Testing Dynamic Incentive Compatibility in Display Ad Auctions

Yuan Deng
Duke University
Durham, NC
ericdy@cs.duke.edu

Sébastien Lahaie
Google Research
New York City, NY
slahaie@google.com

ABSTRACT

The question of transparency has become a key point of contention between buyers and sellers of display advertising space: ads are allocated via complex, black-box auction systems whose mechanics can be difficult to model let alone optimize against. Motivated by this concern, this paper takes the perspective of a single advertiser and develops statistical tests to confirm whether an underlying auction mechanism is dynamically incentive compatible (IC), so that truthful bidding in each individual auction and across time is an optimal strategy. The most general notion of dynamic-IC presumes that the seller knows how buyers discount future surplus, which is questionable in practice. We characterize dynamic mechanisms that are dynamic-IC for all possible discounting factors according to two intuitive conditions: the mechanism should be IC at each stage in the usual sense, and expected present utility (under truthful bidding) should be independent of past bids. The conditions motivate two separate experiments based on bid perturbations that can be run simultaneously on the same impression traffic. We provide a novel statistical test of stage-IC along with a test for utility-independence that can detect lags in how the seller uses past bid information. We evaluate our tests on display ad data from a major ad exchange and show how they can accurately uncover evidence of first- or second-price auctions coupled with dynamic reserve prices, among other types of dynamic mechanisms.

KEYWORDS

incentive compatibility, dynamic mechanism, display advertising, hypothesis test

1 INTRODUCTION

The issue of transparency has emerged as a major driver of change in the display advertising industry. Advertisers increasingly want to understand how exactly their advertising dollars flow from their ad campaigns through the ecosystem (demand- and sell-side platforms, exchanges) to reach the publishers who show their ads. One

key aspect of transparency concerns the auction mechanism used to allocate ad impressions, and whether it is incentive compatible (i.e., truthful bidding is optimal) [7]. Traditionally exchanges have used second-price auctions, which are truthful in isolation, but because ad auctions are repeated sellers have the ability to apply various policies that complicate incentives over time. One canonical example is dynamic reserve prices, where past bids are used to set present reserve prices [12]. Another example is throttling: exchanges need to decide who to solicit bids from, and may decline to send calls to bidders with historically low bids. Intentionally or not, these kinds of policies complicate bidding incentives because bids now have bearing on both present and future surplus.

Characterization. This paper takes the perspective of a single advertiser and asks whether it is empirically possible, via experimentation, to test whether an underlying auction mechanism is *dynamically incentive compatible* (dynamic-IC). A mechanism is dynamic-IC if bidding truthfully for each impression optimizes present and future utility combined (assuming continued truthful bidding). One approach to test dynamic-IC would be to list different policies that the seller might use, like dynamic reserves or throttling, and develop tests for each of these, but this can turn into a tedious game of cat and mouse. Instead, our approach is to consider a broad and practically-relevant subclass of dynamic-IC mechanisms, characterize its properties, and develop tests for these properties. Specifically, we consider the class of mechanisms that are dynamic-IC no matter how advertisers might time-discount future surplus. Since ad exchanges are not communicated the individual advertisers' time-discount factors, this plausibly covers the space of dynamic-IC mechanisms they might run.[1] Our first contribution is to show that this class is equivalent to the class of *history-independent mechanisms*, which we characterize by two conditions: each individual auction is incentive compatible in isolation (stage-IC), and expected present utility under truthful bidding is independent of past bids (utility independence).

In fact, the class of history-independent mechanisms does not just cover the mechanisms that exchanges could plausibly run, it also includes mechanisms they may plausibly *want* to run. Mirrokni et al. [16, 17, 19] show that any dynamic-IC and ex-post individually rational mechanism can be converted into a special class of dynamic mechanism, called a *bank account mechanism*, without any loss of efficiency or revenue. In a bank account mechanism, the bid history is summarized by a scalar called the bank account balance, and each stage mechanism is parameterized by the balance only. Moreover, the parameterized stage mechanisms are all stage-IC and the buyer's expected utility is independent of the balance for each stage; the latter is essentially the property of utility independence. As a

---

[1]An exchange could try to model and estimate time-discount factors, but this would be challenging because advertisers typically bid through intermediaries like agencies or demand-side platforms.

result, bank account mechanisms are history-independent, by our characterization, and the class of history-independent mechanisms includes mechanisms that maximize efficiency or revenue.

Testing Methods. Our characterization motivates two tests that can jointly confirm or reject whether the advertiser is participating in a history-independent, dynamic-IC mechanism. By a *test* we mean an experiment where random noise is applied to bids in a systematic way, followed by a statistical comparison of auction results at different noise levels (including the control level where no noise is applied). Other kinds of perturbations would be possible (e.g., random bid throttling), but we only make use of bid perturbations for our tests.

Our second contribution is a novel test of stage-IC based on Myerson's classic characterization of the relationship between allocation and payment rules of IC auctions [20]. The test is based on an experiment with two treatment conditions where the advertiser randomly perturbs its bids upwards or downwards on a query by query basis. The magnitude of the perturbations controls the trade-off between the power of the test to detect stage-IC violations and the cost of running the test in terms of foregone surplus.

Our third contribution is a test of utility independence (UI). To apply the test, the advertiser partitions the auctions it participates in into *buckets* defined by observable features (e.g., publisher, country, device type), and systematically perturbs its bids within buckets upwards or downwards. It then checks whether future utility within buckets correlates with the applied perturbations. Utility independence is a powerful concept because it captures the common channel through which various dynamic policies can impact an advertiser's incentives across time. If a seller uses dynamic reserve prices, then higher present bids will translate to higher future reserves, and lower future utility; the utility independence test will show negative correlation. If a seller throttles the advertiser in the future based on low present bids, then the test will show a positive correlation between bids and utility. In both cases, the advertiser can establish that dynamic-IC is violated with the same test.

To implement our test, the way in which the advertiser buckets queries must be aligned with how the seller segments the query traffic; as Lahaie et al. [14] show, this is necessary to conduct a test with conclusive results. Other than this, our testing methods rely only on the advertiser observing its realized allocation and payment, nothing more. We assume that this is possible at the auction level, which is consistent with real-time bidding standards. For instance, the Open RTB standard includes *win notice* and *billing notice* messages that transmit an auction id and settlement price [13]. Comparing to the testing methods developed in Lahaie et al. [14], we do not rely on observations of other auction parameters like reserve prices, so our tests are robust to various information scenarios where auction parameters may be censored to different extents.

Experiments. We evaluate the effectiveness of our tests using real bid data from display ad auctions logs collected from the Google Ad Exchange over the course of a week. We run counterfactual scenarios where the seller uses second-price auctions, first-price auctions, and mixtures of the two to apply the stage-IC test. Simultaneously, the seller also applies dynamic reserves to demonstrate how the UI test can detect them when they interact with different auction formats. We also apply the tests on a stateful mechanism, inspired by the theory of bank account mechanisms [19], that is

completely different from dynamic reserves, in order to illustrate the time scales over which the UI test can prove effective.

Related Work

Lahaie et al. [14] initiated the study of testing IC in dynamic display ad auctions, and our work builds on many of their ideas. In particular, they introduce the approach of segmenting traffic into buckets and systematically perturbing bids in buckets to examine long-run effects on future auctions; we use the same setup for our UI test. They provide tests for stage-IC and for dynamic reserve pricing. Our tests parallel theirs but differ in crucial respects. Our tests are not just separately interesting, instead we show that they logically go together to test for dynamic IC. Our stage-IC test only requires three treatments, while their stage-IC test requires one treatment per shading factor (they use 10 for their experiments). Most importantly, our UI-test covers a wide range of possible stateful policies, of which dynamic reserves are just one special case.

Dynamic Mechanisms. The field of dynamic mechanism design has attracted a growing body of literature in recent years motivated in part by applications to online advertising markets. We briefly discuss research that is closely related to ours. Readers are encouraged to refer to [5] for a comprehensive survey. Bergemann and Välimäki [6] generalize the VCG mechanism to design a dynamic pivot mechanism that achieves welfare optimality. Kakade et al. [11] extend the idea of dynamic pivot mechanisms to design a virtual-pivot mechanism by incorporating the idea of virtual valuations [20], which is proved to be revenue-optimal in certain dynamic environments that are "separable".

Recent inquiries to study revenue-optimal dynamic mechanisms were initiated by Papadimitriou et al. [21]. They provide an example to show that the revenue gap between static and dynamic mechanisms can be arbitrarily large. Ashlagi et al. [2] and Mirrokni et al. [17] simultaneously and independently design a fully polynomial-time approximation scheme to compute the optimal dynamic mechanism. Mirrokni et al. [18] apply the idea of bank account mechanisms, a special class of dynamic mechanism that is proved to be revenue-optimal [16, 17, 19], to design a practical dynamic mechanism, and provide empirical evidence of huge revenue lifts over static mechanisms.

Strategic Buyers. Our research is also related to the study of the buyer's strategic behavior in online auctions. Edelman and Ostrovsky [8] provide empirical evidence that buyers do act strategically in online advertising markets. From the seller's perspective, he would like to design a dynamic mechanism that is robust against the buyer's strategic behavior. Kanoria and Nazerzadeh [12] design an almost dynamic-IC mechanism that changes the reserve price based on the buyer's historical bids such that the buyer's gain from misreporting is negligible. Amin et al. [1] and Medina and Mohri [15] initiate the study of robust price learning in a setting where the seller needs to design a policy to set the reserve price and learn a private factor related to the buyer's valuations. In such a setting, the buyer may act strategically in order to hide his private information or provide misleading information to the seller. In addition, they design a robust learning policy that achieves sublinear revenue loss (no-regret) against a static mechanism that knows the private factor in advance, when the buyer is less patient than the seller. In

particular, Amin et al. [1] show that no learning policy is no-regret when the buyer is as patient as the seller. All these mechanisms and policies are variants of dynamic reserve pricing mechanisms, while our characterization of history-independent mechanisms can provide guidance to sellers on how to design other kinds of dynamic-IC mechanisms that are robust against strategic buyers.

From the buyer's perspective, she would like to find out the optimal strategy against the mechanism used by the seller when the mechanism is not IC. In practice, most online mechanisms are not dynamic-IC when the buyers are financially constrained, even for standard repeated static mechanisms like the second price auction or the generalized second price auction [9]. Balseiro et al. [3] study the effect of different budget management schemes and Balseiro and Gur [4] propose an asymptotically optimal bidding strategy when the online mechanisms are static. Our testing methods for dynamic-IC can serve as a first step for the buyer to figure out her optimal bidding strategy in dynamic mechanisms.

2 PRELIMINARIES

Consider a buyer participating in a dynamic auction with a single seller, for example, an ad exchange. The dynamic auction consists of a sequence of *queries*, in which the seller has an opportunity to display an ad to a user visiting the website. The queries are filled in an online manner such that the ad slot must be sold once it arrives. We consider a setting where the queries are partitioned into *buckets* depending on the features of the buyer and the queries. The general dynamic auction mechanism is the same in each bucket, but the actual auction state and parameters are kept independent across different buckets. For instance, the dynamic auction might be run independently on a website-by-website basis.

We assume that the buyer's valuation $v_t \in V = [0, 1]$ in each bucket for the t-th query is drawn independently from a continuous distribution F_t with density f_t. The distributions F_t are not necessarily identical within a bucket and the distributions can be different across different buckets. At stage t, the buyer first draws her valuation v_t from the distribution F_t and then submits a bid b_t to the seller. After receiving the bid b_t, the seller implements an outcome with an allocation probability and a payment.

In general, a dynamic auction $\langle x, p \rangle$ is represented by a family of allocation functions $x_t : V^t \rightarrow [0, 1]$ that maps the buyer's reported bids from stage 1 to stage t to an allocation probability for the t-th query and a family of payment functions $p_t : V^t \rightarrow \mathbb{R}$ that maps the buyer's reported bids from stage 1 to stage t to a payment at stage t. We say that $\langle x_t, p_t \rangle$ is the *stage mechanism* of the dynamic mechanism for stage t. Since our study is from the perspective of a single buyer, the allocation function x_t and the payment function p_t here subsume the other players' information and bidding behavior. This does not mean that the other players are assumed to play some equilibrium, or that their behavior is controlled or held fixed in some way, only that the randomness in the x_t and p_t includes mechanism randomness on the part of the seller and any randomness in the opposing players' strategies.[2]

[2]In particular, this means that any violations of dynamic-IC could be attributed to the opposing player's strategies rather than the seller's mechanism itself. For instance, our tests would detect "bid jamming" in a second-price auction where an opponent tries to bid just below the winning bid to increase the price, based on past bidding data [10]. Without observing opponent bids, it is not possible to disentangle violations

In line with the literature, we assume that the buyer's utility is quasi-linear such that the utility u_t of the buyer with true valuation v_t at stage t is $u_t(b_{(1,t)}; v_t) = v_t \cdot x_t(b_{(1,t)}) - p_t(b_{(1,t)})$, where we use the notation $b_{(t', t'')}$ to represent the buyer's bids from stages t' to t'', and similarly for $v_{(t', t'')}$. Usually the buyer is less patient than the seller, so we make the standard assumption that at stage t the buyer aims to maximize her utility discounted by a factor $\gamma \in [0, 1]$. Therefore, for a bidding sequence $b_{(1, T)}$, the buyer's maximization objective at stage t is given by $\sum_{\tau=t}^{T} \gamma^{\tau-t} \cdot u_\tau(b_{(1, \tau)}; v_\tau)$.

2.1 Dynamic Incentive-Compatibility

A mechanism is *dynamic incentive-compatible* (dynamic-IC) if reporting truthfully at every stage is the optimal strategy that maximizes the buyer's utility. Dynamic incentive-compatibility can be defined by backward induction: in the last stage T, the buyer must be incentivized to report truthfully no matter what the historical bids $b_{(1, T-1)}$ are, or more precisely:

$$v_T \in \arg\max_{b_T} u_T(b_{(1, T-1)}, b_T; v_T)$$

for all v_t and $b_{(1, T-1)}$. For the induction step, at stage t the buyer should have the incentive to report truthfully no matter what the historical bids $b_{(1, t-1)}$ are, conditioned on the fact that the buyer reports truthfully in the remaining stages:

$$v_t \in \arg\max_{b_t} u_t(b_{(1, t-1)}, b_t; v_t) + U_t(b_{(1, t-1)}, b_t)$$

for all v_t and $b_{(1, t-1)}$, where

$$U_t(b_{(1,t)}) = \sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot \mathbb{E}_{v_\tau} \left[u_\tau(b_{(1,t)}, v_{(t+1, \tau)}; v_\tau) \right]$$

represents the buyer's expected *continuation utility* assuming she reports truthfully in the future.

2.2 Stage Incentive-Compatibility

In particular, for a myopic buyer with $\gamma = 0$ the dynamic-IC criterion simply requires that the stage mechanism $\langle x_t, p_t \rangle$ is stage incentive-compatible (stage-IC) for every stage:

$$v_t \in \arg\max_{b_t} u_t(b_{(1, t-1)}, b_t; v_t)$$

for all t, v_t, and $b_{(1, t-1)}$.

Consider a stage-IC mechanism $\langle x, p \rangle$ with the allocation rule $x : V \rightarrow [0, 1]$, payment rule $p : V \rightarrow [0, 1]$, yielding the utility function $u(v) = v \cdot x(v) - p(v)$. The celebrated Myerson's Lemma [20] provides a characterization of the relationship between utility, allocation probability, and payment for a stage-IC mechanism:

$$\frac{dx(v)}{dv} \geq 0, \ \frac{du(v)}{dv} = x(v) \text{ and } \frac{dp(v)}{dv} = v \frac{dx(v)}{dv}.$$

These relationships will form the basis of our analysis.

of dynamic-IC due to the seller or due to opposing buyers based solely on observations of x_t and p_t.

3 HISTORY INDEPENDENT MECHANISMS

Because the buyer participates in repeated auctions, our goal is to develop tests not just for stage-IC, but dynamic-IC. In principle, there are two kinds of mechanisms that are dynamic-IC. One kind is a mechanism that is stage-IC and static, i.e., the stage mechanism is independent of the historical bids. Since the stage mechanisms do not rely on the buyer's reported bids in the past, the buyer's misreport has no effect on the future. Therefore, once the mechanism can ensure stage-IC for every stage mechanism, the entire mechanism is dynamic-IC. On the other hand, for a dynamic mechanism in which the stage mechanisms do depend on the historical bids, the mechanism must use the historical bids in an *incentive-compatible way*. Based on these observations, we introduce the concept of a *history-independent mechanism*:

Definition 3.1 (History Independent Mechanism). A dynamic mechanism $\langle x, p \rangle$ is history independent if for each stage t,

- the stage mechanism $\langle x_t, p_t \rangle$ is stage-IC;
- assuming the buyer reports truthfully at stage t, the expected utility of the buyer at stage t is independent of her historical bids: for all $b_{(1,t-1)}$, $\mathbb{E}_{v_t}[u_t(b_{(1,t-1)}, v_t; v_t)] = c_t$ where c_t is a constant independent of $b_{(1,t-1)}$.

For convenience, we will refer to the second property as *utility independence*, abbreviated UI.

Recall that a general dynamic mechanism could depend on the buyer's discounting factor γ, which can play an important role in the mechanism's incentive properties. In practice, however, a seller may have little information on how specific buyers weigh future utility—an ad exchange can interact with thousands of buyers, and via intermediaries like demand-side platforms. In this environment, the only way for the seller to induce truthful bidding is to use a mechanism that is dynamic-IC no matter what the buyer's discounting factor might be.

We show that the class of dynamic mechanisms that are dynamic-IC for all discounting factors γ corresponds to the class of history-independent mechanisms. In light of this result, our tests will test for the defining properties of this class of mechanisms.

THEOREM 3.2. *A dynamic mechanism $\langle x, p \rangle$ is dynamic-IC for all $\gamma \in [0, 1]$ if and only if $\langle x, p \rangle$ is a history independent mechanism.*

PROOF. **Sufficiency**: For a history independent mechanism $\langle x, p \rangle$, since the mechanism is stage-IC, the buyer is incentivized to report truthfully in the last stage. By a backward induction, for stage t, dynamic-IC requires that

$$v_t \in \arg\max_{b_t} u_t\left(b_{(1,t-1)}, b_t; v_t\right) + U_t\left(b_{(1,t-1)}, b_t\right)$$

for all v_t and $b_{(1,t-1)}$, where $U_t(b_{(1,t)})$ is the continuation utility. Since the mechanism is UI for every stage, we have for any $b_{(1,t)}$,

$$U_t(b_{(1,t)}) = \sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot \mathbb{E}_{v_\tau \sim F_\tau}\left[u_\tau\left(b_{(1,t)}, v_{(t+1,\tau)}; v_\tau\right)\right]$$

$$= \sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot c_\tau$$

Notice that $\sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot c_\tau$ depends on γ but is independent of $b_{(1,t)}$. Therefore, for a buyer with discounting factor γ, dynamic-IC requirement is equivalent to $v_t \in \arg\max_{b_t} u_t\left(b_{(1,t-1)}, b_t; v_t\right)$, which is exactly the requirement of stage-IC. Since the mechanism is stage-IC for stage t, the buyer is incentivized to report truthfully at stage t. Thus, a history independent mechanism is dynamic-IC for all $\gamma \in [0, 1]$.

Necessity: First, notice that for a dynamic mechanism $\langle x, p \rangle$ to be dynamic-IC for $\gamma = 0$, the dynamic mechanism must be stage-IC. We prove by a backward induction on t from $t = T$ to $t = 1$ to show that for all $\tau \geq t$ and $b_{(1,\tau-1)}$,

$$\mathbb{E}_{v_\tau}[u_\tau(b_{(1,\tau-1)}, v_\tau; v_\tau)] = \zeta_\tau^t(b_{(1,t-1)})$$

where ζ_τ^t is a function depending on $b_{(1,t-1)}$ only. In other words, the induction hypothesis at stage t assumes that for all stage $\tau \geq t$, the expected utility of the buyer at stage τ is dependent on $b_{(1,t-1)}$ but independent of $b_{(t,\tau-1)}$. The mechanism is UI when the induction hypothesis is true for $t = 1$.

The base case when $t = T$ is clearly true by defining

$$\zeta_T^T(b_{(1,T-1)}) = \mathbb{E}_{v_T}[u_T(b_{(1,T-1)}, v_T; v_T)].$$

Assume the induction hypothesis is true for all $t' > t$ and we consider the case t. Notice that, for stage t, the dynamic-IC requirement is equivalent to a first order condition such that for all v_t,

$$0 = \left.\frac{\partial\left\{u_t\left(b_{(1,t)}; v_t\right) + U_t\left(b_{(1,t)}\right)\right\}}{\partial b_t}\right|_{b_t=v_t}$$

For a stage-IC mechanism, by Myerson's lemma, the payment function and the allocation rule must satisfy $\frac{\partial p_t(b_{(1,t)})}{\partial b_t} = b_t \cdot \frac{\partial x_t(b_{(1,t)})}{\partial b_t}$. Therefore, we have

$$\frac{\partial u_t\left(b_{(1,t)}; v_t\right)}{\partial b_t} = v_t \cdot \frac{\partial x_t\left(b_{(1,t)}\right)}{\partial b_t} - \frac{\partial p_t\left(b_{(1,t)}\right)}{\partial b_t}$$

$$= (v_t - b_t) \cdot \frac{\partial x_t\left(b_{(1,t)}\right)}{\partial b_t}$$

Plugging the above formula to the first order condition, we have

$$0 = \left.\frac{\partial\left\{u_t\left(b_{(1,t)}; v_t\right) + U_t\left(b_{(1,t)}\right)\right\}}{\partial b_t}\right|_{b_t=v_t} = \left.\frac{\partial U_t(b_{(1,t)})}{\partial b_t}\right|_{b_t=v_t}$$

for all v_t. Thus, to satisfy the first order condition, we have,

$$\forall b_{(1,t)}, \frac{\partial U_t(b_{(1,t)})}{\partial b_t} = 0$$

The continuation utility $U_t(b_{(1,t)})$ can be written as

$$U_t(b_{(1,t)}) = \sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot \mathbb{E}_{v_{(t+1,\tau)}}\left[u_\tau\left(b_{(1,t)}, v_{(t+1,\tau)}; v_\tau\right)\right]$$

$$= \sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot \zeta_\tau^{t+1}(b_{(1,t)})$$

where the second equality follows the fact that the induction hypothesis is true for $t+1$. Combining with the fact that $\frac{\partial U_t(b_{(1,t)})}{\partial b_t} = 0$ for all $b_{(1,t)}$, we have for all $\gamma \in [0, 1]$,

$$\sum_{\tau=t+1}^{T} \gamma^{\tau-t} \cdot \frac{\partial \zeta_\tau^{t+1}(b_{(1,t)})}{\partial b_t} = 0$$

which implies that for all $\tau > t$ and $b_{(1,t)}$, $\frac{\partial \zeta_\tau^{t+1}(b_{(1,t)})}{\partial b_t} = 0$. Therefore, we can define $\zeta_\tau^t(b_{(1,t-1)})$ such that for $\tau > t$,

$$\zeta_\tau^t(b_{(1,t-1)}) \doteq \zeta_\tau^{t+1}(b_{(1,t-1)}, 0) = \zeta_\tau^{t+1}(b_{(1,t)})$$

and moreover, for $\tau = t$, $\zeta_t^t(b_{(1,t-1)}) = \mathbb{E}_{v_t}[u_t(b_{(1,t-1)}, v_t; v_t)]$, which concludes the proof of the induction. $\qquad\square$

As mentioned previously, the class of history-independent mechanisms includes bank account mechanisms [16, 17]. By the results of Mirrokni et al. [19], this implies that the class includes mechanisms that optimize revenue or social welfare.

4 EXPERIMENTAL DESIGN

Following the characterization in Section 3, we focus on designing two separate tests, one for stage-IC and one for UI, to determine whether the seller is running a history-independent mechanism. More specifically, our approach is to assume that the conditions hold as the null hypothesis, and we test for violations of the conditions at desired significance levels.

Note that to test stage-IC for a stage mechanism, it is necessary to interact with the same stage mechanism multiple times. However, even when the format of the stage mechanisms is fixed in a period of time (e.g., a second price auction), the stage mechanism $\langle x_t, p_t \rangle$ is different query by query from the buyer's perspective since $\langle x_t, p_t \rangle$ depends on the other buyers' bids on the query. Instead of checking whether the stage mechanism is stage-IC for each query individually, we design a test to check whether for a longer timescale, such as an hour or a day, the mechanism is stage-IC over the queries in aggregate. In particular, we design a statistical test to test whether the seller is running a mechanism that is stage-IC for each day and UI across days. Formally, let $\{B_1, \cdots, B_n\}$ denote the partition of queries into n buckets and let I_k^j be the set of queries on day k in bucket j. We represent the stage mechanism for the i-th query on day k in the j-th bucket as $\langle x_i^k(\cdot; j), p_i^k(\cdot; j) \rangle$ and the buyer's valuation for this query as $v_i^k(j)$.

Following Lahaie et al. [14], our tests are based on bid perturbations within buckets of traffic. Lahaie et al. [14] uncover two important pitfalls of these kinds of tests. First, additive perturbations can have counter-intuitive effects, such as optimal reserve prices that decrease even though bids are perturbed upwards. Second, if the test buckets are not aligned with the way the seller segments traffic to run its mechanism, results will be inconclusive. We therefore apply multiplicative bid perturbations for our tests and assume that the partition of the query traffic is consistent (i.e., the same or coarser) with the one used by the seller to run independent versions of the mechanism.

4.1 Stage Incentive Compatibility Test

We propose our testing method to test stage-IC for a specific day k. For simplicity, we will omit the dependence on the bucket in this section. Let $\alpha \in [0, 1]$ denote the bid perturbation magnitude. For the i-th query, we perturb the true valuation v_i^k by a multiplicative factor α_i uniformly drawn from $\{1 - \alpha, 1, 1 + \alpha\}$ and submit the bid $b_i^k = \alpha_i v_i^k$. Let $\hat{u}_i^k(b_i^k) = b_i^k \cdot x_i^k(b_i^k) - p_i^k(b_i^k)$ be the utility with respect to the submitted bid rather than the utility with respect to the

buyer's true valuation, where we drop $\langle x_i^k(\cdot), p_i^k(\cdot) \rangle$'s dependence on the bids before the i-th query on day k for simplicity.

LEMMA 4.1. *For a stage-IC mechanism $\langle x_i^k, p_i^k \rangle$, for all v_i^k:*

$$\frac{\hat{u}_i^k(v_i^k) - \hat{u}_i^k((1-\alpha)v_i^k)}{\alpha} \leq v_i^k \cdot x_i^k(v_i^k) \leq \frac{\hat{u}_i^k((1+\alpha)v_i^k) - \hat{u}_i^k(v_i^k)}{\alpha}.$$

PROOF. For a stage-IC mechanism, by Myerson's lemma:

$$\hat{u}_i^k(v_i^k) - \hat{u}_i^k((1-\alpha)v_i^k) = \int_{(1-\alpha)v_i^k}^{v_i^k} x_i^k(v) dv$$

$$\leq \int_{(1-\alpha)v_i^k}^{v_i^k} x_i^k(v_i^k) dv = \alpha \cdot v_i^k \cdot x_i^k(v_i^k)$$

where the first equality follows that $\frac{d\hat{u}_i^k(v)}{dv} = x_i^k(v)$ and the inequality is due to $x_i^k(v)$ is non-decreasing. Similarly, for the right hand side of the inequality, we have

$$\hat{u}_i^k((1+\alpha)v_i^k) - \hat{u}_i^k(v_i^k) = \int_{v_i^k}^{(1+\alpha)v_i^k} x_i^k(v) dv$$

$$\geq \int_{v_i^k}^{(1+\alpha)v_i^k} x_i^k(v_i^k) dv = \alpha \cdot v_i^k \cdot x_i^k(v_i^k) \qquad \square$$

By Lemma 4.1, taking the summation over all the queries on day k and the expectation of v_i^k, we have

$$\frac{\sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k(v_i^k)\right] - \sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k((1-\alpha)v_i^k)\right]}{\alpha|I_k|}$$

$$\leq \frac{\sum_i \mathbb{E}_{v_i^k}\left[v_i^k \cdot x_i^k(v_i^k)\right]}{|I_k|}$$

$$\leq \frac{\sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k((1+\alpha)v_i^k)\right] - \sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k(v_i^k)\right]}{\alpha|I_k|} \qquad (1)$$

where I_k is the set of queries on day k. We test the property of stage-IC by checking whether (1) holds. To do so, we estimate

$$\hat{u}_0 = \frac{\sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k(v_i^k)\right]}{|I_k|}, \quad \hat{w}_0 = \frac{\sum_i \mathbb{E}_{v_i^k}\left[v_i^k \cdot x_i^k(v_i^k)\right]}{|I_k|},$$

$$\hat{u}_- = \frac{\sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k((1-\alpha)v_i^k)\right]}{|I_k|}, \quad \hat{u}_+ = \frac{\sum_i \mathbb{E}_{v_i^k}\left[\hat{u}_i^k((1+\alpha)v_i^k)\right]}{|I_k|}$$

separately. Denote sets $I_k^-, I_k^\perp$ and I_k^+ such that $I_k^- = \{i \in I_k | \alpha_i = 1 - \alpha\}$, $I_k^\perp = \{i \in I_k | \alpha_i = 1\}$, and $I_k^+ = \{i \in I_k | \alpha_i = 1 + \alpha\}$. The estimates of $\hat{u}_0, \hat{w}_0, \hat{u}_-, \hat{u}_+$ are obtained as follows

$$\tilde{u}_0 = \frac{1}{|I_k^\perp|} \sum_{i \in I_k^\perp} \hat{u}_i^k(b_i^k), \quad \tilde{w}_0 = \frac{1}{|I_k^\perp|} \sum_{i \in I_k^\perp} b_i^k \cdot x_i^k(b_i^k),$$

$$\tilde{u}_- = \frac{1}{|I_k^-|} \sum_{i \in I_k^-} \hat{u}_i^k(b_i^k), \quad \tilde{u}_+ = \frac{1}{|I_k^+|} \sum_{i \in I_k^+} \hat{u}_i^k(b_i^k). \qquad (2)$$

We will then check whether the following inequalities holds:

$$\frac{\tilde{u}_0 - \tilde{u}_-}{\alpha} \leq \tilde{w}_0 \leq \frac{\tilde{u}_+ - \tilde{u}_0}{\alpha}. \qquad (3)$$

The estimators $\tilde{u}_0, \tilde{w}_0, \tilde{u}_-, \tilde{u}_+$ are unbiased estimators of $\hat{u}_0, \hat{w}_0, \hat{u}_-$, and $\hat{u}_+$, respectively. Moreover, under a mild assumption that

the buyer's maximum valuation for each query is bounded by a constant, the standard error for each estimator is $O(\frac{1}{\alpha} \cdot |I_k|^{-1/2})$ with high probability for sufficiently large $|I_k|$.

The standard error demonstrates a trade-off between the confidence of the test and the buyer's revenue loss from implementing the test. Under the null hypothesis of a stage-IC mechanism, the foregone surplus from perturbing bids should increase with α, while the width of the confidence intervals decreases.

4.2 Utility Independence Test

We next propose a testing method to test whether the expected utility in day k is independent of the bids in day 0. Let $\beta \in [0, 1]$ denote the bid perturbation magnitude. For the j-th bucket, β_j is drawn uniformly from $[1 - \beta, 1 + \beta]$ and we perturb all the bids of bucket j on day 0 systematically by a multiplicative factor β_j; specifically, we submit bid $b_i^0(j) = \beta_j \cdot v_i^0(j)$ for all i. We then report truthfully between day 1 and day k for all queries in all buckets.

Denote the buyer's average utility on day k in bucket j as $\bar{u}_k(j)$:

$$\bar{u}_k(j) = \frac{1}{|I_k^j|} \cdot \sum_i \left[v_i^k(j) \cdot x_i^k \left(v_i^k(j); j \right) - p_i^k \left(v_i^k(j); j \right) \right] \quad (4)$$

where we omit the dependence of $\langle x_i^k(\cdot; j), p_i^k(\cdot; j) \rangle$ on the buyer's bids before the i-th query on day k for simplicity.

LEMMA 4.2. *For a history-independent mechanism, we have*

$$\forall j, \ \mathbb{E}\left[(\beta_j - 1) \cdot \bar{u}_k(j) \right] = 0.$$

PROOF. For a history-independent mechanism, the expected utility of the buyer on day k is independent of the historical bids in the past days. Therefore, the expected utility of the buyer on day k is independent of any bidding strategy in the past days.

From the perspective that our perturbation is a bidding strategy to scale the bids by a factor β_j, $\mathbb{E}[\bar{u}_k(j)|\beta_j]$ should be a constant c_j independent of β_j. Therefore,

$$\mathbb{E}\left[(\beta_j - 1) \cdot \bar{u}_k(j) \right] = \int_{1-\beta}^{1+\beta} \frac{\beta_j - 1}{2\beta} \cdot \mathbb{E}[\bar{u}_k(j)|\beta_j] \, d\beta_j$$

$$= c_j \cdot \mathbb{E}[(\beta_j - 1)] = 0$$

where $\frac{1}{2\beta}$ is the probability density for selecting β_j uniformly from $[1 - \beta, 1 + \beta]$. □

By Lemma 4.2, we have $\mu_k = \frac{1}{n} \sum_j \mathbb{E}\left[(\beta_j - 1) \cdot \bar{u}_k(j) \right] = 0$, where n is the number of buckets. We estimate μ directly from the realization of β_j and $\bar{u}_k(j)$:

$$\tilde{\mu}_k = \frac{1}{n} \sum_j (\beta_j - 1) \cdot \bar{u}_k(j). \quad (5)$$

We will then check whether $\tilde{\mu}_k$ is significantly different from 0, as $\tilde{\mu}_k$ is an unbiased estimator of μ_k. Under a mild assumption that the maximum valuation for each query is bounded by a constant, the standard error of estimator $\tilde{\mu}_k$ is $O(\beta \cdot n^{-1/2})$ with high probability for sufficiently large n. The standard error indicates that the larger β, the larger confidence interval and the larger the buyer's revenue loss from the test. However, it does not follow that the power of the test is higher for a smaller β, because a larger β could potentially increase the absolute value of the estimator when the mechanism is not UI.

4.3 Combining the Tests

We combine our testing methods for stage-IC and UI to obtain a complete test to detect whether a mechanism is history-independent. The buyer can arbitrarily select a day as day 0 and test whether the mechanism is history-independent for the next K days. More precisely, we test whether the mechanism is stage-IC in day k and that the expected utility in day k is independent of the bids in day 0, for any $1 \leq k \leq K$.

The buyer first determines the perturbation magnitude α for the stage-IC test and the perturbation magnitude β for the UI test. On day 0, the buyer draws β_j uniformly from $[1 - \beta, 1 + \beta]$ for each bucket j. The buyer then submits bid $b_i^0(j) = \beta_j \cdot v_i^0(j)$ for the i-th query of bucket j on day 0. For the next K days, the buyer draws $\alpha_i^k(j)$ uniformly from $\{1 - \alpha, 1, 1 + \alpha\}$ for the i-th query on day k in bucket j, and the buyer then submits bid $b_i^k(j) = \alpha_i^k(j) \cdot v_i^k(j)$.

For the stage-IC test for day k, the buyer computes the estimators in (2) by treating the n different buckets as a single giant bucket, and then testing whether (3) holds. As for the UI test for day k, the buyer computes the estimator $\tilde{\mu}_k$ according to (5), where the average utility $\bar{u}_k(j)$ of bucket j is computed according to (4) by using only the bids that are not perturbed. More precisely,

$$\bar{u}_k(j) = \frac{1}{|I_k^{j\perp}|} \cdot \sum_i \left[v_i^k(j) \cdot x_i^k \left(v_i^k(j); j \right) - p_i^k \left(v_i^k(j); j \right) \right]$$

where $I_k^{j\perp} = \{i \in I_k^j | \alpha_i^k(j) = 1\}$.

5 EMPIRICAL EVALUATION

To evaluate our tests we collected bid data from over 20M auction records from the logs of the Google Ad Exchange, focusing on a single large advertiser (i.e., bidder). The data consists of a random sample of second-price auctions that the bidder participated in, including auctions that were won or lost, spanning seven days. For each auction we record the advertiser's bid and all competing bids. We do not record other auction parameters like reserve prices, since our analysis involves simulating mechanisms with potentially different reserves.

We consider a setup where the exchange segments query traffic into buckets according to features like publisher identity, browser type, country, etc. Within each bucket the seller runs a dynamic mechanism to sell impressions; mechanisms in separate buckets are independent of each other. We specifically consider buckets defined according to two features: publisher id (which generally identifies a publisher like nytimes.com or sfchronicle.com), and device type (mobile or desktop). Note that the exchange is under no constraint to bucket the impression traffic uniformly. Comparing to the evaluation in Lahaie et al. [14], we do not consider various information scenarios where auction parameters like reserve prices are censored to different extents. The reason is that the conditions that we test for only rely on the advertiser observing its realized allocation and payment, and we assume that this is possible at the auction level.

In our simulations we number the days $0, 1, \ldots, 6$. On day 0 the advertiser applies bucket perturbations to evaluate UI on future days. We consider two treatments, following Lahaie et al. [14]: *exact* buckets that agree with the publisher's bucketing scheme defined by publisher id and device type, and *coarse* buckets defined only by

publisher id. Coarse buckets can still in principle detect violations of UI, but with less precision. On each following day 1–6 the advertiser runs stage-IC experiments by perturbing bids at the query level.

The cost of an experiment is quantified by its *regret*, defined as the difference between the utility when bidding truthfully and the utility under perturbed bids, as a percentage of the truthful utility. At the query level, regret may be undefined if the advertiser loses the auction and its truthful utility is 0. Therefore, we evaluate regret either at the bucket level (for the UI test), or over the entire realized impression traffic (for the stage-IC test). To ensure precise averages within buckets, the advertiser discards buckets with less than 100 impressions in our evaluation. Note that the foregoing experimental setup is not tailored to any specific kinds of exchange policies (e.g., dynamic reserves, throttling). In the next sections we will see how the setup effectively applies to two very different dynamic mechanisms. For the stage-IC test we report on results with noise levels $\alpha = 1\%, 5\%, 9\%$; for the UI test we report on levels $\beta = 10\%, 20\%, 30\%$.

5.1 Dynamic Reserve Prices

We first consider a mechanism where the exchange computes Myerson-optimal reserve prices within each bucket (publisher id, device type) based on the advertiser's bids on day 0, and then applies these reserve prices going forward on days 1 to 6. On day 0 the exchange runs a second-price auction for each query. In the following days the exchange runs a mixture between a second- and first-price auction. Specifically, let $\kappa \in [0, 1]$ denote a *first-price weight*; the auction allocates the impression to the highest bidder and charges price $\kappa b_1 + (1 - \kappa)b_2$, where b_1, b_2 are the first and second highest bids, respectively. On day 1 the exchange sets $\kappa = 0$ and runs a second price auction, and κ is incremented by 0.2 in each following day until we reach a first-price auction on day 6 with $\kappa = 1$. The purpose of this scenario is to evaluate how the tests respond when dynamic reserves interact with different kinds of stage mechanisms, both truthful and non-truthful.

Stage Incentive Compatibility. The results of the stage-IC test are reported in Figure 1, which plots the three components of inequality (3). In the plot the average allocation value (the middle component of the inequality) has been normalized to 1.0 to respect data confidentiality. We plot 95% confidence intervals throughout, here estimated according to the standard error across queries. Given the volume of data, the allocation value is always very precisely estimated. The noise level impacts the precision of the estimates for other inequality components, which are always unbiased estimates.

We see that for $\kappa \geq 0.4$ all noise levels are able to detect a violation of the upper bound in (3). The only failure to reject truthfulness is for $\kappa = 0.2$ at noise level 1%. At $\kappa = 0$ the test fails to reject the null hypothesis of a truthful auction at all noise levels, as should be the case with a second-price auction. At noise level 9% we can also reject the null hypothesis of an untruthful auction, since the test confirms the inequalities in (3)—confidence intervals do not overlap and the components of the inequality are separated. Note that the confidence intervals narrow as κ increases because the left- and right-hand components of the bounds in (3) tend to 0 as we mix in more of the first-price auction, reducing the variance. In a
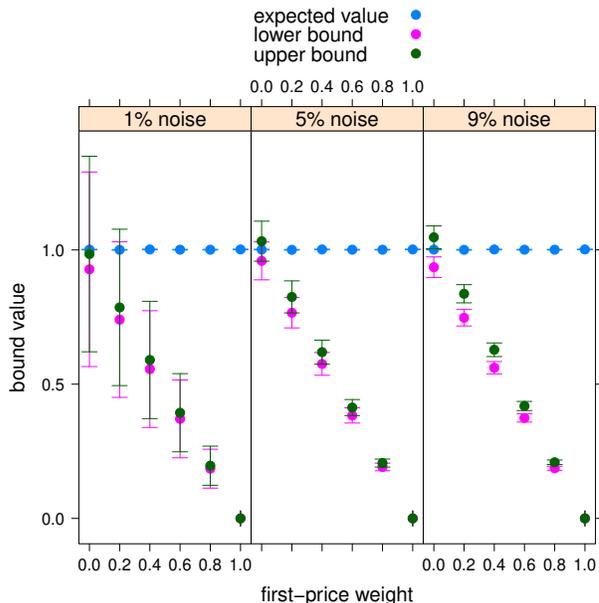

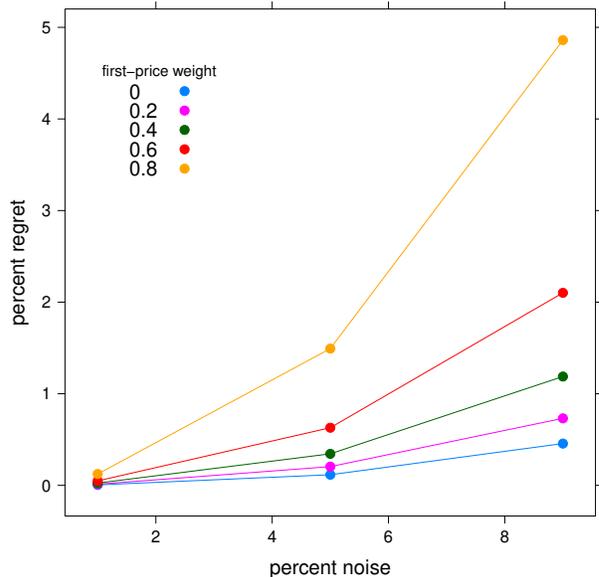Figure 1: Empirical bounds for the stage-IC test.

Figure 2: Regrets for the stage-IC test.

first-price auction, the realized utility with respect to the submitted bid is always 0.

The cost of the stage-IC test in terms of regret is shown in Figure 2. (Regret under the first-price auction with $\kappa = 1.0$ is undefined because truthful utility is always 0, and therefore ommitted.) Regret is computed according to total utility under truthful and perturbed bidding across queries in the day, so there are no associated confidence intervals. We see that the regret cost of running the stage-IC
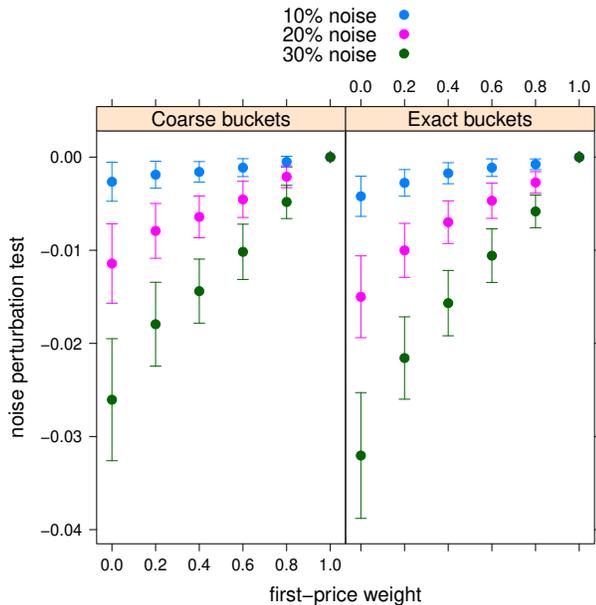
Figure 3: UI test results under different stage mechanisms.

test is very minimal for the 1% noise level, even though that level could detect IC violations of $\kappa \geq 0.4$. We should stress however that this favorable trade-off is due to the fact that we are considering a large advertiser with a wealth of impression traffic. In absolute terms low regret may still be costly. Note that while regret is necessarily non-negative under a second price auction at $\kappa = 0$, regret could in principle be negative for $\kappa > 0$, because the downward bid perturbations applied as part of the test can increase utility for those auctions (i.e., shading bids can be beneficial). On the other hand, upward bid perturbations can be even more detrimental to utility. Since regret is always positive and increasing with κ, it turns out that the cost from upward perturbations outweighs the benefit from downward perturbations over the entire traffic.[3]

Utility Independence. For the utility independence test, recall that the advertiser segments traffic by buckets (either exact or coarse) on day 0 to apply systematic bid perturbations, and then correlates the noise applied with realized utility in each bucket in future days, according to (5). Under our counterfactual scenario the exchange applies Myerson-reserves computed in each *exact* bucket on day 0 to auctions on days 1–6. However, the results of the utility independence test may not be the same for each day because the stage mechanism is evolving from a second price auction on day 1 to a first price auction on day 6.

Figure 3 plots the results of the UI test. The performance of the test is not substantially different between the coarse and exact bucket when bucketing by publisher id and device type, only that

---

[3]Intuitively, although the buyer may be better off if she wins the item at a lower price after downward perturbation, the buyer suffers from downward perturbations if she loses the item when she would have won under truthful reporting. On the other hand, upward perturbations either have no effect on utility, if the buyer does not win the item even after the perturbation, or have negative effect on the utility, if the buyer wins the item with a higher price after the perturbation.

the statistic comes out lower under exact buckets. We see that the test statistic is always significantly different from 0 at all κ at noise levels of 20% and 30%, and with 10% noise the test is effective up to $\kappa = 0.6$. (Confidence intervals are based on standard errors across buckets.) The exception at all noise levels is $\kappa = 1$, where the statistic is essentially 0. In this case, the statistic fails to detect the dynamic reserves because they are immaterial: in a first price auction, reserve prices do not impact utility under truthful bidding. This is also why the variance decreases with κ, as reflected in the narrowing of the confidence intervals.

The cost of the UI test in terms of regret is given in Table 1, where we report the average regret per bucket with standard error in parentheses. (Recall that we only retain buckets of size at least 100, and buckets with utility 0 are discarded.) In this scenario the cost of the test is only incurred on day 0, but for practical applications one would repeat it frequently.

Noise	10%	20%	30%	40%
Coarse	0.89 (0.05)	3.11 (0.08)	7.83 (1.00)	12.70 (1.01)
Exact	0.98 (0.06)	3.51 (0.25)	7.36 (0.28)	12.68 (0.39)

Table 1: Percentage Regrets for the UI test.

5.2 Stateful Mechanism

The previous experiments were based on a dynamic reserve pricing mechanism combined with mixtures of first- and second-price auctions. To illustrate the versatility of our tests, we now simulate a scenario where the exchange runs a mechanism inspired by the recent development of bank account mechanisms, a class of dynamic mechanisms that is efficient, revenue-optimal, and uses historical bids in an incentive-compatible way [16, 17, 19]. We are not aware of this kind of mechanism being used in practice, but it offers a clean way to illustrate how the UI test can uncover negative or positive correlation with future utility, in a mechanism entirely different from dynamic reserves.

Let K be the final day of the mechanism. On days prior to K the exchange simply runs second-price auctions, and records the advertiser's average utility $\bar{u}$ over those days. There are three versions of the mechanism, which we call the *Zero*, *Negative*, and *Positive* mechanisms according to the sign we expect them to produce for the UI test statistic. Let u^* be the exchange's forecast for the advertiser's total utility on the final day K when running second price auctions with no reserve price. Let r be the Myerson-optimal reserve price for the advertiser on the final day, based on a value distribution estimated by the exchange. On day K the mechanisms proceed as follows.

- **Zero.** If $\bar{u} \geq u^*$, the advertiser is included in every auction with 0 reserve and charged a daily participation fee of u^*. If $\bar{u} < u^*$, the advertiser is excluded from every auction.
- **Negative.** If $\bar{u} \geq u^*$, the advertiser is included in every auction with 0 reserve and charged a daily participation fee of u^*. If $\bar{u} < u^*$, the advertiser is included in every auction with r reserve (no participation fee).
- **Positive.** If $\bar{u} \geq u^*$, the advertiser is included in every auction and charged a daily participation fee of $u^*/2$. If $\bar{u} < u^*$, the advertiser is excluded from every auction.
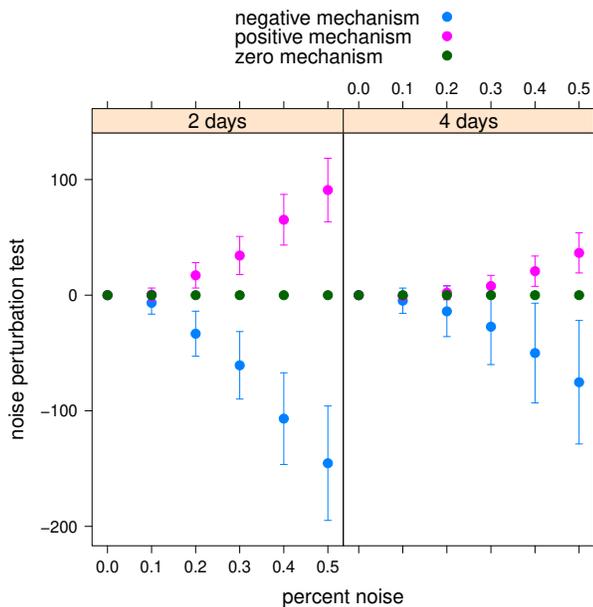
Figure 4: UI test on the stateful mechanisms.

Assuming that total daily utilities are relatively stable, we have $\Pr(\bar{u} \geq u^*) \approx 1/2$. Thus the Zero mechanism extracts $1/2$ of the advertiser's surplus on the final day, in expectation. This may be more or less than the surplus extracted from setting a Myerson-optimal reserve, which makes it an interesting mechanism to consider. The Zero mechanism is utility-independent (with expected utility 0) and stage-IC, so it is dynamic-IC by Theorem 3.2. The Negative mechanism incentivizes the advertiser to lower its bids on previous days to reach the case $\bar{u} < u^*$ where it achieves positive surplus, rather than the case $\bar{u} \geq u^*$ where it achieves zero surplus (in expectation). Although the Positive mechanism grants the advertiser more surplus than the Zero mechanism, it is not utility-independent and therefore not dynamic-IC, because the advertiser will want to raise its bids on previous days to pass threshold u^* on the final day.

The advertiser performs perturbations within buckets for the UI test on day 0 and then report truthfully for the next K days. Because the mechanisms sum up utilities on days prior to K, the larger K is, the more the initial perturbations are diluted, leading to weaker test results. We report on results where the mechanisms sums up utilities over 2 and 4 days before the final day in Figure 4, for coarse buckets (results were similar with exact buckets, with larger test magnitudes). As expected, the test statistic is closely concentrated around 0 for the Zero mechanism. The sign of the test is as expected for the Negative and Positive mechanisms; with 2 days of lag, 20% noise detects violations of UI, and 40% noise detects 4 days of lag. The cost of running the test, in terms of regret, is the same as reported in the previous section, because the test remains the same no matter the underlying mechanism.

6 CONCLUSIONS

This paper examined the question of how an advertiser might test that an auction mechanism is dynamic-IC via bid perturbation experiments. We showed that dynamic mechanisms that are dynamic-IC no matter how advertisers time-discount future surplus are characterized by history-independent mechanisms. We then developed novel statistical tests for the two defining properties of these mechanisms: stage-IC and utility independence. We ran experiments on real ad auction data from the Google Ad Exchange to confirm the practical viability of the tests in uncovering violations of dynamic-IC. Our tests offer new techniques for advertisers to gain insights into the underlying dynamic auction mechanism, which can inform their bidding and add a measure of transparency to the interaction between buyers and sellers in display advertising. An interesting extension of this line of work would be to develop IC tests for multi-item mechanisms, which could for instance distinguish between a generalized second-price auction and a Vickrey auction in search advertising.

REFERENCES

[1] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. 2013. Learning prices for repeated auctions with strategic buyers. In *Advances in Neural Information Processing Systems (NIPS)*. 1169–1177.

[2] Itai Ashlagi, Constantinos Daskalakis, and Nima Haghpanah. 2016. Sequential mechanisms with ex-post participation guarantees. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC)*. ACM, 213–214.

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

[4] Santiago R. Balseiro and Yonatan Gur. 2017. Learning in Repeated Auctions with Budgets: Regret Minimization and Equilibrium. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC)*. ACM, 609–609.

[5] Dirk Bergemann and Maher Said. 2010. *Dynamic Auctions*.

[6] Dirk Bergemann and Juuso Välimäki. 2010. The dynamic pivot mechanism. *Econometrica* 78, 2 (2010), 771–789.

[7] Yuyu Chen. 2017. Programmatic advertising is preparing for the first-price auction era. https://digiday.com/marketing/programmatic- advertising-readying-first-price-auction-era. Accessed: 2018-11-04.

[8] Benjamin Edelman and Michael Ostrovsky. 2007. Strategic bidder behavior in sponsored search auctions. *Decision Support Systems* 43, 1 (2007), 192–198.

[9] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. 2007. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review* 97, 1 (2007), 242–259.

[10] K. Ganchev, A. Kulesza, J. Tan, R. Gabbard, Q. Liu, and M. Kearns. 2007. Empirical Price Modeling for Sponsored Search. In *Internet and Network Economics (WINE)*, Vol. 4858. Springer.

[11] Sham M. Kakade, Ilan Lobel, and Hamid Nazerzadeh. 2013. Optimal dynamic mechanism design and the virtual-pivot mechanism. *Operations Research* 61, 4 (2013), 837–854.

[12] Yash Kanoria and Hamid Nazerzadeh. 2017. Dynamic reserve prices for repeated auctions: Learning from bids. (2017).

[13] IAB Tech Lab. 2016. OpenRTB API Specification Version 2.5. www.iab.com/ wp-content/uploads/2016/03/OpenRTB-API-Specification-Version-2-5-FINAL.pdf. Accessed: 2018-11-04.

[14] Sébastien Lahaie, Andrés Munoz Medina, Balasubramanian Sivan, and Sergei Vassilvitskii. 2018. Testing Incentive Compatibility in Display Ad Auctions. In *Proceedings of the 2018 World Wide Web Conference (WWW)*. 1419–1428.

[15] Andres M. Medina and Mehryar Mohri. 2014. Learning theory and algorithms for revenue optimization in second price auctions with reserve. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*. 262–270.

[16] Vahab Mirrokni, Renato Paes Leme, Pingzhong Tang, and Song Zuo. 2016. Dynamic Auctions with Bank Accounts.. In *IJCAI*. 387–393.

[17] Vahab Mirrokni, Renato Paes Leme, Pingzhong Tang, and Song Zuo. 2016. Optimal dynamic mechanisms with ex-post IR via bank accounts. *arXiv preprint arXiv:1605.08840* (2016).

[18] Vahab Mirrokni, Renato Paes Leme, Rita Ren, and Song Zuo. 2018. Dynamic Mechanism Design in the Field. In *Proceedings of the 2018 World Wide Web Conference (WWW)*. 1359–1368.

[19] Vahab Mirrokni, Renato Paes Leme, Pingzhong Tang, and Song Zuo. 2018. Non-clairvoyant dynamic mechanism design. In *Proceedings of the 2018 ACM Conference on Economics and Computation (EC)*. ACM, 169–169.

[20] Roger B Myerson. 1981. Optimal auction design. *Mathematics of Operations Research* 6, 1 (1981), 58–73.

[21] Christos Papadimitriou, George Pierrakos, Alexandros Psomas, and Aviad Rubinstein. 2016. On the complexity of dynamic mechanism design. In *Proceedings of the 27th annual Symposium on Discrete Algorithms (SODA)*. 1458–1475.