No Free Best-of-Both-Worlds Learning in Repeated Bilateral Trade

Yutian Cheng Canzhe Zhao Jingye Zhao Shuai Li

Shanghai Jiao Tong University (Yutian Cheng, Jingye Zhao) | National University of Singapore (Canzhe Zhao) | The University of Hong Kong and Shanghai Innovation Institute (Shuai Li)

Main message: no free BoBW without side information $\Longrightarrow$ one scalar hint $g_\nu^\star$ restores the classical rates

Contributions

- **Question.** Can one learner adapt between *independent-values* markets and the full *bounded-density* class using only two accept/reject bits?
- **Negative result.** Without extra information, the classical best-of-both-worlds target is impossible.
- **Sharp frontier.** We characterize the optimal trade-off between the easy IV regime and the hard dependent regime.
- **Positive result.** A single scalar hint—the optimal same-price value $g_\nu^\star$—is enough to recover the classical rates.

Setting

At each round $t = 1, 2, \ldots$, a seller and a buyer arrive with private values S_t and B_t. The platform posts prices $P_t \le Q_t$. It only observes whether the seller accepts and whether the buyer accepts; it never sees their values or the realized gain from trade.

The per-round reward is

$$Y_t = \big((P_t - S_t) + (B_t - Q_t)\big)\mathbf{1}\{S_t \le P_t \le Q_t \le B_t\}.$$

The learner observes only the two acceptance bits

$$X_t^S = \mathbf{1}\{S_t \le P_t\}, \qquad X_t^B = \mathbf{1}\{B_t \ge Q_t\}.$$

The same-price gain curve and benchmark are

$$g_\nu(x) = \mathbb{E}_\nu[(B - S)\mathbf{1}\{S \le x \le B\}],$$
$$g_\nu^* = \sup_{x\in[0,1]} g_\nu(x).$$

The expected regret is

$$R_T(A;\nu) = Tg_\nu^* - \mathbb{E}_\nu\left[\sum_{t=1}^T Y_t\right].$$

Seller (S_t) $\xrightarrow{P_t}$ Platform $(P_t \le Q_t)$ $\xrightarrow{Q_t}$ Buyer (B_t)

$$X_t^S = \mathbf{1}\{S_t \le P_t\}, \qquad X_t^B = \mathbf{1}\{B_t \ge Q_t\}.$$

Why adaptation is difficult

- Independent values (IV) admit regret $\widetilde{\Theta}(T^{2/3})$, while the general bounded-density (BD) class admits $\widetilde{\Theta}(T^{3/4})$.
- Hidden dependence can preserve both marginals, so same-price feedback cannot certify independence.
- Detecting hidden structure requires costly two-price exploration, which introduces unavoidable additional regret.

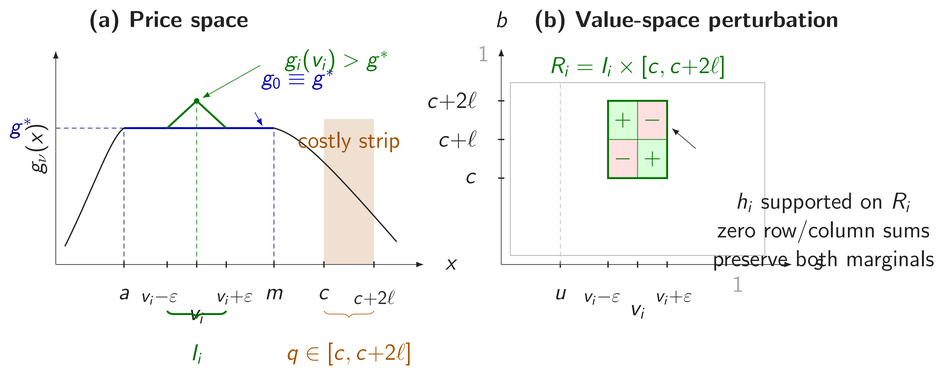
Theorem 1: No-extra-information lower bound

There exists a density level $M_* > 0$ such that, for every $\alpha \in [0,1]$ and every algorithm A,

$$\sup_{\nu\in\mathcal{E}_{\mathrm{IV}}(M_*)} R_T(A;\nu) = O(T^\alpha) \implies \sup_{\mu\in\mathcal{H}(M_*)\setminus\mathcal{E}_{\mathrm{IV}}(M_*)} R_T(A;\mu) = \Omega(T^{1-\alpha/3}).$$

In particular, any algorithm with $O(T^{2/3})$ regret on the IV class must incur $\Omega(T^{7/9})$ regret on some non-IV instance. Hence the target pair $(2/3, 3/4)$ is impossible.

Lower-bound construction



Why this family is hard. The IV base law has a flat optimal plateau, so same-price observations do not reveal a preferred price. The checkerboard perturbation preserves both marginals while changing the joint distribution. Distinguishing the two instances therefore requires costly two-price exploration.

Theorem 2: Matching trade-off

For every $M > 0$ and $\gamma \in (0, 1/3)$, one algorithm simultaneously satisfies

$$\sup_{\nu\in\mathcal{E}_{\mathrm{IV}}(M)} R_T(\mathrm{SCM};\nu) \le C_1(M,\gamma)\,T^{\max\{2/3,3\gamma\}}\log(eT),$$

and

$$\sup_{\nu\in\mathcal{H}(M)} R_T(\mathrm{SCM};\nu) \le C_2(M,\gamma)\,T^{\max\{1-\gamma,3\gamma\}}\log(eT).$$

Algorithm idea. Compare the true same-price gain curve $g_\nu(x)$ with the product-surrogate curve $\widetilde{g}_\nu(x)$. If they agree, temporal decoupling creates virtual product-law feedback and an IV learner is used. Otherwise, the algorithm falls back to a robust bounded-density strategy.

Four-step view.

- Probe a grid of same prices and estimate g_ν and $\widetilde{g}_\nu$.
- Test whether the two curves are uniformly close.
- If consistent, use fresh independent rounds to construct virtual product-law feedback.
- If inconsistent, switch to the robust bounded-density strategy.

Theorem 3: One scalar restores BoBW

Knowing only the optimal same-price value $g_\nu^\star$ is enough for a value-auditing strategy to recover both classical rates.

$$\text{Independent values:} \quad R_T(\mathrm{VAM};\nu) = \widetilde{O}(T^{2/3}),$$
$$\text{General bounded density:} \quad R_T(\mathrm{VAM};\nu) = \widetilde{O}(T^{3/4}).$$

The learner does not need to know the maximizing price and never observes realized trade gains.

Reading the trade-off frontier

If the IV regret exponent is α, Theorem 1 forces some dependent instance to have regret exponent at least $1 - \alpha/3$. Theorem 2 matches this curve up to logarithmic factors, and at the classical IV target $\alpha = 2/3$ the best simultaneous guarantee is $7/9$. Theorem 3 shows that the single scalar hint g_ν^* removes this gap and restores the classical $3/4$ bounded-density rate.

Further discussion

Welfare conventions. We use residual-surplus gain from trade. The regret guarantees also transfer to the gross-surplus convention via the standard conversion.

Lay summary. Online marketplaces receive only limited accept/reject feedback from the seller and buyer. This work identifies the limits of adapting to hidden market structure and shows that one extra scalar can restore the best-known guarantees.

Selected references

[1] N. Cesa-Bianchi, T. Cesari, R. Colomboni, F. Fusco, and S. Leonardi. *Regret Analysis of Bilateral Trade with a Smoothed Adversary*. JMLR 25(234):1–36, 2024.

[2] N. Cesa-Bianchi, T. Cesari, R. Colomboni, F. Fusco, and S. Leonardi. *Bilateral Trade: A Regret Minimization Perspective*. Mathematics of Operations Research 49(1):171–203, 2024.

[3] N. Cesa-Bianchi, T. R. Cesari, R. Colomboni, F. Fusco, and S. Leonardi. *Repeated Bilateral Trade against a Smoothed Adversary*. COLT, 2023.

Hidden structure
Same marginals do not reveal dependence. The learner must probe different seller and buyer prices.

Unavoidable trade-off
Faster learning in the IV case worsens the guarantee on some dependent instances.

One scalar is enough
Knowing $g_\nu^\star$ restores $\widetilde{O}(T^{2/3})$ and $\widetilde{O}(T^{3/4})$ guarantees.

Contact me via cyt2021@sjtu.edu.cn