上海纽约大学
NYU SHANGHAI

CUHK
BUSINESS SCHOOL

Cold Start to Improve Market Thickness on Online Advertising Platforms: Data-Driven Algorithms and Field Experiments

Renyu (Philip) Zhang

(Joint work with Zikun Ye, Dennis J. Zhang, Heng Zhang, Xin Chen)

1. Introduction and Contribution
2. Theory: Model, Algorithm and Analysis
3. Practice: Field Implementation, Experiment Design and Empirical Results

Online Advertising Platform

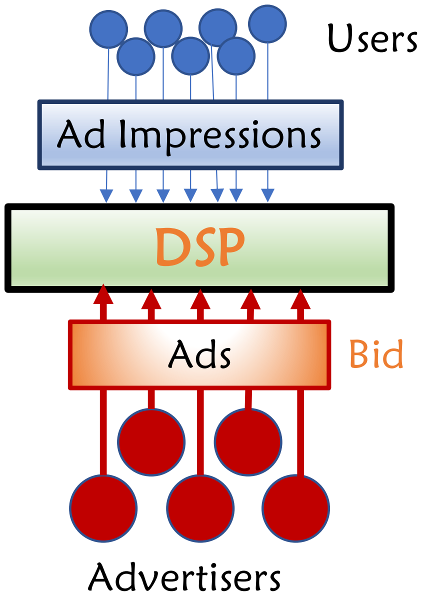
- Online advertising platform: Demand Side Platform (DSP)

- Fundamental operations question of a DSP:

  When an ad request (user impression) arrives, which ad should be displayed to her?

- Core business logic:

  The DSP runs large-scale auctions to determine which ad to display, in order to maximize the advertising revenue of each user ad impression.



Performance-Based In-Feed Ads

Click-Through Rate (CTR)

Conversion Rate (CVR)

Ad Impression on the Platform

Registration form of the Advertiser

DSP and Advertisers

- Performance-based ads: Advertisers want more conversions at a low cost.
  - Mobile games: Activation & deposit
  - eCommerce: Activation & purchase

$$\text{Impression} \xrightarrow{\text{CTR}} \text{Click} \xrightarrow{\text{CVR}} \text{Conversion}$$

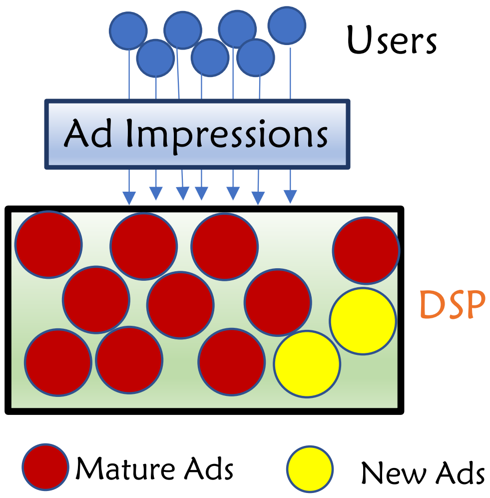
- Auction Mechanism & Billing Option: Optimized Cost-Per-Click (oCPC)
  - The advertisers bid on conversions and pay upon clicks (a compromise between the advertisers and the DSP).
  - The ads are ranked by the expected cost-per-mile (eCPM), which is the expected revenue per unit impression.
  - The impression is allocated to the ad with the highest eCPM.

Mechanism	Charged-upon	Fee-Deduction	Rank-by
oCPC	click	pCVR * bid_convert	$\text{eCPM} = \text{pCTR} * \underbrace{\text{pCVR} * \text{bid\_convert}}_{\text{Cost Per Click (CPC)}}$

pCTR (predicted CTR) and pCVR (predicted CVR conditioned on click) are produced by underlying deep neural networks of the DSP

Cold Start on DSP

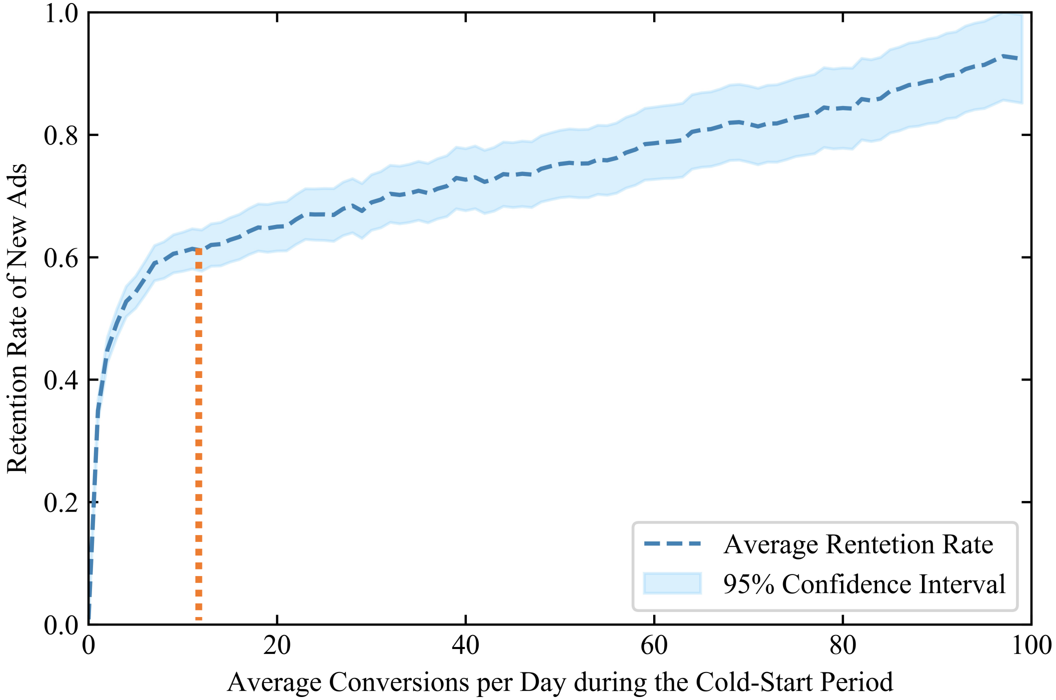
- Cold Start: Learning pCTR and pCVR efficiently with limited data for new ads.
- New ads:
  - No enough data to estimate pCTR and pCVR accurately.
  - Unclear revenue implications for the DSP.
  - Successful cold start of new ads thickens the ad pool and boosts advertiser retention.
- Mature ads: High and stable revenues, minimally affecting user experiences.



Core problem in cold start: How to allocate user impressions between new and mature ads to balance the short-term revenue and the long-term market thickness?

Exploitation vs Exploration

Cold Start on DSP: Ad Retention and Market Thickness



- X-axis: # of conversions per day in the first 3 days.
- Y-axis: Proportion of ads that will stay active on the DSP every single day in the next 2 weeks.
- Everything is rescaled.

- Key observation: If the # of conversions in the cold start period surpasses 10, the long-term retention and value soon flatten.
- Cold start will have a non-linear long-term impact on the market thickness and advertising revenue.
  - Addressing cold start through an operations lens.
- Causality of the figure: PSM and IV analysis.
- The phenomenon of the left figure is robust with respect to different definitions of ad retention.

Key Research (and Business) Questions

- With the inaccurate predictions of CTR and CVR for new ads, how to smartly balance short-term revenue and long-term market thickness/revenue?

Primal-dual based MAB algorithms:
Shadow Bidding with Learning (SBL)

- With different ads competing for the same user impressions, how to unbiasedly estimate the value of our proposed algorithm?

User-ad two-sided experiment framework

Related Literature

- Ad Cold Start: **More accurate** CTR and CVR predictions for new ads.
  - Dave and Varma (2012, 2014), Zhou et al. (2018), Choi et al. (2020), etc.

- Contextual Bandits: Establishing sublinear regret bounds with **optimization oracles**.
  - Langford and Zhang (2007), Chu et al. (2011), Bandanidiyuru et al. (2013), Agrawal et al. (2014), Agrawal et a. (2016), Jacot et al. (2018), Arora et al. (2019), Simchi-Levi and Xu (2020), etc.

- Operations problem with online learning: Bandit learning algorithms applied to **pricing, inventory, and advertising** problems.
  - Besbes and Zeevi (2009), Nambiar et al. (2019), Chen et al. (2019), Golrezaei et al. (2019), etc.

- Experimental evaluations of algorithms on two-sided platforms: Debiasing the estimate when **SUTVA does not hold**.
  - Ha-Thuc et al. (2020), Johari et al. (2020), Bojinov et al. (2020), Candogan et al. (2021), etc.

上海纽约大学 NYU SHANGHAI

Highlight of Main Contributions

CUHK BUSINESS SCHOOL

- **End-to-end Solution:**
  - Implement a novel data-driven algorithm online to address the cold start challenge for a large-scale advertising platform with minimal engineering adjustments
  - Connect learning theory and online advertising practice
- **Theory and Algorithm:**
  - Tackle the cold-start challenge through an operations lens
  - SBL algorithm: Duality + MAB + neural networks
- **Practice and Experiment:**
  - Evaluate our algorithm with two-sided experiments that restore SUTVA
  - Causally demonstrate the significant value of the SBL algorithm to thicken the marketplace (+3.13%) and boost long-term advertising revenue (+5.35%)

1. Introduction and Contribution

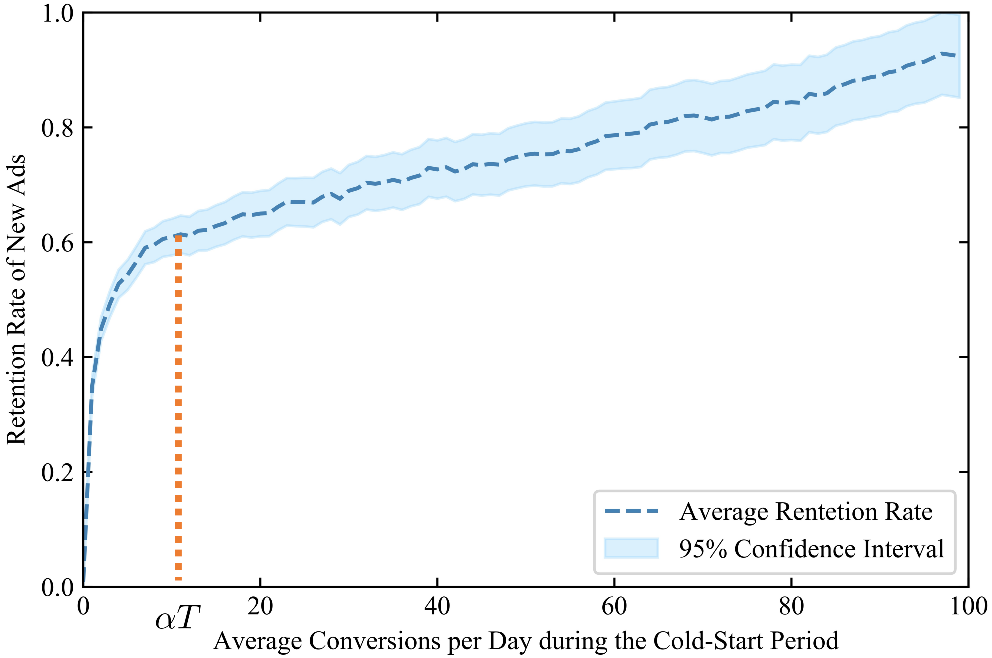
2. **Theory: Model, Algorithm and Analysis**

3. Practice: Field Implementation, Experiment Design and Empirical Results

Model for Cold Start

- Problem setting: First-price auction, CVR=pCVR=1 (oCPC = CPC)
- A set of new ads $A := \{1, 2, \ldots, K\}$, with bids (per conversion) $\{b_1, b_2, \ldots, b_K\}$
- A set of user impressions, arriving at the DSP sequentially: $[T] := \{1, 2, \ldots, T\}$
- Context associated with each user: $x_t \in X$, *i.i.d.* on a countable set X, $x_t \sim \mathcal{D}$
- v_{ij}^t =0,1: Whether a user with context_i clicks ad_j in round t; $y_{tj} = 0, 1$: Whether display ad_j to user_t
- CTR $c_{ij} := \mathbb{E}[v_{ij}^t]$, $\hat{c}_{ij}^t :=$ pCTR estimate by the underlying DNN in round t
- Sequence of events in each round t:

Modeling Cold Start Value



- If the # of conversions in the cold start period is below 10, the long-term retention and value increase linearly.
- If the # of conversions in the cold start period surpasses 10, the long-term retention and value soon flatten.

- The cold start value:

$$\sum_{j=1}^{K} \beta_j \min \left\{ \sum_{t=1}^{T} v_{x_t,j}^{t} y_{tj}, \alpha T \right\}$$

- β_j = the cold start value of ad_j, determined by business sense and simulation
- αT = the threshold for cold start success

Reward Upper Bound and Regret

- Our objective is to identify a policy π to maximize the expected reward $E[\Gamma]$:

$$\Gamma := \underbrace{\sum_{j=1}^{K} b_j \left(\sum_{t=1}^{T} v_{x_t,j}^{t} y_{tj} \right)}_{\text{Short-term revenue}} + \underbrace{\sum_{j=1}^{K} \beta_j \min \left\{ \sum_{t=1}^{T} v_{x_t,j}^{t} y_{tj}, \alpha T \right\}}_{\text{Long-term cold start value}}$$

- π is a non-anticipative (randomized) ad allocation policy.

Lemma (Fluid upper bound). We have the following upper bound for the expected reward:

$$\frac{1}{T} E_{\mathcal{D}^T,\pi}[\Gamma] \leq \text{OPT} := \max_{y_i \in \Delta_A, \forall i} \left\{ \sum_{j=1}^{K} E_{i \sim \mathcal{D}_X}[c_{ij} y_{ij} b_j] + \sum_{j=1}^{K} \beta_j \min \left\{ E_{i \sim \mathcal{D}_X}[c_{ij} y_{ij}], \alpha \right\} \right\}$$

where Δ_A is the probability distribution over all the ads and $y_i = (y_{i1}, y_{i2}, \ldots, y_{iK}) \in \Delta_A$ is the ad assignment distribution for a user with context i.

- The regret of a policy: $\text{Reg}(\pi) := T \cdot \text{OPT} - E_{\mathcal{D}^T,\pi}[\Gamma]$
- We seek to design an algorithm with sublinear regret and implementable on a real DSP.

Empirical Reward and Ad Allocation

- Empirically optimal ad allocation policy at round t:

$$\max_{y_i \in \Delta_A, \forall i} \underbrace{\sum_i \sum_{j \in A} \hat{p}_i^t \overbrace{\hat{c}_{ij}^t b_j}^{\text{eCPM of ad\_j and context i}} y_{ij}}_{\text{Short-term revenue}} + \underbrace{\sum_{j \in A} \beta_j \min \left\{ \sum_i \hat{p}_i^t \hat{c}_{ij}^t y_{ij}, \alpha \right\}}_{\text{Long-term cold start reward}}$$

- Linearize:

$$\max_{y_{ij} \geq 0,\ u_j \geq 0} \sum_i \sum_{j \in A} \hat{p}_i^t \hat{c}_{ij}^t b_j y_{ij} + \sum_{j \in A} \beta_j (\alpha - u_j)$$

$$\text{s.t.} \quad \sum_{j \in A} y_{ij} \leq 1,\ \forall i, \quad \sum_i \hat{p}_i^t \hat{c}_{ij}^t y_{ij} + u_j \geq \alpha,\ \forall j \in A$$

  - u_j is the number of conversions below the threshold.
  - The model has a too high dimension to solve efficiently online.

Duality and Shadow Bidding

$$\min_{\lambda_j \in [0,\beta_j], \forall j} \quad \sum_i \hat{p}_i^t \max_{j \in A} \underbrace{\left\{ \hat{c}_{ij}^t (b_j + \lambda_j) \right\}}_{\text{Adjusted eCPM for ad\_j and context i}} - \alpha \sum_{j \in A} \lambda_j$$

- Only K decision variables: Can be efficiently solved using sub-gradient descent!
- λ_j is the dual variable for the cold start reward constraint: $\sum_i \hat{p}_i^t \hat{c}_{ij}^t y_{ij} + u_j \geq \alpha$
  - We call λ_j the shadow bid of ad_j.
- The shadow bid λ_j is bounded from above by the cold start value of ad_j, β_j
- Actionable Insights:

> - Smartly computing the shadow bid of each new ad and incorporating it into the auctions of the DSP could effectively trade off short-term revenues with long-term cold start rewards!
> - Naturally fit into the ad auction system of a DSP in practice.

Shadow Bidding with Learning (SBL)

Shadow Bidding with Learning (SBL) Algorithm

- Update shadow bids at rounds $\tau_1, \tau_2, \ldots$, with $\tau_{m+1} - \tau_m = \tau_m - \tau_{m-1} = O(T^{\frac{2}{3}})$
- For each round t=1, 2, 3, ⋯, T
  1. Observe the context $x_t = i$. With probability $\epsilon_t = t^{-\frac{1}{3}}(K \log t)^{\frac{1}{3}}$, explore uniformly at random; with probability $1 - \epsilon_t$, display the ad $\mathrm{argmax}_j \hat{c}_{ij}^t (b_j + \lambda_j)$, with an arbitrary tie-breaking rule.
  2. If $t = \tau_m$, solve the empirical dual program to update λ , and update $m \leftarrow m + 1$
  3. Observe the click-through outcome, and update $\hat{c}_{ij}^{t+1}$

Existing approaches in the literature:

- Based on empirical risk minimization oracle.
- Regret benchmarked with the best policy in a policy set, dependent on the policy set size.
- Of theoretical nature, not scalable and implementable on a real DSP.

vs

SBL Algorithm:

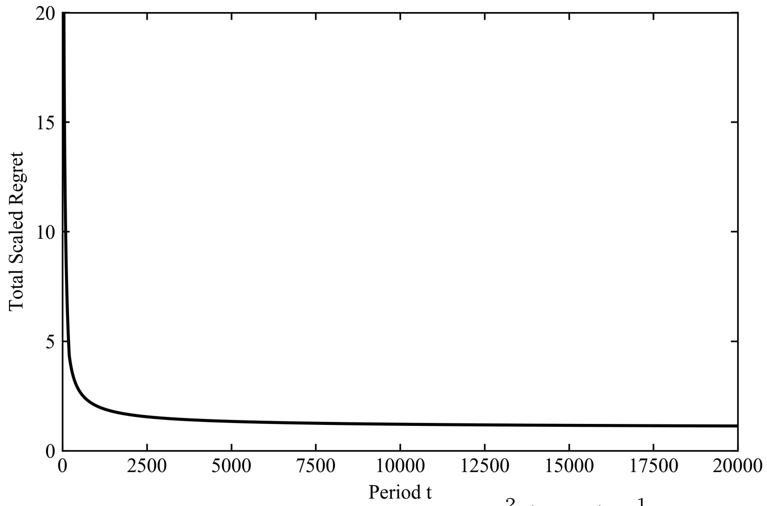
- Dual + MAB (epsilon greedy) + ML Oracle.
  - Optimal primal ad allocation with the dual solution and prediction from the ML Oracle.
- Regret benchmarked with the optimal primal allocation under true CTR.
  - Duality gap + prediction error.
- Implementable on a large-scale DSP in practice with minimal changes to the system.

Theoretical Performance Guarantee of SBL

- Machine learning oracle assumption: With probability at least $1-\delta$, the estimate $\hat{c}_{ij}^t$ satisfies
$$|\hat{c}_{ij}^t - c_{ij}| \leq O\left(\sqrt{\log(1/\delta) d / n_j^t}\right)$$
where d captures the error magnitude of the underlying machine learning oracle to obtain the pCTR $\hat{c}_{ij}^t$, n_j^t is the number of *i.i.d.* impressions for ad_j by round t.

- Satisfied by (i) linear regressions; (ii) regression trees; and (iii) fully connected neural networks.

Theorem (Regret bound). The expected regret of SBL is bounded by $O(T^{\frac{2}{3}} K^{\frac{1}{3}} (\log T)^{\frac{1}{3}} d^{\frac{1}{2}})$.



Total Regret Scaled by $t^{-\frac{2}{3}}(\log t)^{-\frac{1}{3}}$

Remark: The same regret bound holds if the shadow bids are updated via dual-mirror descent (DMD).

Sketched Regret Analysis

- Key challenges:
  - History-dependent cold start reward (i.e., the knapsack bandit setting)
  - Dual-based bidding strategy implemented on the primal space
  - Regret dependent on the underlying machine learning oracle to predict CTR
  - Too high variance with Inverse Propensity Score to estimate the expected reward

- Key ideas and the road map to overcome the challenges:

1. Establish approximate complementary slackness and bound the duality gap between the empirical primal and the empirical dual, due to tie breaking in SBL by $\boxed{O(T^{\frac{1}{3}}(\log T)^{\frac{1}{3}}K^{\frac{1}{3}})}$
2. Build an auxiliary reward process independent of history: Each click of ad j generates a reward of $b_j + \beta_j$, irrespective of whether the threshold αT is met. Under SBL, bound the gap between the auxiliary reward process and the optimal reward by $\boxed{O(T^{\frac{2}{3}}(\log T)^{\frac{1}{3}}K^{\frac{1}{3}}d^{\frac{1}{2}})}$
3. Under SBL, bound the gap between the auxiliary reward process and the true reward process by $\boxed{O(T^{\frac{1}{2}}(\log T)^{\frac{1}{2}}K^{\frac{1}{2}})}$
4. Putting the above bounds together yields the expected regret of SBL is $\boxed{O(T^{\frac{2}{3}}(\log T)^{\frac{1}{3}}K^{\frac{1}{3}}d^{\frac{1}{2}})}$

1. Introduction and Contribution
2. Theory: Model, Algorithm and Analysis
3. Practice: Field Implementation, Experiment Design and Empirical Results
4. Future Research Agenda

Online Implementation of SBL (oSBL)

Online Shadow Bidding with Learning (oSBL) Algorithm

- Update shadow bids at rounds $\tau_1, \tau_2, \ldots$, where $\tau_{m+1} - \tau_m =$ 1 hour; set $\alpha T = 10$, $\beta_j = 2b_j$
- For each round t=1, 2, 3, ⋯, T
  1. Observe the context $x_t = i$. Choose top 150 (new & mature) ads and another 15 randomly picked new ads to join the auction.
  2. Obtain $\hat{c}^t_{ij}$=pCTR*pCVR. Display $\mathrm{argmax}_{j\in[K_t]}\hat{c}^t_{ij}(b_{tj}+\lambda_j)$, where b_{tj} is the system bidding price calculated by the real-time PID system and $[K_t]$ is the set of the 165 ads who join the auction.
  3. If $t = \tau_m$, sample 4% of the auctions in the past hour $\mathcal{H}_t$ and update the shadow bids λ by

$$\min_{\lambda_j\in[0,\beta_j],\forall j\in[K],\lambda_j=0,\forall j\in[K']} \sum_{s\in[\mathcal{H}_t]} \max_{j\in[K_s]} \{\hat{c}^s_{ij}(b_{sj}+\lambda_j)\} - \alpha|\mathcal{H}_t| \sum_{j\in[K_s]} \lambda_j$$

     where [K] is the set of new ads and [K'] is the set of mature ads.

  4. Observe the click-through and conversion outcome, and update $\hat{c}^{t+1}_{ij}$

- Conversions are incorporated into the algorithm (CVR and pCVR are much smaller than 1).
- For a mature ad, the shadow bid is 0.
- Shadow bids are updated every hour based on sampled data, implemented on top of the PID real-time bidding system.
- More like uniform exploration than epsilon-greedy.
- Cold start value is set at twice as much as the target CPA (i.e., bid_convert) of the ad: $\beta_j = 2b_j$

Implementing and Testing SBL

- SBL algorithm implemented on a large-scale video-sharing platform (Platform O).
- How to unbiasedly evaluate the SBL algorithm?
- Naive one-sided experiment designs:
  - Ad-side randomization:

	Treatment New Ads	Control New Ads	Non-Experiment New Ads	Mature Ads
100% UV	Treatment Condition	Control Condition		

  - User-side randomization:

	100% New Ads	Mature Ads
Treatment UV	Treatment Condition	
Control UV	Control Condition	
Non-Experiment UV		

Treatment = oSBL algorithm; Control = baseline algorithm (uniformly increase the bidding price of all new ads)

Violation of SUTVA

- Ad-side randomization:

	Treatment New Ads	Control New Ads	Non-Experiment New Ads	Mature Ads
100% UV	Treatment Condition	Control Condition		

- User-side randomization:

	100% New Ads	Mature Ads
Treatment UV	Treatment Condition	
Control UV	Control Condition	
Non-Experiment UV		

- SUTVA (Stable unit treatment valuation assumption): The assignment of one unit to treatment or control will not affect the outcome of another unit.

- Ad-side randomization: 120% overestimate
  - Violation of SUTVA: New ads from different groups compete on the same user impressions.

- User-side randomization: 40% underestimate
  - Violation of SUTVA: The effect of oSBL spills over to the control group users.

上海纽约大学 NYU SHANGHAI

Two-Sided Experiment Design

CUHK BUSINESS SCHOOL

- A novel two-sided experiment design:

	20% Treatment New Ads	20% Control New Ads	60% Non-Experiment New Ads	Mature Ads
33% Treatment UV	B11	B12	B13	B14
33% Control UV	B21	B22	B23	B24
33% Non-Experiment UV	B31	B32	B33	B34

  - Blue = oSBL, white = baseline algorithm, grey = ad blocked (to remove externalities)
  - SUTVA restored! Able to generate unbiased estimates.
  - Impact on long-term cold start value: Comparing B11 with B22
  - Impact on short-term revenue: Comparing B11+B13+B14 with B22+B23+B24
  - Experiments conducted between May 23, 2020, and May 30, 2020

- A simulation system for cold start and experimentation on online advertising platforms open-sourced @GitHub: https://github.com/zikunye2/cold_start_to_improve_market_thickness_simulation

Short- and Long-term Effects of oSBL

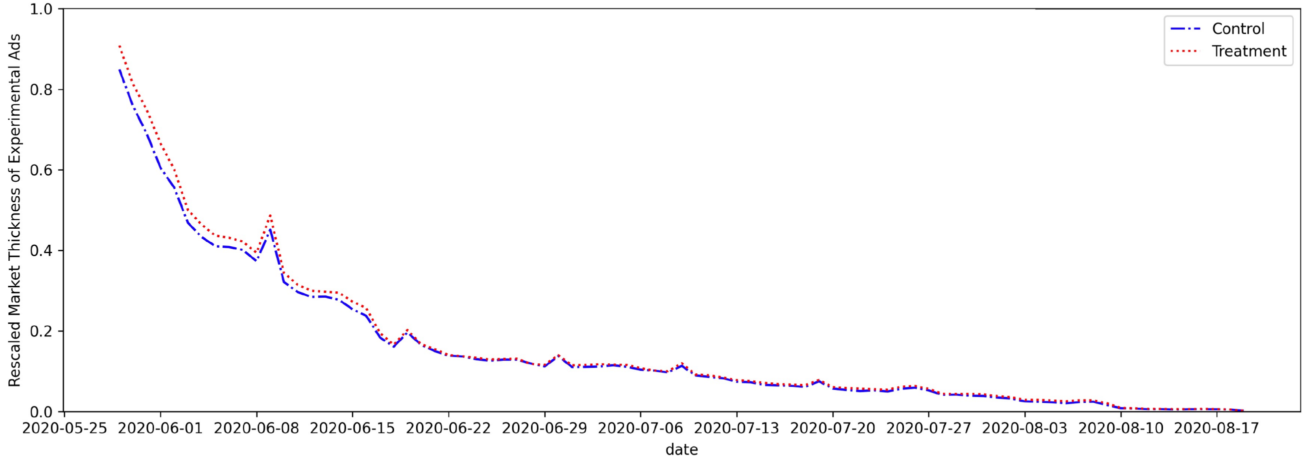
- Short-term effect of oSBL:

Metric of Interest	Cold start success rate	Cold start reward	Total short-term revenue	CTR Prediction AUC
Relative Change	+61.62%***	+47.71%***	-0.717%**	+7.48%*

- Long-term (post-cold start and post-experiment) effect of oSBL:

Metric of Interest	Market Thickness	CTR	Post-Cold-Start Revenue
Relative Change	+3.13%**	+11.14%***	+34.02%***

Note: The above long-term revenue boost is over-estimated because the treated ads compete with all others on impressions,



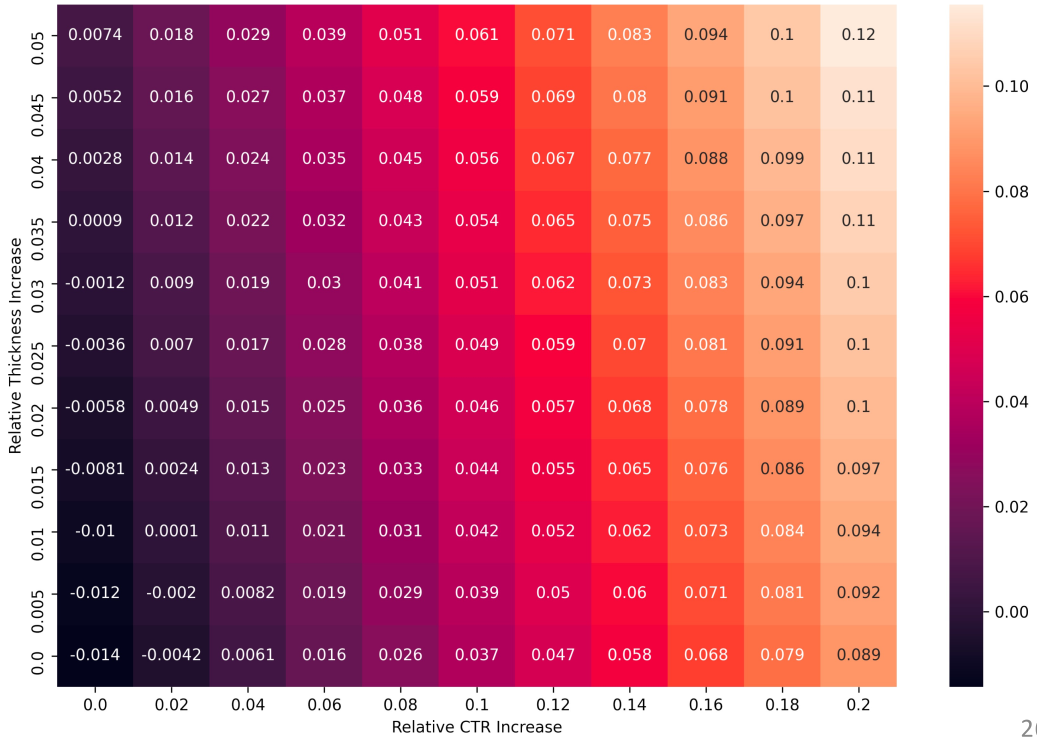
Simulation Results on Long-Term Ad Revenue

- Long-term (post-cold start and post-experiment) effect of oSBL:

Metric of Interest	Market Thickness	CTR	Post-Cold-Start Revenue
Relative Change	+3.13%**	+11.14%***	+34.02%***

Note: The above long-term revenue boost is over-estimated because the treated ads compete with all others on impressions,

- Simulation study on the global treatment effect of oSBL:
  - Based on 12 million auctions sampled from April 9, 2020, to April 30, 2020.
  - oSBL increases the market thickness by +0%~+5%, and the post-cold-start CTR by 0%~20% (i.e., sensitivity analysis).
  - The total number of ad impressions/auctions remain the same.
  - Simulation model validated by accurately predicting the short-term revenue loss during the experiment.
- oSBL boosts the total long-term revenue by +5.35% (at a magnitude of hundred million USD per year for Platform O) if the market thickness is increased by 3.13% and CTR by 11.46%.



Takeaways

- SBL Algorithm: A smart algorithm to connect bandit learning theory and the ad cold start practice.

- Two-sided experiment: Causally estimate the value of SBL for a large-scale online advertising platform (substantial ad revenue improvement for the platform).

- SBL and two-sided experiment have the potential to optimize and evaluate more general recommender systems of online two-sided platforms.

Link to the paper: https://rphilipzhang.github.io/rphilipzhang/Cold_Start_unblinded.pdf

Thank You!

Questions?

renyu.zhang@nyu.edu and philipzhang@cuhk.edu.hk

https://rphilipzhang.github.io/rphilipzhang/index.html