E-companion to "Utility Fairness in Contextual Dynamic Pricing with Demand Learning

Authors: Xi Chen, David Simchi-Levi and Yining Wang.

EC.1. Omitted proofs of technical lemmas in Sec. 4

EC.1.1. Proof of Lemma 1

First, from the definition of $\iota_k(\cdot)$ and the fact that $|j^*_{k+1} - j^*_k| \leq 1$ for all k, the defined $\varpi^*_\epsilon(\cdot)$ is δ_0-Lipschitz continuous. This immediately implies that π^*_ϵ satisfies δ_0-UF.

To prove the lower bound on the expected revenue of π^*_ϵ, we shall construct a sequence $j^\circ_1, \cdots, j^\circ_{M_n}$ feasible to Eq. (5) that is approximately as profitable as the non-discretized optimal policy π^*. Recall that, thanks to Eq. (4), the optimal π^* can be parameterized as $\pi^*(x) = \varpi^*(x^\top\theta_0)$. Let $p^*_1 = \varpi^*(u_1)$, $p^*_2 = \varpi^*(u_2)$, and so on. Let $j^\circ_1 = \arg\min_{j\in[M_p]} |p_j - p^*_1|$, with ties broken arbitrarily. For $k = 2, 3, \cdots, M_u$, determine the values of j°_k in the following ways:

1. If $p_{j^\circ_{k-1}} \leq p^*_{k-1} \leq p^*_k$ then set $j^\circ_k = j^\circ_{k-1} + 1$;
2. If $p_{j^\circ_{k-1}} \geq p^*_{k-1} \geq p^*_k$ then set $j^\circ_k = j^\circ_{k-1} - 1$;
3. In all other scenarios, set $j^\circ_k = j^\circ_{k-1}$.

Because $p_{j+1} = p_j + \delta_0\epsilon$ for every j and $|p^*_{k+1} - p^*_k| \leq \delta_0|u_{k+1} - u_k| \leq \delta_0\epsilon$ for every k, thanks to the fact that π^* satisfies δ_0-UF, it is easy to verify via induction that that $\{j^\circ_k\}_{k=1}^{M_u}$ sequence constructed above satisfies $|p_{j^\circ_k} - p^*_k| \leq \delta_0\epsilon$. Construct a policy $\pi^\circ_\epsilon(x) = \varpi^\circ_\epsilon(x^\top\theta_0)$ with

$$\varpi^\circ_\epsilon(u) = p_{j^\circ_{k(u)}} + \delta_0\iota_{k(u)}(u - u_{k(x)}; \{j^\circ_k\}) \quad \text{where } k(u) = \arg\min_{k\in[M_u]} \left|u_k - u\right|, \tag{EC.1}$$

and the $\iota_k(\cdot; \{j_k\}_k)$ function being defined in Eq. (7). This definition together with the properties that $|p_{j^\circ_k} - p^*_k| \leq \delta_0\epsilon$ for all k and that π^* is δ_0-UF, yields

$$\sup_u \left|\varpi^\circ_\epsilon(u) - \varpi^*(u)\right| \leq 2\delta_0\epsilon. \tag{EC.2}$$

By Assumptions 1 and 2, the expected revenue function $r_u(\cdot)$ is $2L_f$-Lipschitz continuous. Subsequently,

$$\begin{aligned}\mathbb{E}_{x\sim P_\mathcal{X}}\left[r_{x^\top\theta_0}(\pi^*(x)) - r_{x^\top\theta_0}(\pi^\circ_\epsilon(x))\right] &= \int_{-B}^{B} \left(r_u(\varpi^*(u)) - r_u(\varpi^\circ_\epsilon(u))\right)\mathrm{d}P_\mathcal{U}(u)\\ &\leq \int_{-B}^{B} 2L_f\left|\varpi^*(u) - \varpi^\circ_\epsilon(u)\right|\mathrm{d}P_\mathcal{U}(u) \leq 4L_f\delta_0\epsilon,\end{aligned}$$

which is to be proved. □

EC.1.2. Additional technical lemmas

We first define some notations. Let

$$\mathcal{F} := \left\{ g : [-B, B] \to [\underline{p}, \overline{p}] \middle| |g(u) - g(u')| \le \delta_0 |u - u'|, \forall u, u' \in [-B, B] \right\}$$

be the class of all functions mapping from $[-B, B]$ to $[\underline{p}, \overline{p}]$ that are δ_0-Lipschitz continuous. It is clear that for any utility-based contextual pricing policy $\pi(x) = \varpi(x^\top \theta_0)$, π satisfies δ_0-UF if and only if $\varpi \in \mathcal{F}$. For any $g, h \in \mathcal{F}$, define

$$\|g - h\|_\infty := \sup_{u \in [-B, B]} \big|g(u) - h(u)\big|.$$

We then have the following lemma:

Lemma EC.1. *Suppose the optimal δ_0-UF contextual pricing policy $\pi^*(x) = \varpi^*(x^\top \theta_0)$ is unique. Let $\{\varpi_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ be a sequence of functions in $\mathcal{F}$ such that*

$$\lim_{n \to \infty} \mathbb{E}_{P_\mathcal{U}}[r_u(\varpi_n(u))] = \mathbb{E}_{P_\mathcal{U}}[r_u(\varpi^*(u))].$$

Then it holds that

$$\lim_{n \to \infty} \|\varpi_n - \varpi^*\|_\infty = 0.$$

While seemingly intuitive, the proof of Lemma EC.1 is non-trivial and given later. The main idea behind the proof of Lemma EC.1 is the following: the set of all δ_0-UF contextual pricing policies can be completely characterized by the set of all δ_0-Lipschitz continuous functions on $[0, 1]$. By exclusing a small neighborhood around ϖ^* and using the Arzela-Ascoli theorem, the remaining function set is compact with respect to the point-wise convergence norm $\|\cdot\|_\infty$. On the other hand, the expected revenue of any function is continuous in $\|\cdot\|_\infty$. This means that the supreme of the maximal expected revenue can be attained on a compact space which is different from the objective of ϖ^* thanks to the uniqueness of ϖ^*.

Finally, our next lemma shows that the expected revenue of the optimal policy is a left-continuous function of the fairness parameter δ_0.

Lemma EC.2. *For $\delta_0 > 0$, let $R(\delta_0)$ be the expected revenue of optimal δ_0-UF policy. Then for every $\delta_0 > 0$, $\lim_{\delta \to \delta_0^-} R(\delta) = R(\delta_0)$.*

The proof of Lemma EC.2 is given later in this section. The proof is at a higher level similar to the proof of Lemma EC.1, utilizing pointwise convergence of a sequence of pricing policies and the compactness of δ_0-UF policies.

EC.1.3. Proof of Lemma EC.1

For notational simplicity define

$$\varphi(\varpi) := \mathbb{E}_{P_{\mathcal{U}}}[r_u(\varpi(u)]$$

for every $\varpi \in \mathcal{F}$. Because $r_u(\cdot)$ is $2L_f$-Lipschitz continuous for all u, it is easy to verify that, for every pair of $\varpi, \widetilde{\varpi} \in \mathcal{F}$,

$$\left|\varphi(\varpi) - \varphi(\widetilde{\varpi})\right| \le \int_{-B}^{B} 2L_f|\varpi(u) - \widetilde{\varpi}(u)|\mathrm{d}P_{\mathcal{U}}(u) \le 2L_f\|\varpi - \widetilde{\varpi}\|_\infty,$$

meaning that $\varphi(\cdot)$ is a $2L_f$-Lipschitz continuous function in $\|\cdot\|_\infty$.

Fix arbitrary $\epsilon > 0$ and consider

$$\mathcal{F}_\epsilon := \{g \in \mathcal{F} : \|g - \varpi^*\|_\infty \ge \epsilon\}.$$

It is easy to verify that $\mathcal{F}_\epsilon$ is closed and bounded in $\|\cdot\|_\infty$. Furthermore, since $\mathcal{F}_\epsilon \subseteq \mathcal{F}$ and $\mathcal{F}$ only contains δ_0-Lipschitz continuous functions, any function sequence in $\mathcal{F}_\epsilon$ is equi-continuous. By the Arzela-Ascoli theorem, this means that $\mathcal{F}_\epsilon$ is compact with respect to $\|\cdot\|_\infty$. Since φ itself is Lipschitz continuous, this means that the image $\varphi(\mathcal{F}_\epsilon) = \{\varphi(\varpi) : \varpi \in \mathcal{F}_\epsilon\}$ is compact and therefore

$$\beta := \varphi(\varpi^*) - \sup_{\varpi\in\mathcal{F}_\epsilon} \varphi(\varpi) = \varphi(\varpi^*) - \max_{\varpi\in\mathcal{F}_\epsilon} \varphi(\varpi) > 0,$$

because the optimal policy π^* is unique and $\pi^* \notin \mathcal{F}_\epsilon$. Because $\lim_{n\to\infty} \varphi(\varpi_n) = \varphi(\varpi^*)$, there exists $N \in \mathbb{N}$ such that for all $n \ge N$, $\varphi(\varpi^*) - \varphi(\varpi_n) \le \beta/2$. This implies that $\varpi_n \notin \mathcal{F}_\epsilon$ for all $n \ge N$, meaning that $\|\varpi_n - \varpi^*\|_\infty \le \epsilon$ for all $n \ge N$. The lemma is thus proved because $\epsilon > 0$ is arbitrary. $\square$

EC.1.4. Proof of Lemma EC.2

First, note that $R(\cdot)$ is a monotonically decreasing function because a policy that satisfies δ-UF also satisfies δ'-UF for any $\delta' > \delta$. Let $\pi^*(x) = \varpi^*(x^\top\theta_0)$ be the δ_0-UF policy that maximizes its expected revenue, which exists because the expected revenue is continuous in $\|\cdot\|_\infty$, and the set of δ_0-UF policies coincide with the set of δ_0-Lipschitz continuous functions on $[-B, B]$, which is compact in $\|\cdot\|_\infty$ thanks to the Arzela-Ascoli theorem. For $\epsilon > 0$ being arbitrarily small, consider the policy $\widetilde{\varpi}(u) = \max\{\underline{p}, (1 - \epsilon/\delta_0)\varpi^*(u)\}$. It is easy to verify that $\widetilde{\varpi}(\cdot)$ satisfies $(\delta_0 - \epsilon)$-UF. On the other hand, the expected revenue of $\widetilde{\varpi}$ and ϖ^* differs only by $O(\epsilon)$ because the expected revenue function $r_u(\cdot)$ is $2L_f$-Lipschitz continuous, for all u. Taking the limit of $\epsilon \to 0^+$ we have proved Lemma EC.2. $\square$

EC.1.5. Proof of Lemma 2

For every $p \in [\underline{p}, \overline{p}]$ and $u \in [-B, B]$, define $r(u,p) := pf(u - \alpha_0 p)$. Because $p^*(u), p^*(u')$ are interior minimizers, the first-order KKT condition asserts that

$$\partial_p r(u, p^*(u)) = \partial_p r(u', p^*(u')) = 0. \tag{EC.3}$$

Additionally, using the chain rule and Eq. (9), we have for every $u \in [-B, B]$ and $p \in [\underline{p}, \overline{p}]$ that

$$\partial_u \partial_p r(u,p) = \partial_u \left(f(u - \alpha_0 p) - \frac{\alpha_0 p}{2} f'(u - \alpha_0 p) \right) = f'(u - \alpha_0 p) - \frac{\alpha_0 p}{2} f''(u - \alpha_0 p) \geq \sigma_u. \tag{EC.4}$$

Incorporating Eq. (EC.4) into Eq. (EC.3) and using the mean-value theorem, there exists $\widetilde{u} \in [u, u']$ such that

$$\partial_p r(u', p^*(u)) = \partial_p r(u, p^*(u)) + \partial_u \partial_p r(\widetilde{u}, p^*(u))(u' - u) \geq \sigma_u (u' - u) > 0. \tag{EC.5}$$

Because $r(u', \cdot)$ is unimodal and peaks at $p^*(u')$, the fact that $\partial_p r(u', p^*(u)) > 0$ implies that $p^*(u) < p^*(u')$. Furthermore, because $r(u', \cdot)$ is twice continuously differentiable with its second-order derivatives being uniformly upper bounded by M_r, it holds that

$$\left|\partial_p r(u', p^*(u))\right| \leq M_r \left|p^*(u) - p^*(u')\right|. \tag{EC.6}$$

Combining Eqs. (EC.5,EC.7) we obtain a lower bound on $p^*(u') - p^*(u)$, which is the first inequality in Lemma 2.

To prove the second inequality, note that $r(u,p)$ is L_f-Lipschitz continuous in u for every p because $f(\cdot)$ is L_f-Lipschitz continuous thanks to Assumption 2. This means that

$$r(u', p^*(u')) \leq r(u, p^*(u')) + L_f(u' - u) \leq r(u, p^*(u)) + L_f(u' - u) \leq r(u', p^*(u)) + 2L_f(u' - u). \tag{EC.7}$$

On the other hand, the strong uni-modality property of r_u stated in Assumption 2 implies that

$$r(u', p^*(u)) \leq r(u', p^*(u')) - \frac{\sigma_r}{2}(p^*(u') - p^*(u))^2. \tag{EC.8}$$

Combining Eqs. (EC.7,EC.8) we obtain the second inequality in Lemma 2. □

EC.2. Proof of Theorem 1

To prove this theorem we rely again on a discretization idea similar to the one developed in the previous section, but for mathematical proof purposes instead of efficient computation. Let $\epsilon > 0$ be a small positive discretization error parameter and $\{u_k\}_{k=1}^{M_u}$ be the discretized grids of $[-B, B]$, so that $u_{k+1} = u_k + \epsilon$ and $[-B, B] = \bigcup_{k=1}^{M_u} [u_k - \epsilon/2, u_k + \epsilon/2]$. Consider the following finite-dimensional, continuous optimization problem:

$$q_1^*, \cdots, q_{M_u}^* = \arg \max_{q_1, \cdots, q_{M_u} \in [\underline{p}, \overline{p}]} \sum_{k=1}^{M_u} \gamma_k r_{u_k}(q_k) \tag{EC.9}$$

$$s.t. \quad \left| q_{k+1}^* - q_k^* \right| \le \delta_0 \epsilon, \;\; k = 1, 2, \cdots, M_u - 1. \tag{EC.10}$$

Given the optimal solution $q_1^*, \cdots, q_{M_u}^*$, a contextual policy $\pi_\epsilon^\dagger(x) = \varpi_\epsilon^\dagger(x^\top \theta_0)$ is constructed as a piecewise linear interpolation:

$$\varpi_\epsilon^\dagger(u) = q_k^* + \frac{u - u_k}{u_{k+1} - u_k} q_{k+1}^* \quad \text{ if } u_k \le u < u_{k+1},$$

with the understanding that $u_0 = -B$, $q_0^* = q_1^*$, $u_{M_u+1} = B$ and $q_{M_u=1}^* = q_{M_u}^*$. It is easy to verify that π_ϵ^* satisfies δ_0-UF. Furthermore, the optimal solution $j_1^*, \cdots, j_{M_u}^*$ feasible to Eq. (5) can be easily converted to a feasible solution $q_1 = p_{j_1^*}$, $q_2 = p_{j_2^*}$, etc. so that the resulting policies are exactly the same. Therefore, Lemma 1 impllies that

$$\mathbb{E}_{P_\mathcal{U}}[r_u(\varpi_\epsilon^\dagger(u))] \ge \mathbb{E}_{P_\mathcal{U}}[r_u(\varpi_\epsilon^*(u))] \ge \mathbb{E}_{P_\mathcal{U}}[r_u(\varpi^*(u))] - 4L_f \delta_0 \epsilon. \tag{EC.11}$$

We next use Lagrangian multipliers to analyze the properties of $q_1^*, \cdots, q_{M_u}^*$. For $k \in [M_u]$, let $\lambda_k^+ \ge 0$ be the Lagrangian multiplier associated with the constraint $q_{k+1} - q_k - \delta_0 \epsilon \le 0$ and $\lambda_k^- \ge 0$ be the Lagrangian multiplier associated with the constraint $-q_{k+1} + q_k - \delta_0 \epsilon \le 0$. Let $q^* = (q_1^*, \cdots, q_{M_u}^*)$ be the vectorized optimal price solution and $\lambda^+ = (\lambda_1^+, \cdots, \lambda_{M_u}^+)$, $\lambda^- = (\lambda_1^-, \cdots, \lambda_{M_u}^-)$ be the vectorized optimal Lagrangian multipliers. The Lagrangian function (when expressing the original problem as a minimization problem) can then be written as

$$\mathcal{L}(q^*, \lambda^+, \lambda^-) = \sum_{k=1}^{M_u} -\gamma_k r_{u_k}(q_k^*) + \lambda_k^+ (q_{k+1}^* - q_k^* - \delta_0 \epsilon) + \lambda_k^- (q_k^* - q_{k+1}^* - \delta_0 \epsilon).$$

For every $k \in [M_u]$, define

$$\Delta_k := \lambda_k^+ - \lambda_k^-.$$

Taking the partial derivative of $\mathcal{L}$ with respect to q_k^* for each $k \in [M_u]$ and using the first-order KKT condition [4], it holds that

$$\begin{aligned} \partial_{q_k} \mathcal{L}(q^*, \lambda^+, \lambda^-) &= -\gamma_k r'_{u_k}(q_k) - (\lambda_k^+ - \lambda_k^-) + \mathbf{1}\{k > 1\}(\lambda_{k-1}^+ - \lambda_{k-1}^-) \\ &= -\gamma_k r'_{u_k}(q_k^*) - \Delta_k + \mathbf{1}\{k > 1\} \Delta_{k-1} = 0. \end{aligned} \tag{EC.12}$$

[4] Strong duality holds here because all constraints are affine in the primal variables.

Subsequently, we have the recursion that

$$\Delta_k = \mathbf{1}\{k>1\}\Delta_{k-1} - \gamma_k r'_{u_k}(q_k^*), \quad k=1,2,\cdots,M_u. \tag{EC.13}$$

Additionally, by complementary slackness (the constraints associated with λ_k^+,λ_k^- cannot be binding simultaneously if $\delta_0\epsilon>0$, and therefore at least one of λ_k^+,λ_k^- must be zero) and the fact that $\lambda^+,\lambda^-\geq 0$, the following facts are true for every k:

$$\Delta_k>0 \Longrightarrow \lambda_k^+>0, \lambda_k^-=0 \Longrightarrow q_{k+1}^* = q_k^*+\delta_0\epsilon; \tag{EC.14}$$

$$\Delta_k<0 \Longrightarrow \lambda_k^->0, \lambda_k^+=0 \Longrightarrow q_{k+1}^* = q_k^*-\delta_0\epsilon; \tag{EC.15}$$

$$\Delta_k=0 \Longrightarrow \lambda_k^+=\lambda_k^-=0 \Longrightarrow q_k^*-\delta_0\epsilon < q_{k+1}^* < q_k^*+\delta_0\epsilon. \tag{EC.16}$$

We are now ready to establish several properties of the optimal primal and dual solution q^*,λ^+,λ^- that would eventually prove Theorem 1. Without loss of generality, we shall assume that $P_{\mathcal{U}}$ is supported on $[-B,B]$. If $P_{\mathcal{U}}$ is instead supported on a closed sub-interval of $[-B,B]$ (Assumption 3), then the rest of the proof remains valid by simply restricting the utility valuation and discretization grid to the support of $P_{\mathcal{U}}$.

Property 1: $\Delta_k\geq 0$ for all $k<M_u$. Proof: assume by way of contradiction that there exists $k<M_u$ such that $\Delta_k<0$. Let k be the smallest integer such that $\Delta_k<0$. Because $\Delta_{k-1}\geq 0$ (with the understanding that $\Delta_0=0$), Eq. (EC.13) implies that $r'_{u_k}(q_k^*)>0$, which means $q_k^*<p^*(u_k)$ thanks to the uni-modality of $r_{u_k}(\cdot)$. On the other hand, Eq. (EC.15) shows that $q_{k+1}^*=q_k^*-\delta_0\epsilon<q_k^*$, which combined with the fact that $p^*(u_{k+1})>p^*(u_k)$ (Lemma 2) shows $r'_{u_{k+1}}(q_{k+1}^*)>0$, which combined with $\Delta_k<0$ and Eq. (EC.13) yields $\Delta_{k+1}<0$. Continuing this argument until $k=M_u$, we have that $\Delta_\ell<0$ for all $\ell=k,k+1,\cdots,M_u$, and therefore $q_k^*>q_{k-1}^*>\cdots>q_{M_u}^*$. Note that, on the other hand, $q_k^*<p^*(u_k)<p^*(u_{k+1})<\cdots p^*(u_{M_n})$. This means that if we set $q_{k+1}^*,\cdots,q_{M_u}^*$ to be q_k^* we obtain a solution that has strictly larger objective while remaining feasible to Eq. (EC.11), which contradicts the optimality of $\{q_k^*\}_{k=1}^{M_u}$. ■

Property 2: if $\Delta_{k-1}>0$, $\Delta_k=0$ for some $k<M_u-1$ then $q_{k+1}^*\geq p^*(u_k)>q_k^*$. Proof: because $\Delta_{k-1}<0$ and $\Delta_k=0$, Eq. (EC.13) implies that $r'_{u_k}(q_k^*)<0$, meaning that $q_k^*<p^*(u_k)$ thanks to the uni-modality of r'_{u_k}. Assume by way of contradiction that $q_{k+1}^*<p^*(u_k)$. Because $p^*(u_{k+1})>p^*(u_k)$ thanks to Lemma 2, this assumption means that $r'_{u_{k+1}}(q_{k+1}^*)>0$ and therefore from Eq. (EC.13) one has that $\Delta_{k+1}=\Delta_k-\gamma_{k+1}r'_{u_{k+1}}(q_{k+1}^*)=-\gamma_{k+1}r'_{u_{k+1}}(q_{k+1}^*)<0$, contradicting property 1. ■

Properties 1 and 2 immediately imply that ϖ^* is a monotonically non-decreasing function in u, because in both cases of $\Delta_k>0$ and $\Delta_k=0$ the q_k^* prices will not drop, and by taking $\epsilon\to 0^+$ we obtain the monotonicity of the entire curve ϖ^*.

Property 3: if $\Delta_k = \Delta_{k-1} = 0$ then $q^*_k = p^*(u_k)$. Proof: it is immediate from Eq. (EC.13) that $r'_{u_k}(q^*_k) = 0$ and therefore $q^*_k = p^*(u_k)$. ■

Property 3 can be used to establish the second property for the general optimal ϖ^* solution. Because ϖ^* is δ_0-Lipschitz continuous (the utility fairness constraint), Radamacher's theorem shows that ϖ^* is differentiable almost everywhere on $[-B, B]$. Consider arbitrary $u \in (-B, B)$ such that ϖ^* is differentiable at u and $\partial_u \varpi^*(u) < \delta_0$. By definition of differentials, there exist $\zeta, \beta \in (0, \delta_0)$ such that for every $u' \in [u - \zeta, u + \zeta]$, $|\varpi^*(u') - \varpi^*(u)| \leq (\delta_0 - \beta)|u' - u|$. Because π^* is unique and $\lim_{\epsilon \to 0^+} \mathbb{E}_{P_\mathcal{U}}[r_u(\varpi^\dagger_\epsilon(u))] = \mathbb{E}_{P_\mathcal{U}}[r_u(\varpi^*(u))]$, Lemma EC.1 shows that $\lim_{\epsilon \to 0^+} \|\varpi^\dagger_\epsilon - \varpi^*\|_\infty = 0$, meaning that for any small $\eta \in (0, \zeta/5]$, there exists $\epsilon_0 > 0$ such that for any $\epsilon \leq \epsilon_0$ and $u \in [-B, B]$, $|\varpi^\dagger_\epsilon(u) - \varpi^*(u)| \leq \eta\beta$. For such a $\varpi^\dagger_\epsilon$, there must exist a discretized grid point $u_k \in [u - 4\eta, u + 4\eta]$ such that $\Delta_k = 0$: otherwise, with $\Delta_k > 0$ for all grid points in $[u - 4\eta, u + 4\eta]$ (the case of $\Delta_k < 0$ is ruled out in property 1), the slope of $\varpi^\dagger_\epsilon$ is fixed locally at δ_0 in $[u - 4\eta, u + 4\eta]$ and therefore $\varpi^\dagger_\epsilon(u - 4\eta) = \varpi^\dagger_\epsilon(u) - 4\delta_0\eta \leq \varpi^*_\epsilon(u) + \eta\beta - 4\delta_0\eta$, while on the other hand $\varpi^\dagger_\epsilon(u - 4\eta) \geq \varpi^*(u - 4\eta) - \eta\beta \geq \varpi^*(u) - 4(\delta_0 - \beta)\eta - \eta\beta = \varpi^*(u) + 3\eta\beta - 4\delta_0\eta$, which is a contradiction. The existence of $u_k \in [u - 4\eta, u + 4\eta]$ such that $\Delta_k = 0$ implies, by invoking Property 3, that $\varpi^\dagger_\epsilon(u_k + \epsilon) \geq p^*(u_k) > \varpi^\dagger_\epsilon(u_k)$. By continuity of $\varpi^\dagger_\epsilon$ and picking $\epsilon \leq \eta$, there exists $\widetilde{u} \in [u - 5\eta, u + 5\eta]$ such that $\varpi^\dagger_\epsilon(\widetilde{u}) = p^*(u_k)$. Now consider $\{\eta_n\}_{n \in \mathbb{N}}$ converging to zero and let $\{\epsilon_n\}_{n \in \mathbb{N}}$ also converging to zero such that $\|\varpi^\dagger_{\epsilon_n} - \varpi^*\|_\infty \leq \eta_n\beta$ for all n. The above argument shows that there exist two sequences $\{v_n, \widetilde{v}_n\}_{n \in \mathbb{N}}$ both converging to u such that $\varpi^\dagger_{\epsilon_n}(v_n) = p^*(\widetilde{v}_n)$, and $|\varpi^*(v_n) - p^*(\widetilde{v}_n)| \leq \epsilon_n$. Because ϖ^* is δ_0-Lipschitz continuous, $\lim_{n \to \infty} \varpi^*(v_n) = \varpi^*(\lim_{n \to \infty} v_n) = \varpi^*(u)$. By Lemma 2, $p^*(\cdot)$ is continuous and therefore $\lim_{n \to \infty} p^*(\widetilde{v}_n) = p^*(\lim_{n \to \infty} \widetilde{v}_n) = p^*(u)$. Noting also that $\lim_{n \to \infty} \epsilon_n = 0$, we have that $\lim_{n \to \infty} |\varpi^*(v_n) - p^*(\widetilde{v}_n)| = 0$, which implies $\varpi^*(u) = p^*(u)$. This proves the second property in Theorem 1 for general δ_0 values.

In the remainder of this proof, we focus on the case when $\delta_0 < \sigma_u/M_r$.

Property 4. if $\delta_0 < \sigma_u/M_r$ then $\Delta_k > 0$ for all $k < M_u - 1$. Proof: assume by way of contradiction that $\Delta_k = 0$ for some $k < M_u - 1$ (the case of $\Delta_k < 0$ has already been ruled out in property 1). Let k be the smallest integer such that this happens. Because $\Delta_{k-1} \geq 0$ (with the understanding that $\Delta_0 = 0$), Eq. (EC.13) implies that $r'_{u_k}(q^*_k) \geq 0$, which means that $q^*_k \leq p^*(u_k)$ because $r_{u_k}(\cdot)$ is uni-modal. Since q^* is primal feasible, this means that $q^*_{k+1} \leq q^*_k + \delta_0\epsilon \leq p^*(u_k) + \delta_0\epsilon$. On the other hand, Lemma 2 asserts that $p^*(u_{k+1}) \geq \frac{\sigma_r}{M_r}(u_{k+1} - u_k) = \frac{\sigma_r}{M_r}\epsilon$. Because $\delta_0 < \sigma_r/M_r$, it holds that $q^*_{k+1} < p^*(u_{k+1})$ and therefore $r'_{u_{k+1}}(q^*_{k+1}) > 0$. Applying Eq. (EC.13) again, we obtain $\Delta_{k+1} = \Delta_k - \gamma_{u_{k+1}} r'_{u_{k+1}}(q^*_{k+1}) = -\gamma_{u_{k+1}} r'_{u_{k+1}}(q^*_{k+1}) < 0$, which contradicts property 1. ■

Property 4 shows that with the additional condition of $\delta_0 < \sigma_u/M_r$, we must have $\Delta_k > 0$ for all $k < M_u - 1$, implying that ϖ^* must have the slope fixed to δ_0 throughout. Finally, this structure can be extended to $\delta_0 = \sigma_u/M_r$ by considering the sequence of optimal policies with the desirable linear structure for $\{\delta_{0n}\}_{n\in\mathbb{N}}$, $\lim_{n\to\infty}\delta_{0n}=\sigma_u/M_r$. By Lemma EC.2, the expected revenues of these optimal policies converge to the expected revenue of the optimal policy for $\delta_0=\sigma_u/M_r$, which then implies uniform convergence of the policies themselves when the optimal policy is unique, thanks to Lemma EC.1.

EC.3. Proof of Theorem 2

In the purely linear demand case the optimal price (without fairness constraints) of a given baseline utility value u is $p^*(u)=u/\alpha_0$, which has a slope of $1/\alpha_0$. Therefore, if $\delta_0\geq 1/\alpha_0$, $R(\delta_0)$ would be exactly equal to $R(+\infty)$ by applying policy p^* and therefore $\rho(\delta_0)=1$ for all $\delta_0\geq 1/\alpha_0$.

We next focus on the case of $\delta_0<1/\alpha_0$. Theorem 1 shows that in this case, the optimal δ_0-UF policy ϖ^* takes the form of

$$\varpi^*(u)=\pi_0+\delta_0 u,$$

with the only parameter that is free to vary being $\pi_0=\varpi^*(0)$. Let

$$\phi(\pi_0):=\mathbb{E}_{P_{\mathcal{U}}}[r_u(\pi_0+\delta_0 u)],$$

where $r_u(p)=p(u-\alpha_0 p/2)$ and $r_u'(p)=u-\alpha_0 p/2+p(-\alpha_0/2)=u-\alpha_0 p$. Taking derivative of ϕ with respect to π_0, we have that

$$\phi'(\pi_0)=\mathbb{E}_{P_{\mathcal{U}}}[r_u'(\pi_0+\delta_0 u)]=\mathbb{E}[u-\alpha_0(\pi_0+\delta_0 u)]=-\alpha_0\pi_0+(1-\alpha_0\delta_0)\mu_u. \tag{EC.17}$$

Equating $\phi'(\pi_0)=0$ we have that $\pi_0=\alpha_0^{-1}(1-\alpha_0\delta_0)\mu_u$. Consequently,

$$\begin{aligned}
R(\delta_0)&=\mathbb{E}_{P_{\mathcal{U}}}[r_u(\alpha_0^{-1}(1-\alpha_0\delta_0)\mu_u+\delta_0 u)]\\
&=\mathbb{E}_{P_{\mathcal{U}}}\left[-\frac{\alpha_0}{2}\left(\frac{1-\alpha_0\delta_0}{\alpha_0}\mu_u+\delta_0 u\right)^2+u\left(\frac{1-\alpha_0\delta_0}{\alpha_0}\mu_u+\delta_0 u\right)\right]\\
&=-\frac{(1-\alpha_0\delta_0)^2}{2\alpha_0}\mu_u^2-\delta_0(1-\alpha_0\delta_0)\mu_u^2-\frac{\alpha_0\delta_0^2}{2}\nu_u^2+\frac{1-\alpha_0\delta_0}{\alpha_0}\mu_u^2+\delta_0\nu_u^2\\
&=\frac{(1-\alpha_0\delta_0)^2}{2\alpha_0}\mu_u^2+\delta_0\left(1-\frac{\alpha_0\delta_0}{2}\right)\nu_u^2.
\end{aligned} \tag{EC.18}$$

On the other hand, when $\delta_0=1/\alpha_0$, the first term in Eq. (EC.18) is zero and therefore

$$R(1/\alpha_0)=\frac{1}{2\alpha_0}\nu_u^2. \tag{EC.19}$$

Note that for all $\delta_0\geq 1/\alpha_0$, the optimal expected revenue is the same because the utility fairness constraint is mute here, or more specifically $R(1/\alpha_0)=R(+\infty)$. Taking the ratio betwen Eqs. (EC.18,EC.19) we complete the proof of Theorem 2.

EC.4. Proof of Theorem 3

Fix $\delta_0 \in (0,1/2]$. Set $d=1$, $\mathcal{X}=[1/4,3/4]$, $[\underline{p},\overline{p}]=[0,3\delta_0/4]$ and let $P_{\mathcal{X}}$ be the uniform distribution on $\mathcal{X}$. Let $\Delta_\theta,\Delta_\alpha \in [0,\delta_0^2/10]$ be positive parameters to be specified later. Consider two hypothesis H_0,H_1 that differ only in θ and α:

$$\begin{aligned} H_0: &\quad \theta=1,\ \alpha=\delta_0^{-1};\\ H_1: &\quad \theta=1-\Delta_\theta,\ \alpha=\delta_0^{-1}-\Delta_\alpha. \end{aligned}$$

The demand model is $\mathbb{E}[y|x,p]=x\theta-0.5\alpha p$, with $y\in\{0,1\}$ being Bernoulli random variables. The expected revenue is $r_x(p)=p(x\theta-0.5\alpha p)$. It is easy to verify that, under both problem instances, $\mathbb{E}[y|x,p]\in[0,1]$ for all $x\in\mathcal{X}$, $p\in[\underline{p},\overline{p}]$. Furthermore, the condition $\delta_0\le\alpha_0^{-1}$ is true for both problem instances.

EC.4.1. Optimal policy and deviation analysis

In this section we derive the optimal policy under problem instances parameterized by $\Delta_\theta,\Delta_\alpha$, and also analyze the property of sub-optimal policies that deviate from the optimal policy. Our first lemma gives the properties of the optimal policy.

LEMMA EC.3. *Let $\mathcal{F}_{\theta,\alpha}(\delta_0)$ be the set of all δ_0-UF policies under problem instance θ,α. Let $\pi^*\in\mathcal{F}_{\theta,\alpha}(\delta_0)$ be the policy that maximizes the expected revenue under problem instance θ,α, and suppose that $\delta_0\le 1/\alpha$. For every $x\in\mathcal{X}$, let $p^*(x)=\arg\max_{p\in[\underline{p},\overline{p}]} r_x(p)$. Then for every $x\in\mathcal{X}$,*

1. $p^*(x)=\frac{\theta}{\alpha}$;
2. $\pi^*(x)=\frac{\theta(1-\alpha\delta_0)}{2\alpha}+\delta_0 x\theta$;
3. $\mathbb{E}[r_x(\pi^*(x))]=\frac{\theta^2}{8}(\frac{1}{\alpha}+\frac{\delta_0}{6}+\frac{\delta_0^2}{12})$.

Proof of Lemma EC.3. For the first property, taking $r_x'(p)=-\alpha p+x\theta=0$ we obtain $p^*(x)=x\theta/\alpha$. For the second property, note that π^* must take the form of $\pi^*(x)=\pi_0^*+\delta_0 x\theta$ for some $\pi_0^*\in\mathbb{R}$, thanks to Theorem 1. Taking the derivative of the expected revenue with respect to π_0 and equating it to zero, we obtain

$$\partial_\pi \mathbb{E}[r_x(\pi_0^*+\delta_0 x\theta)]=\mathbb{E}[r_x'(\pi_0^*+\delta_0 x\theta)]=\mathbb{E}[-\alpha(\pi_0^*+\delta_0 x\theta)+x\theta]=0, \tag{EC.20}$$

which implies, together with the fact that $\mathbb{E}[x]=1/2$,

$$\pi_0^*=\frac{(1-\alpha\delta_0)\mathbb{E}[x\theta]}{\alpha}=\frac{(1-\alpha\delta_0)\theta}{2\alpha},$$

proving the second property. For the third property, simply plugging the form of π^* into the expected revenue evaluation we obtain

$$\begin{aligned}\mathbb{E}[r_x(\pi^*(x))] &= \mathbb{E}[(\pi_0^* + \delta_0 x\theta)(x\theta - 0.5\alpha(\pi_0^* + \delta_0 x\theta))] \\ &= -0.5\alpha(\pi_0^*)^2 + (1-\delta_0\alpha)\pi_0\mathbb{E}[x\theta] + \delta_0(1+0.5\alpha\delta_0)\mathbb{E}[x^2\theta^2] \\ &= -\frac{\alpha(\pi_0^*)^2}{2} + \frac{(1-\delta_0\alpha)\pi_0^*\theta}{2} + \frac{13}{48}\delta_0(1+0.5\alpha\delta_0)\theta^2 \\ &= \frac{\theta^2}{8}\left(\frac{1}{\alpha} + \frac{\delta_0}{6} + \frac{\delta_0^2}{12}\right),\end{aligned}$$

which is to be proved. □

To simplify our notations, we shall use $\mathcal{F}_0(\delta_0)$ and $\mathcal{F}_1(\delta_0)$ to denote the class of δ_0-UF policies under problem instances H_0 and H_1, respectively. We also use $\mathbb{E}_0$ and $\mathbb{E}_1$ for the laws under H_0 and H_1. For $x \in \mathcal{X}$, $b \in \{0,1\}$ and $p \in [\underline{p}, \overline{p}]$, write

$$r_{bx}(p) := r_{x^\top\theta}(p) = p(x^\top\theta - 0.5\alpha p), \quad \theta, \alpha \text{ associated with } H_b.$$

Our next lemma shows that, any policy that is δ_0-UF with respect to H_1 is going to incur a gap of $\Omega(\Delta_\theta)$ under H_0, compared with the optimal policy that is δ_0-UF with respect to H_0.

Lemma EC.4. *For any $\pi \in \mathcal{F}_1(\delta_0)$, it holds that*

$$\max_{\pi' \in \mathcal{F}_0(\delta_0)} \mathbb{E}_0[r_{0x}(\pi'(x))] - \mathbb{E}_0[r_{0x}(\pi(x))] \geq \Delta_\theta/32.$$

Proof of Lemma EC.4. Because the constructed problem instances are one-dimensional and $\theta = 1-\Delta_\theta$ in H_1 for some $\Delta_\theta > 0$, it is easy to verify that $\mathcal{F}_1(\delta_0) \subseteq \mathcal{F}_0(\delta_0 - \Delta_\theta)$. Consequently, the third property of Lemma EC.3 implies that, the left-hand side of the inequality in Lemma EC.4 is lower bounded by

$$\begin{aligned}&\frac{\theta^2}{8}\left(\frac{1}{\alpha} + \frac{\delta_0}{6} + \frac{\delta_0^2}{12}\right) - \frac{\theta^2}{8}\left(\frac{1}{\alpha} + \frac{\delta_0 - \Delta_\theta}{6} + \frac{(\delta_0-\Delta_\theta)^2}{12}\right) \\ &\quad \geq \frac{\theta^2}{8}\left(\frac{\Delta_\theta}{3} - \frac{\Delta_\theta^2}{12}\right) \geq \frac{\Delta_\theta}{32},\end{aligned}$$

which is to be proved. □

Our next lemma shows that, under problem instance H_0, if a δ_0-UF contextual pricing policy $\pi \in \mathcal{F}_0(\delta_0)$ deviates significantly from the optimal δ_0-UF policy π^*, then it must also have a significantly smaller expected revenue.

Lemma EC.5. *Let $\pi^* = \arg\max_{\pi\in\mathcal{F}_0(\delta_0)} \mathbb{E}_0[r_{0x}(\pi(x))]$ and π be arbitrary. Then*

$$\mathbb{E}_0[r_{0x}(\pi^*(x))] - \mathbb{E}_0[r_{0x}(\pi(x))] = \int_{1/4}^{3/4} \left|\pi(x) - \pi^*(x)\right|^2 \mathrm{d}x.$$

Proof of Lemma EC.5. Immediate from the fact that, under H_0, $\pi^*(x) = p^*(x)$ and $r_{0x}(p^*(x)) - r_{0x}(p) = (p - p^*(x))/2$ for all $x \in \mathcal{X}$. □

EC.4.2. Bounding KL-Divergence

For $b \in \{0, 1\}$, let $\pi_b^* \in \arg\min_{\pi\in\mathcal{F}_b(\delta_0)} \mathbb{E}[r_{bx}(\pi(x))]$ be the optimal δ_0-UF pricing policy under problem instance H_b. Also define the following events:

$$\text{Event } \mathcal{E}_b: \quad \pi_1, \cdots, \pi_T \in \mathcal{F}_b(\delta_0);$$
$$\text{Event } \mathcal{W}_b: \quad \sum_{t=1}^{T} \int_{1/4}^{3/4} r_{bx}(\pi_b^*(x)) - r_{bx}(\pi_t(x)) \le 2\underline{C} \times T^{2/3}.$$

Intuitively, event $\mathcal{E}_b$ captures the good event that implemented policies are δ_0-UF under problem instance H_b, and event $\mathcal{W}_b$ captures the good event that the cumulative regret of implemented policies over T time periods are small under problem instance H_b.

Our next lemma shows that, when Δ_θ is not too small, the events defined cannot co-exist, meaning that a sequence of policies $\pi_1, \cdots, \pi_T$ cannot be δ_0-UF with respect to H_1 while still achieving small regret under H_0 simultaneously.

Lemma EC.6. *Suppose $\Delta_\theta > 128\underline{C} \times T^{-2/3}$. Then $\mathcal{E}_1 \cap \mathcal{W}_0 = \emptyset$.*

Proof of Lemma EC.6. By Lemma EC.4, $\mathcal{E}_1$ implies that, under H_0,

$$\sum_{t=1}^{T} \int_{1/4}^{3/4} [r_{0x}(\pi_0^*(x)) - r_{0x}(\pi_t(x))]\mathrm{d}x \ge \frac{\Delta_\theta T}{64}. \tag{EC.21}$$

On the other hand, the definition of $\mathcal{W}_0$ implies that, under H_0,

$$\sum_{t=1}^{T} \int_{1/4}^{3/4} [r_{0x}(\pi_0^*(x)) - r_{0x}(\pi_t(x))]\mathrm{d}x \le 2\underline{C} \times T^{2/3}. \tag{EC.22}$$

With $\Delta_\theta > 128\underline{C} \times T^{-2/3}$, Eqs. (EC.21,EC.22) are contradictory to each other and therefore the lemma is proved. □

The following lemma upper bounds the Kullback-Leibler (KL) divergence between the observables under H_0 and H_1, using the fact that $\mathcal{A}$ is near-optimal under H_0.

LEMMA EC.7. *Suppose $\Delta_\alpha = 2\Delta_\theta/\delta_0$. For any algorithm $\mathcal{A}$ that satisfies*

$$\mathbb{E}_0^{\mathcal{A}}\left[\sum_{t=1}^{T} r_{0x}(\pi_0^*(x)) - r_{0x}(\pi_t(x))\right] \leq \underline{C} \times T^{2/3}, \tag{EC.23}$$

it holds that

$$\mathrm{KL}(P_0^{\mathcal{A}} \| P_1^{\mathcal{A}}) \leq \underline{C} \times \frac{44\Delta_\theta^2 T^{2/3}}{\delta_0^2},$$

where $P_b^{\mathcal{A}}$ is the law of $\{(x_t, p_t, y_t)\}_{t=1}^{T}$ under problem instance H_b and algorithm $\mathcal{A}$, for $b \in \{0,1\}$.

Proof of Lemma EC.7. By Lemma EC.5, Eq. (EC.23) implies that

$$\sum_{t=1}^{T} \int_{1/4}^{3/4} \left|\pi_t(x) - \pi_0^*(x)\right|^2 \mathrm{d}x \leq 0.5\underline{C} \times T^{2/3}. \tag{EC.24}$$

Note that π_0^* takes the closed-form of $\pi_0^*(x) = \delta_0 x$. With $\Delta_\alpha = 2\Delta_\theta/\delta_0$, it holds that

$$\mathbb{E}_0[y|x, \pi_0^*(x)] = \mathbb{E}_1[y|x, \pi_0^*(x)], \quad \forall x \in \mathcal{X}, \tag{EC.25}$$

because $x - 0.5\delta_0^{-1}\pi_0^*(x) = x(1-\Delta_\theta) - 0.5(\delta_0^{-1} - \Delta_\alpha)\pi_0^*(x)$ for all x. Note also that $\mathbb{E}_0[y|x, \pi_0(x^*)] \in [1/8, 3/8]$ for all x. Subsequently, by (Chen & Wang 2018a, Lemma 3),

$$\begin{aligned}\mathrm{KL}(P_0(\cdot|x,p) \| P_1(\cdot|x,p)) &\leq \Delta_\alpha^2 (p - \pi_0^*(x))^2 \left(8 + \frac{8}{3}\right) \leq 11\Delta_\alpha^2 (p - \pi_0^*(x)) \\ &= \frac{44\Delta_\theta^2}{\delta_0^2} \times (p - \pi_0^*(x))^2, \quad \forall x \in \mathcal{X},\end{aligned} \tag{EC.26}$$

where $P_b(\cdot|x,p)$ is the law of y conditioned on x, p under problem instance H_b, $b \in \{0,1\}$. Combining Eqs. (EC.24,EC.26) and using the Markovian structure of the observations across T time periods, we have that

$$\begin{aligned}\mathrm{KL}(P_0^{\mathcal{A}} \| P_1^{\mathcal{A}}) &= \mathbb{E}_0^{\mathcal{A}}\left[\sum_{t=1}^{T} 2\int_{1/4}^{3/4} \mathrm{KL}(P_0(\cdot|x, \pi_t(x)) \| P_1(\cdot|x, \pi_t(x)))\mathrm{d}x\right] \\ &\leq \frac{44\Delta_\theta^2}{\delta_0^2} \mathbb{E}_0^{\mathcal{A}}\left[\sum_{t=1}^{T} \int_{1/4}^{3/4} (\pi_t(x) - \pi_0^*(x))^2 \mathrm{d}x\right] \\ &\leq \underline{C} \times \frac{44\Delta_\theta^2 T^{2/3}}{\delta_0^2}.\end{aligned} \tag{EC.27}$$

This proves Lemma EC.7. □

EC.4.3. Completing the proof of Theorem 3

Instantiate $\Delta_\theta = 130\underline{C} \times T^{-2/3}$ and $\Delta_\alpha = 2\Delta_\theta/\delta_0^2$. For sufficiently large T, such $\Delta_\theta, \Delta_\alpha$ instantiated will satisfy all range conditions at the beginning of this proof. Consider an algorithm $\mathcal{A}$ such that

$\min_{b\in\{0,1\}} \Pr_b^{\mathcal{A}}[\mathcal{E}_b] \geq 0.95$. Assume by way of contradiction that

$$\mathbb{E}_0^{\mathcal{A}}\left[\sum_{t=1}^{T} r_{0x}(\pi_0^*(x)) - r_{0x}(\pi_t(x))\right] \leq \underline{C} \times T^{2/3}. \tag{EC.28}$$

If Eq. (EC.28) does not hold, then the conclusion of Theorem 3 automatically holds and the proof is done.

By Markov inequality, Eq. (EC.28) implies that $\Pr_0^{\mathcal{A}}[\mathcal{W}_0^c] \leq 1/2$, and therefore $\Pr_0^{\mathcal{A}}[\mathcal{W}_0] \geq 1/2$. Because $\mathcal{W}_0 \cap \mathcal{E}_1 = \emptyset$ thanks to Lemma EC.6, $\Pr_0^{\mathcal{A}}[\mathcal{W}_0] \geq 1/2$ implies that $\Pr_0^{\mathcal{A}}[\mathcal{E}_1] \leq 1/2$. Now apply Lemma EC.7 and Pinsker's inequality; we obtain

$$\begin{aligned} Pr_1^{\mathcal{A}}[\mathcal{E}_1] &\leq Pr_0^{\mathcal{A}}[\mathcal{E}_1] + \|P_0^{\mathcal{A}} - P_1^{\mathcal{A}}\|_{\mathrm{TV}} \leq \frac{1}{2} + \sqrt{\frac{\mathrm{KL}(P_0^{\mathcal{A}}\|P_1^{\mathcal{A}})}{2}} \\ &\leq \frac{1}{2} + \sqrt{\underline{C} \times \frac{22\Delta_\theta^2 T^{2/3}}{\delta_0^2}} = \frac{1}{2} + \sqrt{\frac{371800\underline{C}^3}{\delta_0^2}}. \end{aligned} \tag{EC.29}$$

Set

$$\underline{C} = \sqrt[3]{\frac{\delta_0^2}{371800 \times 5}} = \delta_0^{2/3}/123,$$

a strictly positive constant that only depends on δ_0. The right-hand side of Eq. (EC.29) is then upper bounded by $\frac{1}{2} + \frac{1}{\sqrt{5}} = 0.947 < 0.95$, which contradicts $\Pr_1^{\mathcal{A}}[\mathcal{E}_1] \geq 0.95$. This completes the proof of Theorem 3.

EC.5. Proof of Theorem 4

In this section we prove both fairness and regret claims in Theorem 4. Throughout this proof we assume that all assumptions 1-5 are valid, and will not explicitly invoke or state them when deriving results.

EC.5.1. Analysis of $\widehat{\theta}$ and $\widehat{\alpha}$

We first establish a technical lemma shows that $\widehat{\theta}$ and $\widehat{\alpha}$ estimate θ_0 and α_0 in ℓ_2 norm, with the estimation errors on the order of $\widetilde{O}(1/\sqrt{T_0}) = \widetilde{O}(T^{-1/3})$.

Lemma EC.8. *For T being sufficiently large, with probability $1 - \widetilde{O}(T^{-2})$ it holds that*

$$\|\widehat{\theta} - \theta_0\|_2 \leq \frac{4M_r}{\sigma_1}\sqrt{\frac{\ln(dT)}{QT_0}}.$$

Proof of Lemma EC.8. Let $p_L = \underline{p}$ and $p_U = \overline{p}$ be the two exploration prices. Let $z = (x, -0.5p)$ be the extended context vector, and $\beta_0 = (\theta_0, \alpha_0)$ be the extended model. Let $P_{\mathcal{Z}}$ be the distribution of z over the first T_0 time periods. To simplify notations, for every $\beta \in \mathbb{R}^d$, define

$$\varphi(\beta) := \sum_{t=1}^{T_0} \sum_{\ell=1}^{Q} \ln \mathcal{L}(y_{t\ell} | z_{t\ell}, \beta).$$

Because $\widehat{\beta}$ is an unconstrained maximizer of φ, it holds that $\nabla\varphi(\widehat{\beta}) = 0$. Subsequently,

$$-\langle \widehat{\beta} - \beta_0, \nabla\varphi(\beta_0)\rangle = -\int_0^1 (\widehat{\beta} - \beta_0)^\top \nabla^2\varphi(\beta_0 + s(\widehat{\beta} - \beta_0))(\widehat{\beta} - \beta_0)\mathrm{d}s \tag{EC.30}$$

$$\geq -\int_0^{\min\{1, \rho_L/\|\widehat{\beta} - \beta_0\|_2\}} (\widehat{\beta} - \beta_0)^\top \nabla^2\varphi(\beta_0 + s(\widehat{\beta} - \beta_0))(\widehat{\beta} - \beta_0)\mathrm{d}s \tag{EC.31}$$

$$\geq \min\{1, \rho_L/\|\widehat{\beta} - \beta_0\|_2\} \times \sigma_L(\widehat{\beta} - \beta_0)^\top \Lambda(\widehat{\beta} - \beta_0), \tag{EC.32}$$

where

$$\Lambda := \sum_{t=1}^{T_0} \sum_{\ell=1}^{Q} z_{t\ell} z_{t\ell}^\top. \tag{EC.33}$$

Here, Eq. (EC.30) is a consequence of the fundamental theorem of calculus and the fact that $\nabla\varphi(\widehat{\beta}) = 0$; Eq. (EC.31) holds because φ is concave thanks to the second property of Assumption 5, and therefore $\nabla^2\varphi(\cdot)$ is negative semi-definite; Eq. (EC.32) holds by the third property of Assumption 5.

Define $\Lambda_0 := \mathbb{E}_{P_{\mathcal{Z}}}[zz^\top]$. We then have that $\Lambda_0 \succeq \mathrm{Cov}(P_{\mathcal{Z}}) \succeq \min\{\sigma_x, 0.25(\overline{p} - \underline{p})^2\} I_{d+1} =: \sigma_0 I_{d+1}$. On the other hand, because z is bounded almost surely, using the matrix Hoeffding's inequality we have with probability $1 - \widetilde{O}(T^{-2})$ that $\Lambda = \sum_{t=1}^{T} \sum_{\ell=1}^{Q} z_{t\ell} z_{t\ell}^\top \succeq 0.5\sigma_0 T_0 Q I_{d+1}$, provided that T is sufficiently large. This combined together with Eqs. (EC.32,EC.33) yields that

$$-\langle \widehat{\beta} - \beta_0, \nabla\varphi(\beta_0)\rangle \geq \sigma_1 \times \min\left\{1, \frac{\rho_L}{\|\widehat{\beta} - \beta_0\|_2}\right\} \times T_0 Q \|\widehat{\beta} - \beta_0\|_2^2, \tag{EC.34}$$

where $\sigma_1 := 0.5\sigma_0\sigma_r > 0$.

On the other hand, by Hoeffding's inequality and the fact that $\mathbb{E}[\nabla\varphi(\beta_0)] = 0$ (property 1 of Assumption 5), it holds with probability $1 - \widetilde{O}(T^{-2})$ that

$$\|\nabla\varphi(\beta_0)\|_2 \leq 4M_r\sqrt{T_0 Q \ln(dT)}.$$

Consequently,

$$\left|\langle \widehat{\beta} - \beta_0, \nabla\varphi(\beta_0)\rangle\right| \leq 4M_r\sqrt{T_0 Q \ln(dT)} \times \|\widehat{\beta} - \beta_0\|_2. \tag{EC.35}$$

Combine Eqs. (EC.34,EC.35). We obtain

$$\min\{\rho_L, \|\widehat{\beta}-\beta_0\|_2\} \le \frac{4M_r}{\sigma_1}\sqrt{\frac{\ln(dT)}{QT_0}}. \tag{EC.36}$$

With $\rho_L > 0$ being a constant and T_0 being sufficiently large, the left-hand side of the above inequality reduces to $\|\widehat{\beta}-\beta_0\|_2$. This then proves Lemma EC.8. □

With Lemma EC.8 and the value of κ_1, we have with high probability that $\|\widehat{\theta}-\theta_0\|_2 \le \kappa_1/\sqrt{T_0}$. This immediately proves the fairness claim in Theorem 4, because for every $x, x' \in \mathcal{X}$,

$$\begin{aligned}\widetilde{\delta}_0\big|(x-x')^\top\widehat{\theta}\big| &\le \widetilde{\delta}_0(|(x-x')^\top\theta_0| + |(x-x')^\top(\widehat{\theta}-\theta_0)|) \le (\widetilde{\delta}_0 + \mathrm{diam}(\mathcal{X})\|\widehat{\theta}-\theta_0\|_2)|(x-x')^\top\theta_0| \\ &\le (\widetilde{\delta}_0 + \kappa_1/\sqrt{T_0})|(x-x')^\top\theta_0| \le \delta_0|(x-x')^\top\theta_0|.\end{aligned}$$

EC.5.2. Analysis of the MAB-UCB procedure

Given $\widehat{\theta}$, the estimate of the linear model θ_0 from the price experimentation phase, the expected revenue of a policy with a "starting price" π_0 can be written as

$$\Phi(\pi_0) := \mathbb{E}[r_x(\pi_0 + \widetilde{\delta}_0 x^\top\widehat{\theta})],$$

where $r_x(p) = pf(x^\top\theta_0 - \alpha_0 p)$. Recall also the definition that $R(\delta_0)$ is the expected revenue of the optimal δ_0-UF policy, with full information of θ_0 and α_0. The following lemma describes several properties of $\Phi(\cdot)$.

Lemma EC.9. *Let $k^* = \arg\max_{k\in[K]}\Phi(\pi_k)$. Then $\Phi(\pi_{k^*}) \ge R(\delta_0) - 6L_f\widetilde{B}\kappa_1/\sqrt{T_0} - 2L_f(\overline{p} - \underline{p})/K$.*

Proof of Lemma EC.9. Let π^* be the optimal δ_0-UF policy that maximizes the expected revenue, which admits the form of $\pi^*(x) = \pi_0^* + \delta_0 x^\top\theta_0$ thanks to Theorem 1. Let π_k be the discretized price such that $|\pi_k - \pi_0^*| \le (\overline{p}-\underline{p})/K$. Because $r_x(\cdot)$ is $2L_f$-Lipschitz continuous, we have that

$$\begin{aligned}\Phi(\pi_{k^*}) \ge \Phi(\pi_k) &= \mathbb{E}[r_x(\mathrm{Trim}_{[\underline{p},\overline{p}]}(\pi_k + \widetilde{\delta}_0 x^\top\widehat{\theta}))] \\ &\ge R(\delta_0) - 2L_f\big(\big|\pi_k - \pi_0^*\big| + \big|\widetilde{\delta}_0 x^\top\widehat{\theta} - \delta_0 x^\top\theta_0\big|\big) \\ &\ge R(\delta_0) - 2\delta_0 L_f\mathrm{diam}(\mathcal{X}) \times \|\widehat{\theta}-\theta_0\|_2 - 4L_f(\delta_0 - \widetilde{\delta}_0)\times\widetilde{B} - 2L_f(\overline{p}-\underline{p})/K \\ &\ge R(\delta_0) - 6L_f\widetilde{B}\kappa_1/\sqrt{T_0} - 2L_f(\overline{p}-\underline{p})/K,\end{aligned} \tag{EC.37}$$

which proves Lemma EC.9. □

EC.5.3. Regret analysis

The regret cumulated in the first price experimentation phase is upper bounded by $QT_0 = QT^{2/3}$.

For the second MAB phase, note that with the parameter κ_2 chosen as $\kappa_2 = 4\sqrt{\ln T}$, the Hoeffding's inequality and a union bound over all $K = \lceil T^{1/3}\rceil$ arms and $T - T_0$ time periods imply that, with probability $1 - \widetilde{O}(T^{-2})$,

$$\left|\frac{r_k}{n_k} - \Phi(\pi_k)\right| \le \frac{\kappa_2}{\sqrt{n_k}}, \quad \forall k \in [K]. \tag{EC.38}$$

Subsequently, using the standard upper-confidence bound analysis, it holds that

$$\sum_{t=T_0+1}^{T} \Phi(\pi_{k^*}) - \Phi(\pi_{k_t}) \le 2\kappa_2\sqrt{KQT}, \tag{EC.39}$$

where $k^* = \arg\max_{k\in[K]} \Phi(\pi_k)$. Eq. (EC.39) together with Lemma EC.9 shows that the cumulative regret of the second phase is upper bounded by

$$\begin{aligned} 2\kappa_2\sqrt{KQT} + QT \times (6L_f\widetilde{B}\kappa_1/\sqrt{T_0} + 2L_f(\overline{p} - \underline{p})/K) \\ &= 2\kappa_2\sqrt{Q}T^{2/3} + 6L_f\widetilde{B}\kappa_1 QT^{2/3} + 2L_f(\overline{p} - \underline{p})QT^{2/3} \\ &\le (6L_f\widetilde{B}\kappa_1 + 4L_f\kappa_2)QT^{2/3}, \end{aligned} \tag{EC.40}$$

which proves the regret upper bound.

EC.6. Experimental settings of real online shopping data

We use the online shopping data from https://www.kaggle.com/datasets/jacksondivakarr/online-shopping-dataset/data which consists of 52955 data entries (rows) and 42 features (columns). After removing rows with missing entries, the data set has 52924 rows each corresponding to a user purchase activity.

We first estimate a Logistic demand models as follows. We focus on one particular product *Google Men's 100% Cotton Short Sleeve Hero Tee Black*, which is relatively popular and has 595 associated purchase activities in the data set. Because of the lack of no-purchase activities in the data set, we simply set the price elasticity parameter as $\alpha_0 = 2.0$. We identified 11 covariates that would be informative to model users' heterogeneous purchase behaviors, as summarized below:

- The *gender* of the user, described by a binary variable;
- The *geographic location* of the user, described by three dummy variables indicating whether the user is from New York/New Jersey, Chicago, or California;
- The *tenure* of a user in the system, measured by months and divided by 10 for normalization purposes;
- The *offline spend* and *online spend* of a users, measured in dollars and divided by \$1000 for normalization purposes;

- The *season* at which a transaction occurs, described by three dummy variables indicating whether the transaction occurred during the first, the second or the third quarters of a year;
- A binary variable indicating whether a *discount coupon* has been applied to the transaction.

Each one of the 595 customers who have purchased the focal T-shirt product is associated with a 11-dimensional contextual vector described above. When then fit a linear regression model with independent variables being the 11 covariates and an additional intercept parameter, and the dependent variable being the price at which the customers purchased the focal product. The 12-dimensional linear model $\theta_0 \in \mathbb{R}^{12}$ (including the intercept parameter) is then used together with the synthetic price elasticity parameter $\alpha_0 = 2$ to form a Logistic demand model $\mathbb{E}[y|x,p] = f(x^\top \theta_0 - p)$ where $f(u) = e^u/(1+e^u)$. With the demand model, the contextual vectors of all the 52924 users are then used (subject to a random permutation to ensure identically and independently distributed contexts) to carry out the bandit algorithm and performance analysis.