Online Appendix Companion to Optimal Dynamic Assortment Planning with Demand Learning

Appendix C: Proof of Main Results

Proof of Theorem 1. The lower bound is trivial when $\widetilde{\mathcal{N}} = S^*(\mu)$, so assume $S^*(\mu) \subset \widetilde{\mathcal{N}}$. For $i \in \mathcal{N}$ define $T_i(t)$ as the number of customers product i has been offered to, before customer t's arrival,

$$T_i(t) := \sum_{u=1}^{t-1} \mathbf{1}\{i \in S_u\}, \; t \geq 1.$$

Similarly, for $n \geq 1$ define $t_i(n)$ as the customer to whom product i is offered for the n-th time,

$$t_i(n) := \inf\{t \geq 1 : T_i(t+1) = n\}, \; n \geq 1.$$

For $i \in \widetilde{\mathcal{N}} \setminus S^*(\mu)$, define Γ_i as the set of mean utility vectors for which product i is in the optimal assortment, but that differs from μ only on its i-th coordinate. That is,

$$\Gamma_i := \left\{\gamma \in \mathbb{R}^N : \gamma_i \neq \mu_i, \, \gamma_j = \mu_j \quad \forall j \in \mathcal{N} \setminus \{i\}, \, i \in S^*(\gamma)\right\}.$$

We will use $\mathbb{E}_\pi^\gamma$ and $\mathbb{P}_\pi^\gamma$ to denote expectations and probabilities of random variables, when the assortment policy $\pi \in \mathcal{P}$ is used, and the mean utilities are given by the vector γ. Let $\mathcal{I}_i(\mu\|\gamma)$ denote the Kullback-Leibler divergence between $F(\cdot - \mu_i)$ and $F(\cdot - \gamma_i)$,

$$\mathcal{I}_i(\mu\|\gamma) := \int_{-\infty}^{\infty} \left[\log\left(dF(x-\mu_i)/dF(x-\gamma_i)\right)\right] dF(x-\mu_i).$$

This quantity measures the "distance" between $\mathbb{P}_\pi^\mu$ and $\mathbb{P}_\pi^\gamma$. We have that $0 < \mathcal{I}_i(\mu\|\gamma) < \infty$ for all $\gamma \neq \mu$, $i \in \widetilde{\mathcal{N}} \setminus S^*(\mu)$. Fix $i \in \widetilde{\mathcal{N}}$ and consider a configuration $\gamma \in \Gamma_i$. For $n \geq 1$ define the log-likelihood function

$$\mathcal{L}_i(n) := \sum_{u=1}^{n} \left[\log(dF(U_i^{t_i(u)} - \mu_i)/dF(U_i^{t_i(u)} - \gamma_i))\right].$$

Note that $\mathcal{L}_i(\cdot)$ is defined in terms of utility realizations that are unobservable to the retailer. Define $\delta(\eta)$ as the minimum (relative) optimality gap when the mean utility vector is given by $\eta \in \mathbb{R}^N$,

$$\delta(\eta) := \inf\{1 - r(S,\eta)/r(S^*(\eta),\eta) > 0 : S \in \mathcal{S}\}. \tag{13}$$

Fix $\alpha \in (0,1)$. For any consistent policy π one has that for any $\epsilon > 0$,

$$\begin{aligned}\mathcal{R}^\pi(T,\gamma) &\geq \delta(\gamma)\mathbb{E}_\pi^\gamma\{T - T_i(T)\} \\ &\geq \delta(\gamma)\left(T - \frac{(1-\epsilon)}{\mathcal{I}_i(\mu\|\gamma)}\log T\right)\mathbb{P}_\pi^\gamma\{T_i(T) < (1-\epsilon)\log T/\mathcal{I}_i(\mu\|\gamma)\},\end{aligned}$$

and by assumption on π $\mathcal{R}^\pi(T,\gamma) = o(T^\alpha)$. From the above, we have that

$$\mathbb{P}_\pi^\gamma\{T_i(T) < (1-\epsilon)\log T/\mathcal{I}_i(\mu\|\gamma)\} = o(T^{\alpha-1}). \tag{14}$$

Define the event

$$\beta_i := \left\{T_i(T) \leq \frac{(1-\epsilon)}{\mathcal{I}_i(\mu\|\gamma)}\log T, \, \mathcal{L}_i(T_i(T)) \leq (1-\alpha)\log T\right\}.$$

From the independence of utilities across products and the definition of β_i, we have that

$$\begin{aligned}
\mathbb{P}^{\gamma}_{\pi}\{\beta_i\} &= \int_{\omega\in\beta_i} d\mathbb{P}^{\gamma}_{\pi} \\
&= \int_{\omega\in\beta_i} \prod_{u=1}^{T-1}\prod_{i\in S_u} dF(U^u_i-\gamma_i) \\
&= \int_{\omega\in\beta_i} \prod_{u=1}^{T-1}\prod_{i\in S_u} \frac{dF(U^u_i-\gamma_i)}{dF(U^u_i-\mu_i)} d\mathbb{P}^{\mu}_{\pi} \\
&= \int_{\omega\in\beta_i} \prod_{n=1}^{T_i(T)} \frac{dF(U^{t_i(n)}_i-\gamma_i)}{dF(U^{t_i(n)}_i-\mu_i)} d\mathbb{P}^{\mu}_{\pi} \\
&= \int_{\omega\in\beta_i} \exp(-\mathcal{L}_i(T_i(T)))d\mathbb{P}^{\mu}_{\pi} \\
&\geq \exp(-(1-\alpha)\log T)\mathbb{P}^{\mu}_{\pi}\{\beta_i\}.
\end{aligned}$$

From (14) one has that $\mathbb{P}^{\gamma}_{\pi}\{\beta_i\} = o(T^{\alpha-1})$. It follows by (14) that as $T\to\infty$

$$\mathbb{P}^{\mu}_{\pi}\{\beta_i\} \leq \mathbb{P}^{\gamma}_{\pi}\{\beta_i\}/T^{\alpha-1} \to 0. \tag{15}$$

Indexed by n, $\mathcal{L}_i(n)$ is the sum of finite mean identically distributed independent random variables, therefore, by the strong law of large numbers (SLLN).

$$\limsup_{n\to\infty} \frac{\max\{\mathcal{L}_i(l) : l\leq n\}}{n} \leq \frac{\mathcal{I}_i(\mu\|\gamma)}{(1-\alpha)} \quad \mathbb{P}^{\mu}_{\pi}\, a.s.,$$

i.e., the log-likelihood function grows no faster than linearly with slope $\mathcal{I}_i(\mu||\gamma)$. This implies that

$$\limsup_{n\to\infty} \mathbb{P}^{\mu}_{\pi}\{\exists\, l\leq n\,,\, \mathcal{L}_i(l) > n\mathcal{I}_i(\mu\|\gamma)/(1-\epsilon)\} = 0.$$

In particular,

$$\lim_{T\to\infty} \mathbb{P}^{\mu}_{\pi}\left\{T_i(T) < \frac{(1-\epsilon)}{I_i(\mu\|\gamma)}\log T\,,\, \mathcal{L}_i(T_i(T)) > \frac{(1-\epsilon)}{1-\alpha}\log T\right\} = 0.$$

Taking $\alpha < \epsilon$ small enough, and combining with (15) one has that

$$\lim_{T\to\infty} \mathbb{P}^{\mu}_{\pi}\left\{T_i(T) < \frac{(1-\epsilon)}{I_i(\mu\|\gamma)}\log T)\right\} = 0.$$

Finally, defining the positive finite constant $H^{\mu}_i := \inf\{\mathcal{I}(\mu\|\gamma) : \gamma\in\Gamma_i\}$, it follows that

$$\lim_{T\to\infty} \mathbb{P}^{\mu}_{\pi}\{T_i(T)\geq(1-\epsilon)\log T/H^{\mu}_i)\} = 1.$$

For $i\in\mathcal{N}$, let $\overline{T}_i$ denote the largest $T\geq 0$ such that $\mathbb{P}^{\mu}_{\pi}\{T_i(T)\geq(1-\epsilon)\log T/H^{\mu}_i\} < 1/2$. By Markov's inequality, and letting ϵ shrink to zero we get

$$\mathbb{E}^{\mu}_{\pi}\{T_i(T)\} \geq (2H^{\mu}_i)^{-1}\log T, \tag{16}$$

for $T > \overline{T}_i$. By the definition of the regret, we have that for any policy $\pi\in\mathcal{P}'$,

$$\begin{aligned}
\mathcal{R}^{\pi}(T,\mu) &\overset{(a)}{\geq} \delta(\mu)\,\mathbb{E}^{\mu}_{\pi}\left[\sum_{t=1}^{T}\mathbb{P}^{\mu}_{\pi}\mathbf{1}\{S_t\neq S^*(\mu)\}\right] \\
&\overset{(b)}{\geq} \delta(\mu)\frac{1}{C}\sum_{i\in\widetilde{\mathcal{N}}\setminus S^*(\mu)} \mathbb{E}^{\mu}_{\pi}[T_i(T)].
\end{aligned}$$

where (a) follows from the non-optimal assortments contributing at least $\delta(\mu)$ to the regret, and (b) follows by assuming non-optimal products are always tested in batches of size C, considering only products in $\widetilde{\mathcal{N}}$. Thus

$$\sum_{u=1}^{T} \mathbf{1}\left\{S_u \neq S^*(\mu)\right\} \geq \sum_{u=1}^{T} \mathbf{1}\left\{S_u \cap \widetilde{\mathcal{N}} \setminus S^*(\mu) \neq \emptyset\right\} \geq \frac{1}{C} \sum_{i \in \widetilde{\mathcal{N}} \setminus S^*(\mu)} \sum_{u=1}^{T} \mathbf{1}\left\{i \in S_u\right\} = \frac{1}{C} \sum_{i \in \widetilde{\mathcal{N}} \setminus S^*(\mu)} T_i(T).$$

Combining the above with (16) we have that

$$\mathcal{R}^{\pi}(T,\mu) \geq \delta(\mu) \frac{1}{C} \left(\sum_{i \in \widetilde{\mathcal{N}} \setminus S^*(\mu)} (2H_i^{\mu})^{-1} \right) \log T + \delta(\mu) \overline{T},$$

for all T, where $\overline{T} := \|\overline{T}_i\|_{\mathcal{N}}$. Taking $\underline{K} := \delta(\mu) \min_{i \in \widetilde{\mathcal{N}} \setminus S^*(\mu)} \left\{(2H_i^{\mu})^{-1}\right\}$ and $\underline{K}' := \delta(\mu)\overline{T}$ gives the desired result.

We now comment on the fact that the reasoning above extends to the case when $|S^*(\mu)| > 1$. Note that (16) remains valid for products in $\widetilde{\mathcal{N}} \setminus \mathcal{N}^*$, where $\mathcal{N}^* := \{i \in \mathcal{N} : i \in S \text{ for some } S \in \mathcal{S}^*(\mu)\}$ and

$$\mathcal{S}^*(\mu) := \arg\max\left\{r(S,\mu) : S \in \mathcal{S}\right\}.$$

In particular, one has that

$$\sum_{u=1}^{T} \mathbf{1}\left\{S_u \notin \mathcal{S}^*(\mu)\right\} \geq \sum_{u=1}^{T} \mathbf{1}\left\{S_u \cap \widetilde{\mathcal{N}} \setminus \mathcal{N}^* \neq \emptyset\right\} \geq \frac{1}{C} \sum_{i \in \widetilde{\mathcal{N}} \setminus \mathcal{N}^*} \sum_{u=1}^{T} \mathbf{1}\left\{i \in S_u\right\} = \frac{1}{C} \sum_{i \in \widetilde{\mathcal{N}} \setminus \mathcal{N}^*} T_i(T),$$

The result follows from the bound on the expectation over $T_i(T)$ for products in $\widetilde{\mathcal{N}} \setminus \mathcal{N}^*$. ∎

Proof of Theorem 3. The proof follows the arguments in the proof of Theorem 2. Steps 1 and 2 are identical.

Step 3. Let $NO(t)$ denote the event that a non-optimal assortment is offered to customer t, and $G(t)$ the event that there is no forced testing for customer t. That is,

$$\begin{aligned} NO(t) &:= \left\{S_t \neq S^*(\mu)\right\}, \\ G(t) &:= \left\{T^j(t) \geq \kappa_2 \log t\,,\, j \leq |\mathcal{A}| \text{ such that } \|w\|_{A_j} \geq \omega(\hat{\mu}_t)\right\}. \end{aligned} \tag{17}$$

Define $\xi := (2\|w\|_{\infty} K\, C)^{-1} \delta(\mu) r(S^*(\mu),\mu)$. We have that

$$\begin{aligned} \mathbb{P}\left\{NO(t)\,,\, G(t)\right\} &\overset{(a)}{\leq} \mathbb{P}\left\{\|\mu - \hat{\mu}_t\|_{(S^*(\mu) \cup S^*(\hat{\mu}_t))} > \xi\,,\, G(t)\right\} \\ &\leq \mathbb{P}\left\{\|\mu - \hat{\mu}_t\|_{S^*(\mu)} > \xi\,,\, G(t)\right\} + \mathbb{P}\left\{\|\mu - \hat{\mu}_t\|_{S^*(\hat{\mu}_t)} > \xi\,,\, G(t)\right\} \\ &\overset{(b)}{\leq} \sum_{j : A_j \cap S^*(\hat{\mu}_t) \neq \emptyset} \mathbb{P}\left\{\|\mu - \hat{\mu}_t\|_{A_j} > \xi\,,\, T^j(t) > \kappa_2 \log t\right\} + \\ &\qquad \sum_{j\,:\, A_j \cap \in S^*(\mu) \neq \emptyset} \mathbb{P}\left\{\|\mu - \hat{\mu}_t\|_{A_j} > \xi\,,\, G(t)\right\} \\ &\overset{(c)}{\leq} \sum_{j : A_j \cap S^*(\hat{\mu}_t) \neq \emptyset} 2\,|A_j|\, t^{-c(\xi/\kappa(\xi))\kappa_2} + \sum_{j\,:\, A_j \cap \in S^*(\mu) \neq \emptyset} \mathbb{P}\left\{\|\mu - \hat{\mu}_t\|_{A_j} > \xi\,,\, G(t)\right\}, \end{aligned}$$

where: (a) follows from (12); (b) follows from the fact that $w_i \geq \omega(\nu)$ trivially for all $i \in S^*(\nu)$, for any vector $\nu \in \mathbb{R}^N$; and (c) follows from (10).

Fix j such that $A_j \cap S^*(\mu) \neq \emptyset$. For such an assortment we have that

$$\begin{aligned}\mathbb{P}\left\{\|\mu-\hat{\mu}_t\|_{A_j} > \xi\,,\, G(t)\right\} \leq\ & \mathbb{P}\left\{\|\mu-\hat{\mu}_t\|_{A_j} > \xi\,,\, T^j(t) \geq \kappa_2 \log t\,,\, G(t)\right\} + \\ & \mathbb{P}\left\{T^j(t) < \kappa_2 \log t\,,\, G(t)\right\}.\end{aligned}$$

The first term on the right-hand-side above can be bounded using (10). For the second one, note that $\{T^j(t) < \kappa_2 \log t\,,\, G(t)\} \subseteq \left\{\|w\|_{A_j} < \omega(\hat{\mu}_t), G(t)\right\}$. Let $\tilde{\mu} \in \mathbb{R}^N$ be such that $\tilde{\mu}_i = \mu_i$ for all $i \in S^*(\mu) \setminus S^*(\hat{\mu}_t)$, and $\tilde{\mu}_i = \hat{\mu}_{t,i}$ otherwise. We have that

$$\left\{\|\mu-\hat{\mu}_t\|_{S^*(\hat{\mu}_t)} \leq \xi\right\} \overset{(a)}{\subseteq} \left\{r(S^*(\hat{\mu}_t),\hat{\mu}_t) < r(S^*(\mu),\tilde{\mu})\right\} \overset{(b)}{\subseteq} \left\{\|w\|_{A_j} \geq \omega(\hat{\mu}_t)\right\},$$

where (a) follows from (11), and (b) follows from noting that $\tilde{\mu}$ makes $S^*(\mu)$ optimal, hence products in $A_j \cap S^*(\mu)$ are potentially optimal under $\hat{\mu}_t$. This implies that $\left\{\|w\|_{A_j} < \omega(\hat{\mu}_t)\right\} \subseteq \left\{\|\mu-\hat{\mu}_t\|_{S^*(\hat{\mu}_t)} > \xi\right\}$, i.e.,

$$\begin{aligned}\mathbb{P}\left\{T^j(t) < \kappa_2 \log t\,,\, G(t)\right\} &\leq \mathbb{P}\left\{\|w\|_{A_j} < \omega(\hat{\mu}_t),\, G(t)\right\} \\ &\leq \mathbb{P}\left\{\|\mu-\hat{\mu}_t\|_{S^*(\hat{\mu}_t)} > \xi\,,\, G(t)\right\} \\ &\leq \sum_{k\,:\,A_k \cap S^*(\hat{\mu}_t) \neq \emptyset} \mathbb{P}\left\{\|\mu-\hat{\mu}_t\|_{A_k} > \xi\,,\, G(t)\right\} \\ &\leq \sum_{k\,:\,A_k \cap S^*(\hat{\mu}_t) \neq \emptyset} 2\,|A_k|\, t^{-c(\xi/\kappa(\xi))\kappa_2},\end{aligned}$$

where the last step follows from (10). Using the above we have that

$$\begin{aligned}\mathbb{P}\left\{NO(t)\,,\, G(t)\right\} \leq & \sum_{j:A_j \cap S^*(\hat{\mu}_t) \neq \emptyset} 2\,|A_j|\, t^{-c(\xi/\kappa(\xi))\kappa_2} + \\ & \sum_{j\,:\,A_j \cap \in S^*(\mu) \neq \emptyset} \left(2\,|A_j|\, t^{-c(\xi/\kappa(\xi))\kappa_2} + \sum_{k\,:\,A_k \cap S^*(\hat{\mu}_t) \neq \emptyset} 2\,|A_k|\, t^{-c(\xi/\kappa(\xi))\kappa_2} \right) \\ \leq & \; 2C^2(2+C) t^{-c(\xi/\kappa(\xi))\kappa_2}.\end{aligned} \tag{18}$$

On the other hand, we have that

$$\mathbb{P}\left\{NO(t)\,,\, G(t)^c\right\} \leq \sum_{j\,:\,\|w\|_{A_j} \geq \omega(\mu)} \mathbb{P}\left\{S_t = A_j\,,\, G(t)^c\right\} + \sum_{j\,:\,\|w\|_{A_j} < \omega(\mu)} \mathbb{P}\left\{S_t = A_j\,,\, G(t)^c\right\}.$$

For the first term above, we have from the policy specification that

$$\sum_{u=1}^{T} \sum_{j\,:\,\|w\|_{A_j} \geq \omega(\mu)} \mathbb{P}\left\{S_u = A_j\,,\, G(u)^c\right\} \leq \lceil \overline{N}/C \rceil \left(\kappa_2 \log T + 1\right). \tag{19}$$

To analyze the second term, fix j such that $\|w\|_{A_j} < \omega(\mu)$, and define $L(t)$ as the last customer (previous to customer t) to whom the empirical optimal assortment (according to estimated mean utilities) was offered. That is

$$L(t) := \sup\left\{u \leq t-1 \,:\, G(u)\right\},$$

with $G(u)$ given in (17). Note that $L(t) \in \{t - \lfloor |\mathcal{A}|\,\kappa_2 \log t \rfloor, \ldots, t-1\}$ for $t \geq \tau$, where τ is given by

$$\tau := \inf\left\{u \geq 1 \,:\, \log(u - \lfloor |A|\,\kappa_2 \log u \rfloor) + \kappa_2^{-1} > \log u\right\}.$$

Consider $t\geq\tau$ and $u\in\{t-\lfloor|\mathcal{A}|\,\kappa_2\log t\rfloor,\dots,t-1\}$. Then

$$\begin{aligned}\mathbb{P}\{S_t=A_j\,,\,G(t)^c\,,\,L(t)=u\}&\leq \mathbb{P}\left\{\|w\|_{A_j}\geq\omega(\hat{\mu}_t)\,,\,G(t)^c\,,\,G(u)\right\}\\&\leq \mathbb{P}\{G(u),NO(u)\}+\mathbb{P}\left\{\|w\|_{A_j}\geq\omega(\hat{\mu}_t)\,,\,G(t)^c\,,\,G(u),NO(u)^c\right\}\\&\leq \mathbb{P}\{G(u),NO(u)\}+\mathbb{P}\{G(u)\,,\,NO(u)^c\,,\,\underline{\mathcal{N}}(\hat{\mu}_u)\neq\underline{\mathcal{N}}(\mu)\}+\\&\quad\ \mathbb{P}\{\|w\|_{A_j}\geq\omega(\hat{\mu}_t)\,,\,G(u)\,,\,NO(u)^c\,,\,\underline{\mathcal{N}}(\hat{\mu}_u)=\underline{\mathcal{N}}(\mu)\}.\end{aligned}\tag{20}$$

The first term in (20) can be bounded using (18). For the second, let $\tilde{\nu}\in\mathbb{R}^N$ be such that $\tilde{\nu}_i=\hat{\mu}_{u,i}$ for $i\in\mathcal{S}^*(\mu)$. Then, for any $S\in\mathcal{S}$, one has that

$$\begin{aligned}r(S^*(\tilde{\nu}),\tilde{\nu})-r(S,\tilde{\nu})&\geq r(S^*(\nu),\tilde{\nu})-r(S,\tilde{\nu})\\&\geq r(S^*(\nu),\nu)-r(S,\tilde{\nu})-\|w\|_\infty KC\|\mu-\hat{\mu}_u\|_{S^*(\mu)}\\&\geq r(S^*(\nu),\nu)-r(S,\nu)-2\|w\|_\infty KC\|\mu-\hat{\mu}_u\|_{S^*(\mu)},\end{aligned}$$

where $\nu\in\mathbb{R}^N$ is such that $\nu_i=\mu_i$ for $i\in\mathcal{S}^*(\mu)$ and $\nu_i=\tilde{\nu}_i$ otherwise. The last two inequalities make use of (11). Define

$$\underline{\delta}:=(2\|w\|_\infty KC)^{-1}\inf\left\{r(S^*(\nu),\nu)-r(S,\nu):v\in\mathbb{R}^N\,,\,\nu_i=\mu_i\text{ for }i\in S^*(\mu)\,,\,S\cap\underline{\mathcal{N}}(\mu)\neq\emptyset\right\}>0.$$

We conclude that $\{\underline{\mathcal{N}}(\hat{\mu}_u)\subset\underline{\mathcal{N}}(\mu)\,,\,NO(u)^c\}\subseteq\left\{\|\mu-\hat{\mu}_u\|_{S^*(\mu)}>\underline{\delta}\,,\,NO(u)^c\right\}$. Repeating the argument above, one has that $\{\underline{\mathcal{N}}(\mu)\subset\underline{\mathcal{N}}(\hat{\mu}_u)\,,\,NO(u)^c\}\subseteq\left\{\|\mu-\hat{\mu}_u\|_{S^*(\mu)}>\overline{\delta}\,,\,NO(u)^c\right\}$, where

$$\overline{\delta}:=(2\|w\|_\infty KC)^{-1}\sup\left\{\inf\{r(S^*(\nu),\nu)-r(S,\nu):S\neq S^*(\nu)\}:v\in\mathbb{R}^N\,,\,\nu_i=\mu_i\text{ for }i\in S^*(\mu)\right\}>0.$$

Define $\delta:=\min\left\{\underline{\delta},\overline{\delta}\right\}$. (Note that $\delta>0$, provided that $\{i\in\mathcal{N}:w_i=\omega(\mu)\}=\emptyset$.) One has that

$$\begin{aligned}\mathbb{P}\{G(u)\,,\,NO(u)^c\,,\,\underline{\mathcal{N}}(\mu)\neq\underline{\mathcal{N}}(\hat{\mu}_u)\}&\overset{(a)}{\leq}\mathbb{P}\left\{\|\mu-\hat{\mu}_u\|_{S^*(\mu)}\geq\delta\,,\,T^k(t)+1\geq\kappa_2\log t\,,\,\forall k\text{ s.t. }A_k\cap S^*(\mu)\neq\emptyset\right\}\\&\leq\sum_{k:A_k\cap S^*(\mu)\neq\emptyset}\mathbb{P}\left\{\|\mu-\hat{\mu}_u\|_{A_k}\geq\delta\,,\,T^k(t)+1\geq\kappa_2\log t\right\}\\&\overset{(b)}{\leq}D\ t^{-c(\delta/\kappa(\delta))\kappa_2},\end{aligned}$$

where (a) follows from noting that $T^k(t)\geq T^k(u)\geq\kappa_2\log u\geq\kappa_2\log t-1$ for k such that $A_k\cap S^*(\mu)\neq\emptyset$, (b) follows from (10), and D is a positive and finite constant. For the third term in (20), define the event $\Xi:=\left\{\|w\|_{A_j}\geq\omega(\hat{\mu}_t)\,,\,G(u)\,,\,NO(u)^c\,,\,\underline{\mathcal{N}}(\hat{\mu}_u)=\underline{\mathcal{N}}(\mu)\right\}$.

$$\mathbb{P}\{\Xi\}=\mathbb{P}\{\Xi\,,\,S^*(\hat{\mu}_t)=S^*(\mu)\}+\mathbb{P}\{\Xi\,,\,S^*(\hat{\mu}_t)\neq S^*(\mu)\}\,.$$

From the arguments leading to the bound for the second term in (20), one has that

$$\mathbb{P}\{\Xi\,,\,S^*(\hat{\mu}_t)=S^*(\mu)\}\leq D\ t^{-c(\underline{\delta}/\kappa(\underline{\delta}))\kappa_2}.$$

On the other hand, one has that

$$\begin{aligned}\mathbb{P}\{\Xi\,,\,S^*(\hat{\mu}_t)\neq S^*(\mu)\}&\leq\mathbb{P}\{\Xi\,,\,S^*(\hat{\mu}_t)\neq S^*(\mu)\,,\,\underline{\mathcal{N}}(\mu)\cap S^*(\hat{\mu}_t)=\emptyset\}+\\&\quad\ \mathbb{P}\{\Xi\,,\,S^*(\hat{\mu}_t)\neq S^*(\mu)\,,\,\underline{\mathcal{N}}(\mu)\cap S^*(\hat{\mu}_t)\neq\emptyset\}\,.\end{aligned}\tag{21}$$

For bounding the first term above, note that $\underline{\mathcal{N}}(\mu)\cap S^*(\hat\mu_t)=\emptyset$ implies that $T^k(t)\geq \kappa_2\log t-1$ for all k such that $A_k\cap S^*(\hat\mu_t)\neq\emptyset$. Thus, the arguments leading to (18) imply that

$$\begin{aligned}\mathbb{P}\{\Xi\,,\,S^*(\hat\mu_t)\neq S^*(\mu)\,,\,\underline{\mathcal{N}}(\mu)\cap S^*(\hat\mu_t)=\emptyset\} \leq& \sum_{k:A_k\cap S^*(\hat\mu_t)\neq\emptyset}\mathbb{P}\left\{\|\mu-\hat\mu_t\|_{A_k}>\xi\,,\,T^k(t)>\kappa_2\log t-1\right\}+\\ &\sum_{k:A_k\cap S^*(\mu)\neq\emptyset}\mathbb{P}\left\{\|\mu-\hat\mu_t\|_{A_k}>\xi\,,\,T^k(t)>\kappa_2\log t-1\right\}\\ \leq& \, D't^{-c(\xi/\kappa(\xi))\kappa_2},\end{aligned}$$

for a finite and positive constant D'. For the second term in (21), note that

$$\left\{S^*(\nu)\cap\underline{\mathcal{N}}(\mu)\neq\emptyset,\,v\in\mathbb{R}^N\text{ such that }\nu_i=\hat\mu_{t,i}\,\forall\,i\in S^*(\mu)\right\}\subseteq\left\{\|\hat\mu_t-\mu\|_{s^*(\mu)}>\underline{\delta}\right\},$$

hence, one has that

$$\begin{aligned}\mathbb{P}\{\Xi\,,\,S^*(\hat\mu_t)\neq S^*(\mu)\,,\,\underline{\mathcal{N}}(\mu)\cap S^*(\hat\mu_t)\neq\emptyset\} \leq& \,\mathbb{P}\left\{\|\hat\mu_t-\mu\|_\infty s^*(\mu)>\underline{\delta}\,,\,T^k(t)\geq\kappa_2\log t-1\,,\,\forall\,k\text{ s.t. }A_k\cap S^*(\mu)\neq\emptyset\right\}\\ \leq& \sum_{k:A_k\cap S^*(\mu)\neq\emptyset}\mathbb{P}\left\{\|\hat\mu_t-\mu\|_\infty A_k>\underline{\delta}\,,\,T^k(t)\geq\kappa_2\log t-1\right\}\\ \leq& \, D''t^{-c(\underline{\delta}/\kappa(\underline{\delta}))\kappa_2},\end{aligned}$$

for some finite and positive constant D'', where the last inequality follows from (10). Putting the bounds above together, (20) becomes

$$\mathbb{P}\{S_t=A_j\,,\,G(t)^c\}\leq D'''t^{-\tilde\xi\kappa_2},$$

for some finite and positive constant D''', where $\tilde\xi:=\min\{c(\xi)/\kappa(\xi),c(\delta)/\kappa(\delta),c(\underline{\delta})/\kappa(\underline{\delta})\}$. The bound above, and those in (19) and (18), imply that for $\kappa_2>\tilde\xi^{-1}$, one has

$$\begin{aligned}\mathcal{R}^\pi(T,\mu) \leq& \sum_{t=1}^T\mathbb{P}\{NO(t),G(t)\}+\sum_{t=1}^T\mathbb{P}\{NO(t)\,,\,G(t)^c\}\\ \leq& \sum_{t=1}^T\mathbb{P}\{NO(t),G(t)\}+\\ &\sum_{t=1}^T\sum_{j\,:\,\|w\|_{A_j}\geq r(S^*(\mu),\mu)}\mathbb{P}\{S_t=A_j\,,\,G(t)^c\}+\\ &\sum_{t=1}^T\sum_{j\,:\,\|w\|_{A_j}<r(S^*(\mu),\mu)}\mathbb{P}\{S_t=A_j\,,\,G(t)^c\}\\ \leq& \sum_{t=1}^\infty C^2(2+C)u^{-c(\xi/\kappa(\xi))\kappa_2}+\lceil\overline{\mathcal{N}}/C\rceil(\kappa_2\log(T)+1)+\\ &\sum_{t=1}^\infty\sum_{j\,:\,\|w\|_{A_j}<r(S^*(\mu),\mu)}D'''t^{-\tilde\xi\kappa_2}\\ \overset{(a)}{\leq}& \lceil|\overline{\mathcal{N}}|/C\rceil\kappa_2\log T+\overline{K}_2,\end{aligned}$$

for a finite constant $\overline{K}_2<\infty$, where (b) uses the summability of the series implied by (20). Taking $\overline{\kappa}_2>\tilde\xi^{-1}$ provides the desired result. ■

Proof of Corollary 1. Fix $i\in\underline{\mathcal{N}}$, and fix $j=\{k\leq|\mathcal{A}|\,:\,i\in A_k\}$. We have that

$$\mathbb{E}_\pi[T_i(T)]\leq\tau+\sum_{t=\tau+1}^T\mathbb{P}[NO(t)\,,\,G(t)]+\mathbb{P}[S_t=A_j\,,\,G(t)^c]$$

$$\leq K_2,$$

for a finite constant K_2, where we have used the summability of the series implied by (20). This completes the proof. ■

Proof of Theorem 4. The proof is an adaptation of the one for Theorem 3, customized for the MNL choice model. However, we provide a explanation version of each step with the objective of highlighting how the structure of the MNL model is exploited.

Step 1. We will need the following side lemma, whose proof is deferred to Appendix D.

LEMMA 2. *Fix $i \in \mathcal{N}$. For any $n \geq 1$ and $\epsilon > 0$ one has*

$$\mathbb{P}\left\{\left|\sum_{u=1}^{t-1}\left(Z_j^u - \mathbb{E}\left\{Z_j^u\right\}\right)\mathbf{1}\left\{i \in S_u\right\}\right| \geq \epsilon T_i(t)\,,\, T_i(t) \geq n\right\} \leq 2\exp(-c(\epsilon)n),$$

for $j \in \{i, 0\}$ and a positive constant $c(\epsilon) < \infty$.

Consider $\epsilon > 0$ and fix $t \geq 1$ and $i \in \mathcal{N}$. Define $\varrho = 1/2\,(1 + C\|w\|_\infty)^{-1}$: one has that $p_0(S,\mu) \geq 2\varrho$, for all $S \in \mathcal{S}$. For $n \geq 1$ define the event $\Xi := \{|\nu_i - \hat{\nu}_{i,t}| > \epsilon\,,\, T_i(t) \geq n\}$. We have that

$$\begin{aligned}
\mathbb{P}\{\Xi\} &= \mathbb{P}\left\{\left|\frac{\sum_{u=1}^{t-1} Z_i^u \mathbf{1}\{i \in S_u\}}{\sum_{u=1}^{t-1} Z_0^u \mathbf{1}\{i \in S_u\}} - \nu_i\right| > \epsilon\,,\, T_i(t) \geq n\right\} \\
&\leq \mathbb{P}\left\{\left|\frac{\sum_{u=1}^{t-1} Z_i^u \mathbf{1}\{i \in S_u\}}{\sum_{u=1}^{t-1} Z_0^u \mathbf{1}\{i \in S_u\}} - \nu_i\right| > \epsilon\,,\right. \\
&\qquad \left.\left|\sum_{u=1}^{t-1}(Z_0^u - \mathbb{E}\{Z_0^u\})\mathbf{1}\{i \in S_u\}\right| < \varrho T_i(t)\,,\, T_i(t) \geq n\right\} + \\
&\quad \mathbb{P}\left\{\left|\sum_{u=1}^{t-1}(Z_0^u - \mathbb{E}\{Z_0^u\})\mathbf{1}\{i \in S_u\}\right| \geq \varrho T_i(t)\,,\, T_i(t) \geq n\right\} \\
&\overset{(a)}{\leq} \mathbb{P}\left\{\left|\sum_{u=1}^{t-1}(Z_i^u - Z_0^u\nu_i)\mathbf{1}\{i \in S_u\}\right| > \epsilon\varrho T_i(t)\,,\, T_i(t) \geq n\right\} + 2\exp(-c(\varrho)n) \\
&\overset{(b)}{\leq} \mathbb{P}\left\{\left|\sum_{u=1}^{t-1}(Z_i^u - E[Z_i^u])\mathbf{1}\{i \in S_u\}\right| > \epsilon\varrho/2\, T_i(t)\,,\, T_i(t) \geq n\right\} + \\
&\quad \mathbb{P}\left\{\left|\sum_{u=1}^{t-1}(Z_0^u - E[Z_0^u])\mathbf{1}\{i \in S_u\}\right| > \epsilon\varrho/(2\nu_i)\, T_i(t)\,,\, T_i(t) \geq n\right\} + 2\exp(-c(\varrho)n) \\
&\leq 2\exp(-c(\epsilon\varrho/2)n) + 2\exp(-c(\epsilon\varrho/(2\nu_i))n) + 2\exp(-c(\varrho)n).
\end{aligned}$$

where: (a) follows from Lemma 2 and from the fact that

$$\left|\sum_{u=1}^{t-1} Z_0^u \mathbf{1}\{i \in S_u\}\right| \geq \left|\sum_{u=1}^{t-1} E[Z_0^u]\mathbf{1}\{i \in S_u\}\right| - \left|\sum_{u=1}^{t-1}(Z_0^u - E[Z_0^u])\mathbf{1}\{i \in S_u\}\right| \geq \varrho T_i(t),$$

when $\left|\sum_{u=1}^{t-1}(Z_0^u - \mathbb{E}\{Z_0^u\})\mathbf{1}\{i \in S_u\}\right| < \varrho T_i(t)$; and (b) follows from the fact that $\mathbb{E}Z_i^u = \nu_i \mathbb{E}Z_0^u$, for all $u \geq 1$ such that $i \in S_u$, $i \in \mathcal{N}$. For $\epsilon > 0$ define

$$\tilde{c}(\epsilon) := \min\{c(\epsilon\varrho/2)\,,\, c(\epsilon\varrho/(2\|\nu\|_{\mathcal{N}}))\,,\, c(\varrho)\}\,.$$

From above we have that for $\epsilon > 0$

$$\mathbb{P}\{|\nu_i - \hat{\nu}_{i,t}| > \epsilon\,,\, T_i(t) \geq n\} \leq 6\exp(\tilde{c}(\epsilon)n), \tag{22}$$

for all $i \in \mathcal{N}$.

Step 2. Consider two vectors $\upsilon, \eta \in \mathbb{R}_+^N$, and define $\tilde{\upsilon} := \ln \upsilon$ and $\tilde{\eta} := \ln \eta$. From (7), for any $S \in \mathcal{S}$ one has

$$\begin{aligned}\sum_{i\in S} \upsilon_i(w_i - r(S,\tilde{\upsilon})) &= r(S,\tilde{\upsilon}) \\ \sum_{i\in S} \eta_i(w_i - r(S,\tilde{\upsilon})) &\geq r(S,\tilde{\upsilon}) - C\|w\|_\infty \|\upsilon-\eta\|_S \\ \sum_{i\in S} \eta_i(w_i - (r(S,\tilde{\upsilon}) - C\|w\|_\infty\|\upsilon-\eta\|_S) &\geq r(S,\tilde{\upsilon}) - C\|w\|_\infty \|\upsilon-\eta\|_S\end{aligned}$$

This implies that

$$r(S,\tilde{\eta}) \geq r(S,\tilde{\upsilon}) - C\|w\|_\infty \|\eta - \upsilon\|_S. \tag{23}$$

From the above we conclude that

$$\{S^*(\mu) \neq S^*(\hat{\mu}_t)\} \subseteq \left\{ \|\nu - \hat{\nu}_t\|_{S^*(\mu)\cup S^*(\hat{\mu}_t)} \geq (2\|w\|_\infty C)^{-1}\delta(\nu) r(S^*(\mu),\mu) \right\}, \tag{24}$$

where with a slight abuse of notation $\delta(\nu)$ refers to the minimum optimality gap, in terms of the adjusted terms $\exp(\mu)$.

Step 3. Let $NO(t)$ denote the event that a non-optimal assortment is offered to customer t, and $G(t)$ the event that there is no "forced testing" on customer t. That is

$$\begin{aligned} NO(t) &:= \{S_t \neq S^*(\mu)\}, \\ G(t) &:= \{T_i(t) \geq \kappa_3 \log t\,,\ \forall\, i \in \mathcal{N} \text{ such that } w_i \geq r(S^*(\hat{\mu}_t),\hat{\mu}_t)\}. \end{aligned}$$

Define $\xi := (2\|w\|_\infty C)^{-1}\delta(\nu) r(S^*(\mu),\mu)$. We have that

$$\begin{aligned} \mathbb{P}\{NO(t)\,,G(t)\} &\overset{(a)}{\leq} \mathbb{P}\left\{\|\nu-\hat{\nu}_t\|_{(S^*(\mu)\cup S^*(\hat{\mu}_t))} > \xi\,,G(t)\right\} \\ &\leq \mathbb{P}\left\{\|\nu-\hat{\nu}_t\|_{S^*(\hat{\mu}_t)} > \xi\,,G(t)\right\} + \mathbb{P}\left\{\|\nu-\hat{\nu}_t\|_{S^*(\mu)} > \xi\,,G(t)\right\} \\ &\overset{(b)}{\leq} \sum_{i\in S^*(\hat{\mu}_t)} \mathbb{P}\{|\nu_i - \hat{\nu}_{i,t}| > \xi\,, T_i(t) \geq \kappa_3 \log t\} + \sum_{i\in S^*(\mu)} \mathbb{P}\{|\nu_i-\hat{\nu}_{i,t}|>\xi\,,G(t)\} \\ &\overset{(c)}{\leq} 6Ct^{-\kappa_3\tilde{c}(\xi)} + \sum_{i\in S^*(\mu)} \mathbb{P}\{|\nu_i-\hat{\nu}_{i,t}|>\xi\,,G(t)\} \end{aligned}$$

where: (a) follows from (24); (b) follows from the fact that $w_i \geq r(S^*(\eta),\eta)$ for all $i \in S^*(\eta)$ and for any vector $\eta \in \mathbb{R}^N$ (see Step 2 above); and (c) follows from (22). Fix $i \in S^*(\mu)$. We have that

$$\mathbb{P}\{|\nu_i - \hat{\nu}_{i,t}| > \xi, G(t)\} \leq \mathbb{P}\{|\nu_i - \hat{\nu}_{i,t}| > \xi\,, T_i(t) \geq \kappa_3\log t\} + \mathbb{P}\{G(t),\, T_i(t) < \kappa_3 \log t\}.$$

The first term above can be bounded using (22). Regarding the second one, note that $\{G(t)\,, T_i(t) < \kappa_3 \log t\} \subseteq \{w_i < r(S^*(\hat{\mu}_t),\hat{\mu}_t)\}$, and that

$$\begin{aligned} w_i - r(S^*(\mu),\mu)\delta(\nu)/2 &\geq r(S^*(\mu),\mu)(1-\delta(\nu)/2) \\ &\overset{(a)}{\geq} r(S^*(\hat{\mu}_t),\mu) \\ &\overset{(b)}{\geq} r(S^*(\hat{\mu}_t),\hat{\mu}_t) - \|w\|_\infty C\|\nu-\hat{\nu}_t\|_{S^*(\hat{\mu}_t)}, \end{aligned}$$

where (a) follows from the definition of $\delta(\nu)$, and (b) follows from (23). The above implies that $\{w_i < r(S^*(\hat{\mu}_t), \hat{\mu}_t)\} \subseteq \{\|\nu - \hat{\nu}_t\|_{S^*(\hat{\mu}_t)} > \xi\}$, i.e.,

$$\begin{aligned}\mathbb{P}\{T_i(t) < \kappa_3 \log t\,,\, G(t)\} &\le \mathbb{P}\{w_i < r(S^*(\hat{\mu}_t), \hat{\mu}_t), G(t)\}\\ &\le \mathbb{P}\left\{\|\nu - \hat{\nu}_t\|_{S^*(\hat{\mu}_t)} > \xi\,,\, G(t)\right\}\\ &\le \sum_{j\in S^*(\hat{\mu}_t)} \mathbb{P}\{|\nu_j - \hat{\nu}_{j,t}| > \xi\,,\, G(t)\}\\ &\le 6Ct^{-\kappa_3 \tilde{c}(\xi)},\end{aligned}$$

where the last step follows from (22). Using the above we have that

$$\mathbb{P}\{NO(t)\,,\, G(t)\} \le 6C(1+C)t^{-\kappa_3 \tilde{c}(\xi)}. \tag{25}$$

From here, we have that

$$\begin{aligned}\mathbb{P}\{NO(t)\,,\, G(t)^c\} &\le \sum_{i\,:\,w_i < r^*(S^*(\mu),\mu)} \mathbb{P}\{i \in S_t\,,\, G(t)^c\} + \sum_{i\,:\,w_i \ge r^*(S^*(\mu),\mu)} \mathbb{P}\{i \in S_t\,,\, G(t)^c\}\\ &\overset{(a)}{\le} \sum_{i\,:\,w_i < r^*(S^*(\mu),\mu)} \mathbb{P}\{i \in S_t\,,\, G(t)^c\} + \quad \left|\overline{\mathcal{N}}\right| (\kappa_3 \log T + 1)\,. + \\ &\sum_{i\in S^*(\mu)} \mathbb{P}\{i \in S_t\,,\, G(t)^c\}\end{aligned}$$

where (a) follows from the specification of the policy. Fix i such that $w_i < r(S^*(\mu), \mu)$, and define $L(t)$ as the last customer (previous to customer t) to whom the empirical optimal assortment, according to estimated mean utilities, was offered. That is

$$L(t) := \sup\{u \le t-1 \,:\, G(u)\}\,.$$

Note that $L(t) \in \{t - \lfloor N\kappa_3 \log t\rfloor, \ldots, t-1\}$ for $t \ge \tau$, where τ is given by

$$\tau := \inf\left\{u \ge 1 \,:\, \log(u - \lfloor N\kappa_3 \log u\rfloor) + \kappa_3^{-1} > \log u\right\}.$$

Consider $t \ge \tau$ and $u \in \{t - \lfloor N\kappa_3 \log t\rfloor, \ldots, t-1\}$. Then

$$\begin{aligned}\mathbb{P}\{i \in S_t\,,\, G(t)^c\,,\, L(t) = u\} &\le \mathbb{P}\{w_i \ge r(S^*(\hat{\mu}_t), \hat{\mu}_t)\,,\, G(t)^c\,,\, L(t) = u\}\\ &\overset{(a)}{\le} \mathbb{P}\{w_i \ge r(S^*(\hat{\mu}_t), \hat{\mu}_t)\,,\, G(t)^c\,,\, G(u)\}\\ &\le \mathbb{P}\{G(u), NO(u)\} + \mathbb{P}\{w_i \ge r(S^*(\hat{\mu}_t), \hat{\mu}_t)\,,\, G(t)^c\,,\, G(u), NO(u)^c\}\\ &\overset{(b)}{\le} 6C(1+C)u^{-\kappa_3 \tilde{c}(\xi)} +\\ &\quad \mathbb{P}\{w_i \ge r(S^*(\hat{\mu}_t), \hat{\mu}_t)\,,\, T_j(t) \ge \kappa_3 \log t \;\forall\, j \in S^*(\mu)\}\,,\end{aligned}$$

where (a) follows from $\{L(t) = u\} \subseteq \{G(u)\}$, and (b) from (25) and the fact that offering $S^*(\mu)$ to customer u implies (from $G(u)$) that $T_j(u) \ge \kappa_3 \log u$ and therefore (from $t \ge \tau$) that $T_j(t) \ge \kappa_3 \log t$, for all $j \in S^*(\mu)$. From (23) we have that

$$r(S^*(\mu), \hat{\mu}_t) - w_i \ge r(S^*(\mu), \mu) - \|w\|_\infty C \|\nu - \hat{\nu}_t\|_{S^*(\mu)} - w_i.$$

Define $\delta := \inf\{(\|w\|_\infty C)^{-1}(1 - w_i / r(S^*(\mu), \mu)) > 0 \,:\, i \in \mathcal{N}\}$. From the above, we have that

$$\{w_i \ge r(S^*(\hat{\mu}_t), \hat{\mu}_t)\} \subseteq \{w_i \ge r(S^*(\mu), \hat{\mu}_t)\} \subseteq \left\{\|\nu - \hat{\nu}_t\|_{S^*(\mu)} > \delta r(S^*(\mu), \mu)\right\}.$$

Define $\bar{\delta} := \delta r(S^*(\mu),\mu)$, and the event $\Xi = w_i \geq r(S^*(\hat{\mu}_t),\hat{\mu}_t)\,,\, T_j(t) \geq \kappa_3 \log t \;\forall\, j \in S^*(\mu)$. It follows that

$$\begin{aligned}\mathbb{P}\{\Xi\} &\leq \mathbb{P}\left\{\|\nu - \hat{\nu}_t\|_{S^*(\mu)} > \bar{\delta}\,,\, T_j(t) \geq \kappa_3 \log t \;\forall\, j \in S^*(\mu)\right\}\\ &\leq \sum_{i \in S^*(\mu)} \mathbb{P}\left\{|\nu_i - \hat{\nu}_{t,i}| > \bar{\delta}\,,\, T_i(t) \geq \kappa_3 \log t\right\}\\ &\leq 6Ct^{-\kappa_3 \tilde{c}(\bar{\delta})}.\end{aligned}$$

Using the above one gets that, when $\kappa_3 > \tilde{c}(\xi)^{-1}$

$$\begin{aligned}\mathbb{P}\{i \in S_t\,,\, G(t)^c\,,\, L(t) = u\} &\leq 6C(1+C)u^{-\kappa_3\tilde{c}(\xi)} + 6Ct^{-\kappa_3\tilde{c}(\bar{\delta})}\\ &\leq 6C(1+C)(t - \lfloor N\kappa_3 \log t\rfloor)^{-\kappa_3\tilde{c}(\xi)} + 6Ct^{-\kappa_3\tilde{c}(\bar{\delta})}.\end{aligned}$$

Since the right hand side above is independent of u, one has that

$$\mathbb{P}\{i \in S_t\,,\, G(t)^c\} \leq 6C(1+C)(t - \lfloor N\kappa_3 \log t\rfloor)^{-\kappa_3\tilde{c}(\xi)} + 6Ct^{-\kappa_3\tilde{c}(\bar{\delta})}, \tag{26}$$

for all $i \in \mathcal{N}$ such that $w_i < r(S^*(\mu),\mu)$, and $t \geq \tau$.

Now fix $i \in S^*(\mu)$, and consider $t \geq \tau$, $u \in \{t - \lfloor N\kappa_3 \log t\rfloor, \ldots, t-1\}$ and $\kappa_3 > \tilde{c}(\xi)^{-1}$. Then

$$\begin{aligned}\mathbb{P}\{i \in S_t\,,\, G(t)^c\,,\, L(t) = u\} &\leq \mathbb{P}\{T_i(t) < \kappa_3 \log t\,,\, G(t)^c\,,\, L(t) = u\}\\ &\overset{(a)}{\leq} \mathbb{P}\{T_i(t) < \kappa_3 \log t\,,\, G(u)\}\\ &\leq \mathbb{P}\{G(u), NO(u)\} + \mathbb{P}\{T_i(t) < \kappa_3 \log t\,,\, G(u)\,,\, NO^c(u)\}\\ &\overset{(b)}{\leq} 6C(1+C)u^{-\kappa_3\tilde{c}(\xi)}\\ &\leq 6C(1+C)(t - \lfloor N\kappa_3 \log t\rfloor)^{-\kappa_3\tilde{c}(\xi)},\end{aligned}$$

where (a) follows from $\{L(t) = u\} \subseteq \{G(u)\}$, and (b) from (25) and the fact that offering $S^*(\nu)$ to customer u implies (from $G(u)$) that $T_i(u) \geq \kappa_3 \log u$ and therefore (from $t \geq \tau$) that $T_i(t) \geq \kappa_3 \log t$. Since the right hand side above is independent of u, one has that

$$\mathbb{P}\{i \in S_t\,,\, G(t)^c\} \leq 6C(1+C)(t - \lfloor N\kappa_3 \log t\rfloor)^{-\kappa_3\tilde{c}(\xi)}, \tag{27}$$

for all $i \in S^*(\mu)$ and $t \geq \tau$.

Considering $\kappa_3 > \max\left\{\tilde{c}(\xi)^{-1}\,,\, \tilde{c}(\bar{\delta})^{-1}\right\}$ results in the following bound for the regret

$$\begin{aligned}\mathcal{R}^\pi(T,\nu) &\leq \sum_{t=1}^{T} \mathbb{P}\{NO(t), G(t)\} + \sum_{t=1}^{T} \mathbb{P}\{NO(t)\,,\, G(t)^c\}\\ &\overset{(a)}{\leq} 6C(1+C)\sum_{t=1}^{\infty} t^{-\kappa_3\tilde{c}(\xi)} + \left|\overline{\mathcal{N}} \setminus S^*(\mu)\right| \kappa_3(\log T + 1) + \tau +\\ &\quad 6C\,|\underline{\mathcal{N}} \cup S^*(\mu)| \sum_{t=\tau}^{\infty} (1+C)(t^{-\kappa_3\tilde{c}(\xi)} + (t - \lfloor N\kappa_3 \log t\rfloor)^{-\kappa_3\tilde{c}(\xi)}) + t^{-\kappa_3\tilde{c}(\bar{\delta})}\\ &\overset{(b)}{\leq} \left|\overline{\mathcal{N}} \setminus S^*(\mu)\right| \kappa_3 \log T + \overline{K}_3,\end{aligned}$$

for a finite constant $\overline{K}_3 < \infty$, where (a) follows from (25), (26) and (27), and (b) uses the summability of the series, implied by the terms in (25), (26) and (27). Taking $\overline{\kappa}_3 > \max\left\{\tilde{c}(\xi)^{-1}\,,\, \tilde{c}(\bar{\delta})^{-1}\right\}$ provides the desired result. ■

Proof of Corollary 2. Fix $i \in \underline{\mathcal{N}}$. We have that

$$\begin{aligned}\mathbb{E}_\pi[T_i(T)] &\leq \tau + \sum_{t=\tau+1}^{T} \mathbb{P}[NO(t)\,,\, G(t)] + \mathbb{P}[i \in S_t\,,\, G(t)^c] \\ &\leq K_3 < \infty,\end{aligned}$$

for a finite constant K_3, where we have used the summability of the terms in (25) and (26). This concludes the proof. ■

Appendix D: Proof of Auxiliary Results

Proof of Lemma 1. Fix $i \in \mathcal{N}$. For $\theta > 0$ consider the process $\{M_t(\theta) : t \geq 1\}$, defined as

$$M_t(\theta) := \exp\left(\sum_{u=1}^{t} \mathbf{1}\{S_u = A_j\}\left[\theta(Z_i^u - p_i(A_j,\mu)) - \phi(\theta)\right]\right),$$

where

$$\phi(\theta) := \log \mathbb{E}\{\exp(\theta\,(Z_i^u - p_i(A_j,\mu)))\} = -\theta p_i(A_j,\mu) + \log(p_i(A_j,\mu)\exp(\theta) + 1 - p_i(A_j,\mu)),$$

and $A_j \in \mathcal{A}$ such that $i \in A_j$. One can check that $M_t(\theta)$ is an $\mathcal{F}_t$-martingale, for any $\theta > 0$ (see Section 3 for the definition of $\mathcal{F}_t$). Note that

$$\exp\left(\theta \sum_{u=1}^{t} \mathbf{1}\{S_u = A_j\}\left((Z_i^u - p_i(A_j,\mu)) - \epsilon\right)\right) = \sqrt{M_t(2\theta)}\exp\left(\sum_{u=1}^{t} \mathbf{1}\{S_u = A_j\}\left(\phi(2\theta)/2 - \theta\epsilon\right)\right). \qquad (28)$$

Let χ_i denote the event we are interested in. That is

$$\chi_i := \left\{\sum_{u=1}^{t-1} (Z_i^u - p_i(A_j,\mu))\,\mathbf{1}\{S_u = A_j\} \geq T^j(t)\epsilon\,,\, T^j(t) \geq n\right\}.$$

Let $\psi(t)$ denote the choice made by the t-th user. Using the above one has that

$$\begin{aligned}\mathbb{P}\{\chi_i\} &\overset{(a)}{\leq} \mathbb{E}\left\{\exp\left(\theta \sum_{u=1}^{t-1} \mathbf{1}\{S_u = A_j\}\left(Z_i^u - p_i(A_j,\mu) - \epsilon\right)\right) ; T_i(t) \geq n\right\} \\ &\overset{(b)}{\leq} \left(\mathbb{E}\{M_{t-1}(2\theta)\}\,\mathbb{E}\left\{\exp\left(\sum_{u=1}^{t-1} \mathbf{1}\{\psi(u) = i\}\left(\phi(2\theta) - 2\theta\epsilon\right)\right) ; T_i(t) \geq n\right\}\right)^{1/2} \\ &\overset{(c)}{\leq} \left(\mathbb{E}\left\{\exp\left(\sum_{u=1}^{t-1} \mathbf{1}\{\psi(u) = i\}\left(\phi(2\theta) - 2\theta\epsilon\right)\right) ; T_i(t) \geq n\right\}\right)^{1/2},\end{aligned}$$

where: (a) follows from Chernoff's inequality; (b) follows from the Cauchy-Schwartz inequality and (28); and (c) follows from the properties of $M_t(\theta)$. Note that when $\epsilon < (1 - p_i(A_j,\mu))$ minimizing $\phi(\theta) - \theta\epsilon$ over $\theta > 0$ results on

$$\theta^* := \log\left(1 + \frac{\epsilon}{p_i(A_j,\mu)(1 - p_i(A_j,\mu) - \epsilon)}\right) > 0,$$

with

$$c(\epsilon) := \phi(2\theta^*)/2 - \theta^*\epsilon < 0.$$

Using this we have

$$\mathbb{P}\left\{\sum_{u=1}^{t-1} (Z_i^u - p_i)\,\mathbf{1}\{S_u = A_j\} \geq T^j(t)\epsilon\,,\, T^j(t) \geq n\right\} \leq \sqrt{\mathbb{E}\{\exp(-2c(\epsilon)T_i(t)); T_i(t) \geq n\}}$$

$$\leq \exp(-c(\epsilon)n).$$

Using the same arguments one has that

$$\mathbb{P}\left\{\sum_{u=1}^{t-1}(Z_i^u - p_i)\,\mathbf{1}\{S_u = A_j\} \leq -T^j(t)\epsilon\,,\, T^j(t)\geq n\right\} \leq \exp(-c(\epsilon)n).$$

The result follows from the union bound. ■

Proof of Lemma 2. The proof follows almost verbatim the steps in the proof of Lemma 1. Fix $i\in\mathcal{N}$. For $\theta>0$ consider the process $\left\{M_t^j(\theta) : t\geq 1\right\}$, defined as

$$M_t^j(\theta) := \exp\left(\sum_{u=1}^{t}\mathbf{1}\{i\in S_u\}\left[\theta(Z_j^u - p_j(S_u,\mu)) - \phi_u^j(\theta)\right]\right) \qquad j\in\{i,0\}\,,$$

where

$$\begin{aligned}\phi_u^j(\theta) &:= \log\mathbb{E}\left\{\exp(\theta(Z_j^u - p_j(S_u,\mu)))\right\}\\ &= \log\mathbb{E}\left\{\exp(-\theta p_j(S_u,\mu))\left(\exp(\theta)p_j(S_u,\mu)+1-p_j(S_u,\mu)\right)\right\}.\end{aligned}$$

One can verify that $M_t^j(\theta)$ is an $\mathcal{F}_t$-martingale, for any $\theta>0$ and $j\in\{i,0\}$ (see §3 for the definition of $\mathcal{F}_t$). Fix $j\in\{i,0\}$ and note that

$$\exp\left(\theta\sum_{u=1}^{t}\mathbf{1}\{i\in S_u\}\left((Z_j^u - p_j(S_u,\mu))-\epsilon\right)\right) = \sqrt{M_t^j(2\theta)}\exp\left(\sum_{u=1}^{t}\mathbf{1}\{i\in S_u\}\left(\phi_u^j(2\theta)/2-\theta\epsilon\right)\right). \tag{29}$$

Put

$$\chi_j := \left\{\sum_{u=1}^{t-1}\left(Z_j^u - p_j(S_u,\mu)\right)\mathbf{1}\{i\in S_u\}\geq T_i(t)\epsilon\,,\, T_i(t)\geq n\right\}.$$

Let $\psi(t)$ denote the choice made by the t-th customer. Using the above one has that

$$\begin{aligned}\mathbb{P}\{\chi_j\} &\overset{(a)}{\leq} \mathbb{E}\left\{\exp\left(\theta\sum_{u=1}^{t-1}\mathbf{1}\{i\in S_u\}\left(Z_j^u - p_j(S_u,\mu)-\epsilon\right)\right);\, T_i(t)\geq n\right\}\\ &\overset{(b)}{\leq} \left(\mathbb{E}\left\{M_{t-1}^j(2\theta)\right\}\mathbb{E}\left\{\exp\left(\sum_{u=1}^{t-1}\mathbf{1}\{\psi(u)=j\,,\, i\in S_u\}\left(\phi_u^j(2\theta)-2\theta\epsilon\right)\right);\, T_i(t)\geq n\right\}\right)^{1/2}\\ &\overset{(c)}{\leq} \left(\mathbb{E}\left\{\exp\left(\sum_{u=1}^{t-1}\mathbf{1}\{\psi(u)=j\,,\, i\in S_u\}\left(\phi_u^j(2\theta)-2\theta\epsilon\right)\right);\, T_i(t)\geq n\right\}\right)^{1/2},\end{aligned}$$

where; (a) follows from Chernoff's inequality; (b) follows from the Cauchy-Schwartz inequality and (28); and (c) follows from the properties of $M_t^j(\theta)$. Note that $\phi_s^j(\cdot)$ is continuous, $\phi_s^j(0)=0$, $(\phi_s^j)'(0)=0$, and $\phi_s^j(\theta)\to\infty$ when $\theta\to\infty$, for all $s\geq 1$. This implies that there exists a positive constant $c(\epsilon)<\infty$ (independent of n), and a $\theta^*>0$, such that $\phi_s^j(2\theta^*)-2\theta^*\epsilon < -2c(\epsilon)$ for all $s\geq 1$. Using this we have that

$$\begin{aligned}\mathbb{P}\left\{\sum_{u=1}^{t-1}\left(Z_j^u - p_j(S_u,\mu)\right)\mathbf{1}\{i\in S_u\}\geq T_i(t)\epsilon\,,\, T_i(t)\geq n\right\} &\leq \sqrt{\mathbb{E}\left\{\exp(-2c(\epsilon)T_i(t));\, T_i(t)\geq n\right\}}\\ &\leq \exp(-c(\epsilon)n).\end{aligned}$$

Using the same arguments one has that

$$\mathbb{P}\left\{\sum_{u=1}^{t-1}\left(Z_j^u - p_j(S_u,\mu)\right)\mathbf{1}\{i\in S_u\}\leq -T_i(t)\epsilon\,,\, T_i(t)\geq n\right\}\leq \exp(-c(\epsilon)n).$$

The result follows from the union bound. ■