subseteq\mathcal{N}_2\cap S_1\setminus\{i\}, S_u \text{is RO}} \max_{S\subseteq\mathcal{N}_2\setminus S_1} \frac{e^{-d}(1+V(S_1))r_i}{\bar{D}(S_u)+V(S)} + \frac{(1-e^{-d})(1+V(S_1))W(S_u)}{(\bar{D}(S_u)+V(S))\cdot(\bar{Z}+V(S))} + \frac{(1-e^{-d})(1+V(S_u\setminus\{i\}))+\bar{Z}+V(S)}{(\bar{D}(S_u)+V(S))\cdot(\bar{Z}+V(S))}\cdot W(S).$$

For each inner problem, we can find an assortment with approximation ratio $\frac{1}{3-e^{-d}}$. Since the total possible number of assortments S_u is $|S_1|$, we can find an assortment with approximation ratio $\frac{1}{3-e^{-d}}$ in polynomial time. □

Proof of Theorem 8. We first build an upper bound for the expected revenue under the optimal adaptive policy. Let $\mathbb{P}_A(i)$ be the expected number of product i that is purchased under the optimal adaptive policy. Since each product can be purchased at most once in each period, $\mathbb{P}_A(i)$ is maximized when product i is provided solely in each period. We have

$$\mathbb{P}_A(i) \le \frac{v_i}{1+v_i} + \frac{e^{-d}v_i}{1+e^{-d}v_i}. \tag{F.32}$$

We denote the right-hand part as p_i. Moreover, since at most two products are purchased, the expected total number of purchased products is upper bounded by 2, that is

$$\sum_{i\in\mathcal{N}} \mathbb{P}_A(i) \le 2. \tag{F.33}$$

Then, for any positive value τ, we have

$$\begin{aligned}
\mathscr{R}_A^* &= \sum_{i\in\mathcal{N}} \mathbb{P}_A(i) r_i \\
&= \tau \sum_{i\in\mathcal{N}} \mathbb{P}_A(i) + \sum_{i\in\mathcal{N}} \mathbb{P}_A(i)(r_i-\tau) \\
&\le 2\tau + \sum_{i\in\mathcal{N}} \mathbb{P}_A(i)(r_i-\tau)^+ \\
&\le 2\tau + \sum_{i\in\mathcal{N}} p_i (r_i-\tau)^+.
\end{aligned} \tag{F.34}$$

Where $(\cdot)^+ = \max\{\cdot, 0\}$. The first inequality is from the inequality (F.33) and $r_i - \tau \le (r_i-\tau)^+$. The second inequality is due to the upper bound in (F.32).

Since (F.34) is satisfied for all positive value τ, we have

$$\mathscr{R}_A^* \le \min_{\tau>0} 2\tau + \sum_{i\in\mathcal{N}} p_i (r_i-\tau)^+. \tag{F.35}$$

Then we further build a lower bound for expected revenue under the restricted fixed policy (S,S). Let $\mathbb{P}_F(i|S)$ denote the expected number of product i that is purchased under the fixed policy $\mathscr{R}_F(S,S)$. From (22), we have

$$\mathbb{P}_F(i|S) = \left[\sum_{j\in S\setminus\{i\}} \frac{(1-e^{-d})v_j}{(1+V(S)-(1-e^{-d})v_j)(1+V(S))} + \frac{e^{-d}}{1+V(S)-(1-e^{-d})v_i} + \frac{1}{1+V(S)}\right]\cdot v_i \tag{F.36}$$

Moreover, we use $\mathbb{P}_F^k(S)$ represent the probability that the customer will purchase at most k products under the fixed policy (S,S). The probability $\mathbb{P}_F^1(S)$ can be presented as follows:

$$\mathbb{P}_F^1(S) = \frac{1}{1+V(S)} + \sum_{i\in S} \frac{v_i}{1+V(S)} \cdot \frac{1-e^{-d}}{1+V(S)-(1-e^{-d})v_i},$$

where the first part represents the probability of no purchase, and each term in the second part represents the probability that the customer purchases product i in the first period and nothing in the second period. Denote $S\setminus\{i\}$ as S_{-i}, we then show $\mathbb{P}_F(i|S)$ is lower bounded by $\frac{p_i}{1+e^{-d}}\mathbb{P}_F^1(S)$ in two steps.

Step 1: We first show $\mathbb{P}_F(i|S)\geq\alpha\cdot\frac{v_i}{1+v_i}\mathbb{P}_F^1(S)$ for any $\alpha\leq\min\{1+v_i,1+e^{-d}v_i+(1+v_i)e^{-d}\}$. It is equal to show the following expression is larger than 0.

$$
\begin{aligned}
&\frac{1}{1+V(S)}+\frac{e^{-d}}{1+V(S_{-i})+e^{-d}v_i}+\sum_{j\neq i}\frac{v_j(1-e^{-d})}{(1+V(S))(1+V(S_{-j})+e^{-d}v_j)}-\frac{\alpha}{1+v_i}\left[\frac{1}{1+V(S)}+\sum_{j\in S}\frac{v_j}{1+V(S)}\cdot\frac{1-e^{-d}}{1+V(S_{-j})+e^{-d}v_j}\right]\\
=&\frac{1}{(1+v_i)(1+V(S))}\cdot\left[1+v_i+\frac{(e^{-d}(1+v_i)-\alpha)(1+V(S))}{1+V(S_{-i})+e^{-d}v_i}+\sum_{j\neq i}\frac{(1+v_i-\alpha)(1-e^{-d})v_j}{1+V(S_{-j})+e^{-d}v_j}\right].
\end{aligned}
$$

Since $\alpha\leq 1+v_i$, a sufficient condition is to show that $1+v_i+\frac{(e^{-d}(1+v_i)-\alpha)(1+V(S))}{1+V(S_{-i})+e^{-d}v_i}$ is larger than 0.

$$
\begin{aligned}
1+v_i+\frac{(e^{-d}(1+v_i)-\alpha)(1+V(S))}{1+V(S_{-i})+e^{-d}v_i}&=\frac{1+V(S)}{1+V(S_{-i})+e^{-d}v_i}\cdot\left[\frac{(1+v_i)(1+V(S_{-i})+e^{-d}v_i)}{1+V(S)}+e^{-d}(1+v_i)-\alpha\right]\\
&=\frac{1+V(S)}{1+V(S_{-i})+e^{-d}v_i}\cdot\left[(1+v_i)\cdot\left(1-\frac{(1-e^{-d})v_i}{1+V(S)}\right)+e^{-d}(1+v_i)-\alpha\right]\\
&\geq\frac{1+V(S)}{1+V(S_{-i})+e^{-d}v_i}\cdot\left[(1+v_i)\cdot\left(1-\frac{(1-e^{-d})v_i}{1+v_i}\right)+e^{-d}(1+v_i)-\alpha\right]\\
&=\frac{1+V(S)}{1+V(S_{-i})+e^{-d}v_i}\cdot\left[1+v_i-(1-e^{-d})v_i+e^{-d}(1+v_i)-\alpha\right]\\
&=\frac{1+V(S)}{1+V(S_{-i})+e^{-d}v_i}\cdot\left[1+e^{-d}v_i+e^{-d}(1+v_i)-\alpha\right]\\
&\geq 0.
\end{aligned}
$$

The first inequality is due to the fact that $V(S)=\sum_{j\in S}v_j\geq v_i$ and the last inequality is from the definition of α.

Step 2: Notice that $p_i=\frac{v_i}{1+v_i}+\frac{e^{-d}v_i}{1+e^{-d}v_i}=\frac{v_i}{1+v_i}\cdot\frac{1+e^{-d}v_i+(1+v_i)e^{-d}}{1+e^{-d}v_i}$, we only need to show $\frac{1+e^{-d}v_i+(1+v_i)e^{-d}}{1+e^{-d}v_i}\cdot\frac{1}{1+e^{-d}}\leq\min\{(1+v_i),1+e^{-d}v_i+(1+v_i)e^{-d}\}$. Firstly, we have $(1+e^{-d}v_i)(1+e^{-d})>1$, then $\frac{1+e^{-d}v_i+(1+v_i)e^{-d}}{1+e^{-d}v_i}\cdot\frac{1}{1+e^{-d}}\leq 1+e^{-d}v_i+(1+v_i)e^{-d}$. Moreover, since $(1+e^{-d}v_i)(1+e^{-d})(1+v_i)\geq 1+e^{-d}v_i+(1+e^{-d}v_i)(1+v_i)e^{-d}\geq 1+e^{-d}v_i+(1+v_i)e^{-d}$, we have $\frac{1+e^{-d}v_i+(1+v_i)e^{-d}}{1+e^{-d}v_i}\cdot\frac{1}{1+e^{-d}}\leq 1+v_i$. We conclude that $\frac{p_i}{1+e^{-d}}\leq\alpha\cdot\frac{v_i}{1+v_i}$.

Based on the above two steps, we prove that $\mathbb{P}_F(i|S)\geq\alpha\cdot\frac{v_i}{1+v_i}\mathbb{P}_F^1(S)\geq\frac{p_i}{1+e^{-d}}\mathbb{P}_F^1(S)$. Then let S_τ be the revenue-ordered assortment in which the revenue of each product is higher than τ, and we consider τ satisfying the following condition:

$$2\tau=\sum_{i\in S_\tau}\frac{p_i}{1+e^{-d}}(r_i-\tau)^+. \tag{F.37}$$

Finally we could lower bound the expected revenue under the fixed policy (S_τ,S_τ) as follows:

$$
\begin{aligned}
\mathscr{R}_F(S_\tau,S_\tau)&=\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)r_i\\
&=\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)\tau+\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)(r_i-\tau)\\
&=\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)\tau+\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)(r_i-\tau)^+ &\text{(F.38)}\\
&\geq(1-\mathbb{P}_F^1(S_\tau))2\tau+\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)(r_i-\tau)^+ &\text{(F.39)}
\end{aligned}
$$

$$
\begin{aligned}
&\geq (1-\mathbb{P}_F^1(S_\tau))2\tau + \sum_{i\in S_\tau}\frac{p_i}{1+e^{-d}}\mathbb{P}_F^1(S_\tau)(r_i-\tau)^+ &\text{(F.40)}\\
&= 2\tau &\text{(F.41)}\\
&= \frac{2\tau}{2\tau+2(1+e^{-d})\tau}\cdot(2\tau+2(1+e^{-d})\tau)\\
&= \frac{2\tau}{2\tau+2(1+e^{-d})\tau}\cdot(2\tau+\sum_{i\in S_\tau}p_i(r_i-\tau)^+)\\
&\geq \frac{1}{2+e^{-d}}\mathscr{R}_A^*. &\text{(F.42)}
\end{aligned}
$$

In (F.38), we use the definition of the assortment S_τ that $r_i \geq \tau$ for $i \in S_\tau$. In (F.39), notice that $\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)$ can be represented as $\mathbb{P}_F^1(S_\tau)-\mathbb{P}_F^0(S_\tau)+2(1-\mathbb{P}_F^1(S_\tau))$, where $\mathbb{P}_F^1(S_\tau)-\mathbb{P}_F^0(S_\tau)$ and $1-\mathbb{P}_F^1(S_\tau)$ are the probability that the consumer will purchase 1 product and 2 products under the fixed policy (S_τ, S_τ) separately, we have $\sum_{i\in S_\tau}\mathbb{P}_F(i|S_\tau)\geq 2(1-\mathbb{P}_F^1(S_\tau))$. We use the lower bound for $\mathbb{P}_F(i|S_\tau)$ in the inequality (F.40). (F.41) is from the definition of τ in (F.37). The last inequality (F.42) is from the upper bound we build in (F.35).

Finally, we give the sufficient conditions that guarantee the approximation ratio for the revenue-ordered assortment is at least $\frac{1}{2}$. Based on the above analysis, if $\mathbb{P}_F(i|S_\tau)$ is lower bounded by $p_i\mathbb{P}_F^1(S_\tau)$, we could consider τ that satisfying $2\tau = \sum_{i\in\mathcal{N}}p_i(r_i-\tau)^+$ and show that $\mathscr{R}_F(S_\tau,S_\tau)\geq\frac{1}{2}\mathscr{R}_A^*$. After simplifying the expression of $\mathbb{P}_F(i|S_\tau)-p_i\mathbb{P}_F^1(S_\tau)$, it is equivalent to the following condition.

$$
\left(\sum_{j\neq i}\frac{v_j(1-e^{-d})}{1+V(S_{-j})+e^{-d}v_j}+\frac{(1+V(S))e^{-d}}{1+V(S_{-i})+e^{-d}v_i}+1\right)(1+v_i)-\left(1+\sum_{j\in S}\frac{v_j(1-e^{-d})}{1+V(S_{-j})+e^{-d}v_j}\right)\cdot\left(1+\frac{(1+v_i)e^{-d}}{1+e^{-d}v_i}\right)\geq 0,
$$

which can be rewritten as follows,

$$
\frac{e^{-d}v_i(1+v_i)}{1+e^{-d}v_i}\cdot\frac{1+e^{-d}V(S)}{1+V(S_{-i})+e^{-d}v_i}+\frac{V(S_{-i})(1+e^{-d}v_i)+e^{-2d}(1+v_i)^2}{(1+V(S_{-i})+e^{-d}v_i)(1+e^{-d}v_i)}v_i\geq\frac{(1+v_i-v_i^2)e^{-d}-v_i}{1+e^{-d}v_i}\sum_{j\neq i}\frac{(1-e^{-d})v_j}{1+V(S_{-j})+e^{-d}v_j}.
$$

Since the left-hand part is larger than 0, a sufficient condition for $\mathbb{P}_F(i|S_\tau)-p_i\mathbb{P}_F^1(S_\tau)\geq 0$ is that $(1+v_i-v_i^2)e^{-d}-v_i\leq 0$. When $1\leq v_i\leq\frac{\sqrt{5}+1}{2}$, we have $(1+v_i-v_i^2)e^{-d}-v_i\leq 1+v_i-v_i^2-v_i=1-v_i^2\leq 0$. If $v_i\geq\frac{\sqrt{5}+1}{2}$, we have $(1+v_i-v_i^2)e^{-d}-v_i\leq -v_i\leq 0$. And if $e^{-d}\leq\min_{i\in\mathcal{N}}\frac{v_i}{1+v_i-v_i^2}$, this condition is also satisfied.

□

Proof of Proposition C.1. To prove the result, we first introduce the following useful lemma.

LEMMA F.12 **(Property 2 in Bai et al. (2024))**. *Given that j has the highest utility in assortment T, the ranking of all other alternatives in $T\setminus\{j\}$ is independent of this fact and follows the same distribution as if j had never been in the choice set.*

Let $T=S\cup S_1\cup\{0\}$ and define $w_T=\arg\max_{k\in T}U_k$. we can rewrite the joint choice probability $\mathbb{P}_m(i,S,A,B)$ as follows:

$$
\begin{aligned}
&\sum_{j\in T}\Pr(j=w_T)\cdot\Pr\left(\{\text{rank}(i,S)\leq m\}\cap\{U_i>U_0\}\cap\{\underline{U}_A>\max(U_0,\overline{U}_B)\}|j=w_T\right)\\
&=\sum_{j\in T}\frac{v_j}{V(T)}\cdot\Pr\left(\{\text{rank}(i,S)\leq m\}\cap\{U_i>U_0\}\cap\{\underline{U}_A>\max(U_0,\overline{U}_B)\}|j=w_T\right),
\end{aligned}
$$

By definition, w_T must be in A or in $S \setminus S_1$ to satisfy the chain $\{\underline{U}_A > \max(U_0, \overline{U}_B)\}$. Now we consider the following three possible cases.

Case 1: $j = i$**.** This case is possible if $i \in A$ or if $i \in S \setminus S_1$. Given U_i is maximum, the rank condition is satisfied (rank $= 1$) and $U_i > U_0$ holds. The remaining condition becomes $\underline{U}_{A\setminus\{i\}} > \max(U_0, \overline{U}_B)$. Based on Lemma F.12, the choice probability in this case becomes

$$\frac{v_i}{V(T)} \mathcal{Q}(A \setminus \{i\}, B).$$

Case 2: $j \in (S \setminus B) \setminus \{i\}$**.** Given U_j is the maximum, the rank condition becomes $\mathrm{rank}(i, S \setminus \{j\}) \leq m-1$ and the event $\underline{U}_A > \max(U_0, \overline{U}_B)$ reduces to $\underline{U}_{A\setminus\{j\}} > \max(U_0, \overline{U}_B)$. Then choice probability in this case becomes

$$\frac{v_j}{V(T)} \mathbb{P}_{m-1}(i, S \setminus \{j\}, A \setminus \{j\}, B).$$

Case 3: $j \in A \setminus S$**.** Since j is in A but not S, the rank condition is unchanged but A is reduced, the choice probability becomes:

$$\frac{v_j}{V(T)} \mathbb{P}_m(i, S, A \setminus \{j\}, B).$$

The function $\mathcal{Q}(A', B)$ follows a similar logic. For the event $\underline{U}_{A'} > \max(U_0, \overline{U}_B)$ to hold, the maximum of $A' \cup B \cup \{0\}$ must be in A'. We sum over all $k \in A'$ as the potential maximum and recurse until A' is empty. When $A' = \emptyset$, the condition is true with probability 1.

□