Appendix: Assortment Optimization and Pricing under the Multinomial Logit Model with Impatient Customers: Sequential Recommendation and Selection

Appendix A: Maximum of Independent Gumbel Random Variables

Let X and Y be independent Gumbel random variables with location-scale parameters $(\mu, 1)$ and $(\eta, 1)$, respectively. So, the probability density function of X is $f(x) = \exp(-(x-\mu+e^{-(x-\mu)}))$ and the cumulative distribution function of Y is $G(x) = \exp(-e^{-(x-\eta)})$. Thus, we have

$$\begin{aligned}
&\mathbb{P}\big\{\mathbf{1}(X \ge Y) = 1,\ \max\{X,Y\} \ge u\big\} \;=\; \mathbb{P}\big\{X \ge Y,\ \max\{X,Y\} \ge u\big\} \;=\; \mathbb{P}\big\{X \ge Y,\ X \ge u\big\} \\
&\quad = \int_u^\infty \mathbb{P}\big\{Y \le x\big\} \cdot \mathbb{P}\big\{X \in dx\big\} \;=\; \int_u^\infty \exp(-e^{-(x-\eta)})\exp(-(x-\mu+e^{-(x-\mu)}))\,dx \\
&\quad = e^{\mu}\int_u^\infty \exp(-(x+e^{-x}(e^{\mu}+e^{\eta})))\,dx \;=\; e^{\mu}\int_u^\infty \exp(-(x+e^{-(x-\log(e^{\mu}+e^{\eta}))}))\,dx \\
&\quad = \frac{e^{\mu}}{e^{\mu}+e^{\eta}}\int_u^\infty \exp(-(x-\log(e^{\mu}+e^{\eta})+e^{-(x-\log(e^{\mu}+e^{\eta}))}))\,dx \overset{(a)}{=} \mathbb{P}\{X \ge Y\}\cdot\mathbb{P}\{\max\{X,Y\} \ge u\},
\end{aligned}$$

where (a) holds because $\mathbb{P}\{X \ge Y\} = \frac{e^{\mu}}{e^{\mu}+e^{\eta}}$, and $\max\{X,Y\}$ has the Gumbel distribution with location-scale parameters $(\log(e^{\mu}+e^{\eta}), 1)$ by the first and second properties in the proof of Theorem 2.1. Thus, $\mathbb{P}\{\mathbf{1}(X \ge Y) = 1,\ \max\{X,Y\} \ge u\} = \mathbb{P}\{\mathbf{1}(X \ge Y) = 1\}\cdot\mathbb{P}\{\max\{X,Y\} \ge u\}$, as desired.

Appendix B: Proof of Lemma 3.2

In (3), we have one term for each stage. Considering $\Pi(S_1,\ldots,S_{k-1}\cup\{i\},S_k\setminus\{i\},\ldots,S_m)$ and $\Pi(S_1,\ldots,S_m)$, the terms for stages other than $k-1$ and k are identical. Thus, fixing $(S_1,\ldots,S_m)$ and letting $\widehat{W}_k = W(S_k)$, $\widehat{\theta}_k = \sum_{\ell=1}^k V(S_\ell)$ and $\widehat{R}_k = R_k(S_1,\ldots,S_m)$, we have

$$\begin{aligned}
&\Pi(S_1,\ldots,S_{k-1}\cup\{i\},S_k\setminus\{i\},\ldots,S_m) - \Pi(S_1,\ldots,S_m) \\
&\overset{(a)}{=} \frac{\lambda_{k-1}(\widehat{W}_{k-1}+v_i\,r_i)}{(1+\widehat{\theta}_{k-2})(1+\widehat{\theta}_{k-1}+v_i)} + \frac{\lambda_k(\widehat{W}_k-v_i\,r_i)}{(1+\widehat{\theta}_{k-1}+v_i)(1+\widehat{\theta}_k)} - \frac{\lambda_{k-1}\widehat{W}_{k-1}}{(1+\widehat{\theta}_{k-2})(1+\widehat{\theta}_{k-1})} - \frac{\lambda_k\widehat{W}_k}{(1+\widehat{\theta}_{k-1})(1+\widehat{\theta}_k)} \\
&= \frac{v_i\,r_i}{1+\widehat{\theta}_{k-1}+v_i}\left(\frac{\lambda_{k-1}}{1+\widehat{\theta}_{k-2}} - \frac{\lambda_k}{1+\widehat{\theta}_k}\right) + \frac{\lambda_{k-1}\widehat{W}_{k-1}}{1+\widehat{\theta}_{k-2}}\left(\frac{1}{1+\widehat{\theta}_{k-1}+v_i} - \frac{1}{1+\widehat{\theta}_{k-1}}\right) \\
&\qquad + \frac{\lambda_k\widehat{W}_k}{1+\widehat{\theta}_k}\left(\frac{1}{1+\widehat{\theta}_{k-1}+v_i} - \frac{1}{1+\widehat{\theta}_{k-1}}\right),
\end{aligned}$$

where we follow the convention that $\widehat{\theta}_0 = 0$. In the chain of equalities above, (a) uses the fact that $W(S_{k-1}\cup\{i\}) = W(S_{k-1}) + v_i\,r_i$ and $W(S_k\setminus\{i\}) = W(S_k) - v_i\,r_i$, along with $\sum_{\ell=1}^{k-2} V(S_\ell) + V(S_{k-1}\cup\{i\}) = \sum_{\ell=1}^{k-1} V(S_\ell) + v_i$ and $\sum_{\ell=1}^{k-2} V(S_\ell) + V(S_{k-1}\cup\{i\}) + V(S_k\setminus\{i\}) = \sum_{\ell=1}^{k} V(S_\ell)$.

Arranging the terms on the right side of the chain of equalities above, the right side of the chain of equalities above is equivalent to

$$\begin{aligned}
&\frac{v_i\, r_i}{1+\widehat{\theta}_{k-1}+v_i}\left(\frac{\lambda_{k-1}}{1+\widehat{\theta}_{k-2}}-\frac{\lambda_k}{1+\widehat{\theta}_k}\right)-\left(\frac{\lambda_{k-1}\,\widehat{W}_{k-1}}{1+\widehat{\theta}_{k-2}}+\frac{\lambda_k\,\widehat{W}_k}{1+\widehat{\theta}_k}\right)\left(\frac{v_i}{(1+\widehat{\theta}_{k-1}+v_i)\,(1+\widehat{\theta}_{k-1})}\right)\\
&\stackrel{(b)}{=}\ \frac{v_i\, r_i}{1+\widehat{\theta}_{k-1}+v_i}\left(\frac{\lambda_{k-1}}{1+\widehat{\theta}_{k-2}}-\frac{\lambda_k}{1+\widehat{\theta}_k}\right)-\frac{v_i}{1+\widehat{\theta}_{k-1}+v_i}\,(\widehat{R}_{k-1}+\widehat{R}_k)\\
&=\ \frac{\frac{\lambda_{k-1}}{1+\widehat{\theta}_{k-2}}-\frac{\lambda_k}{1+\widehat{\theta}_k}}{1+\widehat{\theta}_{k-1}+v_i}\,v_i\left(r_i-\frac{\widehat{R}_{k-1}+\widehat{R}_k}{\frac{\lambda_{k-1}}{1+\widehat{\theta}_{k-2}}-\frac{\lambda_k}{1+\widehat{\theta}_k}}\right)\ =\ \frac{\frac{\lambda_{k-1}}{1+\widehat{\theta}_{k-2}}-\frac{\lambda_k}{1+\widehat{\theta}_k}}{1+\widehat{\theta}_{k-1}+v_i}\,v_i\,(r_i-t_k(S_1,\ldots,S_m)),
\end{aligned}$$

where (b) uses the definition of $R_k(S_1,\ldots,S_m)$. The two chains of equalities above show that the first identity in the lemma holds. The proofs for the second and third identities are similar. ■

Appendix C: Nonnegative Optimal Prices

In the next lemma, we show that the prices are nonnegative in any optimal solution to the PRICING problem, which allows us to use first-order conditions to characterize the optimal prices.

Lemma C.1 *Letting $\boldsymbol{p}^*$ be an optimal solution to the* PRICING *problem, the prices in the optimal solution satisfies $p_i^* \geq 0$ for all $i \in \mathcal{N}$.*

Proof: Using $\boldsymbol{p}^*$ to denote an optimal solution to the PRICING problem, let $\mathcal{N}^+$ and $\mathcal{N}^-$ be such that $p_i^* \geq 0$ for all $i \in \mathcal{N}^+$ and $p_i^* < 0$ for all $i \in \mathcal{N}^-$. To get a contradiction, assume that $\mathcal{N}^- \neq \varnothing$. Let $\widehat{\boldsymbol{p}}$ be defined as

$$\widehat{p}_i = p_i^* \quad \forall\, i \in \mathcal{N}^+ \ \text{ and } \ \widehat{p}_i = 0 \quad \forall\, i \in \mathcal{N}^-.$$

We claim that the choice probabilities corresponding to the prices $\widehat{\boldsymbol{p}}$ satisfy $\phi_i^k(\widehat{\boldsymbol{p}}) \geq \phi_i^k(\boldsymbol{p}^*)$ for all $i \in S_k \cap \mathcal{N}^+$, $k \in \mathcal{M}$. In particular, by the definition of $\widehat{\boldsymbol{p}}$, we have $\widehat{p}_i \geq p_i^*$ for all $i \in \mathcal{N}$, so $e^{\alpha_i - \beta\, \widehat{p}_i} \leq e^{\alpha_i - \beta\, p_i^*}$ for all $i \in \mathcal{N}$. Thus, we get $V_k(\widehat{\boldsymbol{p}}) \leq V_k(\boldsymbol{p}^*)$ for all $k \in \mathcal{M}$. In this case, since $\widehat{p}_i = p_i^*$ for all $i \in \mathcal{N}^+$, we have $\phi_i^k(\widehat{\boldsymbol{p}}) = \lambda_k \frac{e^{\alpha_i - \beta\, \widehat{p}_i}}{((1+\sum_{\ell=1}^{k-1} V_\ell(\widehat{\boldsymbol{p}}))\,(1+\sum_{\ell=1}^{k} V_\ell(\widehat{\boldsymbol{p}})))} \geq \lambda_k \frac{e^{\alpha_i - \beta\, p_i^*}}{((1+\sum_{\ell=1}^{k-1} V_\ell(\boldsymbol{p}^*))\,(1+\sum_{\ell=1}^{k} V_\ell(\boldsymbol{p}^*)))} = \phi_i^k(\boldsymbol{p}^*)$ for all $i \in S_k \cap \mathcal{N}^+$, $k \in \mathcal{M}$. Thus, the claim holds. Letting $S_k^+ = S_k \cap \mathcal{N}^+$ and $S_k^- = S_k \cap \mathcal{N}^-$ for notational brevity, we have $\widehat{p}_i = p_i^*$ for all $i \in S_k^+$, as well as $\widehat{p}_i = 0 > p_i^*$ for all $i \in S_k^-$. Furthermore, since $\mathcal{N}^- \neq \varnothing$, we have $S_k^- \neq \varnothing$ for some $k \in \mathcal{M}$. In this case, we obtain

$$\begin{aligned}
\Pi(\boldsymbol{p}^*) = \sum_{k\in\mathcal{M}}\sum_{i\in S_k^+} p_i^*\,\phi_i^k(\boldsymbol{p}^*) + \sum_{k\in\mathcal{M}}\sum_{i\in S_k^-} p_i^*\,\phi_i^k(\boldsymbol{p}^*) &\stackrel{(a)}{<} \sum_{k\in\mathcal{M}}\sum_{i\in S_k^+} p_i^*\,\phi_i^k(\boldsymbol{p}^*) \stackrel{(b)}{\leq} \sum_{k\in\mathcal{M}}\sum_{i\in S_k^+} \widehat{p}_i\,\phi_i^k(\widehat{\boldsymbol{p}})\\
&= \sum_{k\in\mathcal{M}}\sum_{i\in S_k^+} \widehat{\boldsymbol{p}}_i\,\phi_i^k(\widehat{\boldsymbol{p}}) + \sum_{k\in\mathcal{M}}\sum_{i\in S_k^-} \widehat{\boldsymbol{p}}_i\,\phi_i^k(\widehat{\boldsymbol{p}}) = \Pi(\widehat{\boldsymbol{p}}),
\end{aligned}$$

where (a) holds since $p_i^* < 0$ for all $i \in S_k^-$ and $S_k^- \neq \varnothing$ for some $k \in \mathcal{M}$, whereas (b) holds since $p_i^* = \widehat{p}_i$ and $\phi_i^k(\boldsymbol{p}^*) \leq \phi_i^k(\widehat{\boldsymbol{p}})$ for all $i \in S_k \cap \mathcal{N}^+$, $k \in \mathcal{M}$. The chain of inequalities above contradicts the fact that $\boldsymbol{p}^*$ is an optimal solution to the PRICING problem. ■

Appendix D: Additive Performance Guarantee for Optimal Prices under Fixed Assortments

In this section, we give a dynamic programming approach to obtain a solution to the PRICING problem with an additive performance guarantee. Letting π^* be the optimal objective value of the PRICING problem, for any $\theta > 0$, our approach comes up with a solution that provides an expected revenue of at least $\pi^* - \theta$ and the number of operations required to obtain this solution is polynomial in $1/\theta$. At the end of this section, we explain that we can easily use a lower bound on the optimal expected revenue to numerically evaluate the multiplicative performance guarantee of the solution that has an additive performance guarantee. To construct a solution with an additive performance guarantee, fixing an integer $K > 0$, we construct the grid points $\mathsf{Grid} = \{\ell/K : \ell = 1, \ldots, K\}$ over the interval $[0,1]$. Noting the expected revenue expression in (5), we use the no-purchase probabilities over different numbers of stages as the decision variables in the PRICING problem. We focus on only the values of the no-purchase probabilities that take values in Grid. Let $\Theta_k(q_{k-1})$ denote the maximum expected revenue that can be obtained from stages $k, k+1, \ldots, m$, given that the no-purchase probability over the first $k-1$ stages is q_{k-1}. In this case, letting $\Theta_{m+1}(\cdot) = 0$ and recalling that $q_0 = 1$, by (5), for all $k \in \mathcal{M}$ and $q_{k-1} \in \mathsf{Grid}$, we have the recursion

$$\Theta_k(q_{k-1}) \;=\; \max_{\substack{q_k \,\in\, \mathsf{Grid}\,:\\ q_{k-1} \geq q_k}} \left\{ \frac{\lambda_k}{\beta}\,(q_{k-1} - q_k) \left\{ \log\left(\sum_{i \in S_k} e^{\alpha_i} \right) - \log\left(\frac{1}{q_k} - \frac{1}{q_{k-1}} \right) \right\} + \Theta_{k+1}(q_k) \right\}. \qquad (14)$$

Thus, $\Theta_1(q_0) = \Theta_1(1)$ corresponds to the largest expected revenue provided by no-purchase probabilities of the form $\boldsymbol{q} = (q_1, \ldots, q_m)$ with $q_k \in \mathsf{Grid}$ and $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$.

In the next theorem, we show that we can use the dynamic program above to come up with a solution to the PRICING problem with an additive performance guarantee.

Theorem D.1 *Letting π^* be the optimal objective value of the* PRICING *problem and $\{\Theta_k(\cdot) : k \in \mathcal{M}\}$ be obtained through* (14) *with $\mathsf{Grid} = \{\ell/K : \ell = 1, \ldots, K\}$ and $K \geq 3$, we have*

$$\Theta_1(1) \;\geq\; \pi^* - \frac{1}{\beta K}\left(\sum_{k \in \mathcal{M}} \left| \log \sum_{i \in S_k} e^{\alpha_i} \right| + m + 3m \log K \right).$$

Proof: Let $\boldsymbol{q}^*$ be an optimal solution to problem (6). By Theorem 4.2, we have $\pi^* = \widehat{\Pi}(\boldsymbol{q}^*)$, where $\widehat{\Pi}(\boldsymbol{q})$ is as in (5). As discussed immediately before the theorem, $\Theta_1(1)$ corresponds to the largest expected revenue provided by no-purchase probabilities of the form $\boldsymbol{q} = (q_1, \ldots, q_m)$ with $q_k \in \mathsf{Grid}$ and $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$. Thus, we have $\Theta_1(1) \geq \widehat{\Pi}(\widehat{\boldsymbol{q}})$ for any $\widehat{\boldsymbol{q}}$ that satisfies $\widehat{q}_k \in \mathsf{Grid}$ and $\widehat{q}_{k-1} \geq \widehat{q}_k$ for all $k \in \mathcal{M}$. We show that there exists some $\widehat{\boldsymbol{q}} = (\widehat{q}_1, \ldots, \widehat{q}_m)$ that satisfies $\widehat{q}_k \in \mathsf{Grid}$ and $\widehat{q}_{k-1} \geq \widehat{q}_k$ for all $k \in \mathcal{M}$ such that $\widehat{\Pi}(\widehat{\boldsymbol{q}}) \geq \widehat{\Pi}(\boldsymbol{q}^*) - \frac{1}{\beta K}\big(\sum_{k \in \mathcal{M}} |\log \sum_{i \in S_k} e^{\alpha_i}| + m + 3m \log K \big)$, in which case, the desired result follows by noting that $\Theta_1(1) \geq \widehat{\Pi}(\widehat{\boldsymbol{q}})$ and $\pi^* = \widehat{\Pi}(\boldsymbol{q}^*)$. Define $\widehat{\boldsymbol{q}}$

as $\widehat{\boldsymbol{q}}_k = \min\{q_k \in \mathsf{Grid} : q_k \geq q_k^*\}$ for all $k \in \mathcal{M}$, so $\widehat{q}_k$ is obtained by rounding q_k^* up to the nearest point in Grid. In this case, since $1 \geq q_1^* \geq \ldots \geq q_m^* \geq 0$, we have $1 \geq \widehat{q}_1 \geq \ldots \geq \widehat{q}_m \geq 0$. Furthermore, since $\widehat{q}_k$ is obtained by rounding q_k^* up to the nearest point in Grid, we have $1 \geq \widehat{q}_k \geq q_k^*$ and $\widehat{q}_k \geq 1/K$ for all $k \in \mathcal{M}$. Lastly, since the two successive points in Grid are separated by $1/K$, we have $0 \leq \widehat{q}_k - q_k^* \leq 1/K$ and $-1/K \leq q_{k-1}^* - \widehat{q}_{k-1} \leq 0$, in which case, adding the two yields $-1/K \leq (q_{k-1}^* - q_k^*) - (\widehat{q}_{k-1} - \widehat{q}_k) \leq 1/K$. For notational brevity, let $\Delta_k^* = q_{k-1}^* - q_k^*$ and $\widehat{\Delta}_k = \widehat{q}_{k-1} - \widehat{q}_k$, so we write the last chain of inequalities as $-1/K \leq \Delta_k^* - \widehat{\Delta}_k \leq 1/K$. Noting that $q_{k-1}^* \geq q_k^*$ and $\widehat{q}_{k-1} \geq \widehat{q}_k$, we have $\Delta_k^* \geq 0$ and $\widehat{\Delta}_k \geq 0$. By the definition of $\widehat{\Pi}(\boldsymbol{q})$ in (5), we have

$$\begin{aligned} \widehat{\Pi}(\boldsymbol{q}^*) - \widehat{\Pi}(\widehat{\boldsymbol{q}}) \;=\; & \sum_{k \in \mathcal{M}} \frac{\lambda_k}{\beta} (\Delta_k^* - \widehat{\Delta}_k) \log\left(\sum_{i \in S_k} e^{\alpha_i} \right) + \sum_{k \in \mathcal{M}} \frac{\lambda_k}{\beta} \Big\{ \Delta_k^* \log q_{k-1}^* - \widehat{\Delta}_k \log \widehat{q}_{k-1} \Big\} \\ & + \sum_{k \in \mathcal{M}} \frac{\lambda_k}{\beta} \Big\{ \Delta_k^* \log q_k^* - \widehat{\Delta}_k \log \widehat{q}_k \Big\} - \sum_{k \in \mathcal{M}} \frac{\lambda_k}{\beta} \Big\{ \Delta_k^* \log \Delta_k^* - \widehat{\Delta}_k \log \widehat{\Delta}_k \Big\}. \end{aligned} \tag{15}$$

We bound each one of the four sums above separately. To bound the first sum, noting that $|\Delta_k^* - \widehat{\Delta}_k| \leq 1/K$ by the discussion at the beginning of the proof, we obtain

$$(\Delta_k^* - \widehat{\Delta}_k) \log\left(\sum_{i \in S_k} e^{\alpha_i} \right) \;\leq\; |\Delta_k^* - \widehat{\Delta}_k| \left| \log \sum_{i \in S_k} e^{\alpha_i} \right| \;\leq\; \frac{1}{K} \left| \log \sum_{i \in S_k} e^{\alpha_i} \right|. \tag{16}$$

To bound the second sum, note that $\Delta_k^* \geq 0$ and $q_{k-1}^* \leq \widehat{q}_{k-1}$, so $\Delta_k^* \log q_{k-1}^* \leq \Delta_k^* \log \widehat{q}_{k-1}$. Also, $\Delta_k^* - \widehat{\Delta}_k \geq -1/K$ and $\log \widehat{q}_{k-1} \leq 0$. Lastly, since $\widehat{q}^k \geq 1/K$, $-\log \widehat{q}_k \leq \log K$. Thus, we get

$$\Delta_k^* \log q_{k-1}^* - \widehat{\Delta}_k \log \widehat{q}_{k-1} \;\leq\; (\Delta_k^* - \widehat{\Delta}_k) \log \widehat{q}_{k-1} \;\leq\; -\frac{1}{K} \log \widehat{q}_{k-1} \;\leq\; \frac{1}{K} \log K. \tag{17}$$

Similarly, we have $\Delta_k^* \log q_k^* - \widehat{\Delta}_k \log \widehat{q}_k \leq \frac{1}{K} \log K$, bounding the third sum. To bound the fourth sum, consider the case $\widehat{\Delta}_k \geq 1/K$. Since $x \log x$ convex in x, the subgradient inequality yields

$$\begin{aligned} \Delta_k^* \log \Delta_k^* - \widehat{\Delta}_k \log \widehat{\Delta}_k \; & \stackrel{(a)}{\geq} \; (1 + \log \widehat{\Delta}_k)(\Delta_k^* - \widehat{\Delta}_k) \\ & = \; \Delta_k^* - \widehat{\Delta}_k + \log \widehat{\Delta}_k (\Delta_k^* - \widehat{\Delta}_k) \; \stackrel{(b)}{\geq} \; -\frac{1}{K} + \frac{1}{K} \log \widehat{\Delta}_k \; \stackrel{(c)}{\geq} \; -\frac{1}{K} - \frac{1}{K} \log K, \end{aligned} \tag{18}$$

where (a) holds since the derivative of $x \log x$ is $1 + \log x$, (b) holds since $\widehat{\Delta}_k \leq 1$, so $\log \Delta_k \leq 0$ and $-1/K \leq \Delta_k^* - \widehat{\Delta}_k \leq 1/K$, and (c) holds since $\widehat{\Delta}_k \geq 1/K$, so $\log \widehat{\Delta}_k \geq -\log K$.

Consider the case $\widehat{\Delta}_k < 1/K$. Since $\widehat{q}_{k-1}$ and $\widehat{q}_k$ are, respectively, obtained by rounding q_{k-1}^* and q_k^* up to the nearest point in Grid, if $\widehat{\Delta}_k = \widehat{q}_{k-1} - \widehat{q}_k < 1/K$, then we must have $\widehat{\Delta}_k = \widehat{q}_{k-1} - \widehat{q}_k = 0$, in which case, we must have $q_{k-1}^* - q_k^* \leq 1/K$. Furthermore, $x \log x$ is decreasing in x for $x \in (0, e^{-1})$, so since $K \geq 3$, we have $e^{-1} > 1/K \geq q_{k-1}^* - q_k^* = \Delta_k^*$, which implies that $-\frac{1}{K} \log K \leq \Delta_k^* \log \Delta_k^*$. Thus, noting that $\lim_{x \to 0} x \log x = 0$, we get $\Delta_k^* \log \Delta_k^* - \widehat{\Delta}_k \log \widehat{\Delta}_k = \Delta_k^* \log \Delta_k^* \geq -\frac{1}{K} \log K$, indicating that the inequality in (18) holds under the case $\widehat{\Delta}_k < 1/K$ as well. Adding up the inequalities in (16), (17) and (18), recalling that we also have $\Delta^*_k \log q^*_k - \widehat{\Delta}_k \log \widehat{q}_k \leq \frac{1}{K} \log K$ for the third sum and noting that $\lambda_k \leq 1$, by (15), we get $\widehat{\Pi}(\boldsymbol{q}^*) - \widehat{\Pi}(\widehat{\boldsymbol{q}}) \leq \frac{1}{\beta K} \sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + \frac{2m}{\beta K} \log K + \frac{m}{\beta K} + \frac{m}{\beta K} \log K = \frac{1}{\beta K}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m \log K\big)$. ■

We can follow the optimal state-action trajectory in the dynamic program in (14) to obtain the no-purchases probabilities $\widehat{\boldsymbol{q}} = (\widehat{q}_1, \ldots, \widehat{q}_m)$ that provide the additive performance guarantee of $\frac{1}{\beta K}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m \log K\big)$. In particular, after computing $\{\Theta_k(\cdot) : k \in \mathcal{M}\}$ through the dynamic program (14), we set $\widehat{q}_0 = 1$. For each $k \in \mathcal{M}$, we compute $\widehat{q}_k$ as an optimal solution to the problem on the right side of (14) when we solve this problem with $q_{k-1} = \widehat{q}_{k-1}$. Once we have these no-purchase probabilities, noting the expression right before (5), we can compute the corresponding stage-specific prices $\widehat{\boldsymbol{\rho}} = (\widehat{\rho}_1, \ldots, \widehat{\rho}_m)$ as $\widehat{\rho}_k = \frac{1}{\beta}\big\{\log\big(\sum_{i\in S_k} e^{\alpha_i}\big) - \log\big(\frac{1}{\widehat{q}_k} - \frac{1}{\widehat{q}_{k-1}}\big)\big\}$ for all $k \in \mathcal{M}$. These stage-specific prices yield the same additive performance guarantee.

The number of operations to obtain an additive performance guarantee of $\theta > 0$ is polynomial in $1/\theta$. By the theorem above, to get an additive performance guarantee $\theta > 0$, we need to choose K such that $\frac{1}{\beta K}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m \log K\big) \leq \theta$, but since $1/K \leq 1/\sqrt{K}$ and $\log K/K \leq 1/\sqrt{K}$ for $K \geq 3$, it is enough to choose K such that $\frac{1}{\beta\sqrt{K}}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m\big) \leq \theta$, so we can set $K = \frac{1}{(\theta\beta)^2}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m\big)^2$. The dynamic program in (14) has K possible states, K possible actions, and m decision epochs. Thus, we can solve this dynamic program in $m K^2$ operations, so noting the choice of K, we can obtain a solution that provides an additive performance guarantee of θ in $\frac{m}{(\theta\beta)^4}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m\big)^4$ operations. Our analysis for the number of operations is rather loose. For any $\epsilon > 0$, we have $1/K \leq 1/K^{1-\epsilon}$ and $\log K/K \leq 1/K^{1-\epsilon}$ for large enough K. In this case, following the same line of reasoning in this paragraph, we can obtain a solution that provides an additive performance guarantee of θ in $\frac{m}{(\theta\beta)^{2/(1-\epsilon)}}\big(\sum_{k\in\mathcal{M}} |\log \sum_{i\in S_k} e^{\alpha_i}| + m + 3m\big)^{2/(1-\epsilon)}$ operations, so the number of operations scale a bit faster than quadratically in $1/\theta$ as long as the number of points K in Grid is large, which is more aligned with our experience with the dynamic program in (14), as we report shortly.

If we have a lower bound on the optimal objective value of the Pricing problem, then we can always use the theorem above to numerically obtain a multiplicative performance guarantee. In particular, for any $\delta \in (0, 1)$, letting $\widehat{\pi}$ be a lower bound on the optimal objective value of the Pricing problem, we can use Theorem D.1 to obtain a solution that provides an additive performance guarantee of $\delta\,\widehat{\pi}$. In other words, using π^* to denote the optimal objective value of the Pricing problem, this solution provides an expected revenue of at least $\pi^* - \delta\,\widehat{\pi}$. Noting that $\pi^* - \delta\,\widehat{\pi} \geq (1 - \delta)\,\pi^*$, the solution that provides an additive performance guarantee of $\delta\,\widehat{\pi}$ also provides an expected revenue of at least $(1 - \delta)\,\pi^*$, corresponding to a solution with a multiplicative

$(m,C)=(6,3)$		
a	Avg. Gap	CPU Secs.
$+\infty$	$7.2\cdot 10^{-5}\%$	34
0.5	$9.6\cdot 10^{-5}\%$	30
0.0	$1.2\cdot 10^{-4}\%$	31
-0.1	$1.4\cdot 10^{-4}\%$	33

$(m,C)=(6,5)$		
a	Avg. Gap	CPU Secs.
$+\infty$	$1.5\cdot 10^{-4}\%$	43
0.5	$1.6\cdot 10^{-4}\%$	31
0.0	$1.8\cdot 10^{-4}\%$	30
-0.1	$1.7\cdot 10^{-4}\%$	32

$(m,C)=(10,3)$		
a	Avg. Gap	CPU Secs.
$+\infty$	$3.9\cdot 10^{-5}\%$	423
0.5	$5.3\cdot 10^{-5}\%$	358
0.0	$5.0\cdot 10^{-5}\%$	278
-0.1	$8.0\cdot 10^{-5}\%$	260

$(m,C)=(10,5)$		
a	Avg. Gap	CPU Secs.
$+\infty$	$7.5\cdot 10^{-5}\%$	520
0.5	$6.4\cdot 10^{-5}\%$	426
0.0	$1.0\cdot 10^{-4}\%$	340
-0.1	$9.2\cdot 10^{-5}\%$	336

Table EC.1 Optimality gaps of the prices obtained by using the dynamic program in (14).

performance guarantee of $1-\delta$. To obtain a lower bound on the optimal objective value of the Pricing problem, we can, for example, charge the same price for all products in all stages, in which case, we have a single decision variable in the Pricing problem. We can carry out a numerical search to find the best single price to charge in all stages.

Computational Experiments: We give a small set of computational experiments to understand the quality of the additive performance guarantee given in Theorem D.1. We randomly generate a number of test problems. For each problem instance, we use the approach in Section 4.1 to compute the optimal prices, as well as the dynamic program in (14) to compute prices with an additive performance guarantee. To choose the number of points K in Grid, we compute a lower bound on the optimal objective value of the Pricing problem by charging the same price for all products in all stages and finding the best single price to charge through numerical search. Letting $\widehat{\pi}$ be this lower bound, we choose the value of K such that we obtain an additive performance guarantee of $\frac{1}{2}\widehat{\pi}$. By the discussion in the previous paragraph, this approach yields a multiplicative performance guarantee of 50%. The approach to generate our test problems closely follows the one in Section 6.2. We briefly describe our approach and refer to Section 6.2 for details.

The number of products is $n=20$ and the price sensitivity is $\beta=1$. We come up with the parameters $\{\alpha_i : i\in\mathcal{N}\}$ as follows. We have C product clusters. We randomly assign each product to a cluster. If products i and j are in the same cluster, then the values of α_i and α_j are close. Specifically, cluster c has the centroid γ_c. We set the centroid of cluster c as $\gamma_c = c-0.5$ for all $c=1,\dots,C$. If product i belongs to cluster c, then we generate κ_i from the normal distribution with mean γ_c and standard deviation one. We set $\alpha_i=\kappa_i-\Delta$, where $\Delta=\log\sum_{i\in\mathcal{N}} e^{\kappa_i}-\log 9$. Thus, if all products were offered in the first stage at zero price, then a customer would leave without a purchase with probability 0.1. We randomly assign each product to one of the assortments $(S_1,\dots,S_m)$. Letting the random variable Y be the patience level of a customer, the probability mass function of Y is $\mathbb{P}\{Y=k\}=\frac{e^{a\cdot k}}{\sum_{\ell\in\mathcal{M}} e^{a\cdot\ell}}$, where a is a parameter that we vary.

Varying $m\in\{6,10\}$, $C\in\{3,5\}$, and $a\in\{+\infty,0.5,0.0,-0.1\}$, we get 16 parameter configurations. In each parameter configuration, we generate 25 problem instances. For each problem instance, we compute the optimal expected revenue by solving the convex program in (6) through the `fmincon` routine in Matlab. In Table EC.1, we show the average percent gap between the optimal expected revenue and the expected revenue from the prices obtained through the dynamic program in (14), averaged over the 25 instances in a parameter configuration. The second column in the table shows the runtime to solve the dynamic program in (14). Note that we choose the value of K for a performance guarantee of 50%, but the optimality gap of the prices that we obtain is less than $1.8 \cdot 10^{-4}$%. Although we report average optimality gaps, each optimality gap deviates from the average by no more than $0.2 \cdot 10^{-4}$%. For our problem instances, the runtime to solve the dynamic program in (14) ranges between half a minute to nine minutes. For comparison, although we do not report in Table EC.1, the runtime to solve problem (6) through the `fmincon` routine in Matlab takes a few seconds. Thus, solving (14) does not require convex optimization tools, but solving problem (6) through convex optimization software is faster.

Appendix E: First-Order Conditions for Optimal Prices

The proof of Theorem 4.3 uses the following lemma, which gives a characterization of the optimal stage-specific prices for the Pricing problem by using first order conditions.

Lemma E.1 *Letting $\boldsymbol{\rho}^* = (\rho_1^*, \ldots, \rho_m^*)$ be the optimal stage-specific prices in the* Pricing *problem and $q_k^* = q_k(\boldsymbol{\rho}^*)$ for all $k \in \mathcal{M}$ with $q_0^* = 1$, we have*

$$\frac{1}{\beta} - \frac{q_\ell^*}{q_{\ell-1}^*}\,\rho_\ell^* + \frac{1}{\lambda_\ell\, q_\ell^*\, q_{\ell-1}^*} \sum_{k=\ell+1}^{m} \rho_k^*\, \lambda_k \Big\{ (q_{k-1}^*)^2 - (q_k^*)^2 \Big\} = 0.$$

Proof: Since $\widehat{V}_k(\boldsymbol{\rho}) = e^{-\beta\,\rho_k} \sum_{i \in S_k} e^{\alpha_i}$, we have $\frac{\partial \widehat{V}_k(\boldsymbol{\rho})}{\partial \rho_k} = -\beta\, \widehat{V}_k(\boldsymbol{\rho})$ and $\frac{\partial \widehat{V}_k(\boldsymbol{\rho})}{\partial \rho_\ell} = 0$ for all $\ell \neq k$. In this case, noting that $q_k(\boldsymbol{\rho}) = \frac{1}{1 + \sum_{\ell=1}^k \widehat{V}_\ell(\boldsymbol{\rho})}$, for $k \geq \ell$, we get $\frac{\partial q_k(\boldsymbol{\rho})}{\partial \rho_\ell} = \beta\, \widehat{V}_\ell(\boldsymbol{\rho})\, q_k(\boldsymbol{\rho})^2$. Also, we have

$$q_{k-1}(\boldsymbol{\rho}) - q_k(\boldsymbol{\rho}) = \frac{1}{1 + \sum_{\ell=1}^{k-1} \widehat{V}_\ell(\boldsymbol{\rho})} - \frac{1}{1 + \sum_{\ell=1}^{k} \widehat{V}_\ell(\boldsymbol{\rho})} = \frac{\widehat{V}_k(\boldsymbol{\rho})}{(1 + \sum_{\ell=1}^{k-1} \widehat{V}_\ell(\boldsymbol{\rho}))\,(1 + \sum_{\ell=1}^{k} \widehat{V}_\ell(\boldsymbol{\rho}))},$$

so $q_{k-1}^* - q_k^* = \widehat{V}_k(\boldsymbol{\rho}^*)\, q_{k-1}^*\, q_k^*$. Lastly, the optimal prices are finite, since decreasing all infinite prices to the largest finite price charged in any stage improves the expected revenue.

By (5), $\Pi(\boldsymbol{\rho}) = \sum_{k \in \mathcal{M}} \lambda_k\, \rho_k\, (q_{k-1}(\boldsymbol{\rho}) - q_k(\boldsymbol{\rho}))$ is the expected revenue as a function of stage-specific prices. Note that $q_k(\boldsymbol{\rho})$ depends on ρ_ℓ only if $k \geq \ell$. Thus, differentiating $\Pi(\boldsymbol{\rho})$, we get

$$\frac{\partial \Pi(\boldsymbol{\rho})}{\partial \rho_\ell} \;=\; \lambda_\ell\, (q_{\ell-1}(\boldsymbol{\rho}) - q_\ell(\boldsymbol{\rho})) - \lambda_\ell\, \rho_\ell\, \beta\, \widehat{V}_\ell(\boldsymbol{\rho})\, q_\ell(\boldsymbol{\rho})^2 + \sum_{k=\ell+1}^{m} \lambda_k\, \rho_k\, \beta\, \widehat{V}_\ell(\boldsymbol{\rho}) \Big\{ q_{k-1}(\boldsymbol{\rho})^2 - q_k(\boldsymbol{\rho})^2 \Big\},$$

where we use the fact that $\frac{\partial q_k(\boldsymbol{\rho})}{\partial \rho_\ell} = \beta\, \widehat{V}_\ell(\boldsymbol{\rho})\, q_k(\boldsymbol{\rho})^2$ for $k \geq \ell$, but $q_k(\boldsymbol{\rho})$ does not depend on ρ_ℓ for $k < \ell$, so we have $\frac{\partial q_k(\boldsymbol{\rho})}{\partial \rho_\ell} = 0$ for $k < \ell$. The optimal stage-specific prices $\boldsymbol{\rho}^*$ satisfies the first order condition $\frac{\partial \Pi(\boldsymbol{\rho})}{\partial \rho_\ell}\Big|_{\boldsymbol{\rho}=\boldsymbol{\rho}^*} = 0$. Therefore, using the equality above, along with the fact that $q^*_{\ell-1} - q^*_\ell = \widehat{V}_\ell(\boldsymbol{\rho}^*)\, q^*_{\ell-1}\, q^*_\ell$, we get

$$\frac{\partial \Pi(\boldsymbol{\rho})}{\partial \rho_\ell}\Bigg|_{\boldsymbol{\rho}=\boldsymbol{\rho}^*} = \lambda_\ell\, \beta\, \widehat{V}_\ell(\boldsymbol{\rho}^*)\, q^*_{\ell-1}\, q^*_\ell \left\{ \frac{1}{\beta} - \frac{q^*_\ell}{q^*_{\ell-1}}\, \rho^*_\ell + \frac{1}{\lambda_\ell\, q^*_\ell\, q^*_{\ell-1}} \sum_{k=\ell+1}^m \rho^*_k\, \lambda_k \Big\{ (q^*_{k-1})^2 - (q^*_k)^2 \Big\} \right\} = 0.$$

Since the optimal prices are finite, $\widehat{V}_\ell(\boldsymbol{\rho}^*) \neq 0$, along with $q^*_{\ell-1} = q_{\ell-1}(\boldsymbol{\rho}^*) \neq 0$ and $q^*_\ell = q_\ell(\boldsymbol{\rho}^*) \neq 0$, in which case $\boldsymbol{\rho}^*$ satisfies the equality in the lemma. ■

Appendix F: Proof of Theorem 4.4 and Tightness of the Performance Guarantee of 87.8%

The proof uses a chain of upper bounds. Setting $\lambda_k = 1$ for all $k \in \mathcal{M}$ enlarges the objective function of the Pricing-Assortment problem. By (5), we can express the expected revenue as a function of the no-purchase probabilities. So, setting $\lambda_k = 1$ for all $k \in \mathcal{M}$ in (5), as a function of no-purchase probabilities $\boldsymbol{q}$ and assortments $(S_1, \ldots, S_m)$, we can upper bound the expected revenue by

$$\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m) = \frac{1}{\beta} \sum_{k \in \mathcal{M}} (q_{k-1} - q_k) \left\{ \log\left(\sum_{i \in S_k} e^{\alpha_i} \right) - \log\left(\frac{1}{q_k} - \frac{1}{q_{k-1}} \right) \right\}. \tag{19}$$

So, we can upper bound the optimal objective value of the Pricing-Assortment problem by maximizing $\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m)$ over all $(\boldsymbol{q}, S_1, \ldots, S_m) \in \mathbb{R}^m_+ \times \mathcal{F}$ such that $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$.

Throughout this section, we set $\beta = 1$ for notational brevity, which simply scales the expected revenue by β. Also, recall that $q_0 = 1$. Letting $T = \sum_{i \in \mathcal{N}} e^{\alpha_i}$, we define $R^{(\ell)}(q_1, \ldots, q_\ell)$ as

$$R^{(\ell)}(q_1, \ldots, q_\ell) = \sum_{k=1}^{\ell} (q_{k-1} - q_k) \log(q_{k-1}\, q_k) + (1 - q_\ell) \log\left(\frac{T}{1 - q_\ell} \right). \tag{20}$$

We have the superscript (ℓ) in $R^{(\ell)}(q_1, \ldots, q_\ell)$ since we will work with different numbers of stages. In the next lemma, we show that $R^{(m)}(q_1, \ldots, q_m)$ is an upper bound on $\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m)$.

Lemma F.1 *If $(\boldsymbol{q}, S_1, \ldots, S_m) \in \mathbb{R}^m_+ \times \mathcal{F}$ satisfies $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$, then we have $R^{(m)}(q_1, \ldots, q_m) \geq \widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m)$.*

Proof: If $(S_1, \ldots, S_m) \in \mathcal{F}$, then we have $\sum_{k \in \mathcal{M}} \sum_{i \in S_k} e^{\alpha_i} \leq T$. Thus, noting the definition of $\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m)$ and using the decision variables $\boldsymbol{x} = (x_1, \ldots, x_m)$, we get

$$\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m) \leq \max_{\boldsymbol{x} \in \mathbb{R}^m_+} \left\{ \sum_{k \in \mathcal{M}} (q_{k-1} - q_k) \left\{ \log x_k - \log\left(\frac{1}{q_k} - \frac{1}{q_{k-1}} \right) \right\} \; : \; \sum_{k \in \mathcal{M}} x_k \leq T \right\}. \tag{21}$$

Since $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$ with $q_0 = 1$, if $q_m = 1$, then we have $q_k = 1$ for all $k \in \mathcal{M}$, so using the fact that $\lim_{x \to 0} x \log x = 0$, we get $\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m) = 0 = R^{(m)}(q_1, \ldots, q_m)$.

In the rest of the proof, we consider the case $q_m < 1$. We can solve the problem on the right side of (21) by using Lagrangian relaxation. For $(a_1, \ldots, a_m) \in \mathbb{R}^m_+$, the optimal solution $\boldsymbol{x}^*$ to the problem $\max_{\boldsymbol{x}\in\mathbb{R}^m_+}\left\{\sum_{k\in\mathcal{M}} a_k \log x_k : \sum_{k\in\mathcal{M}} x_k \le T\right\}$ is obtained by setting $x^*_k = \frac{T}{\sum_{\ell\in\mathcal{M}} a_\ell}\, a_k$ for all $k\in\mathcal{M}$. To show this result, we can relax the constraint $\sum_{k\in\mathcal{M}} x_k \le T$ by using a Lagrange multiplier and compute the optimal value of the Lagrange multiplier by noting that this constraint must be tight at optimality. Using this result with $a_k = q_{k-1} - q_k$, since $\sum_{k\in\mathcal{M}}(q_{k-1}-q_k) = q_0 - q_m = 1-q_m$, the optimal solution $\boldsymbol{x}^*$ to the problem on the right side of (21) is obtained by setting $x^*_k = \frac{T}{1-q_m}(q_{k-1}-q_k)$ for all $k\in\mathcal{M}$. Plugging this optimal solution into (21), we get

$$\begin{aligned}\widehat{\Pi}(\boldsymbol{q}, S_1,\ldots,S_m) &\le \sum_{k\in\mathcal{M}}(q_{k-1}-q_k)\left\{\log\left(\frac{T}{1-q_m}\right) + \log(q_{k-1}-q_k) - \log\left(\frac{1}{q_k}-\frac{1}{q_{k-1}}\right)\right\}\\ &= (1-q_m)\log\left(\frac{T}{1-q_m}\right) + \sum_{k\in\mathcal{M}}(q_{k-1}-q_k)\left\{\log(q_{k-1}-q_k) - \log\left(\frac{q_{k-1}-q_k}{q_{k-1}\,q_k}\right)\right\}\\ &= (1-q_m)\log\left(\frac{T}{1-q_m}\right) + \sum_{k\in\mathcal{M}}(q_{k-1}-q_k)\log(q_{k-1}\,q_k).\end{aligned}$$

The desired result follows by noting that the expression on the right side of the chain of inequalities above corresponds to $R^{(m)}(q_1,\ldots,q_m)$. ■

To get an upper bound on the optimal objective value of the Pricing-Assortment problem, we can maximize the upper bound on the objective function, yielding the problem

$$\widehat{z}^{(\ell)} = \max_{(q_1,\ldots,q_\ell)\in\mathbb{R}^\ell_+}\left\{R^{(\ell)}(q_1,\ldots,q_\ell)\ :\ 1\ge q_1\ge\ldots\ge q_\ell\ge 0\right\}. \tag{22}$$

In the next lemma, we show that an optimal solution to the problem above occurs in the strict interior of the feasible set and the optimal objective value in (22) is strictly increasing in ℓ.

Lemma F.2 *Letting $(q^*_1,\ldots,q^*_\ell)$ be an optimal solution to problem (22), for all $\ell = 1,2,\ldots,$ we have $1 > q^*_1 > \ldots > q^*_\ell > 0$ and $\widehat{z}^{(\ell)} > \widehat{z}^{(\ell-1)}$ with the convention that $\widehat{z}^{(0)} = 0$.*

Proof: We show the result by using induction on the number of stages. For $\ell = 1$, we have $R^{(1)}(q_1) = (1-q_1)\log q_1 + (1-q_1)\log\frac{T}{1-q_1}$, so that $R^{(1)}(0) = -\infty$, $R^{(1)}(1) = 0$ and $\frac{\partial R^{(1)}(q_1)}{\partial q_1}\Big|_{q_1=1} = -\infty$. Thus, the value of $R^{(1)}(q_1)$ at $1-\epsilon$ is strictly greater than zero for small enough $\epsilon > 0$, which implies that $\widehat{z}^{(1)} > 0 = \widehat{z}^{(0)}$ and the maximizer of $R^{(1)}(q_1)$ over the interval $[0,1]$ is in the strict interior of the interval $[0,1]$. Therefore, the result holds for $\ell = 1$. Assuming that the result holds for ℓ stages, we show that the result holds for $\ell+1$ stages. We have

$$\begin{aligned}R^{(\ell+1)}(q,q_1,\ldots,q_\ell) &= (1-q)\log q + (q-q_1)\log(q\,q_1)\\ &\quad + \sum_{k=2}^{\ell}(q_{k-1}-q_k)\log(q_{k-1}\,q_k) + (1-q_\ell)\log\left(\frac{T}{1-q_\ell}\right),\end{aligned}$$

which follows by using the definition of $R^{(\ell)}(q_1,\ldots,q_\ell)$. In this case, subtracting the expression above from $R^{(\ell)}(q_1,\ldots,q_\ell)$, we have $R^{(\ell+1)}(q,q_1,\ldots,q_\ell) = R^{(\ell)}(q_1,\ldots,q_\ell) + f(q,q_1)$, where $f(q,q_1)$ is given by $f(q,q_1) = (1-q)\log q + (q-q_1)\log(q\,q_1) - (1-q_1)\log q_1$. By the subgradient inequality, we have $\log x < x-1$ for all $x \in (0,1)$. Also, using the definition of $f(q,q_1)$, we have $f(1,q_1) = 0$ and $\frac{\partial f(q,q_1)}{\partial q}\big|_{q=1} = 1 - q_1 + \log q_1$. Let $(r_1^*,\ldots,r_\ell^*)$ be an optimal solution to problem (22) when we solve this problem with ℓ stages. By the induction assumption, we have $1 > r_1^* > \ldots > r_\ell^* > 0$. Since $r_1^* \in (0,1)$, we get $f(1,r_1^*) = 0$ and $\frac{\partial f(q,r_1^*)}{\partial q}\big|_{q=1} = 1 - r_1^* + \log r_1^* < 0$. Therefore, letting q^* be an optimal solution to the problem $\max_{q\in[r_1^*,1]} f(q,r_1^*)$, the objective value of this problem at $q=1$ is zero, but since the derivative of the objective function at $q=1$ is strictly negative, the objective value of this problem at $q = 1-\epsilon$ is strictly greater than zero for small enough $\epsilon > 0$. Therefore, it follows that $f(q^*,r_1^*) > 0$. In this case, we get

$$\widehat{z}^{(\ell)} \;=\; R^{(\ell)}(r_1^*,\ldots,r_\ell^*) \;<\; R^{(\ell)}(r_1^*,\ldots,r_\ell^*) + f(q^*,r_1^*) \;=\; R^{(\ell+1)}(q^*,r_1^*,\ldots,r_\ell^*) \;\stackrel{(a)}{\leq}\; \widehat{z}^{(\ell+1)},$$

where (a) holds since $1 \geq q^* \geq r_1^* > \ldots > r_\ell^* \geq 0$, so $(q^*,r_1^*,\ldots,r_\ell^*)$ is a feasible, but not necessarily an optimal, solution to problem (22) with $\ell+1$ stages. Thus, we have $\widehat{z}^{(\ell+1)} > \widehat{z}^{(\ell)}$.

Let $(q_1^*,\ldots,q_{\ell+1}^*)$ be an optimal solution to problem (22) with $\ell+1$ stages. We show that $1 > q_1^* > \ldots > q_{\ell+1}^* > 0$. To get a contradiction, assume that $q_{\tau-1}^* = q_\tau^*$ for some $\tau \leq \ell+1$, so

$$\begin{aligned}
\widehat{z}^{(\ell+1)} \;&=\; R^{(\ell+1)}(q_1^*,\ldots,q_{\ell+1}^*) \\
&=\; \sum_{k=1}^{\ell+1} (q_{k-1}^* - q_k^*)\log(q_{k-1}^*\,q_k^*) \;+\; (1-q_{\ell+1}^*)\log\left(\frac{T}{1-q_{\ell+1}^*}\right) \\
&\stackrel{(b)}{=}\; \sum_{k=1}^{\tau-1} (q_{k-1}^* - q_k^*)\log(q_{k-1}^*\,q_k^*) + \sum_{k=\tau+1}^{\ell+1} (q_{k-1}^* - q_k^*)\log(q_{k-1}^*\,q_k^*) + \; (1-q_{\ell+1}^*)\log\left(\frac{T}{1-q_{\ell+1}^*}\right) \\
&\stackrel{(c)}{=}\; R^{(\ell)}(q_1^*,\ldots,q_{\tau-1}^*,q_{\tau+1}^*,\ldots,q_{\ell+1}^*) \;\stackrel{(d)}{\leq}\; \widehat{z}^{(\ell)},
\end{aligned}$$

where (b) and (c) hold since $q_{\tau-1}^* = q_\tau^*$ and (d) holds since $(q_1^*,\ldots,q_{\ell+1}^*)$ is a feasible solution to problem (22) with $\ell+1$ stages so $1 \geq q_1^* \geq \ldots \geq q_{\ell+1}^* \geq 0$, in which case, $(q_1^*,\ldots,q_{\tau-1}^*,q_{\tau+1}^*,\ldots,q_{\ell+1}^*)$ is a feasible, but not necessarily an optimal, solution to problem (22). The chain of inequalities above contradict the fact that $\widehat{z}^{(\ell+1)} > \widehat{z}^{(\ell)}$. Therefore, we have $q_{k-1}^* > q_k^*$ for all $k = 1,\ldots,\ell+1$. Noting the convention that $q_0 = 1$, we get $1 > q_1^* > \ldots > q_{\ell+1}^*$. Lastly, if we have $q_{\ell+1}^* = 0$, then there must exist some $k = 1,\ldots,\ell+1$ such that $q_{k-1}^* > q_k^* = 0$, which implies that $(q_{k-1}^* - q_k^*)\log(q_{k-1}^*\,q_k^*) = -\infty$. Thus, by the definition of $R^{(\ell)}(q_1,\ldots,q_\ell)$, we get $\widehat{z}^{(\ell+1)} = R^{(\ell+1)}(q_1^*,\ldots,q_{\ell+1}^*) = -\infty$, contradicting the fact that $\widehat{z}^{(\ell+1)} > \widehat{z}^{(\ell)} > \ldots > \widehat{z}^{(0)} = 0$. Therefore, we have $1 > q_1^* > \ldots > q_{\ell+1}^* > 0$. In the previous paragraph, we also had $\widehat{z}^{(\ell+1)} > \widehat{z}^{(\ell)}$, so the result holds for $\ell+1$ stages. ■

In the next lemma, we build on the lemma above to give a simple expression for the objective function of problem (22) when evaluated at its optimal solution.

Lemma F.3 *Letting* $(q_1^*, \ldots, q_\ell^*)$ *be an optimal solution to problem* (22)*, this solution satisfies the two identities given by*

$$R^{(\ell)}(q_1^*, \ldots, q_\ell^*) = \log q_1^* - (1 + q_1^*) + \frac{q_{\ell-1}^*}{q_\ell^*} - \log(q_{\ell-1}^*\, q_\ell^*) + q_\ell^*,$$
$$\frac{1 - q_\ell^*}{q_{\ell-1}^*\, q_\ell^*} \exp\left(\frac{q_{\ell-1}^*}{q_\ell^*}\right) = T.$$

Proof: Using the definition of $R^{(\ell)}(q_1, \ldots, q_\ell)$ in (20), directly by differentiating this function, we have the partial derivatives

$$\frac{\partial R^{(\ell)}(q_1, \ldots, q_\ell)}{\partial q_k} = \begin{cases} \log\left(\frac{q_{k+1}}{q_{k-1}}\right) + \frac{q_{k-1} - q_{k+1}}{q_k} & \text{if } k = 1, \ldots, \ell - 1 \\ \frac{q_{\ell-1}}{q_\ell} - \log(q_{\ell-1}\, q_\ell) - \log\left(\frac{T}{1 - q_\ell}\right) & \text{otherwise.} \end{cases} \tag{23}$$

By Lemma F.2, $(q_1^*, \ldots, q_\ell^*)$ is in the strict interior of the feasible set of problem (22), so it satisfies the first order condition $\frac{\partial R^{(\ell)}(q_1,\ldots,q_\ell)}{\partial q_k}\big|_{(q_1,\ldots,q_\ell)=(q_1^*,\ldots,q_\ell^*)} = 0$ for all $k = 1, \ldots, \ell$. By (23), we get $\log\left(\frac{q_{k+1}^*}{q_{k-1}^*}\right) = -\frac{q_{k-1}^* - q_{k+1}^*}{q_k^*}$ for all $k = 1, \ldots, \ell - 1$ and $\frac{q_{\ell-1}^*}{q_\ell^*} = \log\left(\frac{T}{1-q_\ell^*}\, q_{\ell-1}^*\, q_\ell^*\right)$. Solving for T in the last equality yields the second identity in the lemma. By the definition of $R^{(\ell)}(q_1, \ldots, q_\ell)$, we get

$$\begin{aligned} R^{(\ell)}(q_1^*, \ldots, q_\ell^*) &= \sum_{k=1}^{\ell} q_{k-1}^* \log(q_{k-1}^*\, q_k^*) - \sum_{k=1}^{\ell} q_k^* \log(q_{k-1}^*\, q_k^*) + (1 - q_\ell^*) \log\left(\tfrac{T}{1-q_\ell^*}\right) \\ &= \log q_1^* + \sum_{k=1}^{\ell-1} q_k^* \log(q_k^*\, q_{k+1}^*) - \sum_{k=1}^{\ell} q_k^* \log(q_{k-1}^*\, q_k^*) + (1 - q_\ell^*) \log\left(\tfrac{T}{1-q_\ell^*}\right) \\ &= \log q_1^* + \sum_{k=1}^{\ell-1} q_k^* \log\left(\tfrac{q_{k+1}^*}{q_{k-1}^*}\right) - q_\ell^* \log(q_{\ell-1}^*\, q_\ell^*) + (1 - q_\ell^*) \log\left(\tfrac{T}{1-q_\ell^*}\right) \\ &= \log q_1^* + \sum_{k=1}^{\ell-1} q_k^* \log\left(\tfrac{q_{k+1}^*}{q_{k-1}^*}\right) - q_\ell^* \left\{ \log(q_{\ell-1}^*\, q_\ell^*) + \log\left(\tfrac{T}{1-q_\ell^*}\right) \right\} + \log\left(\tfrac{T}{1-q_\ell^*}\right) \\ &\overset{(a)}{=} \log q_1^* - \sum_{k=1}^{\ell-1} (q_{k-1}^* - q_{k+1}^*) - q_{\ell-1}^* + \frac{q_{\ell-1}^*}{q_\ell^*} - \log(q_{\ell-1}^*\, q_\ell^*) \\ &\overset{(b)}{=} \log q_1^* - (1 + q_1^*) + q_\ell^* + \frac{q_{\ell-1}^*}{q_\ell^*} - \log(q_{\ell-1}^*\, q_\ell^*), \end{aligned}$$

where (a) holds since $\log\left(\frac{q_{k+1}^*}{q_{k-1}^*}\right) = -\frac{q_{k-1}^* - q_{k+1}^*}{q_k^*}$ for all $k = 1, \ldots, \ell - 1$ and $\frac{q_{\ell-1}^*}{q_\ell^*} = \log\left(\frac{T}{1-q_\ell^*}\, q_{\ell-1}^*\, q_\ell^*\right)$, and (b) follows by cancelling the telescoping terms, so the first identity in the lemma holds. ■

In the next lemma, we give a simple inequality that will allow us to upper bound $\widehat{z}^{(\ell)}$.

Lemma F.4 *If* $s, t > 1$ *satisfies* $\log(st) + \frac{1}{t} - s = 0$*, then we have* $s \geq 2 - \frac{1}{t}$.

Proof: Letting $h(s) = 2(1 - s) + \log s - \log(2 - s)$, we have $h'(s) = -2 + \frac{1}{s} + \frac{1}{2-s} = 2\frac{(s-1)^2}{s(2-s)}$. Thus, $h(s)$ is strictly increasing in s for all $s \in (1, 2)$, which implies that $h(s) > h(1) = 0$ for all $s \in (1, 2)$.

Also, letting $f(x) = x - \log x$, we have $f'(x) = 1 - \frac{1}{x}$, so $f(x)$ is strictly decreasing in x for all $x \in (0,1)$. To get a contradiction, assume that $s, t > 1$ satisfies $\log(st) + \frac{1}{t} - s = 0$ and we have $s < 2 - \frac{1}{t}$. Since $0 < \frac{1}{t} < 2 - s < 1$ and $f(x)$ is is strictly decreasing in x for all $x \in (0,1)$, we get $f(\frac{1}{t}) > f(2-s)$. Noting the definition of $f(x)$, the last inequality is equivalent to

$$\frac{1}{t} + \log(st) - s \;>\; (2-s) - \log(2-s) + \log s - s \;=\; h(s).$$

Since $1 < s < 2 - \frac{1}{t}$, we have $s \in (1,2)$. Noting that $h(s) > 0$ for all $s \in (1,2)$, the inequality above yields $\log(st) + \frac{1}{t} - s > 0$, which contradicts the fact that $\log(st) + \frac{1}{t} - s = 0$. ■

In the next proposition, we use Lemmas F.3 and F.4 to upper bound $\widehat{z}^{(\ell)}$ with a closed-form.

Proposition F.5 *Defining the function $G_T(x) = \frac{1}{2}\left(\sqrt{1 + 4T/e^x} + 1\right)$ and noting that $\widehat{z}^{(\ell)}$ is the optimal objective value of problem* (22)*, we have*

$$\widehat{z}^{(\ell)} \;\le\; 2\log(G_T(1)) + \frac{1}{G_T(1 + \frac{1}{\ell}\,G_T(1))} - 1 + \frac{G_T(1)}{\ell}.$$

Proof: Let $(q_1^*, \ldots, q_\ell^*)$ be an optimal solution to problem (22). First, we give an upper bound on $1/q_\ell^*$. Letting $f(x) = \frac{1}{x}\exp(x/q_\ell^*)$, we have $f'(x) = f(x)\left(\frac{1}{q_\ell^*} - \frac{1}{x}\right)$. Therefore, $f(x)$ is increasing in x for all $x \ge q_\ell^*$. Since $q_{\ell-1}^* > q_\ell^*$, we obtain $\frac{1}{q_{\ell-1}^*}\exp(q_{\ell-1}^*/q_\ell^*) = f(q_{\ell-1}^*) \ge f(q_\ell^*) = e/q_\ell^*$. In this case, since $(q_1^*, \ldots, q_\ell^*)$ satisfies the second identity in Lemma F.3, we have

$$T \;=\; \frac{1 - q_\ell^*}{q_\ell^*}\left\{\frac{1}{q_{\ell-1}^*}\exp\left(\frac{q_{\ell-1}^*}{q_\ell^*}\right)\right\} \;\ge\; e\,\frac{1 - q_\ell^*}{(q_\ell^*)^2} \;=\; e\left\{\frac{1}{(q_\ell^*)^2} - \frac{1}{q_\ell^*}\right\}.$$

For fixed $x \in \mathbb{R}$, the only positive root of the quadratic equation $z^2 - z - \frac{T}{e^x} = 0$ is $G_T(x)$. Since $\frac{1}{(q_\ell^*)^2} - \frac{1}{q_\ell^*} - \frac{T}{e} \le 0$ by the chain of inequalities above, we get $\frac{1}{q_\ell^*} \le G_T(1)$.

Second, we give a lower bound on $1/q_\ell^*$. Letting $t_k^* = q_{k-1}^*/q_k^*$ for all $k = 1, \ldots, \ell$, by Lemma F.2, $t_k^* > 1$ for all $k = 1, \ldots, \ell$. Also, we can write the first order condition in the first case in (23) as $-\log(t_k^*\, t_{k+1}^*) + t_k^* - \frac{1}{t_{k+1}^*} = 0$ for all $k = 1, \ldots, \ell - 1$. In this case, by Lemma F.4, we get $t_k^* \ge 2 - \frac{1}{t_{k+1}^*}$ for all $k = 1, \ldots, \ell - 1$. Thus, letting $V_k^* = \frac{T}{1 - q_\ell^*}(q_{k-1}^* - q_k^*)$, we obtain

$$\frac{V_k^*}{V_{k+1}^*} \;=\; \frac{q_{k-1}^* - q_k^*}{q_k^* - q_{k+1}^*} \;=\; \frac{t_k^* - 1}{1 - \frac{1}{t_{k+1}^*}} \;\ge\; 1,$$

where the last inequality holds since $t_k^* \ge 2 - \frac{1}{t_{k+1}^*}$. Thus, we get $V_k^* \ge V_{k+1}^*$. By the definition of V_k^*, we have $\sum_{k=1}^{\ell} V_k^* = T$. Since $V_1^* \ge \ldots \ge V_\ell^*$ and $\sum_{k=1}^{\ell} V_k^* = T$, it follows that $V_\ell^* \le T/\ell$.

We have $V_\ell^* = \frac{T}{1 - q_\ell^*}(q_{\ell-1}^* - q_\ell^*) \le \frac{T}{\ell}$, which implies that $q_{\ell-1}^* - q_\ell^* \le \frac{1}{1 - q_\ell^*}(q_{\ell-1}^* - q_\ell^*) \le \frac{1}{\ell}$, in which case, dividing both sides of the last inequality by q_ℓ^*, we get $\frac{q_{\ell-1}^*}{q_\ell^*} \le 1 + \frac{1}{\ell\, q_\ell^*} \le 1 + \frac{G_T(1)}{\ell}$, where the last inequality holds due to the fact that $\frac{1}{q^*_\ell} \leq G_T(1)$. Since $(q^*_1,\ldots,q^*_\ell)$ satisfies the second identity in Lemma F.3, noting that $q^*_{\ell-1} \geq q^*_\ell$ and $\frac{q^*_{\ell-1}}{q^*_\ell} \leq 1 + \frac{G_T(1)}{\ell}$, we have

$$T = \frac{1-q^*_\ell}{q^*_{\ell-1}\,q^*_\ell}\exp\left(\frac{q^*_{\ell-1}}{q^*_\ell}\right) \;\leq\; \frac{1-q^*_\ell}{(q^*_\ell)^2}\exp\left(1+\frac{G_T(1)}{\ell}\right) \;=\; \exp\left(1+\frac{G_T(1)}{\ell}\right)\left\{\frac{1}{(q^*_\ell)^2} - \frac{1}{q^*_\ell}\right\}.$$

Since the only positive root of the quadratic equation $z^2 - z - \frac{T}{e^x} = 0$ is $G_T(x)$, noting that $\frac{1}{(q^*_\ell)^2} - \frac{1}{q^*_\ell} - \frac{T}{e^{1+G_T(1)/\ell}} \geq 0$ by the chain of inequalities above, we get $\frac{1}{q^*_\ell} \geq G_T(1+\frac{1}{\ell}\,G_T(1))$.

By the subgradient inequality, we have $\log x \leq x - 1$ for all $x > 0$. Noting that $(q^*_1,\ldots,q^*_\ell)$ satisfies the first identity in Lemma F.3, we get

$$\begin{aligned}
\widehat{z}^{(\ell)} \;=\; R^{(\ell)}(q^*_1,\ldots,q^*_\ell) \;&=\; \log q^*_1 - (1+q^*_1) + \frac{q^*_{\ell-1}}{q^*_\ell} - \log(q^*_{\ell-1}\,q^*_\ell) + q^*_\ell \\
&\overset{(a)}{\leq}\; \log q^*_1 - (1+q^*_1) + 1 + \frac{G_T(1)}{\ell} - 2\log(q^*_\ell) + \frac{1}{G_T(1+\frac{1}{\ell}\,G_T(1))} \\
&\overset{(b)}{\leq}\; -1 + \frac{G_T(1)}{\ell} + 2\,\log(G_T(1)) + \frac{1}{G_T(1+\frac{1}{\ell}\,G_T(1))},
\end{aligned}$$

where (a) holds since we have $\frac{q^*_{\ell-1}}{q^*_\ell} \leq 1 + \frac{G_T(1)}{\ell}$, $q^*_{\ell-1} \geq q^*_\ell$ and $\frac{1}{q^*_\ell} \geq G_T(1+\frac{1}{\ell}\,G_T(1))$, whereas (b) holds since $\log q^*_1 \leq q^*_1 - 1$ and $\frac{1}{q^*_\ell} \leq G_T(1)$. ■

If we offer all products in the first stage, then the PRICING-ASSORTMENT problem reduces to the standard pricing problem under the multinomial logit model with the same price sensitivity for all products. In this case, using $W(\cdot)$ to denote the Lambert-W function, it is a standard result that the optimal price to charge for all products is $\frac{1}{\beta}\,(1+W(T/e))$, yielding the optimal expected revenue $\frac{1}{\beta}\,W(T/e)$; see Proposition 3.2 in Zhang et al. (2018). Recalling that we set $\beta = 1$, if we offer all products in the first stage, then the optimal expected revenue is $W(T/e)$. In the next theorem, we compare the optimal expected revenue that we obtain when we offer all products in the first stage with the optimal expected revenue in the PRICING-ASSORTMENT problem.

Theorem F.6 *Noting that* $\Pi(\boldsymbol{p}, S_1,\ldots,S_m)$ *is the objective function of the* PRICING-ASSORTMENT *problem as a function of the prices* $\boldsymbol{p}$ *and the assortments* $(S_1,\ldots,S_m)$*, we have*

$$\frac{\max_{\boldsymbol{p}\in\mathbb{R}^n}\Pi(\boldsymbol{p},\mathcal{N},\varnothing,\ldots,\varnothing)}{\max_{(\boldsymbol{p},S_1,\ldots,S_m)\in\mathbb{R}^n\times\mathcal{F}}\Pi(\boldsymbol{p},S_1,\ldots,S_m)} \;\geq\; \min_{x\geq 0}\left\{\frac{(1+x)\,W(x(1+x))}{2\,(1+x)\,\log(1+x) - x}\right\}.$$

Proof: Let π^* be the optimal objective value of the PRICING-ASSORTMENT problem, corresponding to the denominator of the first fraction in the theorem. Note that $\widehat{\Pi}(\boldsymbol{q}, S_1,\ldots,S_m)$ in (19) is an upper bound on the expected revenue from the no-purchase probabilities $\boldsymbol{q}$ and the assortments $(S_1,\ldots,S_m)$. Thus, we have $\pi^* \leq \max_{(\boldsymbol{q},S_1,\ldots,S_m)\in\mathbb{R}^m_+\times\mathcal{F}}\{\widehat{\Pi}(\boldsymbol{q},S_1,\ldots,S_m) : q_{k-1}\geq q_k\ \forall\, k\in\mathcal{M}\}$, so by Lemma F.1, we obtain $\pi^* \leq \max_{(q_1,\ldots,q_m)\in\mathbb{R}^m_+}\{R^{(m)}(q_1,\ldots,q_m) : q_{k-1}\geq q_k\ \forall\, k\in\mathcal{M}\} = \widehat{z}^{(m)}$, where the equality uses the fact that $\widehat{z}^{(m)}$ is the optimal objective value of problem (22) with $\ell = m$. By the last chain of inequalities, for all $\ell \geq m$, we get

$$\pi^* \;\leq\; \widehat{z}^{(m)} \;\overset{(a)}{\leq}\; \widehat{z}^{(\ell)} \;\overset{(b)}{\leq}\; 2\,\log(G_T(1)) + \frac{1}{G_T(1+\frac{1}{\ell}\,G_T(1))} - 1 + \frac{G_T(1)}{\ell},$$

where (a) holds since $\widehat{z}^{(\ell)} \geq \widehat{z}^{(m)}$ for all $\ell \geq m$ by Lemma F.2 and (b) uses Proposition F.5. Thus, for any $\ell \geq m$, we have $\pi^* \leq 2\,\log(G_T(1)) + \frac{1}{G_T(1+G_T(1)/\ell)} - 1 + \frac{G_T(1)}{\ell}$.

The last inequality holds for all $\ell \geq m$. Taking the limit as $\ell \to \infty$ and noting that $G_T(x)$ is continuous in x, we can upper bound π^* as

$$\begin{aligned}\pi^* \;&\leq\; \lim_{\ell\to\infty}\left\{2\,\log(G_T(1)) + \frac{1}{G_T(1+\frac{1}{\ell}\,G_T(1))} - 1 + \frac{G_T(1)}{\ell}\right\}\\ &\leq\; 2\,\log(G_T(1)) + \frac{1}{G_T(1)} - 1 \;=\; 2\,\log\left(\frac{\sqrt{1+4T/e}+1}{2}\right) + \frac{2}{\sqrt{1+4T/e}+1} - 1,\end{aligned} \tag{24}$$

where the last equality uses the definition of $G_T(x)$ in Proposition F.5. Also, we know that if we offer all products in the first stage, then the optimal expected revenue is $W(T/e)$.

The function $f(x) = e\,x\,(1+x)$ is strictly increasing in x for $x \geq 0$. Making the change of variables $T = e\,x\,(x+1)$, we have $\frac{1}{2}(\sqrt{1+4T/e}+1) = 1+x$. Therefore, we obtain

$$\begin{aligned}\frac{\max_{\boldsymbol{p}\in\mathbb{R}^n}\Pi(\boldsymbol{p},\mathcal{N},\varnothing,\ldots,\varnothing)}{\max_{(\boldsymbol{p},S_1,\ldots,S_m)\in\mathbb{R}^n\times\mathcal{F}}\Pi(\boldsymbol{p},S_1,\ldots,S_m)} \;&=\; \frac{W(T/e)}{\pi^*} \;\overset{(c)}{\geq}\; \frac{W(T/e)}{2\,\log\left(\frac{\sqrt{1+4T/e}+1}{2}\right) + \frac{2}{\sqrt{1+4T/e}+1} - 1}\\ &\overset{(d)}{\geq}\; \min_{x\geq 0}\left\{\frac{W(x\,(1+x))}{2\,\log(1+x) + \frac{1}{1+x} - 1}\right\} \;=\; \min_{x\geq 0}\left\{\frac{(1+x)\,W(x(1+x))}{2\,(1+x)\,\log(1+x) - x}\right\},\end{aligned}$$

where (c) follows from (24), and (d) follows by making the change of variables $T = e\,x\,(1+x)$ and minimizing the lower bound over all $x \geq 0$. ■

Here is the proof of Theorem 4.4.

Proof of Theorem 4.4: We argue that the optimal objective value of the minimization problem on the right side of the inequality in Theorem F.6 is at least 0.878. Let $H(x) = \frac{(1+x)\,W(x(1+x))}{2\,(1+x)\,\log(1+x)-x}$. In Figure EC.1, we plot $H(x)$ as a function of x over the interval $[0, 1827]$, which is at least 0.878. It remains to demonstrate that $\min_{x\geq 1827} H(x) \geq 0.878$. For two functions $f(\cdot)$ and $g(\cdot)$ that take nonnegative values over the interval $[1827, +\infty)$, fixing $\beta = 0.878$, we have $\min_{x\geq 1827}\frac{f(x)}{g(x)} \geq \beta$ if and only if $f(x) - \beta\,g(x) \geq 0$ for all $x \geq 1827$. So, having $\min_{x\geq 1827} H(x) = 0.878$ is equivalent to having $(1+x)\,W(x(1+x)) - \beta\,(2\,(1+x)\,\log(1+x) - x) \geq 0$ for all $x \geq 1827$.

For $x \geq e$, we have $W(x) \geq \log(x) - \log\log(x)$; see Hoorfar and Hassani (2008) Therefore, it is enough to argue that $(1+x)\,(\log(x\,(1+x)) - \log\log(x\,(1+x))) - \beta\,(2\,(1+x)\,\log(1+x) - x) \geq 0$ for

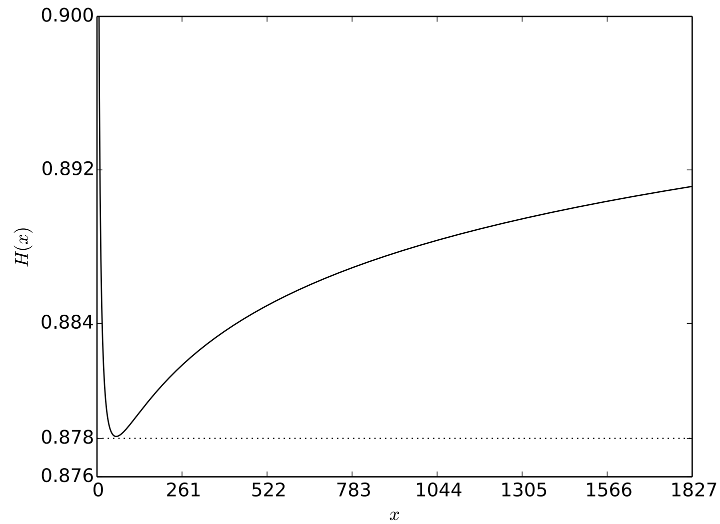

Figure EC.1 Plot of $H(x)$ as a function of x.

all $x \geq 1827$. We use $F(x)$ to denote the expression on the left side of the last inequality. By direct computation with $\beta = 0.878$, $F(1827) \geq 0.31$. Next, we check that $F'(x) \geq 0$ for $x \geq 1827$. We have

$$\begin{aligned} F'(x) &= \log x + \log(1+x) + \frac{1+x}{x} + 1 - \log\log(x\,(1+x)) \\ &\qquad - \frac{1+x}{\log x + \log(1+x)} \left(\frac{1}{x} + \frac{1}{1+x} \right) - \beta - 2\,\beta\,\log(1+x) \\ &\geq 2\log x + 2 - \log 2 - \log\log(1+x) - \frac{3655/3654}{\log x} - \beta - 2\,\beta\,\log(1+x) \\ &= (2-2\beta)\log x + 2\beta\,\log\left(\frac{x}{x+1}\right) + 2 - \log 2 - \log\log(1+x) - \frac{3655/3654}{\log x} - \beta, \qquad (25) \end{aligned}$$

where the first inequality holds since $\log\log(x\,(1+x)) \leq \log(2\,\log(1+x))$ and $\frac{1+x}{x} + 1 \leq \frac{3655}{1827}$ for all $x \geq 1827$. We split the expression on the right side above into two expressions.

First, consider the function $P(x) = 2\beta\,\log(\frac{x}{x+1}) + 2 - \log 2 - \frac{3651/3650}{\log x} - \beta$, which is increasing in x. Thus, for all $x \geq 1827$, we have $P(x) \geq P(1827) \geq 0.29$, where the second inequality is by direct computation. Second, consider the function $Q(x) = (2-2\beta)\log x - \log\log(1+x)$. By direct computation, we have $Q(1827) \geq -0.19$. Also, noting that $2 - 2\beta - \frac{1}{\log 1828} \geq 0.11$, for all $x \geq 1827$, we have $Q'(x) = (2-2\beta)\frac{1}{x} - \frac{1}{(1+x)\,\log(1+x)} \geq \frac{1}{1+x}(2 - 2\beta - \frac{1}{\log(1+x)}) \geq \frac{1}{1+x}\,(2 - 2\beta - \frac{1}{\log 1828}) \geq 0$. Thus, $Q(x)$ is increasing in x for all $x \geq 1827$, so $Q(x) \geq Q(1827) \geq -0.19$. The expression on the right side of (25) is $P(x) + Q(x)$, so $F'(x) = P(x) + Q(x) \geq 0.29 - 0.19 \geq 0$ for all $x \geq 1827$. ■

Tightness of the Performance Guarantee of 87.8%:

Intuitively speaking, we can reverse-engineer the sequence of steps in our proof of the performance guarantee of 87.8% to come up with a problem instance to demonstrate that this performance guarantee is tight. In particular, we use the following steps. First, we find the value of x^* that minimizes the function $H(x)$ in the proof of Theorem 4.4. This value is approximately 58.83. Noting the change of variables $T = e\,x\,(1+x)$ in the proof of Theorem F.6, we set $T^* = e\,x^*\,(1+x^*)$, which is approximately 9567.33. We fix the number of stages m to any positive integer. Second, setting $T = T^*$ in (20), we solve problem (22) with $\ell = m$. We let $(q_1^*, \ldots, q_m^*)$ be an optimal solution to this problem. Third, noting that the optimal value of x_k^* in problem (21) in the proof of Lemma F.1 is $x_k^* = \frac{T}{1-q_m}\,(q_{k-1} - q_k)$, we set $T_k^* = \frac{T^*}{1-q_m^*}\,(q_{k-1}^* - q_k^*)$ for all $k = 1, \ldots, m$, where T^* is as obtained in the first step and $(q_1^*, \ldots, q_m^*)$ is as obtained in the second step. Once we compute the values of $(T_1^*, \ldots, T_m^*)$, we construct our problem instance as follows. We have $\lambda_1 = \ldots = \lambda_m = 1$. The price sensitivity is $\beta = 1$. There is one product associated with each stage, so we index the products by $\{1, \ldots, m\}$. The parameter α_i for product i is such that $e^{\alpha_i} = T_i^*$.

Noting that $\widehat{z}^{(m)}$ is the optimal objective value of problem (22) with $\ell = m$, by Lemma F.1, $\widehat{z}^{(m)}$ is an upper bound on the optimal objective value of the Pricing-Assortment problem. In this case, for the problem instance that we constructed as in the previous paragraph, we can follow the proof of Lemma F.1 line by line to show that if we offer product i in stage i and optimize only over the prices of the products in the Pricing-Assortment problem, then the optimal objective value that we obtain is equal to $\widehat{z}^{(m)}$, achieving the upper bound of $\widehat{z}^{(m)}$. Thus, for the problem instance that we constructed, the optimal solution for the Pricing-Assortment problem involves offering each product i in stage i. In other words, for the problem instance that we constructed, we can solve the Pricing-Assortment problem efficiently. We offer each product i in stage i and use the approach in Section 4.1 to find the optimal prices to charge for the products.

In this case, we can follow the proofs of Proposition F.5 and Theorem F.6 line by line to show that the performance guarantee of 87.8% is tight for the problem instance that we constructed, as long as the number of stages m gets arbitrarily large. The number of stages needs to get arbitrarily large due to the limit in (24). In Figure EC.2, we numerically verify the tightness of the performance guarantee of 87.8%. For each $m \in \mathbb{Z}_+$, we construct a problem instance as described in this section. For each problem instance that we construct, we solve the Pricing-Assortment problem to get the optimal objective value, which we denote by $\overline{\pi}^{(m)}$. Also, we offer all products in the first stage and compute the optimal prices to charge for the products. We denote the corresponding optimal objective value by $\underline{\pi}^{(m)}$. In the figure, as a function of the number of stages m in the problem instance that we construct, we plot the ratio $\underline{\pi}^{(m)}/\overline{\pi}^{(m)}$.

Naturally, by Theorem 4.4, $\underline{\pi}^{(m)}/\overline{\pi}^{(m)}$ never falls below 87.8%. For smaller values of m, the ratio $\underline{\pi}^{(m)}/\overline{\pi}^{(m)}$ can be noticeably far from 87.8%, but m does not need to get too large for the ratio to be close to 87.8%. Once m reaches about 15, $\underline{\pi}^{(m)}/\overline{\pi}^{(m)}$ gets remarkably close to 87.8%, verifying that our analysis in the proof of Theorem 4.4 is tight for the problem instances constructed by using the approach discussed in this section, as long as m gets large.

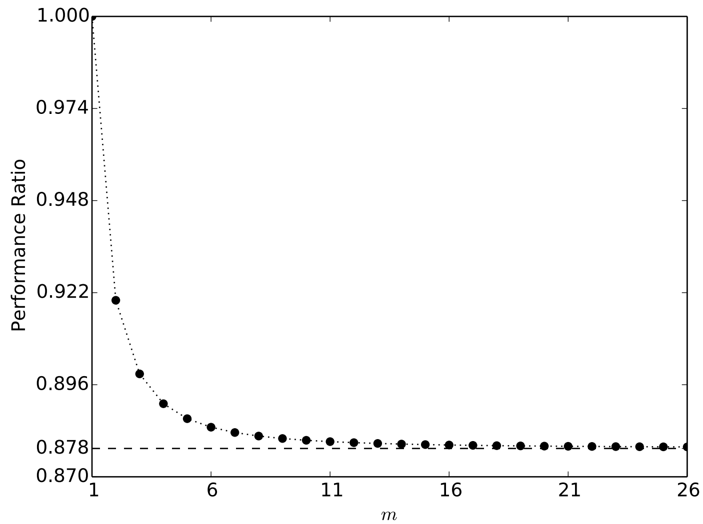

Figure EC.2 Plot of $\underline{\pi}^{(m)}/\overline{\pi}^{(m)}$ as a function of m.

Appendix G: Performance Guarantee of 50% and its Tightness

In the following lemma, we give a 50% performance guarantee for the Assortment problem by offering a nonempty assortment only in the first stage.

Lemma G.1 *Letting π^* be the optimal objective value of the* Assortment *problem, we have* $\max_{S\subseteq\mathcal{N}} \Pi(S,\varnothing,\ldots,\varnothing) \geq \frac{1}{2}\pi^*$.

Proof: Let $(S_1^*,\ldots,S_m^*)$ be an optimal solution to the Assortment problem and $T_k^* = S_1^* \cup \ldots \cup S_k^*$ with $T_0^* = \varnothing$. Noting the definition of $\Pi(S_1,\ldots,S_m)$, we get

$$\begin{aligned}
\pi^* &= \Pi(S_1^*,\ldots,S_m^*) = \sum_{k\in\mathcal{M}} \frac{\lambda_k W(T_k^*) - \lambda_k W(T_{k-1}^*)}{(1+V(T_{k-1}^*))\,(1+V(T_k^*))} \\
&\overset{(a)}{=} \sum_{k=1}^{m-1} \frac{W(T_k^*)}{1+V(T_k^*)}\left\{\frac{\lambda_k}{1+V(T_{k-1}^*)} - \frac{\lambda_{k+1}}{1+V(T_{k+1}^*)}\right\} + \frac{\lambda_m W(T_m^*)}{(1+V(T_{m-1}^*))\,(1+V(T_m^*))} \\
&\leq \max_{S\subseteq\mathcal{N}}\left\{\frac{W(S)}{1+V(S)}\right\}\left(\sum_{k=1}^{m-1}\left\{\frac{\lambda_k}{1+V(T_{k-1}^*)} - \frac{\lambda_{k+1}}{1+V(T_{k+1}^*)}\right\} + \frac{\lambda_m}{1+V(T_{m-1}^*)}\right) \\
&\overset{(b)}{=} \max_{S\subseteq\mathcal{N}}\left\{\Pi(S,\varnothing,\ldots,\varnothing)\right\}\left(\lambda_1 + \sum_{k=2}^{m}\lambda_k\left\{\frac{1}{1+V(T_{k-1}^*)} - \frac{1}{1+V(T_k^*)}\right\}\right) \\
&\leq \max_{S\subseteq\mathcal{N}}\left\{\Pi(S,\varnothing,\ldots,\varnothing)\right\}\left(1 + \sum_{k=2}^{m}\left\{\frac{1}{1+V(T_{k-1}^*)} - \frac{1}{1+V(T_k^*)}\right\}\right) \\
&= \max_{S\subseteq\mathcal{N}}\left\{\Pi(S,\varnothing,\ldots,\varnothing)\right\}\left(1 + \frac{1}{1+V(T_1^*)} - \frac{1}{1+V(T_m^*)}\right) \leq 2\max_{S\subseteq\mathcal{N}}\left\{\Pi(S,\varnothing,\ldots,\varnothing)\right\},
\end{aligned}$$

where (a) follows by by arranging the terms and (b) follows by noting that $\Pi(S,\varnothing,\ldots,\varnothing) = \lambda_1\frac{W(S)}{1+V(S)}$ and $\lambda_1 = 1$, as well as arranging the terms in the sum on the left side of the equality. ■

(S_1, S_2)	Exp. Rev
$(\{1\}, \varnothing)$	$\frac{r_1\, v_1}{1+v_1} = \frac{(1+1/\epsilon)\epsilon}{1+\epsilon} = 1 \xrightarrow{\epsilon \to 0} 1$
$(\{2\}, \varnothing)$	$\frac{r_2\, v_2}{1+v_2} = \frac{1/\epsilon}{1+1/\epsilon} = \frac{1}{1+\epsilon} \xrightarrow{\epsilon \to 0} 1$
$(\{1,2\}, \varnothing)$	$\frac{r_1\, v_1 + r_2\, v_2}{1+v_1+v_2} = \frac{(1+1/\epsilon)\epsilon + 1/\epsilon}{1+\epsilon+1/\epsilon} = 1 \xrightarrow{\epsilon \to 0} 1$
$(\{1\}, \{2\})$	$\frac{r_1\, v_1}{1+v_1} + \frac{r_2\, v_2}{(1+v_1)\,(1+v_1+v_2)} = \frac{(1+1/\epsilon)\epsilon}{1+\epsilon} + \frac{1/\epsilon}{(1+\epsilon)\,(1+\epsilon+1/\epsilon)} = 1 + \frac{1}{(1+\epsilon)\,(1+\epsilon+\epsilon^2)} \xrightarrow{\epsilon \to 0} 2$
$(\{2\}, \{1\})$	$\frac{r_2\, v_2}{1+v_2} + \frac{r_1\, v_1}{(1+v_2)\,(1+v_2+v_1)} = \frac{1/\epsilon}{1+1/\epsilon} + \frac{(1+1/\epsilon)\,\epsilon}{(1+1/\epsilon)\,(1+1/\epsilon+\epsilon)} = \frac{1}{1+\epsilon} + \frac{\epsilon^2(1+\epsilon)}{(1+\epsilon)\,(1+\epsilon+\epsilon^2)} \xrightarrow{\epsilon \to 0} 1$

Table EC.2 Expected revenue from non-dominated assortments.

Tightness of the Performance Guarantee of 50%:

We give a problem instance to demonstrate that the performance guarantee of 50% that we give for the Assortment problem in Lemma G.1 is tight. We consider a problem instance with two products and two stages. The revenues and preference weights of the products are $r_1 = 1 + 1/\epsilon$, $r_2 = 1$, $v_1 = \epsilon$, and $v_2 = 1/\epsilon$. The distribution of the patience level is given by $\lambda_1 = \lambda_2 = 1$. In Table EC.2, we give the expected revenue from each non-dominated solution, along with the limit of the expected revenue as $\epsilon \to 0$. If we offer the empty assortment in all stages except for the first one, then the largest expected revenue that we can obtain is the expected revenue from one of the solutions $(\{1\}, \varnothing)$, $(\{2\}, \varnothing)$ and $(\{1,2\}, \varnothing)$, all of which get arbitrarily close to one, as we choose ϵ arbitrarily small. On the other hand, as we choose ϵ arbitrarily small, noting the solution $(\{1\}, \{2\})$, the largest expected revenue from any solution is arbitrarily close to two. Thus, the performance guarantee of 50% that we give for the Assortment problem in Lemma G.1 is tight. To make the contrast, for the Pricing-Assortment problem, offering the empty assortment in all stages except for the first one and finding the revenue-maximizing prices in the first stage provides a tight performance guarantee of 87.8%. On the other hand, for the Assortment problem, offering the empty assortment in all stages except for the first one and finding the revenue-maximizing assortment in the first stage provides a tight performance guarantee of 50%.

Appendix H: Complexity of Joint Pricing and Assortment Optimization

We consider the Pricing-Assortment problem when the prices of the products take values only over a finite set. We show that the problem is NP-hard even when we have only two possible price levels for the products and the choice process of the customers involves only two stages with $\lambda_1 = \lambda_2 = 1$. Consider the following instance. The set of products is $\mathcal{N} = \{1, 2, \ldots, n\}$. We have two stages with $\lambda_1 = \lambda_2 = 1$. We have two price levels, which we denote by p_H and p_L with $p_H > p_L$. For each product i, if we offer it at price $q \in \{p_H, p_L\}$, then its preference weight is given by v_{iq}. We want to find the sequence of assortments to offer in the two stages and the prices to charge for the products to maximize the expected revenue. Using the vector $\boldsymbol{p} = (p_1, \ldots, p_n)$ to denote the prices that we charge for the products and (S_1, S_2) to denote the assortments that we offer in the two stages, noting the expected revenue expression in (3), we want to solve the problem

$$\max_{(\boldsymbol{p},S_1,S_2)\in\{p_L,p_H\}^n\times\mathcal{F}} \left\{ \frac{\sum_{i\in S_1} p_i\, v_{i,p_i}}{1+\sum_{i\in S_1} v_{i,p_i}} + \frac{\sum_{i\in S_2} p_i\, v_{i,p_i}}{(1+\sum_{i\in S_1} v_{i,p_i})\,(1+\sum_{i\in S_1\cup S_2} v_{i,p_i})} \right\}. \tag{26}$$

In the next lemma, we give a structural property of an optimal solution to the problem above to express it in a simpler fashion. We defer the proofs of auxiliary lemmas to the end of this section.

Lemma H.1 *There exists an optimal solution $(\boldsymbol{p}^*, S_1^*, S_2^*)$ to problem (26) such that all products are offered; that is, $S_1^*\cup S_2^* = \mathcal{N}$. All products in the first stage have the high price and all products in the second stage have the low price; that is, $p_i^* = p_H$ for all $i\in S_1^*$ and $p_i^* = p_L$ for all $i\in S_2^*$.*

By Lemma H.1, the critical decision is the assortment offered S in the first stage, in which case, we offer the assortment $\mathcal{N}\setminus S$ in the second stage. Thus, problem (26) is equivalent to

$$\max_{S\subseteq\mathcal{N}} \left\{ \frac{p_H\sum_{i\in S} v_{i,H}}{1+\sum_{i\in S} v_{i,H}} + \frac{p_L\sum_{i\notin S} v_{i,L}}{(1+\sum_{i\in S} v_{i,H})\,(1+\sum_{i\in S} v_{i,H}+\sum_{i\notin S} v_{i,L})} \right\}, \tag{27}$$

where we let $v_{i,H} = v_{i,p_H}$ and $v_{i,L} = v_{i,p_L}$ for notational brevity. To establish the computational complexity, we will consider the following decision-theoretic version of the problem above.

<u>Two Stages and Two Price Levels</u>:

Inputs: A set of products index by $\mathcal{N} = \{1,\ldots,n\}$, two price levels p_H and p_L with $p_H > p_L > 0$, two preference weights $v_{i,H}$ and $v_{i,L}$ for each product $i\in\mathcal{N}$, and an expected revenue target T.

Question: Does there exist a subset of products $S\subseteq\mathcal{N}$ that provides an expected revenue of T or more in problem (27)?

The main result of this section is given in the following theorem, showing that the Two Stages and Two Price Levels problem is NP-complete.

Theorem H.2 *The* Two Stages and Two Price Levels *problem is NP-complete.*

We will use two auxiliary lemmas in the proof of the theorem above. The first lemma focuses on the complexity of the following variant of the subset sum problem.

<u>Three-Quarters Subset Sum</u>:

Inputs: A collection of weights $w_1, w_2, \ldots, w_n$ such that $w_i\in\mathbb{Q}_{++}$ for all $i = 1,\ldots,n$.

Question: Does there exist a subset $S\subseteq\{1,\ldots,n\}$ such that $\sum_{i\in S} w_i = \frac{3}{4}\sum_{i=1}^n w_i$.

In the following lemma, we show that the Three-Quarters Subset Sum problem is NP-complete. We give the proof of this lemma also at the end of this section.

Lemma H.3 *The* Three-Quarters Subset Sum *problem is NP-complete.*

Lastly, in the next lemma, we characterize the maximizer of a function that is crucial in our NP-completeness proof. The proof of this lemma is at the end of this section as well.

Lemma H.4 *For each* $\pi > 1$ *and* $\alpha > 1$ *such that* $1 < \frac{\alpha-1}{\pi-1} < (1+\alpha)^2$, *define* $f_{\pi,\alpha} : [0,1] \to \mathbb{R}_+$ *as*

$$f_{\pi,\alpha}(x) \;=\; \frac{\pi x}{1+x} + \frac{\alpha(1-x)}{(1+x)(1+\alpha-(\alpha-1)x)}.$$

Then, $f_{\pi,\alpha}$ *achieves its unique maximum at* $x^* = \dfrac{1+\alpha-\sqrt{(\alpha-1)/(\pi-1)}}{-1+\alpha+\sqrt{(\alpha-1)/(\pi-1)}}$.

Here is the proof of Theorem H.2.

Proof of Theorem H.2: We will use a reduction from the Three-Quarters Subset Sum problem, which is NP-complete by Lemma H.3. Consider an arbitrary instance of the Three-Quarters Subset Sum problem with the weights $w_1, w_2, \ldots, w_n$. Without loss of generality, we assume that $\sum_{i=1}^n w_i = 1$, because we can normalize all of the weights by dividing them by $\sum_{i=1}^n w_i$ without changing the answer to the problem. The Three-Quarters Subset Sum problem asks whether there exists a subset $S \subseteq \{1, \ldots, n\}$ such that $\sum_{i\in S} w_i = \frac{3}{4}$.

We construct an instance of the Two Stages and Two Price Levels problem as follows. The set of products is $\{1, \ldots, n\}$. The two price levels are $p_H = \frac{5}{2}$ and $p_L = 1$ with the corresponding preference weights $v_{i,H} = w_i$ and $v_{i,L} = 7w_i$ for each product i. Considering the function $f_{\pi,\alpha}$ in Lemma H.4 with $\pi = \frac{5}{2}$ and $\alpha = 7$, we set the expected revenue target as $T = f_{\frac{5}{2},7}(3/4)$. Let $\textsc{Rev}(S)$ be the expected revenue in this Two Stages and Two Price Levels problem. Noting that $\sum_{i\notin S} w_i = \sum_{i=1}^n w_i - \sum_{i\in S} w_i = 1 - \sum_{i\in S} w_i$, by (27), we have

$$\begin{aligned}\textsc{Rev}(S) \;&=\; \frac{\frac{5}{2}\sum_{i\in S} w_i}{1+\sum_{i\in S} w_i} \;+\; \frac{7\sum_{i\notin S} w_i}{(1+\sum_{i\in S} w_i)\,(1+\sum_{i\in S} w_i + 7\sum_{i\notin S} w_i)} \\ &=\; \frac{\frac{5}{2}\sum_{i\in S} w_i}{1+\sum_{i\in S} w_i} \;+\; \frac{7(1-\sum_{i\in S} w_i)}{(1+\sum_{i\in S} w_i)\,(8-6\sum_{i\in S} w_i)}.\end{aligned}$$

We will show that there exists a subset $A \subseteq \{1, \ldots, n\}$ such that $\sum_{i\in A} w_i = \frac{3}{4}$ in the Three-Quarters Subset Sum Problem if and only if there exists a subset $S \subseteq \{1, \ldots, n\}$ in the Two-Stages and Two Price Levels problem such that $\textsc{Rev}(S) \geq T$.

By the definitions of $\textsc{Rev}(S)$ above and $f_{\pi,\alpha}$ in Lemma H.4, $\textsc{Rev}(S) = f_{\frac{5}{2},7}(\sum_{i\in S} w_i)$. Also, our choice of $\pi = \frac{5}{2}$ and $\alpha = 7$ satisfies $\pi > 1$, $\alpha > 1$ and $1 < \frac{\alpha-1}{\pi-1} < (1+\alpha)^2$, so by Lemma H.4, the function $f_{\frac{5}{2},7}$ achieves its unique maximum at $x^* = \frac{1+7-\sqrt{(7-1)/(\frac{5}{2}-1)}}{-1+7+\sqrt{(7-1)/(\frac{5}{2}-1)}} = \frac{3}{4}$. Thus, for any subset $S \subseteq \{1, \ldots, n\}$, we have $\textsc{Rev}(S) = f_{\frac{5}{2},7}(\sum_{i\in S} w_i) \leq f_{\frac{5}{2},7}(3/4) = T$ and the inequality holds as an equality if and only if $\sum_{i\in S} w_i = 3/4$. Therefore, there exists a subset $S \subseteq \{1, \ldots, n\}$ such that $\textsc{Rev}(S) \geq T$ if and only if there exists a subset $S \subseteq \{1, \ldots, n\}$ such that $\sum_{i\in S} w_i = \frac{3}{4}$. ■

Proofs of Auxiliary Lemmas:

In the rest of this section, we give the proofs for the auxiliary lemmas that we used to show Theorem H.2. Here is the proof of Lemma H.1.

Proof of Lemma H.1: Let $(\boldsymbol{p}^*, S_1^*, S_2^*)$ be an optimal solution to problem (26). If $i \notin S_1^* \cup S_2^*$, then offering product i in the second stage at price level p_H does not degrade the expected revenue, so we can assume that $S_1^* \cup S_2^* = \mathcal{N}$. Considering the prices $\boldsymbol{p}^*$, let $H^* = \{i \in \mathcal{N} : p_i^* = p_H\}$ and $L^* = \{i \in \mathcal{N} : p_i^* = p_L\}$ be the sets of products for which we charge the two price levels. Fixing the prices at $\boldsymbol{p}^*$ and optimizing over the sequence of assortments $(S_1, S_2) \in \mathcal{F}$, problem (26) becomes equivalent to the ASSORTMENT problem. Thus, by the revenue-ordered property in Theorem 3.1, one of the three solutions $(H^* \cup L^*, \varnothing)$, (H^*, L^*), and $(H^*, \varnothing)$ is optimal to this problem. In particular, we have two stages, so noting the revenue thresholds $+\infty = t_1^* \geq t_2^* \geq t_3^*$ in Theorem 3.1, the solutions $(H^* \cup L^*, \varnothing)$, (H^*, L^*), and $(H^*, \varnothing)$, respectively, correspond to the cases $+\infty = t_1^* > p_H > p_L \geq t_2^* \geq t_3^*$, $+\infty = t_1^* > p_H \geq t_2^* > p_L \geq t_3^*$, and $+\infty = t_1^* > p_H \geq t_2^* \geq t_3^* > p_L$. The solution (H^*, L^*) does not degrade the expected revenue from the solution $(H^*, \varnothing)$, since offering some product at the second stage provides additional expected revenue without changing the expected revenue from the first stage. Thus, it is enough to show that the solution (H^*, L^*) does not degrade the expected revenue from the solution $(H^* \cup L^*, \varnothing)$. Let $V_{H^*} = \sum_{i \in H^*} v_{i,H}$ and $V_{L^*} = \sum_{i \in L^*} v_{i,L}$, so

$$\begin{aligned}
\frac{p_H V_{H^*} + p_L V_{L^*}}{1 + V_{H^*} + V_{L^*}} &= \frac{p_H V_{H^*}}{1 + V_{H^*}} + \frac{p_H V_{H^*}}{1 + V_{H^*} + V_{L^*}} - \frac{p_H V_{H^*}}{1 + V_{H^*}} \\
&\quad + \frac{p_L V_{L^*}}{(1 + V_{H^*})(1 + V_{H^*} + V_{L^*})} + \frac{p_L V_{L^*}}{1 + V_{H^*} + V_{L^*}} - \frac{p_L V_{L^*}}{(1 + V_{H^*})(1 + V_{H^*} + V_{L^*})} \\
&= \frac{p_H V_{H^*}}{1 + V_{H^*}} - \frac{p_H V_{H^*} V_{L^*}}{(1 + V_{H^*})(1 + V_{H^*} + V_{L^*})} \\
&\quad + \frac{p_L V_{L^*}}{(1 + V_{H^*})(1 + V_{H^*} + V_{L^*})} + \frac{p_L V_{L^*} V_{H^*}}{(1 + V_{H^*})(1 + V_{H^*} + V_{L^*})} \\
&\leq \frac{p_H V_{H^*}}{1 + V_{H^*}} + \frac{p_L V_{L^*}}{(1 + V_{H^*})(1 + V_{H^*} + V_{L^*})},
\end{aligned}$$

where the inequality uses the fact that $p_H^* > p_L^*$. The first and last expressions above are, respectively, the expected revenues from the solutions $(H^* \cup L^*, \varnothing)$ and (H^*, L^*). ■

Next, we give a proof for Lemma H.3.

Proof of Lemma H.3: We use a reduction from the standard PARTITION problem, which is a well-known NP-complete problem; see Garey and Johnson (1979). Consider an arbitrary instance of the PARTITION problem, where we have a collection of n items with weights $\{s_1, s_2, \ldots, s_n\} \in \mathbb{Q}_{++}$. Letting $T = \sum_{i=1}^n s_i$, the question is to determine whether there exists a subset $S \subseteq \{1, \ldots, n\}$ such that $\sum_{i \in S} s_i = T/2$. Given an instance of the PARTITION problem, we construct instance of the THREE-QUARTERS SUBSET SUM problem with $n+1$ items, where $w_1 = s_1, \ldots, w_n = s_n$ and

$w_{n+1} = T$. We will show there exists a subset $S \subseteq \{1,\ldots,n\}$ in the Partition problem such that $\sum_{i\in S} w_i = T/2$ if and only if there exists a subset $A \subseteq \{1,\ldots,n,n+1\}$ in the Three-Quarters Subset Sum problem such that $\sum_{i\in A} w_i = \frac{3}{4}\sum_{i=1}^{n+1} w_i$.

Assume that there exists a subset $S \subseteq \{1,\ldots,n\}$ such that $\sum_{i\in S} s_i = T/2$. In this case, the subset $A = S \cup \{n+1\} \subseteq \{1,\ldots,n\}$ satisfies $\sum_{i\in A} w_i = w_{n+1} + \sum_{i\in S} w_i = T + (T/2) = \frac{3}{2}T = \frac{3}{4}(2T) = \frac{3}{4}\sum_{i=1}^{n+1} w_i$. On the other hand, assume that there exists a subset $A \subseteq \{1,\ldots,n,n+1\}$ such that $\sum_{i\in A} w_i = \frac{3}{4}\sum_{i=1}^{n+1} w_i = \frac{3}{2}T$. In this case, note that we must have $n+1 \in A$, because $\sum_{i=1}^{n} w_i = T$, but $\sum_{i\in A} w_i = \frac{3}{2}T$. Since $w_{n+1} = T$ and $n+1 \in A$, it follows that the subset $S = A \setminus \{n+1\} \subseteq \{1,\ldots,n\}$ satisfies $\sum_{i\in S} w_i = \sum_{i\in A} w_i - w_{n+1} = \frac{3}{2}T - T = T/2$. ∎

Here is the proof of Lemma H.4.

Proof of Lemma H.4: Letting $G(x) = 1 + \alpha - (\alpha - 1)\,x$ for notational brevity, we have $G(x) - 2\alpha = (1-\alpha)\,(1+x)$ and $G(x) - \alpha\,(1-x) = 1 + x$. Furthermore, we can express $f_{\pi,\alpha}(x)$ as $f_{\pi,\alpha}(x) = \frac{\pi\,x}{1+x} + \alpha\,\frac{1-x}{1+x}\cdot\frac{1}{G(x)}$. Noting that $G'(x) = -(\alpha-1)$, differentiating $f_{\pi,\alpha}(x)$, we get

$$\begin{aligned}
f'_{\pi,\alpha}(x) &= \frac{\pi}{(1+x)^2} - \alpha\,\frac{2}{(1+x)^2}\cdot\frac{1}{G(x)} + \alpha\,\frac{1-x}{1+x}\cdot\frac{\alpha-1}{G(x)^2} \\
&= \frac{1}{G(x)^2}\left\{\frac{\pi\,G(x)^2 - 2\alpha\,G(x) + \alpha(\alpha-1)\,(1-x)\,(1+x)}{(1+x)^2}\right\} \\
&= \frac{1}{G(x)^2}\left\{\frac{(\pi-1)\,G(x)^2 + G(x)\,(G(x)-2\alpha) + \alpha(\alpha-1)\,(1-x)\,(1+x)}{(1+x)^2}\right\} \\
&\overset{(a)}{=} \frac{1}{G(x)^2}\left\{\frac{(\pi-1)\,G(x)^2 + (1-\alpha)\,(1+x)\,(G(x)-\alpha\,(1-x))}{(1+x)^2}\right\} \\
&\overset{(b)}{=} \frac{1}{G(x)^2}\left\{\frac{(\pi-1)\,G(x)^2 + (1-\alpha)\,(1+x)^2}{(1+x)^2}\right\} \\
&= \frac{\pi-1}{G(x)^2}\left\{\left(\frac{G(x)}{1+x}\right)^2 - \frac{\alpha-1}{\pi-1}\right\} \overset{(c)}{=} \frac{\pi-1}{(1+\alpha-(\alpha-1)\,x)^2}\left\{\left(1+\alpha\,\frac{1-x}{1+x}\right)^2 - \frac{\alpha-1}{\pi-1}\right\},
\end{aligned}$$

where (a) holds since $G(x) - 2\alpha = (1-\alpha)\,(1+x)$, (b) holds since $G(x) - \alpha\,(1-x) = 1 + x$ and (c) holds by the definition of $G(x)$. Therefore, defining the function $g : [0,1] \to \mathbb{R}$ as $g(x) = \left(1 + \alpha\,\frac{1-x}{1+x}\right)^2 - \frac{\alpha-1}{\pi-1}$, the sign of $f'_{\pi,\alpha}(x)$ is determined by the sign of $g(x)$.

We have $g'(x) = -\frac{4\alpha}{(1+x)^2}\left(1 + \alpha\frac{1-x}{1+x}\right) < 0$ for all $x \in [0,1]$, so $g(x)$ is strictly decreasing in x over the interval $[0,1]$. Also, since $1 < \frac{\alpha-1}{\pi-1} < (1+\alpha)^2$, we get $(1+\alpha)^2 - \frac{\alpha-1}{\pi-1} = g(0) > 0 > g(1) = 1 - \frac{\alpha-1}{\pi-1}$, which implies that $g(x)$ crosses zero at a unique point over the interval $[0,1]$.

By the discussion in the previous paragraph, $f_{\pi,\alpha}(x)$ is strictly increasing, then strictly decreasing in x over the interval $[0,1]$, so it has a unique maximizer over this interval. To find the maximizer x^* of $f_{\pi,\alpha}(x)$, we set $g(x^*) = \left(1 + \alpha\,\frac{1-x^*}{1+x^*}\right)^2 - \frac{\alpha-1}{\pi-1} = 0$, yielding $\alpha\,\frac{1-x^*}{1+x^*} = \sqrt{\frac{\alpha-1}{\pi-1}} - 1$. For constants

a and b, the value of x that solves $a\frac{1-x}{1+x}=b$ is $\frac{a-b}{a+b}$. Thus, setting $a=\alpha$ and $b=\sqrt{\frac{\alpha-1}{\pi-1}}-1$ in the last equality, we get $x^*=\frac{1+\alpha-\sqrt{(\alpha-1)/(\pi-1)}}{-1+\alpha+\sqrt{(\alpha-1)/(\pi-1)}}$, as desired. Lastly, we check that this value of x^* is in the interval $[0,1]$. Noting that $\sqrt{\frac{\alpha-1}{\pi-1}}<1+\alpha$ and $\alpha>1$, we get $x^*\geq 0$. Also, $1-\sqrt{\frac{\alpha-1}{\pi-1}}<-1+\sqrt{\frac{\alpha-1}{\pi-1}}$, so adding α to both sides of this inequality and dividing by $-1+\alpha+\sqrt{\frac{\alpha-1}{\pi-1}}$, we get $x^*\leq 1$. ■

Appendix I: Proof of Lemma 5.1

For $i\in S_k^*$ and $j\in S_{k+1}^*$, we must have $r_i\geq t_{k+1}(S_1^*,\dots,S_m^*)>r_j$. In particular, by the first part of Lemma 3.2, if $r_j\geq t_{k+1}(S_1^*,\dots,S_m^*)$, then we can move product j from stage $k+1$ to stage k without degrading the expected revenue provided by the solution $(S_1^*,\dots,S_m^*)$, which contradicts the fact that $(S_1^*,\dots,S_m^*)$ is non-dominated. By the second part of Lemma 3.2, if $r_i<t_{k+1}(S_1^*,\dots,S_m^*)$, then we can move product i from stage k to stage $k+1$ to obtain a solution strictly better than the solution $(S_1^*,\dots,S_m^*)$, which contradicts the fact that $(S_1^*,\dots,S_m^*)$ is an optimal solution. Thus, if $r_i\in S_k^*$ and $j\in S_{k+1}^*$, then we must have $r_i>r_j$. ■

Appendix J: Proof of Lemma 5.3

In this section, we give a proof for Lemma 5.3. We need the next intermediate lemma, where we show a monotonicity property for the value functions computed through (10).

Lemma J.1 *If the value functions* $\{\Theta_i^\ell(x,y):(x,y)\in\textsc{Dom}^2,\ i=1,\dots,\ell+1\}$ *are computed through the dynamic program in* (10)*, then* $\Theta_i^\ell(x,y)$ *is increasing in* x *and decreasing in* y*.*

Proof: We show the result by using induction over the decision epochs. By the boundary condition, $\Theta_{\ell+1}^\ell(x,y)$ is increasing in x and decreasing in y. Assuming that $\Theta_{i+1}^\ell(x,y)$ is increasing in x and decreasing in y, we proceed to showing that $\Theta_i^\ell(x,y)$ is increasing in x and decreasing in y. Since $\lfloor a\rfloor$ and $\lceil a\rceil$ are increasing in a, $\lfloor x-v_i\,r_i\,u_i\rfloor$ and $\lceil y-v_i\,u_i\rceil$ are increasing in x and y, in which case, by the induction hypothesis, $\Theta_{i+1}^\ell(\lfloor x-v_i\,r_i\,u_i\rfloor,\lceil y-v_i\,u_i\rceil)$ is increasing in x and decreasing in y. Thus, for a fixed value of u_i, the objective function of the minimization problem in (10) is increasing in x and decreasing in y. So, the optimal objective value of this minimization problem, which is equal to $\Theta_i^\ell(x,y)$, must be increasing in x and decreasing in y as well. ■

Here is the proof of Lemma 5.3.

Proof of Lemma 5.3: Throughout the proof, let $S\subseteq\{j+1,\dots,\ell\}$ be such that $W(S)\geq x$ and $V(S)\leq y$. Our proof proceeds in three parts.

Part 1: First, assuming that such an assortment S exists, we show that $\Theta_{j+1}^\ell(x,y)<+\infty$. For notational brevity, we let $S^i=S\cap\{i,\dots,\ell\}$. Also, we define $\widetilde{u}_i\in\{0,1\}$ as $\widetilde{u}_i=1$ if and only if

$i \in S$. Since $\widetilde{u}_i$ is a feasible but not necessarily an optimal solution to the minimization problem on the right side of (10) with $x = W(S^i)$ and $y = V(S^i)$, we have

$$\begin{aligned}\Theta^\ell_i(W(S^i), V(S^i)) \;&\le\; c_i\,\widetilde{u}_i + \Theta^\ell_{i+1}(\lfloor W(S^i) - v_i\, r_i\, \widetilde{u}_i\rfloor, \lceil V(S^i) - v_i\,\widetilde{u}_i\rceil)\\ &=\; c_i\,\widetilde{u}_i + \Theta^\ell_{i+1}(\lfloor W(S^{i+1})\rfloor, \lceil V(S^{i+1})\rceil)\\ &\le\; c_i\,\widetilde{u}_i + \Theta^\ell_{i+1}(W(S^{i+1}), V(S^{i+1})),\end{aligned}$$

where the last inequality uses the fact that $\Theta^\ell_{i+1}(x,y)$ is increasing in x and decreasing in y by Lemma J.1, along with the fact that $\lfloor W(S^{i+1})\rfloor \le W(S^{i+1})$ and $\lceil V(S^{i+1})\rceil \ge V(S^{i+1})$.

Since $S \subseteq \{j+1,\ldots,\ell\}$, we have $S^{j+1} = S$ and $S^{\ell+1} = \varnothing$ by the definition of S^i in the previous paragraph. Thus, adding the chain of inequalities above over all $i = j+1,\ldots,\ell$, we obtain

$$\Theta^\ell_{j+1}(W(S), V(S)) \le \sum_{i=j+1}^{\ell} c_i\,\widetilde{u}_i + \Theta^\ell_{\ell+1}(W(\varnothing), V(\varnothing)) = C(S),$$

where the equality uses the fact that $\Theta^\ell_{\ell+1}(0,0) = 0$. Since $W(S) \ge x$, $V(S) \le y$, using Lemma J.1 once again, we get $\Theta^\ell_{j+1}(x,y) \le \Theta^\ell_{j+1}(W(S), V(S)) \le C(S) < +\infty$.

Part 2: Second, we show that $C(\widehat{S}_{x,y}) \le C(S)$. Noting the last chain of inequalities at the end of the previous paragraph, it is enough to show that $\Theta^\ell_{j+1}(x,y) = C(\widehat{S}_{x,y})$.

Consider executing the candidate construction algorithm with $(x,y) \in \textsc{Dom}^2$. By Steps 2 and 3 in the candidate construction algorithm, along with the dynamic program in (10), we have

$$\Theta^\ell_i(\widehat{x}_i, \widehat{y}_i) = c_i\,\widehat{u}_i + \Theta^\ell_{i+1}(\widehat{x}_{i+1}, \widehat{y}_{i+1}).$$

Adding this equality over all $i = j+1,\ldots,\ell$, we get $\Theta^\ell_{j+1}(\widehat{x}_{j+1}, \widehat{y}_{j+1}) = C(\widehat{S}_{x,y}) + \Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1})$. Since we start the candidate construction algorithm with $\widehat{x}_{j+1} = x$ and $\widehat{y}_{j+1} = y$, the last equality yields $\Theta^\ell_{j+1}(x,y) = C(\widehat{S}_{x,y}) + \Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1})$. By Part 1, $\Theta^\ell_{j+1}(x,y) < +\infty$. Also, $\Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1})$ takes the value $+\infty$ or zero. If $\Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1}) = +\infty$, then we get a contradiction to the fact that $\Theta^\ell_{j+1}(x,y) < +\infty$ and $\Theta^\ell_{j+1}(x,y) = C(\widehat{S}_{x,y}) + \Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1})$. Thus, we have $\Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1}) = 0$, in which case, having $\Theta^\ell_{j+1}(x,y) = C(\widehat{S}_{x,y}) + \Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1})$ yields $\Theta^\ell_{j+1}(x,y) = C(\widehat{S}_{x,y})$.

Part 3: Third, we show that $W(\widehat{S}_{x,y}) \ge x/(1+\rho)^n$. Letting $\widehat{S}^i_{x,y} = \widehat{S}_{x,y} \cap \{i,\ldots,\ell\}$, we use induction over the decision epochs to show that $W(\widehat{S}^i_{x,y}) \ge \widehat{x}_i/(1+\rho)^{\ell+1-i}$, where $\widehat{x}_i$ is as in the candidate construction algorithm. By the discussion in the previous paragraph, we have $\Theta^\ell_{\ell+1}(\widehat{x}_{\ell+1}, \widehat{y}_{\ell+1}) = 0$, in which case, by the boundary condition of the dynamic program in (10), we must have $\widehat{x}_{\ell+1} \le 0$. Also, $\widehat{S}^{\ell+1}_{x,y} = \varnothing$. Therefore, we get $W(\widehat{S}^{\ell+1}_{x,y}) = 0 \ge \widehat{x}_{\ell+1}$, so the result holds for decision epoch $\ell+1$. Assuming that $W(\widehat{S}^{i+1}_{x,y}) \ge \widehat{x}_{i+1}/(1+\rho)^{\ell-i}$, we proceed to showing that

$W(\widehat{S}^i_{x,y}) \geq \widehat{x}_i/(1+\rho)^{\ell+1-i}$. Since $\widehat{x}_{i+1} = \lfloor \widehat{x}_i - v_i\, r_i\, \widehat{u}_i \rfloor$, we have $\widehat{x}_{i+1} \geq \frac{1}{1+\rho}\,(\widehat{x}_i - v_i\, r_i\, \widehat{u}_i)$. Noting that $\lfloor a \rfloor = 0$ for $a < 0$, the last inequality holds when $\widehat{x}_i - v_i\, r_i\, \widehat{u}_i < 0$ as well. The last inequality yields $(1+\rho)\,\widehat{x}_{i+1} + v_i\, r_i\, \widehat{u}_i \geq \widehat{x}_i$. Since $\widehat{S}^i_{x,y} \setminus \{i\} = \widehat{S}^{i+1}_{x,y}$ and $\widehat{u}_i = 1$ if and only if $i \in \widehat{S}^i_{x,y}$, we get

$$W(\widehat{S}^i_{x,y}) \;=\; W(\widehat{S}^{i+1}_{x,y}) + v_i\, r_i\, \widehat{u}_i \;\geq\; \frac{\widehat{x}_{i+1}}{(1+\rho)^{\ell-i}} + v_i\, r_i\, \widehat{u}_i$$
$$\geq\; \frac{1}{(1+\rho)^{\ell+1-i}} \left\{ (1+\rho)\,\widehat{x}_{i+1} + v_i\, r_i\, \widehat{u}_i \right\} \;\geq\; \frac{\widehat{x}_i}{(1+\rho)^{\ell+1-i}},$$

where the first inequality uses the induction hypothesis. Thus, the induction argument is complete. Since $\widehat{S}^{j+1}_{x,y} = \widehat{S}_{x,y}$ and $\widehat{x}_{j+1} = x$, we get $W(\widehat{S}_{x,y}) = W(\widehat{S}^{j+1}_{x,y}) \geq \widehat{x}_{j+1}/(1+\rho)^{\ell-j} \geq x/(1+\rho)^n$. Lastly, we can follow a similar argument to also show that $V(\widehat{S}_{x,y}) \leq (1+\rho)^n\, y$. ■

Appendix K: Bounds on the State Variable for Constructing Candidate Assortments

To construct the collection of candidate assortments as in (11), we need the value functions $\Theta^\ell_i(x,y)$ through the dynamic program in (10) for $(x,y) \in \textsc{Dom}^2$ such that $x \in [\lfloor w_{\min} \rfloor, \lceil n\, w_{\max} \rceil] \cup \{0\}$, $y \in [\lfloor v_{\min} \rfloor, \lceil n\, v_{\max} \rceil] \cup \{0\}$, and $i \in \mathcal{N}$, $\ell \in \{0,\ldots,n\}$ with $i \leq \ell + 1$. Therefore, the largest values of x and y in the state variable (x,y) are, respectively, $\lceil n\, w_{\max} \rceil$ and $\lceil n\, v_{\max} \rceil$. Since $\lfloor a - b \rfloor \leq a$ and $\lceil a - b \rceil \leq a$ for $a \in \textsc{Dom}$ and $a, b \in \mathbb{R}_+$, from one decision epoch to another, the values of x and y in the state variable (x,y) in (10) go down. Moreover, the boundary condition in (10) depends only on the sign of x and y. Thus, if the value of the state variable x goes below $\lfloor w_{\min} \rfloor$ but it is still strictly positive, then without loss of generality, we can bump the value of the state variable x up to $\lfloor w_{\min} \rfloor$, because offering any of the products would immediately turn the value of the state variable x to negative. Similarly, if the value of the state variable y goes below $\lfloor v_{\min} \rfloor$ but is still strictly positive, then we can bump the value of the state variable y up to $\lfloor v_{\min} \rfloor$. Lastly, once the value of x and y in the state variable (x,y) turns negative, we do not need to keep their exact values, since each component of the state variable can only go down and the boundary condition at state (x,y) with $y < 0$ always yields a value function of $+\infty$. Thus, the smallest nonzero values of x and y in the state variable $(x,y) \in \textsc{Dom}^2$ are, respectively, $\lfloor w_{\min} \rfloor$ and $\lfloor v_{\min} \rfloor$.

Appendix L: Proof of Lemma 5.5

In this section, we give a proof for Lemma 5.5. We need the next intermedia lemma, where we give two monotonicity properties of the value functions $\{\Psi_k(\ell,u,z) : \ell = 0,\ldots,n,\ (u,z) \in \textsc{Dom}^2,\ k \in \mathcal{M}\}$ computed through the dynamic program in (13). Intuitively speaking, the second one of these properties states that we can compensate for an increase by a factor of $(1+\rho)^2$ in the state variable z by an increase by a factor of $1+\rho$ in the state variable u. This result becomes critical in ultimately proving the performance guarantee of our FPTAS.

Lemma L.1 *If the value functions* $\{\Psi_k(j,u,z) : j = 0,\ldots,n,\ (u,z) \in \textsc{Dom}^2,\ k \in \mathcal{M}\}$ *are computed through the dynamic program in* (13)*, then* $\Psi_k(j,u,z)$ *is increasing in* j*,* u *and* z*. Furthermore, we have* $\Psi_k(j,(1+\rho)\,u,z) \le \Psi_k(j,u,(1+\rho)^2\,z)$*.*

Proof: The fact that $\Psi_k(j,u,z)$ is increasing in j, u and z follows from an induction argument that is similar to the one in the proof of Lemma J.1. To show that $\Psi_k(j,(1+\rho)\,u,z) \le \Psi_k(j,u,(1+\rho)^2\,z)$, we use induction over the decision epochs. Since $\Psi_{m+1}(j,(1+\rho)\,u,z)$ depends only on the sign of z and the signs of z and $(1+\rho)^2\,z$ are the same, we have $\Psi_{m+1}(j,(1+\rho)\,u,z) = \Psi_{m+1}(j,u,(1+\rho)^2\,z)$. Assuming that $\Psi_{k+1}(j,(1+\rho)\,u,z) \le \Psi_{k+1}(j,u,(1+\rho)^2\,z)$, we proceed to showing that $\Psi_k(j,(1+\rho)\,u,z) \le \Psi_k(j,u,(1+\rho)^2\,z)$. We have $(1+\rho)\,u + V(S) \le (1+\rho)\lceil u+V(S)\rceil$. Since $(1+\rho)\lceil u+V(S)\rceil \in \textsc{Dom}$, the last inequality implies that $\lceil (1+\rho)\,u+V(S)\rceil \le (1+\rho)\lceil u+V(S)\rceil$. In this case, we have

$$
\begin{aligned}
&C(S) + \Psi_{k+1}\left(\ell,\, \lceil (1+\rho)\,u+V(S)\rceil,\, \left\lceil z - \frac{\lambda_k\,W(S)}{(1+(1+\rho)\,u)\,(1+(1+\rho)\,u+V(S))}\right\rceil\right)\\
&\qquad \le C(S) + \Psi_{k+1}\left(\ell,\, (1+\rho)\,\lceil u+V(S)\rceil,\, \left\lceil z - \frac{\lambda_k\,W(S)}{(1+(1+\rho)\,u)\,(1+(1+\rho)\,u+V(S))}\right\rceil\right)\\
&\qquad \le C(S) + \Psi_{k+1}\left(\ell,\, \lceil u+V(S)\rceil,\, (1+\rho)^2\left\lceil z - \frac{\lambda_k\,W(S)}{(1+(1+\rho)\,u)\,(1+(1+\rho)\,u+V(S))}\right\rceil\right), \qquad (28)
\end{aligned}
$$

where the first inequality follows from the fact that $\Psi_k(\ell,u,z)$ is increasing in u and the second inequality follows from the induction argument.

Note that $(1+\rho)^2\lceil a\rceil \le \lceil (1+\rho)^2\,a\rceil$. If $a<0$, then the inequality is trivial. For $a \ge 0$, $a \le \frac{1}{(1+\rho)^2}\lceil (1+\rho)^2 a\rceil$. Since $\frac{\lceil (1+\rho)^2 a\rceil}{(1+\rho)^2} \in \textsc{Dom}$, the last inequality yields $\lceil a\rceil \le \frac{1}{(1+\rho)^2}\lceil (1+\rho)^2 a\rceil$. So,

$$
\begin{aligned}
&(1+\rho)^2\left\lceil z - \frac{\lambda_k\,W(S)}{(1+(1+\rho)u)\,(1+(1+\rho)u+V(S))}\right\rceil\\
&\qquad \le \left\lceil (1+\rho)^2 z - \frac{(1+\rho)^2\,\lambda_k\,W(S)}{(1+(1+\rho)u)\,(1+(1+\rho)u+V(S))}\right\rceil \le \left\lceil (1+\rho)^2 z - \frac{\lambda_k\,W(S)}{(1+u)\,(1+u+V(S))}\right\rceil,
\end{aligned}
$$

where the second inequality uses the fact that $\lceil a\rceil$ is increasing in a. Note that $\lceil a\rceil$ is increasing in a even with the convention that $\lceil a\rceil = -\infty$ for $a<0$.

Using the chain of inequalities above and the fact that $\Psi_{k+1}(j,u,z)$ is increasing in z, we can bound the expression on the right side of (28) as

$$
\begin{aligned}
&C(S) + \Psi_{k+1}\left(\ell,\, \lceil u+V(S)\rceil,\, (1+\rho)^2\left\lceil z - \frac{\lambda_k\,W(S)}{(1+(1+\rho)\,u)\,(1+(1+\rho)\,u+V(S))}\right\rceil\right)\\
&\qquad\qquad \le C(S) + \Psi_{k+1}\left(\ell,\, \lceil u+V(S)\rceil,\, \left\lceil (1+\rho)^2 z - \frac{\lambda_k\,W(S)}{(1+u)\,(1+u+V(S))}\right\rceil\right). \qquad (29)
\end{aligned}
$$

By (28) and (29), we have $C(S) + \Psi_{k+1}\big(\ell,\, \lceil (1+\rho)\,u+V(S)\rceil,\, \lceil z - \frac{\lambda_k\,W(S)}{(1+(1+\rho)\,u)\,(1+(1+\rho)\,u+V(S))}\rceil\big) \le C(S) + \Psi_{k+1}\big(\ell,\, \lceil u+V(S)\rceil,\, \lceil (1+\rho)^2 z - \frac{\lambda_k\,W(S)}{(1+u)\,(1+u+V(S))}\rceil\big)$ for all S and ℓ. In this case, minimizing both sides of the inequality over (ℓ, S) with $\ell \geq j$ and $S \in \textsc{Cand}(j,\ell)$, the inequality is still preserved, but noting (13), the left side of the inequality gives $\Psi_k(j,(1+\rho)\,u,z)$, whereas the right side gives $\Psi_k(j,u,(1+\rho)^2\,z)$. Thus, we have $\Psi_k(\ell,(1+\rho)\,u,z) \leq \Psi_k(\ell,u,(1+\rho)^2\,z)$. ■

Next, we give a proof for Lemma 5.5.

Proof of Lemma 5.5: Let $(\widehat{S}_1,\ldots,\widehat{S}_m)$ be the output of the candidate stitching algorithm, $\widetilde{z}$ be the optimal objective value of problem (9), and $\widehat{z}_{\textsc{App}} = \max\{z \in \textsc{Dom} : \Psi_1(0,0,z) \leq b\}$. Our proof proceeds in three parts.

Part 1: First, we show that $\sum_{k\in\mathcal{M}} C(\widehat{S}_k) \leq b$. Noting Steps 1 and 2 in the candidate stitching algorithm, along with the dynamic program in (13), we have $\Psi_k(\widehat{j}_k,\widehat{u}_k,\widehat{z}_k) = C(\widehat{S}_k) + \Psi_{k+1}(\widehat{j}_{k+1},\widehat{u}_{k+1},\widehat{z}_{k+1})$. Adding this equality over all $k\in\mathcal{M}$ and noting that we start the candidate stitching algorithm with $\widehat{j}_1 = 0$, $\widehat{u}_1 = 0$ and $\widehat{z}_1 = \widehat{z}_{\textsc{App}}$, we obtain $\Psi_1(0,0,\widehat{z}_{\textsc{App}}) = \sum_{k\in\mathcal{M}} C(\widehat{S}_k) + \Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1})$. By the initialization of candidate stitching, we have $\Psi_1(0,0,\widehat{z}_{\textsc{App}}) \leq b$, in which case, the last equality implies that $\sum_{k\in\mathcal{M}} C(\widehat{S}_k) + \Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1}) \leq b$. By the boundary condition of the dynamic program in (13), $\Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1})$ takes the value $+\infty$ or zero. If we have $\Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1}) = +\infty$, then we get a contradiction to the fact that $\sum_{k\in\mathcal{M}} C(\widehat{S}_k) + \Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1}) \leq b$. Thus, we must have $\Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1}) = 0$, so having $\sum_{k\in\mathcal{M}} C(\widehat{S}_k) + \Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1}) \leq b$ implies that $\sum_{k\in\mathcal{M}} C(\widehat{S}_k) \leq b$.

Part 2: Second, we show that $\textsc{Rev}(\widehat{S}_1,\ldots,\widehat{S}_m) \geq \widehat{z}_{\textsc{App}}$. By Step 2 of the candidate stitching algorithm, we have $\widehat{u}_{k+1} \geq \widehat{u}_k + V(\widehat{S}_k)$. Adding this inequality over all $k = 1,\ldots,q-1$ and noting that $\widehat{u}_1 = 0$ in the initialization of the algorithm, we get $\widehat{u}_q \geq \sum_{k=1}^{q-1} V(\widehat{S}_k)$. For notational brevity, we let $\widehat{R}_k = \sum_{q=k}^m \frac{\lambda_q\, W(\widehat{S}_q)}{(1+\sum_{r=1}^{q-1} V(\widehat{S}_r))\,(1+\sum_{r=1}^{q} V(\widehat{S}_r))}$ with the convention that $\widehat{R}_{m+1} = 0$. We use induction over the stages to show that $\widehat{R}_k \geq \widehat{z}_k$ for all $k = 1,\ldots,m+1$. By the discussion in the previous paragraph, $\Psi_{m+1}(\widehat{j}_{m+1},\widehat{u}_{m+1},\widehat{z}_{m+1}) = 0$, in which case, by the boundary condition in (13), we must have $\widehat{z}_{m+1} \leq 0$. Thus, we have $\widehat{R}_{m+1} = 0 \geq \widehat{z}_{m+1}$. Assuming that $\widehat{R}_{k+1} \geq \widehat{z}_{k+1}$, we proceed to showing that $\widehat{R}_k \geq \widehat{z}_k$. Noting Step 2 of the candidate stitching algorithm and using the induction hypothesis, if $\widehat{z}_{k+1} \geq 0$, then $\widehat{z}_k - \frac{\lambda_k\, W(\widehat{S}_k)}{(1+\widehat{u}_k)\,(1+\widehat{u}_k+V(\widehat{S}_k))} \leq \widehat{z}_{k+1} \leq \widehat{R}_{k+1}$. Also, if $\widehat{z}_{k+1} < 0$, then $\widehat{z}_k - \frac{\lambda_k\, W(\widehat{S}_k)}{(1+\widehat{u}_k)\,(1+\widehat{u}_k+V(\widehat{S}_k))} < 0 \leq \widehat{R}_{k+1}$. So, $\widehat{z}_k \leq \widehat{R}_{k+1} + \frac{\lambda_k\, W(\widehat{S}_k)}{(1+\widehat{u}_k)\,(1+\widehat{u}_k+V(\widehat{S}_k))}$ in both cases. Thus, we get

$$\widehat{R}_k \;=\; \widehat{R}_{k+1} + \frac{\lambda_k\, W(\widehat{S}_k)}{(1+\sum_{q=1}^{k-1} V(\widehat{S}_q))\,(1+\sum_{q=1}^{k} V(\widehat{S}_q))} \;\geq\; \widehat{R}_{k+1} + \frac{\lambda_k\, W(\widehat{S}_k)}{(1+\widehat{u}_k)\,(1+\widehat{u}_k+V(\widehat{S}_k))} \;\geq\; \widehat{z}_k,$$

where we use the fact that $\widehat{u}_k \geq \sum_{q=1}^{k-1} V(\widehat{S}_q)$. The induction argument is complete, in which case, we have $\widehat{R}_1 \geq \widehat{z}_1$. Noting that $\widehat{R}_1 = \textsc{Rev}(\widehat{S}_1,\ldots,\widehat{S}_m)$ and $\widehat{z}_1 = \widehat{z}_{\textsc{App}}$, the result follows.

Part 3: Third, we show that $\widehat{z}_{\textsc{App}} \geq \widetilde{z}/(1+\rho)^{3m+1}$. Let $(\widetilde{S}_1,\ldots,\widetilde{S}_m,\widetilde{j}_1,\ldots,\widetilde{j}_m)$ be an optimal solution to problem (9). For notational brevity, we let $\widetilde{C}_k = \sum_{q=k}^m C(\widetilde{S}_q)$, $\widetilde{u}_k = \sum_{q=1}^{k-1} V(\widetilde{S}_q)$ and

$\widetilde{z}_k = \sum_{q=k}^m \frac{\lambda_q W(\widetilde{S}_q)}{(1+\widetilde{u}_q)(1+\widetilde{u}_{q+1})}$ with the convention that $\widetilde{C}_{m+1} = 0$, $\widetilde{u}_1 = 0$ and $\widetilde{z}_{m+1} = 0$. We use induction over the stages to show that $\Psi_k(\widetilde{j}_k, \widetilde{u}_k, \widetilde{z}_k/(1+\rho)^{3(m+1-k)}) \leq \widetilde{C}_k$. We have $\widetilde{z}_{m+1} = 0$ and $\widetilde{C}_{m+1} = 0$, in which case, noting the boundary condition in (13), we have $\Psi_{m+1}(\widetilde{j}_{m+1}, \widetilde{u}_{m+1}, \widetilde{z}_{m+1}) = \Psi_{m+1}(\widetilde{j}_{m+1}, \widetilde{u}_{m+1}, 0) = 0 = \widetilde{C}_{m+1}$. Assuming that $\Psi_{k+1}(\widetilde{j}_{k+1}, \widetilde{u}_{k+1}, \widetilde{z}_{k+1}/(1+\rho)^{3(m-k)}) \leq \widetilde{C}_{k+1}$, we proceed to showing that $\Psi_k(\widetilde{j}_k, \widetilde{u}_k, \widetilde{z}_k/(1+\rho)^{3(m+1-k)}) \leq \widetilde{C}_k$. We have

$$
\begin{aligned}
&(1+\rho)^2 \left\lceil \frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}} - \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} \right\rceil \\
&\leq (1+\rho)^3 \left(\frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}} - \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} \right) \\
&\overset{(a)}{=} \frac{\widetilde{z}_{k+1}}{(1+\rho)^{3(m-k)}} + \frac{\lambda_k W(\widetilde{S}_k)}{(1+\rho)^{3(m-k)}(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} - \frac{(1+\rho)^3 \lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} \overset{(b)}{\leq} \frac{\widetilde{z}_{k+1}}{(1+\rho)^{3(m-k)}}, \quad (30)
\end{aligned}
$$

where (a) follows from the fact that $\widetilde{z}_k = \widetilde{z}_{k+1} + \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})}$ by the definition of $\widetilde{z}_k$ and (b) holds because we have $k \leq m$.

In (13), the action $(\widetilde{j}_{k+1}, \widetilde{S}_k)$ is feasible when the state of the system at decision epoch k is $(\widetilde{j}_k, \widetilde{u}_k, \widetilde{z}_k/(1+\rho)^{3(m+1-k)})$. In particular, since $(\widetilde{S}_1, \ldots, \widetilde{S}_m, \widetilde{j}_1, \ldots, \widetilde{j}_m)$ is a feasible solution to problem (9), we have $\widetilde{j}_{k+1} \geq \widetilde{j}_k$ and $\widetilde{S}_k \in \textsc{Cand}(\widetilde{j}_k, \widetilde{j}_{k+1})$. Since, the action $(\widetilde{j}_{k+1}, \widetilde{S}_k)$ is feasible to the minimization problem in (13) with $(j, u, z) = (\widetilde{j}_k, \widetilde{u}_k, \widetilde{z}_k/(1+\rho)^{3(m+1-k)})$, we get

$$
\begin{aligned}
&\Psi_k\left(\widetilde{j}_k, \widetilde{u}_k, \frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}}\right) \\
&\quad \leq C(\widetilde{S}_k) + \Psi_{k+1}\left(\widetilde{j}_{k+1}, \lceil \widetilde{u}_k + V(\widetilde{S}_k) \rceil, \left\lceil \frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}} - \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_k + V(\widetilde{S}_k))} \right\rceil \right) \\
&\quad = C(\widetilde{S}_k) + \Psi_{k+1}\left(\widetilde{j}_{k+1}, \lceil \widetilde{u}_{k+1} \rceil, \left\lceil \frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}} - \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} \right\rceil \right) \\
&\quad \overset{(c)}{\leq} C(\widetilde{S}_k) + \Psi_{k+1}\left(\widetilde{j}_{k+1}, (1+\rho)\,\widetilde{u}_{k+1}, \left\lceil \frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}} - \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} \right\rceil \right) \\
&\quad \overset{(d)}{\leq} C(\widetilde{S}_k) + \Psi_{k+1}\left(\widetilde{j}_{k+1}, \widetilde{u}_{k+1}, (1+\rho)^2 \left\lceil \frac{\widetilde{z}_k}{(1+\rho)^{3(m+1-k)}} - \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} \right\rceil \right) \\
&\quad \overset{(e)}{\leq} C(\widetilde{S}_k) + \Psi_{k+1}\left(\widetilde{j}_{k+1}, \widetilde{u}_{k+1}, \frac{\widetilde{z}_{k+1}}{(1+\rho)^{3(m-k)}}\right) \\
&\quad \overset{(f)}{\leq} C(\widetilde{S}_k) + \widetilde{C}_{k+1} = \widetilde{C}_k,
\end{aligned}
$$

where (c) follows from the fact that $\Psi_k(\ell, u, z)$ is increasing in u and $(1+\rho)\,u \geq \lceil u \rceil$, (d) follows by the second part of Lemma L.1, (e) follows by noting the fact that $\Psi_k(j, u, z)$ is increasing in z and using the inequality in (30), and (f) is by the induction hypothesis. Thus, the induction argument is complete, so it follows that $\Psi_k(\widetilde{j}_k, \widetilde{u}_k, \widetilde{z}_k/(1+\rho)^{3(m+1-k)}) \leq \widetilde{C}_k$.

By the definition of $\widetilde{z}_k$ and $\widetilde{u}_k$, $\widetilde{z}_1 = \sum_{k \in \mathcal{M}} \frac{\lambda_k W(\widetilde{S}_k)}{(1+\widetilde{u}_k)(1+\widetilde{u}_{k+1})} = \sum_{k \in \mathcal{M}} \frac{\lambda_k W(\widetilde{S}_k)}{(1+\sum_{q=1}^{k-1} V(\widetilde{S}_q))(1+\sum_{q=1}^{k} V(\widetilde{S}_q))} = \textsc{Rev}(\widetilde{S}_1, \ldots, \widetilde{S}_m) = \widetilde{z}$, where the last equality uses the fact that $(\widetilde{S}_1, \ldots, \widetilde{S}_m, \widetilde{j}_1, \ldots, \widetilde{j}_m)$ is an optimal solution to problem (9), so $\widetilde{z}_1 = \widetilde{z}$. Thus, using the inequality $\Psi_k(\widetilde{j}_k, \widetilde{u}_k, \widetilde{z}_k/(1+\rho)^{3(m+1-k)}) \leq \widetilde{C}_k$ with $k=1$, we get $\Psi_1(\widetilde{j}_1, 0, \widetilde{z}/(1+\rho)^{3m}) \leq \widetilde{C}_1 \leq b$, where the last inequality uses the fact that $(\widetilde{S}_1, \ldots, \widetilde{S}_m, \widetilde{j}_1, \ldots, \widetilde{j}_m)$ is an optimal solution to problem (9) so that $\widetilde{C}_1 = \sum_{k\in\mathcal{M}} C(\widetilde{S}_k) \leq b$. Since $\Psi_k(j,u,z)$ is increasing in j and z by Lemma L.1, we obtain

$$\Psi_1(0,0,\lfloor\widetilde{z}\rfloor/(1+\rho)^{3m}) \leq \Psi_1(\widetilde{j}_1, 0, \widetilde{z}/(1+\rho)^{3m}) \leq b,$$

which implies that $\lfloor\widetilde{z}\rfloor/(1+\rho)^{3m} \in \textsc{Dom}$ is a feasible solution to the problem $\widehat{z}_{\text{APP}} = \max\{z \in \textsc{Dom} : \Psi_1(0,0,z) \leq b\}$. Therefore, $\widehat{z}_{\text{APP}} \geq \lfloor\widetilde{z}\rfloor/(1+\rho)^{3m} \geq \widetilde{z}/(1+\rho)^{3m+1}$. ■

Appendix M: Bound on the State Variable for Combining Candidate Assortments

To solve the dynamic program in (13), we argue that that the largest values of u and z that we need to consider in the state variable $(j,u,z) \in \mathcal{N} \times \textsc{Dom}^2$ are, respectively, $\lceil 2n\, v_{\max}\rceil$ and $\lceil n w_{\max}\rceil$. Similarly, the smallest nonzero values of u and z that we need to consider in the state variable $(j,u,z) \in \mathcal{N} \times \textsc{Dom}^2$ are, respectively, $\lfloor v_{\min}\rfloor$ and $\lfloor \lambda_m \frac{w_{\min}}{(1+2n\, v_{\max})^2}\rfloor$. In particular, a simple lemma, given as Lemma M.1 at the end of this section, shows that if we compute $\{\widehat{u}_k : k = 1, \ldots, m+1\}$ as $\widehat{u}_{k+1} = \lceil \widehat{u}_k + V(S_k)\rceil$ with $\widehat{u}_1 = 0$ and $S_k \cap S_q = \varnothing$ for all $k \neq q$, then $\widehat{u}_k \leq 2n\, v_{\max}$ for all $k \in \mathcal{M}$. Thus, the value of u in the state variable (j,u,z) in the dynamic program in (13) is at most $\lceil 2n\, v_{\max}\rceil$. A strictly positive value of u in the state variable (j,u,z) is at least $\lfloor v_{\min}\rfloor$, as the initial value of this state variable is zero and the preference weight of any product is at least $v_{\min}$. Therefore, the desired upper and lower bounds for u in the state variable (j,u,z) follow.

If the initial state variable (j,u,z) satisfies $z > n\, w_{\max}$, then since $\sum_{k\in\mathcal{M}} W(S_k) \leq n\, w_{\max}$ for any $(S_1, \ldots, S_m)$ with $S_k \cap S_q = \varnothing$ for all $k \neq q$, no matter which assortments we offer, the final state variable (j,u,z) satisfies $z > 0$, in which case the value function $\Psi_1(0,0,z)$ takes the value $+\infty$. Thus, we do not need to consider the values of z that exceed $n\, w_{\max}$ in the state variable (j,u,z). So, we can assume that the value of z in the state variable (j,u,z) is at most $\lceil n\, w_{\max}\rceil$. Finally, if the value of z in the state variable goes below $\lfloor \lambda_m\, w_{\min}/(1+2n\, v_{\max})^2\rfloor$ but is still strictly positive, then without of loss generality, we can bump the value of z up to $\lfloor \lambda_m\, w_{\min}/(1+2n\, v_{\max})^2\rfloor$, since offering any nonempty candidate assortment would immediately turn the value of the state variable to negative. Therefore, it follows that we can assume that a strictly positive value of z in the state variable (j,u,z) is at least $\lfloor \lambda_m\, w_{\min}/(1+2n\, v_{\max})^2\rfloor$.

We used the next lemma in our discussion earlier in this section. Recall that we choose the accuracy parameter for the geometric grid as $\rho = \frac{1}{8(3m+1)}\epsilon$ for $\epsilon \in (0,1)$, so $\rho \leq \frac{1}{2m}$.

Lemma M.1 *For $\rho \leq \frac{1}{2m}$, if we compute $\{\widehat{u}_k : k = 1, \ldots, m+1\}$ as $\widehat{u}_{k+1} = \lceil \widehat{u}_k + V(S_k)\rceil$ with $\widehat{u}_1 = 0$ and $S_k \cap S_q = \varnothing$ for all $k \neq q$, then $\widehat{u}_{m+1} \leq 2n\, v_{\max}$.*

Proof: We use induction to show that $\widehat{u}_k \leq (1+\rho)^{k-1}\,(V(S_1)+\ldots+V(S_{k-1}))$. For $k=1$, we have $\widehat{u}_1 = 0$. Therefore, the result holds for $k=1$. Assuming that $\widehat{u}_k \leq (1+\rho)^{k-1}\,(V(S_1)+\ldots+V(S_{k-1}))$, we proceed to showing that $\widehat{u}_{k+1} \leq (1+\rho)^k\,(V(S_1)+\ldots+V(S_k))$. We have

$$\begin{aligned}\widehat{u}_{k+1} \;=\; \lceil \widehat{u}_k + V(S_k) \rceil \;\leq\; (1+\rho)\,(\widehat{u}_k + V(S_k)) \\ \leq\; (1+\rho)\Big((1+\rho)^{k-1}\,(V(S_1)+\ldots+V(S_{k-1})) + V(S_k)\Big) \\ \leq\; (1+\rho)^k (V(S_1)+\ldots+V(S_{k-1})+V(S_k)),\end{aligned}$$

which completes the induction argument. Thus, we have $\widehat{u}_{m+1} \leq (1+\rho)^m (V(S_1)+\ldots+V(S_m)) \leq (1+\rho)^m\, n v_{\max}$. In this case, the result follows because $(1+\rho)^m \leq \left(1+\frac{1}{2m}\right)^m \leq \exp(1/2) \leq 2$. ∎

Appendix N: Assortment Optimization under a Cardinality Constraint

In this section, we consider a version of the CAPACITATED problem, where each product occupies one unit of space. Therefore, we can express the constraint $\sum_{k\in\mathcal{M}} C(S_k) \leq b$ as $\sum_{k\in\mathcal{M}} |S_k| \leq b$, in which case, we ensure that the total number of products offered over all stages does not exceed b. Note that b is an integer without loss of generality. Otherwise, we can round it down to the nearest integer. In this section, we give three results. First, we give an algorithm that finds an exact solution. The running time of this algorithm is polynomial in the number of products, but exponential in the number of stages. Second, we give a pseudo polynomial-time algorithm that finds an exact solution. Assuming that the preference weight of the products take on integer values, the running time of this algorithm is polynomial in the number of products, number of stages, and $v_{\max}$. Third, we give an FPTAS to get a $(1-\epsilon)$-approximate solution, whose running time is polynomial in all of the input parameters and $1/\epsilon$. Next, we go into the details of each of these results, compare them with each other and explain their common components.

First, we show that we can obtain an optimal solution by checking the expected revenue from $O(b^m\, n^{3m-1})$ possible solutions. The running time of this approach is polynomial in the number of products for a fixed number of stages. In general, since each one of as many as b products in an optimal solution can be offered in one of the m stages, the number of all possible solutions to the CAPACITATED problem under a cardinality constraint is $O(\binom{n}{b}\, b^m) = O(n^b\, b^m)$, which is exponential in the number of products even for a fixed number of stages. Second, treating the preference weights as the problem input, if all of the preference weights take on integer values, then we give a pseudo polynomial-time algorithm that obtains an optimal solution in $O(v_{\max} m n^5 b^2)$ operations. This algorithm is based on a dynamic programming formulation of the problem. If the preference weights take on rational values, then we can ensure that the preference weights take on integer values by scaling all of the preference weights by a constant, since the choice probabilities do not change by doing so. Third, by discretizing the state variable in the dynamic program that we use in the pseudo polynomial-time algorithm through a geometric grid, we obtain an FPTAS. Our FPTAS obtains a $(1-\epsilon)$-approximate solution in $O(m^2\, n^4\, b^2 \log(n v_{\max}/v_{\min})/\epsilon)$ operations. All of these three results, the exact algorithm whose running time is exponential in the number of stages, the pseudo polynomial-time algorithm and the FPTAS, are based on constructing a collection of candidate assortments for each stage so that an optimal assortment to offer in a stage lies within this collection. Therefore, we start by focusing on constructing the collections of candidate assortments for the different stages. Throughout this section, when we refer to the Capacitated problem, we refer to the version where each product occupies one unit of space, so we have a constraint on the number of offered products.

Constructing Collections of Candidate Assortments:

Note that Lemma 5.1 continues to hold when each product occupies one unit of space. Thus, there exists an optimal solution $(S_1^*, \ldots, S_m^*)$ such that $S_k^* \subseteq \{j_k^*+1, \ldots, j_{k+1}^*\}$ and $S_q^* \cap \{j_k^*+1, \ldots, j_{k+1}^*\} = \varnothing$ for all $q \neq k$, for some $j_1^*, \ldots, j_{m+1}^*$ that satisfy $0 = j_1^* \leq j_2^* \leq \ldots \leq j_m^* \leq j_{m+1}^* = n$. To construct the collection of candidate assortments for stage k, we proceed under the assumption that we know the values of j_k^*, j_{k+1}^*, $|\cup_{q\neq k} S_q^*|$ along with $V(S_q^*)$ and $W(S_q^*)$ for all $q \neq k$. In this case, since the assortment that we offer in stage k affects the expected revenue in stages $k, \ldots, m$, we can recover an optimal assortment to offer in stage k by solving

$$\max_{\substack{S \subseteq \{j_k^*+1, \ldots, j_{k+1}^*\},\\ |S| \leq b - |\cup_{q\neq k} S_q^*|}} \left\{ \frac{\lambda_k\, W(S)}{(1+\sum_{q=1}^{k-1} V(S_q^*))\,(1+\sum_{q=1}^{k-1} V(S_q^*)+V(S))} + \sum_{\ell=k+1}^{m} \frac{\lambda_\ell\, W(S_\ell^*)}{(1+\sum_{q=1,q\neq k}^{\ell-1} V(S_q^*)+V(S))\,(1+\sum_{q=1,q\neq k}^{\ell} V(S_q^*)+V(S))} \right\},$$

where we use the fact that if we know the value of $|\cup_{q\neq k} S_q^*|$, then we can offer at most $b - |\cup_{q\neq k} S_q^*|$ products in stage k.

For notational brevity, we let $b_k^* = b - |\cup_{q\neq k} S_q^*|$, $f_\ell^* = \lambda_\ell\, W(S_\ell^*)/V(S_\ell^*)$ and $u_\ell^* = \sum_{q=1,q\neq k}^{\ell} V(S_q^*)$. We write the objective function of the problem above as

$$\begin{aligned} &\frac{\lambda_k\, W(S)}{(1+u_{k-1}^*)\,(1+u_{k-1}^*+V(S))} + \sum_{\ell=k+1}^{m} f_\ell^* \left\{ \frac{1}{1+u_{\ell-1}^*+V(S)} - \frac{1}{1+u_\ell^*+V(S)} \right\} \\ &\quad = \frac{\lambda_k\, W(S)}{(1+u_{k-1}^*)\,(1+u_{k-1}^*+V(S))} + \sum_{\ell=k+1}^{m} (f_\ell^* - f_{\ell+1}^*) \left\{ \frac{1}{1+u_{k-1}^*+V(S)} - \frac{1}{1+u_\ell^*+V(S)} \right\} \end{aligned}$$

with the convention that $f_{m+1}^* = 0$. The equality above follows by noting that the sum on the left side of the equality is equivalent to $f_{k+1}^* \frac{1}{1+u_k^*+V(S)} + \sum_{\ell=k+1}^{m} (f_{\ell+1}^* - f_\ell^*) \frac{1}{1+u_\ell^*+V(S)} =$

$\sum_{\ell=k+1}^{m}(f_\ell^* - f_{\ell+1}^*)\frac{1}{1+u_k^*+V(S)} - \sum_{\ell=k+1}^{m}(f_\ell^* - f_{\ell+1}^*)\frac{1}{1+u_\ell^*+V(S)}$, along with the fact that $u_k^* = u_{k-1}^*$. In this case, to recover an optimal assortment to offer in stage k, we can solve the problem

$$\max_{\substack{S \subseteq \{j_k^*+1,\ldots,j_{k+1}^*\},\\ |S| \le b_k^*}} \left\{ \frac{\lambda_k\, W(S)}{(1+u_{k-1}^*)\,(1+u_{k-1}^*+V(S))} + \sum_{\ell=k+1}^{m}(f_\ell^* - f_{\ell+1}^*)\left\{\frac{1}{1+u_{k-1}^*+V(S)} - \frac{1}{1+u_\ell^*+V(S)}\right\}\right\}. \quad (31)$$

In the next lemma, we show that we can efficiently construct a collection of candidate assortments that includes an optimal solution to problem (31) for any values of $\{(f_\ell^*, u_\ell^*) : \ell \in \mathcal{M},\ \ell \neq k\}$.

Lemma N.1 *Given j_k^*, j_{k+1}^* and b_k^*, there exists a collection of candidate assortments $\textsc{Cand}_k(j_k^*, j_{k+1}^*, b_k^*)$ with $|\textsc{Cand}_k(j_k^*, j_{k+1}^*, b_k^*)| = O(n^2)$ that includes an optimal solution to problem (31) for any values of $\{(f_\ell^*, u_\ell^*) : \ell \in \mathcal{M},\ \ell \neq k\}$.*

Proof: Let $g_\ell^* = (f_\ell^* - f_{\ell+1}^*)\,(u_\ell^* - u_{k-1}^*)$. Multiplying the objective function of problem (31) by the constant $1+u_{k-1}^*$, we can obtain an optimal assortment to offer in stage k by solving

$$\max_{\substack{S \subseteq \{j_k^*+1,\ldots,j_{k+1}^*\},\\ |S| \le b_k^*}} \left\{ \frac{1}{1+u_{k-1}^*+V(S)}\left\{\lambda_k\, W(S) + (1+u_{k-1}^*)\sum_{\ell=k+1}^{m}\frac{g_\ell^*}{1+u_\ell^*+V(S)}\right\}\right\}.$$

Letting t^* be the optimal objective value of the problem above, t^* is no smaller than the objective function of the problem above at each S such that $S \subseteq \{j_k^*+1,\ldots,j_{k+1}^*\}$ and $|S| \le b_k^*$.

Therefore, letting $\mathcal{G} = \{S \subseteq \{j_k^*+1,\ldots,j_{k+1}^*\} : |S| \le b_k^*\}$, we can obtain an optimal solution to the problem above by using the so-called dual formulation, which is given by

$$\begin{aligned}
&\min\left\{t \;:\; t \;\ge\; \frac{1}{1+u_{k-1}^*+V(S)}\left\{\lambda_k\, W(S) + (1+u_{k-1}^*)\sum_{\ell=k+1}^{m}\frac{g_\ell^*}{1+u_\ell^*+V(S)}\right\}\quad \forall\, S \in \mathcal{G}\right\}\\
&= \min\left\{t \;:\; t \;\ge\; \frac{\lambda_k\, W(S)}{1+u_{k-1}^*} - \frac{t\,V(S)}{1+u_{k-1}^*} + \sum_{\ell=k+1}^{m}\frac{g_\ell^*}{1+u_\ell^*+V(S)}\quad \forall\, S \in \mathcal{G}\right\}\\
&= \min\left\{t \;:\; t \;\ge\; \max_{S\in\mathcal{G}}\left\{\frac{\lambda_k\, W(S)}{1+u_{k-1}^*} - \frac{t\,V(S)}{1+u_{k-1}^*} + \sum_{\ell=k+1}^{m}\frac{g_\ell^*}{1+u_\ell^*+V(S)}\right\}\right\}
\end{aligned}$$

where the first equality follows by multiplying both sides of the constraint in the first minimization problem above by $1+u_{k-1}^*+V(S)$ and arranging the terms.

By the discussion so far, if t^* is an optimal solution to the last minimization problem above, then we can recover an optimal assortment to offer in stage k by replacing t in the maximization problem on the right side of the constraint with t^* and solving this maximization problem. Thus, we can obtain an optimal assortment to offer in stage k by solving the problem

$$\max_{S\in\mathcal{G}}\left\{\frac{\lambda_k\,W(S)}{1+u^*_{k-1}}-\frac{t\,V(S)}{1+u^*_{k-1}}+\sum_{\ell=k+1}^{m}\frac{g^*_\ell}{1+u^*_\ell+V(S)}\right\} \tag{32}$$

for some value of t. We will construct a collection of $O(n^2)$ candidate assortments that includes an optimal solution to the problem above for any values of $\{(g^*_\ell,u^*_\ell):\ell\in\mathcal{M},\ \ell\neq k\}$ and t.

Note that $\lambda_k W(S)=\lambda_k\sum_{i\in S}r_i\,v_i$ and $t\,V(S)=t\sum_{i\in S}v_i$. In this case, using the decision variables $\boldsymbol{x}=(x_1,\ldots,x_n)$ and noting the definition of $\mathcal{G}$, we write problem (32) equivalently as

$$\max_{\boldsymbol{x}\in\{0,1\}^n}\left\{\frac{\lambda_k}{1+u^*_{k-1}}\sum_{i\in\mathcal{N}}r_i\,v_i\,x_i-\frac{t}{1+u^*_{k-1}}\sum_{i\in\mathcal{N}}v_i\,x_i+\sum_{\ell=k+1}^{m}\frac{g^*_\ell}{1+u^*_\ell+\sum_{i\in\mathcal{N}}v_i\,x_i}\right.$$
$$\left.:\ \sum_{i\in\mathcal{N}}x_i\le b^*,\ x_i=0\ \ \forall\, i\notin\{j^*_k+1,\ldots,j^*_{k+1}\}\right\}. \tag{33}$$

If $g^*_\ell\ge 0$, then the objective function of the problem above is convex in $\boldsymbol{x}$, in which case, an optimal solution occurs at an extreme point, so we can relax $\boldsymbol{x}\in\{0,1\}^n$ to $\boldsymbol{x}\in[0,1]^n$.

Indeed, we have $g^*_\ell\ge 0$. Note that $W(S^*_\ell)/V(S^*_\ell)$ is the weighted average of the revenues of the products in S^*_ℓ. By Lemma 5.1, the revenues of the products in S^*_ℓ are larger than those of the products in $S^*_{\ell+1}$, so we have $W(S^*_\ell)/V(S^*_\ell)\ge W(S^*_{\ell+1})/V(S^*_{\ell+1})$. Furthermore, we have $\lambda_\ell\ge\lambda_{\ell+1}$, in which case, we get $f^*_\ell=\lambda_\ell\,W(S^*_\ell)/V(S^*_\ell)\ge\lambda_{\ell+1}\,W(S^*_{\ell+1})/V(S^*_{\ell+1})=f^*_{\ell+1}$. We have $u^*_\ell\ge u^*_{k-1}$ for all $\ell\ge k+1$ as well, so $g^*_\ell=(f^*_\ell-f^*_{\ell+1})\,(u^*_\ell-u^*_{k-1})\ge 0$. We solve problem (33) with $\boldsymbol{x}\in[0,1]^n$ in two stages. First, intuitively speaking, we guess the value of $\sum_{i\in\mathcal{N}}v_i\,x_i$. Second, we find solution $\boldsymbol{x}$ that maximizes the objective function, while satisfying our guess.

Using w to denote our guess of $\sum_{i\in\mathcal{N}}v_i\,x_i$, we can write the last problem in two stages. In particular, problem (33) is equivalent to the problem

$$\max_{w\in\mathbb{R}_+}\max_{\boldsymbol{x}\in[0,1]^n}\left\{\frac{\lambda_k}{1+u^*_{k-1}}\sum_{i\in\mathcal{N}}r_i\,v_i\,x_i-\frac{t\,w}{1+u^*_{k-1}}+\sum_{\ell=k+1}^{m}\frac{g^*_\ell}{1+u^*_\ell+w}\right.$$
$$\left.:\ \sum_{i\in\mathcal{N}}x_i\le b^*,\quad\sum_{i\in\mathcal{N}}v_i\,x_i\le w,\quad x_i=0\ \ \forall\, i\notin\{j^*_k+1,\ldots,j^*_{k+1}\}\right\}$$
$$=\ \max_{w\in\mathbb{R}_+}\left\{-\frac{t\,w}{1+u^*_{k-1}}+\sum_{\ell=k+1}^{m}\frac{g^*_\ell}{1+u^*_\ell+w}+\frac{\lambda_k}{1+u^*_{k-1}}\max_{\boldsymbol{x}\in[0,1]^n}\left\{\sum_{i\in\mathcal{N}}r_i\,v_i\,x_i\right.\right.$$
$$\left.\left.:\ \sum_{i\in\mathcal{N}}x_i\le b^*,\quad\sum_{i\in\mathcal{N}}v_i\,x_i\le w,\quad x_i=0\ \ \forall\, i\notin\{j^*_k+1,\ldots,j^*_{k+1}\}\right\}\right\}. \tag{34}$$

The first problem above is equivalent to problem (33) since $g^*_\ell\ge 0$, in which case, the objective function of the first problem above is decreasing in w. Therefore, w takes the value $\sum_{i\in\mathcal{N}}v_i\,x_i$

in an optimal solution to the first problem above. Considering the second problem above, the inner maximization problem is a linear program with two constraints. We let $Q(w)$ be the optimal objective value and $\boldsymbol{x}^*(w)$ be an optimal solution of this linear program as a function of w. It is a standard result in linear programming theory that $Q(w)$ is a piecewise linear function of w with $O(n^2)$ points of nondifferentiability. Furthermore, these points of nondifferentiability for $Q(\cdot)$ do not depend on the values of $\{(f^*_\ell, u^*_\ell) : \ell \in \mathcal{M},\ \ell \neq k\}$ and t.

Letting $T = \sum_{i\in\mathcal{N}} v_i$, $\sum_{i\in\mathcal{N}} v_i\, x_i \in [0,T]$. We use $\{\widehat{w}_s : s \in \mathcal{Q}\}$ to denote the points of nondifferentiability of $Q(\cdot)$ with the convention that $0, T \in \mathcal{Q}$. We write problem (34) as

$$\max_{w\in\mathbb{R}_+}\left\{ -\frac{t\,w}{1+u^*_{k-1}} + \sum_{\ell=k+1}^{m} \frac{g^*_\ell}{1+u^*_\ell + w} + \frac{\lambda_k}{1+u^*_{k-1}}\, Q(w)\right\} = \max_{s\in\mathcal{Q}}\left\{ -\frac{t\,\widehat{w}_s}{1+u^*_{k-1}} + \sum_{\ell=k+1}^{m} \frac{g^*_\ell}{1+u^*_\ell + \widehat{w}_s} + \frac{\lambda_k}{1+u^*_{k-1}}\, Q(\widehat{w}_s)\right\},$$

where the equality holds since the objective function of the first problem above is convex in w, in which case, an optimal solution must occur at a point of nondifferentiability.

Thus, the collection $\{\boldsymbol{x}^*(\widehat{w}_s) : s \in \mathcal{Q}\}$ with $|\mathcal{Q}| = O(n^2)$ includes an optimal solution to problem (32) for any value of $\{(g^*_\ell, u^*_\ell) : \ell \in \mathcal{M},\ \ell \neq k\}$ and t. ■

The main computational effort in constructing the collection of candidate assortments $\textsc{Cand}_k(j_k, j_{k+1}, b_k)$ is to solve a parametric linear program with $O(n^2)$ points of nondifferentiability.

A Polynomial-Time Algorithm for Fixed Number of Stages:

We can solve the $\textsc{Capacitated}$ problem as follows. We construct the collection of candidate assortments $\textsc{Cand}_k(j_k, j_{k+1}, b_k)$ for all $j_k, j_{k+1} \in \mathcal{N}$, $b_k \le b$, $k \in \mathcal{M}$. There are $O(n^{m-1})$ choices of $(j_1, \ldots, j_m)$ such that $0 = j_1 \le j_2 \le \ldots \le j_m \le j_{m+1} = n$, as well as $O(b^m)$ choices of $(b_1, \ldots, b_m)$ such that $\sum_{k\in\mathcal{M}} b_k = b$. For each choice of $(j_1, \ldots, j_m)$ and $(b_1, \ldots, b_m)$, since $|\textsc{Cand}_k(j_k, j_{k+1}, b_k)| = O(n^2)$, there are $O(n^{2m})$ ways of picking an assortment from the collection for each stage to construct a possible solution to the $\textsc{Capacitated}$ problem. Thus, we get the next result.

Theorem N.2 *We can construct a collection of $O(b^m\, n^{3m-1})$ possible solutions to the* $\textsc{Capacitated}$ *problem that is guaranteed to include an optimal solution to this problem. Letting* LP *be the number of operations to solve a parametric linear program with $O(n^2)$ points of nondifferentiability, constructing these solutions requires $O(b\, n^2 \mathsf{LP} + b^m\, n^{3m-1})$ operations.*

A Pseudo Polynomial-Time Algorithm:

Noting the objective function of the $\textsc{Capacitated}$ problem, knowing the value of j^*_k such that $S^*_1 \cup \ldots \cup S^*_{k-1} \subseteq \{1, \ldots, j^*_k\}$, the value of b^*_k such that $|S^*_1 \cup \ldots \cup S^*_{k-1}| = b^*_k$, and the value of

u^*_{k-1} such that $\sum_{q=1}^{k-1} V(S^*_q) = u^*_{k-1}$ is enough to compute the optimal expected revenue in stages $k+1,\ldots,m$. Thus, we can solve the CAPACITATED problem by using dynamic programming. The decision epochs are the stages. The state variable at decision epoch k is (j_k, b_k, u_{k-1}) such that the assortments $S_1,\ldots,S_{k-1}$ offered in the previous stages satisfy $S_1 \cup \ldots \cup S_{k-1} \subseteq \{1,\ldots,j_k\}$, $|S_1 \cup \ldots \cup S_{k-1}| = b_k$ and $\sum_{q=1}^{k-1} V(S_q) = u_{k-1}$. The action at decision epoch k is the value of j_{k+1} such that the assortment offered in stage k satisfies $S_k \subseteq \{j_k+1,\ldots,j_{k+1}\}$, along with the assortment $S_k \in \cup_{d=0}^{b} \textsc{Cand}_k(j_k, j_{k+1}, d)$ offered in stage k. So, we consider the dynamic program

$$J_k(j,c,u) \quad = \max_{\substack{(\ell,S)\,:\,\ell \in \{j,\ldots,n\} \\ S \in \cup_{d=0}^{b} \textsc{Cand}_k(j,\ell,d)}} \left\{ \frac{\lambda_k\, W(S)}{(1+u)\,(1+u+V(S))} + J_{k+1}(\ell\,,\, c+|S|\,,\, u+V(S)) \right\}$$

with the boundary condition that $J_{m+1}(j,c,u) = -\infty$ if $c > b$. If $c \le b$, then $J_{m+1}(j,c,u) = 0$. Solving the dynamic program above requires constructing the collections of candidate assortments a priori.

Since $|\textsc{Cand}_k(j,\ell,d)| = O(n^2)$, at each decision epoch, there are $O(v_{\max}\, b\, n^2)$ possible values of the state variable and $O(bn^3)$ possible values of the action. So, we have the next result.

Theorem N.3 *Letting* LP *be as in Theorem N.2, we can obtain an optimal solution to the* CAPACITATED *problem in* $O(b\, n^2 \mathsf{LP} + v_{\max}\, m\, n^5\, b^2)$ *operations.*

Fully Polynomial-Time Approximation Scheme:

To obtain an FPTAS, we discretize the state variable in the dynamic program that we use to construct a pseudo polynomial-time algorithm. We consider the dynamic program

$$\Psi_k(j,c,u) \quad = \max_{\substack{(\ell,S)\,:\,\ell \in \{j,\ldots,n\} \\ S \in \cup_{d=0}^{b} \textsc{Cand}_k(j,\ell,d)}} \left\{ \frac{\lambda_k\, W(S)}{(1+u)\,(1+u+V(S))} + \Psi_{k+1}(\ell\,,\, c+|S|\,,\, \lceil u+V(S) \rceil) \right\}$$

with the boundary condition that $\Psi_{m+1}(j,c,u) = -\infty$ if $c > b$. If $c \le b$, then $\Psi_{m+1}(j,c,u) = 0$. In the dynamic program above, the roundup operator $\lceil \cdot \rceil$ is as in Section 5.1.

Building on the dynamic program above, we can give an FPTAS by using an argument similar to the one in Section 5. In particular, once we compute the value functions $\{\Psi_k(j,c,u) : j = 0,\ldots,n+1,\ c = 0,\ldots,b,\ u \in \textsc{Dom},\ k \in \mathcal{M}\}$ through the dynamic program above, starting from state $(0,0,0)$, we follow the sequence of optimal state-action pairs to obtain the assortments $(\widehat{S}_1,\ldots,\widehat{S}_m)$ over m stages. We can show that expected revenue from the assortments $(\widehat{S}_1,\ldots,\widehat{S}_m)$ deviate from the optimal expected revenue by at most a factor of $(1+\rho)^{2m}$, where ρ is the size of the geometric grid. For given $\epsilon \in (0,1)$, setting $\rho = \epsilon/(2m)$, we get the next result.

Theorem N.4 *Letting* LP *be as in Theorem N.2, for each* $\epsilon \in (0,1)$*, we can obtain a* $(1-\epsilon)$*-approximate solution to the* CAPACITATED *problem in* $O(b\, n^2 \mathsf{LP} + \frac{m^2\, n^4\, b^2}{\epsilon} \log(\frac{n\, v_{\max}}{v_{\min}}))$ *operations.*

P_0	$b=1$	$b=3$	$b=5$	$b=10$	$b=20$
0.5	251.65	126.49	108.12	86.39	73.53
0.7	281.75	143.99	119.20	90.04	74.66
0.9	278.77	143.25	117.72	85.93	70.85

Table EC.3 CPU seconds to estimate the parameters of our choice model.

Appendix O: Preprocessing the Dataset from Expedia

We explain our approach for preprocessing the dataset from Expedia and give a full description of the columns. The raw dataset includes about ten million rows and 54 columns. In some of the search queries, the price is given as the total amount over the whole length of the stay, whereas in some others, the price is given as the amount per night. It is not possible to reliably tell which approach is used in each search query. To avoid ambiguity, we focused our attention on the search queries for a single night stay and dropped the remaining search queries. Furthermore, we dropped the columns for which the entries are missing for more than 25% of the rows. Considering the remaining columns, we dropped the search queries for which the entries were missing in one of the remaining columns. Lastly, some rows in the dataset included entries that are too large or too small. We dropped all search queries which had an entry in a column that falls outside the 0.5-th and 99.5-th percentile band of the entries in the corresponding column. After preprocessing the dataset, we end up with 595,965 rows representing 34,561 search queries and 15 columns. We describe the first three columns in the main text.

The remaining 12 columns give the star rating and the average review score for the hotel, an indicator for whether the hotel is part of a chain, two location desirability scores, the average price of the hotel over the last trading period, the displayed price, an indicator for whether the hotel is on promotion, the number of days until the day of stay, the number of adults and children in the search query, and an indicator for whether the stay is over the weekend.

Appendix P: Running Time for Fitting the Choice Models

We used the routine `fmincon` in Matlab to maximize the log-likelihood functions for both choice models under consideration. In Table EC.3, we give the average CPU seconds to estimate the parameters of our multinomial logit model with impatient customers for different values of P_0 and b, where the average is computed over the 50 datasets. We observe that the CPU seconds to estimate the parameters of our choice model increase as b gets smaller so that we have more stages in the choice model. For a fixed value of b, the CPU seconds showed less than 20% variation from one dataset to another. For comparison purposes, we note that the average CPU seconds to estimate the parameters of the standard multinomial logit model is 18.34 seconds.

Appendix Q: Upper Bound for Joint Pricing and Assortment Optimization

We give an upper bound on the optimal objective value of the PRICING-ASSORTMENT problem. For given assortments $(S_1, \ldots, S_m)$ and no-purchase probabilities $\boldsymbol{q}$ satisfying $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$, the expected revenue is given by (5). Making its dependence on the assortments explicit, we use $\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m)$ to denote the expected revenue in (5). We construct an upper bound on the expected revenue by treating $\sum_{i \in S_k} e^{\alpha_i}$ in (5) as a continuous quantity.

Specifically, letting $T = \sum_{i \in \mathcal{N}} e^{\alpha_i}$, for each $(S_1, \ldots, S_m) \in \mathcal{F}$, we have $\sum_{k \in \mathcal{M}} \sum_{i \in S_k} e^{\alpha_i} \leq T$. In this case, using the decision variables $\boldsymbol{x} = (x_1, \ldots, x_m)$, by (5), we have

$$\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m) \;\leq\; \frac{1}{\beta} \max_{\boldsymbol{x} \in \mathbb{R}_+^m} \left\{ \sum_{k \in \mathcal{M}} \lambda_k \, (q_{k-1} - q_k) \left\{ \log x_k - \log \left(\frac{1}{q_k} - \frac{1}{q_{k-1}} \right) \right\} \;:\; \sum_{k \in \mathcal{M}} x_k \leq T \right\},$$

where we use the fact that $(\sum_{i \in S_1} e^{\alpha_i}, \ldots, \sum_{i \in S_m} e^{\alpha_i})$ is a feasible but not necessarily an optimal solution to the problem on the right side above.

Using the Lagrange multiplier $\alpha \geq 0$, we relax the constraint $\sum_{k \in \mathcal{M}} x_k \leq T$. Thus, for each $\alpha \geq 0$, we can upper bound the optimal objective value of the problem on the right side above by

$$\frac{1}{\beta} \max_{\boldsymbol{x} \in \mathbb{R}_+^m} \left\{ \sum_{k \in \mathcal{M}} \lambda_k \, (q_{k-1} - q_k) \left\{ \log x_k - \log \left(\frac{1}{q_k} - \frac{1}{q_{k-1}} \right) \right\} - \sum_{k \in \mathcal{M}} \alpha \, x_k + \alpha \, T \right\}.$$

This problem decomposes by the stages. By the first-order condition for the problem $\max_{x_k \in \mathbb{R}_+} \lambda_k \, (q_{k-1} - q_k) \log x_k - \alpha x_k$, the optimal value of x_k is $\lambda_k \, (q_{k-1} - q_k) / \alpha$.

Plugging the optimal value of x_k into the objective function of the problem presented immediately above, the optimal objective value of the problem is

$$\frac{1}{\beta} \sum_{k \in \mathcal{M}} \lambda_k \, (q_{k-1} - q_k) \left\{ \log \frac{\lambda_k \, (q_{k-1} - q_k)}{\alpha} - \log \left(\frac{1}{q_k} - \frac{1}{q_{k-1}} \right) - 1 \right\} + \frac{\alpha \, T}{\beta}.$$

By the discussion so far, for any $\alpha \geq 0$, the quantity shown above provides an upper bound on $\widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m)$, as long as $(S_1, \ldots, S_m) \in \mathcal{F}$ and $q_{k-1} \geq q_k$ for all $k \in \mathcal{M}$. We simplify this quantity by noting that $\log \frac{\lambda_k \, (q_{k-1} - q_k)}{\alpha} - \log(\frac{1}{q_k} - \frac{1}{q_{k-1}}) - 1 = \log(q_{k-1} \, q_k) + \log \frac{\lambda_k}{\alpha} - 1$. Thus, we can upper bound the optimal expected revenue in the PRICING-ASSORTMENT problem as

$$\begin{aligned} &\max_{(\boldsymbol{q}, S_1, \ldots, S_m) \in \mathbb{R}_+^m \times \mathcal{F}} \left\{ \widehat{\Pi}(\boldsymbol{q}, S_1, \ldots, S_m) \;:\; q_{k-1} \geq q_k \quad \forall \, k \in \mathcal{M} \right\} \\ &\quad \leq \; \frac{1}{\beta} \max_{\boldsymbol{q} \in \mathbb{R}_+^m} \left\{ \sum_{k \in \mathcal{M}} \lambda_k \, (q_{k-1} - q_k) \, (\log(q_{k-1} \, q_k) + \log \tfrac{\lambda_k}{\alpha} - 1) \;:\; q_{k-1} \geq q_k \; \forall \, k \in \mathcal{M} \right\} + \frac{\alpha \, T}{\beta}. \qquad (35) \end{aligned}$$

In the problem shown on the right side above, intuitively speaking, only the no-purchase probabilities in two successive stages k and $k-1$ interact, which indicates that we can solve this problem using dynamic programming. To obtain a dynamic program with a finite number of possible states, we discretize the state variable. It is never optimal to charge negative prices in the joint pricing and assortment problem, since dropping a product with a negative price always increases the expected revenue. Thus, we can lower bound the no-purchase probability in any stage as $q_k(\boldsymbol{\rho}) = 1/(1+\widehat{V}_k(\boldsymbol{\rho})) = 1/(1+\sum_{i \in S_k} e^{\alpha_i - \beta\,\rho_k}) \geq 1/(1+\sum_{i\in\mathcal{N}} e^{\alpha_i}) = \frac{1}{1+T}$. We divide the interval $[\frac{1}{1+T}, 1]$ into $L+1$ subintervals using $\nu_0, \ldots, \nu_{L+1}$ that satisfy $\frac{1}{1+T} = \nu_0 < \nu_1 < \ldots < \nu_L < \nu_{L+1} = 1$. Let $G^\alpha_k(p,r)$ be such that $G^\alpha_k(p,r) \geq \lambda_k\,(q_{k-1} - q_k)\,(\log(q_{k-1}\,q_k) + \log\frac{\lambda_k}{\alpha} - 1)$ all $q_{k-1} \in [\nu_p, \nu_{p+1}]$ and $q_k \in [\nu_r, \nu_{r+1}]$. Coming up with such an upper bound $G^\alpha_k(p,r)$ is not difficult. The first derivatives of $(q_{k-1} - q_k)\log(q_{k-1}\,q_k)$ with respect to q_{k-1} and q_k are, respectively, negative and positive, so $(q_{k-1} - q_k)\log(q_{k-1}\,q_k)$ is decreasing in q_{k-1} and increasing in q_k Thus, if $\log\frac{\lambda_k}{\alpha} - 1 \geq 0$, then we set $G^\alpha_k(p,r) = \lambda_k\,(\nu_p - \nu_{r+1})\log(\nu_p\,\nu_{r+1}) + \lambda_k\,(\nu_{p+1} - \nu_r)\,(\log\frac{\lambda_k}{\alpha} - 1)$. If $\log\frac{\lambda_k}{\alpha} - 1 < 0$, then we set $G^\alpha_k(p,r) = \lambda_k\,(\nu_p - \nu_{r+1})\log(\nu_p\,\nu_{r+1}) + \lambda_k\,(\nu_p - \nu_{r+1})\,(\log\frac{\lambda_k}{\alpha} - 1)$. In our dynamic program, we focus on the possible intervals that can include the no-purchase probabilities $(q_1, \ldots, q_m)$. The decision epochs are the stages. The state at decision epoch k is the interval that includes q_{k-1}. The action at decision epoch k is the interval that includes q_k. Since the no-purchase probabilities in problem (35) satisfy $q_{k-1} \geq q_k$, we impose the constraint that the interval that includes q_k should not lie to the right of the interval that includes q_{k-1}. Thus, we consider the dynamic program

$$J^\alpha_k(p) \;=\; \max_{r \in \{0,\ldots,p\}} \Big\{ G^\alpha_k(p,r) + J^\alpha_{k+1}(r) \Big\} \tag{36}$$

with the boundary condition that $J^\alpha_{m+1}(p) = \alpha\,T$. Next, we show that $\frac{1}{\beta}\,J^\alpha_1(L)$ is an upper bound on the optimal objective value of the Pricing-Assortment problem.

Proposition Q.1 *For each $\alpha \geq 0$, $\frac{1}{\beta}\,J^\alpha_1(L)$ is an upper bound on the optimal expected revenue of the* Pricing-Assortment *problem.*

Proof: By the discussion earlier in this section, it is enough to show that $J^\alpha_1(L)$ is an upper bound on the optimal objective value of the problem

$$\max_{\boldsymbol{q} \in \mathbb{R}^m_+} \left\{ \sum_{k \in \mathcal{M}} \lambda_k\,(q_{k-1} - q_k)\,(\log(q_{k-1}\,q_k) + \log\tfrac{\lambda_k}{\alpha} - 1) \;:\; q_{k-1} \geq q_k \;\; \forall\, k \in \mathcal{M} \right\} + \alpha T.$$

Let $\boldsymbol{q}^*$ be an optimal solution to the problem above and p^*_k be such that $q^*_k \in [\nu_{p^*_k}, \nu_{p^*_k+1}]$. Since $q^*_{k-1} \geq q^*_k$, we have $p^*_{k-1} \geq p^*_k$. Also, since $q^*_0 = 1$, we have $p^*_0 = L$.

Let $Z_k = \sum_{\ell=k}^m \lambda_\ell\,(q^*_{\ell-1} - q^*_\ell)\,(\log(q^*_{\ell-1}\,q^*_\ell) + \log\frac{\lambda_\ell}{\alpha} - 1) + \alpha T$ with $Z_{m+1} = \alpha T$. We use induction over the stages to show that $J^\alpha_k(p^*_{k-1}) \geq Z_k$. Since $J^\alpha_{m+1}(p) = \alpha T$, the result holds for stage $m+1$. Assuming that $J^\alpha_{k+1}(p^*_k) \geq Z_{k+1}$, we proceed to showing that $J^\alpha_k(p^*_{k-1}) \geq Z_k$. Since $p^*_k \leq p^*_{k-1}$, when computing $J_k^\alpha(p_{k-1}^*)$ though the dynamic program in (36), p_k^* is a feasible but not necessarily an optimal decision. Therefore, we get

$$\begin{aligned} J_k^\alpha(p_{k-1}^*) \;&\geq\; G_k^\alpha(p_{k-1}^*, p_k^*) + J_{k+1}^\alpha(p_k^*) \\ &\geq\; \lambda_k\,(q_{k-1}^* - q_k^*)\,(\log(q_{k-1}^*\,q_k^*) + \log\tfrac{\lambda_k}{\alpha} - 1) + Z_{k+1} \;=\; Z_k, \end{aligned}$$

where the second inequality uses the fact that $J_{k+1}^\alpha(p_k^*) \geq Z_{k+1}$ by the induction hypothesis, along with the fact that $G_k^\alpha(p,r) \geq \lambda_k\,(q_{k-1} - q_k)\,(\log(q_{k-1}\,q_k) + \log\frac{\lambda_k}{\alpha} - 1)$ for all $q_{k-1} \in [\nu_p, \nu_{p+1}]$ and $q_k \in [\nu_r, \nu_{r+1}]$ by the definition of $G_k^\alpha(p,r)$, as well as noting that $q_{k-1}^* \in [\nu_{p_{k-1}^*}, \nu_{p_{k-1}^*+1}]$ and $q_k^* \in [\nu_{p_k^*}, \nu_{p_k^*+1}]$. Thus, the induction argument is complete. Therefore, we have $J_1^\alpha(L) = J_1^\alpha(p_0^*) \geq Z_1$, in which case, the desired result follows by observing that Z_1 is the optimal objective value of the problem at the beginning of the proof. ■

By the proposition above, the quantity $\frac{1}{\beta}\,J_1^\alpha(L)$ is an upper bound on the optimal objective value of the Pricing-Assortment problem for any $\alpha \geq 0$, so computing $\frac{1}{\beta}\,J_1^\alpha(L)$ for any $\alpha \geq 0$ provides an upper bound on the optimal expected revenue. To get a reasonably tight upper bound on the optimal expected revenue, we use a few iterations of the golden ratio search to find an approximate solution to the problem $\frac{1}{\beta}\min_{\alpha\geq 0} J_1^\alpha(L)$. This approach amounts to computing $J_1^\alpha(L)$ for a few different values of α. To obtain the results reported in our computational experiments in Section 6.2, we choose the end points $\nu_0, \ldots, \nu_{L+1}$ of the intervals $\{[\nu_p, \nu_{p+1}] : p = 0, \ldots, L\}$ such that $\nu_{p+1} - \nu_p$ is approximately 0.001 for all $p = 0, \ldots, L$.

Appendix R: Upper Bound under a Space Constraint

To obtain an upper bound on the optimal expected revenue in the assortment problem under a space constraint, we consider the linear program

$$\textsc{Cap}(j,\ell,x,y) \;=\; \min_{\boldsymbol{w}\in[0,1]^{\ell-j}} \left\{ \sum_{i=j+1}^{\ell} c_i\,w_i \;:\; \sum_{i=j+1}^{\ell} v_i\,r_i\,w_i \geq x, \quad \sum_{i=j+1}^{\ell} v_i\,w_i \leq y \right\}. \tag{37}$$

If we impose the constraints $\boldsymbol{w} \in \{0,1\}^{\ell-j}$ in the problem above and drop the round down and up operators in (10), then the problem above and (10) solve the same knapsack problem.

If the problem above is infeasible, then we set $\textsc{Cap}(j,\ell,x,y) = +\infty$. Note that $W(S) \leq n\,w_{\max}$ and $V(S) \leq n\,v_{\max}$ for all $S \subseteq \mathcal{N}$. Also, letting $r_{\max} = \max\{r_i : i \in \mathcal{N}\}$, we have $\Pi(S_1, \ldots, S_m) \leq r_{\max}$ for all $(S_1, \ldots, S_m) \in \mathcal{F}$. Letting $B = \max\{n\,w_{\max}, n\,v_{\max}, r_{\max}\}$, we divide the interval $[0, B]$ into $L+1$ subintervals using $\nu_0, \ldots, \nu_{L+1}$ that satisfy $0 = \nu_0 < \nu_1 < \ldots < \nu_L < \nu_{L+1} = B$. Throughout this section, we define the round down operator "$\lfloor\cdot\rfloor$" that rounds its argument down to the closest point in $\{\nu_p : p = 0, \ldots, L+1\}$ when the argument is positive. That is, if $a \geq 0$, then $\lfloor a \rfloor = \max\{\nu_r : \nu_r \leq a,\ r = 0, \ldots, L+1\}$. If $a < 0$, then $\lfloor a \rfloor = -\infty$. We consider the dynamic program

$$\overline{\Psi}_k(j,u,z) = \min_{\substack{(\ell,p,r)\,:\, \ell \in \{j,\ldots,n\},\\ p \in \{0,\ldots,L\},\\ r \in \{1,\ldots,L+1\}}} \Bigg\{ \textsc{Cap}(j,\ell,\nu_p,\nu_r) + \overline{\Psi}_{k+1}\left(\ell,\, \lfloor u + \nu_{r-1} \rfloor,\, \left\lfloor z - \frac{\lambda_k\, \nu_{p+1}}{(1+u)\,(1+u+\nu_{r-1})} \right\rfloor \right) \Bigg\} \quad (38)$$

with the boundary condition that $\overline{\Psi}_{m+1}(j,u,z) = 0$ if $z \leq 0$. Otherwise, we have $\overline{\Psi}_{m+1}(j,u,z) = +\infty$. Note that the dynamic program above is analogous to the one in (13).

In the next proposition, we show that we obtain an upper bound on the optimal expected revenue in the Capacitated problem by solving the dynamic program above.

Proposition R.1 *Letting* $\overline{z}_{\text{APP}} = \max\{z \in \mathbb{R}_+ : \overline{\Psi}_1(0,0,z) \leq b\}$, $\overline{z}$ *is an upper bound on the optimal expected revenue in the* Capacitated *problem.*

Proof: Using an induction argument that is similar to the one in the proof of Lemma J.1, it follows that $\overline{\Psi}_k(j,u,z)$ is increasing in j, u and z. Let $(S_1^*, \ldots, S_m^*)$ be an optimal solution to the Capacitated problem. By Lemma 5.1, there exist $j_1^*, \ldots, j_{m+1}^*$ satisfying $0 = j_1^* \leq j_2^* \leq \ldots \leq j_m^* \leq j_{m+1}^* = n$ such that $S_k^* \subseteq \{j_k^*+1, \ldots, j_{k+1}^*\}$. Also, let $p_k^* = 0, \ldots, L$ and $r_k^* = 1, \ldots, L+1$ be such that $W(S_k^*) \in [\nu_{p_k^*}, \nu_{p_k^*+1}]$ and $V(S_k^*) \in [\nu_{r_k^*-1}, \nu_{r_k^*}]$. Consider solving problem (37) with $j = j_k^*$, $\ell = j_{k+1}^*$, $x = \nu_{p_k^*}$ and $y = \nu_{r_k^*}$. Setting $w_i = 1$ if $i \in S_k$ and $w_i = 0$ if $i \notin S_k$ provides a feasible solution to this problem with the objective value $C(S_k^*)$. Thus, $\textsc{Cap}(j_k^*, j_{k+1}^*, \nu_{p_k^*}, \nu_{r_k^*}) \leq C(S_k^*)$.

For notational brevity, we let $C_k^* = \sum_{q=k}^m C(S_q^*)$, $u_k^* = \sum_{q=1}^{k-1} V(S_q^*)$ and $z_k^* = \sum_{q=k}^m \frac{\lambda_q\, W(S_q^*)}{(1+u_q^*)\,(1+u_{q+1}^*)}$ with the convention that $C_{m+1}^* = 0$, $u_1^* = 0$ and $z_{m+1}^* = 0$. Observe that z_1^* corresponds to the optimal objective value of the Capacitated problem. We use induction over the stages to show that $\overline{\Psi}_k(j_k^*, u_k^*, z_k^*) \leq C_k^*$. Since $z_{m+1}^* = 0$, we have $\overline{\Psi}_{m+1}(j_{m+1}^*, u_{m+1}^*, z_{m+1}^*) = 0 = C_{m+1}^*$. Therefore, the result holds for the base case. Assuming that $\overline{\Psi}_{k+1}(j_{k+1}^*, u_{k+1}^*, z_{k+1}^*) \leq C_{k+1}^*$, we proceed to showing that $\overline{\Psi}_k(j_k^*, u_k^*, z_k^*) \leq C_k^*$. Using the fact that $\overline{\Psi}(j,u,z)$ is increasing in u and z along with $\lfloor a \rfloor \leq a$ and noting that $W(S_k^*) \leq \nu_{p_k^*+1}$ and $V(S_k^*) \geq \nu_{r_k^*-1}$, we have

$$\overline{\Psi}_{k+1}\left(j_{k+1}^*,\, \lfloor u_k^* + \nu_{r_k^*-1} \rfloor,\, \left\lfloor z_k^* - \frac{\lambda_k\, \nu_{p_k^*+1}}{(1+u_k^*)\,(1+u_k^*+\nu_{r_k^*-1})} \right\rfloor \right) \leq \overline{\Psi}_{k+1}\left(j_{k+1}^*,\, u_k^* + V(S_k^*),\, z_k^* - \frac{\lambda_k\, W(S_k^*)}{(1+u_k^*)\,(1+u_k^*+V(S_k^*))} \right) = \overline{\Psi}_{k+1}(j_{k+1}^*,\, u_{k+1}^*,\, z_{k+1}^*),$$

where the equality above uses the definition of u_k^* and z_k^*. Consider computing $\overline{\Psi}_k(j_k^*, u_k^*, z_k^*)$ through the dynamic program in (38). Since $j_{k+1}^* \geq j_k^*$, the solution $(j_{k+1}^*, p_k^*, r_k^*)$ is feasible but not necessarily optimal to the minimization problem on the right side of (38) when we solve this problem with $(j,u,z)=(j_k^*,u_k^*,z_k^*)$. Therefore, we have the chain of inequalities

$$\begin{aligned}\overline{\Psi}_k(j_k^*,u_k^*,z_k^*) &\le \textsc{Cap}(j_k^*,j_{k+1}^*,\nu_{p_k^*},\nu_{r_k^*})+\overline{\Psi}_{k+1}\left(j_{k+1}^*\,,\,\lfloor u_k^*+\nu_{r_k^*-1}\rfloor\,,\,\left\lfloor z_k^*-\frac{\lambda_k\,\nu_{p_k^*+1}}{(1+u_k^*)\,(1+u_k^*+\nu_{r_k^*-1})}\right\rfloor\right)\\ &\overset{(a)}{\le} C(S_k^*)+\overline{\Psi}_{k+1}(j_{k+1}^*\,,\,u_{k+1}^*\,,\,z_{k+1}^*)\\ &\overset{(b)}{\le} C(S_k^*)+C_{k+1}^* \overset{(c)}{=} C_k^*,\end{aligned}$$

where (a) follows from the inequality that we give earlier in this paragraph and the fact that $\textsc{Cap}(j_k^*,j_{k+1}^*,\nu_{p_k^*},\nu_{r_k^*})\le C(S_k^*)$, (b) uses the induction hypothesis and (c) is by the definition of C_k^*. Thus, the induction argument is complete, in which case, noting that $j_1^*=0$ and $u_1^*=0$, we obtain $\overline{\Psi}_1(0,0,z_1^*)\le C_1^*=\sum_{k\in\mathcal{M}}C_k(S_k^*)\le b$, where the last inequality uses the fact that $(S_1^*,\ldots,S_m^*)$ is a feasible solution to the Capacitated problem. Therefore, z_1^* is a feasible solution to the problem $\overline{z}_{\text{APP}}=\max\{z\in\mathbb{R}_+:\overline{\Psi}_1(0,0,z)\le b\}$, which implies that the optimal objective value of this problem is at least as large as z_1^*. In other words, we have $\overline{z}_{\text{APP}}\ge z_1^*$. In this case, the result follows by noting that z_1^* is the optimal objective value of the Capacitated problem. ■

Note that the upper bound in the proposition above holds for any choice of $\nu_0,\ldots,\nu_{L+1}$ that satisfy $0=\nu_0<\nu_1<\ldots<\nu_L<\nu_{L+1}=B$.

Online Appendix References

Garey, M. R., D. S. Johnson. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, New York, NY.

Hoorfar, A., M. Hassani. 2008. Inequalities on the Lambert-W function and hyperpower function. *Journal of Inequalities in Pure & Applied Mathematics* **9**(2) 51.1–51.11.

Zhang, H., P. Rusmevichientong, H. Topaloglu. 2018. Technical note: Multiproduct pricing under the generalized extreme value models with homogeneous price sensitivity parameters. *Operations Research* **66**(6) 1559–1570.