INTERFERENCE AMONG FIRST-PRICE PACING EQUILIBRIA: A BIAS AND VARIANCE ANALYSIS

Anonymous authors
Paper under double-blind review

ABSTRACT

A/B testing is widely used in the internet industry. For online marketplaces (such as advertising markets), standard approaches to A/B testing may lead to biased results when buyers have budget constraints, as budget consumption in one arm of the experiment impacts performance of the other arm. This is often addressed using a budget-split design. Yet such splitting may degrade statistical performance as budgets become too small in each arm. We propose a *parallel budget-controlled A/B testing* design where we use market segmentation to identify submarkets in the larger market, and we run parallel budget-split experiments in each submarket. We demonstrate the effectiveness of this approach on real experiments on advertising markets at Meta. Then, we formally study interference that derives from such experimental designs, using the first-price pacing equilibrium framework as our model of market equilibration. We propose a debiased surrogate that eliminates the first-order bias of FPPE, and derive a plug-in estimator for the surrogate and establish its asymptotic normality. We then provide an estimation procedure for submarket parallel budget-controlled A/B tests. Finally, we present numerical examples on semi-synthetic data, confirming that the debiasing technique achieves the desired coverage properties.

1 INTRODUCTION

Online A/B testing is widely used in the internet industry to inform decisions on new feature rollouts. For online marketplaces (such as advertising markets), standard approaches to A/B testing may lead to biased results when buyers operate under a budget constraint, as budget consumption in one arm of the experiment impacts performance of the other arm. To counteract this interference, one can use a budget-split design where the budget constraint operates on a per-arm basis and each arm receives an equal fraction of the budget, leading to "budget-controlled A/B testing," see e.g. Basse et al. (2016); Liu et al. (2021).

Despite clear advantages of budget-controlled A/B testing, companies are extremely constrained by the number of such experiments they can run. While it's possible to create more budget splits, this will lower the budget per group substantially, which could lead to different equilibrium outcomes and may disproportionately affect smaller buyers. Additionally, a common approach to increase experimentation throughput is to run orthogonal experiments (with their own orthogonal randomization), but this would either suffer from the same interference as the vanilla A/B test setup, or also require further budget splits.

In this paper, we propose a *parallel budget-controlled A/B test* design where we use market segmentation to identify submarkets in the larger market, and we run parallel experiments on each submarket. When the overall market can be divided into several relatively isolated submarkets, budget-controlled A/B tests can be conducted in parallel within these submarkets. However, this method also presents some challenges. First, submarkets are rarely completely isolated; certain items may attract buyers from multiple submarkets, resulting in interference across submarkets when conducting tests in parallel. Second, submarkets differ in terms of buyer (and user) composition, which might cause the local treatment effect estimates to not be representative of the global treatment effect where all buyers are included in the market. The second challenge is relatively easy to address in practice by imposing balancing constraints in the clustering algorithm used to define submarkets, while the first challenge is more fundamentally important and requires deeper understanding.

Before the theoretical exposition, we consider a comparison of results for paired experiments between a parallel budget-controlled A/B test setup, and that of a traditional budget-split design; where the latter is considered the gold standard. Sec. 1 on the left shows comparisons of 99 experiments where the point estimate and CIs are plotted on the vertical axis for the parallel design, and on the horizontal axis for the budget-split design. The most important feature is whether the two experiments agree between (negative, neutral, positive), as a change would result in a launch reversal. The two experiment designs agree in 75% of cases (at 90% confidence level, hence the optimal agreement is 81.5%), which increases to 79% after the introducing a guardrail metric, see Sec. 1 on the right. These results are quite satisfactory, but do point at the existence of remaining interference bias. In the remainder of this paper, correcting the interference bias is the main objective.

Contributions:

1. We formally define market interference in first-price auction markets using the first price pacing equilibrium (FPPE) framework (Conitzer et al., 2022a) (Sec. 3)
2. We propose a debiased surrogate that eliminates the first-order bias of FPPE, and derive a plug-in estimator for the surrogate and establish its asymptotic normality. (Sec. 4)
3. We run semi-synthetic experiments, confirming that the debiasing technique achieves the desired coverage properties. (Sec 5).

1.1 Parallel A/B Testing in Practice

In this section we describe the real-world problem of A/B testing with congestion that we wish to model, and our proposed solution of parallel A/B tests in carefully balanced submarkets. We start by describing the market environment. There is a set of n advertisers, and each advertiser i has a budget b_i. Whenever a user shows up on the platform an *impression opportunity* occurs, and an auction is conducted in order to determine which ad will be shown to the user. Each advertiser i has some stated value $v_i(\theta)$ of being shown to a particular impression opportunity θ. The advertiser submits a bid which is determined based on $v_i(\theta)$, as well as the expenditure of the advertiser so far. For example, in *multiplicative pacing* Balseiro et al. (2017); Conitzer et al. (2022b), the platform adaptively learns a *pacing multiplier* $\beta_i \in [0, 1]$ such that the bid is formed as $\beta_i v_i(\theta)$. The budget-management system then adaptively controls β_i over time, in order to ensure the correct rate of budget expenditure on behalf of the advertiser. We consider first-price auctions, which is the predominant way display advertising is sold online.

Now that we have discussed budget management, we describe the A/B testing problem. Suppose that a platform wants to run K A/B tests, which may affect, e.g., the valuations that advertisers have for impression slots, revenue, etc. We construct the market-segmented experimental setup as follows: We first define a bipartite graph between advertisers and users based on targeting criteria. Subsequently, we cluster the advertisers into K clusters, where the objective is to minimize the sum of weighted edges between clusters subject to traffic balancing constraints to make the resulting

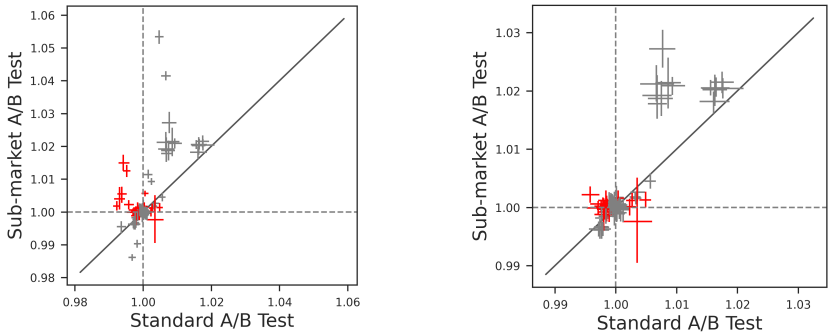

Figure 1: Parallel vs. standard budget-controlled A/B test, daily treatment effect. We denote neutral treatment effects with a value of 1.0. Red crosses indicate instances of sign inconsistencies. Left are all datapoints, on the right, datapoints that fail a guardrail metric are removed.

clusters as similar to the whole market as possible. The edge weight between a pair of advertisers is the number of impressions (or users) where they are both within the top-k bids. The choice of k is a parameter that must be chosen based on experience with the specific application setting. If the clustering achieves a small objective function value, then each cluster is a mostly isolated submarket, in the sense that each user will mostly receive bids from advertisers in a single cluster. Then, we run an A/B test within each of the K submarkets. Every user is randomly assigned to either "A" or "B" in each submarket. The main challenge is that while submarkets are relatively isolated, there is remaining interference from users who are targeted by advertisers from different submarkets, leading to a slightly different equilibrium. Our main contribution is to define a framework for analyzing such interference, and giving an estimator that removes the bias from these users. We survey related works in App. D.

2 Review of FPPE Theory

Notation. For a measurable space $(\Theta, \mathrm{d}\theta)$, we let L^p (and L^p_+, resp.) denote the set of (nonnegative, resp.) L^p functions on Θ w.r.t the base measure $\mathrm{d}\theta$ for any $p \in [1, \infty]$ (including $p = \infty$). Given $x \in L^\infty$ and $v \in L^1$, we let $\langle v, x \rangle = \int_\Theta v(\theta)x(\theta)\,\mathrm{d}\theta$. We treat all functions that agree on all but a measure-zero set as the same. For a sequence of random variables $\{X_n\}$, we say $X_n = O_p(1)$ if for any $\epsilon > 0$ there exists a finite M_ϵ and a finite N_ϵ such that $\mathbb{P}(|X_n| > M_\epsilon) < \epsilon$ for all $n \geq N_\epsilon$. We say $X_n = o_p(1)$ if X_n converges to zero in probability. For a subset $\Theta' \subset \Theta$, let $1_{\Theta'}(\cdot) : \Theta \to \{0, 1\}$ be the indicator function of Θ'. Convergence in distribution and probability is denoted by $\xrightarrow{d}$ and $\xrightarrow{p}$. Given a vector $a = [a_1, \ldots, a_n]^\intercal$, let $\mathrm{Diag}(a)$ denote the diagonal matrix with (i, i)-th entry being a_i; sometimes we write $\mathrm{Diag}(a_i)$ when it is convenient to define each a_i inline. Let $\boldsymbol{A}^\dagger$ denote the Moore–Penrose inverse of the matrix $\boldsymbol{A}$, $\boldsymbol{e}_j$ the j-th unit vector, and $[n] = \{1, \ldots, n\}$.

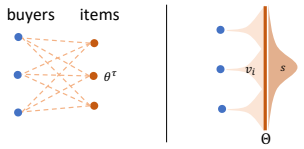

Figure 2: Finite FPPE (left) and limit FPPE (right). In a finite FPPE, there are a finite number of items. In a limit FPPE, the item set is a continuum.

Limit FPPE. We first introduce our notion of a limit market and two regularity conditions on the market, which models the underlying market structure that we sample from. We have n buyers and a possibly continuous set of items Θ with an integrating measure $\mathrm{d}\theta$. For example, one could take $\Theta = [0, 1]$ and $\mathrm{d}\theta =$ the Lebesgue measure on $[0, 1]$. Each buyer has a *budget* b_i; let $\boldsymbol{b} = (b_1, \ldots, b_n)$. The *valuation* for buyer i is a function $v_i \in L^1_+$ such that buyer i has valuation $v_i(\theta)$ for one unit of item $\theta \in \Theta$; let $\boldsymbol{v} : \Theta \to \mathbb{R}^n$, $\boldsymbol{v}(\theta) = [v_1(\theta), \ldots, v_n(\theta)]$. We assume $\bar{v} = \max_i \sup_\theta v_i(\theta) < \infty$. The *supplies* of items are given by a function $s \in L^\infty_+$, i.e., item $\theta \in \Theta$ has $s(\theta)$ units of supply. Without loss of generality, we assume a unit total supply $\int_\Theta s\,\mathrm{d}\theta = 1$. Given $g : \Theta \to \mathbb{R}$, we let $\mathbb{E}[g] = \int g(\theta)s(\theta)\,\mathrm{d}\theta$ and $Var[g] = \mathbb{E}[g^2] - (\mathbb{E}[g])^2$. Given t i.i.d. draws $\{\theta^1, \ldots, \theta^t\}$ from s, let $P_t g(\cdot) = \frac{1}{t}\sum_{\tau=1}^t g(\theta^\tau)$.

Next we introduce the market equilibrium concept that is the foundation of our study. For that we leverage the *first-price pacing equilibrium* (FPPE) (Conitzer et al., 2022a). FPPE models equilibrium outcomes under budget-management systems employed in several practical settings. Each buyer is assigned a *pacing multiplier* β_i, which is used to control their budget expenditure. For each individual auction θ, the buyer bids $\beta_i v_i(\theta)$, which can be seen as their adjusted valuation after factoring in their budget constraint (in practice, the valuation $v_i(\theta)$ may not be the buyer's true valuation, but instead their unscaled bid reported to the platform). The goal of the budget management system is to achieve an equilibrium, meaning that it must ensure that buyers spend their budget exactly by appropriately choosing β_i. If the budget cannot be fully spent, then no pacing must occur (i.e., $\beta_i = 1$). Below we formally define the pacing equilibrium concept in the continuous setting (see Fig 2, right), and give the finite version in the next section (Fig 2, left).

Definition 1 (Limit FPPE, Gao & Kroer (2022); Conitzer et al. (2022a))**.** *A limit FPPE, denoted as* $\mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s, \Theta)$*, is the unique tuple* $(\boldsymbol{\beta}, p(\cdot)) \in [0,1]^n \times L^1_+(\Theta)$ *such that there exist an allocation* $x_i : \Theta \to [0,1]$ *for all* $i \in [n]$ *such that,*

1. *(First-price) Prices and allocations are determined by first price auctions: for all items* $\theta \in \Theta$*,* $p(\theta) = \max_i \beta_i v_i(\theta)$*, and only the highest bidders obtain items: for all* i *and* θ*,* $x_i(\theta) > 0$ *implies* $\beta_i v_i(\theta) = \max_k \beta_k v_k(\theta)$
2. *(Feasiblity, market clearing) Budget are respected: for all* i*,* $\int x_i(\theta) p(\theta) s(\theta)\, \mathrm{d}\theta \leq b_i$*. There is no overselling: for all* θ*,* $\sum_{i=1}^n x_i(\theta) \leq 1$*. Items with nonzero price are fully allocated: for all* θ*,* $p(\theta) > 0$ *implies* $\sum_{i=1}^n x_i(\theta) = 1$*.*
3. *(No unnecessary pacing) For all* i*,* $\int x_i(\theta) p(\theta) s(\theta)\, \mathrm{d}\theta < b_i$ *implies* $\beta_i = 1$*.*

Let $\boldsymbol{\beta}^*$ and p^* be the equilibrium pacing multipliers and prices. Revenue in the limit FPPE is $REV^* = \int p^*(\theta) s(\theta)\, \mathrm{d}\theta$. It measures the profitability of the auction platform. The leftover budgets for buyers are denoted by $\delta_i^* = b_i - \int p^*(\theta) s(\theta) x_i^*(\theta)\, \mathrm{d}\theta$.

The first two conditions simply describe the possible outcomes of a first-price auction system that uses pacing as the budget management strategy. The last condition, no unnecessary pacing, ensures that we only scale down a buyer's bids in case their budget constraint is binding. FPPE has many nice properties, including that they are competitive equilibria and that they are revenue-maximizing among budget-feasible pacing multipliers (Conitzer et al., 2022a).

In a pacing auction market $\mathscr{M} = \mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s, \Theta)$ the following two regularity conditions are important for the study of its statistical properties.

Definition 2 (SMO)**.** *We say the smoothness condition holds if the map* $\boldsymbol{\beta} \mapsto \mathbb{E}_s[\max_i \beta_i v_i(\theta)]$ *is twice continuously differentiable in a neighborhood of* $\boldsymbol{\beta}^*$*.*

Definition 3 (SCS)**.** *We say strict complementary slackness holds if, whenever a buyer is unpaced* ($\beta_i^* = 1$)*, then her leftover budget is strictly positive* ($\delta_i^* > 0$)*.*

The condition SMO ensures that in the limit market, items that incur a tie are measure zero. The condition SCS rules out degenerate buyers that spend their budget exactly at β_i^*. See Liao & Kroer (2023) and Liao et al. (2023) for an extensive discussion about these conditions.

Finite FPPE. Next we introduce the finite FPPE, which models the auction data we observe in practice. Let $\gamma = (\theta^1, \ldots, \theta^t)$ be a sequence of items. Assume each item has the same supply of $\sigma \in \mathbb{R}_+$ units. A finite FPPE is a limit FPPE where the supply is a discrete measure supported on the observed items γ. Let $v_i^\tau = v_i(\theta^\tau)$.

Definition 4 (Finite FPPE, informal)**.** *A finite FPPE,* $\widehat{\mathsf{FPPE}}(\boldsymbol{b}, \boldsymbol{v}, \sigma, \gamma)$*, is a limit FPPE where the item set is the finite set of observed items* γ*. See App. C for the full definition.*

In Liao & Kroer (2023), it is shown that if γ consists of t i.i.d. draws from distribution s, and one takes $\sigma = 1/t$, then the pacing multiplier in $\widehat{\mathsf{FPPE}}(\boldsymbol{b}, \boldsymbol{v}, 1/t, \gamma)$ converge to the pacing multiplier in $\mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s, \Theta)$ in probability. Also, note that the FPPE $\widehat{\mathsf{FPPE}}(t\boldsymbol{b}, \boldsymbol{v}, 1, \gamma)$ converges to the same limit FPPE $\mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s, \Theta)$ because the pacing multipliers of a finite FPPE do not change when budgets and supplies are multiplied by the same scalar.

The Eisenberg-Gale Program. Both the limit FPPE and the finite FPPE have convex program characterizations (Chen et al., 2007; Conitzer et al., 2022a; Gao & Kroer, 2022). We define the dual Eisenberg-Gale (EG) objective for a single item θ as

$$F(\theta, \boldsymbol{\beta}) = f(\theta, \boldsymbol{\beta}) - \sum_{i=1}^n b_i \log \beta_i \ , \quad f(\theta, \boldsymbol{\beta}) = \max_{i \in [n]} \beta_i v_i(\theta) \ . \tag{1}$$

The population and sample (dual) EG objectives are then defined as

$$H(\boldsymbol{\beta}) = \mathbb{E}[F(\theta, \boldsymbol{\beta})] \ , \quad H_t(\boldsymbol{\beta}) = \frac{1}{t} \sum_{\tau=1}^t F(\theta^\tau, \boldsymbol{\beta}) \ . \tag{2}$$

We say $\boldsymbol{H}$ is the Hessian of market $\mathscr{M}$ when $\boldsymbol{H} = \nabla^2_{\boldsymbol{\beta}\boldsymbol{\beta}} \int F(\theta, \boldsymbol{\beta}) s \, \mathrm{d}\theta|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}$.

The equilibrium pacing multipliers $\boldsymbol{\beta}^*$ in $\mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s, \Theta)$ can be recovered through the population dual EG program

$$\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta} \in (0,1]^n} H(\boldsymbol{\beta}) \ . \tag{3}$$

The pacing multiplier vector $\boldsymbol{\beta}^*$ is the unique solution to Eq. (3). Let $\boldsymbol{\beta}^\gamma$ be the equilibrium pacing multiplier in $\widehat{\mathsf{FPPE}}(\boldsymbol{b}, \boldsymbol{v}, 1/t, \gamma)$. Then $\boldsymbol{\beta}^\gamma$ solves the sample analogue of Eq. (3):

$$\boldsymbol{\beta}^\gamma = \arg\min_{\boldsymbol{\beta} \in (0,1]^n} H_t(\boldsymbol{\beta}) \ . \tag{4}$$

Let us briefly comment on the differential structure of f, since it plays a role in later sections. The function $f(\boldsymbol{\beta}, \theta)$ is a convex function of $\boldsymbol{\beta}$ and its subdifferential $\partial_{\boldsymbol{\beta}} f(\boldsymbol{\beta}, \theta)$ is the convex hull of $\{v_i \boldsymbol{e}_i \ : \text{index } i \text{ such that } \beta_i v_i(\theta) = \max_k \beta_k v_k(\theta)\}$, with $\boldsymbol{e}_i$ being the base vector in $\mathbb{R}^n$. When $\max_i \beta_i v_i(\theta)$ is attained by a unique i^*, the function $f(\cdot, \theta)$ is differentiable. In that case, all entries of $\nabla_{\boldsymbol{\beta}} f(\boldsymbol{\beta}, \theta)$ are zero expect that the i^*-th entry is filled with the value $v_{i^*}(\theta)$.

3 Interference as Contamination

In this section, we discuss how to estimate market equilibria when there is *contamination* in the supply, meaning that items are generated from a mixture of two distributions, when in reality we wish to estimate equilibrium quantities from one of the two distributions. Then, we show that interference from other markets can be viewed as a form of contamination, and so the problem of removing interference bias can be analyzed via our contamination framework.

3.1 FPPE with Contaminated Supply

We assume that we are in the same FPPE setting as before: there are n buyers, each with budget b_i, and an item set Θ which is now partitioned in to Θ_{bad} and Θ_{good}. However, now we assume that the supply s is contaminated by s', another supply distribution. We define the α-contaminated market as $\mathscr{M}_\alpha = \mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s_\alpha, \Theta)$ and the uncontaminated market as $\mathscr{M}_0 = \mathsf{FPPE}(\boldsymbol{b}, \boldsymbol{v}, s, \Theta)$, where $s_\alpha = \alpha s' + (1-\alpha)s$, distribution s' is supported on Θ_{bad}, and s on Θ_{good}. Our goal is to perform inference about FPPE properties in the limit FPPE with the supply s. However, we are given access to finite FPPEs sampled from s_α instead. In particular, let γ be t i.i.d. draws from s_α and let $\widehat{\mathscr{M}_\alpha} = \widehat{\mathsf{FPPE}}(\boldsymbol{b}, \boldsymbol{v}, 1/t, \gamma)$. We assume α is known throughout the paper. In practice, this can often be estimated from historical data; in the parallel A/B test setting, this can be estimated directly from the sampled set of items, since we know whether an item is drawn from s or s'. Let $\boldsymbol{\beta}^*$ and $\boldsymbol{\beta}^*_\alpha$ be the limit pacing multipliers in $\mathscr{M}_0$ and $\mathscr{M}_\alpha$, respectively. Let $\boldsymbol{\beta}^\gamma_\alpha$ be the pacing multipliers in the sampled market $\widehat{\mathscr{M}_\alpha}$ and let $H_{\alpha,t}(\boldsymbol{\beta}) = \frac{1}{t}\sum_{\tau=1}^t F(\theta^\tau, \boldsymbol{\beta})$ be the sample EG objective.

If we wanted to make inferences about $\mathscr{M}_\alpha$ then we could use existing statistical inference theory on how to use data in a finite FPPE $\widehat{\mathscr{M}_\alpha}$ to make inferences about the limit FPPE $\mathscr{M}_\alpha$ (Liao & Kroer, 2023). However, the supply contamination prevents the application of these tools to our problem.

Our central research question is then on how to use data in the finite contaminated market $\widehat{\mathscr{M}_\alpha}$ to make inferences about the uncontaminated limit market $\mathscr{M}_0$.

In Sec. 4 we propose an estimator for this problem and derive its properties. The results there apply to general item space Θ and supplies s and s'. By imposing structure on Θ, s and s', we show that the contamination model captures the interference among FPPEs.

3.2 Application: Modeling Interference among FPPEs

Now we show how the contamination model from the previous section can be used to model interference. Consider K separate auction markets, which together form a global market. In the global market there are n buyers, each with budget b_i, and an item set Θ, partitioned into Θ_{good} and Θ_{bad}. Let $C_1, \ldots, C_K$ be a partition of the buyers, $\Theta_1, \ldots, \Theta_K$ be a partition of the good item set Θ_{good},

and $s_1, \dots s_K$ be a set of supply functions, supported on $\Theta_1, \dots, \Theta_K$ respectively. The k-th submarket consists of buyers in C_k, the item set Θ_k and supply s_k. Let $s = \frac{1}{K}\sum_k s_k$ be the average mixture and s' be a supply supported on Θ_{bad}. Let the contaminated supply be $s_\alpha = \alpha s' + (1-\alpha)s$.

By imposing structure on Θ_{good} and Θ_{bad} the contamination model can capture interference among auction markets. We assume that submarkets are separated, which models the ideal case where there is no interference. A buyer $i \in C_k$ is only interested in items from the submarket he belongs to: $v_i(\theta) = 0$ for $\theta \in \Theta_{k'}, k' \neq k$. Next, we let Θ_{bad} represent items that cause outbound edges from submarkets; see the green edges in Fig 3 left panel. An item is referred to as *bad* if it has positive values for buyers from at least two different submarkets. Formally, $\theta \in \Theta_{\text{bad}}$ if there exist $i \in C_k$, $j \in C_{k'}$, $k \neq k'$, such that $v_i(\theta) > 0$ and $v_j(\theta) > 0$. Combining these assumptions, we have that a buyer i from submarket k has positive values *only* for items from the sets Θ_k and possibly Θ_{bad}. Now we have fully specified a contaminated market setup: we wish to make inferences on the market consisting of only Θ_{good} (which is really K fully separate submarkets), but we observe an actual market containing items from $\Theta_{\text{good}} \cup \Theta_{\text{bad}}$. With this setup, we can use the results developed in the following section to model interference in parallel submarkets.

In Fig 3 we present an example of interference among $K = 3$ submarkets. The market of interest is the perfectly separated market (right). This is because, in parallel A/B testing, submarkets are explicitly created such that each submarket resembles the global market. Then when a submarket receives a treatment, the observed quantities in that submarket, such as revenues and social welfare, are considered surrogates for the treatment effect in the global market. However, in practice we only observe the interfered finite market (left), which converges to the interfered limit market (middle). In App. F we show how to analyze parallel A/B testing using this framework.

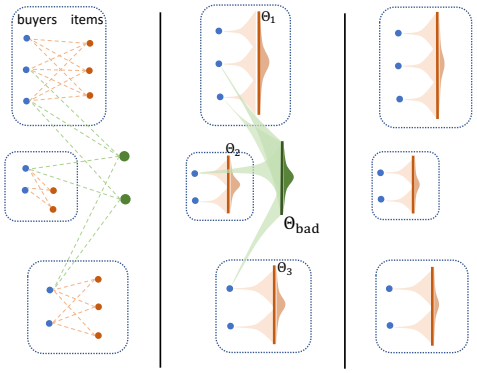

Figure 3: The model. Left ($\widehat{\mathscr{M}_\alpha}$): the observed market where interference is present among submarkets. Middle ($\mathscr{M}_\alpha$): the limit market with interference from bad item set. Right ($\mathscr{M}_0$): the limit market with perfectly separated submarkets. We use data from the left panel to make inferences about the market in the right panel.

4 A Debiased Estimator and Its Properties

This section develops a methodology for making inferences about the uncontaminated limit FPPE. Since the interference setting is a special case of the contamination setting, we develop theories for the latter. We introduce a surrogate for pacing multipliers, based on the notion of directional derivatives, and establish its debiasing property in Sec. 4.1. Then, we focus on estimating this surrogate quantity in Sec. 4.2, and develop asymptotic normality results in Sec. 4.3. Secondly, we consider estimating revenue, which can be thought of as a smooth function of pacing multipliers. We discuss debiased revenue estimation and inference based on our pacing multiplier results in App. E.

4.1 A Debiased Surrogate for Pacing Multipliers

If we view $\boldsymbol{\beta}^*_\alpha$ as a function of the level of contamination α, then one can imagine that under sufficient regularity conditions, the pacing multipliers in the perfectly separated market, $\boldsymbol{\beta}^*_0$, can be

approximated by some form of Taylor expansion of $\alpha \mapsto \boldsymbol{\beta}^*_\alpha$ at α. This can be made rigorous by the notion of directional derivatives. We define

$$\mathrm{d}\boldsymbol{\beta}^*(\alpha) = \lim_{\epsilon \to 0^+} \frac{\boldsymbol{\beta}^*_{\alpha-\epsilon} - \boldsymbol{\beta}^*_\alpha}{\epsilon} \ . \tag{5}$$

if the limit exists. We will show in Thm. 1 that $\boldsymbol{\beta}^*_\alpha + \alpha\, \mathrm{d}\boldsymbol{\beta}^*(\alpha)$ serves as a good approximation to $\boldsymbol{\beta}^* = \boldsymbol{\beta}^*_0$.

Thanks to the convex program characterization of FPPE, the directional derivative $\mathrm{d}\boldsymbol{\beta}^*(\alpha)$ has a closed-form expression under certain regularity conditions (the conditions are given in Thm. 1; the full proof is given in the appendix). We need a few notations for this expression. Define $\boldsymbol{\delta}_\alpha = \int \nabla f(\theta, \boldsymbol{\beta}^*_\alpha)(s - s')\, \mathrm{d}\theta$. Let $\boldsymbol{H}_\alpha = \nabla^2_{\boldsymbol{\beta}\boldsymbol{\beta}} \int F(\theta, \boldsymbol{\beta}^*_\alpha) s_\alpha \, \mathrm{d}\theta$ be the Hessian matrix in the market $\mathscr{M}_\alpha$ and $\boldsymbol{P}_\alpha = \mathrm{Diag}(1(\boldsymbol{\beta}^*_{\alpha,i} < 1))$. Then, under the regularity conditions given in Thm. 1 below,

$$\mathrm{d}\boldsymbol{\beta}^*(\alpha) = -(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \boldsymbol{\delta}_\alpha \ . \tag{6}$$

We present a heuristic derivation in App. G. Given the closed-form expression of $\mathrm{d}\boldsymbol{\beta}^*(\alpha)$, we define the following debiased pacing multiplier

$$\widetilde{\boldsymbol{\beta}}^* = \boldsymbol{\beta}^*_\alpha + \alpha \cdot (-(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \boldsymbol{\delta}_\alpha) \ . \tag{7}$$

Theorem 1 (Analysis of Bias)**.** *Suppose that in the market $\mathscr{M}_0$ conditions* SMO *and* SCS *hold, and assume that $\boldsymbol{\beta} \mapsto \nabla^2 \int F(\theta, \boldsymbol{\beta}) s' \, \mathrm{d}\theta$ is twice continuously differentiable in a neighborhood of $\boldsymbol{\beta}^*$. Then the directional derivative $\mathrm{d}\boldsymbol{\beta}^*(\cdot)$ is well-defined in a neighborhood of zero, and is given by Eq. (6). Moreover, as $\alpha \downarrow 0$,*

$$\|\widetilde{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^*\|_2 = o(\alpha) \ .$$

The proof is given in App. B.1. Thm. 1 indicates that the debiased surrogate $\widetilde{\boldsymbol{\beta}}^*$ removes first-order bias caused by contamination. The limit pacing multipliers $\boldsymbol{\beta}^*_\alpha$ of the contaminated market $\mathscr{M}_\alpha$ will have bias of order $\boldsymbol{\beta}^*_\alpha - \boldsymbol{\beta}^* = \Theta(\alpha)$. In contrast, Thm. 1 shows that the debiased surrogate only incurs a bias of order $o(\alpha)$.

4.2 The Estimator

In this section we introduce a plug-in estimator for the debiased surrogate $\widetilde{\boldsymbol{\beta}}^*$ and introduce a consistency theorem. The next section discusses constructing confidence intervals.

To estimate $\mathrm{d}\boldsymbol{\beta}^*(\alpha)$ in Eq. (6) we need estimates of its three components: the Hessian $\boldsymbol{H}_\alpha = \nabla^2_{\boldsymbol{\beta}\boldsymbol{\beta}} \int F(\theta, \boldsymbol{\beta}^*_\alpha) s_\alpha \, \mathrm{d}\theta$, the diagonal matrix $\boldsymbol{P}_\alpha$ and the vector $\boldsymbol{\delta}_\alpha = \int \nabla f(\theta, \boldsymbol{\beta}^*_\alpha)(s - s')\, \mathrm{d}\theta$.

The Hessian. For simplicity in our theoretical results, we will simply assume a generic Hessian estimator $\widehat{\boldsymbol{H}}_\alpha$ such that for some $\eta_t \downarrow 0$ we have $\widehat{\boldsymbol{H}}_\alpha - \boldsymbol{H}_\alpha = O_p(\eta_t)$. Hong et al. (2015) discuss the estimation of the derivative in detail. Different kinds of statistical guarantees require different rate conditions on η_t; see Theorems 2 and 3. We then introduce two Hessian estimators: one is applicable for general FPPE, while the other requires an extra market regularity condition. The first Hessian estimator is the finite difference method. Let $\boldsymbol{e}_i, \boldsymbol{e}_j$ be basis vectors and ε_t be a step-size. Then the estimator is

$$\widehat{\boldsymbol{H}}_\alpha[i,j] = [H_{\alpha,t}(\boldsymbol{\beta}^\gamma_{++}) - H_{\alpha,t}(\boldsymbol{\beta}^\gamma_{+-}) - H_{\alpha,t}(\boldsymbol{\beta}^\gamma_{-+}) + H_{\alpha,t}(\boldsymbol{\beta}^\gamma_{--})]/(4\varepsilon_t^2) \ ,$$

where $\boldsymbol{\beta}^\gamma_{\pm\pm} = \boldsymbol{\beta}^\gamma_\alpha \pm \boldsymbol{e}_i \varepsilon_t \pm \boldsymbol{e}_j \varepsilon_t$, and $H_{\alpha,t}(\boldsymbol{\beta}) = \frac{1}{t}\sum_{\tau=1}^t F(\theta^\tau, \boldsymbol{\beta})$, with $\{\theta^\tau\}_\tau$ being the items in $\widehat{\mathscr{M}_\alpha}$. In practice, a diagonal approximation of the Hessian suffices. The second method relies on an additional regularity condition, in which case we derive a simplified formula for the Hessian, thereby enabling a simpler estimation procedure (see Thm. 3).

The vector $\boldsymbol{\delta}_\alpha$. Let g be the Radon-Nikodym ratio $g(\theta) = (\mathrm{d}(s - s')/\,\mathrm{d}s_\alpha)(\theta) = \frac{1}{1-\alpha} 1_{\Theta_{\text{good}}}(\theta) - \frac{1}{\alpha} 1_{\Theta_{\text{bad}}}(\theta)$. With the ratio g, the true vector $\boldsymbol{\delta}_\alpha$ can be written as $\boldsymbol{\delta}_\alpha = \int g(\theta) \nabla f(\theta, \boldsymbol{\beta}^*_\alpha) s_\alpha(\theta)\, \mathrm{d}\theta$, which is easy to estimate given i.i.d. draws from s_α. In particular, our estimator is then $\widehat{\boldsymbol{\delta}}_\alpha = \frac{1}{t}\sum_{\tau=1}^t g(\theta^\tau) \boldsymbol{\mu}^\tau$. Here $\boldsymbol{\mu}^\tau = [x_1^\tau v_1^\tau, \ldots, x_n^\tau v_n^\tau]^\intercal$ is a subgradient of $f(\theta^\tau, \boldsymbol{\beta}^\gamma_\alpha)$ w.r.t. $\boldsymbol{\beta}$.

The diagonal matrix $\boldsymbol{P}_\alpha$. Recall $\boldsymbol{P}_\alpha = \mathrm{Diag}(1(\beta^*_{\alpha,i} < 1))$. So a natural estimator is $\widehat{\boldsymbol{P}}_\alpha = \mathrm{Diag}(1(\beta^\gamma_{\alpha,i} < 1 - \iota_t))$, where the slackness $\iota_t \asymp \frac{1}{\sqrt{t}}$.

With all three components estimated, using $\boldsymbol{\beta}^\gamma_\alpha$ is the pacing multiplier in the market $\widehat{\mathscr{M}_\alpha}$ we define the plug-in estimator for $\mathrm{d}\boldsymbol{\beta}^*(\alpha)$ in Eq. (6) as

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}^\gamma_\alpha - \alpha \cdot (\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger \widehat{\boldsymbol{\delta}}_\alpha \ . \tag{8}$$

Theorem 2 (Consistency)**.** *Suppose that in the market $\mathscr{M}_\alpha$ conditions* SMO *and* SCS *hold. If the Hessian estimation error satisfies $\eta_t = o(1)$, then $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \widetilde{\boldsymbol{\beta}}^*$. The proof is in App. B.2.*

4.3 Asymptotic Normality and Inference

We present two asymptotic normality results. In the first result, we require a stronger condition on the Hessian error rate η_t. In particular, as will be shown in Thm. 3, the rate condition $\eta_t = o(1/\sqrt{t})$ is sufficient for normality. One could use a separate large historical dataset to obtain a good estimate of the Hessian matrix. In the second result, we impose an additional condition on market structure which simplifies the Hessian expression and facilitates efficient Hessian estimation. To describe the additional market structure, we define the gap between the highest and the second-highest bid for an item θ under pacing $\boldsymbol{\beta}$ by $\mathsf{bidgap}(\boldsymbol{\beta}, \theta) = \max\{\beta_i v_i(\theta)\} - \mathrm{secondmax}\{\beta_i v_i(\theta)\}$, where $\mathrm{secondmax}$ is the second-highest entry potentially equal to the highest; e.g., $\mathrm{secondmax}([1, 1, 2]) = 1$. When there is a tie for an item θ under pacing $\boldsymbol{\beta}$, we have $\mathsf{bidgap}(\boldsymbol{\beta}, \theta) = 0$. When there is no tie for an item θ, the gap $\mathsf{bidgap}(\boldsymbol{\beta}, \theta)$ is strictly positive.

For any $g : \Theta \to \mathbb{R}$, let $\mathbb{E}_\alpha[g] = \int g s_\alpha \, \mathrm{d}\theta$ and $Cov_\alpha(g) = \mathbb{E}_\alpha[(g - \mathbb{E}_\alpha[g])(g - \mathbb{E}_\alpha[g])^\intercal]$. Recall $\widetilde{\boldsymbol{\beta}}^*$ is the debiased surrogate in Eq. (7) and $\widehat{\boldsymbol{\beta}}$ is its estimator defined in Eq. (8).

We need to introduce a few more notations to describe the normality results. First, let $\boldsymbol{d}_\alpha = -(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \nabla f(\cdot, \boldsymbol{\beta}^*_\alpha)$. As mentioned previously, the pacing multipliers in the contaminated market converge to the limit counterpart and have the representation

$$\sqrt{t}(\boldsymbol{\beta}^\gamma_\alpha - \boldsymbol{\beta}^*_\alpha) = \frac{1}{\sqrt{t}} \sum_{\tau=1}^{t} (\boldsymbol{d}_\alpha(\theta^\tau) - \mathbb{E}_\alpha[\boldsymbol{d}_\alpha(\theta^\tau)]) + o_p(1) \ .$$

In the statistics literature, the function $\boldsymbol{d}_\alpha(\cdot) - \mathbb{E}[\boldsymbol{d}_\alpha]$ is called the influence function (van der Vaart, 2000). For our debiased estimators, we need the following (uncentered) influence functions.

$$\boldsymbol{d}_1(\theta) = \big(\tfrac{1}{1-\alpha} 1_{\Theta_{\mathsf{good}}}(\theta)\big) \boldsymbol{d}_\alpha(\theta) \ , \ \boldsymbol{d}_2(\theta) = \boldsymbol{d}_1(\theta) - 2\alpha \mathrm{Diag}(\beta^*_{\alpha,i} \delta_{\alpha,i} b_i^{-1}) \boldsymbol{d}_\alpha(\theta) \ .$$

Theorem 3. *Let* SCS *and* SMO *hold in $\mathscr{M}_\alpha$. 1. Asymptotic Normality. It holds that $\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^* = \boldsymbol{z}_t + O_p(\eta_t) + o_p(\frac{1}{\sqrt{t}})$, where $\sqrt{t}\boldsymbol{z}_t \xrightarrow{d} \mathscr{N}(\mathbf{0}, \boldsymbol{\Sigma}_1)$ with $\boldsymbol{\Sigma}_1 = Cov_\alpha(\boldsymbol{d}_1)$, and η_t is the Hessian estimation error. 2. Asymptotic Normality under a Bid Gap Condition. If in addition $\mathbb{E}_\alpha[1/\mathsf{bidgap}(\boldsymbol{\beta}^*_\alpha, \theta)] < \infty$, then $\boldsymbol{H}_\alpha = \mathrm{Diag}(b_i/(\beta^*_{\alpha,i})^2)$. Suppose we estimate $\boldsymbol{H}_\alpha$ with $\widehat{\boldsymbol{H}}_\alpha = \mathrm{Diag}(b_i/(\beta^\gamma_{\alpha,i})^2)$. Then $\sqrt{t}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^*) \xrightarrow{d} \mathscr{N}(0, \boldsymbol{\Sigma}_2)$ where $\boldsymbol{\Sigma}_2 = Cov_\alpha(\boldsymbol{d}_2)$. The proof is in App. B.3.*

Thm. 3 part 1 shows how the error of the Hessian estimate affects the distribution of the estimator $\widehat{\boldsymbol{\beta}}$. If $\eta_t = o(1/\sqrt{t})$, then the decomposition becomes $\sqrt{t}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^*) = \sqrt{t}\boldsymbol{z}_t + o_p(1)$, implying asymptotic normality, i.e. $\sqrt{t}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^*) \xrightarrow{d} \mathscr{N}(\mathbf{0}, \boldsymbol{\Sigma}_1)$, in which case one can construct an ellipsoidal confidence region for $\widetilde{\boldsymbol{\beta}}^*$. Thm. 3 part 2 shows direct asymptotic normality under the extra condition, with a simpler Hessian estimator that avoids finite differences.

To perform inference, we need to construct a consistent estimate of the covariance matrix. Now we describe a plug-in estimate of $\boldsymbol{\Sigma}_1$. Let the estimator $\widehat{\boldsymbol{d}_1}^\tau$ be $\widehat{\boldsymbol{d}_1}^\tau = -\big(\frac{1}{1-\alpha} 1_{\Theta_{\mathsf{good}}}(\theta^\tau)\big)(\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger \boldsymbol{\mu}^\tau$, where $\boldsymbol{\mu}^\tau = [x_1^\tau v_1^\tau, \ldots, x_n^\tau v_n^\tau]^\intercal$, and $\widehat{\boldsymbol{P}}_\alpha, \widehat{\boldsymbol{H}}_\alpha$ have been defined in Sec. 4.2. The plug-in estimator is $\widehat{\boldsymbol{\Sigma}}_1 = \frac{1}{t}\sum_{\tau=1}^{t} (\widehat{\boldsymbol{d}_1}^\tau - \overline{\boldsymbol{d}_1})(\widehat{\boldsymbol{d}_1}^\tau - \overline{\boldsymbol{d}_1})^\intercal$ with $\overline{\boldsymbol{d}_1} = \frac{1}{t}\sum_{\tau=1}^{t} \widehat{\boldsymbol{d}_1}^\tau$. By similar arguments as in Liao et al. (2023), the plug-in estimates of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are consistent. Algorithm 1 summarizes the debiasing procedure.

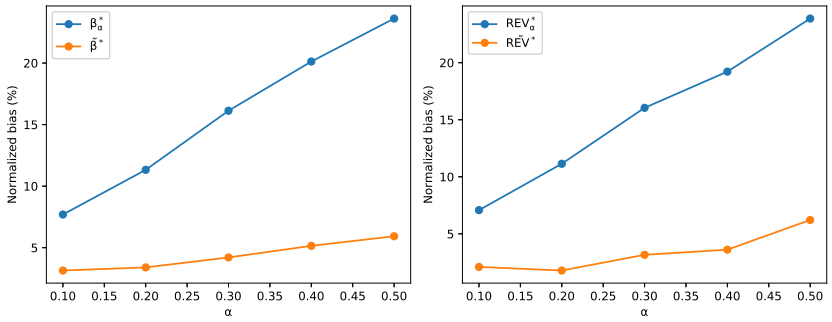

Figure 4: Normalized bias (in percent of true value) as a function of α in semi-synthetic experiments. $\widetilde{\boldsymbol{\beta}}^*$ and $\widetilde{REV}^*$ are the debiased surrogates for pacing multiplier and revenue in the limit market with interference $\mathscr{M}_\alpha$.

In App. E we present a similar debiased estimator for revenue and its bias and variance properties. In App. F we specialize the debiased estimator to parallel budget-controlled A/B testing.

5 Semi-synthetic experiment

To evaluate our proposed framework and debiased estimator, we run semi-synthetic simulations to check if the proposed estimator for beta and revenue are indeed less biased, and we test the coverage of the proposed estimator. Fully synthetic experiments are presented in App. H.

In the semi-synthetic experiments, we simulate 40 buyers and 10000 *good* items in two submarkets, with a varying number of *bad* items (up to 5000) in order to study the effect of the contamination parameter α. For each α, we randomly sample a budget for each buyer, and compute $\boldsymbol{\beta}^*$ and REV^* from the limit pure market $\mathscr{M}_0$ with a value function in each submarket. Both the budget and values are sampled from historical bidding data, making the budget and value distributions heavy-tailed as in the real-world applications. More specifically, we first sample a certain number of auctions. For each auction, we sample a given number of advertisers with their per-impression bids. Advertisers that are sampled across different auctions are treated as the same buyers and their budgets are determined by aggregating their values over auctions up to a scalar to calibrated to get the percentage of budget-constrained buyers equal to what was observed in the real-world auction market, along the same lines as the experiments of Conitzer et al. (2022b).

To check if the debiased surrogate reduces bias, we compute $\boldsymbol{\beta}^*_\alpha$ and REV^*_α from the limit market with interference $\mathscr{M}_\alpha$ and their surrogates, $\widetilde{\boldsymbol{\beta}}^*$ and $\widetilde{REV}^*$ ($\widetilde{\boldsymbol{\beta}}^*$ is defined Eq. (7), $\widetilde{REV}^*$ is defined in App. E. We look at the normalized bias for the surrogate, defined as $\|\widetilde{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^*\|_2 / \|\boldsymbol{\beta}^*\|_2$ for pacing multipliers and as $|\widetilde{REV}^*/REV^* - 1|$ for revenue, and similarly defined for the limit quantities. Fig 4 shows the normalized bias curves as a function of α. The magnitude of the bias increases with α, for both the variables in the limit market with interference $\mathscr{M}_\alpha$ and their debiased surrogates. The bias of the debiased surrogates is indeed much smaller than the contaminated limit quantities.

Next, we check the coverage of the proposed variance estimator. For each α and each budget sample, we run 100 simulations in the following way: We sample items (or their values for each buyer) considering two submarkets and bad items. We then run the finite FPPE with bad items and obtain a baseline estimate for pacing multiplier and revenue without applying the debiasing procedure. Then, we apply the debiasing procedure to compute the debiased estimates. For each simulation, we check if the debiased surrogate is within the confidence interval of the debiased estimator. Finally, we aggregate them to compute the estimated coverage of the estimator. The results for both pacing multiplier and revenue are shown in Table 1. For the coverage of $\widehat{\boldsymbol{\beta}}$, we first compute the coverage of each component and report only the average in the table. For revenue, we construct the CI using the two approaches as mentioned in App. E: one based on Eq. (22) and the other using parametric bootstrap based on the estimated asymptotic distribution of $\widehat{\boldsymbol{\beta}}$ (with "(b)" in the column names).

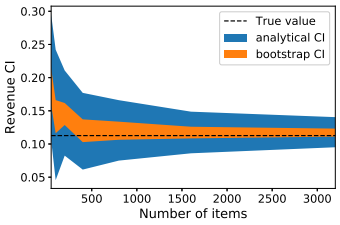

Figure 5: Revenue confidence intervals as a function of the number of items in semi-synthetic experiments. The analytic CI comes from Eq. (22). The true value is the debiased surrogates for revenue in the limit market with interference $\mathscr{M}_\alpha$.

Firstly, Fig 5 shows that both CIs converges to the true value from the limit market with interfrence $\mathscr{M}_\alpha$ as the number of items goes to infinity. Then, in Table 1, we show that the coverage for $\widehat{\boldsymbol{\beta}}$ is slightly smaller than the nominal level (95%), as well as the coverage of the bootstrap CI of revenue. The under-coverage for $\widehat{\boldsymbol{\beta}}$ is mainly driven by the under-estimation of the variance of $\widehat{\boldsymbol{\beta}}$, while the under-coverage of the bootstrap CI for revenue can also be partially attributed to the higher dimensionality (with 40 buyers), making the bootstrap resampling harder to explore the whole space.

Although the proposed variance estimator has good asymptotic properties, the results from our synthetic experiments suggest that it can perform badly, in either direction, for finite markets. Constructing more accurate variance estimators for our debiased estimator in finite settings would certainly mitigate the over- or under-coverage issues that we observe here and deserve more future research. One promising alternative is to construct the CI for $\widehat{\boldsymbol{\beta}}$ and $\widehat{REV}$ by directly bootstrapping the observed value matrix, though this might work best for independent valuations across buyers.

	$\widehat{\boldsymbol{\beta}}$	$\widehat{REV}$			
α	coverage	CI width	CI width (b)	coverage	coverage (b)
1000/11000	0.877	0.244	0.044	1.0	0.95
2000/12000	0.849	0.225	0.043	1.0	0.87
3000/13000	0.852	0.210	0.041	1.0	0.81
4000/14000	0.828	0.200	0.040	1.0	0.90
5000/15000	0.826	0.191	0.039	1.0	0.90

Table 1: Coverage of $\widehat{\boldsymbol{\beta}}$ and revenue estimates in the semi-synthetic experiments. All quantities are averaged over 100 simulations for each α (the ratio of the number of bad items and the total items). The coverage of $\widehat{\boldsymbol{\beta}}$ is averaged over all components of $\widehat{\boldsymbol{\beta}}$. For revenue estimates, the columns with "(b)" represent the quantities from the bootstrap CI, while the columns without "(b)" are for the CI from Eq. (22). The CI widths are normalized by the revenue from the limit market $\mathscr{M}_\alpha$.

6 Conclusion

We have proposed a practical experimental design for performing concurrent A/B tests in large-scale ad auction markets, using a submarket clustering approach, and showed that in production experiments, this submarket clustering approach leads to strong sign consistency performance, as compared to A/B testing on the full market, while allowing significantly-higher A/B test throughput. In order to model the potential for interference between submarket A/B tests, we introduced a theoretical model of statistical inference in first-price pacing equilibrium problems, under settings with supply contamination. We showed how one can perform statistical inference in such a setting using a debiased estimation procedure, and studied the statistical properties of this procedure. We then showed how our model of statistical inference in FPPE with contamination can be used to model the submarket clustering parallel A/B test design, and gave theoretical performance guarantees. Finally, we presented numerical experiments on fully synthetic and semi-synthetic data derived from Meta ad auctions. The experiments showed that our proposed debiased estimator achieves smaller biases and its statistical coverage on realistic data is generally in line with the predictions from our theory.

REFERENCES

Gagan Aggarwal, Ashwinkumar Badanidiyuru, and Aranyak Mehta. Autobidding with constraints. In *Web and Internet Economics: 15th International Conference, WINE 2019, New York, NY, USA, December 10–12, 2019, Proceedings 15*, pp. 17–30. Springer, 2019.

Peter M. Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11 (4):1912 – 1947, 2017. doi: 10.1214/16-AOAS1005. URL https://doi.org/10.1214/16-AOAS1005.

Susan Athey, Dean Eckles, and Guido W. Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018. doi: 10.1080/01621459.2016.1241178. URL https://doi.org/10.1080/01621459.2016.1241178.

Moshe Babaioff, Richard Cole, Jason Hartline, Nicole Immorlica, and Brendan Lucier. Non-quasi-linear agents in quasi-linear mechanisms. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185, pp. 84, 2021.

Santiago Balseiro, Anthony Kim, Mohammad Mahdian, and Vahab Mirrokni. Budget management strategies in repeated auctions. In *Proceedings of the 26th International Conference on World Wide Web*, pp. 15–23, 2017.

Santiago Balseiro, Yuan Deng, Jieming Mao, Vahab Mirrokni, and Song Zuo. Robust auction design in the auto-bidding world. *Advances in Neural Information Processing Systems*, 34:17777–17788, 2021a.

Santiago Balseiro, Anthony Kim, Mohammad Mahdian, and Vahab Mirrokni. Budget-management strategies in repeated auctions. *Operations research*, pp. 859–876, 2021b.

Santiago Balseiro, Christian Kroer, and Rachitesh Kumar. Contextual standard auctions with budgets: Revenue equivalence and efficiency guarantees. *Management Science*, 2023.

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

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.

Guillaume W. Basse, Hossein Azari Soufiani, and Diane Lambert. Randomization and the pernicious effects of limited budgets on auction experiments. In Arthur Gretton and Christian C. Robert (eds.), *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pp. 1412–1420, Cadiz, Spain, 09–11 May 2016. PMLR. URL https://proceedings.mlr.press/v51/basse16b.html.

Christian Borgs, Jennifer Chayes, Nicole Immorlica, Kamal Jain, Omid Etesami, and Mohammad Mahdian. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th international conference on World Wide Web*, pp. 531–540, 2007.

Lihua Chen, Yinyu Ye, and Jiawei Zhang. A note on equilibrium pricing as convex optimization. In *International Workshop on Web and Internet Economics*, pp. 7–16. Springer, 2007.

Xi Chen, Christian Kroer, and Rachitesh Kumar. Throttling equilibria in auction markets. In *Web and Internet Economics - 17th International Conference, WINE Proceedings*, 2021.

Xi Chen, Christian Kroer, and Rachitesh Kumar. The complexity of pacing for second-price auctions. *Mathematics of Operations Research*, 2023.

Alex Chin. Regression adjustments for estimating the global treatment effect in experiments with interference. *Journal of Causal Inference*, 7, 05 2019. doi: 10.1515/jci-2018-0026.

Richard Cole, Nikhil R Devanur, Vasilis Gkatzelis, Kamal Jain, Tung Mai, Vijay V Vazirani, and Sadra Yazdanbod. Convex program duality, fisher markets, and Nash social welfare. In *18th ACM Conference on Economics and Computation, EC 2017*. Association for Computing Machinery, Inc, 2017.

Vincent Conitzer, Christian Kroer, Debmalya Panigrahi, Okke Schrijvers, Nicolas E Stier-Moses, Eric Sodomka, and Christopher A Wilkens. Pacing equilibrium in first price auction markets. *Management Science*, 2022a.

Vincent Conitzer, Christian Kroer, Eric Sodomka, and Nicolas E Stier-Moses. Multiplicative pacing equilibria in auction markets. *Operations Research*, 70(2):963–989, 2022b.

Yuan Deng, Jieming Mao, Vahab Mirrokni, and Song Zuo. Towards efficient auctions in an auto-bidding world. In *Proceedings of the Web Conference 2021*, pp. 3965–3973, 2021.

Yuan Deng, Jieming Mao, Vahab Mirrokni, Yifeng Teng, and Song Zuo. Non-uniform bid-scaling and equilibria for different auctions: An empirical study. *arXiv preprint arXiv:2311.10679*, 2023.

Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1):20150021, 2017. doi: doi: 10.1515/jci-2015-0021. URL https://doi.org/10.1515/jci-2015-0021.

Jason Gaitonde, Yingkai Li, Bar Light, Brendan Lucier, and Aleksandrs Slivkins. Budget pacing in repeated auctions: Regret and efficiency without convergence. In *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251, pp. 52, 2023.

Yuan Gao and Christian Kroer. Infinite-dimensional fisher markets and tractable fair division. *Operations Research (forthcoming)*, 2022.

Yuan Gao, Christian Kroer, and Alex Peysakhovich. Online market equilibrium with application to fair division. *arXiv preprint arXiv:2103.12936*, 2021.

Shumpei Goke, Gabriel Y Weintraub, Ralph Mastromonaco, and Sam Seljan. Bidders' responses to auction format change in internet display advertising auctions. *arXiv preprint arXiv:2110.13814*, 2021.

Christopher Harshaw, Fredrik Sävje, David Eisenstat, Vahab Mirrorkni, and Jean Pouget-Abadie. Design and analysis of bipartite experiments under a linear exposure-response model. *arXiv*, 2103.06392v3, 12 2021.

Han Hong, Aprajit Mahajan, and Denis Nekipelov. Extremum estimation and numerical derivatives. *Journal of Econometrics*, 188(1):250–263, 2015.

Chinmay Karande, Aranyak Mehta, and Ramakrishnan Srikant. Optimizing budget constrained spend in search advertising. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 697–706, 2013.

Brian Karrer, Liang Shi, Monica Bhole, Matt Goldman, Tyrone Palmer, Charlie Gelman, Mikael Konutgan, and Feng Sun. Network experimentation at scale. In *Proceedings of the 27th acm sigkdd conference on knowledge discovery & data mining*, pp. 3106–3116, 2021.

Shuangning Li and Stefan Wager. Random graph asymptotics for treatment effect estimation under network interference. *The Annals of Statistics*, 50(4):2334 – 2358, 2022. doi: 10.1214/22-AOS2191. URL https://doi.org/10.1214/22-AOS2191.

Luofeng Liao and Christian Kroer. Statistical inference and a/b testing for first-price pacing equilibria. *International Conference on Machine Learning*, 2023.

Luofeng Liao and Christian Kroer. Bootstrapping fisher market equilibrium and first-price pacing equilibrium. In *International Conference on Machine Learning*, 2024.

Luofeng Liao, Yuan Gao, and Christian Kroer. Nonstationary dual averaging and online fair allocation. *Advances in Neural Information Processing Systems*, 35:37159–37172, 2022.

Luofeng Liao, Yuan Gao, and Christian Kroer. Statistical inference for fisher market equilibrium. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=KemSBwOYJC.

Min Liu, Jialiang Mao, and Kang Kang. Trustworthy and powerful online marketplace experimentation with budget-split design. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, pp. 33193329, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383325. doi: 10.1145/3447548.3467193. URL https://doi.org/10.1145/3447548.3467193.

David Rolnick, Kevin Aydin, Jean Pouget-Abadie, Shahab Kamali, Vahab Mirrokni, and Amir Najmi. Randomized experimental design via geographic clustering. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2745–2753, 2019.

Paul R Rosenbaum. Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477):191–200, 2007. doi: 10.1198/016214506000001112. URL https://doi.org/10.1198/016214506000001112.

Alexander Shapiro. Sensitivity analysis of nonlinear programs and differentiability properties of metric projections. *SIAM Journal on Control and Optimization*, 26(3):628–645, 1988.

Alexander Shapiro. On differential stability in stochastic programming. *Mathematical Programming*, 47:107–116, 1990.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

Gilbert W Stewart. On the perturbation of pseudo-inverses, projections and linear least squares problems. *SIAM review*, 19(4):634–662, 1977.

Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, pp. 329337, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450321747. doi: 10.1145/2487575.2487695. URL https://doi.org/10.1145/2487575.2487695.

Aad W van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Davide Viviano, Lihua Lei, Guido Imbens, Brian Karrer, Okke Schrijvers, and Liang Shi. Causal clustering: design of cluster experiments under network interference. *arXiv preprint arXiv:2310.14983*, 2023.

Stefan Wager and Kuang Xu. Experimenting in equilibrium. *Management Science*, 67(11): 66946715, 2021. doi: 10.1287/mnsc.2020.3844. URL https://doi.org/10.1287/mnsc.2020.3844.

Corwin M. Zigler and Georgia Papadogeorgou. Bipartite causal inference with interference. *arXiv*, 1807.08660v1, 07 2018.

A Technical Lemmas

A.1 A perturbation result for constrained stochastic programs

We introduce a theorem from Shapiro (1990). Theorem 1 in that paper handles the case where the constraints are also defined as expectations of some random functions, and constraints are also perturbed in the analysis. For simplicity we specialize the theorem to deterministic constraints and do not perturb constraints. Note that Theorem 7.27 from Shapiro et al. (2021) can also be used to prove our Thm. 1.

Let Θ be a probability space equipped with an appropriate σ-algebra. Consider $F : \Theta \times \mathbb{R}^n \to \mathbb{R}$ and a set $B \subset \mathbb{R}^n$, given by

$$B = \{\boldsymbol{\beta} : g_i(\boldsymbol{\beta}) = 0, i \in I; g_i(\boldsymbol{\beta}) \leq 0, i \in J\}.$$

where I and J are finite index sets. Let P and Q be two probability measures on Θ. Let $\phi(\boldsymbol{\beta}, \alpha) = (P + \alpha(Q - P))F(\cdot, \boldsymbol{\beta})$ and $\phi(\boldsymbol{\beta}) = \phi(\boldsymbol{\beta}, 0)$. Here $\min_B \phi$ is the main program of interest and Q is a perturbation measure. The amount of perturbation is measured by $\alpha \in [0, 1]$.

Let $\boldsymbol{\beta}^*$ be the unique minimizer of $\phi(\boldsymbol{\beta})$ over B. Let $J^* \subset J$ be the inequality constraints active at $\boldsymbol{\beta}^*$, meaning $g_i(\boldsymbol{\beta}^*) = 0$ for all $i \in J^*$. Let $\boldsymbol{\beta}^*(\alpha)$ be the unique optimal solution to $\min_B \phi(\cdot, \alpha)$.

Lemma 1 (Theorem 1 from Shapiro (1990))**.** *Define the Lagrangian function by $L(\boldsymbol{\beta}, \lambda, \alpha) = \phi(\boldsymbol{\beta}, \alpha) + \sum_{i \in I \cup J} \lambda_i g_i(\boldsymbol{\beta})$ and $L_0(\boldsymbol{\beta}, \lambda) = L(\boldsymbol{\beta}, \lambda, 0)$. Let Λ_0 be the set of optimal Lagrangian multipliers* [1]*. Define the critical cone $C = \{\boldsymbol{u} : \boldsymbol{u}^\top \nabla g_i(\boldsymbol{\beta}^*) = 0, i \in I; \boldsymbol{u}^\top \nabla g_i(\boldsymbol{\beta}^*) \leq 0$ for $i \in J^*; \boldsymbol{u}^\top \nabla \phi(\boldsymbol{\beta}^*) \leq 0\}$. Define $\Lambda_0^* = \arg\max_{\lambda \in \Lambda_0} \sum_{i \in I \cup J} \lambda_i g_i(\boldsymbol{\beta}^*)$. Assume the following conditions.*

1. *Differentiability. The functions $\phi(\cdot, 0)$, $\phi(\cdot, 1)$, and $g_i, i \in I \cup J$, are continuously differentiable in a neighborhood of $\boldsymbol{\beta}^*$.*
2. *Constraint Qualification. The gradients $\nabla g_i(\boldsymbol{\beta}^*)$, $i \in I$ are linearly independent. And there exists $b \in \mathbb{R}^n$ such that $b^\top \nabla g_i(\boldsymbol{\beta}^*) = 0$ for $i \in I$ and $b^\top \nabla g_i(\boldsymbol{\beta}^*) < 0$ for $i \in J^*$.*
3. *Differentiability. The functions $\phi(\cdot, 0)$, g_i, $i \in I \cup J^*$ are twice continuously differentiable in a neighbourhood of $\boldsymbol{\beta}^*$.*
4. *Second-order sufficient condition. Assume for all nonzero $\boldsymbol{u} \in C$, $\max_{\lambda \in \Lambda_0^*} \boldsymbol{u}^\top \nabla^2_{xx} L_0(\boldsymbol{\beta}^*, \lambda) \boldsymbol{u} > 0$.*

Define the function b and the set $\boldsymbol{\Sigma}$ by

$$b(\boldsymbol{u}) = \max_{\lambda \in \Lambda_0^*} \boldsymbol{u}^\top \nabla^2_{xt} L_0(\boldsymbol{\beta}^*, \lambda) + \frac{1}{2} \boldsymbol{u}^\top \nabla^2_{xx} L_0(\boldsymbol{\beta}^*, \lambda) \boldsymbol{u} \tag{9}$$

$$\begin{aligned} \Sigma = \arg\min\, & \boldsymbol{u}^\top \nabla \phi(\boldsymbol{\beta}^*) \\ \text{s.t.} \quad & \boldsymbol{u}^\top \nabla g_i(\boldsymbol{\beta}^*) = 0, i \in I; \boldsymbol{u}^\top \nabla g_i(\boldsymbol{\beta}^*) \leq 0, i \in J^*; \end{aligned} \tag{10}$$

Then (i) there exists a positive K such that $\|\boldsymbol{\beta}^*(\alpha) - \boldsymbol{\beta}^*\| \leq K\alpha$ for all positive α in a neighborhood of zero. (ii) If, in addition, the function b has a unique minimizer $\boldsymbol{u}$ over Σ, then the limit $\lim_{\alpha \downarrow 0} (\boldsymbol{\beta}^*(\alpha) - \boldsymbol{\beta}^*)/\alpha$ exists and equals to $\boldsymbol{u}$.

B Proofs

B.1 Proof of Closed-Form Expression for Bias and Thm. 1

Now we prove Thm. 1.

To begin with, we define directional derivatives. For a probability measure P on the item set Θ, let $\boldsymbol{\beta}(P)$ be the unique optimal solution to the Eisenberg-Gale program: $\boldsymbol{\beta}(P) = \min_{\boldsymbol{\beta} \in (0,1]^n} \int F(\theta, \boldsymbol{\beta}) \, \mathrm{d}P(\theta)$. Let P and Q be two measures. When exists, the directional derivative is defined as

$$\mathrm{d}\boldsymbol{\beta}(P; Q - P) = \lim_{t \downarrow 0} \frac{\boldsymbol{\beta}(P + t(Q - P)) - \boldsymbol{\beta}(P)}{t}.$$

Step 1. Show there exist $K > 0$ and $\bar{\alpha}_1 > 0$ such that $\|\boldsymbol{\beta}^*_\alpha - \boldsymbol{\beta}^*\| \leq K\alpha$ for $\alpha \in [0, \bar{\alpha}_1]$. And $\mathrm{d}\boldsymbol{\beta}(s; s - s') = -(\boldsymbol{P}\boldsymbol{H}_0\boldsymbol{P})^\dagger \boldsymbol{\delta}_0$.

We apply lemma 1 with $F(\theta, \boldsymbol{\beta}) = \max_i v_i(\theta)\beta_i - \sum_i b_i \log(\beta_i)$, $g_i(\boldsymbol{\beta}) = \beta_i - 1$, $B = (0, 1]^n$, $\mathrm{d}P = s \, \mathrm{d}\theta$, $\mathrm{d}(Q - P) = (s - s') \, \mathrm{d}\theta$.

By SCS in the market $\mathscr{M}_0$, the set Σ (defined in Eq. (10)) becomes the plane $\{\boldsymbol{u} : u_i = 0, i \in I_+\}$ where we recall $I_+ = \{i : \beta_i^* = 0\}$. By SMO in $\mathscr{M}_0$, we know the Lagrangian multiplier of the EG program in $\mathscr{M}_0$ is unique, and thus $b(\cdot)$ defined in Eq. (9) becomes $b(\boldsymbol{u}) = \boldsymbol{u}^\top \boldsymbol{\delta}_0 + \frac{1}{2} \boldsymbol{u}^\top \boldsymbol{H}_0 \boldsymbol{u}$. We conclude the directional derivative is $\mathrm{d}\boldsymbol{\beta}(s; s - s') = \lim_{\epsilon \to 0^+} (\boldsymbol{\beta}_0^* - \boldsymbol{\beta}_{-\epsilon}^*)/\epsilon = -(\boldsymbol{P}\boldsymbol{H}_0\boldsymbol{P})^\dagger \boldsymbol{\delta}_0$, where $\boldsymbol{P} = \mathrm{Diag}(1(\beta_i^* < 1))$ and $\boldsymbol{\delta}_0 = \int \nabla f(\cdot, \boldsymbol{\beta}^*)(s - s') \, \mathrm{d}\theta$

[1] $\lambda \in \Lambda_0$ iff $\nabla_x L_0(\beta_0, \lambda) = 0$ and $\lambda_i \geq 0, i \in J^*$ and $\lambda_i = 0$ for all $j \in J \setminus J^*$.

Step 2. For all $\alpha \geq 0$ small enough, it holds $I_\alpha = I$, where we recall $I_\alpha = \{i : \beta^*_{\alpha,i} = 1\}$ and $I = \{i : \beta^*_i = 0\}$.

First we show $I \subseteq I_\alpha$, i.e., $\{i : \beta^*_i = 1\} \subseteq \{i : \beta^*_{\alpha,i} = 1\}$ for α small enough. This is saying, if a constraint $\beta_i \leq 1$ is strongly active in the EG program of $\mathscr{M}_0$, then it is also strongly active in the EG program of the market $\mathscr{M}_\alpha$. This holds by Lemma 2.2 from Shapiro (1988) and the SCS condition in $\mathscr{M}_0$. Next, we show, for small α, $I_\alpha \subseteq I$ by showing $[n] \setminus I \subseteq [n] \setminus I_\alpha$, i.e., $\{i : \beta^*_i < 1\} \subseteq \{i : \beta^*_{\alpha,i} < 1\}$. This holds by $\|\boldsymbol{\beta}^*_\alpha - \boldsymbol{\beta}^*\| \leq K\alpha$ for all $\alpha \leq \bar{\alpha}_1$.

Step 3. For all $\alpha \geq 0$ small enough, the directional derivative $\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s') = \lim_{\epsilon \to 0^+}(\boldsymbol{\beta}^*_{\alpha-\epsilon} - \boldsymbol{\beta}^*_\alpha)/\epsilon$ exists and $\alpha \mapsto \mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s')$ is continuous.

Applying lemma 1 with $\mathrm{d}P = s_\alpha \,\mathrm{d}\theta$ and $\mathrm{d}(Q - P) = (s - s')\,\mathrm{d}\theta$ shows that the directional derivative exists for all α small enough. Now we find an expression for $\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s')$. By SMO in $\mathscr{M}_0$, twice continuously differentiability of $\boldsymbol{\beta} \mapsto \nabla^2 \int F(\boldsymbol{\beta}, \theta) s' \,\mathrm{d}\theta$ in a neighborhood of $\boldsymbol{\beta}^*$, and the Lipschitzness result from step 1, we know $\boldsymbol{\beta} \to \nabla^2 \int F(\theta, \boldsymbol{\beta}) s_\alpha \,\mathrm{d}\theta$ is twice continuously differentiable at $\boldsymbol{\beta}^*_\alpha$, which implies uniqueness of Lagrangian multiplier in the market $\mathscr{M}_\alpha$, for all α small enough (Lemma 2.2 from Shapiro (1988)). And thus $b(\cdot)$ defined in Eq. (9) becomes $b(\boldsymbol{u}) = \boldsymbol{u}^\intercal \boldsymbol{\delta}_\alpha + \frac{1}{2}\boldsymbol{u}^\intercal \boldsymbol{H}_\alpha \boldsymbol{u}$. Next, the SCS condition in the market $\mathscr{M}_0$ implies SCS in $\mathscr{M}_\alpha$ for all α small enough. The set Σ (defined in Eq. (10)) becomes the plane $\{\boldsymbol{u} : u_i = 0, i \in I_\alpha\}$. The quadratic program $\min_\Sigma b$ has a closed-form solution: $\boldsymbol{u} = -(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \boldsymbol{\delta}_\alpha$ where $\boldsymbol{P}_\alpha = \mathrm{Diag}(1(\beta^*_{\alpha,i} < 1))$, $\boldsymbol{\delta}_\alpha = \int \nabla f(\cdot, \boldsymbol{\beta}^*_\alpha)(s - s')\,\mathrm{d}\theta$. And we apply $I_\alpha = I$ for all α small enough from step 2, and so $\boldsymbol{P}_\alpha = \boldsymbol{P}$. We conclude

$$\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s') = -(\boldsymbol{P}\boldsymbol{H}_\alpha \boldsymbol{P})^\dagger \boldsymbol{\delta}_\alpha$$

for all α small enough.

Next, we show that the map $\alpha \mapsto (\boldsymbol{P}\boldsymbol{H}_\alpha \boldsymbol{P})^\dagger \boldsymbol{\delta}_\alpha$ is continuous for all α small enough. To see this, note $\boldsymbol{H}_\alpha = \nabla^2 \int F(\theta, \boldsymbol{\beta}^*_\alpha) s_\alpha \,\mathrm{d}\theta = (1-\alpha)\nabla^2 \int F(\cdot, \boldsymbol{\beta}^*_\alpha) s \,\mathrm{d}\theta + \alpha \nabla^2 \int F(\cdot, \boldsymbol{\beta}^*_\alpha) s' \,\mathrm{d}\theta$ is continuous in α. The $\boldsymbol{P}$ matrix is fixed, and so $\alpha \mapsto \boldsymbol{P}\boldsymbol{H}_\alpha \boldsymbol{P}$ is continuous. Without loss of generality, suppose $I^c = \{i : \beta^*_i < 0\} = [k]$ are the first k buyers. Then $\boldsymbol{P}\boldsymbol{H}_\alpha \boldsymbol{P}$ creates a matrix with upper left k-by-k block equal to the upper left k-by-k block of $\boldsymbol{H}_\alpha$, denoted $\boldsymbol{H}_{\alpha, I^c I^c}$, and zeros everywhere else. Then $(\boldsymbol{P}\boldsymbol{H}_\alpha \boldsymbol{P})^\dagger$ is a matrix with upper left block being $(\boldsymbol{H}_{\alpha, I^c I^c})^{-1}$ and zeros everywhere else. Since $\boldsymbol{H}_\alpha$ is positive definite for all small α, the submatrix $\boldsymbol{H}_{\alpha, I^c I^c}$ must also be positive definite. And so $\alpha \mapsto (\boldsymbol{H}_{\alpha, I^c I^c})^{-1}$ is continuous, implying continuity of $\alpha \mapsto (\boldsymbol{P}\boldsymbol{H}_\alpha \boldsymbol{P})^\dagger$. Finally, $\boldsymbol{\delta}_\alpha = \int \nabla f(\cdot, \boldsymbol{\beta}^*_\alpha)(s - s')\,\mathrm{d}\theta$ is continuous in α.

Step 4. The desired claim: $\widetilde{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^* = o(\alpha)$. Note $\mathrm{d}\boldsymbol{\beta}^*(\alpha)$ defined in Eq. (6) is exactly $\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s')$. And so

$$\begin{aligned}
\widetilde{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^* &= \boldsymbol{\beta}^*_\alpha + \alpha\,\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s') - \boldsymbol{\beta}^* \\
&= (\boldsymbol{\beta}^* - \alpha\,\mathrm{d}\boldsymbol{\beta}(s_0; s - s') + o(\alpha)) + \alpha\,\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s') - \boldsymbol{\beta}^* \\
&= \alpha(\mathrm{d}\boldsymbol{\beta}(s_\alpha; s - s') - \mathrm{d}\boldsymbol{\beta}(s_0; s - s')) + o(\alpha) = o(\alpha).
\end{aligned}$$

where the last line uses continuity in α of the directional derivative. This completes the proof of Thm. 1.

B.2 Proof of Thm. 2

Recall the estimator $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}^\gamma_\alpha - \alpha(\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger \widehat{\boldsymbol{\delta}}_\alpha$ and the debiased surrogate is $\widetilde{\boldsymbol{\beta}}^* = \boldsymbol{\beta}^*_\alpha - \alpha(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \boldsymbol{\delta}_\alpha$. By results from Liao & Kroer (2023) we know that SCS and SMO in the market $\mathscr{M}_\alpha$ and $\eta_t = o(1)$ imply $\boldsymbol{\beta}^\gamma_\alpha \xrightarrow{p} \boldsymbol{\beta}^*_\alpha$ and $(\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger \xrightarrow{p} (\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger$. Finally, $\widehat{\boldsymbol{\delta}}_\alpha \xrightarrow{p} \boldsymbol{\delta}_\alpha$ holds by the law of large numbers. We complete the proof of Thm. 2.

B.3 PROOF OF THM. 3

Recall the estimator $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_\alpha^\gamma - \alpha(\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger \widehat{\boldsymbol{\delta}}_\alpha$ and the debiased surrogate is $\widetilde{\boldsymbol{\beta}}^* = \boldsymbol{\beta}_\alpha^* - \alpha(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \boldsymbol{\delta}_\alpha$. Define

$$\boldsymbol{z}_t = \boldsymbol{\beta}_\alpha^\gamma - \boldsymbol{\beta}_\alpha^* - \alpha(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger (\widehat{\boldsymbol{\delta}}_\alpha - \boldsymbol{\delta}_\alpha) \tag{11}$$

$$\boldsymbol{\xi}_t = -\alpha((\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger - (\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger)\boldsymbol{\delta}_\alpha \tag{12}$$

$$\boldsymbol{\zeta}_t = -\alpha((\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger - (\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger)(\widehat{\boldsymbol{\delta}}_\alpha - \boldsymbol{\delta}_\alpha) \tag{13}$$

Then clearly $\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^* = \boldsymbol{z}_t + \boldsymbol{\xi}_t + \boldsymbol{\zeta}_t$.

Step 1. We show $\sqrt{t}\boldsymbol{z}_t$ converges to a normal distribution. By Liao & Kroer (2023), it holds

$$\sqrt{t}(\boldsymbol{\beta}_\alpha^\gamma - \boldsymbol{\beta}_\alpha^*) = \frac{1}{\sqrt{t}} \sum_{\tau=1}^{t} \boldsymbol{d}_\alpha(\theta^\tau) - \mathbb{E}[\boldsymbol{d}_\alpha] + o_p(1) \tag{14}$$

where $\boldsymbol{d}_\alpha(\theta) = -(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \nabla f(\theta, \boldsymbol{\beta}_\alpha^*)$ is the influence function. Next, we define the likelihood ratio function $g(\theta) = (\mathrm{d}(s - s')/\mathrm{d}s_\alpha)(\theta) = -\frac{1}{\alpha} 1_{\Theta_{\text{bad}}}(\theta) + \frac{1}{1-\alpha} 1_{\Theta_{\text{good}}}(\theta)$. Then $\boldsymbol{\delta}_\alpha = \int \nabla f(\cdot, \boldsymbol{\beta}_\alpha^*)(s - s')\,\mathrm{d}\theta = \int g(\cdot) \nabla f(\cdot, \boldsymbol{\beta}_\alpha^*) s_\alpha(\cdot)\,\mathrm{d}\theta$. Then by definition $\widehat{\boldsymbol{\delta}}_\alpha$ can be written as $\widehat{\boldsymbol{\delta}}_\alpha = P_t(g(\cdot)\nabla f(\cdot, \boldsymbol{\beta}_\alpha^\gamma))$ where $P_t h(\cdot) = \frac{1}{t}\sum_{\tau=1}^{t} h(\theta^\tau)$ and θ^τ are i.i.d. draws from s_α. Also let $\nu_t h = \sqrt{t}(P_t h - \int h s_\alpha\,\mathrm{d}\theta)$. By the decomposition,

$$\sqrt{t}(\widehat{\boldsymbol{\delta}}_\alpha - \boldsymbol{\delta}_\alpha) \tag{15}$$

$$= \sqrt{t}\left(P_t\big(g(\cdot)\nabla f(\cdot, \boldsymbol{\beta}_\alpha^\gamma)\big) - \int g(\cdot)\nabla f(\cdot, \boldsymbol{\beta}_\alpha^*) s_\alpha\,\mathrm{d}\theta \right) \tag{16}$$

$$= \sqrt{t}\left(P_t\big(g(\cdot)\nabla f(\cdot, \boldsymbol{\beta}_\alpha^*)\big) - \int g(\cdot)\nabla f(\cdot, \boldsymbol{\beta}_\alpha^*) s_\alpha\,\mathrm{d}\theta \right) + \nu_t\big(g(\cdot)(\nabla f(\cdot, \boldsymbol{\beta}_\alpha^\gamma) - \nabla f(\cdot, \boldsymbol{\beta}_\alpha^*))\big) \tag{17}$$

$$= \nu_t(g(\cdot)\nabla f(\cdot, \boldsymbol{\beta}_\alpha^*)) + o_p(1), \tag{18}$$

where the last line follows by a stochastic equicontinuity argument as in Liao & Kroer (2023). Finally, $\boldsymbol{d}_\alpha(\cdot) - \alpha g(\cdot)(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \nabla f(\cdot, \boldsymbol{\beta}_\alpha^*)$ is exactly the influence function $\boldsymbol{d}_1(\cdot)$ defined in the theorem.

Step 2. Show $\boldsymbol{\xi}_t = O_p(\eta_t)$.

We need a matrix perturbation result.

Lemma 2 (Theorem 2.2 from Stewart (1977))**.** *Let $\boldsymbol{A}$ and $\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{E}$ be nonsingular square matrices, and $\|\boldsymbol{A}^{-1}\|\|\boldsymbol{E}\| < 1$. Here $\|\cdot\|$ is the operator norm. Then*

$$\|\boldsymbol{B}^{-1} - \boldsymbol{A}^{-1}\| \le \frac{\|\boldsymbol{A}^{-1}\|^2}{1 - \|\boldsymbol{E}\|\|\boldsymbol{A}^{-1}\|}\|\boldsymbol{E}\|. \tag{19}$$

By Liao & Kroer (2023), we know $\mathbb{P}(\widehat{\boldsymbol{P}}_\alpha = \boldsymbol{P}_\alpha) \to 1$. Without loss of generality, suppose $I_\alpha^c = \{i : \beta_{\alpha,i}^* < 0\} = [k]$ are the first k buyers. Then $(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)$ creates a matrix with upper left k-by-k block equal to the upper left k-by-k block of $\boldsymbol{H}_\alpha$, denoted $\boldsymbol{H}_{\alpha, I^c I^c}$, and zeros everywhere else. Under the event $\{\widehat{\boldsymbol{P}}_\alpha = \boldsymbol{P}_\alpha\}$, the matrix $(\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)$ is one with upper left k-by-k block being upper left k-by-k block of $\widehat{\boldsymbol{H}}_\alpha$, denoted $\widehat{\boldsymbol{H}}_{\alpha, I^c I^c}$, and zeros everywhere else. Now let $\boldsymbol{A} = \widehat{\boldsymbol{H}}_{\alpha, I^c I^c}$ and $\boldsymbol{B} = \boldsymbol{H}_{\alpha, I^c I^c}$. It is clear that $\|\boldsymbol{\xi}_t\|_2 \le \alpha \|\boldsymbol{\delta}_\alpha\|_2 \|\boldsymbol{A}^{-1} - \boldsymbol{B}^{-1}\| = O_p(\eta_t)$ by lemma 2.

Step 3. Show $\boldsymbol{\zeta}_t = o_p(\frac{1}{\sqrt{t}})$. From previous derivation, $\widehat{\boldsymbol{\delta}}_\alpha - \boldsymbol{\delta}_\alpha = O_p(1/\sqrt{t})$ and $(\widehat{\boldsymbol{P}}_\alpha \widehat{\boldsymbol{H}}_\alpha \widehat{\boldsymbol{P}}_\alpha)^\dagger - (\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger = O_p(\eta_t) = o_p(1)$. We conclude $\boldsymbol{\zeta}_t = O_p(1/\sqrt{t}) o_p(1) = o_p(1/\sqrt{t})$.

We complete the proof of Thm. 3 part 1.

Now we prove part 2. The claim that $\boldsymbol{H}_\alpha = \mathrm{Diag}(b_i/(\beta_{\alpha,i}^*)^2)$ follows from Liao et al. (2023). Next we derive the asymptotic distribution. We only need to handle the $\boldsymbol{\xi}_t$ term defined as before. Under the event $\{\widehat{\boldsymbol{P}}_\alpha = \boldsymbol{P}_\alpha\}$, we have $\boldsymbol{\xi}_t = -\mathrm{Diag}(\alpha \delta_{\alpha,i} 1(\beta_{\alpha,i}^* < 1) b_i^{-1})(\boldsymbol{\beta}_\alpha^\gamma \circ \boldsymbol{\beta}_\alpha^\gamma - \boldsymbol{\beta}_\alpha^* \circ \boldsymbol{\beta}_\alpha^*)$. By Liao

& Kroer (2023) we know $\sqrt{t}(\boldsymbol{\beta}_\alpha^\gamma - \boldsymbol{\beta}_\alpha^*) = \frac{1}{\sqrt{t}}\sum_{\tau=1}^t(\boldsymbol{d}_\alpha(\theta^\tau) - \mathbb{E}_\alpha[\boldsymbol{d}_\alpha]) + o_p(1)$. So by the delta method,

$$\sqrt{t}(\boldsymbol{\beta}_\alpha^\gamma \circ \boldsymbol{\beta}_\alpha^\gamma - \boldsymbol{\beta}_\alpha^* \circ \boldsymbol{\beta}_\alpha^*) = \mathrm{Diag}(2\boldsymbol{\beta}_\alpha^*)\frac{1}{\sqrt{t}}\sum_{\tau=1}^t(\boldsymbol{d}_\alpha(\theta^\tau) - \mathbb{E}_\alpha[\boldsymbol{d}_\alpha]) + o_p(1).$$

Summarizing,

$$\sqrt{t}\boldsymbol{\xi}_t = -\mathrm{Diag}(2\beta_{\alpha,i}^*\alpha\delta_{\alpha,i}1(\beta_{\alpha,i}^* < 1)b_i^{-1})\frac{1}{\sqrt{t}}\sum_{\tau=1}^t(\boldsymbol{d}_\alpha(\theta^\tau) - \mathbb{E}_\alpha[\boldsymbol{d}_\alpha]) + o_p(1)$$

Note that for i such that $\beta_{\alpha,i}^* = 1$, the i-th entry of $\boldsymbol{d}_\alpha$ will be zero almost surely, so the indicator can be dropped in the diagonal matrix. Now let $\boldsymbol{d}_2(\theta) = -\mathrm{Diag}(2\alpha\beta_{\alpha,i}^*\delta_{\alpha,i}b_i^{-1})\boldsymbol{d}_\alpha + \boldsymbol{d}_1$, where we recall $\boldsymbol{d}_1$ is defined in Thm. 3. Then under the condition that $\mathbb{E}_\alpha[\frac{1}{\mathsf{bidgap}(\boldsymbol{\beta}_\alpha^*,\theta)}] < \infty$, we have $\sqrt{t}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^*) \xrightarrow{d} \mathcal{N}(0, \mathbb{E}_\alpha[(\boldsymbol{d}_2 - \mathbb{E}_\alpha[\boldsymbol{d}_2])(\boldsymbol{d}_2 - \mathbb{E}_\alpha[\boldsymbol{d}_2])^\intercal])$.

B.4 Proof of Thm. 4

Proof. Note

$$\begin{aligned}
&\left|\int \max_i(v_i(\theta)\widetilde{\boldsymbol{\beta}}_i^*)s(\theta)\,\mathrm{d}\theta - \int \max_i(v_i(\theta)\boldsymbol{\beta}_i^*)s(\theta)\,\mathrm{d}\theta\right| \\
&\le \int |\max_i(v_i(\theta)\widetilde{\boldsymbol{\beta}}_i^*)s(\theta) - \max_i(v_i(\theta)\boldsymbol{\beta}_i^*)s(\theta)|\,\mathrm{d}\theta \\
&\le \max_i \sup_\theta v_i(\theta)\cdot\|\boldsymbol{\beta}^* - \widetilde{\boldsymbol{\beta}}^*\|_\infty = o(\alpha)
\end{aligned}$$

as $\alpha \downarrow 0$.

Let $P_t h(\cdot) = \frac{1}{t}\sum_{\tau=1}^t h(\theta^\tau)$ and θ^τ are i.i.d. draws from s_α, and $\boldsymbol{P}h(\cdot) = \int h s_\alpha$. Let $g(\cdot) = \mathrm{d}s/\mathrm{d}s_\alpha = \frac{1}{1-\alpha}1_{\Theta_{\mathsf{good}}}$. Also let $\nu_t h = \sqrt{t}(P_t h - \int h s_\alpha\,\mathrm{d}\theta)$. We change the measure and write $\widetilde{REV}^* = \int f(\theta, \widetilde{\boldsymbol{\beta}}^*)s\,\mathrm{d}\theta = \int f(\theta, \widetilde{\boldsymbol{\beta}}^*)g(\theta)s_\alpha\,\mathrm{d}\theta$. For the claim regarding asymptotic normality, note

$$\begin{aligned}
&\sqrt{t}(\widehat{REV} - \widetilde{REV}^*) \\
&= \sqrt{t}[P_t f(\cdot, \widehat{\boldsymbol{\beta}})g(\cdot) - Pf(\cdot, \widetilde{\boldsymbol{\beta}}^*)g(\cdot)] \\
&= \sqrt{t}[P_t f(\cdot, \widetilde{\boldsymbol{\beta}}^*)g(\cdot) - Pf(\cdot, \widetilde{\boldsymbol{\beta}}^*)g(\cdot)] + \sqrt{t}[Pf(\cdot, \widehat{\boldsymbol{\beta}})g(\cdot) - Pf(\cdot, \widetilde{\boldsymbol{\beta}}^*)g(\cdot)] \\
&\quad + \sqrt{t}(P_t - P)(f(\cdot, \widehat{\boldsymbol{\beta}})g(\cdot) - f(\cdot, \widetilde{\boldsymbol{\beta}}^*)g(\cdot)) \\
&= \nu_t(f(\cdot, \widetilde{\boldsymbol{\beta}}^*)g(\cdot)) + \nu_t(\mathbb{E}_\alpha[g(\cdot)\nabla f(\cdot, \widetilde{\boldsymbol{\beta}}^*)]^\intercal \boldsymbol{d}(\cdot)) + o_p(1)
\end{aligned}$$

where the last line follows by the delta method and a stochastic equicontinuity argument. We complete the proof of the theorem. □

C Ommited Maintext

C.1 Formal Definition of Finite FPPE

Definition 5 (Finite FPPE, Conitzer et al. (2022a))**.** *Given (b, v, σ, γ), the finite FPPE, $\widehat{FPPE}(\boldsymbol{b}, \boldsymbol{v}, \sigma, \gamma)$, is the unique tuple $(\boldsymbol{\beta}, p) \in [0,1]^n \times \mathbb{R}_+^t$ such that there exist $x_i^\tau \in [0,1]$: (First-price) For all τ, $p^\tau = \max_i \beta_i v_i^\tau$. For all i and τ, $x_i^\tau > 0$ implies $\beta_i v_i^\tau = \max_k \beta_k v_k^\tau$. (Supply and budget feasible) For all i, $\sigma\sum_{\tau=1}^t x_i^\tau p^\tau \le b_i$. For all τ, $\sum_{i=1}^n x_i^\tau \le 1$. (Market clearing) For all τ, $p^\tau > 0$ implies $\sum_{i=1}^n x_i^\tau = 1$. (No unnecessary pacing) For all i, $\sigma\sum_{\tau=1}^t x_i^\tau p^\tau < b_i$ implies $\beta_i = 1$.*

D RELATED WORK

Pacing equilibrium. Pacing is a budget management strategy where bids are scaled down by a multiplicative factor in order to smooth out and control spending. In the first-price setting, Borgs et al. (2007) study first price auctions with budget constraints in a perturbed model, where the limit prices converge to those of an FPPE. Building on that work, Conitzer et al. (2022a) introduce the FPPE framework to model autobidding in repeated auctions, and discover several properties of FPPE such as shill-proofness monotonicity properties, and a close relationship between FPPE and the quasilinear Fisher market equilibrium (Chen et al., 2007; Cole et al., 2017). Gao & Kroer (2022) propose an infinite-dimensional variant of the quasilinear Fisher market, which lays the probability foundations of the current paper. Gao et al. (2021) and Liao et al. (2022) study online computation of the infinite-dimensional Fisher market equilibrium, and we utilize their methods to compute equilibria. In the second-price setting, a variety of models have been proposed for modeling the outcome of budget management: Balseiro et al. (2015) study budget-management in second-price auctions through a fluid mean-field approximation; Balseiro & Gur (2019) investigate adaptive pacing strategy from buyers' perspective in a stochastic continuous setting; and Balseiro et al. (2021b) study several budget smoothing methods including multiplicative pacing in a stochastic context. A second-price analogue of FPPE was explored by Conitzer et al. (2022b). Modeling statistical inference in the second-price setting is most likely harder, due to equilibrium multiplicity issues (Conitzer et al., 2022b), and hardness of even computing equilibria (Chen et al., 2023).

Recently, a series of papers (e.g., Aggarwal et al. (2019); Deng et al. (2021); Balseiro et al. (2021a)) investigated an "autobidding" model for value-maximizing bidders, who aim to maximize their value subject to budget or return-on-spend constraints, which may include buying individual ad impressions that have negative utility. Several extensions provide results for more general settings, such as the quality of equilibria for generalized rationality models for bidders Babaioff et al. (2021), non-uniform bidding strategies Deng et al. (2023), and liquid welfare guarantees when budget-constrained buyers fail to reach equilibrium Gaitonde et al. (2023). Statistical inference for parallel A/B testing under such models represent an interesting future direction of work.

Interference in A/B Testing for Auction Markets. Experimentation under interference has been extensively studied in the past decade within the context of social networks Ugander et al. (2013); Aronow & Samii (2017); Eckles et al. (2017); Athey et al. (2018); Li & Wager (2022) and two-sided marketplaces Zigler & Papadogeorgou (2018); Wager & Xu (2021). Under interference, the outcome of an experiment unit, either a user in a social network or a buyer in an online marketplace, can be affected by the treatment status of other units in the networks, violating the Stable Unit Treatment Values Assumption (SUTVA) assumption commonly employed in analyzing online A/B tests Rosenbaum (2007). To take into account this type of interference, novel designs and analysis approaches have been proposed, e.g., cluster randomization Ugander et al. (2013); Karrer et al. (2021), regression-based estimators Chin (2019), and bipartite analysis Harshaw et al. (2021). A/B testing in the presence of seller-side budget constraints has been discussed by Basse et al. (2016) and Liu et al. (2021) where the budget-splitting design spans the full market. However, most of the above works do not consider the particular type of interference caused by competitive equilibrium effects. Recent work has also addressed equilibrium effects (Wager & Xu, 2021; Liao et al., 2023; Liao & Kroer, 2024). Most relevant to our results, Liao & Kroer (2023) consider A/B testing via budget splitting in Fisher markets and FPPE. However, none of these works consider the effects caused by parallel A/B tests that interfere with each other. Our method builds on top of the budget-splitting design; our approach additionally segments the full market into submarkets, where different submarkets receive different treatment. In the formal sections of our paper we assume that the submarket structure is known. In practice, the submarkets may not be known; Rolnick et al. (2019) and Viviano et al. (2023) present methods to identify submarkets through clustering approaches that aim to limit the interference between clusters.

Algorithm 1: Debiasing procedure

Input. Budgets $[b_1, \ldots, b_n]$. Values v_i^τ. First-price allocation $x_i^\tau \in [0,1]$. The observed equilibrium pacing multipliers $\boldsymbol{\beta}_\alpha^\gamma = [\boldsymbol{\beta}_1^\gamma, \ldots, \boldsymbol{\beta}_n^\gamma]$. Proportion of bad items α. Desired confidence level $(1-c) \in (0,1)$.

1. Estimate the Hessian matrix $\widehat{\boldsymbol{H}}_\alpha$, e.g. via $\widehat{\boldsymbol{H}}_\alpha = \mathrm{Diag}(b_i/(\beta_{\alpha,i}^\gamma)^2)$, or using finite differences.
2. Compute relevant quantities $\widehat{\boldsymbol{P}}_\alpha$, $\widehat{\boldsymbol{\delta}}_\alpha$, $\boldsymbol{\mu}^\tau$.
3. Compute the debiased estimator $\widehat{\boldsymbol{\beta}}$ and covariance matrix $\widehat{\boldsymbol{\Sigma}}_1$.
4. Output confidence region: $\widehat{\boldsymbol{\beta}} + (\chi_{n,c}/\sqrt{t})\widehat{\boldsymbol{\Sigma}}_1^{1/2}\mathbb{B}$, where $\chi_{n,c}^2$ is the $(1-c)$-quantile of a chi-square distribution with degree n, and $\mathbb{B}$ is the unit ball in $\mathbb{R}^n$.

E Results for Revenue

The revenue in the good market is defined as $REV^* = \int \max_i\{v_i(\theta)\beta_i^*\}s(\theta)\,\mathrm{d}\theta$. We define the debiased revenue surrogate and its estimator

$$\widetilde{REV}^* = \int \max_i\{v_i(\theta)\widetilde{\boldsymbol{\beta}}_i^*\}s(\theta)\,\mathrm{d}\theta\ , \tag{20}$$

$$\widehat{REV} = \frac{1}{1-\alpha}\frac{1}{t}\sum_{\tau=1}^t 1_{\Theta_{\mathsf{good}}}(\theta^\tau)\max_i\{v_i(\theta^\tau)\widehat{\boldsymbol{\beta}}_i\}\ . \tag{21}$$

In words, we approximate REV^* with the revenue generated in the market where buyers bid according to the pacing profile $\widetilde{\boldsymbol{\beta}}^*$, and then we estimate this via plug-in estimation on the subset of items that are good.

Theorem 4. *Let the conditions from Thm. 1 and Thm. 3 hold. The debiased revenue removes first-order bias: $\widetilde{REV}^* = REV^* + o(\alpha)$. Assume that $\sqrt{t}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^*)$ has influence function $\boldsymbol{d}$, i.e., $\sqrt{t}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}}^*) = \frac{1}{\sqrt{t}}\sum_{\tau=1}^t \boldsymbol{d}(\theta^\tau) + o_p(1)$, then $\sqrt{t}(\widehat{REV} - \widetilde{REV}^*) \xrightarrow{d} \mathcal{N}(0, \sigma_{REV}^2)$ where $\sigma_{REV}^2 = Cov_\alpha(\boldsymbol{d}_{REV})$, $\boldsymbol{d}_{REV}(\cdot) = \frac{1}{1-\alpha}\big(1_{\Theta_{\mathsf{good}}}(\cdot)f(\cdot,\widetilde{\boldsymbol{\beta}}^*) + \mathbb{E}_\alpha[\nabla f(\theta,\widetilde{\boldsymbol{\beta}}^*)1_{\Theta_{\mathsf{good}}}(\theta)]^\intercal \boldsymbol{d}(\cdot)\big)$. The proof is in App. B.4.*

Examples of influence functions compatible with Thm. 4 are the influence functions in Thm. 3.

We have two methods to construct CIs for the debiased revenue surrogate in Eq. (7). The first method is based on the asymptotic normality results in Thm. 4. In the second method, a revenue CI is constructed based on a CI for $\boldsymbol{\beta}_\alpha^*$. Suppose $\mathbb{CI}$ is some confidence region for $\boldsymbol{\beta}_\alpha^*$. Then $\{\frac{1}{1-\alpha}\frac{1}{t}\sum_{\tau=1}^t 1_{\Theta_{\mathsf{good}}}(\theta^\tau)\max_i v_i(\theta^\tau)\beta_i : \boldsymbol{\beta} \in \mathbb{CI}\}$ serves as a CI for revenue. Another way to utilize a pacing multiplier CI is based on the observation that revenue is a monotone function of the pacing multiplier. Let $\underline{\beta_i} = \min\{\beta_i : \boldsymbol{\beta} \in \mathbb{CI}\}$ and $\bar{\beta}_i = \max\{\beta_i : \boldsymbol{\beta} \in \mathbb{CI}\}$. Then another natural (potentially wider) CI for $\widetilde{REV}^*$ is $[\underline{REV}, \overline{REV}]$, where

$$\underline{REV} = \frac{1}{t}\sum_{\tau=1}^t \frac{1_{\Theta_{\mathsf{good}}}(\theta^\tau)}{1-\alpha}\max_i v_i(\theta^\tau)\underline{\beta_i}\ , \quad \overline{REV} = \frac{1}{t}\sum_{\tau=1}^t \frac{1_{\Theta_{\mathsf{good}}}(\theta^\tau)}{1-\alpha}\max_i v_i(\theta^\tau)\bar{\boldsymbol{\beta}}_i\ . \tag{22}$$

F Parallel A/B Test under Interference

In this section, we use the theory from the previous sections to formulate a statistical inference theory for parallel A/B tests under interference.

Consider K auction markets in which we want to run parallel A/B tests. There are n buyers, each with budget b_i corresponding to one unit supply of items, and an item set Θ, partitioned into Θ_{good} and Θ_{bad}. Let $C_1, \ldots, C_K$ be a partition of the buyers, $\Theta_1, \ldots, \Theta_K$ be a partition of the good item set Θ_{good}, and $s_1, \ldots s_K$ be a set of supply functions, supported on $\Theta_1, \ldots, \Theta_K$ respectively. The k-th submarket consists of buyers in C_k, the item set Θ_k and supply s_k. Let $s = \frac{1}{K}\sum_k s_k$ be the average mixture and s' be a supply supported on Θ_{bad}.

To model treatment application we introduce the *potential value functions*

$$v(0) = (v_1(0,\cdot), \ldots, v_n(0,\cdot)),\ v(1) = (v_1(1,\cdot), \ldots, v_n(1,\cdot))\ .$$

If item θ is exposed to treatment $w \in \{0,1\}$, then its value to buyer i will be $v_i(w,\theta)$.

We extend the structure on Θ_{bad} and Θ_{good} introduced in Sec. 3.2 to A/B testing here. We assume these submarkets are separated. This is to model the ideal case where there is no interference. A buyer $i \in C_k$ is only interested in items from the submarket he belongs to: $v_i(w,\theta) = 0$ for $\theta \in C_{k'}, k' \neq k$, and $w \in \{0,1\}$. An item is called a bad item if it has positive values for buyers from different submarkets. Formally, $\theta \in \Theta_{\mathsf{bad}}$ if there exist $w \in \{0,1\}$, $i \in C_k$, $j \in C_{k'}$, $k \neq k'$, so that $v_i(w,\theta) > 0$ and $v_j(w,\theta) > 0$. A buyer i from submarket k has a (potentially) positive value only for items in $\Theta_k \cup \Theta_{\mathsf{bad}}$.

Step 1. The treatment effects. Let $\mathscr{M}(w,k) = \mathsf{FPPE}(\{b_i\}_{i\in C_k}, \{v_i(\cdot,w)\}_{i\in C_k}, s_k, \Theta_k)$be the k-th limit submarket under treatment $w \in \{0,1\}$, and let revenue be $REV_k^*(w)$. Then the treatment effect in the submarket k is $\tau_k^* = REV_k^*(1) - REV_k^*(0)$. Equivalently, all K submarkets can be formed simultaneously, $\mathscr{M}(w) = \mathsf{FPPE}\left(\frac{1}{K}\boldsymbol{b}, \boldsymbol{v}(w), s, \Theta\right)$. The scaling $1/K$ in budgets ensures that a buyer in C_k has budget b_i/K to bid for items in Θ_k, whose supply in $\mathscr{M}(w)$ is $\frac{1}{K}s_k$. The k-th component of $\mathscr{M}(w)$ corresponds to $\mathscr{M}(w,k)$.

Step 2. The experiment. A practical parallel A/B test scheme is as follows.

Step 2.1. Budget splitting. Decide on a budget split ratio π_0, π_1 satisfying $\pi_0 + \pi_1 = 1$, and split budgets accordingly. Choose A/B test size t, the total number of impressions in each submarket summed across treatments. Then, the budget of buyer i for treatment 0 will be $\pi_0 t b_i$, and $\pi_1 t b_i$ for treatment 1. This is because b_i is normalized to correspond to one unit of total supply.

Step 2.2. Observe markets with interference. Draw items/impressions from $\alpha s' + (1-\alpha)s$. Draw $\lfloor \pi_0 t \rfloor$ items for each submarket under treatment 0 and $\lfloor \pi_1 t \rfloor$ for treatment 1. A total of $t_w = \lfloor \pi_w t K \rfloor$ items are drawn for the whole market under treatment w. Since items come from the contaminated distribution, roughly a small fraction α of the t_w items cause interference.

The data observed following the described A/B test scheme can be compactly represented by two finite FPPEs (one for each treatment) (Def. 5): $\widehat{\mathscr{M}_\alpha}(w) = \widehat{\mathsf{FPPE}}(\pi_w t \boldsymbol{b}, \boldsymbol{v}(w), 1, \gamma)$, where $\gamma \subset \Theta$ consists of $t_w = \lfloor \pi_w t K \rfloor$ i.i.d. draws from the mixture $\alpha s' + (1-\alpha)s$.

It can be shown that the observed market $\widehat{\mathscr{M}_\alpha}(w)$ converges to $\mathscr{M}_\alpha(w) = \mathsf{FPPE}\left(\frac{1}{K}\boldsymbol{b}, \boldsymbol{v}(w), s_\alpha, \Theta\right)$.

Step 3. Debiasing $\boldsymbol{\beta}$. Following previous section, we can debias $\widehat{\mathscr{M}_\alpha}(w)$ to approximate $\mathscr{M}(w)$. Let $\widehat{\boldsymbol{\beta}}(w)$ be the debiased pacing multiplier.

Step 4. Revenue estimator. Estimate $\widehat{REV}_k(w)$ by

$$\frac{1}{\pi_w(1-\alpha)t} \sum_{\tau=1}^{t_w} 1_{\Theta_k}(\theta^\tau) \max_{i\in C_k} v_i(\theta^\tau)\widehat{\beta}_i(w)\ .$$

Then $\widehat{\tau}_k = \widehat{REV}_k(1) - \widehat{REV}_k(0)$ and a variance estimate can be obtained by either asymptotic normality, confidence region of $\boldsymbol{\beta}$, or bootstrap.

G Heuristic Derivation of the Debiased Surrogate

We present a heuristic argument for Eq. (6) here; for a rigorous treatment we need a perturbation theory result which is given in lemma 1 in the appendix. Define the contaminated EG objective $H_\alpha(\boldsymbol{\beta}) = \int F(\boldsymbol{\beta},\theta)s_\alpha \,\mathrm{d}\theta$. Then $\boldsymbol{H}_\alpha = \nabla^2 H_\alpha(\boldsymbol{\beta}_\alpha^*)$. The map $(\boldsymbol{\beta},\alpha) \mapsto H_\alpha(\boldsymbol{\beta})$ has gradient $[\nabla H(\boldsymbol{\beta}^*), \int F(\boldsymbol{\beta}_\alpha^*,\theta)(s'-s)\,\mathrm{d}\theta]^\intercal$ and Hessian $[\boldsymbol{H}_\alpha, -\boldsymbol{\delta}_\alpha; -\boldsymbol{\delta}_\alpha^\intercal, 0]$ at $(\boldsymbol{\beta}_\alpha^*,\alpha)$. Then we have the

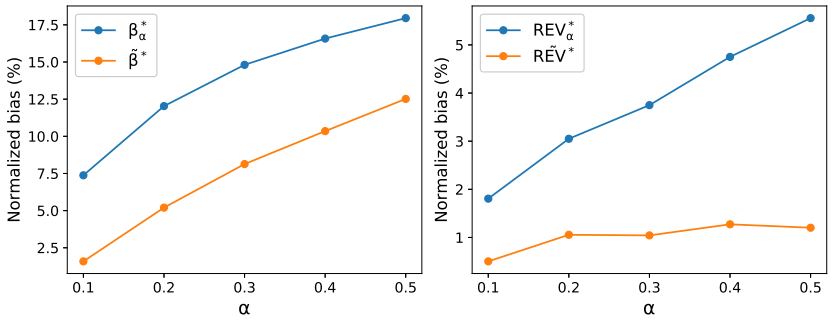

Figure 6: Normalized bias (in percent of true value) as a function of α in fully synthetic experiments. $\widetilde{\boldsymbol{\beta}}^*$ and $\widetilde{REV}^*$ are the debiased surrogates for pacing multiplier and revenue in the limit market with interference $\mathscr{M}_\alpha$.

following quadratic approximation of $H_{\alpha-\epsilon}(\boldsymbol{\beta})$

$$H_{\alpha-\epsilon}(\boldsymbol{\beta}) \approx H_\alpha(\boldsymbol{\beta}^*_\alpha) + \begin{pmatrix} \boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha \\ -\epsilon \end{pmatrix}^\intercal \begin{pmatrix} \nabla H_\alpha(\boldsymbol{\beta}^*_\alpha) \\ \int F(\boldsymbol{\beta}^*_\alpha, \theta)(s' - s)\, \mathrm{d}\theta \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha \\ -\epsilon \end{pmatrix}^\intercal \begin{pmatrix} \boldsymbol{H}_\alpha & -\boldsymbol{\delta}_\alpha \\ -\boldsymbol{\delta}_\alpha^\intercal & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha \\ -\epsilon \end{pmatrix} .$$

And so

$$\boldsymbol{\beta}^*_{\alpha-\epsilon} = \operatorname*{arg\,min}_{\boldsymbol{\beta}\in(0,1]^n} H_{\alpha-\epsilon}(\boldsymbol{\beta}) \approx \operatorname*{arg\,min}_{\boldsymbol{\beta}\in(0,1]^n} (\nabla H_\alpha(\boldsymbol{\beta}^*_\alpha) + \epsilon\boldsymbol{\delta}_\alpha)^\intercal \boldsymbol{\beta} + \tfrac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha)^\intercal \boldsymbol{H}_\alpha (\boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha) ,$$

where we dropped terms that do not involve $\boldsymbol{\beta}$.

Next we need two observations. For ϵ small enough, we have $\boldsymbol{\beta}^*_{\alpha-\epsilon,i} = 1$ if $\beta^*_{\alpha,i} = 1$. Intuitively, buyers in the market $\mathscr{M}_\alpha$ with $\beta^*_{\alpha,i} = 1$ have some nonzero leftover budget, which acts as a buffer to small changes in supply, and thus their pacing multipliers remain one. Next, we also note $\boldsymbol{\beta} \mapsto \nabla H_\alpha(\boldsymbol{\beta}^*_\alpha)^\intercal \boldsymbol{\beta}$ is constant on the set $\{\boldsymbol{\beta} : \beta_i = 1 \text{ if } \beta^*_{\alpha,i} = 1\}$ by complementary slackness. Then we further simplify and obtain

$$\boldsymbol{\beta}^*_{\alpha-\epsilon} \approx \operatorname*{arg\,min}_{\beta_i = 1 \text{ if } \beta^*_{\alpha,i}=1} \epsilon\boldsymbol{\delta}_\alpha^\intercal \boldsymbol{\beta} + \tfrac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha)^\intercal \boldsymbol{H}_\alpha (\boldsymbol{\beta} - \boldsymbol{\beta}^*_\alpha) .$$

The right-hand side is just a linearly constrained quadratic optimization problem, which admits a closed-form solution $\boldsymbol{\beta}^*_\alpha - \epsilon(\boldsymbol{P}_\alpha \boldsymbol{H}_\alpha \boldsymbol{P}_\alpha)^\dagger \boldsymbol{\delta}_\alpha$.

H Fully synthetic experiments

In the fully synthetic experiments, we simulate 10 buyers and 1000 *good* items in two submarkets, with a varying number of *bad* items (up to 500) in order to study the effect of the contamination parameter α. For each α, we randomly sample a budget from a uniform distribution for each buyer, and compute $\boldsymbol{\beta}^*$ and REV^* from the limit pure market $\mathscr{M}_0$ with a uniformly distributed value function in each submarket. Fig 6 shows the normalized bias curves as a function of α for the fully synthetic experiment, which is similar to that from semi-synthetic experiments in Fig 4.

We also check the coverage of the proposed variance estimator in fully synthetic experiments. The coverage of $\widehat{\boldsymbol{\beta}}$ and revenue are estimated in the same way as in the semi-synthetic ones. It is shown in Fig 7 that both CIs converges to the true value from the limit market with interfrence $\mathscr{M}_\alpha$ as the number of items goes to infinity. Then, in Table 2, we show that the bootstrap CI width is much smaller compared to that from Eq. (22). However, unlike in the semi-synthetic experiments, both CIs are over-covered, higher than the nominal level (95%). It is expected though that the CI from Eq. (22) is relatively large and hence conservative since it uses the min and max of each β_i,

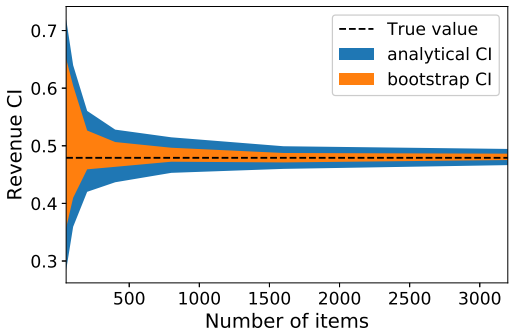

Figure 7: Revenue confidence intervals as a function of the number of items in fully synthetic experiments. The analytic CI comes from Eq. (22). The true value is the debiased surrogates for revenue in the limit market with interference $\mathscr{M}_\alpha$.

hence likely living outside of the confidence region (n-dimension ellipsoid) for $\widehat{\boldsymbol{\beta}}$ and leading to over-coverage. For the bootstrap CI, it requires enough bootstrap samples to fully explore the whole boundary of the confidence region of $\widehat{\boldsymbol{\beta}}$, as well as an accurate estimate of the asymptotic distribution of $\widehat{\boldsymbol{\beta}}$. In our fully synthetic simulation, our hypothesis is that the over-coverage of the bootstrap CI is mainly due to the over-estimation of the variance for $\widehat{\boldsymbol{\beta}}$.

	$\widehat{\boldsymbol{\beta}}$	$\widehat{REV}$			
α	coverage	CI width	CI width (b)	coverage	coverage (b)
100/1100	0.993	0.132	0.043	1.0	1.0
200/1200	0.994	0.141	0.059	1.0	1.0
300/1300	0.980	0.107	0.059	1.0	1.0
400/1400	0.983	0.098	0.064	1.0	1.0
500/1500	0.971	0.108	0.069	1.0	1.0

Table 2: Coverage of $\widehat{\boldsymbol{\beta}}$ and revenue estimates in fully synthetic experiments. All quantities are averaged over 100 simulations for each α. The coverage of $\widehat{\boldsymbol{\beta}}$ is averaged over all components of $\widehat{\boldsymbol{\beta}}$. For revenue estimates, the columns with "(b)" represent the quantities from the bootstrap CI, while the columns without "(b)" are for the CI from Eq. (22). The CI widths are normalized by the revenue from the limit market $\mathscr{M}_\alpha$. α is expressed as the ratio of the number of bad items and the total items.