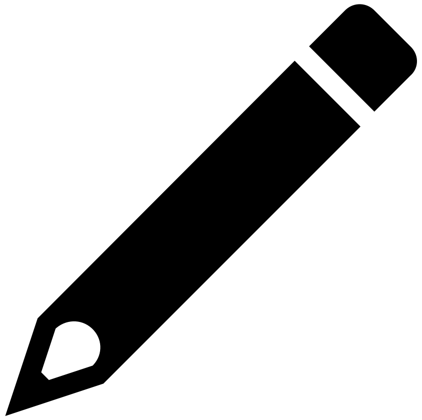
Learning Thresholds with Latent Values and Censored Feedback

Jiahao Zhang[1] Tao Lin[2] Weiqiang Zheng[3]
Zhe Feng[4] Yifeng Teng[4] Xiaotie Deng[1]

[1]Peking University [2]Harvard University [3]Yale University [4]Google Research

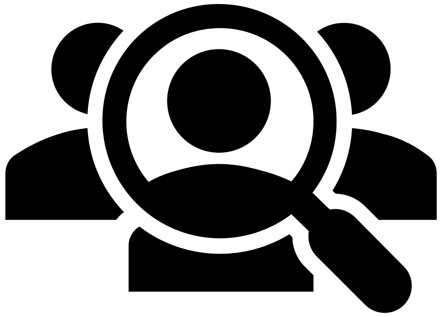
ICLR 2024

Introduction



P/F exams

Auctions



Employment

Model

- A one-shot interaction between a learner and an agent is:
  1. The learner sets a threshold γ,
  2. The agent draws a value v (latent value) from a distribution F,
  3. If $v \geq \gamma$, the learner will observe a reward feedback $g(\gamma, v)$; if $v < \gamma$, the learner will observe nothing. (censored feedback).

In the posted price auctions, $g(\gamma, v) = \gamma$.

- Note: The value distribution F and reward function g are both **unknown** to the learner.

Model

- But the learner knows that
  - the value distribution F is from a set C,
  - the reward function g is from a set G.
- For any threshold γ, the learner's expected reward is

$$U(\gamma) = E_{v \sim F}[g(\gamma, v) \cdot \mathbf{1}_{v \geq \gamma}].$$

- **How many queries are needed to learn a threshold such that the expected reward is at most $\boldsymbol{\varepsilon}$ smaller than the optimum?**

$\mathrm{QC}_{C,G}(\varepsilon) =$ the minimum queries needed to learn such a threshold for any $F \in C$ and $g \in G$.

Main Result: Overview

Reward/Value	General		Lipschitz	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
Monotone	Infinite		$\Omega\left(\frac{1}{\varepsilon^3}\right)$	$\tilde{O}\left(\frac{1}{\varepsilon^3}\right)$
Right-Lipschitz	$\Omega\left(\frac{1}{\varepsilon^3}\right)$	$\tilde{O}\left(\frac{1}{\varepsilon^3}\right)$		

We also extend this model to an online learning setting and demonstrate a **tight** $\Theta\left(T^{2/3}\right)$ regret based on these offline results.

Outline

- Introduction
- Model
- Main Result
  - **Overview**
  - Impossibility
  - Upper Bound
  - Lower Bound

Main Result: Impossibility

TL;DR: "Needle in the Haystack"

- Parameterized pairs of monotone reward function and value distribution (g_α, F_α).
- Given (g_α, F_α), the optimal threshold is $\gamma^* = \alpha$.
- The optimum $U(\gamma^*)$ is larger than any expected reward $U(\gamma)$ $(\gamma \neq \gamma^*)$ by a constant.
- The learner must know the exact value of α!
- Our construction allows any $\alpha \epsilon \left(\frac{1}{2}, \frac{9}{16}\right) \Rightarrow$ infinite queries.

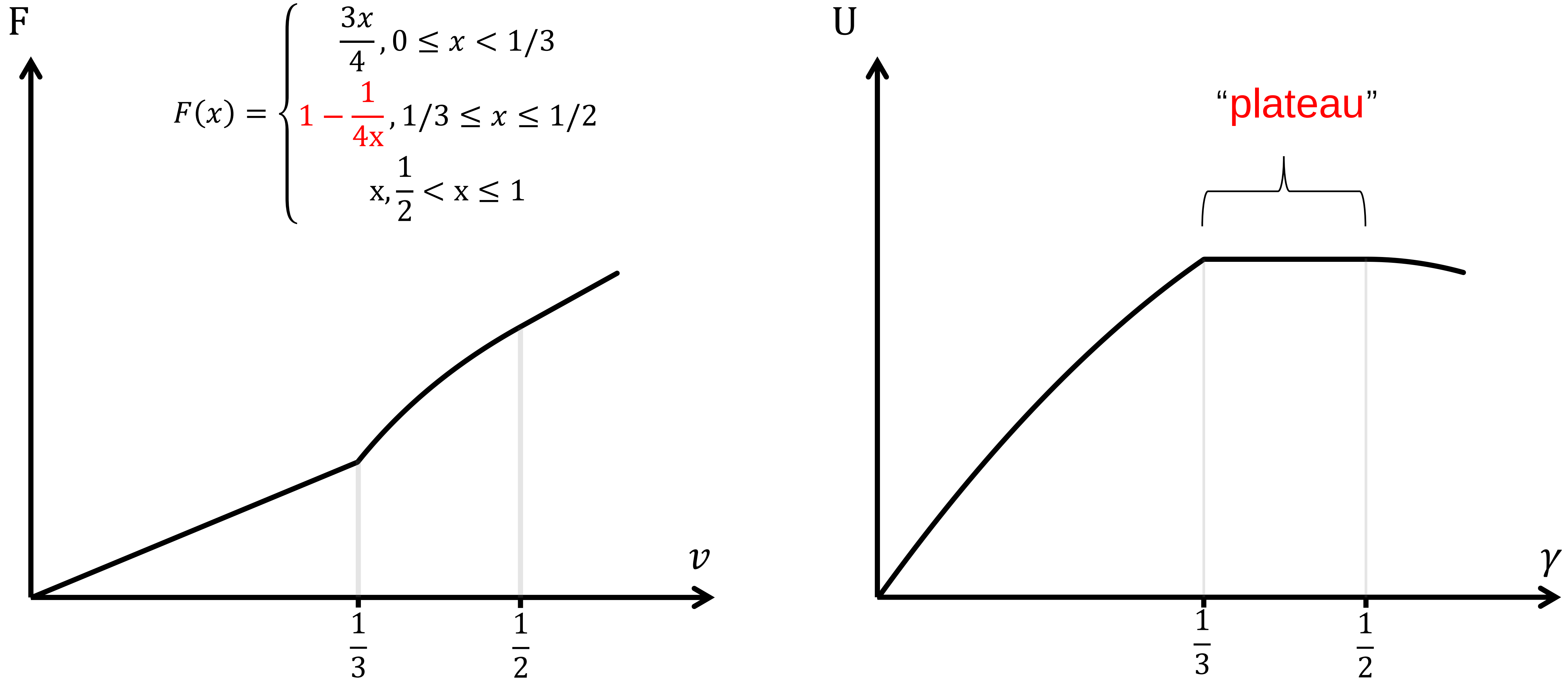
Main Result: Upper Bound

- Lipschitzness of G or C allows us to uniformly discretize on the range $[0,1]$ of threshold γ: $\Gamma = \{0, \frac{1}{\varepsilon}, \ldots, 1\}$, $|\Gamma| = O(\frac{1}{\varepsilon})$.
- For $\gamma \in \Gamma$, $\tilde{O}\left(\frac{1}{\varepsilon^2}\right)$ queries are sufficient to learn the corresponding expected reward $U(\gamma)$ with ε error.
- Choose the threshold that has the largest estimated expected reward.
- Totally $\tilde{O}\left(\frac{1}{\varepsilon^3}\right)$ queries are sufficient.

TL;DR:
Discretization

Main Result: Lower Bound

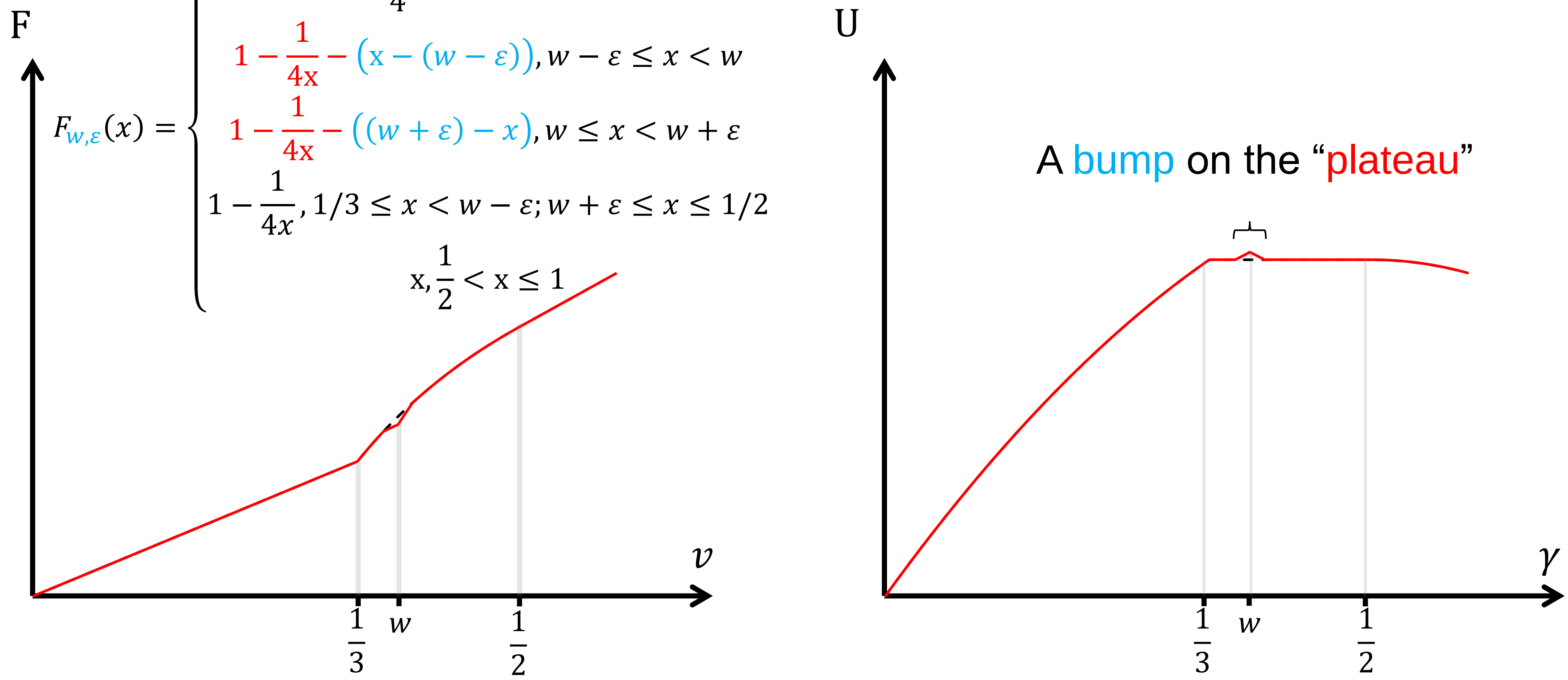
- Simple reward function $g(\gamma, v) = \gamma$.
- Base value distribution satisfies that any $\gamma \in [\frac{1}{3}, \frac{1}{2}]$ is optimal.
- Perturb the base value distribution so that
  - Perturbed value distribution satisfies that a unique $\gamma^* \in [\frac{1}{3}, \frac{1}{2}]$ is optimal
  - One perturbation is operated on a subinterval of $[\frac{1}{3}, \frac{1}{2}]$ that has length $O(\varepsilon)$.
  - Subintervals of different perturbations are required to be disjoint $\Rightarrow O\left(\frac{1}{\varepsilon}\right)$ perturbations.



$$F(x) = \begin{cases} \frac{3x}{4}, 0 \leq x < 1/3 \\ 1 - \frac{1}{4x}, 1/3 \leq x \leq 1/2 \\ x, \frac{1}{2} < x \leq 1 \end{cases}$$

Step 1: base distribution

$$F_{w,\varepsilon}(x) = \begin{cases} \frac{3x}{4}, 0 \le x < 1/3 \\ 1 - \frac{1}{4x} - \left(x - (w - \varepsilon)\right), w - \varepsilon \le x < w \\ 1 - \frac{1}{4x} - \left((w + \varepsilon) - x\right), w \le x < w + \varepsilon \\ 1 - \frac{1}{4x}, 1/3 \le x < w - \varepsilon; w + \varepsilon \le x \le 1/2 \\ x, \frac{1}{2} < x \le 1 \end{cases}$$



Step 2: perturbed distribution

Main Result: Lower Bound

- We prove that $\Omega\left(\frac{1}{\varepsilon^2}\right)$ queries are needed to distinguish base distribution and one perturbed distribution.
- Totally $\Omega\left(\frac{1}{\varepsilon^3}\right)$ queries are necessary.

TL;DR: Reduce to distinguishing similar distributions

Thank you for your attention.