Two-Sided Pricing and Learning with Buyer Choice: A Generalized SGD Method

Online Appendix

Appendix A Dynamic Programming Formulation

The clairvoyant's problems (1) can be solved by dynamic programming. We provide below a formulation for a single product (i.e., $M = 1$), and the product index is omitted. The multiple-product problem can be formulated in a similar manner. Denote by $V(t, I)$ the total expected profit when there are t periods left and I units of inventory. Then, we obtain the following Bellman equations: for each $t \geq 1$ and $I \geq 1$, define

$$
\begin{aligned}
V(t, I) = \max_{b,\ p_+,\ p_0} \ & \mu(b) \left\{ -b + \lambda(p_+)[p_+ + V(t-1, I)] + (1 - \lambda(p_+))V(t-1, I+1) \right\} \\
& + (1 - \mu(b)) \left\{ \lambda(p_0)[p_0 + V(t-1, I-1)] + (1 - \lambda(p_0))V(t-1, I) \right\},
\end{aligned}
$$

where $V(0, I) = 0$ for any $I \geq 1$, and for each $t \geq 1$,

$$
\begin{aligned}
V(t, 0) \ = \max_{b,\ p_+} \ & \mu(b) \left\{ -b + \lambda(p_+)[p_+ + V(t-1, 0)] + (1 - \lambda(p_+))V(t-1, 1) \right\} \\
& + (1 - \mu(b))V(t-1, 0).
\end{aligned}
$$

The optimal policy can be found by solving $V(T, I_0)$ recursively. Similar to the literature on the trade-in programs (e.g., Xiao and Zhou 2020), we can show structural properties based on the formulation above. For example, the optimal purchase and selling prices decrease with more inventory. This is intuitive: More inventory allows us to lower the supply rate and raise the demand rate to match supply and demand better.

Appendix B Omitted Proofs in Section 3

In this section, we prove Proposition 1 and Proposition 2. Note that the fixed-price policy F can be viewed as a special case of $\hat{F}$ with the estimated matching rate to be exactly the same as the underlying rate. As such, we prove Proposition 2 below, which also proves Proposition 1.

B.1 Proof of Proposition 2

Given $\hat{\boldsymbol{\omega}}$, the total expected profit under the corresponding fixed-price policy $\hat{F}$ is

$$
\begin{aligned}
V^{\hat{F}}(T, \boldsymbol{I}_0) = & \sum_{t=1}^{T} \mathbb{E} \left[\left(\sum_{i=1}^{M} r_i(\hat{\boldsymbol{\omega}}) \right) \mathbb{1}\{ I_{j,t}^{\hat{F}} + S_{j,t}^{\hat{F}} > 0, \forall j \} - \sum_{i=1}^{M} c_i(\hat{\omega}_i - I_{i,0}/T) \right] \\
= & \sum_{t=1}^{T} \mathbb{E} \left[f(\hat{\boldsymbol{\omega}}) - \left(\sum_{i=1}^{M} r_i(\hat{\boldsymbol{\omega}}) \right) \cdot \mathbb{1}\{ \exists\ j : I_{j,t}^{\hat{F}} + S_{j,t}^{\hat{F}} = 0 \} \right]. \qquad \text{(B.1)}
\end{aligned}
$$

where f represents the single-period expected profit defined in (3). Then, the performance gap between the optimal policy and $\hat{F}$ can be bounded as

$$\begin{aligned} V(T,\boldsymbol{I}_0) - V^{\hat{F}}(T,\boldsymbol{I}_0) \le \ & V^D(T,\boldsymbol{I}_0) - V^{\hat{F}}(T,\boldsymbol{I}_0) \\ \le \ & T[f(\boldsymbol{\omega}^*) - f(\hat{\boldsymbol{\omega}})] + \left(\sum_{i=1}^M r_i(\hat{\boldsymbol{\omega}})\right) \cdot \sum_{t=1}^T \mathbb{P}\left(\bigcup_i \{I_{i,t}^{\hat{F}} + S_{i,t}^{\hat{F}} = 0\}\right), \end{aligned}$$

where $V^D(T,\boldsymbol{I}_0) = Tf(\boldsymbol{\omega}^*)$ is the optimal profit in the deterministic variant. As shown in (4), the first term in the rightmost expression can be bounded using the squared difference between $\boldsymbol{\omega}^*$ and $\hat{\boldsymbol{\omega}}$, implying $f(\boldsymbol{\omega}^*) - f(\hat{\boldsymbol{\omega}}) \le u_f \|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}\|_2^2$. The second term is the expected revenue loss from stockouts under $\hat{F}$. We claim that

$$\sum_{t=1}^T \mathbb{P}\left(\bigcup_i \{I_{i,t}^{\hat{F}} + S_{i,t}^{\hat{F}} = 0\}\right) \le C_m M \sqrt{T \log T}, \tag{B.2}$$

where C_m is a finite constant independent of M and T. Suppose that the claim holds (which we will show later). Since $\sum_{i=1}^M r_i(\hat{\boldsymbol{\omega}}) = \sum_{i=1}^M p_i(\hat{\boldsymbol{\omega}})\hat{\omega}_i \le \sum_{i=1}^M p_{\max}\hat{\omega}_i \le p_{\max}$, we obtain

$$V(T,\boldsymbol{I}_0) - V^{\hat{F}}(T,\boldsymbol{I}_0) \le \ T \cdot u_f \|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}\|_2^2 + p_{\max} C_m M \sqrt{T \log T}.$$

Thus, the expected performance loss of $\hat{F}$ satisfies

$$\mathbb{E}\left[V(T,\boldsymbol{I}_0) - V^{\hat{F}}(T,\boldsymbol{I}_0)\right] \le \ T \cdot u_f \mathbb{E}[\|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}\|_2^2] + p_{\max} C_m M \sqrt{T \log T}.$$

If the estimated matching rate satisfies $\mathbb{E}[\|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}\|_2^2] \le c_\omega M \sqrt{\log T / T}$, it follows $\mathbb{E}\left[V(T,\boldsymbol{I}_0) - V^{\hat{F}}(T,\boldsymbol{I}_0)\right] \le C_F^{(2)} M \sqrt{T \log T}$, where $C_F^{(2)} := u_f c_\omega + p_{\max} C_m$ is a finite constant. This proves Proposition 2.

In the rest of the proof, we establish the stockout bound in (B.2) directly. Conditional on $\hat{\boldsymbol{\omega}}$, let $\widetilde{\boldsymbol{D}}_t = (\widetilde{D}_{1,t}, \ldots, \widetilde{D}_{M,t})$ denote the potential MNL demand vector in period t with choice probabilities $(\hat{\omega}_1, \ldots, \hat{\omega}_M)$. These potential demand vectors are generated in every period, including periods in which $\hat{F}$ applies the null selling price vector, and are independent over time and independent of supply. The realized demand under $\hat{F}$ can therefore be represented as

$$D_{i,t}^{\hat{F}} = \widetilde{D}_{i,t} \mathbb{1}\{I_{j,t}^{\hat{F}} + S_{j,t}^{\hat{F}} > 0,\ \forall j \in [M]\}.$$

For each product i, construct on the same probability space the one-dimensional process

$$Q_{i,1} = I_{i,0}, \qquad Q_{i,t+1} = \left[Q_{i,t} + S_{i,t}^{\hat{F}} - \widetilde{D}_{i,t}\right]^+.$$

We claim that $I_{i,t}^{\hat{F}} \ge Q_{i,t}$, for any $i \in [M]$ and any $t \in [T]$. This follows by induction. When all products are available, $\hat{F}$ and the comparison process use the same potential demand. When at least one product is unavailable, $\hat{F}$ suppresses all demand, whereas the comparison process may still remove one unit. Therefore, the inventory under $\hat{F}$ cannot be smaller than the inventory in the comparison process. Consequently,

$$\mathbb{1}\left\{\bigcup_{i=1}^M \{I_{i,t}^{\hat{F}} + S_{i,t}^{\hat{F}} = 0\}\right\} \le \sum_{i=1}^M \mathbf{1}\left\{Q_{i,t} + S_{i,t}^{\hat{F}} = 0\right\}. \tag{B.3}$$

Fix a product i. Let $X_t := S^{\hat{F}}_{i,t} - \widetilde{D}_{i,t}$. The comparison process satisfies the Lindley recursion $Q_{i,t+1} = [Q_{i,t} + X_t]^+$. Its cumulative regulator can be written as

$$L_n := \max_{0\le k\le n} [-I_{i,0} - \textstyle\sum_{t=1}^k X_t]^+,$$

so that $Q_{i,n+1} = I_{i,0} + \sum_{t=1}^n X_t + L_n$. Because $X_t \in \{-1, 0, 1\}$, the regulator increases precisely when $Q_{i,t} = 0$, $S^{\hat{F}}_{i,t} = 0$, and $\widetilde{D}_{i,t} = 1$. Hence, $L_t - L_{t-1} = \widetilde{D}_{i,t}\mathbb{1}\left\{Q_{i,t} + S^{\hat{F}}_{i,t} = 0\right\}$. Since the current potential demand is independent of $(Q_{i,t}, S^{\hat{F}}_{i,t})$, it follows that

$$\mathbb{E}[L_T \,|\, \hat{\boldsymbol{\omega}}] = \hat{\omega}_i \sum_{t=1}^{T} \mathbb{P}\Big(Q_{i,t} + S^{\hat{F}}_{i,t} = 0 \,\Big|\, \hat{\boldsymbol{\omega}}\Big). \tag{B.4}$$

Define the martingale $M_k := \sum_{t=1}^k (X_t + I_{i,0}/T)$. Since $\mathbb{E}[X_t \,|\, \hat{\boldsymbol{\omega}}] = (\hat{\omega}_i - I_{i,0}/T) - \hat{\omega}_i = -I_{i,0}/T$, we have, for every $k \le T$,

$$I_{i,0} + \sum_{t=1}^{k} X_t = I_{i,0}\left(1 - \frac{k}{T}\right) + M_k \ge M_k.$$

Therefore, $L_T \le \max_{0\le k\le T}(-M_k)^+ \le \max_{0\le k\le T}|M_k|$. By Doob's L^2 maximal inequality, we obtain

$$\mathbb{E}[L_T \,|\, \hat{\boldsymbol{\omega}}] \le 2\sqrt{\mathbb{E}[M_T^2 \,|\, \hat{\boldsymbol{\omega}}]} = 2\sqrt{T\left[(\hat{\omega}_i - I_{i,0}/T)(1 - \hat{\omega}_i + I_{i,0}/T) + \hat{\omega}_i(1-\hat{\omega}_i)\right]} \le 2\sqrt{2\hat{\omega}_i T},$$

where the last inequality follows from $0 \le \hat{\omega}_i - I_{i,0}/T \le \hat{\omega}_i$. Combining this inequality with (B.4) yields

$$\sum_{t=1}^{T} \mathbb{P}\Big(Q_{i,t} + S^{\hat{F}}_{i,t} = 0 \,\Big|\, \hat{\boldsymbol{\omega}}\Big) \le 2\sqrt{\frac{2T}{\hat{\omega}_i}}.$$

Summing over all products and using (B.3) and $\hat{\omega}_i \ge \lambda_{\min}$, we obtain

$$\sum_{t=1}^{T} \mathbb{P}\left(\bigcup_{i=1}^{M}\left\{I^{\hat{F}}_{i,t} + S^{\hat{F}}_{i,t} = 0\right\} \,\middle|\, \hat{\boldsymbol{\omega}}\right) \le 2\sqrt{2T}\sum_{i=1}^{M}\frac{1}{\sqrt{\hat{\omega}_i}} \le 2M\sqrt{\frac{2T}{\lambda_{\min}}}.$$

Taking expectations with respect to $\hat{\boldsymbol{\omega}}$ proves $N^{\hat{F}}(T, \boldsymbol{I}_0) = O(M\sqrt{T\log T})$. In particular, for $T \ge 2$, (B.2) holds with $C_m = 2\sqrt{2/(\lambda_{\min}\log 2)}$. This completes the proof of Proposition 2.

Appendix C Omitted Proofs in Section 4.3

Recall that our Phase 2 starts with prices estimated based on data from Phase 1. The initial purchase prices are are defined in Section 4.3 as

$$\hat{b}^o_i = \frac{1}{\hat{\gamma}_{2,i}}\left(\hat{\omega}^*_i - \hat{\gamma}_{1,i}\right), \quad \forall\, i \in [M].$$

By definition, we can show that the estimation error satisfies

$$|\hat{b}_i^o - \bar{b}_i|^2 = \left|\hat{\omega}_i^* \left(\frac{1}{\gamma_{2,i}} - \frac{1}{\hat{\gamma}_{2,i}}\right) - \left(\frac{\gamma_{1,i}}{\gamma_{2,i}} - \frac{\hat{\gamma}_{1,i}}{\hat{\gamma}_{2,i}}\right)\right|^2 \leq \kappa_b \|\boldsymbol{\gamma}_i - \hat{\boldsymbol{\gamma}}_i\|_2^2,$$

where κ_b is a finite constant. By the classic results of the MLE estimator, we further have

$$\mathbb{P}\left(|\hat{b}_i^o - \bar{b}_i|^2 > \epsilon\right) \leq \mathbb{P}\left(\|\boldsymbol{\gamma}_i - \hat{\boldsymbol{\gamma}}_i\|_2^2 > \epsilon/\kappa_b\right) \leq \; c_\gamma^{(0)} \exp\left(-c_\gamma^{(1)} N_0 \epsilon/\kappa_b\right), \quad \forall\, i \in [M], \tag{C.1}$$

where $c_\gamma^{(0)}$ and $c_\gamma^{(1)}$ are positive and finite constants independent of M and T. We will use these high-probability bounds in the proof of Proposition 3 below.

Next, we present Theorem 4.1 in Harvey et al. (2019), which will be used to show high-probability bounds for our selected prices in Phase 2.

Lemma C.1 (Theorem 4.1, Harvey et al. 2019)**.** *Let $(X_t)_{t=1}^T$ be a stochastic process and let $(\mathcal{F}_t)_{t=1}^T$ be a filtration such that X_t is $\mathcal{F}_t$-measurable and X_t is non-negative almost surely. Assume that $\mathbb{E}[\exp(\lambda X_1)] \leq \exp(\lambda K)$ for $\lambda \in (0, 1/K]$. Let $\alpha_t \in [0,1)$ and $\beta_t, \gamma_t \geq 0$ for every t. Let $\hat{w}_t$ be a mean-zero random variable conditioned on $\mathcal{F}_t$ such that $|\hat{w}_t| \leq 1$ almost surely for every t. Suppose that $X_{t+1} \leq \alpha_t X_t + \beta_t \hat{w}_t \sqrt{X_t} + \gamma_t$ for every t. Then, the following hold.*

- *For every t, $\mathbb{P}\left(X_t \geq K \log(1/\delta)\right) \leq e\delta$.*
- *More generally, if $\sigma_1, \ldots, \sigma_T \geq 0$, then $\mathbb{P}\left(\sum_{t=1}^T \sigma_t X_t \geq K \log(1/\delta) \sum_{t=1}^T \sigma_t\right) \leq e\delta$,*

where $K = \max_{1 \leq t \leq T}\left(2\gamma_t/(1-\alpha_t), 2\beta_t^2/(1-\alpha_t)\right)$.

Lemma C.1 provides high-probability bounds for X_t satisfying a nonlinear recurrence. Harvey et al. (2019) use it to show that the output $\{x_t\}$ from the SGD method converges to the minimizer x^* with a high-probability bound $\mathbb{P}\left(\sum_{t=2}^T \sigma_t \|x_t - x^*\|^2 = O\left(\log(1/\delta) \sum_{t=2}^T (\sigma_t/t)\right)\right) \geq 1 - \delta$ for all $\sigma_t \geq 0$. To prove (9), which is the high-probability bound stated in Proposition 3, we extend their analysis to our setting with an initial point $b_{i,2N_0+1} = \hat{b}_i^o$, the estimate based on Phase 1, and step sizes of $\eta_t^s = \eta_0^s/(t+1)$ for $t > 2N_0$.

C.1 Proof of Proposition 3

As discussed in Section 4.3, the expectation bound (8) follows from a standard recursive argument for SGD (see, for example, Hazan et al. (2007)). We show below the high-probability bound (9). By the updating rule under SGD, we have for $t > 2N_0$,

$$\begin{aligned}
|b_{i,t+1} - \bar{b}_i|^2 \leq \;& |b_{i,t} - \eta_t^s \cdot [S_i(b_{i,t}) - \hat{\omega}_i^*] - \bar{b}_i|^2 \\
= \;& |b_{i,t} - \bar{b}_i|^2 - 2\eta_t^s \cdot [S_i(b_{i,t}) - \hat{\omega}_i^*](b_{i,t} - \bar{b}_i) + |\eta_t^s \cdot [S_i(b_{i,t}) - \hat{\omega}_i^*]|^2.
\end{aligned}$$

Note that $S_i(b_{i,t}) - \hat{\omega}_i^* = [\mu_i(b_{i,t}) - \hat{\omega}_i^*] - [\mu_i(b_{i,t}) - S_i(b_{i,t})] = \gamma_{2,i}(b_{i,t} - \bar{b}_i) - [\mu_i(b_{i,t}) - S_i(b_{i,t})]$ and $[S_i(b_{i,t}) - \hat{\omega}_i^*]^2 \leq \max\{(1-\hat{\omega}_i^*)^2, (-\hat{\omega}_i^*)^2\} \leq 1$. It follows that

$$\begin{aligned}
|b_{i,t+1} - \bar{b}_i|^2 \leq \;& |b_{i,t} - \bar{b}_i|^2 \left(1 - 2\eta_t^s \gamma_{2,i}\right) + 2\eta_t^s \cdot [\mu_i(b_{i,t}) - S_i(b_{i,t})](b_{i,t} - \bar{b}_i) + (\eta_t^s)^2 \\
\leq \;& |b_{i,t} - \bar{b}_i|^2 \left(\frac{t-1}{t+1}\right) - 2\left(\frac{\eta_0^s}{t+1}\right) \cdot [\mu_i(b_{i,t}) - S_i(b_{i,t})](b_{i,t} - \bar{b}_i) + \left(\frac{\eta_0^s}{t+1}\right)^2,
\end{aligned}$$

where the last inequality holds because $\eta_t^s = \eta_0^s/(t+1)$ with $\eta_0^s > 1/\gamma_{2,\min}$, which implies $1 - 2\eta_t^s\gamma_{2,i} = 1 - 2\eta_0^s\gamma_{2,i}/(t+1) < 1 - 2/(t+1) = (t-1)/(t+1)$. Multiplying by $(t+1)$ on both sides, we obtain

$$
\begin{aligned}
(t+1)|b_{i,t+1} - \bar{b}_i|^2 \leq\ & t|b_{i,t} - \bar{b}_i|^2 \left(\frac{t-1}{t}\right) + 2\eta_0^s \cdot [\mu_i(b_{i,t}) - S_i(b_{i,t})](b_{i,t} - \bar{b}_i) + \frac{(\eta_0^s)^2}{t+1} \\
=\ & t|b_{i,t} - \bar{b}_i|^2 \left(\frac{t-1}{t}\right) + \frac{2\eta_0^s}{\sqrt{t}} \cdot \frac{[\mu_i(b_{i,t}) - S_i(b_{i,t})](b_{i,t} - \bar{b}_i)}{|b_{i,t} - \bar{b}_i|} \cdot \sqrt{t|b_{i,t} - \bar{b}_i|^2} + \frac{(\eta_0^s)^2}{t+1}.
\end{aligned}
$$

Let $X_k := (2N_0 + k)|b_{i,2N_0+k} - \bar{b}_i|^2$ for $k \geq 1$. Then, $X_{k+1} \leq \alpha_k X_k + \beta_k \hat{w}_k \sqrt{X_k} + \gamma_k$ holds for all k with $\alpha_k = (2N_0 + k - 1)/(2N_0 + k)$, $\beta_k = 2\eta_0^s/\sqrt{2N_0 + k}$, $\hat{w}_k = [S_i(b_{i,2N_0+k} - \mu_i(b_{i,2N_0+k})](b_{i,2N_0+k} - \bar{b}_i)/|b_{i,2N_0+k} - \bar{b}_i|$, and $\gamma_k = (\eta_0^s)^2/(2N_0 + k + 1)$.

We want to apply Lemma C.1 to $(X_k)_{k=1}^{T-2N_0}$. Specifically, let $\mathcal{F}_k$ be the σ-algebra generated by the history of purchase prices and realized supplies before period $2N_0 + k$. Then, X_k is $\mathcal{F}_k$-measurable, and moreover, $\mathbb{E}[\hat{w}_k|\mathcal{F}_k] = 0$ holds. Observe that we have $|\hat{w}_k| \leq |S_{i,2N_0+k}^A - \mu_i(b_{i,2N_0+k})| \leq 1$, and $\max_{1\leq k\leq T-2N_0}\left(2\gamma_k/(1-\alpha_k), 2\beta_k^2/(1-\alpha_k)\right) = 8(\eta_0^s)^2$. Furthermore, we claim that we have $\mathbb{E}[\exp(\lambda X_1)] \leq \exp(\lambda K_b)$ for any $\lambda \in (0, 1/K_b]$ and some constant $K_b \geq 8(\eta_0^s)^2$ (we verify below that this claim holds). Then, we can apply Lemma C.1 with $\sigma_k = 1/(2N_0 + k)$ and $\delta = 1/(eT)$, which implies

$$
\mathbb{P}\left(\sum_{k=1}^{T-2N_0} |b_{i,2N_0+k} - \bar{b}_i|^2 \geq K_b \log(eT) \sum_{k=1}^{T-2N_0} (1/k)\right) \leq \frac{1}{T}.
$$

Since $\sum_{k=1}^{T-2N_0}(1/k) \leq 1 + \int_1^{T-2N_0}(1/x)\ dx = 1 + \log(T - 2N_0) \leq \log(eT)$, we have

$$
\begin{aligned}
& \mathbb{P}\left(\sum_{t=2N_0+1}^{T} |b_{i,t} - \bar{b}_i| \geq \sqrt{K_b(T-2N_0)} \cdot \log(eT)\right) \\
\leq\ & \mathbb{P}\left(\left(\sum_{t=2N_0+1}^{T} |b_{i,t} - \bar{b}_i|\right)^2 \geq (T - 2N_0)K_b \cdot \log(eT) \sum_{k=1}^{T-2N_0} (1/k)\right) \\
\leq\ & \mathbb{P}\left(\sum_{t=2N_0+1}^{T} |b_{i,t} - \bar{b}_i|^2 \geq K_b \log(eT) \sum_{k=1}^{T-2N_0} (1/k)\right) \leq \frac{1}{T},
\end{aligned}
$$

where the second inequality follows from Cauchy–Schwarz inequality. Thus, we obtain (9) as required.

It remains to prove the claim that we can define a constant $K_b \geq 8(\eta_0^s)^2$ such that $\mathbb{E}[\exp(\lambda X_1)] \leq \exp(\lambda K_b)$ for $\lambda \in (0, 1/K_b]$. By definition, we have $X_1 = 2N_0|b_{i,2N_0+1} - \bar{b}_i|^2 = 2N_0|\hat{b}_i^o - \bar{b}_i|^2$. Also, $|\hat{b}_i^o - \bar{b}_i|^2$ is bounded by $u_b := (b_{\max} - b_{\min})^2$ almost surely. Thus, for $\lambda \leq c_\gamma^{(1)}/(4\kappa_b)$,

$$
\begin{aligned}
\mathbb{E}[\exp(\lambda X_1)] =\ & \mathbb{E}\left[\exp\left(2\lambda N_0|\hat{b}_i^o - \bar{b}_i|^2\right)\right] \\
=\ & \int_0^{u_b} -\exp\left(2\lambda N_0\epsilon\right)\, d\mathbb{P}\left(|\hat{b}_i^o - \bar{b}_i|^2 > \epsilon\right) \\
=\ & 1 + \int_0^{u_b} 2\lambda N_0 \exp\left(2\lambda N_0\epsilon\right) \cdot \mathbb{P}\left(|\hat{b}_i^o - \bar{b}_i|^2 > \epsilon\right)\, d\epsilon \\
\leq\ & 1 + \int_0^{u_b} 2\lambda N_0 \exp\left(2\lambda N_0\epsilon\right) \cdot c_\gamma^{(0)} \exp\left(-c_\gamma^{(1)} N_0\epsilon/\kappa_b\right)\, d\epsilon,
\end{aligned}
$$

where the inequality follows from (C.1). Thus,

$$\begin{aligned}\mathbb{E}[\exp(\lambda X_1)] \leq\ & 1 + 2\lambda N_0 c_\gamma^{(0)} \int_0^{u_b} \exp\left(-(c_\gamma^{(1)}/\kappa_b - 2\lambda)N_0\epsilon\right) d\epsilon \\ \leq\ & 1 + \frac{2\lambda N_0}{(c_\gamma^{(1)}/\kappa_b - 2\lambda)N_0} \leq \exp\left(\lambda \cdot \frac{2}{c_\gamma^{(1)}/\kappa_b - 2\lambda}\right) \leq \exp\left(\lambda \cdot \frac{4\kappa_b}{c_\gamma^{(1)}}\right),\end{aligned}$$

where the second last inequality follows from $1 + x \leq \exp(x)$, and the last inequality follows from $\lambda < c_\gamma^{(1)}/(4\kappa_b)$ and, thus, $2/(c_\gamma^{(1)}/\kappa_b - 2\lambda) \leq 2/(c_\gamma^{(1)}/(2\kappa_b)) = 4\kappa_b/c_\gamma^{(1)}$. Let $K_b := \max\{8(\eta_0^s)^2, 4\kappa_b/c_\gamma^{(1)}\}$. Then for any $\lambda \leq 1/K_b$, we have $\lambda \leq c_\gamma^{(1)}/(4\kappa_b)$, and the equation above implies $\mathbb{E}[\exp(\lambda X_1)] \leq \exp(\lambda K_b)$, proving the claim. This completes our proof of (9). $\square$

Appendix D Omitted Proofs in Section 4.4

We proceed to provide detailed proofs of Lemma 1, (12), and Proposition 4 when applying the generalized SGD for the demand side. We first provide a complete proof of the expectation bound in Lemma 1.

D.1 Proof of Lemma 1

We have shown in (11) that $\mathbb{E}\left[\|\check{\boldsymbol{p}}_{\tau+1} - \bar{\boldsymbol{p}}\|_2^2\right] \leq \mathbb{E}\left[\|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2^2\right](1 - 2\eta_\tau^d L^d) + (u_G\eta_\tau^d)^2$ holds. Since the step size is set as $\eta_\tau^d = \eta_0^d/(\tau+1)$ with $\eta_0^d > 1/L^d$, we have

$$\mathbb{E}\left[\|\check{\boldsymbol{p}}_{\tau+1} - \bar{\boldsymbol{p}}\|_2^2\right] \leq\ \mathbb{E}\left[\|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2^2\right] \cdot \frac{\tau - 1}{\tau + 1} + \frac{(u_G\eta_0^d)^2}{(\tau+1)^2}.$$

The required result in Lemma 1 is equivalent to

$$\mathbb{E}\left[\|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2^2\right] \leq\ \kappa^d/\tau, \qquad \forall\ \tau \geq 1, \tag{D.1}$$

where we recall $\kappa^d := (u_G\eta_0^d)^2$. We will prove (D.1) by induction on τ. We defer the base case until later and first address the induction step.

Induction Step. Suppose that (D.1) holds for some $\tau \geq 1$. Then, the recursive equation above implies

$$\mathbb{E}\left[\|\check{\boldsymbol{p}}_{\tau+1} - \bar{\boldsymbol{p}}\|_2^2\right] \leq\ \frac{\kappa^d}{\tau} \cdot \frac{\tau-1}{\tau+1} + \frac{(u_G\eta_0^d)^2}{(\tau+1)^2} =\ \frac{\kappa^d}{\tau+1} \cdot \left(1 - \frac{1}{\tau} + \frac{1}{\tau+1}\right) \leq\ \frac{\kappa^d}{\tau+1},$$

proving (D.1) for $\tau + 1$.

Base Case. To prove (D.1) for $\tau = 1$, it remains to check $\mathbb{E}\left[\|\check{\boldsymbol{p}}_1 - \bar{\boldsymbol{p}}\|_2^2\right] \leq \kappa^d = (u_G\eta_0^d)^2$. Indeed, we can show that $\|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2^2 \leq \kappa^d$ holds for any $\boldsymbol{p}$ other than the null price $\boldsymbol{p}_\infty$. Recall that we have found $\mathbf{G}(\boldsymbol{p})$ such that (i) $\mathbb{E}[\mathbf{G}(\boldsymbol{p})|\boldsymbol{p}]^\top(\boldsymbol{p} - \bar{\boldsymbol{p}}) \geq L^d\|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2^2$ and (ii) $\|\mathbf{G}(\boldsymbol{p})\|_2 \leq u_G$ almost surely (see Section 4.4). These two conditions imply $L^d\|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2^2 \leq \mathbb{E}[\mathbf{G}(\boldsymbol{p})|\boldsymbol{p}]^\top(\boldsymbol{p} - \bar{\boldsymbol{p}}) \leq u_G\|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2$. Combining it with $\eta_0^d > 1/L^d$, we obtain

$$\|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2^2 \leq (u_G/L^d)^2 < (u_G\eta_0^d)^2 = \kappa^d, \quad \forall\ \boldsymbol{p} \neq \boldsymbol{p}_\infty. \tag{D.2}$$

Hence, we obtain $\mathbb{E}\left[\|\check{\boldsymbol{p}}_1 - \bar{\boldsymbol{p}}\|_2^2\right] \leq \kappa^d = (u_G\eta_0^d)^2$, completing the proof of the base case. Thus, we prove Lemma 1.

D.2 Proof of (12)

Now we prove (12), i.e., $\left(\hat{\boldsymbol{\beta}} \odot [\boldsymbol{\lambda}(\bar{p}) - \boldsymbol{\lambda}(\boldsymbol{p})]\right)^\top (\boldsymbol{p} - \bar{\boldsymbol{p}}) \geq L^d \|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2^2$, where $L^d := (\beta_{\min}\lambda_{\min})^2/2 > 0$. This justifies our choice of $\mathbf{G}(\boldsymbol{p})$ when generalizing the SGD method.

Recall that under the MNL model, the partial derivatives are given by

$$\frac{\partial}{\partial p_i}\lambda_i(\boldsymbol{p}) \;=\; -\beta_i\lambda_i(\boldsymbol{p})[1-\lambda_i(\boldsymbol{p})] = -\beta_i\lambda_i(\boldsymbol{p})\sum_{j\neq i}\lambda_j(\boldsymbol{p}) \qquad \text{and} \qquad \frac{\partial}{\partial p_j}\lambda_i(\boldsymbol{p}) \;=\; \beta_j\lambda_i(\boldsymbol{p})\lambda_j(\boldsymbol{p}) \quad \forall\, j \neq i.$$

By the mean value theorem, we can rewrite the left hand side of (12) as

$$\left(\hat{\boldsymbol{\beta}} \odot [\boldsymbol{\lambda}(\bar{\boldsymbol{p}}) - \boldsymbol{\lambda}(\boldsymbol{p})]\right)^\top (\boldsymbol{p} - \bar{\boldsymbol{p}}) = (\boldsymbol{p} - \bar{\boldsymbol{p}})^\top \hat{H} (\boldsymbol{p} - \bar{\boldsymbol{p}}),$$

where

$$\hat{H} := \begin{bmatrix} \hat{\beta}_1\beta_1\int_0^1 \lambda_1^u(1-\lambda_1^u)\,du & -\hat{\beta}_2\beta_1\int_0^1 \lambda_1^u\lambda_2^u\,du & \cdots & -\hat{\beta}_M\beta_1\int_0^1 \lambda_1^u\lambda_M^u\,du \\ -\hat{\beta}_1\beta_2\int_0^1 \lambda_1^u\lambda_2^u\,du & \hat{\beta}_2\beta_2\int_0^1 \lambda_2^u(1-\lambda_2^u)\,du & \cdots & -\hat{\beta}_M\beta_2\int_0^1 \lambda_2^u\lambda_M^u\,du \\ \vdots & \vdots & & \vdots \\ -\hat{\beta}_1\beta_M\int_0^1 \lambda_1^u\lambda_M^u\,du & -\hat{\beta}_2\beta_M\int_0^1 \lambda_2^u\lambda_M^u\,du & \cdots & \hat{\beta}_M\beta_M\int_0^1 \lambda_M^u(1-\lambda_M^u)\,du \end{bmatrix}$$

with $\lambda_i^u := \lambda_i(\boldsymbol{p}^u)$ and $\boldsymbol{p}^u := \boldsymbol{p} + u(\bar{\boldsymbol{p}} - \boldsymbol{p})$ for some constant $u \in [0,1]$. To prove (12), it suffices to show $\hat{H} \succeq L^d\mathbf{I}_M = \frac{1}{2}(\beta_{\min}\lambda_{\min})^2\mathbf{I}_M$, i.e., $\boldsymbol{x}^\top\hat{H}\boldsymbol{x} \geq L^d = (\beta_{\min}\lambda_{\min})^2/2$ for any $\boldsymbol{x} = [x_1,\ldots,x_M]^\top \in \mathbb{R}^M \setminus \{\mathbf{0}\}$.

Using $\lambda_0^u + \lambda_1^u + \cdots + \lambda_M^u = 1$, we can write $\hat{H}$ as a sum of two square matrices:

$$\hat{H} := \begin{bmatrix} \hat{\beta}_1\beta_1\int_0^1 \lambda_1^u\lambda_0^u\,du & & & \\ & \hat{\beta}_2\beta_2\int_0^1 \lambda_2^u\lambda_0^u\,du & & \\ & & \ddots & \\ & & & \hat{\beta}_M\beta_M\int_0^1 \lambda_M^u\lambda_0^u\,du \end{bmatrix}$$
$$+ \begin{bmatrix} \hat{\beta}_1\beta_1\int_0^1 \lambda_1^u(\lambda_2^u+\cdots+\lambda_M^u)\,du & -\hat{\beta}_2\beta_1\int_0^1 \lambda_1^u\lambda_2^u\,du & \ldots & -\hat{\beta}_M\beta_1\int_0^1 \lambda_1^u\lambda_M^u\,du \\ -\hat{\beta}_1\beta_2\int_0^1 \lambda_1^u\lambda_2^u\,du & \hat{\beta}_2\beta_2\int_0^1 \lambda_2^u(\lambda_1^u+\lambda_3^u+\cdots+\lambda_M^u)\,du & \ldots & -\hat{\beta}_M\beta_2\int_0^1 \lambda_2^u\lambda_M^u\,du \\ \vdots & \vdots & & \vdots \\ -\hat{\beta}_1\beta_M\int_0^1 \lambda_1^u\lambda_M^u\,du & -\hat{\beta}_2\beta_M\int_0^1 \lambda_2^u\lambda_M^u\,du & \ldots & \hat{\beta}_M\beta_M\int_0^1 \lambda_M^u(\lambda_1^u+\cdots+\lambda_{M-1}^u)\,du \end{bmatrix}.$$

Thus, for any $\boldsymbol{x} = [x_1,\ldots,x_M]^\top \in \mathbb{R}^M \setminus \{\mathbf{0}\}$, we have

$$\boldsymbol{x}^\top\hat{H}\boldsymbol{x} = \sum_{i=1}^M \hat{\beta}_i\beta_i x_i^2 \int_0^1 \lambda_i^u\lambda_0^u\,du + \sum_{i=1}^M\sum_{j=i+1}^M \left[\hat{\beta}_i\beta_i x_i^2 + \hat{\beta}_j\beta_j x_j^2 - (\hat{\beta}_i\beta_j + \hat{\beta}_j\beta_i)x_i x_j\right]\int_0^1 \lambda_i^u\lambda_j^u\,du.$$

Recall that $\beta_i \in [\beta_{\min}, \beta_{\max}] \subset (0,\infty)$ for all i. When the estimate for i satisfies $\hat{\beta}_i \in [\beta_{\min}, \beta_{\max}]$ and

$|\hat{\beta}_i - \beta_i| \le \varepsilon_0$, we obtain $\hat{\beta}_i\beta_i \ge \beta_{\min}^2$ and

$$\begin{aligned}
&\hat{\beta}_i\beta_i x_i^2 + \hat{\beta}_j\beta_j x_j^2 - (\hat{\beta}_i\beta_j + \hat{\beta}_j\beta_i)x_i x_j \\
&\quad = (\beta_i x_i - \beta_j x_j)^2 + (\hat{\beta}_i - \beta_i)(\beta_i x_i^2 - \beta_j x_i x_j) + (\hat{\beta}_j - \beta_j)(\beta_j x_j^2 - \beta_i x_i x_j) \\
&\quad \ge -|\hat{\beta}_i - \beta_i| \cdot (\beta_i x_i^2 + \beta_j |x_i||x_j|) - |\hat{\beta}_j - \beta_j| \cdot (\beta_j x_j^2 + \beta_i |x_i||x_j|) \\
&\quad \ge -\varepsilon_0 \cdot \beta_{\max}(x_i^2 + |x_i||x_j|) - \varepsilon_0 \cdot \beta_{\max}(|x_i||x_j| + x_j^2) \\
&\quad \ge -2\varepsilon_0\beta_{\max}(x_i^2 + x_j^2),
\end{aligned}$$

where the last inequality follows from $|x_i||x_j| \le (x_i^2 + x_j^2)/2$. Thus, we obtain

$$\boldsymbol{x}^\top \hat{H} \boldsymbol{x} \ge \sum_{i=1}^M \beta_{\min}^2 x_i^2 \int_0^1 \lambda_i^u \lambda_0^u \, du - 2\varepsilon_0\beta_{\max} \sum_{i=1}^M \sum_{j=i+1}^M (x_i^2 + x_j^2) \int_0^1 \lambda_i^u \lambda_j^u \, du.$$

Since $\lambda_i(\boldsymbol{p}) \in [\lambda_{\min}, \lambda_{\max}] \subset (0,1)$ for any $\boldsymbol{p} \in \mathcal{P}$ and any $i \in \{0, 1, \ldots, M\}$, we have $\int_0^1 \lambda_i^u \lambda_0^u \, du \ge \lambda_{\min}^2$. Note also that

$$\sum_{i=1}^M \sum_{j=i+1}^M (x_i^2 + x_j^2) \int_0^1 \lambda_i^u \lambda_j^u \, du = \sum_{i=1}^M x_i^2 \int_0^1 \lambda_i^u (1 - \lambda_i^u - \lambda_0^u) \, du \;\le\; \frac{1}{4}(1 - \lambda_{\min})^2 \sum_{i=1}^M x_i^2,$$

where the last inequality follows from $\lambda_i^u(1 - \lambda_i^u - \lambda_0^u) \le \lambda_i^u(1 - \lambda_i^u - \lambda_{\min}) \le (1 - \lambda_{\min})^2/4$ for any $\lambda_i^u \in (0, 1 - \lambda_{\min})$. This implies

$$\boldsymbol{x}^\top \hat{H} \boldsymbol{x} \ge \sum_{i=1}^M x_i^2 \left(\beta_{\min}^2 \lambda_{\min}^2 - \varepsilon_0 \beta_{\max} \cdot \frac{1}{2}(1 - \lambda_{\min})^2 \right).$$

As we define $\varepsilon_0 = \frac{(\beta_{\min}\lambda_{\min})^2}{\beta_{\max}(1-\lambda_{\min})^2}$, we obtain $\hat{H} \succeq \frac{1}{2}(\beta_{\min}\lambda_{\min})^2 \mathbf{I}_M$, as required.

D.3 Proof of Proposition 4

In the updating rule (13), we choose $\mathbf{G}(\boldsymbol{p}) = \hat{\boldsymbol{\beta}} \odot [\hat{\boldsymbol{\omega}}^* - \boldsymbol{D}(\boldsymbol{p})]$, which satisfies (i) $\mathbb{E}[\mathbf{G}(\boldsymbol{p})|\boldsymbol{p}]^\top(\boldsymbol{p} - \bar{\boldsymbol{p}}) \ge L^d \|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_2^2$ and (ii) $\|\mathbf{G}(\boldsymbol{p})\|_2 \le u_G$ under E_2. By extending the argument for (11) conditional on E_2, we can derive the expectation bound (14).

Next, we prove the high-probability bound in (15) using Lemma C.1. By conditions (i) and (ii), we can derive that

$$\begin{aligned}
&\|\check{\boldsymbol{p}}_{\tau+1} - \bar{\boldsymbol{p}}\|_2^2 \;\le\; \|\check{\boldsymbol{p}}_\tau - \eta_\tau^d \mathbf{G}(\check{\boldsymbol{p}}_\tau) - \bar{\boldsymbol{p}}\|_2^2 \\
&= \|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2^2 - 2\eta_\tau^d \mathbf{G}(\check{\boldsymbol{p}}_\tau)^\top (\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}) + (\eta_\tau^d)^2 \|\mathbf{G}(\check{\boldsymbol{p}}_\tau)\|_2^2 \\
&= \|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2^2 - 2\eta_\tau^d \mathbb{E}[\mathbf{G}(\check{\boldsymbol{p}}_\tau)|\check{\boldsymbol{p}}_\tau]^\top (\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}) + 2\eta_\tau^d \left(\mathbb{E}[\mathbf{G}(\check{\boldsymbol{p}}_\tau)|\check{\boldsymbol{p}}_\tau] - \mathbf{G}(\check{\boldsymbol{p}}_\tau)\right)^\top (\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}) + (\eta_\tau^d)^2 \|\mathbf{G}(\check{\boldsymbol{p}}_\tau)\|_2^2 \\
&\le \|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2^2 (1 - 2\eta_\tau^d L^d) + 2\eta_\tau^d \left(\mathbb{E}[\mathbf{G}(\check{\boldsymbol{p}}_\tau)|\check{\boldsymbol{p}}_\tau] - \mathbf{G}(\check{\boldsymbol{p}}_\tau)\right)^\top (\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}) + (u_G \eta_\tau^d)^2.
\end{aligned}$$

Recall that the step size is set as $\eta_\tau^d = \eta_0^d/(\tau+1)$ with $\eta_0^d > 1/L^d$. It follows that $1 - 2\eta_\tau^d L^d < 1 - 2/(\tau+1) =$

$(\tau-1)/(\tau+1)$, and thus,

$$(\tau+1)\|\check{\boldsymbol{p}}_{\tau+1}-\bar{\boldsymbol{p}}\|_2^2 \ \le\ \tau\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2^2\cdot\frac{\tau-1}{\tau} \\ +\frac{4u_G\eta_0^d}{\sqrt{\tau}}\left(\frac{\left(\mathbb{E}[\mathbf{G}(\check{\boldsymbol{p}}_\tau)|\check{\boldsymbol{p}}_\tau]-\mathbf{G}(\check{\boldsymbol{p}}_\tau)\right)^\top(\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}})}{2u_G\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2}\right)\sqrt{\tau\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2^2}+\frac{(u_G\eta_0^d)^2}{\tau+1}.$$

Let $X_\tau := \tau\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2^2$. Then, $X_{\tau+1}\le\alpha_\tau X_\tau+\beta_\tau\hat{w}_\tau\sqrt{X_\tau}+\gamma_\tau$ holds for all $\tau\ge 1$ with $\alpha_\tau=(\tau-1)/\tau$, $\beta_\tau=4u_G\eta_0^d/\sqrt{\tau}$, $\hat{w}_\tau=\left(\mathbb{E}[\mathbf{G}(\check{\boldsymbol{p}}_\tau)|\check{\boldsymbol{p}}_\tau]-\mathbf{G}(\check{\boldsymbol{p}}_\tau)\right)^\top(\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}})/(2u_G\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2)$, and $\gamma_\tau=(u_G\eta_0^d)^2/(\tau+1)$.

Let $\mathcal{F}_\tau$ be the σ-algebra generated by the history of selling prices and realized supplies before applying the price vector $\boldsymbol{p}_\tau$. Then, X_τ is $\mathcal{F}_\tau$-measurable, and moreover, $\mathbb{E}[\hat{w}_\tau|\mathcal{F}_\tau]=0$. Observe that $|\hat{w}_\tau|\le\|\mathbb{E}[\mathbf{G}(\check{\boldsymbol{p}}_\tau)|\check{\boldsymbol{p}}_\tau]-\mathbf{G}(\check{\boldsymbol{p}}_\tau)\|_2/(2u_G)\le 1$, and $\max_{1\le\tau\le\tau_T}\left(2\gamma_\tau/(1-\alpha_\tau),2\beta_\tau^2/(1-\alpha_\tau)\right)=32(u_G\eta_0^d)^2$. Let $K_p:=32(u_G\eta_0^d)^2$. By (D.2), we have $X_1=\|\check{\boldsymbol{p}}_1-\bar{\boldsymbol{p}}\|_2^2\le(u_G\eta_0^d)^2$, implying $\mathbb{E}[\exp(\lambda X_1)]\le\exp(\lambda K_p)$ for any $\lambda\in(0,1/K_p]$. We can apply Lemma C.1 with $\sigma_\tau=1/\tau$ and $\delta=1/(eT)$, which implies

$$\mathbb{P}\left(\sum_{\tau=1}^{\tau_T}\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2^2\ge K_p\log(eT)\sum_{\tau=1}^{\tau_T}(1/\tau)\middle|E_2\right)\le\frac{1}{T}.$$

Since $\sum_{\tau=1}^{\tau_T}(1/\tau)\le 1+\int_1^{\tau_T}(1/x)\,dx=1+\log(\tau_T)\le\log(eT)$, we have

$$\mathbb{P}\left(\sum_{\tau=1}^{\tau_T}\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2\ge\sqrt{K_p\tau_T}\cdot\log(eT)\middle|E_2\right)\le\ \mathbb{P}\left(\sum_{\tau=1}^{\tau_T}\|\check{\boldsymbol{p}}_\tau-\bar{\boldsymbol{p}}\|_2^2\ge K_p\log(eT)\sum_{\tau=1}^{\tau_T}(1/\tau)\middle|E_2\right)\le\ \frac{1}{T},$$

where the first inequality follows from Cauchy–Schwarz inequality. This proves (15) as desired.

Appendix E Omitted Proofs in Section 5

E.1 Proof of (17)

In this section, we prove (17), that is, $\mathbb{E}[\|\boldsymbol{\omega}^*-\hat{\boldsymbol{\omega}}\|_2^2]\le C_{est}\cdot M/N_0$ for some finite constant C_{est}.

We note that the estimation error of the matching rate comes from errors of the estimated demand and supply parameters, that is,

$$\|\boldsymbol{\omega}^*-\hat{\boldsymbol{\omega}}^*\|_2^2\ \le\ c_1\sum_{i=1}^M\left(\|\boldsymbol{\theta}_i-\hat{\boldsymbol{\theta}}_i\|_2^2+\|\boldsymbol{\gamma}_i-\hat{\boldsymbol{\gamma}}_i\|_2^2\right), \tag{E.1}$$

where c_1 is a finite constant, and we will provide a detailed proof of (E.1) below. For the supply side, each purchase price vector is tested for N_0 periods. By the classic results of the MLE estimator (see, e.g., Borokov 1998), we obtain that the estimation error satisfies

$$\mathbb{E}\left[\|\hat{\boldsymbol{\gamma}}_i-\boldsymbol{\gamma}_i\|_2^2\right]\ \le\ c_0^s/N_0,\quad\forall\, i\in[M],$$

where c_0^s is a finite constant independent of M and N_0. For the demand side, we only collect information when all products have positive inventory. Define E_1 as the event that the number of tests at each selling price vector satisfies $|\mathcal{T}_{0,1}| \geq N_0/2$ and $|\mathcal{T}_{0,2}| \geq N_0/2$. The classic results of the MLE estimator imply

$$\mathbb{E}\left[\|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_2^2 \middle| E_1\right] \;\leq\; c_0^d/N_0, \quad \forall\; i \in [M], \tag{E.2}$$

where c_0^d is a finite constant independent of M and N_0. Moreover, we can show that E_1 is a high probability event with its complement satisfying

$$\mathbb{P}\left(E_1^C\right) \leq c_0^{(1)} M/N_0. \tag{E.3}$$

where $c_0^{(1)}$ is a finite constant, and we will prove (E.3) later.

We bound the squared estimation error of the matching rate under the event E_1 and its complement E_1^C. Specifically, under the event E_1, we bound it by estimation errors of the demand and supply parameters as in (E.1). When E_1 does not occur, this occurs with small probability, and we simply use the bound $\|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}^*\|_2^2 = \sum_{i=1}^M (\omega_i^* - \hat{\omega}_i^*)^2 \leq \sum_{i=1}^M (\omega_i^*)^2 + \sum_{i=1}^M (\hat{\omega}_i^*)^2 \leq (\sum_{i=1}^M \omega_i^*)^2 + (\sum_{i=1}^M \hat{\omega}_i^*)^2 \leq 2$. Thus, we obtain the following bound on the expected value of the squared estimation error:

$$\begin{aligned}
\mathbb{E}\left[\|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}^*\|_2^2\right] =&\; \mathbb{E}\left[\|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}^*\|_2^2 \middle| E_1\right] \mathbb{P}\left(E_1\right) + \mathbb{E}\left[\|\boldsymbol{\omega}^* - \hat{\boldsymbol{\omega}}^*\|_2^2 \middle| E_1^C\right] \mathbb{P}\left(E_1^C\right) \\
\leq&\; c_1 \sum_{i=1}^M \mathbb{E}\left[\|\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i\|_2^2 + \|\boldsymbol{\gamma}_i - \hat{\boldsymbol{\gamma}}_i\|_2^2 \middle| E_1\right] \mathbb{P}\left(E_1\right) + 2\mathbb{P}\left(E_1^C\right) \\
\leq&\; c_1 \sum_{i=1}^M \left(c_0^d/N_0 + c_0^s/N_0\right) + 2c_0^{(1)} \cdot M/N_0 \\
=&\; C_{est} \cdot M/N_0,
\end{aligned}$$

where $C_{est} := c_1(c_0^d + c_0^s) + 2c_0^{(1)}$ is a finite constant, and the second inequality follows from the inequalities stated earlier in this proof. Thus, we obtain (17), and it remains to show (E.1) and (E.3).

Proof of (E.1)

We need to show that the estimation error of the matching rate can be bounded by estimation errors of the demand and supply parameters as in (E.1). As discussed in Section 3.2, the one-period profit $f(\boldsymbol{\omega})$ is concave in $\boldsymbol{\omega}$, and thus, the optimal matching rate $\boldsymbol{\omega}^*$ satisfies the first-order condition, that is, $\partial f(\boldsymbol{\omega})/\partial \omega_i = 0$ holds at $\boldsymbol{\omega} = \boldsymbol{\omega}^*$: for any $i \in [M]$,

$$0 = \frac{1}{\beta_i}[\alpha_i + \log(\omega_0) - \log(\omega_i) - 1] - \frac{1}{\omega_0}\sum_{j=1}^M \frac{\omega_j}{\beta_j} + \frac{\gamma_{1,i}}{\gamma_{2,i}} - \frac{2}{\gamma_{2,i}}\left(\omega_i - \frac{I_{i,0}}{T}\right),$$

where $\omega_0 := 1 - \sum_{j=1}^M \omega_j$.

To obtain the estimate $\hat{\boldsymbol{\omega}}^*$, we replace $\theta_i = (\alpha_i, \beta_i)$ and $\boldsymbol{\gamma}_i = (\gamma_{1,i}, \gamma_{2,i})$ by their estimates $\hat{\theta}_i$ and $\hat{\omega}_i$ above and solve for $\hat{\boldsymbol{\omega}}^*$. Let $\boldsymbol{\eta}^\top := [\boldsymbol{\eta}_1^\top, \boldsymbol{\eta}_2^\top, \boldsymbol{\eta}_3^\top, \boldsymbol{\eta}_4^\top]$, where $\boldsymbol{\eta}_1^\top := (\alpha_1/\beta_1, \ldots, \alpha_M/\beta_M)$, $\boldsymbol{\eta}_2^\top := (1/\beta_1, \ldots, 1/\beta_M)$, $\boldsymbol{\eta}_3^\top := (\gamma_{1,1}/\gamma_{2,1}, \ldots, \gamma_{1,M}/\gamma_{2,M})$, and $\boldsymbol{\eta}_4^\top := (1/\gamma_{2,1}, \ldots, 1/\gamma_{2,M})$. With abuse of notation, denote by $\boldsymbol{\omega}^*(\boldsymbol{\eta})$

the solution to the first-order condition for any parameter vector $\boldsymbol{\eta}$. Then, the squared error is

$$\|\boldsymbol{\omega}^*(\boldsymbol{\eta}) - \boldsymbol{\omega}^*(\hat{\boldsymbol{\eta}})\|_2^2 = \|\int_0^1 \sum_{k=1}^4 \left[\frac{\partial \boldsymbol{\omega}^*(\tilde{\boldsymbol{\eta}}^\alpha)}{\partial \boldsymbol{\eta}_k}\right]^\top (\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)\, d\alpha\|_2^2 \leq 4 \int_0^1 \sum_{k=1}^4 \| \left[\frac{\partial \boldsymbol{\omega}^*(\tilde{\boldsymbol{\eta}}^\alpha)}{\partial \boldsymbol{\eta}_k}\right]^\top (\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)\|_2^2\, d\alpha, \quad \text{(E.4)}$$

where $\hat{\boldsymbol{\eta}}$ is the parameter vector corresponding to the estimates $\hat{\theta}_i$ and $\hat{\gamma}_i$, $\tilde{\boldsymbol{\eta}}^\alpha := \alpha\boldsymbol{\eta} + (1-\alpha)\hat{\boldsymbol{\eta}}$, and the last inequality follows from Cauchy-Schwarz inequality.

Taking derivative over $\boldsymbol{\eta}$ on both sides of the first-order condition, we derive that

$$\mathbf{H}_f(\boldsymbol{\omega}^*(\boldsymbol{\eta}))\frac{\partial \boldsymbol{\omega}^*(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}_k} = A^{(k)},$$

where $\mathbf{H}_f(\boldsymbol{\omega})$ is the Hessian matrix of the one-period expected profit $f(\boldsymbol{\omega})$, $A^{(1)} = A^{(3)} = -\boldsymbol{I}_M$, and $A^{(2)}$ and $A^{(4)}$ are M-dimensional matrices with entries

$$\begin{aligned} A_{ii}^{(2)} &= \log\left(\frac{\omega_i^*}{\omega_0^*}\right) + \frac{\omega_i^*}{\omega_0^*} + 1, \quad A_{ij}^{(2)} = \frac{\omega_j^*}{\omega_0^*}, \\ A_{ii}^{(4)} &= 2(\omega_i^* - I_{i,0}/T), \quad A_{ij}^{(4)} = 0, \qquad \forall\ i \neq j. \end{aligned}$$

To bound the terms in (E.4), we bound below $\|[\mathbf{H}_f(\boldsymbol{\omega}^*)]^{-1}A^{(k)}\|_2$, where we fix $\alpha \in [0,1]$ and suppress the argument in $\boldsymbol{\omega}^*(\tilde{\boldsymbol{\eta}}^\alpha)$ to simplify notation. We first consider $A^{(k)}$ for $k \in \{1,3,4\}$. By concavity of the objective, we have $-\mathbf{H}_f(\boldsymbol{\omega}^*) \succeq \ell_f \boldsymbol{I}_M$ for any $\boldsymbol{\eta}$ and some constant $\ell_f > 0$. It follows that

$$\|[H_f(\boldsymbol{\omega})]^{-1}A^{(1)}\|_2 = \|[H_f(\boldsymbol{\omega})]^{-1}A^{(3)}\|_2 \leq \ell_f^{-1} \quad \text{and} \quad \|[H_f(\boldsymbol{\omega})]^{-1}A^{(4)}\|_2 \leq 2\ell_f^{-1}.$$

Next, we consider $A^{(2)}$. This matrix can be decomposed as $A_{\log} + A_-$, where $A_{\log} := \operatorname{diag}(\log(\omega_1^*/\omega_0^*) + 1, \ldots, \log(\omega_i^*/\omega_0^*) + 1)$, and $A_- = \mathbf{1}(\boldsymbol{\omega}^*)^\top/\omega_0^*$ has the same entries ω_j^*/ω_0^* in column j. The first matrix satisfies $\|A_{\log}\|_2 = \max_i |\log(\omega_i^*/\omega_0^*) + 1| \leq \log(\lambda_{\max}/\lambda_{\min}) + 1 \leq \lambda_{\max}/\lambda_{\min}$, implying $\|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}A_{\log}\|_2 \leq \|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}\|_2\|A_{\log}\|_2 \leq \ell_f^{-1}\lambda_{\max}/\lambda_{\min}$. For the second matrix, we derive that

$$\|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}A_-\|_2^2 = \left\|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}\frac{\mathbf{1}(\boldsymbol{\omega}^*)^\top}{\omega_0^*}\right\|_2^2 = \frac{\|\boldsymbol{\omega}^*\|_2^2}{(\omega_0^*)^2}\mathbf{1}^\top(-\mathbf{H}_f(\boldsymbol{\omega}^*))^{-2}\mathbf{1} \quad \text{(E.5)}$$

A direct calculation using the MNL structure gives, for any $\boldsymbol{z} \in \mathbb{R}^M$,

$$\begin{aligned} -\boldsymbol{z}^\top\mathbf{H}_f(\boldsymbol{\omega}^*)\boldsymbol{z} &= \sum_{i=1}^M \tilde{\eta}_{2,i}^\alpha \omega_i^* \left(\frac{z_i}{\omega_i^*} + \frac{\mathbf{1}^\top\boldsymbol{z}}{\omega_0^*}\right)^2 + 2\sum_{i=1}^M \tilde{\eta}_{4,i}^\alpha z_i^2 \\ &\geq \sum_{i=1}^M \frac{\omega_i^*}{\beta_{\max}}\left(\frac{2z_i(\mathbf{1}^\top\boldsymbol{z})}{\omega_i^*\omega_0^*} + \left(\frac{\mathbf{1}^\top\boldsymbol{z}}{\omega_0^*}\right)^2\right) = \frac{1+\omega_0^*}{\beta_{\max}(\omega_0^*)^2}(\mathbf{1}^\top\boldsymbol{z})^2, \end{aligned}$$

where the inequality uses $\tilde{\eta}_{4,i} \geq 1/\beta_{\max}$ and nonnegativity of z_i^2 terms. Choosing $\boldsymbol{z} := -\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}\mathbf{1}$ yields

$$\mathbf{1}^\top(-\mathbf{H}_f(\boldsymbol{\omega}^*))^{-1}\mathbf{1} \geq \frac{1+\omega_0^*}{\beta_{\max}(\omega_0^*)^2}(\mathbf{1}^\top(-\mathbf{H}_f(\boldsymbol{\omega}^*))^{-1})^2,$$

implying that $\mathbf{1}^\top(-\mathbf{H}_f(\boldsymbol{\omega}^*))^{-1}\mathbf{1} \leq \beta_{\max}(\omega_0^*)^2/(1+\omega_0^*)$. Moreover, $-\mathbf{H}_f(\boldsymbol{\omega}^*) \succeq \ell_f I_M$ implies $\mathbf{1}^\top(-\mathbf{H}_f(\boldsymbol{\omega}^*))^{-2}\mathbf{1} \leq$

$\mathbf{1}^\top \ell_f^{-1}(-\mathbf{H}_f(\boldsymbol{\omega}^*))^{-1}\mathbf{1}$. Combining the analysis of the two matrices with (E.5) yields

$$\|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}A_-\|_2^2 \le \frac{\|\boldsymbol{\omega}^*\|_2^2}{(\omega_0^*)^2} \cdot \frac{\beta_{\max}(\omega_0^*)^2}{\ell_f(1+\omega_0^*)} < \frac{\beta_{\max}}{\ell_f},$$

where the last inequality uses $\|\boldsymbol{\omega}^*\|_2^2 \le \|\boldsymbol{\omega}^*\|_1^2 = (1-\omega_0^*)^2 \le 1$ and $\omega_0^* > 0$. Thus, we obtain

$$\|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}A^{(2)}\|_2 \le \|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}A_{\log}\|_2 + \|\mathbf{H}_f(\boldsymbol{\omega}^*)^{-1}A_-\|_2 \le \ell_f^{-1}\lambda_{\max}/\lambda_{\min} + \sqrt{\beta_{\max}/\ell_f}.$$

Define $a_f := \max\{4,\ (\lambda_{\max}/\lambda_{\min} + \sqrt{\beta_{\max}\ell_f})^2\}$, and we have $\|[\mathbf{H}_f(\boldsymbol{\omega}^*)]^{-1}A^{(k)}\|_2^2 \le a_f\ell_f^{-2}$ for all k, which holds uniformly over $\alpha \in [0,1]$.

Combining the above results with (E.4), we obtain

$$\|\boldsymbol{\omega}^*(\boldsymbol{\eta}) - \boldsymbol{\omega}^*(\hat{\boldsymbol{\eta}})\|_2^2 \le\ 4a_f\ell_f^{-2}\sum_{k=1}^{4}\|\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k\|_2^2 =\ 4a_f\ell_f^{-2}\|\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}\|_2^2.$$

By definition of $\boldsymbol{\eta}$, we have

$$\|\boldsymbol{\eta} - \hat{\boldsymbol{\eta}}\|_2^2 =\ \sum_{i=1}^{M}\left[\left(\frac{\alpha_i}{\beta_i} - \frac{\hat{\alpha}_i}{\hat{\beta}_i}\right)^2 + \left(\frac{1}{\beta_i} - \frac{1}{\hat{\beta}_i}\right)^2\right] + \sum_{i=1}^{M}\left[\left(\frac{\gamma_{1,i}}{\gamma_{2,i}} - \frac{\hat{\gamma}_{1,i}}{\hat{\gamma}_{2,i}}\right)^2 + \left(\frac{1}{\gamma_{2,i}} - \frac{1}{\hat{\gamma}_{2,i}}\right)^2\right].$$

We provide upper bounds on each term inside the square brackets above. Note that

$$\begin{aligned}
\left(\frac{\alpha_i}{\beta_i} - \frac{\hat{\alpha}_i}{\hat{\beta}_i}\right)^2 + \left(\frac{1}{\beta_i} - \frac{1}{\hat{\beta}_i}\right)^2 &=\ \left(\frac{\alpha_i - \hat{\alpha}_i}{\beta_i} + \hat{\alpha}_i\left(\frac{1}{\hat{\beta}_i} - \frac{1}{\beta_i}\right)\right)^2 + \left(\frac{1}{\beta_i} - \frac{1}{\hat{\beta}_i}\right)^2 \\
&\le\ 2\left(\frac{\alpha_i - \hat{\alpha}_i}{\beta_i}\right)^2 + (2\hat{\alpha}_i^2 + 1)\left(\frac{1}{\hat{\beta}_i} - \frac{1}{\beta_i}\right)^2 \\
&\le\ \frac{2}{\beta_{\min}^2}(\alpha_i - \hat{\alpha}_i)^2 + \frac{2\alpha_{\max}^2 + 1}{\beta_{\min}^4}\left(\beta_i - \hat{\beta}_i\right)^2.
\end{aligned}$$

where the last inequality hold because $\beta_i, \hat{\beta}_i \ge \beta_{\min}$ and $\hat{\alpha}_i \le \alpha_{\max}$. Similarly, we can derive that

$$\left(\frac{\gamma_{1,i}}{\gamma_{2,i}} - \frac{\hat{\gamma}_{1,i}}{\hat{\gamma}_{2,i}}\right)^2 + \left(\frac{1}{\gamma_{2,i}} - \frac{1}{\hat{\gamma}_{2,i}}\right)^2 \ \le\ \frac{2}{\gamma_{2,\min}^2}(\gamma_{1,i} - \hat{\gamma}_{1,i})^2 + \frac{2\gamma_{1,\max}^2 + 1}{\gamma_{2,\min}^4}(\gamma_{2,i} - \hat{\gamma}_{2,i})^2.$$

Let $c_1 := 4a_f\ell_f^{-2} \cdot \max\left\{2/\beta_{\min}^2, (2\alpha_{\max}^2+1)/\beta_{\min}^4, 2/\gamma_{2,\min}^2, (2\gamma_{1,\max}^2+1)/\gamma_{2,\min}^4\right\}$, which is a finite constant. Then, we obtain

$$\begin{aligned}
\|\boldsymbol{\omega}^*(\boldsymbol{\eta}) - \boldsymbol{\omega}^*(\hat{\boldsymbol{\eta}})\|_2^2 \le&\ c_1\sum_{i=1}^{M}\left[(\alpha_i - \hat{\alpha}_i)^2 + \left(\beta_i - \hat{\beta}_i\right)^2 + (\gamma_{1,i} - \hat{\gamma}_{1,i})^2 + (\gamma_{2,i} - \hat{\gamma}_{2,i})^2\right] \\
=&\ c_1\sum_{i=1}^{M}\left(\|\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i\|_2^2 + \|\boldsymbol{\gamma}_i - \hat{\boldsymbol{\gamma}}_i\|_2^2\right),
\end{aligned}$$

completing the proof of (E.1).

Proof of (E.3)

We need to show that E_1 is a high probability event. We can bound the probability of its complement by

$$\begin{aligned}\mathbb{P}\left(E_1^C\right) \leq\ & \mathbb{P}\left(|\mathcal{T}_{0,1}| < N_0/2\right) + \mathbb{P}\left(|\mathcal{T}_{0,2}| < N_0/2\right) \\ =\ & \mathbb{P}\left(\sum_{t=1}^{N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\} > N_0/2\right) + \mathbb{P}\left(\sum_{t=N_0+1}^{2N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\} > N_0/2\right) \\ \leq\ & \frac{\mathbb{E}\left[\sum_{t=1}^{N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\}\right]}{N_0/2} + \frac{\mathbb{E}\left[\sum_{t=N_0+1}^{2N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\}\right]}{N_0/2},\end{aligned}$$

where the supply $S_{i,t}^{\mathcal{A}}$ is independent with rate $\mu_i(\tilde{\boldsymbol{b}}_1)$ for the first N_0 periods and $\mu_i(\tilde{\boldsymbol{b}}_2)$ for the next N_0 periods, and the last inequality follows from Markov's inequality.

We consider the numerators in each term above separately. For the first term, we have

$$\mathbb{E}\left[\sum_{t=1}^{N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\pi} + S_{i,t}^{\pi} = 0\}\right] = \sum_{t=1}^{N_0} \mathbb{P}\left(\bigcup_{i=1}^{M} \{I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\}\right) \leq \sum_{t=1}^{N_0}\sum_{i=1}^{M} \mathbb{P}\left(I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\right).$$

Let $D_{i,\tau}$ be the i.i.d. demand with rate $\lambda_1(\tilde{\boldsymbol{p}}_1)$. Then, we have $I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = I_{i,0} + \sum_{\tau=1}^{t} S_{i,\tau}^{\mathcal{A}} - \sum_{\tau=1}^{t} D_{i,\tau}\mathbb{1}\{\forall\ j : I_{j,\tau}^{\mathcal{A}} + S_{j,\tau}^{\mathcal{A}} > 0\} \geq \sum_{\tau=1}^{t}(S_{i,\tau}^{\mathcal{A}} - D_{i,\tau})$. It follows that

$$\begin{aligned}\mathbb{P}\left(I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\right) \leq\ & \mathbb{P}\left(\sum_{\tau=1}^{t}(S_{i,\tau}^{\mathcal{A}} - D_{i,\tau}) \leq 0\right) = \mathbb{P}\left(\sum_{\tau=1}^{t}[(S_{i,\tau}^{\mathcal{A}} - D_{i,\tau}) - \mathbb{E}(S_{i,\tau}^{\mathcal{A}} - D_{i,\tau})] \leq -t[\mu_i(\tilde{\boldsymbol{b}}_1) - \lambda_i(\tilde{\boldsymbol{p}}_1)]\right) \\ \leq\ & \exp\left(-t[\mu_i(\tilde{\boldsymbol{b}}_1) - \lambda_i(\tilde{\boldsymbol{p}}_1)]^2/2\right),\end{aligned}$$

where the last inequality follows from Hoeffding's inequality. Combining the two results above, we obtain

$$\mathbb{E}\left[\sum_{t=1}^{N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\}\right] \leq \sum_{t=1}^{N_0}\sum_{i=1}^{M} \exp\left(-t[\mu_i(\tilde{\boldsymbol{b}}_1) - \lambda_i(\tilde{\boldsymbol{p}}_1)]^2/2\right) \leq \sum_{i=1}^{M} \frac{2}{[\mu_i(\tilde{b}_{i1}) - \lambda_i(\tilde{\boldsymbol{p}}_1)]^2},$$

where the last inequality follows from $\sum_{t=1}^{N_0} \exp(-tx) \leq \exp(-x)/[1 - \exp(-x)] = 1/[\exp(x) - 1] \leq 1/x$ for any $x > 0$.

Similarly, we can bound the numerator in the second term by

$$\mathbb{E}\left[\sum_{t=N_0+1}^{2N_0} \mathbb{1}\{\exists\ i : I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = 0\}\right] \leq \sum_{i=1}^{M} \frac{2}{[\mu_i(\tilde{b}_{i2}) - \lambda_i(\tilde{\boldsymbol{p}}_2)]^2}.$$

Let $c_0^{(1)} := 4/\min_{i,k}[\mu_i(\tilde{b}_{ik}) - \lambda_i(\tilde{\boldsymbol{p}}_k)]^2$. By the assumption on the test prices, $c_0^{(1)}$ is a positive constant, and we obtain $\mathbb{P}\left(E_1^C\right) \leq c_0^{(1)} M/N_0$ as in (E.3).

E.2 Proof of (23)

Here we prove the bound (23) on the revenue loss per period. As discussed in Section 5, we separate the revenue loss depending on whether E_2 occurs, and we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^M \left[r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t)) \cdot \mathbb{1}\{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} > 0,\ \forall j\}\right]\right] \\
&\leq\ \mathbb{E}\left[\sum_{i=1}^M \left[r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t)) \cdot \mathbb{1}\{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} > 0,\ \forall j\}\right] \middle| E_2\right] \cdot \mathbb{P}(E_2) + p_{\max}\mathbb{P}(E_2^C) \\
&\leq\ \mathbb{E}\left[\sum_{i=1}^M \left[r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t))\right] \cdot \mathbb{1}\{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} > 0,\ \forall j\} \middle| E_2\right] \\
&\qquad + \mathbb{E}\left[\sum_{i=1}^M r_i(\hat{\boldsymbol{\omega}}^*) \cdot \mathbb{1}\{\exists j : I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0\} \middle| E_2\right] \cdot \mathbb{P}(E_2) + p_{\max}\mathbb{P}(E_2^C),
\end{aligned} \tag{E.6}$$

where the first inequality follows from

$$\sum_{i=1}^M \left[r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t)) \cdot \mathbb{1}\{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} > 0,\ \forall j\}\right] \leq\ \sum_{i=1}^M r_i(\hat{\boldsymbol{\omega}}^*) \leq\ p_{\max} \sum_{i=1}^M \omega_i^* \leq\ p_{\max}.$$

The first term in the rightmost expression of (E.6) comes from implementation of the generalized SGD method. Under the MNL model, we will show below that

$$\sum_{i=1}^M \left[r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t))\right] \leq\ \kappa_r \|\boldsymbol{p}_t - \bar{\boldsymbol{p}}\|_2, \tag{E.7}$$

where κ_r is a finite constant. In this way, we can bound the first term by the price deviation from the generalized SGD method. The second term in (E.6) is incurred by stockouts. Using $\sum_{i=1}^M r_i(\hat{\boldsymbol{\omega}}^*) \leq p_{\max}$ (as shown above), we obtain

$$\mathbb{E}\left[\sum_{i=1}^M r_i(\hat{\boldsymbol{\omega}}^*) \cdot \mathbb{1}\{\exists j : I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0\} \middle| E_2\right] \cdot \mathbb{P}(E_2) \leq\ p_{\max}\mathbb{P}\left(\bigcup_j \{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0\} \middle| E_2\right) \cdot \mathbb{P}(E_2).$$

Applying (E.7) and the above bound, we can derive from (E.6) the bound (23) on the revenue loss:

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^M \left[r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t)) \cdot \mathbb{1}\{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} > 0,\ \forall j\}\right]\right] \\
&\leq \kappa_r \mathbb{E}\left[\|\boldsymbol{p}_t - \bar{\boldsymbol{p}}\|_2 \cdot \mathbb{1}\{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} > 0,\ \forall j\} \middle| E_2\right] + p_{\max}\mathbb{P}\left(\bigcup_j \{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0\} \middle| E_2\right) \cdot \mathbb{P}(E_2) + p_{\max}\mathbb{P}\left(E_2^C\right),
\end{aligned}$$

where we bound the first term in (E.6) by (E.7).

Therefore, it remains to show (E.7), that is, $\sum_{i=1}^M [r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t))] \leq \kappa_r \|\boldsymbol{p}_t - \bar{\boldsymbol{p}}\|_2$ for some finite constant κ_r. By the mean value theorem, we have

$$\begin{aligned}
\sum_{i=1}^M [r_i(\hat{\boldsymbol{\omega}}^*) - r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t))] =&\ \sum_{i=1}^M \int_0^1 \nabla r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t + u(\bar{\boldsymbol{p}} - \boldsymbol{p}_t)))^\top (\bar{\boldsymbol{p}} - \boldsymbol{p}_t)\, du \\
\leq&\ \sum_{i=1}^M \| \int_0^1 \nabla r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t + u(\bar{\boldsymbol{p}} - \boldsymbol{p}_t))) du \|_2 \cdot \|\boldsymbol{p}_t - \bar{\boldsymbol{p}}\|_2.
\end{aligned}$$

where $u \in [0,1]$, and the last inequality follows from the Cauchy–Schwarz inequality. Let $\boldsymbol{p}^u := \boldsymbol{p}_t + u(\bar{\boldsymbol{p}} - \boldsymbol{p}_t)$ and $\lambda_i^u := \lambda_i(\boldsymbol{p}^u)$. We can derive that

$$
\begin{aligned}
&\| \int_0^1 \nabla r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t + u(\bar{\boldsymbol{p}} - \boldsymbol{p}_t)))du\|_2^2 \\
&= \left(\int_0^1 [\lambda_i^u - p_i^u \beta_i \lambda_i^u \sum_{j\neq i} \lambda_j^u] \; du \right)^2 + \sum_{j \neq i} \left(\int_0^1 p_j^u \beta_j \lambda_i^u \lambda_j^u \; du \right)^2 \\
&\leq \; \left(\int_0^1 \lambda_i^u \cdot \max\{1, p_{\max}\beta_{\max}\} \; du \right)^2 + (p_{\max}\beta_{\max})^2 \left(\int_0^1 \lambda_i^u \sum_{j\neq i} \lambda_j^u \; du \right)^2
\end{aligned}
$$

where the first inequality follows from $\lambda_i^u - p_i^u \beta_i \lambda_i^u \sum_{j\neq i} \lambda_j^u \leq \lambda_i^u \cdot \max\{1, p_i^u \beta_i \sum_{j \neq i} \lambda_j^u\}$, $p_i^u \beta_i \leq p_{\max}\beta_{\max}$, and $\sum_{j\neq i} \left(\int_0^1 p_j^u \beta_j \lambda_i^u \lambda_j^u \; du \right)^2 \leq \left(\sum_{j \neq i} \int_0^1 p_j^u \beta_j \lambda_i^u \lambda_j^u \; du \right)^2$. Furthermore, the right-most expression above is bounded above by $2(\max\{1, p_{\max}\beta_{\max}\})^2 \left(\int_0^1 \lambda_i^u \; du \right)^2$. Thus,

$$
\| \int_0^1 \nabla r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t + u(\bar{\boldsymbol{p}} - \boldsymbol{p}_t)))du\|_2^2 \; \leq \; 2(\max\{1, p_{\max}\beta_{\max}\})^2 \left(\int_0^1 \lambda_i^u \; du \right)^2 .
$$

By taking a square root and summing across all products, it follows that

$$
\sum_{i=1}^M \| \int_0^1 \nabla r_i(\boldsymbol{\lambda}(\boldsymbol{p}_t + u(\bar{\boldsymbol{p}} - \boldsymbol{p}_t)))du\|_2 \leq \sqrt{2} \max\{1, p_{\max}\beta_{\max}\} \cdot \int_0^1 \sum_{i=1}^M \lambda_i^u \; du \leq \sqrt{2} \max\{1, p_{\max}\beta_{\max}\},
$$

where the last inequality follows from the definition that $\lambda_i^u = \lambda_i(\boldsymbol{p}^u)$ and $\sum_{i=1}^M \lambda_i(\boldsymbol{p}^u) \leq 1$ under the MNL demand. Denote by $\kappa_r := \sqrt{2}\max\{1, p_{\max}\beta_{\max}\}$. We obtain (E.7), which completes the proof of (23).

E.3 Proof of (24)

Now we prove the probability of a stockout is bounded as in (24) with the choice of $N_0 = \lceil \sqrt{T} (\log T)^2 \rceil$.

We will bound the the probability of a stockout with and without a high-probability event E_3 that we will define later. We have

$$
\begin{aligned}
\mathbb{P}\left(\bigcup_j \{ I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0 \} \cap E_2 \right) &\leq \; \mathbb{P}\left(\bigcup_j \{ I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0 \} \middle| E_2 \cap E_3 \right) \cdot \mathbb{P}(E_3 \cap E_2) + \mathbb{P}\left(E_3^C \cap E_2 \right) \\
&\leq \; \sum_{j=1}^M \mathbb{P}\left(I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0 \middle| E_2 \cap E_3 \right) \cdot \mathbb{P}(E_3 \cap E_2) + \mathbb{P}\left(E_3^C \cap E_2 \right). \qquad \text{(E.8)}
\end{aligned}
$$

We will bound the first term by constructing a martingale. Recall that $I_{i,t}^{\mathcal{A}} + S_{i,t}^{\mathcal{A}} = I_{i,0} + \sum_{k=1}^t S_{i,k}^{\mathcal{A}} - \sum_{k=1}^{t-1} D_{i,k}^{\mathcal{A}}$, where $S_{i,k}^{\mathcal{A}}$ and $D_{i,k}^{\mathcal{A}}$ are the supply and demand under the selected purchase price $b_{i,k}$ and selling price vector $\boldsymbol{p}_k$, respectively. Let $\tilde{\mathcal{F}}_k$ denote the σ-algebra generated by the history up to the supply observation in period k. Then, all $D_{i,k-1}^{\mathcal{A}}$ and $S_{i,k}^{\mathcal{A}}$ are $\tilde{\mathcal{F}}_k$-measurable with $\mathbb{E}[D_{i,k-1}^{\mathcal{A}} | \tilde{\mathcal{F}}_{k-1}] = \lambda_i(\boldsymbol{p}_{k-1})$ and $\mathbb{E}[S_{i,k}^{\mathcal{A}} | \tilde{\mathcal{F}}_{k-1}] = \mu_i(b_{i,k})$ except for $S_{i,2N_0+1}^{\mathcal{A}}$ (as we set the initial prices of Phase 2 according to all prices and observations in Phase 1). Define

$$X_k := \begin{cases} S^{\mathcal{A}}_{i,1} - \mu_i(b_{i,1}), & k = 1, \\ -[D^{\mathcal{A}}_{i,2N_0} - \lambda_i(\boldsymbol{p}_{2N_0})], & k = 2N_0+1, \\ [S^{\mathcal{A}}_{i,k} - \mu_i(b_{i,k})] - [D^{\mathcal{A}}_{i,k-1} - \lambda_i(\boldsymbol{p}_{k-1})], & \forall\ k \in \mathbb{Z}_+ \setminus \{1, 2N_0+1\}. \end{cases}$$

Observe that $\mathbb{E}[X_k|\tilde{\mathcal{F}}_{k-1}] = 0$, and $\{X_k\}$ is a martingale difference sequence with $|X_k| \leq 2$. Moreover, we have for any $t > 2N_0$ that

$$\begin{aligned} \sum_{k=1}^{t} X_k =&\ \sum_{k=1}^{2N_0}[S^{\mathcal{A}}_{i,k} - \mu_i(b_{i,k})] + \sum_{k=2N_0+2}^{t}[S^{\mathcal{A}}_{i,k} - \mu_i(b_{i,k})] - \sum_{k=1}^{t-1}[D^{\mathcal{A}}_{i,k} - \lambda_i(\boldsymbol{p}_k)] \\ \leq&\ \sum_{k=1}^{t} S^{\mathcal{A}}_{i,k} - \sum_{k=1}^{t-1} D^{\mathcal{A}}_{i,k} - \left[\sum_{k=1}^{2N_0}\mu_i(b_{i,k}) + \sum_{k=2N_0+2}^{t}\mu_i(b_{i,k}) - \sum_{k=1}^{t-1}\lambda_i(\boldsymbol{p}_k)\right] \\ \leq&\ (I^{\mathcal{A}}_{i,t} + S^{\mathcal{A}}_{i,t}) - \left[\sum_{k=1}^{2N_0}[\mu_i(b_{i,k}) - \lambda_i(\boldsymbol{p}_k)] + \sum_{k=2N_0+1}^{t-1}[\mu_i(b_{i,k+1}) - \lambda_i(\boldsymbol{p}_k)]\right]. \end{aligned} \tag{E.9}$$

We claim that under the choice of $N_0 = \lceil\sqrt{T}(\log T)^2\rceil$ and a high probability event E_3, we have

$$\sum_{k=1}^{2N_0}[\mu_i(b_{i,k}) - \lambda_i(\boldsymbol{p}_k)] + \sum_{k=2N_0+1}^{t-1}[\mu_i(b_{i,k+1}) - \lambda_i(\boldsymbol{p}_k)] \geq\ N_0 s_0, \quad \forall\ T \geq T_0, \tag{E.10}$$

where s_0 and T_0 are finite constants. Suppose this claim holds. Plugging it into (E.9), we obtain

$$\mathbb{P}\left(I^{\mathcal{A}}_{i,t} + S^{\mathcal{A}}_{i,t} = 0\middle| E_2 \cap E_3\right)\cdot\mathbb{P}(E_3 \cap E_2) \ \leq\ \mathbb{P}\left(\sum_{k=1}^{t} X_k \leq -N_0 s_0\middle| E_2 \cap E_3\right)\cdot\mathbb{P}(E_3 \cap E_2) \ \leq\ \mathbb{P}\left(\sum_{k=1}^{t} X_k \leq -N_0 s_0\right).$$

We can apply Azuma's inequality to the martingale difference sequence $\{X_k\}$, which implies

$$\mathbb{P}\left(\sum_{k=1}^{t} X_k \leq -N_0 s_0\right) \leq\ \exp\left(-\frac{(N_0 s_0)^2}{2\sum_{k=1}^{t} 2^2}\right) =\ \exp\left(-\frac{(N_0 s_0)^2}{8t}\right).$$

Since $N_0 \geq \sqrt{T}(\log T)^2$, we have $(N_0 s_0)^2/(8t) \geq (\log T)^4 s_0^2/8$, which is greater than $\log T$ when T is large (i.e., $\log T \geq 2s_0^{-2/3}$). Let $T_0 \geq \exp(2s_0^{-2/3})$, and we have $\mathbb{P}\left(I^{\mathcal{A}}_{i,t} + S^{\mathcal{A}}_{i,t} = 0\middle| E_2 \cap E_3\right)\cdot\mathbb{P}(E_3 \cap E_2) \leq 1/T$ for any i and all $T \geq T_0$. This provides a bound for the first term in (E.8) when the claim in (E.10) holds, that is, $\sum_{j=1}^{M}\mathbb{P}\left(I^{\mathcal{A}}_{j,t} + S^{\mathcal{A}}_{j,t} = 0\middle| E_2 \cap E_3\right)\cdot\mathbb{P}(E_3 \cap E_2) \leq M/T$.

Next, we define the event E_3 and prove our claim in (E.10). Recall that we set the sell prices as $\boldsymbol{p}_k = \boldsymbol{p}_\infty$ so that $\lambda_i(\boldsymbol{p}_k) = 0$ for all i when there is a stockout for some product. For the first $2N_0$ periods, we set the test prices and have $\mu_i(b_{i,k}) - \lambda_i(\boldsymbol{p}_k) \geq \mu_i(\tilde{b}_{i1}) - \lambda_i(\tilde{\boldsymbol{p}}_1)$ when $k \leq N_0$ and $\mu_i(b_{i,k}) - \lambda_i(\boldsymbol{p}_k) \geq \mu_i(\tilde{b}_{i2}) - \lambda_i(\tilde{\boldsymbol{p}}_2)$ when $N_0 < k \leq 2N_0$. Let $s_0 := \min_{i,\ell}\{\mu_i(\tilde{b}_{i\ell}) - \lambda_i(\tilde{\boldsymbol{p}}_\ell)\}$. Then, $s_0 > 0$ by the assumption on the test prices, and we obtain

$$\sum_{k=1}^{2N_0}[\mu_i(b_{i,k}) - \lambda_i(\boldsymbol{p}_k)] \geq\ 2N_0 s_0.$$

For $k > 2N_0$, we follow the generalized SGD method, and we have

$$\begin{aligned}
\sum_{k=2N_0+1}^{t-1} [\mu_i(b_{i,k+1}) - \lambda_i(\boldsymbol{p}_k)] &= \sum_{k=2N_0+2}^{t} \mu_i(b_{i,k}) - \sum_{\tau=1}^{\tau_{t-1}} \lambda_i(\check{\boldsymbol{p}}_\tau) \\
&\geq \sum_{k=2N_0+2}^{t} \left[\mu_i(\bar{b}_i) - |\mu_i(b_{i,k}) - \mu_i(\bar{b}_i)|\right] - \sum_{\tau=1}^{\tau_{t-1}} [\lambda_i(\bar{\boldsymbol{p}}) + |\lambda_i(\check{\boldsymbol{p}}_\tau) - \lambda_i(\bar{\boldsymbol{p}})|] \\
&\geq -\sum_{k=2N_0+2}^{t} |\mu_i(b_{i,k}) - \mu_i(\bar{b}_i)| - \sum_{\tau=1}^{\tau_{t-1}} |\lambda_i(\check{\boldsymbol{p}}_\tau) - \lambda_i(\bar{\boldsymbol{p}})|,
\end{aligned}$$

where the last inequality follows from $\mu_i(\bar{b}_i) = \hat{\omega}_i^* = \lambda_i(\bar{\boldsymbol{p}})$ and $\tau_{t-1} \leq t - 2N_0 - 1$. Note that $|\mu_i(b_{i,k}) - \mu_i(\bar{b}_i)| = \gamma_{2,i}|b_{i,k} - \bar{b}_i| \leq \gamma_{2,\max}|b_{i,k} - \bar{b}_i|$, and

$$|\lambda_i(\check{\boldsymbol{p}}_\tau) - \lambda_i(\bar{\boldsymbol{p}})| = |\nabla\lambda_i(\tilde{\boldsymbol{p}})^\top(\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}})| \leq \|\nabla\lambda_i(\tilde{\boldsymbol{p}})\|_2 \cdot \|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2$$

with $\tilde{\boldsymbol{p}} := u\check{\boldsymbol{p}}_\tau + (1-u)\bar{\boldsymbol{p}}$ for some constant $u \in (0,1)$. We have

$$\begin{aligned}
\|\nabla\lambda_i(\tilde{\boldsymbol{p}})\|_2 &= \sqrt{\left(\beta_i\lambda_i(\tilde{\boldsymbol{p}})\sum_{j\neq i}\lambda_j(\tilde{\boldsymbol{p}})\right)^2 + \sum_{j\neq i}\left(\beta_j\lambda_i(\tilde{\boldsymbol{p}})\lambda_j(\tilde{\boldsymbol{p}})\right)^2} \\
&\leq \sqrt{2}\beta_{\max}\lambda_i(\tilde{\boldsymbol{p}})\sum_{j\neq i}\lambda_j(\tilde{\boldsymbol{p}}) \leq \sqrt{2}\beta_{\max}\lambda_i(\tilde{\boldsymbol{p}})(1-\lambda_i(\tilde{\boldsymbol{p}})) \leq \beta_{\max}/2\sqrt{2},
\end{aligned}$$

where the last inequality follows from $\lambda_i(\tilde{\boldsymbol{p}})(1-\lambda_i(\tilde{\boldsymbol{p}})) \leq \max_{\lambda\in(0,1)} \lambda(1-\lambda) = 1/4$. Thus, we obtain

$$\begin{aligned}
\sum_{k=2N_0+1}^{t-1} [\mu_i(b_{i,k+1}) - \lambda_i(\boldsymbol{p}_k)] \geq &\ -\gamma_{2,\max}\sum_{k=2N_0+2}^{t} |b_{i,k} - \bar{b}_i| - \frac{\beta_{\max}}{2\sqrt{2}}\sum_{\tau=1}^{\tau_{t-1}}\|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2 \\
\geq &\ -\gamma_{2,\max}\sum_{k=2N_0+1}^{T} |b_{i,k} - \bar{b}_i| - \frac{\beta_{\max}}{2\sqrt{2}}\sum_{\tau=1}^{\tau_T}\|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2.
\end{aligned}$$

Define $E_3 := \bigcap_i\{\sum_{t=2N_0+1}^{T}|b_{i,t} - \bar{b}_i| < \sqrt{K_bT}\cdot\log(eT)\}\bigcap\{\sum_{\tau=1}^{\tau_T}\|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2 < \sqrt{K_pT}\cdot\log(eT)\}$, under which total deviation of our updated prices from the target on each side is $O(\sqrt{T}\log T)$. Then, the right hand side above is bounded under the event E_3, and we obtain

$$\sum_{k=1}^{2N_0}[\mu_i(b_{i,k+1}) - \lambda_i(\boldsymbol{p}_k)] + \sum_{k=2N_0+1}^{t-1}[\mu_i(b_{i,k+1}) - \lambda_i(\boldsymbol{p}_k)] \geq \ 2N_0s_0 - C_d\sqrt{T}\log(eT),$$

where $C_d := \gamma_{2,\max}\sqrt{K_b} + \frac{\beta_{\max}}{2\sqrt{2}}\sqrt{K_p} > 0$ is a finite constant. Since we choose $N_0 = \lceil\sqrt{T}(\log T)^2\rceil$, we have $C_d\sqrt{T}\log(eT) < N_0s_0$ when T is large (e.g., $\log T \geq 2C_d/s_0$). Thus, we complete the proof of our claim (E.10) holds by defining $T_0 \geq \exp(2C_d/s_0)$. This claim was used earlier to establish a bound for the first term in (E.8).

Now we address the second term in (E.8), which is $\mathbb{P}\left(E_3^C \cap E_2\right)$. By the high-probability bounds in (9) and (15), we derive from the definition of E_3 that

$$
\begin{aligned}
\mathbb{P}\left(E_3^C \cap E_2\right) \leq\ & \sum_{i=1}^{M} \mathbb{P}\left(\sum_{t=2N_0+1}^{T} |b_{i,t} - \bar{b}_i| \geq \sqrt{K_b T} \cdot \log(eT) \middle| E_2\right) \cdot \mathbb{P}(E_2) \\
& + \mathbb{P}\left(\sum_{\tau=1}^{\tau_T} \|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2 \geq \sqrt{K_p T} \cdot \log(eT) \middle| E_2\right) \cdot \mathbb{P}(E_2) \\
\leq\ & \sum_{i=1}^{M} \mathbb{P}\left(\sum_{t=2N_0+1}^{T} |b_{i,t} - \bar{b}_i| \geq \sqrt{K_b (T-2N_0)} \cdot \log(eT)\right) \\
& + \mathbb{P}\left(\sum_{\tau=1}^{\tau_T} \|\check{\boldsymbol{p}}_\tau - \bar{\boldsymbol{p}}\|_2 \geq \sqrt{K_p \tau_T} \cdot \log(eT) \middle| E_2\right) \\
\leq\ & (M+1)T^{-1}.
\end{aligned}
$$

Summarizing the bounds on the two terms in (E.8), we set $T_0 = \exp(2\max\{s_0^{-2/3}, C_d/s_0\})$ and obtain

$$
\mathbb{P}\Big(\bigcup_j \{I_{j,t}^{\mathcal{A}} + S_{j,t}^{\mathcal{A}} = 0\} \cap E_2\Big) \leq (2M+1)T^{-1}, \quad \forall\, T \geq T_0.
$$

This completes the proof of (24).

E.4 Proof of (25)

We now prove (25), which stats that E_2 is a high-probability event. To bound $\mathbb{P}\left(E_2^C\right)$, we separate it depending on whether price tests are sufficient in the first Phase, that is, $\mathbb{P}\left(E_2^C\right) \leq \mathbb{P}\left(E_2^C \middle| E_1\right) + \mathbb{P}\left(E_1^C\right)$. By definition, $E_2 = \bigcap_{i=1}^{M} \{|\hat{\beta}_i - \beta_i| \leq \varepsilon_0\}$, and we derive that

$$
\mathbb{P}\left(E_2^C \middle| E_1\right) \leq\ \sum_{i=1}^{M} \mathbb{P}\left(|\hat{\beta}_i - \beta_i| > \varepsilon_0 \middle| E_1\right) \leq\ \sum_{i=1}^{M} \frac{1}{\varepsilon_0^2} \mathbb{E}\left[\|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_2^2 \middle| E_1\right] \leq\ \frac{M}{\varepsilon_0^2} \cdot \frac{c_0^d}{N_0},
$$

where the second inequality follows from Markov's inequality and $|\hat{\beta}_i - \beta_i|^2 \leq \|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i\|_2^2$, and the last inequality follows from (E.2). We also have $\mathbb{P}\left(E_1^C\right) \leq c_0^{(1)} M/N_0$ by (E.3). This implies

$$
\mathbb{P}\left(E_2^C\right) \leq\ \left(\frac{c_0^d}{\varepsilon_0^2} + c_0^{(1)}\right)\frac{M}{N_0},
$$

Thus, (25) holds with $c_\epsilon := c_0^d/\varepsilon_0^2 + c_0^{(1)}$, which is a finite constant.

References

Borokov, A. (1998). *Mathematical statistics.* Gordon and Breach Science Publishers, Amsterdam.

Harvey, N. J., Liaw, C., Plan, Y., and Randhawa, S. (2019). Tight analyses for non-smooth stochastic gradient descent. In *Conference on Learning Theory*, pages 1579–1613. PMLR.

Hazan, E., Agarwal, A., and Kale, S. (2007). Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192.

Xiao, Y. and Zhou, S. X. (2020). Trade-in for cash or for upgrade? Dynamic pricing with customer choice. *Production and Operations Management*, 29(4):856–881.