

# Testing Incentive Compatibility in Display Ad Auctions

Sebastien Lahaie  
Google Research  
New York, NY, USA  
slahaie@google.com

Andrés Muñoz Medina  
Google Research  
New York, NY, USA  
ammedina@google.com

Balasubramanian Sivan  
Google Research  
New York, NY, USA  
balusivan@google.com

Sergei Vassilvitskii  
Google Research  
New York, NY, USA  
sergeiv@google.com

## ABSTRACT

Consider a buyer participating in a repeated auction, such as those prevalent in display advertising. How would she test whether the auction is incentive compatible? To bid effectively, she is interested in whether the auction is *single-shot* incentive compatible—a pure second-price auction, with fixed reserve price—and also *dynamically* incentive compatible—her bids are not used to set future reserve prices. In this work we develop tests based on simple bid perturbations that a buyer can use to answer these questions, with a focus on dynamic incentive compatibility.

There are many potential A/B testing setups that one could use, but we find that many natural experimental designs are, in fact, flawed. For example, we show that additive perturbations can lead to paradoxical results, where higher bids lead to lower optimal reserve prices. We precisely characterize this phenomenon and show that reserve prices are only guaranteed to be monotone for distributions satisfying the Monotone Hazard Rate (MHR) property. The experimenter must also decide how to split traffic to apply systematic perturbations. It is tempting to have this split be randomized, however, we demonstrate empirically that unless the perturbations are aligned with the partitions used by the seller to compute reserve prices, the results are guaranteed to be inconclusive.

We validate our results with experiments on real display auction data and show that a buyer can quantify both single-shot and dynamic incentive compatibility even under realistic conditions where only the cost of the impression is observed (as opposed to the exact reserve price). We analyze the cost of running such experiments, exposing trade-offs between test accuracy, cost, and underlying market dynamics.

## ACM Reference Format:

Sebastien Lahaie, Andrés Muñoz Medina, Balasubramanian Sivan, and Sergei Vassilvitskii. 2017. Testing Incentive Compatibility in Display Ad Auctions. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*Conference'17, July 2017, Washington, DC, USA*

© 2017 Association for Computing Machinery.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM. . . \$15.00

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 INTRODUCTION

The display advertising industry could be on the verge of a major shift in its trading practices as many large exchanges, including AppNexus and Index Exchange, are currently considering a switch from second- to first-price auctions [19]. The push towards first-pricing has been driven in large part by ad buyers who no longer trust that exchanges are properly honoring second-pricing rules. There is concern that exchanges could be using other mechanisms to extract higher margins, such as setting reserve prices after bids are collected rather than before. As a result, the main motivation behind second-pricing, namely truthfulness, is lost [4]. In principle, market participants could audit an exchange's practices by comparing data between the ad buyers and publishers. In practice, this is only possible in one-off cases, and much too time-consuming to represent a systematic, scalable solution. Instead, ad buyers need to gain insights into the underlying auction dynamics based on their own trading data, in a way that can inform their bidding strategies.

Motivated by these concerns, this paper develops statistical tests that a buyer can use to estimate the degree to which a repeated auction is incentive compatible. We first address the question of *single-shot* incentive compatibility: detecting whether an auction is purely second-price, or whether the payment depends to any extent on the winning bid. We propose a simple experimental design where the advertiser randomly shades her value for the impression by some discrete fraction (10%, 20%, etc.), and then verifies that the slice of traffic that saw no shading yields the highest surplus. The core of our paper is devoted to the more subtle question of *dynamic* reserve pricing: detecting whether and to what extent a buyer's historical bids are used to set her future reserve prices<sup>1</sup>.

Expanding on the question of dynamic reserve pricing, our main contributions in this paper are two-fold:

- (1) To detect and assess the extent of dynamic reserve pricing, we introduce an experimental design where the advertiser partitions query traffic according to information provided in the bid callout, and systematically perturbs bids in a randomly chosen direction (up or down) within each partition.

<sup>1</sup>Note that detecting dynamic incentive compatibility (defined in Section 2) is more general than detecting dynamic reserve pricing: irrespective of how the buyer utilities in different rounds are tied together, the buyer in a dynamic incentive compatible mechanism should be incentivized to report truth in every round. Dynamic reserve pricing is the chief means to link incentives in different rounds, and violate dynamic incentive compatibility. In this paper we focus on this important special case of detecting dynamic reserve pricing.

In the next time period, the advertiser observes the correlation between the noise applied and the change in the reserve price. Positive correlation indicates that reserves depend on past bids. More formally, we provide a statistic that allows one to test the null hypothesis that the auction does not use past bids to set reserves. Note that dynamic reserve pricing is a relevant notion in both first- and second-price auctions.

- (2) In the process of analyzing various experimental designs, we give a new and intuitive characterization of the monotone hazard rate (MHR) family of distributions that is widely used in the economics and algorithmic game theory literature.

We investigate two key aspects of an experiment that impact its effectiveness. First, the nature of the perturbations: additive or multiplicative. We observe that a positive additive shift in the bid distribution does not necessarily lead to a higher optimal (i.e., revenue-maximizing) reserve price, and prove that this simple property holds if and only if the bid distribution has a monotone hazard rate (MHR). This is a new and intuitive characterization of MHR distributions that is of independent interest. Since many relevant distributions used in empirical modeling (e.g., the lognormal) are non-MHR distributions, additive perturbations in these cases will cause the buyer to observe confusing results—lower reserve prices in the presence of higher bids. We therefore use multiplicative perturbations, and explain how this leads to a test that can detect any scale-invariant dynamic reserve pricing scheme based on past bids, which includes optimal reserves.

The next key aspect is traffic partitioning. We show that unless the traffic split is aligned with the partitions used by the exchange/seller to compute reserves, the results are *guaranteed* to be inconclusive. This highlights a practical trade-off between using fewer coarse partitions to align with the seller, and more fine partitions to increase statistical power.

We validate our approach by running counterfactual simulations of Google's DoubleClick AdExchange on two days of data, perturbing the bids of a major advertiser on the first day according to our test. We implement a reserve-pricing scheme that partitions the inventory and computes Myerson-optimal reserve prices for each partition based on bids of this advertiser, and we simulate how this reserve-pricing logic may be detected by the advertiser. We also find that, using our test, the advertiser can detect if reserve prices are computed based on past bids, even if the advertiser uses a much coarser traffic partition than that used by our reserve pricing scheme. Furthermore, although at a significant cost for the buyer, detection is still possible assuming the advertiser can only observe the final cost of an impression, rather than the exact reserve price (which may be censored by the second-highest bid).

The paper is organized as follows. In Section 2 we describe the kinds of reserve pricing schemes that our statistical test is suited for: reserves set according to a buyer's history of bids on fixed partitions of query traffic, such that the reserves scale with the bids. We also provide formal definitions for the notions of single-shot and dynamic incentive compatibility. Section 3 provides our experimental design based on bid perturbations, and proves that additive perturbations can be problematic unless stringent conditions on bid distributions are met. In Section 4 we validate our test on real auction data from Google's DoubleClick AdExchange, and

demonstrate the relationship between the quality of the reserve price observations (e.g., exact or censored) and the ability to detect the underlying reserve pricing scheme. We also measure the cost-effectiveness of running the test for various bid perturbations magnitudes. Section 5 concludes.

## Related Work

*Single-shot reserve pricing.* The idea that reserve prices significantly increase the revenue of the seller is central in auction theory. In fact Myerson's seminal result [16] says that the single shot revenue-optimal auction with  $n$  bidders and values independently drawn from a (regular) distribution  $F$  just runs the welfare optimal second price auction with a reserve price for each bidder. The power of reserve prices is well understood and a lot of research has gone into (approximately) optimizing them under various constraints. Hartline and Roughgarden [10] construct second price AND VCG auctions with reserve prices in a variety of settings and show that they approximate the optimal revenue. Chawla et al. [5, 6] and Yan [20] take this further by showing that just reserve prices (namely, posted price mechanisms) without the presence of any auction competition is enough to approximate the optimal revenue in many single-parameter and some multi-parameter settings as well.

*Sample complexity.* Dhangwatnotai, Roughgarden and Yan [8] highlight the necessity to compute these reserve prices with minimal knowledge of the buyer's value distribution and show how to accomplish this and still get a constant fraction of the optimal mechanism's revenue. This work on sample robust reserve pricing was pushed further by Cole and Roughgarden [7], Huang, Mansour and Roughgarden [11], and Morgenstern and Roughgarden [15] who give bounds on the number of samples one needs to get  $(1 - \epsilon)$ -close to the optimal revenue. When it comes to posting personalized reserve prices (i.e., non-anonymous reserve prices), whether the auctioneer runs a lazy or eager auction makes a significant difference. Paes Leme, Pal and Vassilvitskii [17] study this difference in detail, and show that while optimal lazy auctions may be computationally easy, as opposed to the NP-hard problem of optimal eager auctions, empirically even suboptimal eager reserve prices yield better revenue than optimal lazy reserve prices. The problem of approximating personalized eager reserve prices was studied in more detail by Roughgarden and Wang [18] who show how to get a  $1/2$ -approximation to the optimal revenue, and also how to translate their offline algorithm into an online learning algorithm.

*Dynamic mechanisms.* Another line of work that is related to ours is the literature on sequential, dynamic, and repeated mechanisms. When a buyer repeatedly interacts with the same seller, the seller can gain significantly higher per round revenue than in a single-shot auction by linking the incentive constraints across the different rounds, i.e., making the future allocation and pricing decisions be strongly dependent on a buyer's past bids. A string of recent work (see for eg. [2, 3, 13]) with dynamic incentive compatibility as their central theme, has studied the additional revenue the seller can get under a variety of constraints.

*Strategic buyers.* Finally, our work is related to research in repeated auctions where the buyer knows the full mechanism used

by the seller, and the goal is to understand her optimal strategy in the presence of bid-dependent reserve prices. Kanoria and Nazarzadeh [12] propose a dynamic mechanism that is almost incentive compatible. That is, even though a strategic buyer could achieve higher surplus from playing an optimal bidding strategy, the incremental gain is negligible compared to the cost of playing the optimal strategy. Similarly Amin, Rostamizadeh and Syed [1], Mohri and Muñoz Medina [14] and Drusta [9] study a repeated posted price auction set up where the buyer attempts to maximize her discounted surplus. The main result of this line work is that the buyer's strategy can only deviate from truthfulness for a small number of rounds if the discounting factor is less than 1. However, [1] show that without any form of discounting, a buyer can force the seller to reduce the reserve price to a constant fraction of her true valuation. Devanur, Peres and Sivan [?] study a different setting where the buyer's values are not drawn freshly from a distribution every round. Instead they are initially drawn from a publicly known distribution and continue to remain the same afterwards. The seller can simply post a take-it-or-leave-it reserve price in every round, based on the previous round prices and buyer decisions. In the Perfect Bayesian Equilibrium (PBE) of this game where the seller is unable to make any commitments about how he will increase or decrease reserve prices in the future, the authors show that the seller gains non-trivial revenue only in the last few rounds of the game.

The results of these papers however critically rely on the unrealistic assumption that the buyer has access to the full mechanism. In practice, the buyer may only know the type of auction run by the seller, but not all the details of their pricing function; and even in the case of full information, verifying that the seller is following the mechanism is far from trivial.

## 2 PRELIMINARIES

Consider a buyer participating in a repeated auction with a single seller (such as an ad exchange). A *query* is an opportunity for the seller to show an ad to a user visiting a webpage. The seller runs an auction for each query by soliciting bids from different buyers, and potentially specifying different reserve prices for each of them. We consider a setup where the seller partitions the space of queries into *buckets* to compute reserve prices, where each bucket is defined by a specific combination of the query's and buyer's features. The buyer's goal is to figure out whether the exchange is running a truthful (i.e., second-price) auction, and furthermore, whether her current bids are being used to influence future reserve prices.

In each bucket, the buyer's value is drawn independently in every round from an absolutely continuous distribution  $F$  with density  $f$ . The value distributions in different buckets could be different. Since we study this problem from the perspective of a single buyer, the number of buyers in every auction is not relevant (there can be any number of buyers), and the different buyers can have potentially non-identical and even correlated distributions.

*History based reserve pricing.* The seller divides the auction rounds into several roughly equal time periods called stages. A stage, for instance, could be based on the day, or the week or the month to which these auction rounds belong. Given history of bids  $h_s = (b_1, \dots, b_\ell)$  in stage  $s$  in any given bucket, the seller computes the reserve price

$r_{s+1}$  in stage  $s + 1$  for that bucket as  $r_{s+1}(h_s, \cdot)$  where the “ $\cdot$ ” stands for inputs independent of the buyer's bids, such as the bucket to which the query belongs, other bidders' bids, market factors, that the seller may use in setting the reserve price.

*Scale invariant reserve pricing.* We define  $h'_s = (1 + \epsilon)h_s$  if  $b'_i = (1 + \epsilon)b_i$  for all  $1 \leq i \leq \ell$ . We call  $h'_s$  a  $(1 + \epsilon)$ -scaled version of  $h_s$ . We assume in this paper that the seller's reserve pricing scheme is scale-invariant, i.e., fixing the other inputs “ $\cdot$ ” we have,

$$r_{s+1}((1 + \epsilon)h_s, \cdot) = (1 + \epsilon)r_{s+1}(h_s, \cdot).$$

Note that scale invariance is equivalent to saying that prices should be independent of the whether the bids are specified in dollars or cents. In particular, the natural scheme of computing the Myerson-optimal reserve price out of the distribution induced by the history (i.e., samples)  $(b_1, \dots, b_\ell)$  is scale-invariant.

*Sample robust reserve pricing.* In practice the seller sets reserve prices based on historical bids (namely, samples) in a stage, rather than on the true distribution  $F$ . Nevertheless, in the ad exchange context it is reasonable to assume that there are enough samples in each bucket that the difference between the sampled distribution and true distribution is negligible. In particular, while one can construct corner-case reserve price functions that can be significantly off in the reserves they output even with small noise in the samples, we assume that the seller follows a sample robust reserve pricing algorithm where the noise in reserve prices become negligible with a large enough number of samples. Once again, natural schemes like computing the Myerson-optimal reserve price out of the distribution induced by the samples is sample robust. In the rest of the paper, we therefore ignore this sampling error.

*Four worlds: history-{independent vs dependent}-seller, static vs dynamic market.* A seller is history-independent if  $r_{s+1}(h_s, \cdot) = r_{s+1}(h'_s, \cdot)$  for all histories  $h_s, h'_s$ . Since the reserve price could potentially be a function of market factors, other bidders' bids etc. we model fluctuations in these quantities as static vs dynamic market. We say that a market is static if  $r_{s+1}(h_s, y)$  is the same for all values of  $y$ . The fluctuations in “ $y$ ” that creates a dynamic market is modeled as noise in the computed reserve price, namely, if we fix the history  $h_s$ , the changes in other factors “ $y$ ” will manifest as  $r_{s+1}(h_s, y) = (1 + \eta)r_{s+1}(h_s, y')$  where  $\eta$  is a zero mean random variable drawn from some distribution  $D$ . One could model noise to be additive as well, and the nature of our results would not change with these modifications.

*Buyer utility and incentive compatibility.* We assume that the buyers are standard quasi-linear utility maximizers, i.e., their goal is to maximize the sum over all rounds (spanning several periods) of  $v_x - p$  where  $v$  is their value in any given round,  $x$  denotes whether or not they were allocated the good and  $p$  is the price they were charged.

When it comes to analyzing the pricing schemes practiced by exchanges, there are two main categories: whether the buyer stands to gain in any single auction by misreporting their true value (single-shot incentive compatibility) and whether they stand to gain in their overall utility by misreporting in this round (dynamic incentive compatibility).

*Single-shot incentive compatibility.* Any single auction is incentive compatible if reporting the true value is in the buyer's best interests. Let

$$U_t(v_t, v'_t, h_{s(t)-1}, \cdot) = v_t x_t(v'_t, h_{s(t)-1}, \cdot) - p_t(v'_t, h_{s(t)-1}, \cdot)$$

denote the utility in any given round  $t$  for the buyer by reporting  $v'_t$  when her true value is  $v_t$ . The “ $\cdot$ ” in the allocation and payment functions captures, other buyers' reported valuations and other potential features. Single-shot IC requires that

$$U_t(v_t, v_t, h_{s(t)-1}, \cdot) \geq U_t(v_t, v'_t, h_{s(t)-1}, \cdot) \quad \forall v_t, v'_t \in R. \quad (1)$$

*Dynamic incentive compatibility.* Let there be  $T$  auctions in total, spanning several stages. Dynamic incentive-compatibility requires that:

$$\begin{aligned} \sum_{t=\tau}^T U_t(v_t, v_t, h_{s(\tau)-1} + (v_\tau, \dots, v_{t-1}), \cdot) &\geq \\ \sum_{t=\tau}^T U_t(v_t, v'_t, h_{s(\tau)-1} + (v'_\tau, \dots, v'_{t-1}), \cdot) & \\ \forall \tau \in [1, T], \forall t \in [\tau, T], \forall v_t, v'_t \in R. & \end{aligned}$$

*Buyer's goal.* The buyer's goal will be to design a test for evaluating both single-shot and dynamic incentive compatibility. The significance of these tests is that it can guide the choice of a buyer's bidding strategy.

### 3 EXPERIMENTAL DESIGNS

In this section we describe tests for detecting violations of single-shot and dynamic incentive compatibility.

#### 3.1 Single-shot Incentive Compatibility

While it is tempting to test for single-shot IC using simple statistics such as bid to cost ratio, it is easy to see that such methods fail to distinguish between incentive and non-incentive compatible auctions. Indeed, a second price auction with reserve can have a bid to cost ratio close to 1 if the reserve price happens to be close to the buyer's bid. Instead, we propose a simple test motivated by (1). Let  $k > 0$  and  $0 < \alpha_1 < \dots < \alpha_k = 1$  be a sequence of multipliers. Abusing the notation we let  $U_t(v_t, \alpha_j) = U_t(v_t, \alpha_j v_t, h_{s(t)-1}, \cdot)$ . It follows from the definition of incentive compatibility that

$$\alpha_k \in \operatorname{argmax}_{\alpha \in \{\alpha_1, \dots, \alpha_k\}} U_t(v_t, \alpha_j). \quad (2)$$

Our test is thus defined as follows: we randomly partition the traffic into  $k$  buckets  $B_1, \dots, B_k$ . The buyer then scales bids in bucket  $B_j$  by  $\alpha_j$  and calculates the total surplus  $S_j$  for each bucket. From (2), a single-shot incentive compatible mechanism should verify that  $S_k$  maximizes the surplus across all partitions.

*Why not use same test for dynamic incentive compatibility?* The case of dynamic incentive compatibility however, is significantly more difficult. It is easy to see that one cannot use the same test as in the single shot setting—randomly splitting queries is not well defined, since bucket association needs to persist across days. Other bucketing schemes fail as well—we will show in Section 3.3 that the seller needs to partition her traffic along the *same* dimensions as the buyer is using to compute reserve prices.

In this paper, we consider an important subset of ways in which dynamic incentive compatibility can be violated, namely by yesterday's bids influencing today's reserve prices. We call such a seller a history dependent seller. While theoretically this is a subset, notice that in practice this is not a weaker property than testing dynamic incentive compatibility since history dependent reserves correspond to the most common violation of dynamic incentive compatibility. The test we design for this is based on simple bid perturbations and the intuitive notion that higher bids should yield higher reserves. We first show that additive perturbations to bids can violate even this basic intuition, and yield paradoxical results unless stringent conditions are met by the bid distributions. Next, we show that when the buyer's bucketing scheme is not aligned with the seller's buckets, then no matter the nature of perturbations, the buyer is guaranteed to get inconclusive results. We conclude by showing that multiplicative perturbations on seller aligned buckets are appropriate, and that they produce meaningful shocks to reserve prices under general conditions.

#### 3.2 Additive Perturbations

The setup for the additive version of our test is simple: fix a bucket, and set the bids in stage  $s + 1$  in that bucket by adding an  $\epsilon$  to the bids in that bucket in stage  $s$ . The basic expectation from this perturbation would be that the reserve price in a bucket with such an additive perturbation would be no smaller than the reserve for unmodified bucket. Even this very basic requirement is true only for MHR distributions, i.e., for non-MHR distributions the buyer could observe a smaller reserve price after raising her bids!

Let  $F$  be the cdf of the absolutely continuous distribution of values in the bucket we are interested in, and let  $f$  be its density function. We denote by  $G := 1 - F$  the associated survival function for  $F$  and by  $\mathbf{x}: [0, 1] \rightarrow \mathbb{R}$  the quantile function given by  $\mathbf{x}(q) = G^{-1}(q)$ . Consider a buyer who perturbs her bids by adding  $\epsilon \in \mathbb{R}$  to the value. That is the seller observes a bid following the distribution  $F_\epsilon(x) = F(x - \epsilon)$ . It is not hard to show that the quantile function  $\mathbf{x}_\epsilon$  associated with  $F_\epsilon$  is given by

$$\mathbf{x}_\epsilon(q) = \mathbf{x}(q) + \epsilon. \quad (3)$$

We are interested in the set of distributions for which the function  $\epsilon \mapsto \operatorname{argmax}_r rG_\epsilon(r)$  is strictly monotone. The importance of this family of distributions is clear: this is the family for which a buyer can measure the extent to which bid perturbation affects future reserve prices when a seller learns from her bids. While it seems intuitive that the optimal reserve should be monotone in the size of the perturbation  $\epsilon$ , the next example shows that this is not true in general.

**EXAMPLE 1.** Let  $G(\mathbf{x}) = e^{-\mathbf{x}^{1/2}}$  be a Weibull distribution with shape parameter  $1/2$ . Let us calculate  $r_\epsilon = \operatorname{argmax}_r rG_\epsilon(r)$ . From the definition of  $G_\epsilon$  and taking derivatives of the revenue function we see that:

$$(rG_\epsilon(r))' = (re^{-(r-\epsilon)^{1/2}})' = \frac{e^{-(r-\epsilon)^{1/2}}(2\sqrt{r-\epsilon} - r)}{2\sqrt{r-\epsilon}}.$$

Equating the above expression to 0 and solving for  $r$  shows that  $r_\epsilon = 2 + 2\sqrt{1 - \epsilon}$ . Notice that this is a decreasing function of  $\epsilon$ . Therefore

a buyer that perturbs her bids by increasing them by  $\epsilon$  will in fact observe a lower reserve price!

As the previous example shows not even regular distributions satisfy strict monotonicity in reserves obtained from perturbed distributions. However we will show that functions with the monotone hazard rate property are precisely the ones for which this condition holds.

**DEFINITION 1.** A distribution is said to satisfy the monotone hazard rate (MHR) condition if

$$HR(x) = f(x)/G(x)$$

is a weakly increasing function in  $x$ .

**PROPOSITION 1.** If a distribution satisfies MHR then the following inequality is satisfied for every  $q \in [0, 1]$

$$\frac{qx''(q)}{x'(q)} + 1 \geq 0.$$

**PROOF.** From the definition of MHR we see that  $G(x)/f(x)$  must be weakly decreasing therefore by first order conditions:

$$\frac{-f(x)^2 - f'(x)G(x)}{f(x)^2} \leq 0. \quad (4)$$

This implies that

$$1 + \frac{f'(x)G(x)}{f(x)^2} \geq 0. \quad (5)$$

Let us now calculate the second derivative of  $x$ , w.r.t.  $q$ . By the inverse function theorem we have:

$$x' = -\frac{1}{f(x)} \quad x'' = \frac{f'(x)}{f(x)^2} x'.$$

Therefore  $\frac{qx''(q)}{x'(q)} = \frac{qf'(x)}{f(x)^2}$ . The change of variables  $q = G(x)$  in inequality (5) yields the result.  $\square$

**THEOREM 1.** A regular distribution  $F$  satisfies MHR if and only if  $r_\epsilon$  is an increasing function of  $\epsilon$ .

**PROOF.** Notation: In the proof, we will deal with derivatives w.r.t.  $q$  and  $\epsilon$ . We use the  $'$ -symbol for derivative w.r.t.  $q$  and  $\frac{d}{d\epsilon}$  for derivative w.r.t.  $\epsilon$ .

Let  $\epsilon > 0$  and  $q_\epsilon^*$  denote the optimal revenue quantile for the distribution  $F_\epsilon$ , i.e.,  $q_\epsilon^* = \operatorname{argmax}_{q \in [0, 1]} R_\epsilon(q) = \operatorname{argmax}_{q \in [0, 1]} qx_\epsilon(q)$ . We first prove that the function  $\epsilon \mapsto q_\epsilon^*$  is strictly increasing. By equation (3), it follows that the optimal reserve price  $r_\epsilon^* = x_\epsilon(q_\epsilon^*) = x(q_\epsilon^*) + \epsilon$ . From equation (3) it also follows that  $R_\epsilon(q) = qx_\epsilon(q) = R(q) + \epsilon q$ ; differentiating this equation w.r.t.  $q$ , we get  $R'_\epsilon(q) = R'(q) + \epsilon$ , and in particular, since  $R'(q_0^*) = 0$  by definition, we have  $R'_\epsilon(q_0^*) = \epsilon > 0$ . Since  $F$  is a MHR distribution, it follows that  $R(q)$  is concave function (standard result), and thus it follows that  $R_\epsilon(q)$  is concave as well. We just showed that  $R'_\epsilon(q_0^*) = \epsilon > 0$ , i.e., the concave function  $R_\epsilon(\cdot)$  is increasing at  $q_0^*$ . This means that it is maximized at a point strictly larger than  $q_0^*$ , i.e.,  $q_\epsilon^* > q_0^*$ . We can similarly show that for  $\epsilon < 0$ ,  $q_\epsilon^* < q_0^*$ . Thus, the function  $\epsilon \mapsto q_\epsilon^*$  is strictly increasing.

Let us now analyze the derivative of  $r_\epsilon^*$  w.r.t.  $\epsilon$ . In the previous paragraph we showed that  $r_\epsilon^* = x(q_\epsilon^*) + \epsilon$ . Therefore,

$$\frac{d}{d\epsilon} r_\epsilon^* = x'(q_\epsilon^*) \cdot \frac{d}{d\epsilon} q_\epsilon^* + 1. \quad (6)$$

We are interested in those distributions for which  $\frac{d}{d\epsilon} r_\epsilon^* \geq 0$ , i.e., as the additive perturbation  $\epsilon$  increases, the optimal reserve  $r_\epsilon^*$  weakly increases.

Since  $q_\epsilon^* = \operatorname{argmax}_{q \in [0, 1]} R_\epsilon(q) = \operatorname{argmax}_{q \in [0, 1]} qx_\epsilon(q) = \operatorname{argmax}_{q \in [0, 1]} qx(q) + \epsilon q$ , the optimality condition of  $q_\epsilon^*$  (i.e., first order derivative of  $qx(q) + \epsilon q$  w.r.t.  $q$ ) gives us that:

$$R'_\epsilon(q_\epsilon^*) = q_\epsilon^* x'(q_\epsilon^*) + x(q_\epsilon^*) + \epsilon = 0.$$

Taking derivatives on both sides of the equation w.r.t.  $\epsilon$  yields:

$$(q_\epsilon^* x''(q_\epsilon^*) + x'(q_\epsilon^*)) \frac{d}{d\epsilon} q_\epsilon^* + x'(q_\epsilon^*) \frac{d}{d\epsilon} q_\epsilon^* + 1 = 0$$

Therefore,

$$\begin{aligned} x'(q_\epsilon^*) \frac{d}{d\epsilon} q_\epsilon^* + 1 &= -(q_\epsilon^* x''(q_\epsilon^*) + x'(q_\epsilon^*)) \frac{d}{d\epsilon} q_\epsilon^* \\ &= -x'(q_\epsilon^*) \frac{d}{d\epsilon} q_\epsilon^* \left( \frac{q_\epsilon^* x''(q_\epsilon^*)}{x'(q_\epsilon^*)} + 1 \right). \end{aligned}$$

The LHS is just  $\frac{d}{d\epsilon} r_\epsilon^*$  by equation (6), and since we want  $r_\epsilon^*$  to be weakly increasing in  $\epsilon$  at all  $\epsilon$ , we want the LHS to be non-negative for all  $\epsilon \geq 0$ . In the RHS, note that because we proved that the mapping  $\epsilon \mapsto q_\epsilon^*$  is strictly increasing, we have  $\frac{d}{d\epsilon} q_\epsilon^* > 0$ . Furthermore, since  $x = G^{-1}$  and  $G$  is decreasing by definition, it follows that  $x$  is also decreasing, making  $x' < 0$ . Thus the only way the RHS is non-negative is when  $\frac{q_\epsilon^* x''(q_\epsilon^*)}{x'(q_\epsilon^*)} + 1 \geq 0$  for all  $\epsilon$ . From Proposition 1 this expression is non-negative for all  $\epsilon$  if and only if  $F$  satisfies the MHR condition.  $\square$

### 3.3 Unaligned buckets for perturbations

One of the most natural A/B tests to detect dynamic reserve pricing is to (i) pick an arbitrary feature to split the space of queries into two; (ii) perturb to increase the bids in one of the halves, and decrease bids in the other half, and (iii) verify the change of the reserve prices in these two halves the next day. However, if the feature picked by the buyer is not one of the features picked by the seller to partition the space of queries into buckets, the results are *guaranteed* to be inconclusive.

**THEOREM 2.** The results of a buyer A/B test based on a feature unused by the seller are guaranteed to be inconclusive, regardless of whether the perturbations are additive or multiplicative.

**PROOF.** Let the buyer divide the space of queries into two buckets  $B_1$  and  $B_2$  based on a feature not used by the seller. Consider any seller-defined bucket  $B$ . Let the cdf of the unperturbed distribution in  $B$  be  $F$ , supported in  $[0, 1]$ . Let  $q_1$  be the fraction of queries in  $B$  that are from  $B_1$  and  $q_2 = 1 - q_1$  be the fraction of queries from  $B_2$ . Let  $\widetilde{F}_A$  and  $\widetilde{F}_M$  be the distribution in the same bucket  $B$  after additive and multiplicative distributions respectively. An additive perturbation increases the bids additively by  $\epsilon$  in  $B_1$  and decreases them by  $\epsilon$  in  $B_2$ . A multiplicative perturbation multiplies the bids by  $1 + \epsilon$  in  $B_1$  and by  $1 - \epsilon$  in  $B_2$ . Let  $F_1$  be the cdf of the distribution in the bucket  $B \cap B_1$ , and  $F_2$  in the bucket  $B \cap B_2$ . It is easy to verify that:

$$\widetilde{F}_A(x) = \begin{cases} q_2 F_2(x + \epsilon), & 0 \leq x \leq \epsilon \\ q_1 F_1(x - \epsilon) + q_2 F_2(x + \epsilon), & \epsilon \leq x \leq 1 - \epsilon \\ q_1 F_1(x - \epsilon) + q_2, & 1 - \epsilon \leq x \leq 1 + \epsilon \end{cases}$$

and similarly,

$$\widetilde{F}_M(x) = \begin{cases} q_1 F_1\left(\frac{x}{1+\epsilon}\right) + q_2 F_2\left(\frac{x}{1-\epsilon}\right), & 0 \leq x \leq 1 - \epsilon \\ q_1 F_1\left(\frac{x}{1+\epsilon}\right) + q_2, & 1 - \epsilon \leq x \leq 1 + \epsilon \end{cases}$$

Consider the simple case where  $F(x) = F_1(x) = F_2(x) = x$  (i.e., the  $U[0, 1]$  distribution) and  $q_1 = q_2 = \frac{1}{2}$ . The Myersonian optimal reserve price for  $F$ , without any perturbations is  $\max_x x(1-F(x)) = \frac{1}{2}$ . For  $x \in [\epsilon, 1 - \epsilon]$ , we have  $\widetilde{F}_A(x) = x$ , and thus for small  $\epsilon$  the optimal reserve price after perturbation (namely,  $\max_x x(1-\widetilde{F}_A(x))$ ) will be exactly  $\frac{1}{2}$ , i.e., unchanged by perturbation, *both* in bucket  $B_1$  and in  $B_2$ . This is because the seller sets a single reserve for the whole bucket  $B$ . A buyer observing this would be led to conclude that the seller doesn't do any dynamic reserve pricing. Likewise, for  $x \in [0, 1 - \epsilon]$ , we have  $\widetilde{F}_M(x) = \frac{x}{1-\epsilon^2}$ , and thus for small  $\epsilon$  the optimal reserve price after perturbation (namely,  $\max_x x(1-\widetilde{F}_M(x))$ ) will be  $\frac{1-\epsilon^2}{2}$ , i.e., decreased after perturbation, *both* in bucket  $B_1$  and in  $B_2$ . A buyer observing this would feel puzzled that in a bucket where she increased bids, the seller's algorithm managed to decrease reserve prices.

The above example illustrates that the results of the buyer's A/B test either lead her to draw the incorrect conclusion that the seller is not doing dynamic reserve pricing, or give her paradoxical results. This is not a property of any particular distribution. For any  $F$ ,  $F_1$  and  $F_2$ , the seller's new reserve price is a single number for all of bucket  $B$  (by definition of  $B$  being a seller's bucket). If the new reserve price happens to be identical to the old reserve price, buyer incorrectly concludes that there is no dynamic reserve pricing even if there is. If the new reserve price has increased the buyer feels puzzled about increased reserve price even for queries in  $B \cap B_2$  where she decreased bids. If the new reserve price has decreased, the buyer again feels puzzled about decreased reserve price even for queries in  $B \cap B_1$  where she increased bids.  $\square$

Theorem 2 above underscores the subtlety in constructing an A/B test to detect dynamic reserve pricing. Note that the feature picked by the buyer may indeed be quite a useful feature, i.e., it could indeed be that the distribution in  $B_1$  is significantly different from the distribution in  $B_2$  (or, the distribution  $F_1$  in  $B \cap B_1$  is different from the distribution  $F_2$  in  $B \cap B_2$ ). The failure of the above test is irrespective of whether the feature picked by the buyer is relevant or irrelevant. Whether the seller also considered it to be relevant is the question, and this is what makes the A/B test design complicated for the buyer.

### 3.4 Multiplicative Perturbations

We now introduce our main test based on multiplicative perturbations. Notice that any scale invariant choice of reserve prices will satisfy, by definition, that perturbations resulting in larger bids will lead to higher reserve prices. Let  $\{B_1, \dots, B_n\}$  denote a partition of traffic into  $n$  buckets. Let  $\alpha \in [0, 1]$  denote the bid perturbation magnitude. For every bucket  $B_i$  denote by  $\beta_i$  a uniform random variable in  $\{1 - \alpha, 1 + \alpha\}$ . Our proposed experiment proceeds as follows: in round  $s$ , for every query  $q_j$  falling in bucket  $B_i$  for which buyer has value  $v_{ij}$ , the buyer submits a bid of  $\beta_i v_{ij}$ . At round  $s + 1$  the buyer observes reserve price  $r_{s+1}^{(i)}$ . Our statistical test is motivated by the intuition that if a seller uses past bids to influence

reserve prices then the change in magnitude of reserve price from day  $s$  to  $s + 1$  should be highly correlated with the bid perturbation magnitude  $\beta_i - 1$ . More precisely, we define the quotient

$$Q_i := r_{s+1}^{(i)} / r_s^{(i)}$$

and we denote the average correlation of  $Q_i$  with the noise magnitude  $\beta_i - 1$  by

$$C_\alpha = \mathbb{E}[Q_i(\beta_i - 1)].$$

Using the conditions described in Section 2 we now establish the properties of  $C_\alpha$  for the scenarios considered in this paper.

- **History independent seller.** By definition, the reserve prices of a history independent seller are independent of the bids. Therefore the ratio  $Q_i$  is independent of  $\beta_i$  and  $C_\alpha = 0$ . This equality holds in both a static and dynamic market since the dynamic noise  $\eta$  is independent of  $\beta_i$ .
- **History dependent seller.** Under the assumption of robust and scale invariant pricing we must have  $Q_i = \beta_i$ . Therefore,  $E[Q_i \beta_i] = E[\beta_i(\beta_i - 1)] = \alpha^2$ . The same equality holds under dynamic market conditions since the noise  $\eta$  is independent of  $\beta_i$ .

We can now define a test to discriminate between a history dependent and a history independent seller. Let  $\widehat{C}_\alpha = \frac{1}{n} \sum_{i=1}^n Q_i(\beta_i - 1)$  denote the empirical estimate of  $C_\alpha$ . Then for any  $\delta > 0$ , by Chebyshev's inequality we have with probability at least  $1 - \delta$

$$|C_\alpha - \widehat{C}_\alpha| \leq \sqrt{\sum_{i=1}^n \frac{\text{Var}(Q_i(\beta_i - 1))}{n}}, \quad (7)$$

where we have used Chebyshev's inequality for ease of exposition. Note that one could use any other concentration or bootstrapping technique to obtain a confidence interval for  $C_\alpha$ . Equation (7) implies that we can reject the null hypothesis of a history independent seller with probability of error less than  $1 - \delta$  if the confidence interval for  $C_\alpha$  does not include 0. Similarly we can reject the null hypothesis that the seller is history dependent and using scale invariant reserve prices if said interval does not contain  $\alpha^2$ .

## 4 EXPERIMENTAL RESULTS

For our experimental evaluation we select a large advertiser on DoubleClick AdExchange and consider a random sample of auctions that the advertiser participated in over two consecutive days in October, 2017. We use an auction simulator that is able to replay the exchange auction logic from logged data and simulate auction outcomes under different reserve prices, or adjusted bids from some of the buyers. For measuring single-shot incentive compatibility and dynamic incentive compatibility we implement the following reserve price selection method: we partition traffic by observable features (in the bid callout), and compute Myerson-optimal reserve prices within partitions using bids from the first day; note that Myerson-optimal reserves are scale invariant, so our theory applies to this scheme.

### 4.1 Single-Shot Incentive Compatibility

To demonstrate the simplicity and power of our single-shot incentive compatibility test we carry it out under three different mechanisms. We define a  $\gamma$ -price auction as one that sets a reserve

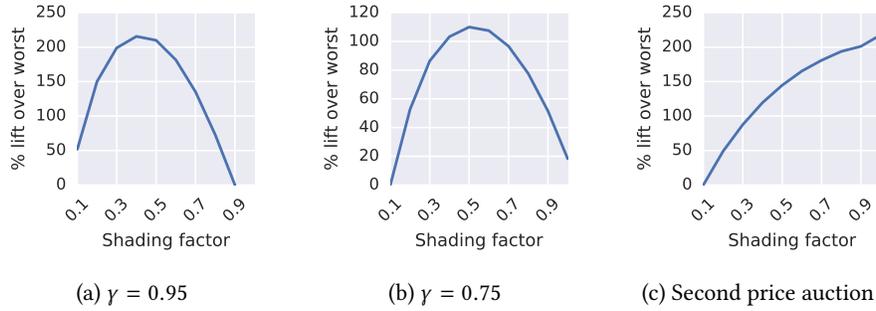


Figure 1: Surplus lift (over worst performing bidding strategy) vs bid shading factor  $\delta$ .

to a  $\gamma$  fraction of the highest bid. Note that  $\gamma = 1$  corresponds to first price auctions. Moreover, a  $\gamma$ -price auction is incentive compatible only if  $\gamma = 0$ ; since otherwise the payment of a buyer may depend on her bid.

The parameters of our test are  $k = 10$  and  $\alpha_i = \frac{i}{10}$ . We modified the exchange auction to simulate a  $\gamma$ -price auction for  $\gamma = 0.95$  and  $\gamma = 0.75$ . We also ran a second price auction using the reserve pricing scheme described above and we show the results in Figure 1. Notice that for both  $\gamma$ -price auctions, the optimal multiplier is bounded away from 1 whereas, for the second price auction, the optimal strategy is given by the multiplier  $\alpha_k$ , even in the presence of reserve prices.

## 4.2 Dynamic Incentive Compatibility

We now turn to our statistical test of dynamic incentive compatibility. We explore the effectiveness of our tests along two dimensions: the information available to the buyer about the exchange-set reserve, and the granularity of the buckets over which bids are perturbed in the test. We then demonstrate the trade-off between the efficacy of the tests and their cost.

For the informational dimension, we distinguish between three scenarios that differ on whether the buyer observes the exact reserve price set by the exchange, or some censored version of it. In the *full information* scenario, the buyer observes the exact reserve; in our implementation, this is the Myerson-optimal reserve price calculated using the bucketing method previously described. This setting is ideal for the buyer, but often unrealistic in practice as the exchange’s reserve can often be overridden by other parameters. For instance, publishers may set their own reserve prices for portions of their inventory that they consider especially valuable. In the *partial information* scenario, the buyer only observes final reserve price that represents the maximum of all these parameters. For our implementation we simulate reserve prices as the maximum between Myerson reserves and the floors submitted by the publishers. Finally, we consider a *limited information* scenario where an advertiser only observes her cost per auction (i.e., the maximum of the reserve price and second-highest bid). This scenario is arguably the most realistic in practice since the cost is always available to the buyer, but the exact reserve price may or may not be shared by the exchange.

The other dimension we explore is the granularity of the buckets over which bids are perturbed and their alignment with partitions

used to set reserve prices. We consider a handful of features associated with each auction, for instance whether the segment of inventory is mobile or desktop, the publisher identity, and so on. In our simulations, the seller uses these features to compute an optimal reserve price for each unique setting, for instance computing a different reserve price for mobile-NYTimes, desktop-NYTimes, mobile-WSJ, desktop-WSJ, and so on (we discard feature combinations that occur fewer than 100 times in the dataset).

For the buyer experiment, we define a *coarse*, *fine* and *random* settings. The *coarse* partition is defined only according to the publisher identity. The *fine* partition is a refinement, and represents the same partitioning scheme used by our reserve price selection algorithm. Finally, in the *random* setting, buckets are defined by the IP address of the machine making the ad request.

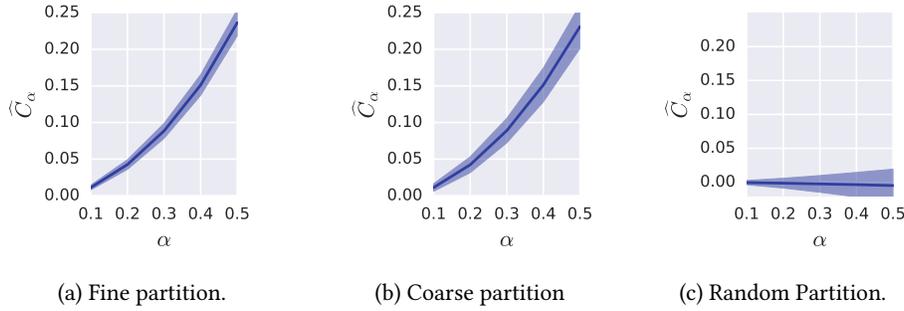
In each setting, we follow the perturbation model described in Section 3.4. For each bucket we select a perturbation magnitude uniformly in  $\{1 - \alpha, 1 + \alpha\}$ . Perturbations vary between buckets but are fixed for all bids within a bucket.

Finally, we track two metrics. The first captures the experimenter’s ability to determine whether the seller is computing Myerson optimal reserves based on the previous bids submitted by the buyer. The second, is the cost of these experiments, and their impact on the utility obtained by the buyer.

*Full Information.* In order to measure how the magnitude of the reserve price changes as a function of the noise we let  $\alpha$  vary in the set  $\{0.1, 0.2, 0.3, 0.4, 0.5\}$ . For every  $\alpha$  we calculate the empirical covariance  $\widehat{C}_\alpha$  and we calculate a 95% confidence interval according to (7). The results of this analysis can be seen in Figure 2.

Observe that when the buyer partitions her traffic randomly, the test fails, even in the full information setting; thus it is important for the buyer partitions to be aligned with those used by the seller. On the other hand, the difference between fine and coarse partitions is relatively small. The granularity only changes the confidence interval, which is a bit tighter for the finer partition.

*Partial and Limited Information.* Recall that in the partial information scenario the bidder does not see the actual reserve price computed by the seller, rather she can only observe a censored reserve price. In the limited information scenario the buyer can only observe her cost per impression. As expected, the inability to observe exact reserve prices computed makes the test less effective.



**Figure 2: Empirical covariance between reserve prices and bid perturbation as a function of  $\alpha$ , in the full information scenario.**

As in the full information setting, we let  $\alpha$  vary in  $\{0.1, 0.2, 0.3, 0.4, 0.5\}$  and compute the correlation between  $\alpha$  and the observed reserve price, as a function of different partitions. Figure 3 shows the results of our statistical test for each traffic partition in the partial information scenario, while Figure 4 shows the results of the limited information scenario. The first thing to notice is that the magnitude of the effect is muted, and the covariance is no longer simply the square of the noise. In fact for the limited information scenario, it is barely possible to detect correlation even for values of  $\alpha$  as high as 0.4. This drastic drop in correlation for the limited information scenario can be explained by the fact that changes in advertisers costs depend not only on the reserve price set by the seller but also on the bids from other buyers which can change from day to day independently of the bid perturbation. This highlights the difficulty of dynamic incentive compatibility testing, recall that single-shot incentive compatibility testing is easy even in the limited information setting (Figure 1).

Finally, notice that unlike the previous two scenarios, the coarse partition detects correlation slightly better than the fine partition. This is mainly due to the fact that each bucket in the fine partition has a smaller number of samples. This makes the estimate of the average cost per bucket noisier, and therefore less correlated with the bid perturbation.

*Experiment Cost.* The statistical power of our test can always be enhanced by increasing the magnitude of the perturbation. However, this comes at a direct cost to the buyer. As we observed above, bidding one’s true value is a surplus maximizing strategy in second price auctions with reserves. Therefore, we define the advertiser’s experimental cost as the loss in surplus compared to her bidding her value. For a given query  $k$ , we assume the bid  $b_k$  as logged in the data represents a buyer’s true value for the impression. Let  $p_k$  be the cost of the query as logged in the data, and let  $\tilde{p}_k$  be the query’s cost under the perturbed bid as calculated by the auction replay. Let  $x_k \in \{0, 1\}$  be a binary variable indicating whether the advertiser won impression  $k$  under its original bid, and define  $\tilde{x}_k$  analogously for the perturbed bid. Note that the cost of the query is necessarily zero if the advertiser loses the auction. It is possible for the advertiser to win the auction under its original bid and lose under its perturbed bid, and vice-versa. The advertiser’s *surplus* under the original and perturbed bids is  $x_k(b_k - p_k)$  and  $\tilde{x}_k(b_k - \tilde{p}_k)$  respectively, and its *regret* for the perturbation is

$x_k(b_k - p_k) - \tilde{x}_k(b_k - \tilde{p}_k)$ . Regret must be non-negative if truthful bidding is a dominant strategy, and we confirm that this holds for each query in our evaluation.

For each experimental condition we report the regret as a percentage of the optimal surplus over all queries, namely

$$\frac{\sum_k x_k(b_k - p_k) - \tilde{x}_k(b_k - \tilde{p}_k)}{\sum_k x_k(b_k - p_k)}.$$

We plot the results in Figure 5. Note that regret only depends on the partition granularity and noise magnitude, and does not depend on whether we are in the full or partial information setting. We computed confidence intervals by separately computing the regret for each hour and using (twice) the standard error across the 24 resulting regret samples. We see that regret can be very precisely estimated at all perturbation levels and for both the coarse and fine partitions, which show very similar trends. At  $\alpha = 10\%$ , regret is no greater than 1%, and from  $\alpha = 30\%$  onwards regret starts to exceed 5% for both partitions. To apply our test, advertisers will want to gradually increase noise magnitude to ensure that the opportunity cost does not exceed the value of the insights into incentive compatibility.

## 5 CONCLUSION

This paper proposed statistical tests to quantify the extent to which a display ad auction is incentive compatible. We considered a test for single-shot IC, to confirm that bid information is not used to set the payment in any individual auction run. Our main focus was on dynamic IC, which asks whether historical bids are used to set future reserve prices in some scale-invariant fashion. We gave theoretical and practical demonstrations that natural tests such as additive perturbations and random traffic splitting can fail at estimating dynamic IC. Moreover, we ran extensive experiments to estimate the tradeoffs between experiment cost, traffic splitting, and the type of pricing information available to the buyer. Our tests offer means for advertisers to gain insights into the underlying auction mechanism that can inform their bidding strategies, without the need for any coordination with publishers.

## REFERENCES

- [1] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. Learning prices for repeated auctions with strategic buyers. In *Proceedings of NIPS*, pages 1169–1177, 2013.

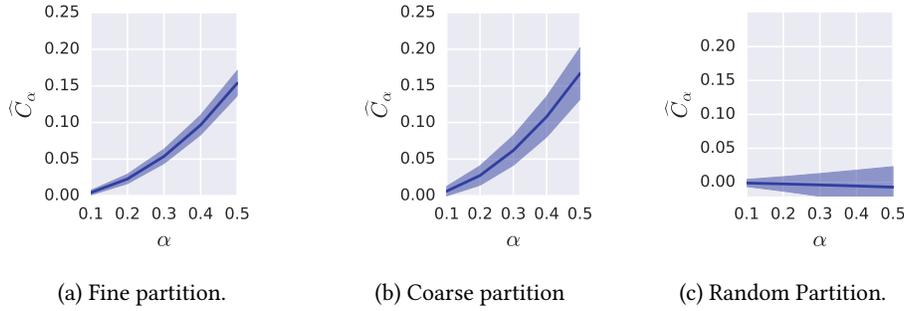


Figure 3: Empirical covariance between reserve prices and bid perturbation as a function of  $\alpha$ , in the partial information scenario.

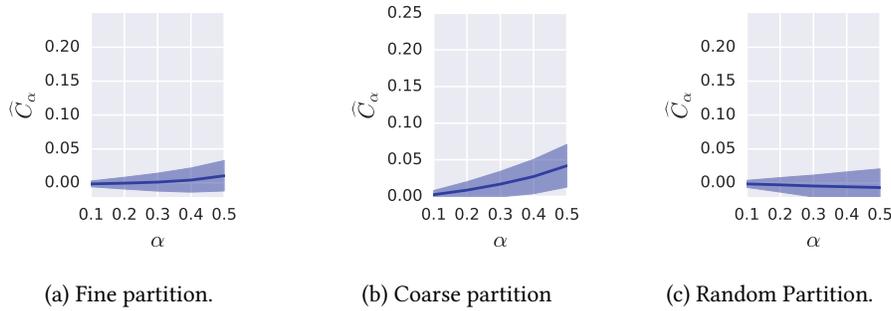


Figure 4: Empirical covariance between reserve prices and bid perturbation as a function of  $\alpha$ , in the limited information scenario.

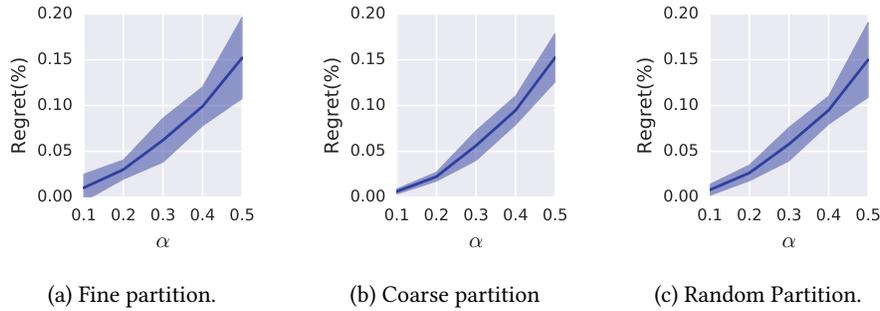


Figure 5: Regret as a function of  $\alpha$ .

[2] Itai Ashlagi, Constantinos Daskalakis, and Nima Haghpahan. Sequential mechanisms with ex-post participation guarantees. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pages 213–214, New York, NY, USA, 2016. ACM.

[3] Santiago Balseiro, Vahab Mirrokni, and Renato Paes Leme. Dynamic mechanisms with martingale utilities. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 165–165, New York, NY, USA, 2017. ACM.

[4] Ross Benes. How SSPs use deceptive price floors to squeeze ad buyers. <https://digiday.com/marketing/ssps-use-deceptive-price-floors-squeeze-ad-buyers>, September 30, 2017. Accessed: 2017-10-20.

[5] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the Forty-second ACM Symposium on Theory of Computing*, STOC '10, pages 311–320, New York, NY, USA, 2010. ACM.

[6] Shuchi Chawla, David L. Malec, and Balasubramanian Sivan. The power of randomness in bayesian optimal mechanism design. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, EC '10, pages 149–158, New York, NY, USA, 2010. ACM.

[7] Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 243–252, New York, NY, USA, 2014. ACM.

[8] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games and Economic Behavior*, 91(C):318–333, 2015.

[9] Alexey Drutsa. Horizon-independent optimal pricing in repeated auctions with truthful and strategic buyers. In *Proceedings of WWW*, pages 33–42, 2017.

[10] Jason D. Hartline and Tim Roughgarden. Simple versus optimal mechanisms. In *Proceedings of the 10th ACM Conference on Electronic Commerce*, EC '09, pages 225–234, New York, NY, USA, 2009. ACM.

- [11] Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, pages 45–60, New York, NY, USA, 2015. ACM.
- [12] Yash Kanoria and Hamid Nazerzadeh. Dynamic reserve prices for repeated auctions: Learning from bids - working paper. In *Proceedings of WINE*, page 232, 2014.
- [13] Vahab Mirrokni, Renato Paes Leme, Pingzhong Tang, and Song Zuo. Dynamic auctions with bank accounts. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16*, pages 387–393, 2016.
- [14] Mehryar Mohri and Andres Muñoz Medina. Revenue optimization against strategic buyers. In *Proceedings of NIPS*, pages 2530–2538, 2015.
- [15] Jamie Morgenstern and Tim Roughgarden. The pseudo-dimension of near-optimal auctions. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, NIPS'15*, pages 136–144, Cambridge, MA, USA, 2015. MIT Press.
- [16] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1):58–73, February 1981.
- [17] Renato Paes Leme, Martin Pal, and Sergei Vassilvitskii. A field guide to personalized reserve prices. In *Proceedings of the 25th International Conference on World Wide Web, WWW '16*, pages 1093–1102, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee.
- [18] Tim Roughgarden and Joshua R. Wang. Minimizing regret with multiple reserves. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pages 601–616, New York, NY, USA, 2016. ACM.
- [19] Sarah Sluis. Big changes coming to auctions, as exchanges roll the dice on first-price. <https://adexchanger.com/platforms/big-changes-coming-auctions-exchanges-roll-dice-first-price>, May 5, 2017. Accessed: 2017-10-20.
- [20] Qiqi Yan. Mechanism design via correlation gap. In *Proceedings of the Twenty-second Annual ACM-SLAM Symposium on Discrete Algorithms, SODA '11*, pages 710–719, Philadelphia, PA, USA, 2011. Society for Industrial and Applied Mathematics.