Robust Pricing in Contextual Auctions

Authors: Negin Golrezaei (Massachusetts Institute of Technology, Sloan School of Management)
Adel Javanmard (University of Southern California, Marshall School of Business)
Vahab Mirrokni (Google Research, New York)

The Neural Information Processing Systems Conference, 2019

(Contextual) Online Markets

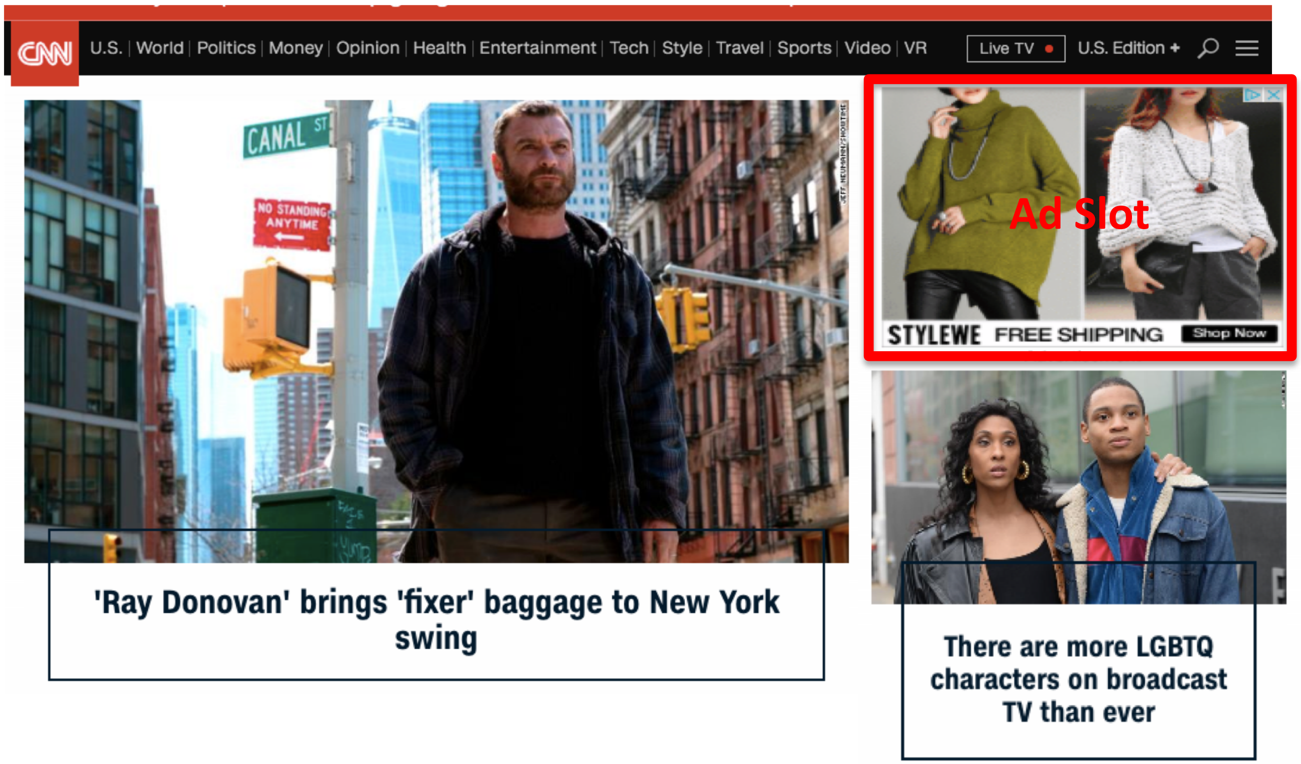
Detailed Contextual Information

- With contextual information, products become highly-differentiated
- Heterogeneous markets: Contextual information changes buyers' willingness-to-pay possibly in a heterogeneous way

Seller can set personalized and contextual prices

Motivation

Display advertising markets



To earn high revenue, setting right prices is crucial [Ostrovsky and Schwarz'11, Beyhaghi, Golrezaei, Paes Leme, Pal, and Sivan'18]

How to set Personalized and Contextual Prices?

Typical approach: Use historical data to learn optimal prices

Challenges:

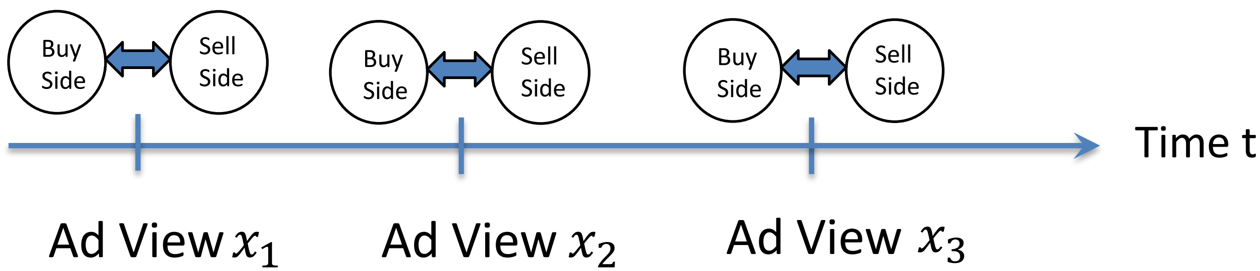
- Billions of auctions every day
- Repeated interactions between advertisers and the platform
- Advertisers are strategic
  - They can have an incentive to manipulate the learning algorithm

Lower bids now → See lower prices later

Goal: Design a low-regret dynamic pricing policy for seller that is "robust" to strategic buyers

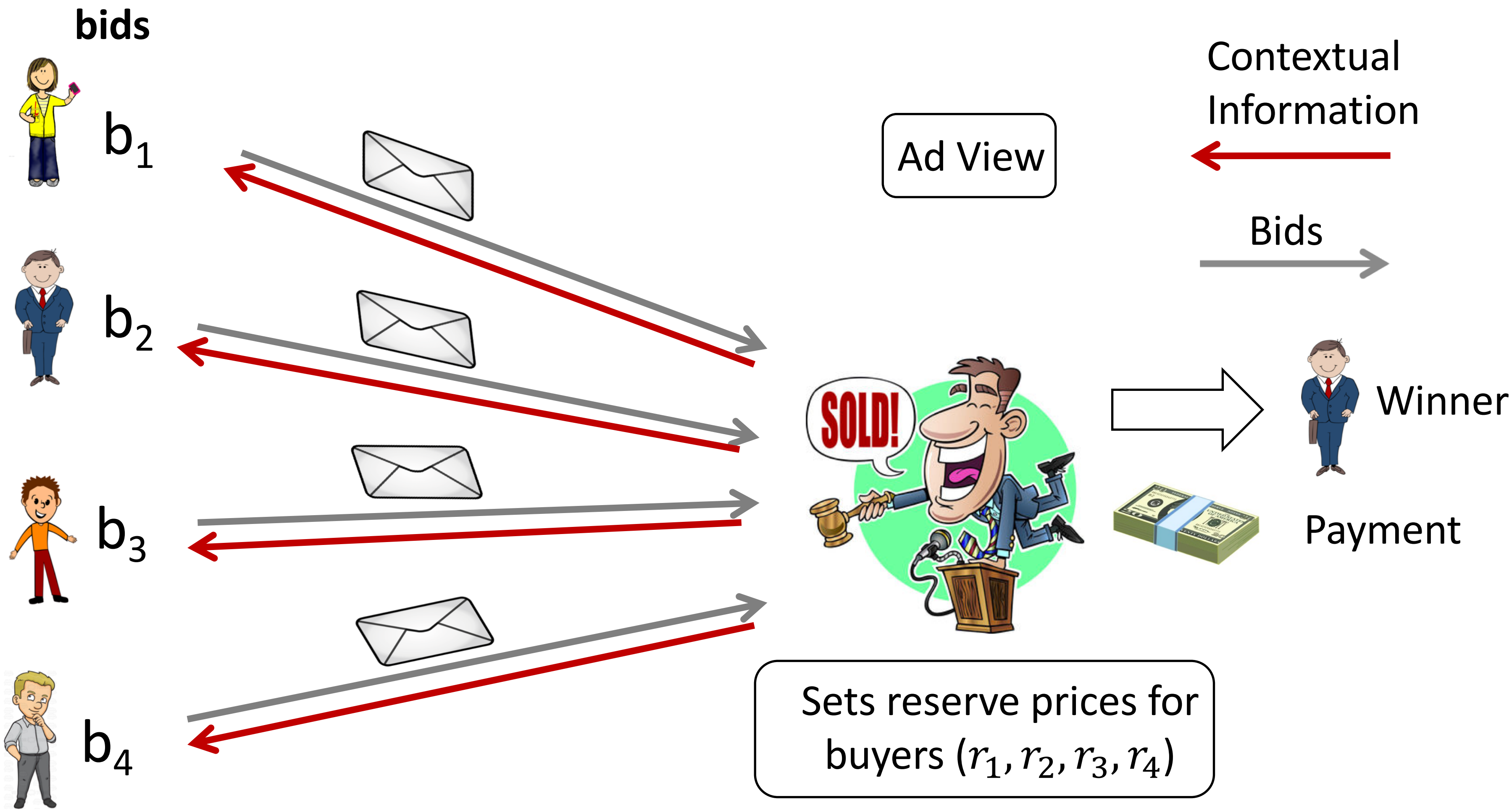
Model

- N buyers (advertisers) and one seller (Ad exchange)
- Items (ad views) are sold over time (one item at the time)
- Each item at time t is described by feature vector $x_t \in \mathbb{R}^d$
  - Features x_t is drawn independently from an unknown distribution
  - Features themselves are known to the buyers and the seller



- The item is sold via a second-price auction with reserves
- Each buyer i has private valuation v_{it} of the item

Second Price Auctions with Reserve



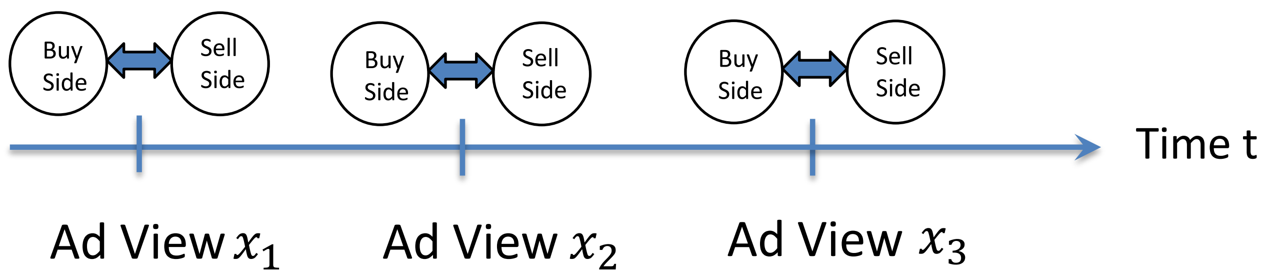
Winner is the buyer with the **highest submitted bid** if he clears his reserve

Payment of winner = **max(second highest bid, winner's reserve)**

Repeated Second Price Auctions

- Widely used in practice because it is simple and truthful

- **With repeated interactions:**



  - Both sides can try to learn their optimal strategy
  - Buyers have incentive to bid untruthfully
  - Buyers may sacrifice their short-term utility to game the seller and lower their future reserve prices (strategic buyers)

Buyer's Valuation

- We focus on a linear model for valuations:

$$v_{it} = \langle x_t, \beta_i \rangle + z_{it}$$

  - Item's feature vector x_t (observable)
  - Preference vectors β_i (unknown to seller a priori, fixed over time)
    - Normalization: $\|\beta_i\| \leq B_p, \|x_t\| \leq 1$
  - Market shocks z_{it} (unobservable)
    - Noise in the valuation model
  - Noise terms z_{it} are drawn i.i.d. from a mean zero distribution
    - $F: [-B_n, B_n] \rightarrow [0,1]$
  - Distribution F and 1- F is log-concave (e.g., normal, Laplace, uniform, etc)

Known F	$\Longrightarrow$	CORP Policy
Unknown F	$\Longrightarrow$	SCORP Policy

Buyers are Utility-maximizer

- Buyer's utility at time t: $u_{it} = v_{it}\boxed{q_{it}} - p_{it}$
  - allocation variables q_{it}: (1 if buyer i gets the item, 0 otherwise.)
- Buyers maximize their time-discounted utility

$$U_i = \sum_{t=1}^{\infty} \gamma^t \, \mathbb{E}[u_{it}]$$

- γ discount factor: Seller is more patient than buyers
  - Buyers would like to target users sooner rather than later

Summary of Contributions and Techniques

Summery of Contributions:

- Known market noise distribution:
  - CORP with regret $O(d\log(Td)\log(T))$
    - d is dimension of contextual information and T is the length of time horizon
- Unknown market noise distribution:
  - SCORP with regret $O\left(\sqrt{d\log(Td)}\ T^{2/3}\right)$

Techniques: to have a low regret policy,

- Using censored bids
- Taking advantage of an episodic structure to lower buyers' incentive for being untruthful

Related Work

- Non-contextual dynamic pricing with learning
  - Bayesian setting: [Farias and Van Roy'10, Harrison et al.'12, Cesa-Bianchi et al.'15, Ferreira et al.'16, Cheung et al. '17]
  - (Frequentist) parametric models: [Broder and Rusmevichientong '12, Besbes and Zeevi '09, den Boer and Zwart '13]
- Contextual dynamic pricing/non-strategic buyers: [Chen et al. 2015, Cohen et al. 2016, Lobel et al. 2016, Javanmard, Nazerzadeh 2016, Ban and Keskin 2017, Javanmard 2017, Shah et ak. 2019]

Pricing with strategic buyers	Contextual	Multiple buyers	Heterogeneity	Noise
Amin et al.'13 and Medina and Mohri'14	✗	✗	NA	✓
Amin et al. 2014	✓	✗	NA	✗
Kanoria and Nazerzadeh'17	✗	✓	✗	✓
Our work	✓	✓	✓	✓

Known Market Noise Distribution:

Contextual Robust Pricing (CORP)

Setting and Benchmark

- **Setting:** The market noise distribution F is known.
- **Benchmark**: A clairvoyant who knows preference vectors β_i

Proposition

If the (clairvoyant) seller knows the preference vectors $\{\beta_i\}_{i\in[N]}$, then the optimal reserve price of buyer $i \in [N]$, for a feature x is given by

$$r_i^*(x) = \arg\max_y \{y(1 - F(y - \langle x, \beta_i \rangle))\}$$

Further, $r_{it}^* = r_i^*(x_t)$.

- Benchmark is measured against truthful buyers
- Optimizing reserve prices becomes decoupled

$$r_{it}^* = \arg\max_y \{y\, \mathbb{P}(v_{it}(x_t) \geq y)\}$$

Seller's Regret against the Benchmark

- Seller does not know the preference vectors

Definition: Regret

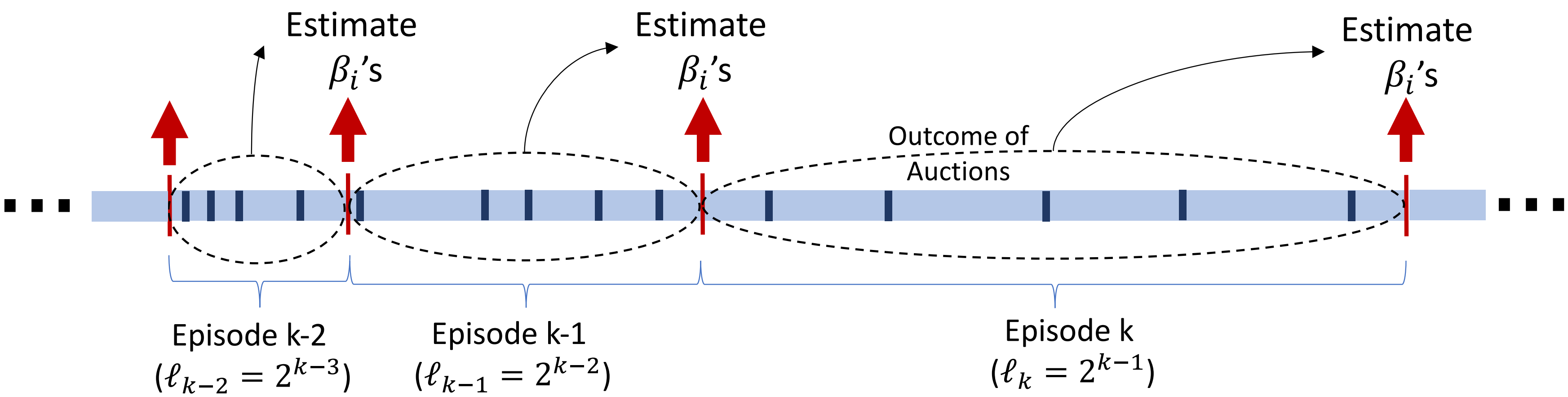
The worst-case cumulative regret of a policy π is defined by

$$\mathrm{Reg}^{\pi}(T) = \max\left\{\sum_{t=1}^{T}(\mathrm{rev}_t^* - \mathrm{rev}_t^{\pi}) : \|\beta_i\| \le B_p, \ \text{ for } i \in [N], \text{ feature distribution}\right\}$$

Here, rev_t^*, rev_t^{π} are the expected revenue of the benchmark and policy π, at time t.

- Getting a low regret is challenge because the benchmark is strong:
  - Under benchmark, buyer are truthful
  - Prices in the benchmark are personalized and contextual

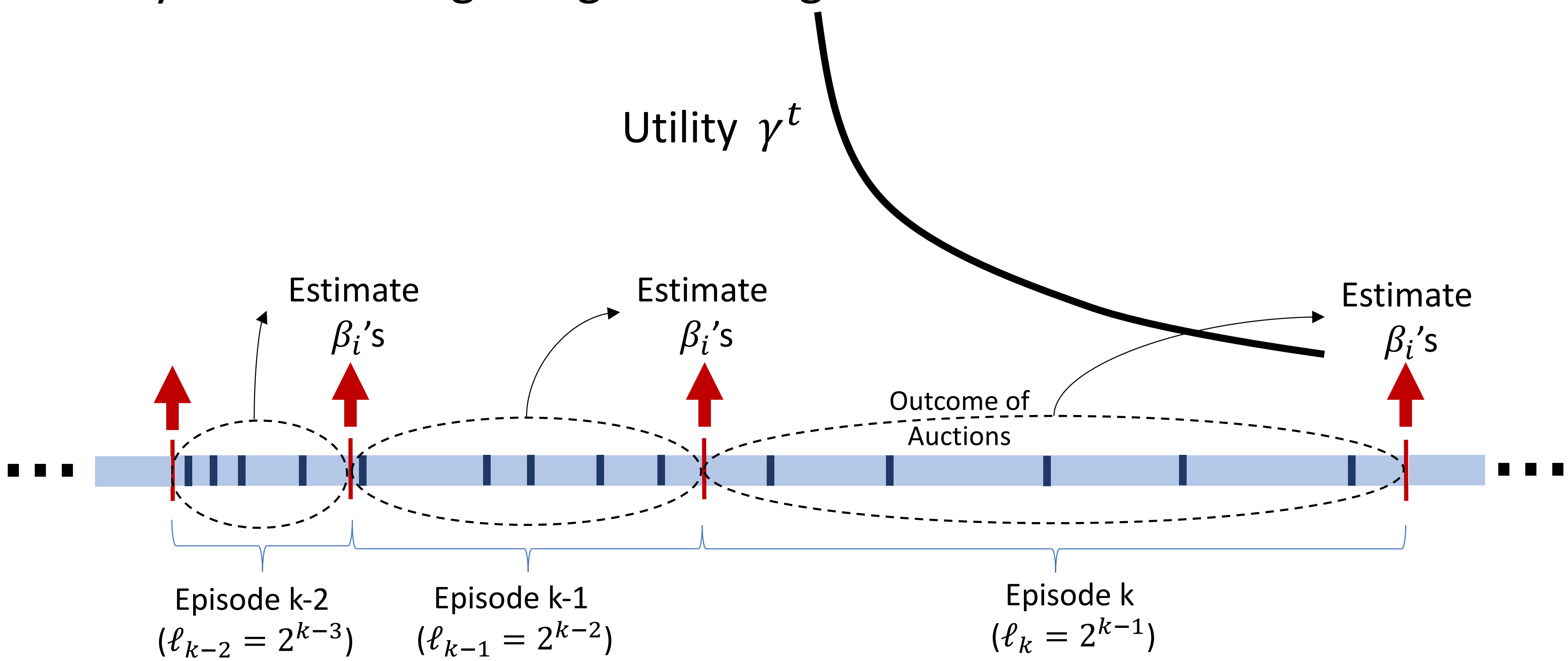
Overview of CORP



- Episodic structure: Updates preference vectors β only at the beginning of each episode.
- Random Exploration: For each period t in episode k, do exploration with Prob. 1/length of episode
  - Choose one buyer uniformly at random and set his reserve price uniformly at random from $[0, B]$ and set other reserves to ∞.
- Exploitation: Use the estimate of β to set prices

Why Episodic Structure?

- Buyers are less patient than the seller (Buyers' utilities are discounted over time)
- Buyers are strategic to get future gain



The episodic structure limits the long-term effects of bids

How Do we Do Exploitation?

- **Q1:** How to estimate preference vectors β_i's?
- **Q2:** How to set reserve prices based on the estimated preference vectors $\hat{\beta}_{ik}$?

Q1: How to Estimate Preference Vectors β_i?

- **Goal**: reduce buyer's incentive to be untruthful
- A Potential approach:

We don't use your bids to set your reserve prices

- The premise is that mechanism remains "truthful" over time.
- Impossible to do this because buyers are heterogeneous

Relaxed statement**:**

We don't rely too much on your bids to set your reserve prices.

- Noisy bids/ randomized algorithm [Mahdian et. al 2018, McSherry and Talwar '07]
  - Large markets
- **Censored bids (We follow this path)**

Using Censored Bids in Our Estimation

- Use bids submitted by other buyers and the outcomes of auctions
- Not the bids submitted by that buyer!

  - Minimize the negative of log likelihood function of outcomes (auction outcome q_{it}) if buyer i bids truthfully

$$\widehat{\beta_{ik}} = argmin_{\beta}\mathcal{L}_{ik}(\beta) \quad i \in [N]$$

$$\mathcal{L}_{ik}(\beta) = -\frac{1}{\ell_{k-1}}\sum_{t\in E_{k-1}} q_{it}\log(1 - F(\max\{b_{-it}^{+}, r_{it}\} - \langle x_t, \beta\rangle)) + (1 - q_{it})\log\left(F\left(\max\{b_{-it}^{+}, r_{it}\} - \langle x_t, \beta\rangle\right)\right)$$

  - b_{-it}^{+}: maximum bids submitted at period t other than b_{it}

If a buyer wants to influence the estimation, he needs to change the outcome of auction! → Very costly

Q2: How to Set Reserve Prices?

- For all periods t in episode k, we set the reserve prices r_{it} as follows

$$r_{it} = \arg\max_{y} \left\{ y \left(1 - F\left(y - \langle x_t, \hat{\beta}_{ik} \rangle \right) \right) \right\}$$

- $\hat{\beta}_{ik}$ is the estimate of β_i computed at the beginning of episode k.

Our Benchmark

If the (clairvoyant) seller knows the preference vectors $\{\beta_i\}_{i \in [N]}$, then the optimal reserve price of buyer $i \in [N]$, for a feature x is given by

$$r_i^*(x) = \arg\max_{y} \{ y(1 - F(y - \langle x, \beta_i \rangle)) \}$$

Further, $r_{it}^* = r_i^*(x_t)$.

Regret Bounds on CORP

Theorem (Regret bound for CROP)

Suppose that the firm knows the market noise distribution F. Then, the T- period worst-case regret of the CORP policy is at most $O(d \log(Td) \log(T))$, where the regret is computed against the benchmark.

Unknown market noise distribution:

Stable Contextual Robust Pricing (SCORP)

What is Different from CORP?

- **Setting**: Seller does not know the market noise distribution F.
  - There is an *ambiguity set* $\mathcal{F}$ of possible distributions and propose a policy that works well for every distribution in the ambiguity set.
- **Benchmark:** A clairvoyant that knows preference vectors β_i , ambiguity set $\mathcal{F}$

Definition **(Stable Benchmark)**

In the stable benchmark, the reserve price of buyer $i \in [N]$, for a feature x is given by

$$r_i^*(x) = \arg\max_{y} \min_{F \in \mathcal{F}} \{y(1 - F(y - \langle x, \beta_i \rangle))\}$$

Further, $r_{it}^* = r_i^*(x_t)$.

Without knowing F, we need to do more exploration

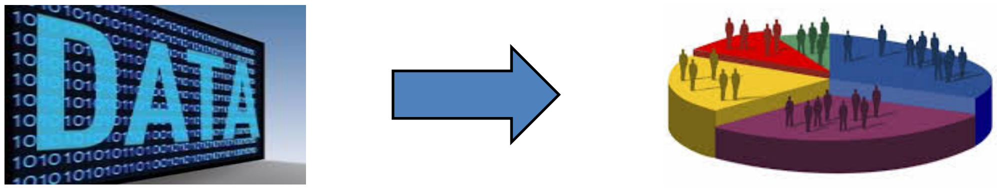
Regret Bounds on SCORP

Theorem (Regret bound for SCROP)

Suppose that the market noise distribution is unknown and belongs to ambiguity set $\mathcal{F}$. Then, the T- period worst-case regret of the SCORP policy is at most $O\left(\sqrt{d \log(Td)}\ T^{2/3}\right)$, where the regret is computed against the stable benchmark.

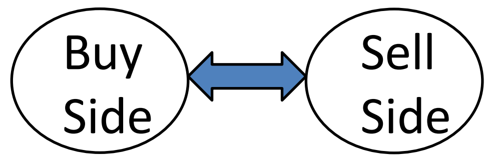
Takeaway

- Optimizing personalized and contextual-based prices



Heterogeneous and contextual markets

- Robust against strategic buyers:
  - Episodic structure of the policy
  - Censored bids



- CORP policy
  - Known market noise distribution--Worst-case regret $O(d\log(Td)\log(T))$

- SCORP policy
  - Unknown market noise distribution--Worst-case regret $O\left(\sqrt{d\log(Td)}\;T^{2/3}\right)$
  - Stable against uncertainty in noise distribution