Optimal Contextual Pricing under Agnostic Non-Lipschitz Demand

Jianyu Xu[1,2] **Yu-Xiang Wang**[3]

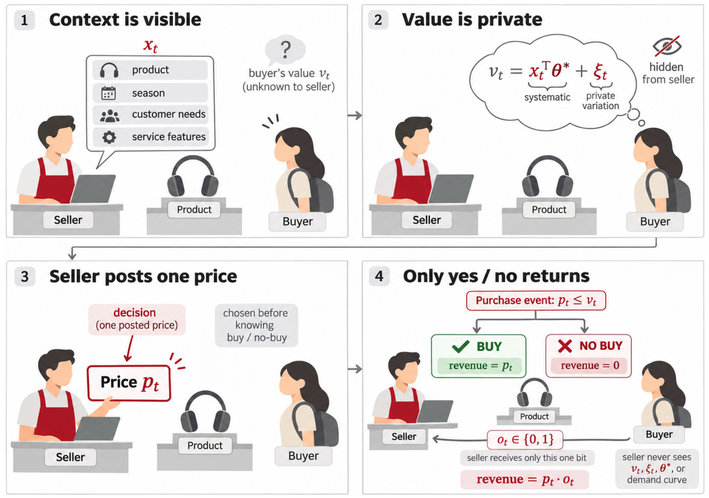
[1]UNC Charlotte [2]Carnegie Mellon University [3]UC San Diego

Contextual Pricing: The Game

A seller sees the context and posts **one price**, but never sees the buyer's private value. The feedback is a **single purchase bit**.

At each round $t = 1, \dots, T$**:**

1. Observe features $x_t \in \mathbb{R}^d$.
2. Buyer has **private value** $v_t = x_t^\top \theta^\star + \xi_t$.
3. Seller posts price $p_t \in [0, B]$.
4. Observe only $o_t = \mathbf{1}\{p_t \le v_t\}$.
5. Receive revenue $p_t o_t$.



What Is the Learning Goal?

Learn the linear value parameter $\theta^\star$ and unknown noise distribution F from **binary feedback**.

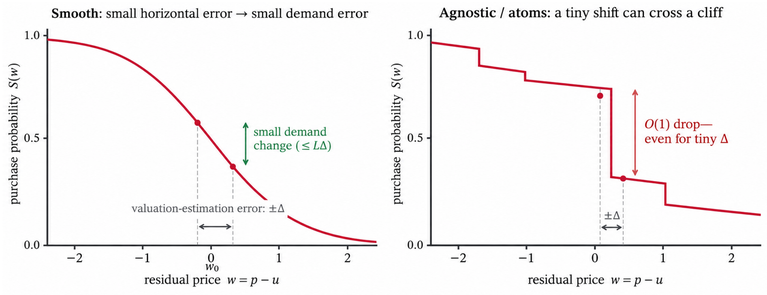
$$R_T = \sum_{t=1}^{T} [\pi_t(p_t^\star) - \pi_t(p_t)]$$

Regret is the revenue gap against the clairvoyant price $p_t^\star$, which knows $\theta^\star$ and F.

Goal: learn while earning.

Why Jumps Change Everything

Valuation error moves residual price along the demand curve: $w = p - u_t$ shifts on the horizontal axis.

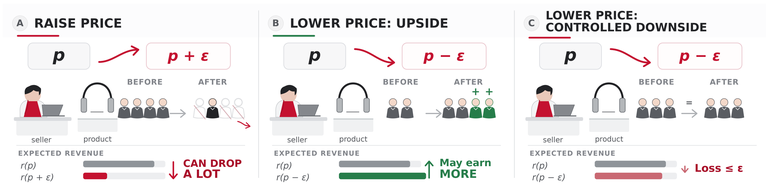


Smooth demand tolerates small shifts; a cliff can cause an $O(1)$ error.

A One-Sided Escape: Markdown

Lower prices **never reduce demand**.
Their revenue downside is at most the markdown:

$$r_x(p - \varepsilon) \;\ge\; r_x(p) - \varepsilon$$



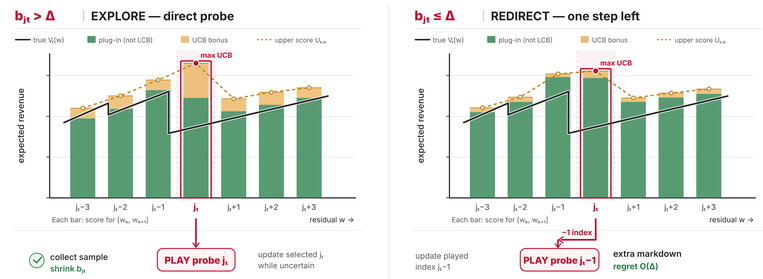
Raising price may cross a cliff. Lowering price has a **controlled downside**.

Regret Landscape

Work	Noise assumption	Regret
Javanmard–Naz. '19	Known log-concave	$\widetilde{O}(\log T)$
Ban–Keskin '21	Parametric	$\widetilde{O}(\sqrt{T})$
Luo et al. '22	2nd-order smooth	$\widetilde{O}(T^{2/3})$
Fan et al. '24	Higher-order smooth	$\widetilde{O}(T^{\frac{2m+1}{4m-1}})$
Tullii et al. '24	Lipschitz	$\widetilde{O}(T^{2/3})$
Xu–Wang '22	Agnostic bounded	$\widetilde{O}(T^{3/4})$
This work	Arbitrary jumps/atoms	$\widetilde{O}(T^{2/3})$

CMRUP: The Algorithm

1. **Estimate value.** Random prices → moment estimate $\hat{\theta} \to \hat{u}_t = x_t^\top \hat{\theta}$.
2. **Conservative grid.** $w_j = -c + 2j\Delta$; probe j at $p_t^{(j)} = \hat{u}_t + w_{j+1} - 3\Delta$.
3. **Confidence-gated UCB.** Explore the selected option until confident; then redirect one grid step left.



Direct probes **learn** the selected option. Redirecting avoids paying for its adjacent demand jump again and again.

$$a_t = \begin{cases} j_t, & b_{j_t} > \Delta \quad \text{(explore)} \\ j_t - 1, & b_{j_t} \le \Delta,\ j_t > 0 \quad \text{(redirect)} \end{cases}$$

Proof Idea: Four Costs

$$R_T \lesssim (T_1 + T_w) + T\Delta + \sqrt{TK \log T} + \frac{\log T}{\Delta^2}$$

initial exploration · grid/markdown · UCB · jump cost

Set $K \asymp 1/\Delta$. Non-overlapping jump mass: $\sum_j \alpha_j \le 1$.

Direct probes per cell: $\widetilde{O}(\Delta^{-2})$; redirected rounds cost only $O(\Delta)$.

$$\Delta \asymp T^{-1/3} \quad \Longrightarrow \quad R_T = \widetilde{O}(T^{2/3})$$

Jianyu Xu and Yu-Xiang Wang. NeurIPS 2026.
Optimal Contextual Pricing under Agnostic Non-Lipschitz Demand.
jxu22@charlotte.edu arXiv:2605.05609