Optimal Bidding and Experimentation for Multi-layer Auctions in Online Advertising

Nian Si*[1], San Gultekin†[2], Jose Blanchet‡[1], and Aaron Flores§[2]

[1]Department of Management Science & Engineering, Stanford University
[2]Yahoo!

Friday 7th October, 2022

Abstract

The digital advertising industry heavily relies on online auctions, which are mostly of first-price type. For first-price auctions, the success of a good bidding algorithm crucially relies on accurately estimating the highest bid distribution based on historical data which is often censored. In practice, a sequence of first-price auctions often takes place through multiple layers, a feature that has been ignored in the literature on data-driven optimal bidding strategies. In this paper, we introduce a two-step algorithmic procedure specifically for this multi-layer first-price auction structure. Furthermore, to examine the performance of our procedure, we develop a novel inference scheme for A/B testing under budget interference (an experimental design which is also often used in practice). Our inference methodology uses a weighted local linear regression estimation to control for the interference incurred by the amount of spending in control/test groups given the budget constraint. Our bidding algorithm and the new testing methodology have been deployed in a major demand-side platform in the United States. Moreover, in such an industrial environment, our tests show that our bidding algorithm outperforms the benchmark algorithm by 0.5% to 1.5%.

1 Introduction

As the main business engine of the current digital economy, the online advertising industry has grown exponentially fast in the past few decades. According to Statista [43], the total value of the global internet advertising market was worth USD 566 billion in 2020 and is expected to reach USD 700 billion by 2025. Along with a fast paced growth, the industry keeps evolving in complexity and sophistication.

*niansi@stanford.edu
†sgultekin@yahooinc.com
‡jose.blanchet@stanford.edu
§aaron.flores@yahooinc.com

Online ads are sold via real-time auctions. Just a few years ago, such auctions used to be largely second-price-type auctions, where the optimal strategy (for single auctions) is to bid truthfully. However, in recent years, the industry has shifted from second-price auctions to first-price auctions [20, 21]. This shift can be attributed to multiple reasons, including the demand for increased transparency [9, 18, 41]; the introduction of so-called header bidding auctions [26], and yield concerns [1, 40, 32]. Unlike in second-price auctions, in first-price auctions, as we discuss below, bidders will no longer bid truthfully. This phenomenon imposes a challenge for bidders to accurately estimate the distribution of the highest competing bid. Naturally, the information used to estimate such distribution depends crucially on what is shared by the different platforms involved in the execution of such high-throughput auctions.

Advertisers rely on demand-side platforms (DSPs) to manage their advertising campaigns. On behalf of advertisers, DSPs submit bids to an auction in online exchanges. If the bid is the highest bid, the DSP wins the opportunity to display their ads. Otherwise, the DSP loses. In a first-price auction, the winners need to pay the full amount they bid. Therefore, the success of a good bidding algorithm crucially relies on accurately estimating the distribution of the highest competition bid in order to bid as low as possible while still wining the auction. Predicting the highest competing bid distribution is a standard machine learning task, where the historical data consists of each bidding request's characteristics, the historical bid price, and a binary label on whether the DSP won/lost the particular request. The goal is to train a model that predicts the highest competing bid distribution given the request's characteristics. Often, besides the binary label, the exchanges also provide the information on the highest competing bid. Naturally, it is conceivable that a more accurate prediction could be obtained by utilizing the additional information.

In current practice, online auctions exhibit a more complicated structure. There are multiple layers of first-price auctions for a given ad request because multiple exchanges may participate in the same ad request. After DSPs submit bids to an exchange, which executes a first level first-price auction, the exchange may further submit a winning bid to a so-called header auction (i.e. a final layer first-price auction). Among all the bids submitted by the exchanges participating, the bid that wins the final auction (i.e. the header auction) wins the opportunity. In fact, there may be multiple layers leading to the header auction. In this setting, the information structure may involve the final binary feedback on whether the DSP won/lost the request in the header auction, and the highest competing bid from the exchange the DSP submits to in the first layer of the auction. However, the DSP will typically not know the competing bid submitted by the other exchanges to the header auction. Therefore, the highest competing bid received by the DSPs may not be the highest competing bid among all competitors but could be viewed only as a lower bound. Thus, if DSPs ignore the multi-layer structure and directly use the historically available data for bidding estimation, they will tend to underestimate the optimal bidding strategy. Please see Figure 1 for the flow chart and information structure of this type of multi-layer auctions.

The first part of the paper studies the bidding problem under this multi-layer scenario. We propose a two-step bidding algorithm, in which we estimate the highest competing bid landscape in both layers of auctions using the highest competing bid and binary feedback information, respectively.

To examine the performance of the two-step algorithm, we rely on online A/B test [34]. Note that the performance evaluation could not be conducted fully offline since we may not know

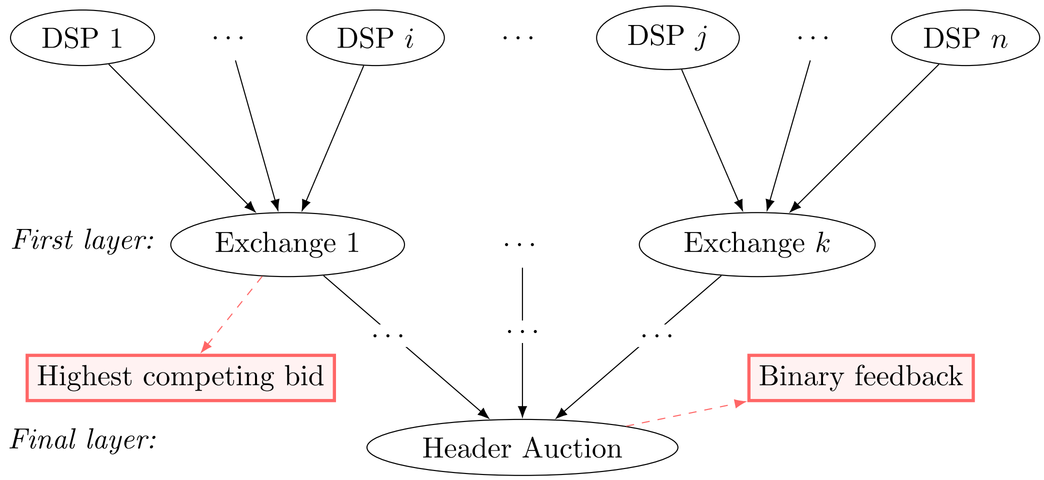


Figure 1: Flow chart and information structure of the multi-layer auction pipeline

whether we win or lose if the bid price suggested by the algorithm is different from the historical bid. In the A/B testing procedure, we allocate half of the online request traffic to the new bidding algorithm and allocate another half to the benchmark algorithm for each ad campaign. Then, we compute and compare a performance metric for the test and control algorithms to obtain the average treatment effect, where performance is measured by cost per click (CPC), cost per action (CPA), cost per impression (CPM), or cost per install (CPI) depending on the campaign goal specified by the advertisers. (CPC means the average cost avertisers pay for the ads based on the number of clicks the ad receives; CPA is the cost divided by the number of actions being measured; CPM is the cost per 1,000 views of a particular ad; and CPI is the total spend divided by the number of new installs.) To be concise, we write CPX to represent those campaign goals and the algorithm with lower CPX is better. If there is no interference between the test and control groups, a direct comparison of CPXs will yield an unbiased estimation of the cost reductions and paired t-test could give valid inferences. However, as the test and control algorithms are bidding differently and compete for the same campaign budget simultaneously, the actual amounts of spending in the treatment and control groups could be unbalanced. This interference could result in a non meaningful comparison as CPX is significantly affected by the total spending in a treatment group. Higher spending usually results in higher CPX, as the marginal return of spending is diminishing. Therefore, if the CPX in the test group is lower than the CPX in the control group while the test group spends less than the control, the test algorithm's effectiveness is inconclusive. For example, if the bidding algorithm always submits 1 cent per thousand impressions to the exchange, the CPX could be extremely low but also with an extremely low win-rate, which generally will not be desirable. Therefore, a fair comparison of CPXs should be based on an appropriately balanced spending in the test and control groups.

To tackle this issue, Liu et al. [36] propose a novel budget-split design. They replicate two identical campaigns with the same characteristics and each having half of the original campaign budget. Then, test and control algorithms run on those two identical campaigns with separate traffic, respectively. In principle, this design is able to solve the aforementioned interference problems caused by the unbalanced spending. However, in practice, it significantly increases engineering overhead, as it duplicates every campaign. Therefore, this design is costly and increases the complexity in implementation.

In the second part of this paper, we propose a novel statistical testing method via weighted local linear regression [14]. This method yields accurate estimates of the treatment effect and provides valid inferences without changing the procedure of the A/B test.

We summarize our contributions as follows:

(1) We model the structure of multi-layer auctions. We propose a two-step offline training algorithm and we demonstrate that our model could be efficiently optimized online in real-time bidding.

(2) We propose a novel statistical testing method for A/B testing under spending and budget interference. We quantify the bias and variance of our estimator and we establish a central limit theorem. We also discuss the choice of key parameters in the inference procedure.

(3) We implement the new bidding algorithm and new testing methodologies in a major DSP in the United States. We demonstrate the effectiveness of the new bidding algorithm based on various metrics. In particular, our testing method suggests that our bidding algorithm achieves 0.3% to 1.5% cost reductions for different goal types.

The remainder of the paper is structured as follows. Section 2 presents preliminaries on online ad auctions and bidding models. In Section 3 we discuss our two-step procedure for optimal bidding. In Section 4, we provide the new statistical methodology for A/B testing with budget interference. The experimental results based on real data are detailed in Section 5. Finally, we conclude with our findings in Section 6. All technical proofs are provided in Appendix Appendix A.

1.1 Related Work

The literature on online ad auctions is rapidly growing. Wu et al. [48] study the winning price prediction problem in a second-price auction to determine the cost of each impression, where the highest competing bid is observed only if the DSP wins. Wu et al. [47] extend their work to a neural-network based method. For the first-price auction, Gligorijevic et al. [19] use factorization machines to estimate the highest competing bid in a non-censorship setting (the historical highest competing bids are observable). Pan et al. [38] propose a logistic model on the setting that only binary feedback is received, while Zhou et al. [51] extend their method to the use of neural networks for estimating the parameters of the highest competing bid distribution. More recently, a non-parametric bidding algorithm is proposed in Zhang et al. [50]. Besides those offline estimation methods, Karlsson and Sang [31] directly optimize the surplus function in the first-price auction using online learning; He et al. [24] and Wu et al. [46] use reinforcement learning methods [44] to optimally bid under a budget constraint; Han et al. [23] use bandit algorithms to bid in adversarial first-price auctions; Baardman et al. [2] also model the bidding problem as a multi-armed bandit with periodic budgets; and Balseiro and Gur [4] model the repeated auctions with budgets as a sequential game of incomplete information. However, to the best of knowledge, we are not aware of any work that discusses the multi-layer structures of online ad auctions.

Online experimentation is a topic of active investigation [45, 33, 10, 49, 34]. In particular, one main challenge in online experimentation is the interference [27]. Interference refers to the violation of the Standard Unit Treatment Value Assumption (SUTVA) that guarantees unbiased estimators [28]. One of the most well-known interference structures involves so-called network effects [42, 11]. Another example of interference arises in online marketplaces, where there is competition and spillover among test and control units [7, 25, 3, 29, 35]. Existing approaches which can be used to overcome the bias caused by interference include experimental design-based methods [36, 6, 8, 29] and model-based correction [7]. The design-based methods usually incur high engineering overhead (e.g. modifying an existing experimental platform), resulting in implementation challenges which are difficult to overcome in practice. Model-based corrections rely on model assumptions, which may be unverifiable in practice. In the context of experiments in online auctions, Basse et al. [6] test different auction formats on the exchange side. In contrast, we test different bidding algorithms on the DSP side. Therefore, the interference structure is different. In this paper, we provide a simple testing method for testing bidding algorithms in DSPs which does not rely on parametric assumptions and does not alter the design of a testing platform which allocates treatment and control to a single stream of bid requests.

2 Preliminaries on Auctions and Bidding Models

We provide an overview of the biding problem in online advertising auctions. We focus on first-price auctions, in which the highest-price bidder wins the auction and pays the exact amount of his bidding price. Specifically, for a specific impression opportunity being auctioned, we denote the bidder's value as U, the bidder's bidding price as b, and the highest competing bid is V, which is a random variable. Let $F_V(\cdot)$ denote the CDF of the random variable V and we assume V has continuous density $f_V(\cdot)$. Although the presence of campaign budget may alter DSPs' decisions in the repeated auctions [5], we only focus on the one-time bidding problem in our algorithm as bidding under budget constraints and the highest competing bid forecasting are usually decoupled in practice [30]. Therefore, as part of the campaign deliver optimization strategy, DSPs have the goal to maximize the surplus, where the optimization problem is defined as

$$\max_b S(b) := (U - b)\, F_V(b). \tag{1}$$

$S(b)$ denotes the expected surplus with the bid b and the bidder's value U is estimated beforehand. So, we assume that U is given in this paper.

From the optimization problem (1), it is clear that (i) the bid price should not exceed U, and (ii) the outcome depends on the distribution of the highest competing bid V. If we know the distribution of V, Problem (1) becomes an one-dimensional optimization problem, which is usually not hard to solve. However, in practice, the distribution of V is unknown. Fortunately, there is sufficient historical data from the past auctions which can be used to infer the unknown distribution.

In our setting, we are given a set of publisher contextual features of the ad placement opportunity being auctioned, such as the device type, website or app, its position, geometry, local time of day, etc. We assume that the features $X \in \mathbb{R}^d$ are a d-dimensional vector. The key task in Problem (1) becomes a machine learning problem, in which we need to predict the conditional distribution of V given observed feature $X = x, U = u$.

Based on the information structure, there are at least two different approaches to address this machine learning task. First, if we only receive binary feedback, i.e., the information whether we win or lose the impression opportunity, we may use a parametric model to estimate the conditional

CDF $F_{V|X,U}(\cdot)$. For example, we may consider a model of the form

$$F_{V|X,U}(b) = F(w^\top \phi(x,u) + \beta g(b)) \tag{2}$$

where F is a chosen cumulative distribution function (CDF) of a given random variable; w are the weights; g is a given bid transformation function; x, u are an observed feature and the bidder's value; ϕ is a feature mapping, and β is the bid coefficient. We estimate the parameters w and β by maximum likelihood estimation (MLE). Namely, for the historical data $\{X_i, I_i, U_i, b_i\}_{i=1}^n$, where $I_i \in \{0,1\}$ represents the binary feedback, we solve

$$\max_{w,\beta} \frac{1}{n} \sum_{i=1}^n I_i \log F(w^\top \phi(X_i,U_i) + \beta g(b_i)) + (1 - I_i) \log \left(1 - F(w^\top \phi(X_i,U_i) + \beta g(b_i))\right). \tag{3}$$

The feature mapping function ϕ could be chosen as the field-weighted factorization machine (FwFM) [37], which captures feature interactions. Moreover, ϕ could be also be parametrized by a neural network as ϕ_θ; see Zhou et al. [51]. A common choice of g and F are $g(b) = \log b$ and the logistic function $F(x) = (1 + e^{-x})^{-1}$ [38]. Under these choices, if we let $\alpha = w^\top \phi(X_i, U_i)$, we have $F_{V|X,U}(\cdot)$ is the CDF of the Loglogistic(α, β) distribution.

A second approach, available if the highest competing bid is provided in the historical data, allows to obtain a more accurate estimation of the distribution of V. Currently, Google's ad exchange (AdX) provides the minimum bid to win values, for which we can directly model the density. In such a case, we consider $f_{V|X,U}(v) = f(v|h_\theta(x,u))$ where $f(\cdot|h_\theta(x,u))$ is a probability density function. For example, for a log-normal model, we let $h_\theta(x,u) = [\mu_\theta(x,u), \sigma_\theta(x,u)]$ and we have

$$f_{V|X,U}(v) = \frac{1}{v\sigma_\theta(x,u)\sqrt{2\pi}} \exp\left(-\frac{(\ln(v) - \mu_\theta(x,u))^2}{2\sigma_\theta^2(x,u)}\right).$$

Then, via maximum likelihood estimation, for the collection of historical data $\{X_i, V_i, U_i\}_{i=1}^n$, we solve

$$\max_\theta \frac{1}{n} \sum_{i=1}^n \log\left(f\left(V_i | h_\theta\left(X_i, U_i\right)\right)\right). \tag{4}$$

The aforementioned two approaches are corresponding to only one-layer auctions with either binary feedback or highest competition bid information. Nevertheless, the focus of this paper is the multi-layer auctions, where both types of information are provided in different stages of multi-layer auctions as depicted in Figure 1. Therefore, the highest competing bid receiving in the first layer of the auctions is only a lower bound of the true highest competing bid. It is important to note that this type of structure prevails in practice. We will present a method that combines the aforementioned two approaches to obtain better estimation of the distribution V in Section 3. Then, we discuss a rigorous inference method to test the effects our new algorithm under A/B testing with budget interference in Section 4.

3 The Two-step Algorithm

In the previous section, we discussed two different machine learning approaches under two different information structures. In this section, we consider the auction and information structures depicted in Figure 1 and combine the merits of both approaches. We will first discuss the offline training and then discuss the online optimization.

3.1 Offline Training

By following the pipeline depicted in Figure 1, suppose we are DSP 1 and submit bids to Exchange 1 in the first layer of auctions. We denote the highest competing bid in Exchange 1 to be $V^{(1)}$, and the highest competing bid among all of other exchanges to be $V^{(2)}$. Then, the highest competing bid in the header auction is $V = \max\left\{V^{(1)}, V^{(2)}\right\}$. The optimal bidding problem (1) is rewritten as

$$\max_b (U-b)\mathbb{P}[b \geq \max\{V^{(1)}, V^{(2)}\}|X,U].$$

For this type of multi-layer auctions, after we submit the bid b, we know the realizations of $V^{(1)}$ and $\mathbb{I}\{b \geq \max\{V^{(1)}, V^{(2)}\}\}$, where $\mathbb{I}\{\cdot\}$ denotes the indicator function. The price $V^{(1)}$ corresponds to the known minimum bid to win in the first layer of the auctions; while the price $V^{(2)}$ is the minimum bid to win in all of other exchanges, which is unknown. Therefore, we have access to the historical data $\left\{X_i, U_i, b_i, V_i^{(1)}, I_i\right\}_{i=1}^n$, where $I_i = \mathbb{I}\left\{b_i \geq \max\left\{V_i^{(1)}, V_i^{(2)}\right\}\right\}$. We further assume the historical bid price b_i is a deterministic function of X_i and U_i.

Assumption 1 below is our key assumption, which states that the highest competing bids in Exchange 1 and other Exchanges are independent conditional on the user and publisher features. Basically, it means that DSPs participating in Exchange 1 are similar to other DSPs that participate in other Exchanges.

Assumption 1. *$V^{(1)}$ and $V^{(2)}$ are independent conditional on the set of features, i.e.,*

$$V^{(1)} \perp\!\!\!\perp V^{(2)}|X,U.$$

Assumption 1 also implies that $\mathbb{I}\left\{b \geq V^{(1)}\right\} \perp\!\!\!\perp \mathbb{I}\left\{b \geq V^{(2)}\right\}|X,U$, which yields

$$\mathbb{P}[b \geq \max\{V^{(1)}, V^{(2)}\}|X,U] = \mathbb{P}[b \geq V^{(1)}|X,U]\mathbb{P}[b \geq V^{(2)}|X,U].$$

We denote $F_{V^{(1)}|X,U}(\cdot)$ and $F_{V^{(2)}|X,U}(\cdot)$ to be the conditional CDF of $V^{(1)}$ and $V^{(2)}$, respectively. Then, the optimal bidding problem becomes

$$\max_b (U-b)\left(F_{V^{(1)}|X,U}(b)\right)\left(F_{V^{(2)}|X,U}(b)\right). \tag{5}$$

We next discuss the estimation of $F_{V^{(1)}|X,U}(\cdot)$ and $F_{V^{(2)}|X,U}(\cdot)$, respectively.

(1) Estimation of the conditional CDF $F_{V^{(1)}|X,U}(\cdot)$: We adopt the second approach in Section 2, where we input the data $\left\{X_i, U_i, V_i^{(1)}\right\}_{i=1}^n$ and solve the optimization problem (4). After doing that, we obtain an estimated density $f\left(\cdot \mid g\left(\hat{\theta}^{(1)}, X, U\right)\right)$, which further gives an estimated $\hat{F}_{V^{(1)}|X,U}(\cdot)$.

(2) Estimation of the conditional CDF $F_{V^{(2)}|X,U}(\cdot)$: To do this, we rely on the following lemma.

Lemma 1. *Under Assumption 1, we have*

$$\mathbb{E}[\mathbb{I}\{b \geq \max\{V^{(1)}, V^{(2)}\}|b \geq V^{(1)}, X, U] = \mathbb{E}[\mathbb{I}\{b \geq V^{(2)}\}|X,U]$$

Proof. Proof of Lemma 1 Note that the event $\mathbb{I}\{b \geq \max\{V^{(1)}, V^{(2)}\}\} = \mathbb{I}\{b \geq V^{(1)}, b \geq V^{(2)}\}$. Thus, we have

$$\begin{aligned}
\mathbb{E}\left[\mathbb{I}\{b \geq \max\{V^{(1)}, V^{(2)}\}|b \geq V^{(1)}, X, U\right] &= \mathbb{E}\left[\mathbb{I}\{b \geq V^{(1)}, b \geq V^{(2)}\}|b \geq V^{(1)}, X, U\right] \\
&= \mathbb{E}\left[\mathbb{I}\{b \geq V^{(2)}\}|b \geq V^{(1)}, X, U\right].
\end{aligned}$$

Then, by the conditional independence of $\mathbb{I}\left\{b \geq V^{(1)}\right\}$ and $\mathbb{I}\left\{b \geq V^{(2)}\right\}$, we have the desired result. ☐

Lemma 1 enables us to estimate the distribution of $V^{(2)}|X, U$ using only the data that $b \geq V^{(1)}$. Therefore, we solve the MLE problem (3) on the data $\{X_i, U_i, b_i, I_i\}_{b_i \geq V_i^{(1)}, i \in [n]}$, where $[n] = \{1, 2, \ldots, n\}$. We summarize the full training procedure in Algorithm 1.

Algorithm 1 The offline training procedure in the two-step algorithm

1: **Input:** $\left\{X_i, U_i, b_i, V_i^{(1)}, I_i\right\}_{i=1}^n$.
2: **Output:** An estimated conditional CDF $\hat{F}_{V|X,U}(\cdot)$.
3: Obtain $\hat{F}_{V^{(1)}|X,U}(\cdot)$ by solving the MLE problem (4) using the data $\left\{X_i, U_i, V_i^{(1)}\right\}_{i=1}^n$.
4: Obtain $\hat{F}_{V^{(2)}|X,U}(\cdot)$ by solving the MLE problem (3) using the data $\{X_i, U_i, b_i, I_i\}_{b_i \geq V_i^{(1)}, i \in [n]}$.
5: **return** $\hat{F}_{V|X,U}(\cdot) = \hat{F}_{V^{(1)}|X,U}(\cdot)\hat{F}_{V^{(2)}|X,U}(\cdot)$.

3.2 Online Optimization

At the online serving period, for each request with features X and estimated value U, we solve the following optimization problem

$$\max_b S(b) = (U - b)\left(\hat{F}_{V^{(1)}|X,U}(b)\right)\left(\hat{F}_{V^{(2)}|X,U}(b)\right). \tag{6}$$

The above optimization problem needs to be solved sufficiently fast since the whole bidding process should be completed in less than a second [22]. Fortunately, S(b) is quasi-concave for certain CDFs as shown in Theorem 1.

Theorem 1. *Suppose there are m CDFs $F_1, F_2, \ldots, F_m$ with continuous differentiable densities $f_1, f_2, \ldots, f_m$. We assume that $b + F_i(b)/f_i(b)$ is non-decreasing. Then, the function*

$$S(b) = (u - b)\prod_{i=1}^m F_i(b),$$

is maximized at some unique $b^* > 0$ and $f(b)$ is non-decreasing in $(0, b^*)$ and non-increasing (b^*, u).

The corollary below provides sufficient conditions to ensure the validity of Theorem 1.

Corollary 1. *Any distributions with a log-concave CDF $F(\cdot)$ and a continuous density $f(\cdot)$ satisfies that $F(b)/f(b)$ is non-decreasing. Further, if $\tilde{F}$ has log-concave CDF, $F(b) = \tilde{F}(\log(b))$ also has log-concave CDF .*

Many commonly-used unimodal distributions have log-concave CDFs, including normal, exponential, uniform, logistic, loglogistic, chi-square, gamma, Weibull, log-normal, extreme value, and Pareto distributions. And some distribution with non-log-concave CDFs also satisfy the assumption in Theorem 1; e.g., Student's t-distributions.

Lemma 2. *Student's t-distributions with degree of freedom $\upsilon \geq 1$ satisfy the assumption that $b + F(b)/f(b)$ is increasing.*

The quasi-concavity enables the efficient optimization of Problem (6). We could implement the same bisection method as Pan et al. [38, Algorithm 4.1] with the starting minimum and maximum bounds 0 and u. However, the interval $[0, u]$ may be sometimes too wide. We explain in Proposition 1 how to subsequently reduce the size of interval at the beginning.

Proposition 1. *Consider $F_i = \tilde{F}_i(\log(b))$ for some distributions $\tilde{F}_i(\cdot)$ with log-concave CDF. Then, we have the optimal solution b^* satisfies*

$$b^* \in [\underline{b}^{(j)}, \bar{b}^{(j)}], \ \textit{for}\ j = 0, 1, 2, \ldots,$$

where $\underline{b}^{(0)} = 0, \bar{b}^{(0)} = u$, and

$$\underline{b}^{(j)} = \Psi\left(\bar{b}^{(j-1)}\right), \ \textit{and}\ \bar{b}^{(j)} = \Psi\left(\underline{b}^{(j)}\right), \ \textit{for}\ j = 1, 2, \ldots,$$

for $\Psi(b)$ defined as

$$\Psi(b) := u\left(\sum_{j=1}^{m} \frac{\tilde{f}_j\left(\log(b)\right)}{\tilde{F}_j\left(\log(b)\right)}\right) \Bigg/ \left(\sum_{j=1}^{m} \frac{\tilde{f}_j\left(\log(b)\right)}{\tilde{F}_j\left(\log(b)\right)} + 1\right).$$

In particular, for the case $m = 2$ with $V^{(1)} \sim \text{Lognormal}(\mu, \sigma^2)$ and $V^{(2)} \sim \text{Loglogistic}(\alpha, \beta)$, we have

$$\Psi(x) = \frac{\left(\frac{1}{\sigma}\phi\left(\frac{\log(x)-\mu}{\sigma}\right)/\Phi\left(\frac{\log(x)-\mu}{\sigma}\right) + \beta\left(1 + x^{\beta}\exp(\alpha)\right)^{-1}\right)u}{\frac{1}{\sigma}\phi\left(\frac{\log(x)-\mu}{\sigma}\right)/\Phi\left(\frac{\log(x)-\mu}{\sigma}\right) + \beta\left(1 + x^{\beta}\exp(\alpha)\right)^{-1} + 1},$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the density and CDF of the standard normal distribution. Inspired by Proposition 1, we propose Algorithm 2 below.

Algorithm 2 Iterative scheme to solve $f(b)$ under the assumption from Proposition 1

1: **Input:** u, the function $\Psi(\cdot)$, and the iteration times T.
2: **Output:** An lower bound and upper bound for b^*.
3: $b_{\min} = 0$ and $b_{\max} = u$;
4: **for** $i \leftarrow 1$ to m **do**
5: $\quad b_{\min} \leftarrow \Psi(b_{\max})$;
6: $\quad b_{\max} \leftarrow \Psi(b_{\min})$;
7: **end for**
8: **return** $b_{\min}$ and $b_{\max}$.

In practice, we usually iterate $T = 2$ or 3 times in Algorithm 2 to obtain a sufficient narrow interval. Then, we apply the bisection method and we find an ϵ-maximizer in at most $\log_2\lceil(b_{\max} - b_{\min})/\epsilon\rceil$ steps.

4 Statistical Test

In this section, we develop a statistically valid test for estimating the effects of a new algorithm using a suitably designed A/B test which is standard in industry. We first describe the A/B test pipeline deployed in the DSP in Section 4.1 and propose our test method via weighted local linear regression in Section 4.2.

4.1 A/B Testing Pipeline

Suppose there are N campaigns. In each campaign, there are two types of treatments: test and control algorithms, and the DSP allocates half of the online traffic (online request) to the control group and another half of the traffic to the test group. In the control group, we run a benchmark bidding algorithm, while in the test group, we run a new test algorithm; for example, the new test algorithm could be the two-step algorithm developed in Section 3. With a slight abuse of notation, we let X_i be the i-th campaign's characteristics and we denote the total budget of this campaign to be B_i. Then, we denote the total amounts spent in the test and control group as S_i^{test}, S_i^{control}. Therefore, we have $S_i^{\text{test}} + S_i^{\text{control}} \leq B_i$. There is an online control algorithm that dynamically adjusts the bids to meet the requirement of exhausting the full budget B within the user-predefined campaign window and appropriately balance the test and control spending, see details in Karlsson [30]. Note that the user-predefined campaign window is usually different from the A/B testing timeframe, so the total spending within the A/B test $S_i^{\text{test}} + S_i^{\text{control}}$ is generally not equal to B_i. However, it is reasonable to assume $S_i^{\text{test}} + S_i^{\text{control}}$ is proportional to the budget B_i as $(S_i^{\text{test}} + S_i^{\text{control}})/B_i$should be approximately the ratio of the time windows of the A/B test to the campaign duration. We further denote the total rewards (e.g., clicks) on the test and control group as R_i^{test}, R_i^{control}. Then the CPX (e.g., cost per click) is

$$\text{CPX}_i^{\text{test}} = S_i^{\text{test}}/R_i^{\text{test}} \text{ and } \text{CPX}_i^{\text{control}} = S_i^{\text{control}}/R_i^{\text{control}}.$$

Thus for an online A/B test, we have i.i.d. data $\left\{\text{CPX}_i^{\text{control}}, \text{CPX}_i^{\text{test}}, S_i^{\text{test}}, S_i^{\text{control}}, B_i, X_i\right\}_{i=1}^{N}$. The CPX is our target metric and we need to test whether CPX^{test} is indeed smaller than $\text{CPX}^{\text{control}}$. Therefore, we introduce the individual treatment effect as an expected log-CPX reduction for each campaign in the following definition.

Definition 1 (Individual treatment effect)**.** The individual treatment effect for the campaign with characteristics x and Budget B is defined as

$$\begin{aligned}\Delta(x) := &\mathbb{E}\left[\log\left(\text{CPX}^{\text{test}}\right)\middle| S^{\text{test}} = B, S^{\text{control}} = 0, X = x\right] \\ &- \mathbb{E}\left[\log\left(\text{CPX}^{\text{control}}\right)\middle| S^{\text{test}} = 0, S^{\text{control}} = B, X = x\right].\end{aligned}$$

However, direct comparison between CPXs in the A/B test is interfered by the spending. As we discussed in Introduction, if we spend a higher amount on a particular treatment group, we are likely to get a higher win rate, which in turn yields more rewards. However, due to diminishing returns of the spending, we also incur higher CPX for such a group. Furthermore, the control and test groups are also competing for the same budget with each other. For example, we consider that the test group has intrinsic better performance with a lower CPX than the control group. In such a case, the spending may shift to the test group. Then, the training group only needs to spend less and thus has lower CPX, which further biases the treatment effect estimation. To better facilitate our settings, we define several necessary auxiliary functions.

Definition 2. We define the test and control effect functions as

$$\begin{aligned}
\varphi^{\text{test}}\left(\log\left(\frac{S^{\text{test}}}{S^{\text{control}}}\right); S^{\text{test}}, X\right) &:= \mathbb{E}\left[\log\left(\text{CPX}^{\text{test}}\right)|S^{\text{test}}, S^{\text{control}}, X\right] \text{ and} \\
\varphi^{\text{control}}\left(\log\left(\frac{S^{\text{control}}}{S^{\text{test}}}\right); S^{\text{control}}, X\right) &:= \mathbb{E}\left[\log\left(\text{CPX}^{\text{control}}\right)|S^{\text{test}}, S^{\text{control}}, X\right].
\end{aligned}$$

The causal relations of different notions are presented in Figure 2. The single direct arrows $A \rightarrow B$ means A causes B. The dashed line with double-direction arrows $A \longleftrightarrow B$ means that A and B are correlated but the causal relationship is undetermined. From the causal graph, the effect function depends on the campaign features and the algorithm (control, test) used. The spending depends on the budget of the campaign and the algorithm used. There is a correlation between the campaign features and the budget, but not a causal relation. The final observed CPX depends on the effect functions and the spending of both test and control groups as defined in Definition 2. The test and control spending are interfered by each other.

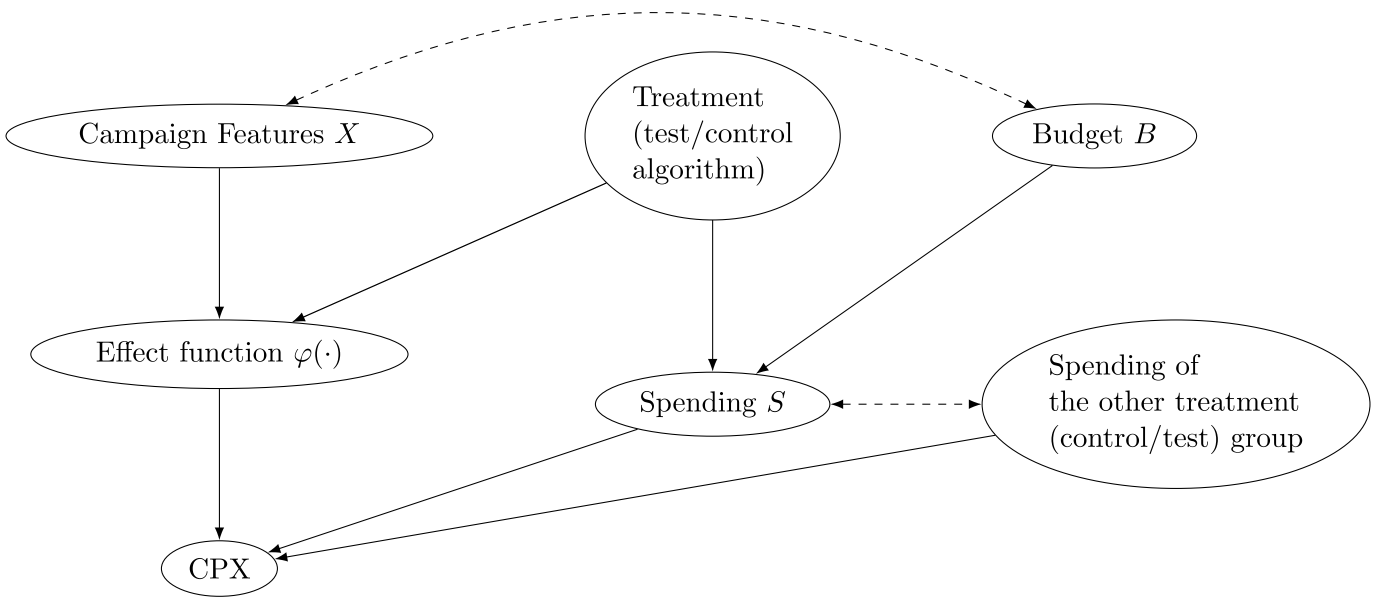


Figure 2: Causal graph of the A/B testing pipeline for each campaign

The key assumption is that the interference is symmetric, which means the interference effects could be offset if we swap the spending in the control and test groups. We formally state this assumption in Assumption 2 which enables us to estimate the individual treatment effect $\Delta(x)$ using the described A/B test.

Assumption 2. *For a given campaign with characteristics x, we assume the difference*

$$\begin{aligned}
&\mathbb{E}\left[\log\left(\text{CPX}^{\text{test}}\right)\middle| S^{\text{test}} = s_1, S^{\text{control}} = s_2, X = x\right] \\
&\quad - \mathbb{E}\left[\log\left(\text{CPX}^{\text{control}}\right)\middle| S^{\text{test}} = s_2, S^{\text{control}} = s_1, X = x\right] \qquad (7)
\end{aligned}$$

remains constant for any non-negative choices of s_1 *and* s_2*.*

Assumption 2 means that for any campaign, the individual treatment effect could be measured unbiasedly if we can observe a counterfactual with reversed test and control spendings.

Assumption 2 implies

$$\begin{aligned}
\Delta(x) &= \mathbb{E}\left[\log\left(\mathrm{CPX}^{\text{test}}\right)\middle| S^{\text{test}} = B, S^{\text{control}} = 0, X = x\right] \\
&\quad -\mathbb{E}\left[\log\left(\mathrm{CPX}^{\text{control}}\right)\middle| S^{\text{test}} = 0, S^{\text{control}} = B, X = x\right] \\
&= \mathbb{E}\left[\log\left(\frac{\mathrm{CPX}^{\text{test}}}{\mathrm{CPX}^{\text{control}}}\right)\middle| S^{\text{test}} = S^{\text{control}} = s, X = x\right] \\
&= \varphi^{\text{test}}\left(0; s, x\right) - \varphi^{\text{control}}\left(0; s, x\right).
\end{aligned} \tag{8}$$

Equation (8) means that if the control and test spending is perfectly balanced, the observed log-CPX ratio is an unbiased estimate for the individual treatment effect.

Then, we define the weighted average treatment effect below.

Definition 3 (Weighted average treatment effect)**.** Let $W = S^{\text{test}} + S^{\text{control}}$ be the weight. The weighted average treatment effect is defined as

$$\Delta_{\text{WATE}} = \frac{\mathbb{E}\left[W\Delta(X)\right]}{\mathbb{E}[W]},$$

where $\Delta(x)$ is the individual treatment effect defined in Definition 1.

The goal of the remainder of this section is to develop a statistical inference approach to estimate Δ_{WATE}.

4.2 Weighted Local Linear Regression

Note that if we denote $\delta = \log\left(S^{\text{test}}/S^{\text{control}}\right)$, then $S^{\text{test}} = e^{\delta} S^{\text{control}}$. This enables us to define

$$\bar{\varphi}\left(\delta; S^{\text{control}}, X\right) = \varphi^{\text{test}}\left(\delta; e^{\delta} S^{\text{control}}, X\right) - \varphi^{\text{control}}\left(-\delta; S^{\text{control}}, X\right). \tag{9}$$

Then, the individual treatment effect is

$$\Delta(x) = \bar{\varphi}(0; s, x).$$

According to Equation (9), conditional on S_{test}, S_{control}, and X, $\log\left(\mathrm{CPX}^{\text{test}}/\mathrm{CPX}^{\text{control}}\right)$ is an unbiased estimate of $\bar{\varphi}\left(\log\left(S^{\text{test}}/S^{\text{control}}\right); S^{\text{control}}, X\right)$. Further, we let $Y_i = \log\left(\mathrm{CPX}_i^{\text{test}}/\mathrm{CPX}_i^{\text{control}}\right)$, $Z_i = \log\left(S_i^{\text{test}}/S_i^{\text{control}}\right)$, and $W_i = S_i^{\text{test}} + S_i^{\text{control}}$ for $i = 1, 2, \ldots, N$. We assume that triplets $\{Y_i, Z_i, W_i\}_{i=1}^N$ are i.i.d. realizations of the random variable $\{Y, Z, W\}$ living on a probability space $\{\Omega, \mathcal{F}, P\}$ and we define a change of measure in this space:

$$\mathbb{E}_W\left[\cdot\right] := \mathbb{E}\left[W \times \cdot\right]/\mathbb{E}\left[W\right],$$

provided that $\mathbb{E}[W] < +\infty$. Therefore, $\Delta_{\text{WATE}} = \mathbb{E}_W\left[\Delta(X)\right]$.

Let $\beta(X, S) = \left[\beta_0(X), \beta_1(X, S)\right]^\top$, where $\beta_0(X) = \Delta(X)$ and $\beta_1(X, S) = \partial\bar{\varphi}\left(\delta; S, X\right)/\partial\delta|_{\delta=0}$. By a Taylor expansion under suitable regularity conditions (discussed in Assumption 3 below), we have

$$
\begin{aligned}
\mathbb{E}[Y|X,S,Z,W] &= \Delta(X) + \left(\partial\bar{\varphi}\left(\delta;S,X\right)/\partial\delta|_{\delta=0}\right)Z + b\left(Z^2;X,S\right) \\
&= \beta_0(X) + \beta_1(X,S)Z + b\left(Z^2;X,S\right),
\end{aligned}
\tag{10}
$$

where $b\left(Z^2;X,S\right)$ is a deterministic function representing the second-order bias and there exist a determinisic function $\psi(X,S)$ such that $b\left(Z^2;X,S\right) \leq \psi(X,S)Z^2$. We can further write

$$
Y = \mathbb{E}[Y|X,S,Z,W] + \epsilon\left(X\right), \tag{11}
$$

where $\mathbb{E}\left[\epsilon(X)|Z,X,S,W\right] = 0$ is the zero-mean residual. Here, for simplicity, we assume the residual only depends on X. By combining equations (10) and (11), we have the full expansion:

$$
Y = \beta_0(X) + \beta_1(X,S)Z + b\left(Z^2;X,S\right) + \epsilon\left(X\right). \tag{12}
$$

The regularity assumptions used above are explicitly defined in Assumption 3 below.

Assumption 3. *We assume $\bar{\varphi}\left(\cdot;s,x\right)$ is twice continuously differentiable and has uniformly bounded second derivatives i.e. $\sup_{s,x,\delta}\partial^2\bar{\varphi}\left(\delta;s,x\right)/\partial\delta^2 < +\infty$. We further assume W is bounded, $\mathrm{Var}(\epsilon(X)) < +\infty$, and*

$$
\sup_{s,x}\partial\bar{\varphi}\left(\delta;S,X\right)/\partial\delta|_{\delta=0} < +\infty, \quad \sup_{x}\Delta(x) < +\infty.
$$

By the Taylor expansion, if $|Z|$ is small, the bias term $b\left(Z^2;X,S\right)$ is also small. Therefore, it naturally leads us to an estimation method based on weighted local linear regression.

Local linear (polynomial) regression belongs to the family of popular non-parametric regression methods [12, 15, 13, 16, 17, 14]. The advantages are that these methods do not rely on strong parametric assumptions and enjoy a relatively fast convergence rate. The idea is to restrict the regression problem to a small neighborhood of the origin $z = 0$. After carefully balancing the bias and variance, one can obtain a reasonably good local approximation.

However, a challenge applying the local linear regression in our setting is that $\beta_0(X)$ and Z are correlated since the total spending $W = S^{\text{test}} + S^{\text{control}}$ is correlated with both $\beta_0(X)$ and Z. Nevertheless, by closely observing the causal graph in Figure 2, we realize that $X \perp\!\!\!\perp Z|W$ if W is proportional to B. Thus, conditional on W, $\beta_0(X)$ and Z are independent.

Assumption 4. *We assume the following two conditions.*

1. *$X \perp\!\!\!\perp Z|W$.*
2. *Z conditional on W has conditional density $f_{Z|W}(\cdot|w)$ around zero. Suppose that $f_{Z|W}(\cdot|w)$ is continuous differentiable with respect to the first argument for every w. We further assume that $f_{Z|W}(0|w)$ is uniformly bounded away from zero and $f'_{Z|W}(\cdot|w)$ is uniformly bounded at a neighborhood 0, i.e., there exists $\delta > 0$ such that*

$$
\sup_{w}\sup_{z\in(-\delta,\delta)}\left|f'_{Z|W}(z|w)\right| < +\infty.
$$

Before formally stating our testing method, we define $\pi_h(w) = \mathbb{P}(|Z| \leq h|W = w)$. $\pi_h(w)$ acts as the propensity score [39]. Algorithm 3 estimates Δ_{WATE} by weighted local linear regression with the inverse propensity weighting technique.

Algorithm 3 Weighted average treatment effect estimation

1: **Input:** The triples of log-CPX difference, log-spending, and weights $\{Y_i, Z_i, W_i\}_{i=1}^N$ of N campaigns, a bandwidth parameter h, and $\hat{\pi}_h(w)$, which is an estimate of $\mathbb{P}(|Z| \le h | W = w])$.
2: **Output:** An estimation of the weighted average treatment effect.
3: Select campaigns $\{Y_{i_j}, Z_{i_j}, W_{i_j}\}_{j=1}^M$ such that $|Z_{i_j}| \le h$ for $j = 1, 2, \ldots, M$.
4: Regress $\{Y_{i_j}\}_{j=1}^M$ on $\{Z_{i_j}\}_{j=1}^M$ with weights $\{W_{i_j}/\hat{\pi}_h(W_{i_j})\}_{j=1}^M$. Obtain the intercept $\hat{\beta}_0$.
5: **return** $\hat{\beta}_0$.

Remark: $\hat{\pi}_h(w)$ can be estimated using standard kernel estimators. However, in practice, weighting only by W would give a sufficiently good estimate since $\pi_h(W)$ is usually increasing with respect to W and the estimation is dominated by the high-weight campaigns. We will discuss practical implementation details in Section 5.2.

Once hyperparameters are specified, Algorithm 3 is easy to implement. In particular, we have closed form formulas

$$\begin{aligned}
\hat{\beta}_0 &= \frac{\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Y_{i_j}\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j} Y_{i_j}\right)}{\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right)\left(\sum_{j=1}^M \tilde{W}_{i_j}\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)^2} \\
\hat{\beta}_1 &= \frac{\left(\sum_{j=1}^M \tilde{W}_{i_j}\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j} Y_{i_j}\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Y_{i_j}\right)}{\left(\sum_{j=1}^M \tilde{W}_{i_j}\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)^2},
\end{aligned}$$

where $\tilde{W} = W/\hat{\pi}_h(W)$. Similar to Fan et al. [17, Theorem 5.2], we quantity the bias and variance of $\hat{\beta}_0$ in Theorem 2.

Theorem 2. *Suppose Assumptions 2, 3, and 4 are enforced and we assume $\hat{\pi}_h(w) = \pi_h(w)$, we have for any $h > 0$, when $N \to +\infty$*

$$\hat{\beta}_0 \to \Delta_{\mathrm{WATE}} + Bias(h) \ a.s.,$$

where the bias term admits

$$\lim_{h\to 0} \frac{1}{h^2} Bias(h) = b^* := \frac{1}{3}\mathbb{E}_W \left[((\beta_0(X) - \beta_0) + (\beta_1(X,S) - \beta_1)) \frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right] + \frac{1}{6}\mathbb{E}_W[\beta_2(X,S)],$$

with $\beta_2(X,S) = \partial^2 \bar{\varphi}(\delta; S, X)/\partial \delta^2\big|_{\delta=0}$. And the variance admits

$$Nh\mathrm{Var}\left(\hat{\beta}_0 | \{Z_i, W_i\}_{i=1}^N\right) \to V^* + o_p(1) \ a.s.,$$

where

$$V^* = \mathbb{E}\left[\frac{W^2(\beta_0(X) - \beta_0 + \epsilon(X))^2}{f(0|W)}\right] \Big/ \left(2(\mathbb{E}[W])^2\right),$$

and a random variable $\omega_h = o_p(1)$ means that $\omega_h \xrightarrow{p} 0$ as $h \to 0$.

Remark 1: If $\frac{f'_{Z|W}(0|w)}{f_{Z|W}(0|w)}$ is a constant (independent of w), we have

$$\mathbb{E}_W\left[\left((\beta_0(X)-\beta_0)+(\beta_1(X,S)-\beta_1)\right)\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right]=0,\text{ and } b^*=\frac{1}{6}\mathbb{E}_W[\beta_2(X,S)],$$

which is exactly the bias in local linear regression [14, Section 3.2]. Actually, $f'_{Z|W}(0|w)$ is often close to zero, as we shall see in Section 5.2. Therefore, b^* should be close to $\frac{1}{6}\mathbb{E}_W[\beta_2(X,S)]$ and we can borrow bias-estimation methods from local linear regression literature directly.

Theorem 2 states that the bias term is $O(h^2)$ and the variance term is $O\left((Nh)^{-1}\right)$. Based on the asymptotic bias and variance derived in Theorem 2, we establish the following central limit theorem.

Theorem 3. *Suppose Assumptions 2, 3, and 4 are enforced and we assume $\hat{\pi}_h(w)=\pi_h(w)$, we have for $Nh_N\to+\infty$ and $h_N\to 0$, we have*

$$\sqrt{Nh_N}\left(\hat{\beta}_0-\Delta_{\text{WATE}}-Bias(h_N)\right)\to\mathcal{N}\left(0,V^*\right),$$

where $Bias(h_N)$ and V^* are defined in Theorem 2.

To balance the bias and variance, we have the optimal bandwidth is

$$h_{\text{OPT}}=\left(\frac{V^*}{4\left(b^*\right)^2}\right)^{1/5}N^{-1/5}.$$

And under this choice of h_{OPT},

$$\sqrt{V^*/\left(Nh_{\text{OPT}}\right)}=2\left|b^*\right|h_{\text{OPT}}^2. \tag{13}$$

We will use techniques in Fan and Gijbels [14, Section 4.2] to estimate h_{OPT}. However, we will show empirically that the estimation of weighted average treatment effect is not sensitive to the specific choice of h as long as h is sufficiently small. The following proposition shows that we can estimate V^* from samples.

Proposition 2. *Suppose Assumptions 2, 3, and 4 are enforced and we assume $1/\hat{\pi}_h(w)\to 1/\pi_h(w)$ uniformly in probability as $N\to+\infty$. For any $h>0$, we have*

$$\frac{Nh\sum_{j=1}^M\tilde{W}_{i_j}^2\hat{\epsilon}_i^2}{\left(\sum_{j=1}^M\tilde{W}_{i_j}\right)^2}\to V^*+o_p(1),$$

where $\hat{\epsilon}_i$s are the estimated regression residuals in Algorithm 3.

Therefore, we could use $\sum_{j=1}^M\tilde{W}_{i_j}^2\hat{\epsilon}_i^2/\left(\sum_{j=1}^M\tilde{W}_{i_j}\right)^2$ as a consistent estimate of the sample variance $V^*/\left(Nh_N\right)$, i.e., the standard error is estimated by

$$\hat{\text{se}}=\sqrt{\sum_{j=1}^M\tilde{W}_{i_j}^2\hat{\epsilon}_i^2/\left(\sum_{j=1}^M\tilde{W}_{i_j}\right)^2}.$$

By using the equality (13), we have $\frac{1}{2}\sqrt{V^*/\left(Nh\right)}(h/h_{\text{OPT}})^{5/2}=\left|b^*\right|h^2$. Therefore, we can construct an approximate $1-\alpha$ confidence interval for Δ_{WATE} as

$$\left[\hat{\beta}_0 - \left(z_{1-\alpha/2} + \left(\frac{1}{2}\left(\frac{h}{h_{\text{OPT}}}\right)^{5/2}\right)\right)\hat{\text{se}}, \hat{\beta}_0 + \left(z_{1-\alpha/2} + \left(\frac{1}{2}\left(\frac{h}{h_{\text{OPT}}}\right)^{5/2}\right)\right)\hat{\text{se}}\right], \tag{14}$$

where $z_{1-\alpha/2}$ is a quantile of a standard normal distribution.

5 Experimental Results

5.1 Datasets and General Descriptions

We implement the two-step algorithm in an online DSP. We compare our algorithm with the benchmark algorithm in an online A/B test. The benchmark algorithm uses the first approach we discussed in Section 2, which only fits a logistic model to the binary feedback without using the additional information, see Pan et al. [38] and Pan et al. [37]. For our two-step algorithm describled in Section 3, $V^{(1)}$ and $V^{(2)}$ follow log-normal and Loglogistic distributions conditional on X, U, respectively. The A/B test was run from October 12th, 2021 to October 26th, 2021. As discussed in Section 4.1, the A/B test was conducted for each campaign, half traffic was allocated to the control group, while the other half was allocated to the test group (after a proper normalization). There are four different campaign delivery optimization goal types in the dataset, including minimizing cost per click (CPC), cost per action (CPA), cost per impression (CPM), and cost per intall (CPI). Table 1 summarizes the number of campaigns in each category.

Table 1: The number of campaigns in each category

	CPC	CPA	CPM	CPI
Number of campaigns	15020	12250	5931	1326

Table 2 shows the total rewards and the total amount of spending in different categories. Test and control groups indicate our algorithm and the benchmark algorithm, respectively. Rewards mean impressions, clicks, actions, and installs for CPC, CPA, CPM, and CPI, respectively. We observe that the total rewards in the test and control groups are almost the same with a difference less than 0.5% except CPI, while the test algorithm spends less. Therefore, those results imply that the test algorithm has lower costs and achieves almost the same reward as the control algorithm.

Table 2: The total rewards and the total amount of spending in different categories (proper normalized)

	Reward			Spending		
	Test	Control	Test/control ratio	Test	Control	Test/control ratio
CPC	1.04E+07	1.04E+07	1.004	4.04E+09	4.09E+09	0.987
CPA	4.88E+06	4.86E+06	1.003	8.61E+09	8.69E+09	0.991
CPM	2.91E+09	2.93E+09	0.995	3.94E+09	4.13E+09	0.954
CPI	5.28E+08	5.35E+08	0.988	5.38E+08	5.50E+08	0.979

We further compare the performance of test and control algorithms qualitatively. Note that under certain cases, we could know the test algorithm is better/worse than the control algorithm. Specifically, the test algorithm is better than the control algorithm for the i-th campaign, if

$$\text{(i) } R_i^{\text{test}} > R_i^{\text{control}} \text{ and } S_i^{\text{test}} < S_i^{\text{control}};\ \text{or}$$
$$\text{(ii) } \text{CPX}_i^{\text{test}} < \text{CPX}_i^{\text{control}} \text{ and } S_i^{\text{test}} > S_i^{\text{control}}.$$

The criteria (i) means that the test algorithm uses less spending to obtain more rewards and the criteria (2) means that the test algorithm achieves less CPX even if it spends more. Similarly, we know that the test algorithm is better than the control algorithm for the i-th campaign, if

$$\text{(i) } R_i^{\text{test}} < R_i^{\text{control}} \text{ and } S_i^{\text{test}} > S_i^{\text{control}};\ \text{or}$$
$$\text{(ii) } \text{CPX}_i^{\text{test}} > \text{CPX}_i^{\text{control}} \text{ and } S_i^{\text{test}} < S_i^{\text{control}}.$$

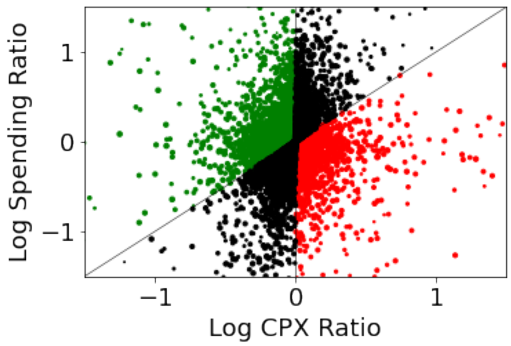
Table 3 shows the percentage of the number of campaigns and the total amount of spending that our algorithm is better/worse than the benchmark algorithm. We observe that for each category, there are more campaigns and more spending in which our algorithm is better than the benchmark algorithm.

Table 3: Percentage of the number of campaigns and the total amount of spending that our algorithm is better/worse than the benchmark algorithm

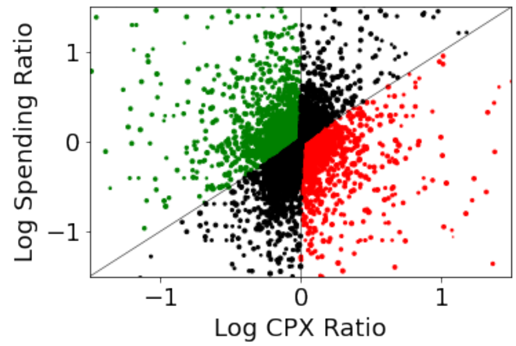
	Counts		Spending	
	Test Better	Test Worse	Test Better	Test Worse
CPC	35.2%	23.0%	50.7%	20.2%
CPA	37.3%	22.9%	47.8%	24.3%
CPM	29.1%	13.5%	30.5%	13.6%
CPI	25.9%	18.6%	32.1%	20.7%

Figure 3 shows the scatter plots of campaigns in each category, where the x-axis is the log CPX ratio and the y-axis is the log Spending ratio. Green points represent campaigns that the test algorithm is better; red points represent campaigns that the control algorithm is better; and black points represent campaigns that are undecidable. The point size is proportional to log of the total spending on this campaign. We observe that there are indeed more green points than red points.

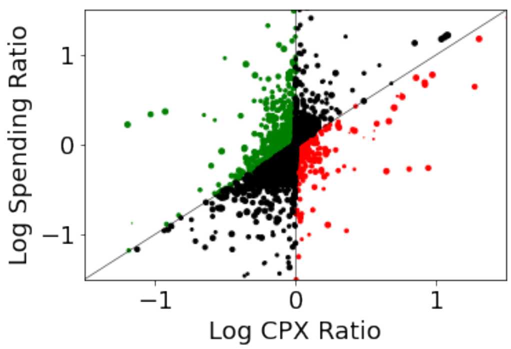
Based on the aforementioned evidences, we have heuristically examined the performance of our algorithm and concluded the superiority of our algorithm qualitatively. In the next subsection, we conduct the statistical test method developed in Section 4.2 and quantitatively compute the average treatment effect and confidence intervals.



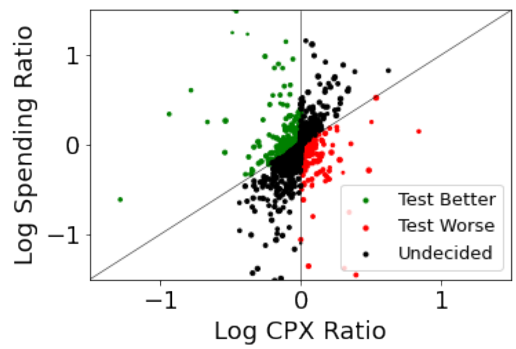
(a) CPC



(b) CPA



(c) CPM



(d) CPI

Figure 3: Log spending ratio v.s. log CPX ratio for each campaign

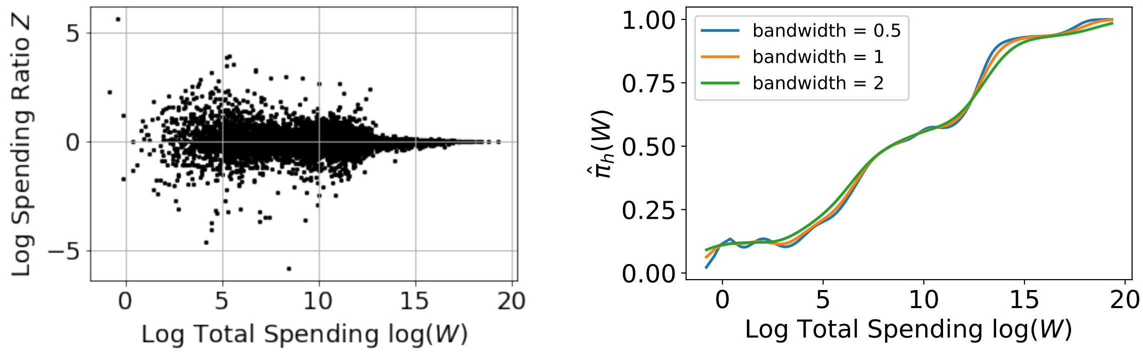
5.2 Statistical Test via Weighted Local Linear Regression

We first discuss the estimation of $\hat{\pi}_h(\cdot)$. We use the kernel estimator. Given a kernel $K(\cdot)$ and a bandwidth bw. We estimate $\pi_h(w)$ as

$$\hat{\pi}_h(w) = \frac{\sum_{i=1}^{N} \mathbb{I}\{|Z_i| \leq h\} K\left(\frac{\log(W_i)-\log(w)}{\mathrm{bw}}\right)}{\sum_{i=1}^{N} K\left(\frac{\log(W_i)-\log(w)}{\mathrm{bw}}\right)}.$$

Figure 4(a) shows the dependence of log spending ratio Z on log total spending $\log(W)$ for the CPC goal type. We see that Z is concentrated around zero and the variance of Z is decreasing with respect to the increase of $\log(W)$. Figure 4(b) shows estimated $\hat{\pi}_h(\cdot)$ using different bandwidths for $h = 0.1$ and the CPC goal type. We see the estimation is not sensitive the specific bandwidth choice and $\hat{\pi}_h(w)$ is increasing with respect to w. From now on, we will stick to the bandwidth choice $\mathrm{bw} = 1$. Additional plots for CPA, CPM, CPI goal types are shown in Appendix Appendix B.

We used the method developed in Fan and Gijbels [14, Section 4.2] to obtain rule of thumb bandwidth estimator $\hat{h}_{\mathrm{ROT}}$. Note that $\hat{h}_{\mathrm{ROT}}$ is only a guidance for the choice of bandwidth and in practice, we recommend a smaller value to reduce bias and obtain valid confidence interval.



(a) Log spending ratio v.s. log total spending

(b) Estimated $\pi_h(\cdot)$

Figure 4: Log spending ratio with respect to log total spending and estimated $\pi_h(\cdot)$ for $h = 0.1$ and the CPC goal type

Table 4 shows the rule of thumb bandwidth estimates for different campaign goals.

Table 4: Rule of thumb choice of h

	CPC	CPA	CPM	CPI
$\hat{h}_{\text{ROT}}$	0.41	0.20	0.72	0.25

By following the CI construction (14), we use $\hat{h}_{\text{ROT}}$ to approximate h_{OPT} and thus we could construct an approximate $1-\alpha$ confidence interval as

$$\left[\hat{\beta}_0 - \left(z_{1-\alpha/2} + \left(\frac{1}{2}\left(\frac{h}{\hat{h}_{\text{ROT}}}\right)^{5/2}\right)\right)\hat{\text{se}}, \hat{\beta}_0 + \left(z_{1-\alpha/2} + \left(\frac{1}{2}\left(\frac{h}{\hat{h}_{\text{ROT}}}\right)^{5/2}\right)\right)\hat{\text{se}}\right]. \tag{15}$$

Table 5 shows the estimation results for $h \in \{\hat{h}_{\text{ROT}}, 0.05, 0.1, 0.2\}$ and four different goal types. We see that the average treatment effect estimators are similar for $h \in \{0.05, 0.1, 0.2\}$. In particular, they are within all of confidence intervals generated by $h \in \{0.05, 0.1, 0.2\}$. In practice, we recommend using $h = 0.1$. Furthermore, all estimation values are significant at 0.1% significance level and the superiority is about 1.5%, 0.5%, 0.3%, and 0.8%, for CPC, CPA, CPM, and CPI, respectively. We remark that those amounts of cost reduction could translate to millions of dollars each year. Therefore, we are able to safely claim that our algorithm is better than the benchmark algorithm. We also report additional results on the statistical tests without inverse weighting by $\hat{\pi}_h\cdot$, i.e., $\tilde{W} = W$ in Appendix Appendix B. And those results are similar to the results in Table 5. Therefore, it is not necessary to estimate $\hat{\pi}_h(\cdot)$ in practice, as campaigns with large weights are usually balanced between the test and control spending.

6 Conclusion

In the paper, we propose a novel bidding algorithm that incorporates additional information provided by the intermediate exchange. To formally examine the effectiveness of the proposed

Table 5: Estimation results via weighted local linear regression

Goal	h	Estimation	95% Confidence interval	t-statistics
CPC	$\hat{h}_{\mathrm{ROT}}$	-0.016	(-0.020, -0.012)	-9.53
	0.05	-0.013	(-0.015, -0.010)	-8.78
	0.10	-0.014	(-0.017, -0.011)	-9.00
	0.20	-0.015	(-0.018, -0.012)	-9.43
CPA	$\hat{h}_{\mathrm{ROT}}$	-0.006	(-0.007, -0.004)	-10.43
	0.05	-0.005	(-0.006, -0.004)	-10.07
	0.10	-0.005	(-0.007, -0.004)	-10.33
	0.20	-0.006	(-0.007, -0.004)	-10.48
CPM	$\hat{h}_{\mathrm{ROT}}$	-0.002	(-0.007, 0.003)	-1.14
	0.05	-0.003	(-0.004, -0.002)	-6.57
	0.10	-0.004	(-0.006, -0.003)	-5.50
	0.20	-0.004	(-0.005, -0.002)	-4.39
CPI	$\hat{h}_{\mathrm{ROT}}$	-0.008	(-0.012, -0.004)	-5.37
	0.05	-0.006	(-0.009, -0.003)	-3.96
	0.10	-0.008	(-0.011, -0.005)	-5.26
	0.20	-0.008	(-0.012, -0.005)	-5.50

algorithm, we further propose a novel statistical testing scheme via weighted local linear regression. We establish the central limit theorem and discuss the choice of the key parameters of our proposed testing method. Using this method, we qualitatively demonstrate that our new bidding algorithm achieves 0.3% to 1.5% cost reductions compared to the benchmark algorithm for different goal types. The new bidding algorithm and the new testing scheme has been deployed in a major DSP in the US.

For the future work, we plan to investigate other aspects of multi-layer auctions including ad ranking and auction design. We also expect to generalize our experimentation methods to other types of online marketplaces, including ride sharing and online booking.

References

[1] AppNexus. Demystifying auction dynamics for digital buyers and sellers, 2018. URL `https://www.appnexus.com/sites/default/files/whitepapers/49344-CM-Auction-Type-Whitepaper-V9.pdf`.

[2] Lennart Baardman, Elaheh Fata, Abhishek Pani, and Georgia Perakis. Learning optimal online advertising portfolios with periodic budgets. *Available at SSRN 3346642*, 2019.

[3] Patrick Bajari, Brian Burdick, Guido W Imbens, Lorenzo Masoero, James McQueen, Thomas Richardson, and Ido M Rosen. Multiple randomization designs. *arXiv preprint arXiv:2112.13495*, 2021.

[4] Santiago R Balseiro and Yonatan Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968, 2019.

[5] Santiago R Balseiro, Omar Besbes, and Gabriel Y Weintraub. Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Science*, 61(4):864–884, 2015.

[6] Guillaume W Basse, Hossein Azari Soufiani, and Diane Lambert. Randomization and the pernicious effects of limited budgets on auction experiments. In *Artificial Intelligence and Statistics*, pages 1412–1420. PMLR, 2016.

[7] Thomas Blake and Dominic Coey. Why marketplace experimentation is harder than it seems: The role of test-control interference. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 567–582, 2014.

[8] Iavor Bojinov and Neil Shephard. Time series experiments and causal estimands: exact randomization tests and trading. *Journal of the American Statistical Association*, 114(528): 1665–1682, 2019.

[9] V Chari and Robert Weber. How the us treasury should auction its debt. *Federal Reserve Bank of Minneapolis Quarterly Review*, 16(4):1–12, 1992. URL `http://kylewoodward.com/blog-data/pdfs/references/chari+weber-quarterly-review-1992A.pdf`.

[10] Alex Deng and Xiaolin Shi. Data-driven metric development for online controlled experiments: Seven lessons learned. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 77–86, 2016.

[11] Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1), 2017.

[12] Jianqing Fan. Local linear regression smoothers and their minimax efficiencies. *The annals of Statistics*, pages 196–216, 1993.

[13] Jianqing Fan and Irene Gijbels. Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(2):371–394, 1995.

[14] Jianqing Fan and Irene Gijbels. *Local polynomial modelling and its applications*. Routledge, 2018.

[15] Jianqing Fan, T Gasser, I Gijbels, M Brockmann, and J Engel. On nonparametric estimation via local polynomial regression. *Institute of Statistics, Catholic University of Louvain*, 1995.

[16] Jianqing Fan, Nancy E Heckman, and Matt P Wand. Local polynomial kernel regression for generalized linear models and quasi-likelihood functions. *Journal of the American Statistical Association*, 90(429):141–150, 1995.

[17] Jianqing Fan, Irène Gijbels, Tien-Chung Hu, and Li-Shan Huang. A study of variable bandwidth selection for local polynomial regression. *Statistica Sinica*, pages 113–127, 1996.

[18] Getintent. Rtb auctions: Fair play?, 2017. URL `https://blog.getintent.com/rtb-auctions-fair-play-3b372d505089`.

[19] Djordje Gligorijevic, Tian Zhou, Bharatbhushan Shetty, Brendan Kitts, Shengjun Pan, Junwei Pan, and Aaron Flores. Bid shading in the brave new world of first-price auctions. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 2453–2460, 2020.

[20] Google. Real time bidding protocol, release notes, 2019. URL `https://developers.google.com/authorized-buyers/rtb/relnotes#updates-2019-03-13`.

[21] Google. Real time bidding protocol protocol buffer v.167, 2019. URL `https://developers.google.com/authorized-buyers/rtb/downloads/realtime-bidding-proto`.

[22] Google. Introduction to real-time bidding (rtb) - authorized buyers help, 2022. URL `https://support.google.com/authorizedbuyers/answer/6136272`.

[23] Yanjun Han, Zhengyuan Zhou, Aaron Flores, Erik Ordentlich, and Tsachy Weissman. Learning to bid optimally and efficiently in adversarial first-price auctions. *arXiv preprint arXiv:2007.04568*, 2020.

[24] Yue He, Xiujun Chen, Di Wu, Junwei Pan, Qing Tan, Chuan Yu, Jian Xu, and Xiaoqiang Zhu. A unified solution to constrained bidding in online display advertising. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 2993–3001, 2021.

[25] David Holtz and Sinan Aral. Limiting bias from test-control interference in online marketplace experiments. *arXiv preprint arXiv:2004.12162*, 2020.

[26] B. Hovaness. Sold for more than you should have paid, 2018. URL `https://www.hearts-science.com/sold-for-more-than-you-should-have-paid/`.

[27] Michael G Hudgens and M Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

[28] Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

[29] Ramesh Johari, Hannah Li, Inessa Liskovich, and Gabriel Y Weintraub. Experimental design in two-sided platforms: An analysis of bias. *Management Science*, 2022.

[30] Niklas Karlsson. Feedback control in programmatic advertising: The frontier of optimization in real-time bidding. *IEEE Control Systems Magazine*, 40(5):40–77, 2020.

[31] Niklas Karlsson and Qian Sang. Adaptive bid shading optimization of first-price ad inventory. In *2021 American Control Conference (ACC)*, pages 4983–4990. IEEE, 2021.

[32] Brendan Kitts. Bidder behavior after shifting from second to first price auctions in online advertising, 2019. URL `http://www.appliedaisystems.com/papers/FPA_Effects33.pdf`.

[33] Ron Kohavi, Alex Deng, Brian Frasca, Roger Longbotham, Toby Walker, and Ya Xu. Trustworthy online controlled experiments: Five puzzling outcomes explained. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 786–794, 2012.

[34] Ron Kohavi, Diane Tang, and Ya Xu. *Trustworthy online controlled experiments: A practical guide to a/b testing*. Cambridge University Press, 2020.

[35] Hannah Li, Geng Zhao, Ramesh Johari, and Gabriel Y Weintraub. Interference, bias, and variance in two-sided marketplace experimentation: Guidance for platforms. In *Proceedings of the ACM Web Conference 2022*, pages 182–192, 2022.

[36] Min Liu, Jialiang Mao, and Kang Kang. Trustworthy online marketplace experimentation with budget-split design. *arXiv preprint arXiv:2012.08724*, 2020.

[37] Junwei Pan, Jian Xu, Alfonso Lobos Ruiz, Wenliang Zhao, Shengjun Pan, Yu Sun, and Quan Lu. Field-weighted factorization machines for click-through rate prediction in display advertising. In *Proceedings of the 2018 World Wide Web Conference*, page 1349–1357, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.

[38] Shengjun Pan, Brendan Kitts, Tian Zhou, Hao He, Bharatbhushan Shetty, Aaron Flores, Djordje Gligorijevic, Junwei Pan, Tingyu Mao, San Gultekin, and Jianlong Zhang. Bid shading by win-rate estimation and surplus maximization. *arXiv preprint arXiv:2009.09259*, 2020.

[39] Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[40] Rubicon. Bridging the gap to first-price auctions: A buyer's guide, 2018. URL `http://go.rubiconproject.com/rs/958-XBX-033/images/Buyers_Guide_to_First_Price_Rubicon_Project.pdf`.

[41] Rubicon. Principles for a better programmatic marketplace, open letter from rubicon, spotx, openx, pubmatic, telaris, and sovr, 2018. URL `https://rubiconproject.com/insights/thought-leadership/principles-better-programmatic-marketplace-open-letter-advertisers-publishers/`.

[42] Martin Saveski, Jean Pouget-Abadie, Guillaume Saint-Jacques, Weitao Duan, Souvik Ghosh, Ya Xu, and Edoardo M Airoldi. Detecting network effects: Randomizing over randomized experiments. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1027–1035, 2017.

[43] Statista. U.S. digital advertising industry - statistics & facts, 2022. URL `https://www.statista.com/topics/1176/online-advertising/`.

[44] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[45] Diane Tang, Ashish Agarwal, Deirdre O'Brien, and Mike Meyer. Overlapping experiment infrastructure: More, better, faster experimentation. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 17–26, 2010.

[46] Di Wu, Xiujun Chen, Xun Yang, Hao Wang, Qing Tan, Xiaoxun Zhang, Jian Xu, and Kun Gai. Budget constrained bidding by model-free reinforcement learning in display advertising. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1443–1451, 2018.

[47] Wush Wu, Mi-Yen Yeh, and Ming-Syan Chen. Deep censored learning of the winning price in the real time bidding. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2526–2535, London, United Kingdom, 2018. KDD '18.

[48] Wush Chi-Hsuan Wu, Mi-Yen Yeh, and Ming-Syan Chen. Predicting winning price in real time bidding with censored data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1305–1314, Sydney, 2015. KDD '15.

[49] Ya Xu, Nanyu Chen, Addrian Fernandez, Omar Sinno, and Anmol Bhasin. From infrastructure to culture: A/B testing challenges in large scale social networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2227–2236, 2015.

[50] Wei Zhang, Brendan Kitts, Yanjun Han, Zhengyuan Zhou, Tingyu Mao, Hao He, Shengjun Pan, Aaron Flores, San Gultekin, and Tsachy Weissman. Meow: A space-efficient nonparametric bid shading algorithm. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 3928–3936, 2021.

[51] Tian Zhou, Hao He, Shengjun Pan, Niklas Karlsson, Bharatbhushan Shetty, Brendan Kitts, Djordje Gligorijevic, San Gultekin, Tingyu Mao, Junwei Pan, et al. An efficient deep distribution network for bid shading in first-price auctions. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 3996–4004, 2021.

Appendix A Proofs

Appendix A.1 Proofs of Results in Section 3

Proof of Theorem 1. Let $F(b) = \prod_{i=1}^{m} F_i(b)$. Then, the first-order condition $f(b) = (u-b)F(b)$ is

$$-F(b) + (u-b)F'(b) = 0.$$

We only need to show $b + F(b)/F'(b)$ is non-decreasing. Note that

$$b + F(b)/F'(b) \quad = \quad b + \frac{\prod_{i=1}^{m} F_i(b)}{\sum_{j=1}^{m} f_j(b) \prod_{i=1, i\neq j}^{m} F_i(b)}$$

$$= \quad b + \frac{1}{\sum_{j=1}^{m} f_j(b)/F_i(b)}.$$

Let $h_i(b) = f_j(b)/F_i(b)$. Then,

$$\frac{\partial \left(b + \left(\sum_{j=1}^{m} h_i(b) \right)^{-1} \right)}{\partial b} = 1 - \left(\sum_{j=1}^{m} h_i(b) \right)^{-2} \sum_{i=1}^{m} h_i'(b).$$

Notice that the monotonicity of $b + F_i(b)/f_i(b)$ implies that

$$\frac{\partial \left(b + h_i(b)^{-1} \right)}{\partial b} = 1 - h_i(b)^{-2} h_i'(b) \geq 0 \Leftrightarrow h_i(b)^2 \geq h_i'(b).$$

Then, recall that $h_i(b) > 0$, we have

$$1 - \left(\sum_{j=1}^{m} h_i(b) \right)^{-2} \sum_{i=1}^{m} h_i'(b) \geq 1 - \left(\sum_{j=1}^{m} h_i(b) \right)^{-2} \sum_{j=1}^{m} h_i(b)^2 \geq 0.$$

□

Proof of Corollary 1. Log-concave $F(\cdot)$ means that

$$\frac{\partial \log(F(b))}{\partial b} = \frac{f(b)}{F(b)} \text{ is decreasing.}$$

If $\tilde{F}(\cdot)$ is log-concave, we have for any $b_1, b_2 > 0$

$$\begin{aligned} \frac{1}{2} \left(\log \left(\tilde{F}(\log(b_1)) \right) + \log \left(\tilde{F}(b_2) \right) \right) &\leq \log \tilde{F} \left(\frac{\log(b_1) + \log(b_2)}{2} \right) \\ &= \log \tilde{F} \left(\log \left(\sqrt{b_1 b_2} \right) \right) \\ &\leq \log \tilde{F} \left(\log \left(\frac{b_1 + b_2}{2} \right) \right). \end{aligned}$$

□

Proof of Lemma 2. Recall that the distribution of student's t-distribution has density

$$f(x) = A_\upsilon (1 + x^2/\upsilon)^{-\frac{\upsilon+1}{2}},$$

where $\upsilon \geq 1$ is the degree of freedom and A_υ is a normalization constant. When $x \geq 0$, $f(x)$ is decreasing and thus $F(x)/f(x)$ is increasing. There, we only need to focus on $x < 0$. Then, we claim

$$F(x) = A_\upsilon \int_{-\infty}^{x} \frac{1}{(1 + b^2/\upsilon)^{(\upsilon+1)/2}} \mathrm{d}b \leq -A_\upsilon (1 + x^2/\upsilon)^{-\frac{\upsilon-1}{2}}/x.$$

This due to

$$\frac{\partial}{\partial b} \left(-(1 + b^2/\upsilon)^{-\frac{\upsilon-1}{2}}/b \right) \quad = \quad \left(1 + \frac{1}{b^2} \right) \frac{1}{(1 + b^2/\upsilon)^{(\upsilon+1)/2}}$$

$$\geq \quad \frac{1}{(1+b^2/\upsilon)^{(\upsilon+1)/2}}.$$

Further, we have

$$f'(x) = -A_\upsilon \left(\frac{\upsilon+1}{\upsilon}\right) x(1+x^2/\upsilon)^{-\frac{\upsilon+3}{2}}.$$

Therefore, we have

$$F(x)f'(x) \leq \frac{\upsilon+1}{\upsilon} A_\upsilon^2 (1+x^2/\upsilon)^{-(\upsilon+1)} = \frac{\upsilon+1}{\upsilon} f(x)^2 \leq 2f(x)^2.$$

On the other hand, we have

$$\left(b + \frac{F(b)}{f(b)}\right)' = \frac{2f(b)f(b) - F(b)f'(b)}{f(b)^2} \geq 0,$$

which completes the proof. □

Proof of Proposition 1. We write

$$\begin{aligned}
S'(b) &= -F(b) + (u-b)F'(b) \\
&= F(b)\left(-1 + (u-b)\frac{F'(b)}{F(b)}\right) \\
&= F(b)\left(-1 + (u-b)\sum_{j=1}^m f_j(b)/F_j(b)\right) \\
&= F(b)\left(-1 + (u-b)\sum_{j=1}^m \frac{\frac{1}{b}\tilde{f}_j(\log(b))}{\tilde{F}_j(\log(b))}\right) \\
&= \frac{F(b)}{b}\left(-b + (u-b)\sum_{j=1}^m \frac{\tilde{f}_j(\log(b))}{\tilde{F}_j(\log(b))}\right).
\end{aligned}$$

Therefore, we have

$$b^* = (u-b^*)\sum_{j=1}^m \frac{\tilde{f}_j(\log(b^*))}{\tilde{F}_j(\log(b^*))}.$$

Since $b^* \in [0,u]$ and $\frac{\tilde{f}_j(\log(b))}{\tilde{F}_j(\log(b))}$ is non-increasing, we have

$$b^* \geq (u-b^*)\sum_{j=1}^m \frac{\tilde{f}_j(\log(u))}{\tilde{F}_j(\log(u))},$$

which leads to

$$b^* \geq \frac{u\sum_{j=1}^m \frac{\tilde{f}_j(\log(u))}{\tilde{F}_j(\log(u))}}{\sum_{j=1}^m \frac{\tilde{f}_j(\log(u))}{\tilde{F}_j(\log(u))} + 1} = \underline{b}^{(1)}.$$

Then, we further have

$$b^* \leq (u-b^*)\sum_{j=1}^m \frac{\tilde{f}_j(\log(\underline{b}^{(1)}))}{\tilde{F}_j(\log(\underline{b}^{(1)}))},$$

which is equivalent to

$$b^* \le \frac{u\sum_{j=1}^m \frac{\tilde{f}_j\left(\log\left(\underline{b}^{(1)}\right)\right)}{\tilde{F}_j\left(\log\left(\underline{b}^{(1)}\right)\right)}}{\sum_{j=1}^m \frac{\tilde{f}_j\left(\log\left(\underline{b}^{(1)}\right)\right)}{\tilde{F}_j\left(\log\left(\underline{b}^{(1)}\right)\right)}+1} = \bar{b}^{(1)}.$$

Then, by induction, we have desired results. □

Appendix A.2 Proofs of Results in Section 4

Lemma A3 is useful in proving Theorem 2.

Lemma A3. *Under Assumptions 3 and 4, we have*

$$\begin{aligned} \lim_{h\to 0}\frac{\mathbb{E}[Z|W=w,|Z|\le h]}{h^2} &= \frac{1}{3}\frac{f'_{Z|W}(0;w)}{f_{Z|W}(0;w)}, \text{ and} \\ \lim_{h\to 0}\frac{\mathbb{E}[Z^2|W=w,|Z|\le h]}{h^2} &= \frac{1}{3}. \end{aligned}$$

Furthermore, $\frac{\mathbb{E}[Z|W=w,|Z|\le h]}{h^2}$ *and* $\frac{\mathbb{E}[Z^2|W=w,|Z|\le h]}{h^2}$ *are uniformly bounded when* h *is sufficiently small, i.e., there exist a* $\bar{h}>0$, *such that*

$$\begin{aligned} \sup_{h<\bar{h}}\sup_{w}\left|\frac{\mathbb{E}[Z|W=w,|Z|\le h]}{h^2}\right| &< +\infty, \text{ and} \\ \sup_{h<\bar{h}}\sup_{w}\left|\frac{\mathbb{E}[Z^2|W=w,|Z|\le h]}{h^2}\right| &< +\infty. \end{aligned}$$

Proof of Lemma A3. For the first limits, we note that

$$\mathbb{E}[Z|W=w,|Z|\le h] = \frac{\mathbb{E}[Z\mathbb{I}\{|Z|\le h\}|W=w]}{\mathbb{P}(|Z|\le h|W=w)}.$$

For the nominator,

$$\mathbb{E}[Z\mathbb{I}\{|Z|\le h\}|W=w] = \int_{-h}^{h} z f_{Z|W}(z;w)\,\mathrm{d}z. \tag{A. 1}$$

By Taylor's theorem, we have

$$f_{Z|W}(z;w) = f_{Z|W}(0;w) + f'_{Z|W}(0;w)\,z + h_1(z)z,$$

where $\lim_{z0} h_1(z)=0$. Then, we have the integral (A. 1) is

$$\begin{aligned} &\int_{-h}^{h} z\left(f_{Z|W}(0;w) + f'_{Z|W}(0;w)\,z + h_1(z)z\right)\mathrm{d}z \\ =\ & \frac{2}{3}f'_{Z|W}(0;w)\,h^3 + \int_{-h}^{h} h_1(z)z^2\mathrm{d}z. \end{aligned}$$

Then, $\left|\frac{1}{h^2}\int_{-h}^{h} h_1(z)z^2 \mathrm{d}z\right| \leq \int_{-h}^{h} h_1(z)\mathrm{d}z \to 0$ as $h \to 0$. For the denominator, we have

$$\begin{aligned}
\mathbb{P}(|Z| &\leq h|W = w) = \int_{-h}^{h} f_{Z|W}(z; w)\,\mathrm{d}z \\
&= \int_{-h}^{h} \left(f_{Z|W}(0; w) + f'_{Z|W}(0; w)\,z + h_1(z)z\right)\mathrm{d}z \\
&= 2f_{Z|W}(0; w)\,h + \int_{-h}^{h} h_1(z)z\mathrm{d}z,
\end{aligned}$$

which yields

$$\lim_{h\to 0} \frac{\mathbb{P}(|Z| \leq h|W = w)}{h} = 2f_{Z|W}(0; w)\,.$$

For the second limits, note that

$$\mathbb{E}[Z^2|W = w, |Z| \leq h] = \frac{\mathbb{E}[Z^2\mathbb{I}\{|Z| \leq h\}|W = w]}{\mathbb{P}(|Z| \leq h|W = w)},$$

and the nominator is

$$\begin{aligned}
\mathbb{E}[Z^2\mathbb{I}\{|Z| \leq h\}|W = w] &= \int_{-h}^{h} z^2\left(f_{Z|W}(0; w) + f'_{Z|W}(0; w)\,z + h_1(z)z\right)\mathrm{d}z \\
&= \frac{2}{3}f_{Z|W}(0; w)\,h^3 + \int_{-h}^{h} z^3 h_1(z)\mathrm{d}z.
\end{aligned}$$

By the same argument as the first limits, we have the desired results.

Finally, for the uniformly boundedness, we observe that

$$f_{Z|W}(0; w) - \bar{c}z \leq f_{Z|W}(z; w) \leq f_{Z|W}(0; w) + \bar{c}z,$$

for z at a neighborhood zero, where

$$\bar{c} = \sup_{w} \sup_{z\in(-\delta,\delta)} f'_{Z|W}(z; w) < +\infty.$$

□

Let $\beta_0 = \mathbb{E}_W[\beta_0(X)] = \Delta_{\text{WATE}}$ and $\beta_1 = \mathbb{E}_W[\beta_1(X, S)]$.

(1) We first deal with the bias part. First, note that as $N \to +\infty$.

$$\hat{\beta}_0 \to \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \leq h\}\right]\mathbb{E}\left[\frac{W}{\pi_h(W)}Y\mathbb{I}\{|Z| \leq h\}\right] - \mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \leq h\}\right]\mathbb{E}\left[\frac{W}{\pi_h(W)}ZY\mathbb{I}\{|Z| \leq h\}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \leq h\}\right]\mathbb{E}[W] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \leq h\}\right]\right)^2} \tag{A. 2}$$

By the tower property of conditional expectation, we observe that

$$\begin{aligned}
\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \leq h\}\right] &= \mathbb{E}\left[W\mathbb{E}\left[Z^2|W, |Z| \leq h\right]\right] \text{ and} \\
\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \leq h\}\right] &= \mathbb{E}\left[W\mathbb{E}\left[Z|W, |Z| \leq h\right]\right].
\end{aligned}$$

Recall that that

$$Y = \beta_0(X) + \beta_1(X,S)Z + b\left(Z^2|X,S\right) + \epsilon\left(Z|X,S\right),$$

where $\mathbb{E}\left[\epsilon\left(Z|X,S\right)|Z,X,S,W\right] = 0$. We have

$$\begin{aligned}
& \mathbb{E}\left[\frac{W}{\pi_h(W)} Y \mathbb{I}\{|Z| \le h\}\right] \\
= \;& \mathbb{E}\left[\frac{W}{\pi_h(W)}\left(\beta_0(X) + \beta_1(X,S)Z + b\left(Z^2|X,S\right)\right)\mathbb{I}\{|Z| \le h\}\right] \\
= \;& \mathbb{E}\left[W\beta_0(X)\right] + \mathbb{E}\left[\frac{W}{\pi(W)}\left(\beta_1(X,S)Z + b\left(Z^2|X,S\right)\right)\mathbb{I}\{|Z| \le h\}\right],
\end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E}\left[\frac{W}{\pi_h(W)} ZY \mathbb{I}\{|Z| \le h\}\right] \\
= \;& \mathbb{E}\left[\frac{W}{\pi_h(W)}\left(\beta_0(X)Z + \beta_1(X,S)Z^2 + b\left(Z^2|X,S\right)Z\right)\mathbb{I}\{|Z| \le h\}\right]. \qquad \text{(A. 3)}
\end{aligned}$$

For the $\beta_0(X)$ term in Equation (A. 3), we observe that

$$\begin{aligned}
& \mathbb{E}\left[\frac{W\beta_0(X)Z\mathbb{I}\{|Z| \le h\}}{\pi_h(W)}\right] \\
= \;& \mathbb{E}\left[W\beta_0(X)\mathbb{E}[Z|W,|Z| \le h]\right] \\
= \;& \mathbb{E}\left[W\left(\beta_0(X) - \beta_0\right)\mathbb{E}[Z|W,|Z| \le h]\right] + \beta_0\mathbb{E}\left[\frac{W}{\pi_h(W)} Z\mathbb{I}\{|Z| \le h\}\right].
\end{aligned}$$

The term in (A. 2) could be decomposed as

$$\begin{aligned}
& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{W}{\pi_h(W)}Y\mathbb{I}\{|Z| \le h\}\right] - \mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{W}{\pi_h(W)}ZY\mathbb{I}\{|Z| \le h\}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\right] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\right)^2} \\
= \;& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\left(\mathbb{E}\left[W\beta_0(X)\right] + \mathbb{E}\left[\frac{W}{\pi(W)}\left(\beta_1(X,S)Z + b\left(Z^2|X,S\right)\right)\mathbb{I}\{|Z| \le h\}\right]\right)}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\right] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\right)^2} - \\
& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{W}{\pi_h(W)}\left(\beta_0(X)Z + \beta_1(X,S)Z^2 + b\left(Z^2|X,S\right)Z\right)\mathbb{I}\{|Z| \le h\}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\right] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\right)} \\
= \;& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\beta_0(X)\right] - \mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{W\beta_0(X)Z\mathbb{I}\{|Z| \le h\}}{\pi_h(W)}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\right] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\right)} + \qquad \text{(A. 4)} \\
& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{W\beta_1(X,S)Z\mathbb{I}\{|Z| \le h\}}{\pi_h(W)}\right] - \mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{W\beta_1(X,S)Z^2\mathbb{I}\{|Z| \le h\}}{\pi_h(W)}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\right] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\right)} \qquad \text{(A. 5)} \\
& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{Wb\left(Z^2|X,S\right)\mathbb{I}\{|Z| \le h\}}{\pi_h(W)}\right] - \mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[\frac{Wb\left(Z^2|X,S\right)\mathbb{I}\{|Z| \le h\}}{\pi(W)}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z| \le h\}\right]\mathbb{E}\left[W\right] - \left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z| \le h\}\right]\right)} \qquad \text{(A. 6)}
\end{aligned}$$

Then, we have for the first term (A. 4), we have

$$
\begin{aligned}
& \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[W\beta_0(X)\right]-\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[\frac{W\beta_0(X)Z\mathbb{I}\{|Z|\le h\}}{\pi_h(W)}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z|\le h\}\right]\right)} \\
= & \frac{\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\beta_0(X)\right]-\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\mathbb{E}\left[\frac{W\beta_0(X)Z\mathbb{I}\{|Z|\le h\}}{\pi(W)}\right]}{\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\right)^2} \\
= & \\
& \frac{\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\beta_0(X)\right]-\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\mathbb{E}\left[\frac{W\beta_0(X)Z\mathbb{I}\{|Z|\le h\}}{\pi_h(W)}\right]}{\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\right)^2} \\
= & \beta_0-\frac{\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\mathbb{E}\left[W\left(\beta_0(X)-\beta_0\right)\mathbb{E}[Z|W,|Z|\le h]\right]}{\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\right)^2}.
\end{aligned}
$$

We denote

$$
Bias_0(h)=-\frac{\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\mathbb{E}\left[W\left(\beta_0(X)-\beta_0\right)\mathbb{E}[Z|W,|Z|\le h]\right]}{\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\right)^2}.
$$

By Lemma A3 and bounded convergence theorem, we have

$$
\begin{aligned}
& \lim_{h\to 0}\frac{1}{h^4}\left(\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\left(\beta_0(X)-\beta_0\right)\mathbb{E}[Z|W,|Z|\le h]\right]\right) \\
= & \frac{1}{9}\mathbb{E}\left[W\right]\mathbb{E}\left[W\left(\beta_0(X)-\beta_0\right)\frac{f'_{Z|W}\left(0;W\right)}{f_{Z|W}\left(0;w\right)}\right]
\end{aligned}
$$

and

$$
\lim_{h\to 0}\frac{1}{h^2}\left(\mathbb{E}\left[W\mathbb{E}\left[Z^2|W,|Z|\le h\right]\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[W\mathbb{E}\left[Z|W,|Z|\le h\right]\right]\right)^2\right)=\frac{1}{3}\left(\mathbb{E}[W]\right)^2.
$$

Therefore, we have $Bias_0(h)$ admits

$$
\begin{aligned}
\lim_{h\to 0}\frac{1}{h^2}Bias_0(h) &= \frac{1}{3}\frac{\mathbb{E}\left[W\left(\beta_0(X)-\beta_0\right)\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right]}{\mathbb{E}\left[W\right]} \\
&= \frac{1}{3}\mathbb{E}_W\left[\left(\beta_0(X)-\beta_0\right)\frac{f'_{Z|W}\left(0|W\right)}{f_{Z|W}\left(0|W\right)}\right].
\end{aligned}
$$

For the $\beta_1(X,S)$ term, we first observe that

$$
\begin{aligned}
\mathbb{E}\left[\frac{W}{\pi(W)}\left(\beta_1(X,S)Z\right)\mathbb{I}\{|Z|\le h\}\right] &= \mathbb{E}\left[W\beta_1(X,S)\mathbb{E}[Z|W,|Z|\le h]\right] \text{ and} \\
\mathbb{E}\left[\frac{W}{\pi(W)}\left(\beta_1(X,S)Z^2\right)\mathbb{I}\{|Z|\le h\}\right] &= \mathbb{E}\left[W\beta_1(X,S)\mathbb{E}[Z^2|W,|Z|\le h]\right].
\end{aligned}
$$

By Lemma A3 and bounded convergence theorem, we have

$$\begin{aligned}
\lim_{h\to 0}\frac{1}{h^2}\mathbb{E}\left[W\beta_1(X,S)\mathbb{E}[Z|W,|Z|\le h]\right] &= \frac{1}{3}\mathbb{E}\left[W\beta_1(X,S)\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right], \text{ and}\\
\lim_{h\to 0}\frac{1}{h^2}\mathbb{E}\left[W\beta_1(X,S)\mathbb{E}[Z^2|W,|Z|\le h]\right] &= \frac{1}{3}\mathbb{E}\left[W\beta_1(X,S)\right]=\mathbb{E}[W]\beta_1.
\end{aligned}$$

We denote $Bias_1(h)$ as the term (A. 5), i.e.,

$$\begin{aligned}
& Bias_1(h)\\
=\; & \frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[\frac{W\beta_1(X,S)Z\mathbb{I}\{|Z|\le h\}}{\pi_h(W)}\right]-\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[\frac{W\beta_1(X,S)Z^2\mathbb{I}\{|Z|\le h\}}{\pi_h(W)}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z|\le h\}\right]\right)}.
\end{aligned}$$

we have

$$\begin{aligned}
\lim_{h\to 0}\frac{1}{h^2}Bias_1(h) &= \frac{\frac{1}{9}\mathbb{E}[W]\mathbb{E}\left[W\beta_1(X,S)\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right]-\frac{1}{9}\mathbb{E}\left[W\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right]\mathbb{E}[W]\beta_1}{\frac{1}{3}\left(\mathbb{E}[W]\right)^2}\\
&= \frac{\frac{1}{3}\mathbb{E}\left[W\left(\beta_1(X,S)-\beta_1\right)\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right]}{\mathbb{E}[W]}\\
&= \frac{1}{3}\mathbb{E}_W\left[\left(\beta_1(X,S)-\beta_1\right)\frac{f'_{Z|W}(0|W)}{f_{Z|W}(0|W)}\right].
\end{aligned}$$

Finally ,For the term $b\left(Z^2;X,S\right)$, we have

$$b\left(Z^2;X,S\right)=\frac{1}{2}\beta_2(X,S)Z^2+o(Z^2).$$

Furthermore, we have

$$\mathbb{E}\left[\frac{W}{\pi(W)}\left(b\left(Z^2|X,S\right)\right)\mathbb{I}\{|Z|\le h\}\right]=\frac{1}{2}\mathbb{E}\left[\mathbb{E}\left[\left(W\beta_2(X,S)Z^2+o(Z^2)\right)|W,|Z|\le h\right]\right].$$

By Lemma A3 and bounded convergence theorem, we have

$$\lim_{h\to 0}\frac{1}{h^2}\mathbb{E}\left[\frac{W}{\pi_h(W)}\left(b\left(Z^2|X,S\right)\right)\mathbb{I}\{|Z|\le h\}\right]=\frac{1}{6}\mathbb{E}\left[W\beta_2(X,S)\right],$$

and

$$\lim_{h\to 0}\frac{1}{h^2}\mathbb{E}\left[\frac{W}{\pi_h(W)}\left(Zb\left(Z^2|X,S\right)\right)\mathbb{I}\{|Z|\le h\}\right]=0.$$

Therefore, by denoting

$$Bias_b(h)=\frac{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[\frac{Wb\left(Z^2|X,S\right)\mathbb{I}\{|Z|\le h\}}{\pi_h(W)}\right]-\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[\frac{Wb\left(Z^2|X,S\right)\mathbb{I}\{|Z|\le h\}}{\pi_h(W)}\right]}{\mathbb{E}\left[\frac{W}{\pi_h(W)}Z^2\mathbb{I}\{|Z|\le h\}\right]\mathbb{E}\left[W\right]-\left(\mathbb{E}\left[\frac{W}{\pi_h(W)}Z\mathbb{I}\{|Z|\le h\}\right]\right)},$$

we have

$$\lim_{h\to 0}\frac{1}{h^2}Bias_b(h)=\frac{1}{6}\mathbb{E}_W[\beta_2(X,S)].$$

Thus, the total bias is

$$Bias(h) = Bias_0(h) + Bias_1(h) + Bias_b(h)$$

$$\begin{aligned}
& \lim_{h\to 0} \frac{1}{h^2} Bias(h) \\
= & \frac{1}{3}\mathbb{E}_W \left[((\beta_0(X) - \beta_0) + (\beta_1(X,S) - \beta_1)) \frac{f'_{Z|W}(0;w)}{f_{Z|W}(0;w)} \right] + \frac{1}{2}\mathbb{E}_W[\beta_2(X,S)].
\end{aligned}$$

We finish the bias part.

(2) We then deal with the variance part. Recall that

$$\begin{aligned}
Y &= \beta_0(X) + \beta_1(X,S)Z + b\left(Z^2|X,S\right) + \epsilon\left(Z|X,S\right), \text{ and} \\
\hat{\beta}_0 &= \frac{\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Y_{i_j}\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j} Y_{i_j}\right)}{\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right)\left(\sum_{j=1}^M \tilde{W}_{i_j}\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)^2}.
\end{aligned}$$

First, note that

$$\lim_{N\to+\infty} \frac{\frac{1}{N}\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right)}{\frac{1}{N^2}\left(\left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}^2\right)\left(\sum_{j=1}^M \tilde{W}_{i_j}\right) - \left(\sum_{j=1}^M \tilde{W}_{i_j} Z_{i_j}\right)^2\right)} = \frac{1}{\mathbb{E}[W]}.$$

Then for any $h > 0$, we have

$$\begin{aligned}
& h\text{Var}\left(\tilde{W}_i Y_i \mathbb{I}\{|Z_i| \le h\} | \{Z_i, W_i\}_{i=1}^N\right) \\
= & h\text{Var}\left(\tilde{W}_i\left(\beta_0(X_i) + \beta_1(X_i,S_i)Z_i + b\left(Z_i^2|X_i,S_i\right) + \epsilon\left(Z_i|X_i,S_i\right)\right)\mathbb{I}\{Z_i \le h\}|Z_i, W_i\right) \\
= & h\text{Var}\left(\tilde{W}_i\left(\beta_0(X_i) + \epsilon\left(Z_i|X_i,S_i\right)\right)\mathbb{I}\{Z_i \le h\}|Z_i, W_i\right) + o_p(1) \\
= & \frac{W_i^2 \mathbb{I}\{Z_i \le h\}\mathbb{E}\left[\left(\beta_0(X_i) - \beta_0 + \epsilon\left(Z_i|X_i,S_i\right)\right)^2 |Z_i, W_i\right]}{\pi_h(W_i)^2} h + o_p(1)
\end{aligned}$$

and

$$h\text{Var}\left(\tilde{W}_i Z_i Y_i \mathbb{I}\{|Z_i| \le h\} | \{Z_i, W_i\}_{i=1}^N\right) = o(1).$$

Therefore, we have

$$\begin{aligned}
& \frac{h}{N}\text{Var}\left(\sum_{j=1}^M \tilde{W}_{i_j} Y_{i_j} | \{Z_i, W_i\}_{i=1}^N\right) \\
\to & \mathbb{E}\left[\frac{W_i^2 \mathbb{I}\{Z_i \le h\}\mathbb{E}\left[\left(\beta_0(X_i) - \beta_0 + \epsilon\left(Z_i|X_i,S_i\right)\right)^2 |Z_i, W_i\right]}{\pi_h(W_i)^2}\right] h + o_p(1) \\
= & \mathbb{E}\left[\frac{W_i^2 \mathbb{I}\{Z_i \le h\}\left(\beta_0(X_i) - \beta_0 + \epsilon\left(Z_i|X_i,S_i\right)\right)^2}{\pi_h(W_i)^2}\right] h + o_p(1)
\end{aligned}$$

$$= \quad \mathbb{E}\left[\frac{W^2\left(\beta_0(X)-\beta_0+\epsilon\left(Z|X,S\right)\right)^2}{\pi_h(W)}\right]h+o_p(1),$$

and

$$\frac{h}{N}\mathrm{Var}\left(\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}Y_{i_j}|\left\{Z_i,W_i\right\}_{i=1}^{N}\right)\to o_p(1).$$

By combining all of the above together, we have

$$\begin{aligned} & Nh\mathrm{Var}\left(\hat{\beta}_0|\left\{Z_i,W_i\right\}_{i=1}^{N}\right) \\ \to \quad & \left(\frac{1}{\mathbb{E}[W]}\right)^2\left(\mathbb{E}\left[\frac{W^2\left(\beta_0(X)-\beta_0+\epsilon\left(Z|X,S\right)\right)^2}{\pi_h(W)}\right]h+o_p(1)\right). \end{aligned}$$

Finally, we have

$$\begin{aligned} & \lim_{h\to 0}h\mathbb{E}\left[\frac{W^2\left(\beta_0(X)-\beta_0+\epsilon\left(Z|X,S\right)\right)^2}{\pi_h(W)}\right] \\ = \quad & \mathbb{E}\left[\frac{W^2\left(\beta_0(X)-\beta_0+\epsilon\left(0|X,S\right)\right)^2}{2f(0|W)}\right]. \end{aligned}$$

Proof of Theorem 3 : Recall that

$$\begin{aligned} \hat{\beta}_0-\beta_0 \quad = \quad & \frac{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\left(Y_{i_j}-\beta_0\right)\right)}{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\right)-\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)^2} \\ & -\frac{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\left(Y_{i_j}-\beta_0\right)\right)}{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\right)-\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)^2}. \end{aligned}$$

Let

$$\begin{aligned} Bias^1(h) \quad &= \quad \mathbb{E}\left[\frac{W}{\pi_h\left(W\right)}\left(Y-\beta_0\right)\mathbb{I}\{|Z|\leq h\}\right] \text{ and} \\ Bias^2(h) \quad &= \quad \mathbb{E}\left[\frac{W}{\pi_h\left(W\right)}\left(Y-\beta_0\right)Z\mathbb{I}\{|Z|\leq h\}\right]. \end{aligned}$$

Then, by Linderberg's CLT and the asymptotic variance derived in Theorem 2, we have

$$\begin{aligned} \frac{\sqrt{Nh_N}}{N}\sum_{j=1}^{M}\left(\tilde{W}_{i_j}\left(Y_{i_j}-\beta_0\right)-Bias^1(h_N)\right) \quad &\to \quad \mathcal{N}\left(0,\mathbb{E}\left[\frac{W^2\left(\beta_0(X)-\beta_0+\epsilon_{X,S}(0)\right)^2}{2f(0|W)}\right]\right), \\ \frac{\sqrt{Nh_N}}{N}\sum_{j=1}^{M}\left(\tilde{W}_{i_j}Z_{i_j}\left(Y_{i_j}-\beta_0\right)-Bias^2(h_N)\right) \quad &\xrightarrow{p} \quad 0. \end{aligned}$$

And we have

$$
\begin{aligned}
\frac{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)}{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\right)-\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)^2} &\to \frac{1}{\mathbb{E}\left[W\right]}, \\
\frac{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)}{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\right)-\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)^2} &\to 0.
\end{aligned}
$$

Finally, by noting that the asymptotic bias admits

$$
\begin{aligned}
&\sqrt{Nh_N}\left(\frac{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)Bias^1(h_N)}{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\right)-\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)^2}\right. \\
+&\left.\frac{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)Bias^2(h_N)}{\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}^2\right)\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}\right)-\left(\frac{1}{N}\sum_{j=1}^{M}\tilde{W}_{i_j}Z_{i_j}\right)^2}-Bias(h_N)\right)\to 0,
\end{aligned}
$$

we complete the proof.

Proof of Proposition 2. The proof is immediate after one observing that $\hat{\beta}_0 \to \beta_0$, and

$$\hat{\epsilon}_i = \beta_0(X) - \hat{\beta}_0 + \epsilon_{X,S}(Z) + O(h).$$

□

Appendix B Additional Numerical Results

Table 6 shows the estimation results without inverse weighting by $\hat{\pi}_h(\cdot)$, where weights are Ws, not $\tilde{W}$s; see Algorithm 4. The confidence interval and the standard error are computed as in Proposition 2 and (15) with $\tilde{W} = W$. see the results are similar to the results in Table 5, which means that the inverse-weighting is not necessary.

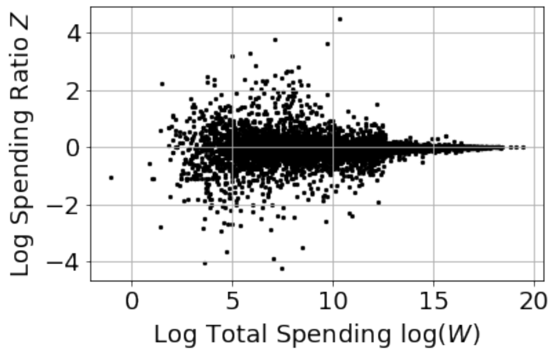
Algorithm 4 Weighted average treatment effect estimation

1: **Input:** The triples of log-CPX difference, log-spending, and weights $\{Y_i, Z_i, W_i\}_{i=1}^N$ of N campaigns, a bandwidth parameter h.
2: **Output:** An estimation of the weighted average treatment effect.
3: Select campaigns $\{Y_{i_j}, Z_{i_j}, W_{i_j}\}_{j=1}^M$ such that $|Z_{i_j}| \le h$ for $j = 1, 2, \ldots, M$.
4: Regress $\{Y_{i_j}\}_{j=1}^M$ on $\{Z_{i_j}\}_{j=1}^M$ with weights $\{W_{i_j}\}_{j=1}^M$. Obtain the intercept $\hat{\beta}_0$.
5: **return** $\hat{\beta}_0$.

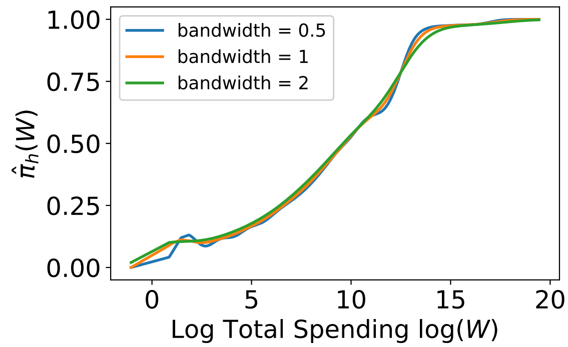
Figure 5 shows the dependence of log spending ratio Z on log total spending $\log(W)$ and estimated $\pi_h(\cdot)$s for CPA, CPM, and CPI goal types. We see similar effects to the one for CPC goal type. Further, We see the estimation is not sensitive the specific bandwidth choice and $\hat{\pi}_h(w)$ is increasing with respect to w.

Table 6: Estimation results via weighted local linear regression without inverse weighting by $\hat{\pi}_h(\cdot)$

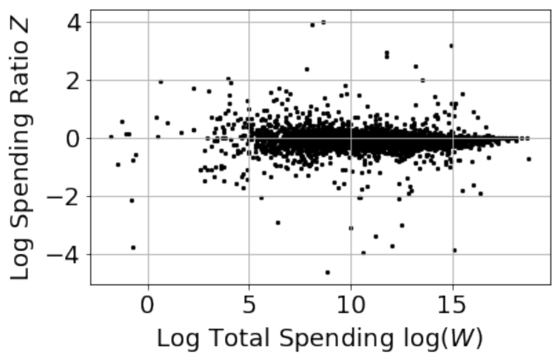
Goal	h	Estimation	Confidence interval	t-statistics
CPC	$\hat{h}_{\mathrm{ROT}}$	-0.016	(-0.020, -0.011)	-9.39
	0.05	-0.013	(-0.016, -0.010)	-7.94
	0.10	-0.014	(-0.017, -0.011)	-8.45
	0.20	-0.015	(-0.018, -0.011)	-9.09
CPA	$\hat{h}_{\mathrm{ROT}}$	-0.005	(-0.007, -0.004)	-10.15
	0.05	-0.005	(-0.006, -0.004)	-9.18
	0.10	-0.005	(-0.006, -0.004)	-9.89
	0.20	-0.005	(-0.007, -0.004)	-10.17
CPM	$\hat{h}_{\mathrm{ROT}}$	-0.002	(-0.007, 0.003)	-1.13
	0.05	-0.003	(-0.004, -0.002)	-5.62
	0.10	-0.004	(-0.006, -0.003)	-5.60
	0.20	-0.004	(-0.005, -0.002)	-4.39
CPI	$\hat{h}_{\mathrm{ROT}}$	-0.008	(-0.012, -0.005)	-5.42
	0.05	-0.007	(-0.010, -0.004)	-4.44
	0.10	-0.007	(-0.010, -0.004)	-5.02
	0.20	-0.008	(-0.012, -0.005)	-5.30



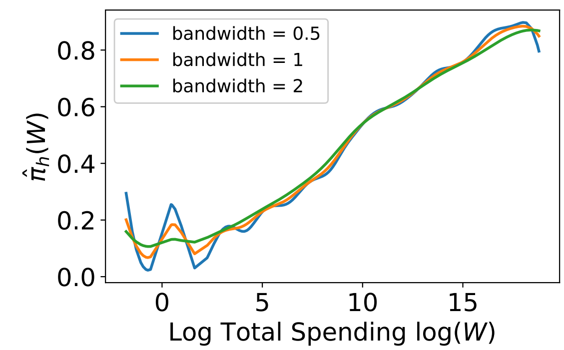
(a) CPA: Log spending ratio v.s. log total spending



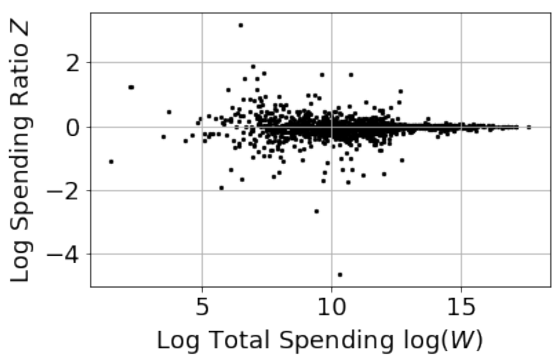
(b) CPA: Estimated $\pi_h(\cdot)$



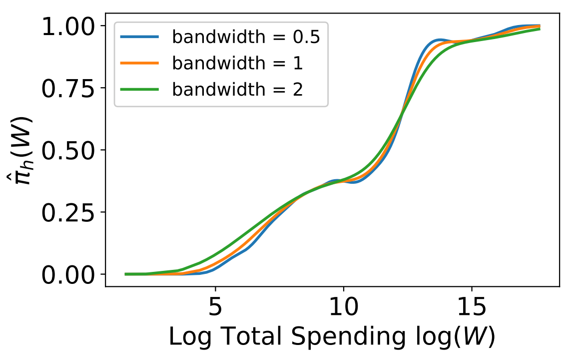
(c) CPM: Log spending ratio v.s. log total spending



(d) CPM: Estimated $\pi_h(\cdot)$



(e) CPI: Log spending ratio v.s. log total spending



(f) CPI: Estimated $\pi_h(\cdot)$

Figure 5: Log spending ratio with respect to log total spending and estimated $\pi_h(\cdot)$ for $h = 0.1$