Boosted Second Price Auctions: Revenue Optimization for Heterogeneous Bidders

Negin Golrezaei
golrezae@mit.edu
Massachusetts Institute of Technology
Cambridge, MA, USA

Max Lin, Vahab Mirrokni
whlin,mirrokni@google.com
Google Inc.
New York, NY, USA

Hamid Nazerzadeh
nazerzad@marshall.usc.edu
University of Southern California
Los Angeles, CA, USA

ABSTRACT

The second price auction has been the prevalent auction format used by advertising exchanges because of its simplicity and desirable incentive properties. However, even with an optimized choice of reserve prices, this auction is not revenue optimal when the bidders are heterogeneous and their valuation distributions differ significantly. In order to optimize the revenue of advertising exchanges, we propose an auction format called the boosted second price auction, which assigns a boost value to each bidder. The auction favors bidders with higher boost values and allocates the item to the bidder with the highest boosted bid. We propose a data-driven approach to optimize boost values using the previous bids of the bidders. Our analysis of auction data from Google's online advertising exchange shows that the boosted second price auction with data-optimized boost values outperforms the second price auction and empirical Myerson auction by up to 6% and 3%, respectively.

CCS CONCEPTS

• **Information systems → Online auctions**; • **Theory of computation → Algorithmic game theory**; **Algorithmic mechanism design**; *Nonconvex optimization*; *Computational pricing and auctions.*

KEYWORDS

boosted second price auctions, heterogeneity, data-driven, online advertising, auction design

ACM Reference Format:
Negin Golrezaei, Max Lin, Vahab Mirrokni, and Hamid Nazerzadeh. 2021. Boosted Second Price Auctions: Revenue Optimization for Heterogeneous Bidders. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14–18, 2021, Virtual Event, Singapore.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3447548.3467454

The design of revenue-maximizing selling mechanisms is crucial for many marketplaces, including online advertising markets. The seminal work of [31] shows that when all the bidders are *homogeneous* and their valuations are drawn from the same *regular* distribution, the second price (SP) auction with the optimal reserve, which allocates the item to bidders with the highest submitted bids above the reserve, is revenue maximizing. However, in practical environments, where bidders are heterogeneous and valuation distributions are irregular, the optimal mechanism is more complicated than that of simple SP auctions. In particular, the optimal mechanism does not always allocate the item to the bidder with the highest value, but rather to the bidder to the highest (non-negative) "ironed virtual value," which we will describe in more detail later. The complicated nature of the optimal mechanism makes its implementation challenging. Motivated by this, we design a simple yet effective auction format, called boosted second price (BSP), that can be easily implemented in realistic environments with heterogeneous bidders.

Heterogeneity among Bidders: Inspired by the structure of the optimal mechanism (i.e., the Myerson auction), the BSP auction distorts the allocation rule by not always allocating the item to the bidder with the highest submitted bid. Rather, submitted bids are transformed into "boosted bids," and the bidder with the highest boosted bid wins the item. The bidder's boosted bid is that bidder's boost value multiplied by his submitted bid. BSP auctions are truthful, and bidders never pay more than their submitted bids.

We note that the idea of boosting has also been used in cost-per-click advertising models, where advertisers are charged only when their ads are clicked. These models have been widely used in search advertising markets, and in particular AdWords. There, boost values do not depend on the bidding behavior of advertisers and are quality scores that measure the relevance of an ad to a keyword [15, 18]. In this work, our main focus is on a cost-per-impression model, under which advertisers are charged each time that their ad is displayed. In such a model, which is mainly used in display advertising markets, the boost values can depend on the heterogeneous bidding behavior of buyers.

With this in mind, let us demonstrate the heterogeneity of bidders/advertisers in online display advertising markets. Analyzing bids submitted to Google's advertising exchange, we can roughly order bidders based on their bidding and targeting behaviors. We have summarized the results of these analyses in Figures 1a and 1b.[1] These figures respectively illustrate the participation rate of bidders versus the average and standard deviation of their submitted bids. The data suggest that bidders can be roughly divided into two groups: *brand* and *retargeting*. Brand bidders participate more often in auctions but submit low average bids and exhibit low variation in their submitted bids. Typically, these bidders have broad targeting criteria, as their goal is to display their ads to as many Internet users as possible to create brand awareness. In contrast, retargeting bidders have more restrictive targeting criteria and participate less often in auctions. In addition, the average and standard deviation of their submitted bids are rather high. These bidders usually submit

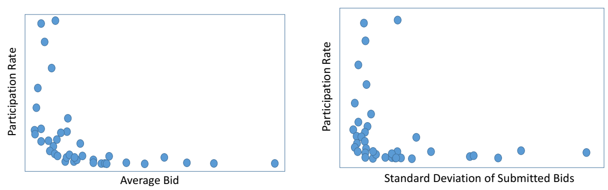
KDD '21, August 14–18, 2021, Virtual Event, Singapore.

ACM ISBN 978-1-4503-8332-5/21/08…$15.00
https://doi.org/10.1145/3447548.3467454

[1]These figures were generated based on the bids submitted for an ad slot with high traffic volume over the course of one day. The dataset was anonymized. Here, an ad slot is a position for a banner ad on a particular publisher's webpage.

low bids unless they are retargeting users who have previously visited their websites.

We note that due to the high variation in the bids submitted by retargeting bidders, SP auctions—even with personalized reserve prices—may not be able to extract high revenue from these bidders. By setting a high reserve, these auctions can yield more revenue from retargeting bidders when these bidders submit high bids. However, the high reserve excludes retargeting bidders from auctions when their bids are low. This can negatively impact revenue, considering that the probability of retargeting a specific Internet user may be low and that, as a result, the bids submitted by these bidders are often low.



(a) **(b)**

Figure 1: (1a) Participation rate of bidders versus their average bids and (1b) participation rate of bidders versus the standard deviation of their submitted bids. Each node represents one bidder. The bidder participation rate is the probability that the bidder clears his reserve price. We do not disclose the values of the x and y axes in these figures because doing so could reveal sensitive dataset information.

To gain insight into how boosts can increase revenue in such environments, let us consider the following example. Assume that there are two bidders. The valuation of the first bidder (brand advertiser) is always 1, and the valuation of the second bidder (retargeting advertiser) is drawn from $F(x) = 1 - \frac{1}{x}$ for any $x \geq 1$. It is easy to see that the optimal personalized reserve price[2] for each bidder is equal to 1, which yields a revenue of 1 for the auction. Now consider a BSP auction mechanism with a (multiplicative) boost of $\beta_1 = 2$ for the first bidder and no boost for the second bidder ($\beta_2 = 1$) and a reserve price of 1 for each bidder. Let b_i, $i = 1, 2$ be the bid of bidder i. We consider the following cases:

- If the second bid, b_2, is in $[1, 2]$, the first bidder, who has the highest boosted bid, equal to $2 = b_1 \times \beta_1 = 1 \times 2$, receives the item at price 1, which equals his reserve price. More specifically, the payment of the boosted bidder is equal to the maximum of the bidder's reserve and the boosted bid of the other bidder divided by his own boost (i.e., $(b_2 \times \beta_2)/\beta_1$); in this case, the maximum will always be equal to 1.
- If the second bid is greater than 2, the second bidder receives the item at a price equal to 2, which is the maximum of his own reserve and the boosted bid of the first bidder divided by the boost of the second bidder, which is equal to $(b_1 \times \beta_1)/\beta_2 = (1\times 2)/1 = 2$; see the description of the BSP auction in Section 2.

Therefore, the expected revenue of the BSP auction would be equal to 1.5, which is a 50% improvement over the SP auction with personalized reserve.

[2]By the optimal personalized reserve price, we mean the monopoly price r that solves $r = \arg\max_x x(1 - F(x))$. For $F(x) = 1 - \frac{1}{x}$, the solution of the aforementioned optimization problem is not unique. However, one can perturb the distribution slightly so that $r = 1$ becomes the unique solution of the optimization problem.

Even though the above example is quite stylized, we will show that the insight extends to practical settings. Specifically, in our counterfactual studies, presented in Section 4, we demonstrate that SP auctions with optimized personalized reserve prices earn 6% less revenue than the BSP auction.

Practical Challenges: As stated earlier, the design of the BSP auction bears some resemblance to the Myerson auction, as they both distort the allocation rule in order to extract more revenue from heterogeneous bidders. However, unlike for the BSP auction, implementing the Myerson auction is very challenging. To implement the Myerson auction, one needs to learn ironed virtual values. From both theoretical and practical perspectives, the process of learning ironed virtual values is complicated and fragile. For instance, [37] show that when distributions are irregular, failure to do ironing can lead to a constant-factor loss in expected revenue. Further, they show that the ironing procedure in Myerson auctions is very sensitive in the sense that a small mistake in the choice of ironing intervals can result in a large loss of expected revenue. [10] empirically evaluate the Myerson auction and highlight the practical challenges of implementing these auctions in online advertising markets. In this work, using a real auction dataset from Google's advertising exchange, we also demonstrate that "empirical Myerson" auctions are not robust and do not perform well in a heterogeneous and realistic environment. To the best of our knowledge, this is the first work that evaluates an empirical Myerson auction for heterogeneous bidders. Our empirical analysis shows that BSP auctions outperform the empirical Myerson auction by 3%. The poor performance of the empirical Myerson auction stems from its sensitivity in estimating the virtual values; see Figure 2.

Considering these challenges in implementing the Myerson auction, we present the BSP auction as a practical yet effective alternative solution for auction design in heterogeneous environments. To implement the BSP auction, one only needs to learn two parameters (boost value and reserve price) per buyer, rather than ironed virtual value "functions." Further, unlike the Myerson auction, the BSP auction does not rely on randomized allocation, which may not plausible from a practical point of view.

To implement the BSP auction, we use a data-driven approach in which we optimize the boost values given a history of bids. Such an approach does not require estimation of the valuation distributions, as the data implicitly reveal such information. Specifically, considering the linearity of the boosted bids, the revenue loss due to the lack of knowledge of valuation distributions vanishes as we get access to more data; c.f. [5, 6]. In addition, unlike the Bayesian classical approach, the data-driven approach does not make any assumption about the valuation distributions and works even if the valuation distributions are irregular.

Our Contributions: We now summarize our contributions.

A Data-Driven Algorithm: Unfortunately, as we show in Theorem 3.1, the data-driven problem of optimizing the boost values is NP-complete. Nonetheless, we propose a data-driven algorithm, called *BSP alternating minimizer* (BSP-AM), that iteratively optimizes the boost for each bidder (c.f. [8, 34]). We show that each iteration of the BSP-AM algorithm can be done in polynomial time.

We observe that the BSP-AM algorithm converges very fast in practice. This is rather surprising because the objective function (revenue function) of the data-driven problem is discontinuous and non-convex. Hence, we analyze the convergence properties of a closely related algorithm with a smooth revenue function(c.f.

[16, 41, 42]). The smooth revenue function is obtained by convolving the (original) revenue function with a smooth probability density function, and it satisfies the following property: the infima of the smooth revenue function converge to that of the revenue function as the support of the probability density function shrinks. We demonstrate in Section 6.1 that the revenue of the BSP auction under the smooth BSP-AM algorithm is very close to that under the BSP-AM algorithm.

We show that the *smooth BSP-AM* algorithm has an exponential convergence rate, justifying the good convergence rate of the BSP-AM algorithm. To obtain this result, we show that the smooth revenue function is analytic and, as a result, satisfies the Kurdyka-Łojasiewicz (KL) inequality [28, 29]. Establishing this property enables us to characterize the convergence rate of the smooth BSP-AM algorithm (c.f. [4]).

Counterfactual Analyses and Comparison with SP and Myerson Auctions: We evaluate the BSP-AM algorithm by conducting counterfactual simulations using data from Google's advertising exchange. We observe that our algorithm converges very quickly and outperforms the SP auctions. Namely, it obtains up to about 30% more revenue than the SP auction with no reserve and up to about 6% more than the SP auction with optimized personalized reserve prices. Moreover, we observe that the BSP auction outperforms the empirical Myerson auction by 3%. To the extent of our knowledge, this is the first paper that evaluates the empirical Myerson auction in a heterogeneous environment.

In order to understand how the bidding behavior of the bidders impacts their boosts, we do a regression analysis of the boosts assigned by our algorithm across different auctions. Our analysis shows that boost values are negatively correlated with the coefficient of variation and standard deviation of the bids of a bidder, confirming our intuition that bidders with more stable bidding patterns would receive higher boost values. We further demonstrate that the BSP auctions maintain their dominance over the SP auctions when bidders strategically shade or compress their bids; see Section B for details. A similar observation has been made by [33].[3]

1 RELATED WORK

Our data-driven approach to design BSP auction is closely related to the literature on the automated mechanism design algorithms, which is originated by [11]. There are two main strands in this literature, where in the first strand, the input is the distribution of the bidders' valuation (e.g., [12, 39]), and in the second strand, the input is the samples from the bidders' valuation distribution (e.g. [7, 13, 27, 38]). Our work is close to the second strand, where the goal is to use the samples to design an auction within a certain class of auctions. While some of the work in this strand including [7, 27], provide sample complexity bounds, which is, roughly speaking, the number of samples required to ensure the designed auction has a good performance, some others, including [13, 38] focus on designing polynomial algorithms that use the available samples to design auctions. The latter is motivated by the fact that the problem of designing an auction using samples is NP-complete in many cases; see Theorem 3.1. Considering this, here in this work, similar to [13, 38], we present a simple and practical algorithm that uses the available samples to optimize the boost values in the BSP auctions.

The design of the BSP auction bears some resemblance to the mechanism in [3]. The aforementioned paper studies adverse selection in online ad markets for the impressions that are sold via auctions versus guaranteed-delivery contracts, where the valuations of the buyers are correlated via a common value component. They show that to address adverse selection, the platform should sometimes allocate the impression to the guaranteed-delivery contracts, even when the bids from the auction are higher. This is similar to assigning higher boosts to those advertisers. In our private-value setting, we do not encounter the adverse selection problem. Nevertheless, we show that assigning boosts based on the bidding patterns of the advertisers can increase revenue.

The BSP auction is designed to deal with the heterogeneity in online advertising market. Another way to deal with heterogeneity is via controlling access of buyers to users' information. Motivated by this fact, several authors have investigated optimal information disclosure in advertising markets; see [9, 17, 20, 23]. In the present paper, however, we assume that all buyers can access information with which to target Internet users. This is usually the case when there are several advertising markets that are in competition with each other, with this competition reducing the markets' abilities to limit buyer access to user information. When all buyers can implement targeting strategies, the resulting market can be very heterogeneous. In this work, we propose a new auction format that enables advertising exchanges to differentiate between buyers based on their targeting behaviors.

2 BOOSTED SECOND PRICE AUCTION

In this section, we present the BSP auction. We assume that there are n bidders, indexed by $i \in [n] = \{1, 2, \ldots, n\}$, who are interested in a good. The BSP auction assigns two parameters to each bidder: reserve price r_i and boost β_i. For each bidder i, the auctioneer computes a boosted bid, which is a function of the submitted bid of bidder i, b_i, and the assigned boost value $\beta_i > 0$. In this paper, we focus on multiplicative boosts, in which the boosted bid of bidder i is given by $b_i\beta_i$. The boosts can also be applied additively. However, to ease the exposition, we only consider one way of applying the boost values. We denote the BSP auction by BSP($\mathbf{r}, \boldsymbol{\beta}$), where $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ is the vector of the reserve prices and $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_n)$ is the vector of the boost values. We now present the BSP auction.

Boosted Second Price Auction (BSP($\mathbf{r}, \boldsymbol{\beta}$))

- First, each bidder i submits his bid b_i.
- Define S as a set of bidders whose bids exceed their reserve prices, i.e., $S = \{i : b_i \geq r_i\}$.
- If set S is empty, the good is not allocated. Otherwise, the good is allocated to bidder i^* with the highest boosted bid, i.e., $i^* = \arg\max_{i \in S}\{b_i\beta_i\}$, and he pays $\max\left\{r_{i^*}, \max_{i \in S, i \neq i^*}\{b_i\beta_i\}/\beta_{i^*}\right\}$. For other bidders, the payment is zero.

In the BSP auction, the winner is the bidder who clears his reserve price and has the highest boosted bid. The winner then pays the maximum of his reserve price and the second-highest boosted bid divided by his own boost value. Note that the payment of the winner, which is the minimum bid that he needs to submit to win the auction, does not exceed the submitted bid of the winner. With standard arguments, since the allocation probability is monotone in

[3] Using a different dataset, [33] show that when the market is thin, which is typically the case in the online advertising markets, the BSP auction is more robust to the strategic behavior of the bidders than the SP auction. Further, they show that the BSP auction is more robust to the strategic behavior of the bidders than the Myerson auction.

the submitted bid and the payment of the winner is not a function of his submitted bid, we obtain the following result.

Proposition 2.1 (BSP Auction Is Truthful). *In BSP auctions, a weakly dominant strategy of each bidder i is to bid truthfully, i.e., b_i is equal to the true value of the bidder, denoted by v_i.*

We now discuss the role of reserve prices in more detail. In the BSP auction, we first discard bidders who do not meet their reserve prices, and then we determine a winner from the rest of the bidders. This way of enforcing reserve prices is called "eager" by [36]. Another way of enforcing reserve prices is the approach known as "lazy," in which the winner is determined first and then the reserve prices are applied. [36] show that eager SP auctions outperform lazy auctions in terms of revenue when the bidders are either independent or symmetric. As such, we assume that reserve prices are applied eagerly in both the BSP and the SP auctions. Note that the SP auction with the vector of reserve prices $\mathbf{r}$ is equivalent to BSP auctions with boosts $\beta_i = \beta$ for $i \in [n]$ and the same vector of reserve prices. Thus, boost values in the BSP auctions allow the seller to further differentiate between bidders.

The fact that boosted bids are linear in submitted bids can simplify the problem of learning parameters of BSP auctions. In particular, the difference in revenue between the BSP auction when boosts are optimized using historical data and the BSP auction when boosts are optimized by having access to the true distribution of valuations is on the order of $U\sqrt{\frac{n \log n}{T}}$, where U is the maximum revenue in an auction, T is the number of previous auctions, and n is the number of bidders; cf., [5, 6]. Motivated by this, in Section 3, we study the problem of learning the optimal boost values using historical data. In addition, in Section B, we discuss the consequence of using historical data in learning the parameters of the BSP auction.

3 DATA-DRIVEN APPROACH

In this section, we present a data-driven approach to optimize the boost values in BSP auctions. Let us define the problem formally. Assume that there are n bidders indexed by $i \in [n]$. Suppose that we have access to the bids submitted by each of these n bidders over the course of T auctions. Let b_i^t be the bid submitted by bidder i in auction $t \in [T]$. We denote the bids submitted by bidders in auction t, i.e., $(b_1^t, b_2^t, \ldots, b_n^t)$, by $\mathbf{b}^t$. Given these submitted bids and reserve prices $\mathbf{r}$, we solve the following optimization problem:

$$\begin{aligned} &\min_{\{\beta_i : i \in [n], \beta_i \in \mathcal{B}\}} -\sum_{t=1}^{T} \mathrm{Rev}(\mathbf{b}^t, \mathbf{r}, \boldsymbol{\beta}) \\ &= \min_{\{\beta_i : i \in [n], \beta_i \in \mathcal{B}\}} \sum_{t=1}^{T} -\max\left\{\frac{b_{(2)}^t \beta_{(2)}^t}{\beta_{(1)}^t}, r_{(1)}\right\}, \quad \text{(Data-Driven BSP)} \end{aligned}$$

where $\mathrm{Rev}(\mathbf{b}^t, \mathbf{r}, \boldsymbol{\beta})$ is the revenue of the BSP auction in period t when the submitted bids, reserve prices, and assigned boost values are respectively $\mathbf{b}^t$, $\mathbf{r}$, and $\boldsymbol{\beta}$. More precisely, $\mathrm{Rev}(\mathbf{b}^t, \mathbf{r}, \boldsymbol{\beta})$ is the second-highest boosted bid divided by the boost value of the winner, i.e., the bidder with the highest boosted bid. Here, $b_{(i)}^t$, $r_{(i)}$, and $\beta_{(i)}$ are respectively the bid, reserve price, and boost value associated with the i^{th} highest boosted bid in auction t. In addition, $\mathcal{B}$ is the feasible region for β_i, $i \in [n]$. We note that in Problem (Data-Driven BSP), we do not optimize over the reserve prices; we determine the optimal boost values given the reserve prices $\mathbf{r}$. However, it is easy to extend our data-driven approach to jointly optimize the reserve prices and boost values. Due to lack of space, we do not provide further details here.

Theorem 3.1. *Data-Driven BSP optimization problem is NP-complete.*

Unless stated otherwise, the proofs of all the results are presented in the appendix. To show Theorem 3.1, we consider an instance of Problem (Data-Driven BSP) in which the feasible region is $\{1, \beta_H\}$. We then reduce this instance of the problem to an edge bipartization problem, which is NP-complete [21]. We note that solving the special case of Problem (Data-Driven BSP), in which the feasible region $\mathcal{B} = \mathbb{R}^+$, is still challenging because its objective function is discontinuous and non-convex in boost values and can have multiple local minima.

In the next section, we present an iterative algorithm, called BSP alternating minimizer (BSP-AM). In this algorithm, we successively optimize one of the boost values while fixing all other boost values (c.f. [8]). A similar approach has shown promise in other applications; see, for example, [34]. As we will show in Theorem 3.2, this simple iterative algorithm is easy to implement because each iteration of the algorithm can be solved effectively. We then study the convergence property of the algorithm in Section 6.

BSP Alternating Minimizer Algorithm

We need a few more definitions to present the algorithm. Let

$$\mathrm{R}(\boldsymbol{\beta}) := -\sum_{t=1}^{T} \mathrm{Rev}(\boldsymbol{b}^t, \boldsymbol{r}, \boldsymbol{\beta}), \tag{1}$$

be the negative of the total revenue of the BSP auction where $\mathrm{Rev}(\boldsymbol{b}^t, \boldsymbol{r}, \boldsymbol{\beta})$ is defined in Problem (Data-Driven BSP). At a discontinuity point $\boldsymbol{\beta}$, we set $\mathrm{R}(\boldsymbol{\beta})$ to $\min(\mathrm{R}(\boldsymbol{\beta}^-), \mathrm{R}(\boldsymbol{\beta}^+))$. We note that the total revenue, $\sum_{t=1}^{T} \mathrm{Rev}(\boldsymbol{b}^t, \boldsymbol{r}, \boldsymbol{\beta})$, is also a function of reserve prices $\boldsymbol{r}$. However, to simplify the notation, here, we do not show this dependency. The BSP-AM algorithm is presented below.

BSP Alternating Minimizer (BSP-AM)
Input: $\lambda > 0$.
Choose $\beta_i = 1$, $i \in [n]$, as our initial values.
Until convergence is reached or for some fixed number of iterations:
For $i \in [n]$, update β_i to

$$\beta_i \;\leftarrow\; \arg\min_{y} \; \mathrm{R}(y, \boldsymbol{\beta}_{-i}) + \lambda(y - \beta_i)^2, \tag{2}$$

where $\mathrm{R}(\boldsymbol{\beta})$ is defined in Equation (1) and $\boldsymbol{\beta}_{-i}$ is the boost value of all the bidders except bidder i.

The algorithm updates the boost value of one bidder at a time. To do so, it solves the one-dimensional optimization problem given in Equation (2). The term $\lambda(y - \beta_i)^2$ in this optimization problem is a regularization term that ensures that the boost values do not change too much in each iteration. Here, $\lambda > 0$ controls the importance of the regularization term. In our empirical studies, we observe that the performance of the BSP-AM algorithm is not sensitive to the choice of λ. [4]

[4]As we show in Theorem 3.2, the objective function of the optimization problem in Equation (2) is one of $O(T)$ discontinuity points or stationary points between any two consecutive discontinuity points. The choice of λ does not impact the discontinuity points and only impacts the stationary points. However, between most pairs of two

To further explain the algorithm, let us set λ to zero. In that case, at any step of the algorithm, the updated boost values increase the revenue of the BSP auction. Thus, when the algorithm terminates, it is guaranteed that the BSP auction outperforms the SP auction with the same vector of reserve prices.

In the following, we show that each iteration of the BSP-AM algorithm is computationally efficient in the sense that the optimization problem in (2) can be solved in polynomial time. To solve problem (2), we evaluate its objective function for $O(T)$ points and return the best one.

Theorem 3.2 (BSP-AM Algorithm). *To compute the optimal solution of problem (2) in the BSP-AM algorithm, it suffices to evaluate its objective function at $O(T)$ points and return the best one.*

The proof of Theorem 3.2 is given in Section A.2. There, we show that for any $\boldsymbol{\beta}_{-i}$, the objective function is piecewise continuous. Specifically, the objective function, $\mathrm{R}(y, \boldsymbol{\beta}_{-i}) + \lambda(y - \beta_i)^2$, has $O(T)$ discontinuity and stationary points, where the discontinuity and stationary points have closed forms.

In Section 6, we study the convergence of the BSP-AM algorithm. We stress that in our empirical studies, the BSP-AM algorithm converges after two or three iterations; see Section 4. We now proceed to evaluate our algorithm using a real auction dataset.

4 EMPIRICAL ANALYSIS

Here, we evaluate the BSP auction by comparing it with the SP auction. Comparison with the empirical Myerson auction is done in Section 5. We use bids submitted to Google's display advertising exchange over the course of one day. The bids in this dataset were submitted to one of the ad slots with the highest traffic volume. The data are provided at the level of an impression, which consists of a set of bidders who participate in an auction run for the impression. For each auction, we have access to the bids submitted by all bidders. We will focus on the auctions in which more than one bidder submitted a positive bid, as in auctions with one bidder, boosting cannot change the auctions' outcome.

We divide our dataset randomly into training and test datasets that have roughly equal size. We use our training dataset to compute the reserve prices and boost values. We then evaluate the computed parameters on the test dataset.

Boost values: Here, we only optimize the boost values of the top bidders[5] with the highest spending using the BSP-AM algorithm with $\lambda = 0.01$. As stated earlier, our algorithm is not sensitive to the value of λ. We note that the revenue gain of the BSP auction (with respect to the SP auction) stays roughly the same when we optimize the boost values of more than 8 top bidders. Thus, we consider the top $k = 2, 4, 6$, and 8 bidders, and we optimize the boost values for these bidders, called *selected bidders*. After optimizing the boost value of a selected bidder, i.e., after solving problem (2) in the BSP-AM algorithm, we normalize the boost values of all the bidders, including the non-selected bidders, such that the minimum boost value is one; that is, without loss of generality, for any $i \in [n]$, we set β_i to $\beta_i / \min_{i' \in [n]} \{\beta_{i'}\}$. Thus, the (normalized) boost values of the bidders can potentially change after updating the boost value of a selected bidder. Then, the algorithm stops at iteration m when either m reaches 200 or $\mathrm{Err} := \sum_{i \in [n]} (\beta_i^m - \beta_i^{m-1})^2 \leq 0.01$. Here, β_i^m is the boost value of bidder i at the end of iteration m.[6]

Reserve prices: We set the reserve prices in two ways. In the first way, the reserve prices of all the bidders are set to zero in both SP and BSP auctions. In the second way, we set the reserves of the top 15 buyers[7] to the monopoly prices, where the monopoly price for bidder i solves the following optimization problem:

$$\bar{r}_i = \arg\max_r \sum_{t \in [T]} r \times \mathbb{I}(b_i^t \geq r)\,. \tag{3}$$

Here, $\mathbb{I}(\cdot)$ is an indicator function; for an event A, $\mathbb{I}(A)$ is one when A happens and zero otherwise. We note that this choice of reserve prices is inspired by how reserve prices are set in this advertising exchange. Further, $\bar{r}_i$ bears some resemblance to the optimal reserve price for the cases where either bidder i is the only bidder or the bidders are all homogeneous in terms of their valuation distribution. There, the optimal reserve price of bidder i is $\arg\max_r \sum_{t \in [T]} r \times \Pr(b_i^t \geq r)$. We use the training dataset to compute $\bar{r}_i$.[8]

Table 1 summarizes the result of our analysis. The first and second columns represent the number of selected bidders and the reserve price, respectively, used in the BSP and SP auctions. As stated earlier, to perform a fair comparison, we use the same reserve prices in the SP and BSP auctions. The third and fourth columns depict the Err and the number of iterations when the algorithm stops, respectively. We observe that the algorithm converges after two or three iterations and that, when it stops, Err is very small. Finally, the last column of the table presents two numbers in the form of (a, b), where a is the average gain of the BSP auction (relative to the SP auction) and b is the standard error of the average gain in the test dataset. To compute these numbers, we sample the test dataset 200 times with a rate of 5%.[9] We then evaluate the revenue gain of the BSP auction in each sampled dataset and take the average.

We now discuss the revenue gain of the BSP auction. When the monopoly reserve prices are used, the revenue gain of the BSP auction can be as low as 2.43% and as high as 6.75%. Overall, the revenue gain increases as the number of selected bidders increases. However, the highest jump in the revenue gain happens when we increase the number of selected bidders from 2 to 4. When increasing the number of selected bidders from 4 to 8, the jump in the revenue gain is non-significant (is less than 0.3%). When the reserve prices are zero, the revenue gain spans from 16.55% to 29.28%. Observe that the revenue gain under zero reserve is much larger than that under the monopoly reserve. This is the case because, with zero reserve, there is more room for improvement. Put differently, when the reserve prices are not set appropriately, then the revenue gain of the BSP auction is higher.

Robustness: Here, we investigate to what extent the revenue gain of the BSP auction is sensitive to the choice of the top bidders. With this aim, we allow the BSP auction to have some flexibility in choosing its selected bidders: the BSP auction randomly chooses its selected bidders from a pool of the top 15 bidders with the highest spending and then employs the BSP-AM algorithm to optimize the boost values of the selected bidders. Here, we focus on the case where the number of selected bidders is 4. This is motivated

consecutive discontinuity points, we observe that there are no stationary points. This explains why our algorithm is not sensitive to the choice of λ.

[5]Optimizing the boost values of these bidders does not necessarily mean that we favor them in the allocation rule.

[6]In each iteration, we update the boost value of each selected bidder once.

[7]For the rest of the bidders, the data are sparse. Thus, we set their reserve to zero.

[8]We obtain similar results when boost values and reserve prices are jointly optimized.

[9]We note that 5% of the test dataset includes the submitted bids in more than 9000 auctions.

# of selected bidders	Reserve price	Err	# of iterations	Average revenue gain & its standard error in test (in %)
2	Monopoly	4e-05	3	(2.43, 0.01)
4		1e-04	2	(6.48, 0.05)
6		5e-04	3	(6.65, 0.04)
8		1e-03	3	(6.75, 0.04)
2	Zero	1e-04	2	(16.55, 0.08)
4		1e-04	3	(18.91, 0.11)
6		6e-03	2	(28.93, 0.14)
8		1e-03	3	(29.28, 0.10)

Table 1: BSP auction versus SP auction

by Table 1, where we show that by updating the boost values of the top 4 bidders, we can obtain a high fraction (more than 90%) of the revenue gain of the BSP auction with monopoly prices.[10] Finally, we restrict our attention to the BSP auction with monopoly reserve prices, as (i) in practice, reserve prices are set to extract more revenue from the bidders [35, 36], and (ii) the revenue gain of the BSP auction is smaller when the reserve prices are set

We observe that the average revenue gain of the BSP auction (relative to the SP auction) is 5.1%[11], where the average is taken over 300 samples. Here, each sample corresponds to a set of 4 bidders (with updated boost values) that has been randomly selected from the top 15 bidders. We note that the average revenue gain is consistent with the result presented in the table.

Who is favored? Next, we would like to provide additional insight into the optimized boost values. We are interested in learning what set of bidders is favored by receiving higher boost values in the BSP auction and how the boost values change. To do so, we use our 300 samples, and for each sample, we choose one of its selected bidders randomly. Let us call him bidder i. Then, we regress the log of the boost value of bidder i, i.e., $\log(\beta_i)$, as follows:

$$\log(\beta_i) \;=\; c_0 + c_1.CV + c_2.\mathrm{STD} + c_3.\mathrm{Scale} + c_4.\mathrm{STD}_{\max} + c_5.\mathrm{STD}_{\min}\,, \tag{4}$$

where CV and STD are the coefficient of variation and standard deviation, respectively, of the submitted bids of bidder i. Scale is defined as $\log(\frac{\beta_{\max}}{\beta_{\min}})$, where $\beta_{\max} = \max_{j\neq i, j\in \mathrm{Sel}}\{\beta_j\}$ and $\beta_{\min} = \min_{j\neq i, j\in \mathrm{Sel}}\{\beta_j\}$. Here, Sel is the set of selected bidders. We consider $\log(\frac{\beta_{\max}}{\beta_{\min}})$ to account for the normalization factor. Recall that the revenue of the BSP auction stays the same, as we scale all the boost values by a constant factor. $\mathrm{STD}_{\max}$ and $\mathrm{STD}_{\min}$ are the standard deviations of the submitted bids of $\beta_{\max}$ and $\beta_{\min}$, respectively.

	Coefficient (c_i's)	t value	p value
Intercept	0.98	5.46	1.03e-07 ***
CV	-0.16	-7.01	1.64e-11 ***
STD	-1.07	-5.57	5.70e-8 ***
Scale	0.25	4.54	7.98e-6 ***
$\mathrm{STD}_{\max}$	1.49	5.05	7.60e-7 ***
$\mathrm{STD}_{\min}$	-0.04	-0.26	0.798

Table 2: Result of regressing the log of a boost value on the CV, STD, Scale, $\mathrm{STD}_{\max}$, and $\mathrm{STD}_{\min}$. Here, R^2 and the adjusted R^2 are respectively 0.29 and 0.28. Significance codes: 0 '*', 0.001 '**' , 0.01 '*'.**

[10] By updating the boost values of the top 15 bidders, the BSP auction with the monopoly reserves outperforms the SP auction by 6.51%. We obtain a similar result for a higher number of selected bidders.

[11] The standard error of the average revenue gain is 0.25%.

The result of the regression is provided in Table 2. All the considered parameters in the regression are significant, except $\mathrm{STD}_{\min}$. We point out that as the volatility (CV and STD) in the submitted bids of bidders increases, the boost values decrease. Thus, the BSP auction favors the bidders with more stable bidding behavior. As another observation, when the standard deviation of a selected bidder with the highest boost value increases, the boost value of the bidder goes up. Put differently, a bidder is favored more when there is higher variation in the submitted bids of the other bidders.

5 EMPIRICAL MYERSON AUCTION

We start by defining the optimal mechanism (Myerson auction) in the the following standard model. Assume that there are n bidders who participate in an auction for a single good valued at zero by the seller. In the context of online advertising, the seller (auctioneer) is an advertising exchange, bidders are advertisers, and the good is an advertisement opportunity (impression). Bidders are risk neutral. The valuation of a bidder i, which is denoted by v_i, is drawn from distribution $F_i : [\underline{v}_i, \bar{v}_i] \to [0, 1]$. The valuations of bidders are private and are independent of each other. However, distributions $F_1, F_2, \ldots, F_n$ are public information. We assume that the distribution F_i, $i \in [n]$ has density $f_i : [\underline{v}_i, \bar{v}_i] \to \mathbb{R}^+$. We denote the IHR associated with distribution F_i by $\alpha_i : [\underline{v}_i, \bar{v}_i] \to \mathbb{R}^+$ and define it as $\alpha_i(x) = \frac{1-F_i(x)}{f_i(x)}$. Furthermore, we denote the virtual value associated with distribution F_i by $\theta_i : [\underline{v}_i, \bar{v}_i] \to \mathbb{R}$, and define it as $\theta_i(x) = x - \alpha_i(x)$.

Optimal Mechanism [31]: First, suppose that distributions F_i are regular, i.e., the virtual value $\theta_i(\cdot)$ is increasing for all bidders. If bidders are symmetric, that is, $F_i = F$ for any $i \in [fn]$, then the optimal mechanism is an SP auction with a reserve price r where r solves $r = \theta^{-1}(0)$ and $\theta(x) = x - \frac{1-F(x)}{f(x)}$. In this case, the optimal mechanism depends on the distribution of valuations only via the reserve price r. This is so because in the SP auction, the good is allocated to the bidder with the highest submitted bid.

When the bidders are not symmetric, SP auctions cannot earn the maximum revenue. In the optimal mechanism, for each bidder i whose submitted bid b_i is greater than his reserve price r_i, a virtual bid $\theta_i(b_i) = b_i - \frac{1-F_i(b_i)}{f_i(b_i)}$ is calculated, where $r_i = \theta_i^{-1}(0)$.

If the virtual bids are non-monotone functions (i.e., irregular distributions), "ironing" needs to be done to make them monotone. Roughly speaking, the ironing procedure makes the non-monotone virtual values monotone by bunching certain values together and treating them the same. Finally, the good is allocated to bidder i^* with the highest non-negative (ironed) virtual bid, and bidder i^* pays $\max\{r_{i^*}, \theta_{i^*}^{-1}(\theta_j(b_j))\}$, where $j = \arg\max_{i\neq i^*}\{\theta_i(b_i)\}$ is the bidder with the second-highest virtual bid.

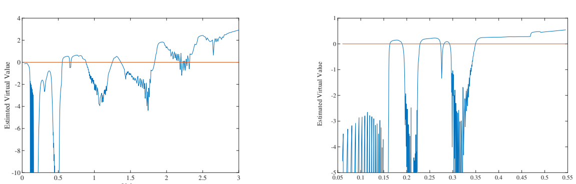
5.1 Empirical Myerson Auction

Now that we have defined the optimal mechanism, one may wonder why we did not work with this auction format from the start, as this auction can also capture the heterogeneity among bidders. One of the reasons behind this choice is the simplicity of the BSP auction: the BSP auction needs to learn two parameters per buyer, while the Myerson auction needs to learn a function (ironed virtual value) per bidder. However, one might be willing to endure the underlying complexity of the Myerson auction in order to improve revenue. In the following, we show that, due to complexities of practical environments, the empirical Myerson auction performs worse than the BSP auction with the monopoly reserve prices. To the best of our knowledge, this is the first work that empirically evaluates the Myerson auction in a realistic environment with heterogeneous bidders.[12]

To run an empirical Myerson auction, we use the following steps, similar to [10], for each of the top 15 bidders: [13]

- We empirically estimate the probability density of the submitted bids of each of these top bidders using the Kernel method. To use this method, we need to adjust its bandwidth. To do so, we use an adaptive bandwidth approach. In particular, we start with Silverman's rule-of-thumb ([40]) as a global bandwidth, and for any grid point for which less than 1% of all data points are available, we locally increase the bandwidth.
- Using the estimated probability density functions, we estimate the virtual value of the top bidders, $\theta_i(v) = v - \frac{1-F_i(v)}{f_i(v)}$. Figure 2 shows the estimated virtual values for two of the top 15 bidders. Observe that the virtual values are non-monotone and need to be ironed.
- We then iron the virtual bids. In particular, we compute $H_i(v) = \int_0^v \theta_i(z)dz$ and its convex hull. The linear segments of the hull get us the ironing regions.

By following this procedure, we obtain (ironed) virtual bids, and this allows us to run an empirical Myerson auction on our dataset. We observe that the empirical Myerson auction only obtains around 97% of the revenue of the BSP auction[14] with the monopoly reserve prices and outperforms the SP auction by 3%-4%. Here, we do not claim that we run the best version of the empirical Myerson auction; this auction format can be very sensitive to ironing regions, and characterizing the best ironing regions is computationally expensive. Our goal here is to demonstrate that a naive but practical way of running a Myerson auction does not necessarily perform well. We note that the inferior performance of the Myerson auction is due to its sensitivity to the estimation error in virtual values.



(a) (b)

Figure 2: Virtual values for two of the (top) buyers. To maintain data confidentiality, we scale both x and y axes by a constant.

[12] We point out that using a different dataset, [10] also assess the Myerson auction in a *homogeneous* environment, where all the bidders have the same distribution of valuations, and they show that a simple randomized auction could outperform empirical Myerson auctions.

[13] For other bidders, the data are sparse, and as a result, the estimation of the probability density functions is rather noisy. Thus, for these bidders, we assume that their virtual bids are equal to their submitted bids.

[14] In the BSP auction, we optimize the boost value of the top 15 bidders, as we compute the virtual values of the top 15 bidders. However, the results are similar even if we optimize the boost value of the top 4 bidders.

6 CONVERGENCE OF BSP-AM

Analyzing the convergence property of the BSP-AM algorithm is very challenging, if not impossible, because the revenue function is non-continuous and non-convex. To the best of our knowledge, the convergence of the alternating minimizer algorithms is mainly analyzed when the objective function is continuous [22, 43, 44]. In fact, beside the continuity assumption, additional assumptions such as convexity or separability are imposed to characterize the convergence property of the alternating minimizer algorithms [22]. Nonetheless, we would like to provide insight into the convergence of the algorithm and justify its good performance in our empirical studies. With this aim, we consider an algorithm that is closely related to the BSP-AM algorithm and analyze its convergence property. In Section 6.1, via empirical studies, we show that performance of this algorithm is very similar to that of the BSP-AM algorithm. In this algorithm, which we call a "smooth BSP-AM" algorithm, the revenue function in Equation (2) is replaced with a "smooth revenue function." Precisely, we smooth out the revenue function using an averaging technique [16, 41, 42] and use the BSP-AM algorithm on the smooth function. To smooth out the revenue function, we convolve it with a (differentiable) probability density function; see Definition 6.1.

We point out that the "smooth revenue function" has nice properties. As we show in Proposition 6.2, the infima of the smooth revenue function converge to that of the revenue function when the support of the probability density function goes to zero. This justifies the study of the BSP-AM algorithm when it applies to the smooth revenue function. We show that the *smooth BSP-AM algorithm* has an exponential convergence rate, rationalizing our empirical results.

Smooth Revenue Functions: The smooth revenue functions are defined using a family of probability density functions $\{\phi_a(\cdot) : \mathbb{R}^n \to \mathbb{R}^+, a \in \mathbb{R}^+\}$. Specifically, a smooth revenue function is obtained by convoluting a probability density function $\phi_a(\cdot)$ with the revenue function R, defined in (1). Later, we specify a particular family of probability density functions that we will use in constructing the smooth revenue functions.

Definition 6.1. Let $\{\phi_a(\cdot) : \mathbb{R}^n \to \mathbb{R}^+, a \in \mathbb{R}^+\}$ satisfy the the following two properties: (1) For any $a \in \mathbb{R}^+$, $\phi_a(\cdot)$ is a continuously differentiable probability density function. (2) The support of $\phi_a(\cdot)$ goes to zero as a converges to zero:[15] That is, supp $\phi_a := \{\boldsymbol{z} \in \mathbb{R}^n | \phi_a(\boldsymbol{z}) > 0\} \subset [-U_a, U_a]$ with $U_a \to 0$ when $a \to 0$. Then, the associated smooth revenue functions $\{\mathrm{R}_a, a \in \mathbb{R}^+\}$ are defined by

$$\mathrm{R}_{a,\phi}(\boldsymbol{\beta}) := \int_{\mathbb{R}^n} \mathrm{R}(\boldsymbol{\beta} - \boldsymbol{z})\phi_a(\boldsymbol{z})d\boldsymbol{z}\,. \tag{5}$$

Note that since ϕ_a's are continuously differentiable functions, the smooth functions R_a's are also continuously differentiable [41].

The following proposition gives a justification for replacing the revenue function with its smooth versions.

Proposition 6.2 (Convergence of Infima). *Let a_k, $k \in \mathbb{N}$, be a sequence of positive real numbers such that a_k goes to zero as k goes to infinity. Then, for any $\boldsymbol{\beta} \in \mathbb{R}^n$ and a family of probability density*

[15] One way to construct such a family of probability density functions is by letting $\phi_a(\boldsymbol{z}) = \frac{\phi(\boldsymbol{z}/\rho_a)}{(\rho_a)^n}$ where $\phi(\cdot)$ is a continuously differentiable probability density function with a bounded support and $\rho_a \to 0$ as $a \to 0$.

functions $\{\phi_a(\cdot) : \mathbb{R}^n \to \mathbb{R}^+, a \in \mathbb{R}^+\}$ *that satisfy the properties in Definition 6.1, we have*

$$\liminf_{k\to\infty} R_{a_k,\phi}(\boldsymbol{\beta}_k) \;\geq\; R(\boldsymbol{\beta}) \text{ for all } \quad \boldsymbol{\beta}_k \to \boldsymbol{\beta} \;;$$

$$\liminf_{k\to\infty} R_{a_k,\phi}(\boldsymbol{\beta}_k) \;=\; R(\boldsymbol{\beta}) \text{ for some } \quad \boldsymbol{\beta}_k \to \boldsymbol{\beta} \;.$$

Further, whenever $\boldsymbol{\beta}$ *is a local minimum of the revenue function, we have*

$$0 \in \partial_\phi R(\boldsymbol{\beta}) \,, \tag{6}$$

where $\partial_\phi R(\boldsymbol{\beta}) = \limsup_{k\to\infty}\{\nabla R_{a_k,\phi}(\boldsymbol{\beta}_k) \mid \boldsymbol{\beta}_k \to \boldsymbol{\beta}\}$.[16]

In the rest of this section, we consider a particular family of the probability density functions. We set $\phi_a(\mathbf{z}) = \prod_{i\in[n]} \psi_a(z_i)$, where

$$\psi_a(z) = \begin{cases} \frac{1}{4a^3}(z+2a)^2 & \text{if } z \in [-2a,-a], \\ \frac{1}{4a^3}(2a^2-z^2) & \text{if } z \in [-a,a], \\ \frac{1}{4a^3}(z-2a)^2 & \text{if } z \in [a,2a], \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

It is easy to verify that $\phi_a(\mathbf{z})$ satisfies the properties in Definition 6.1. Focusing on this family of probability density functions allows us to show that the smooth revenue function $\mathrm{R}_{a,\phi}$ is analytic and, as a result, satisfies the Kurdyka-Łojasiewicz (KL) inequality [28, 29]. Łojasiewicz, in his seminal work on real analytic functions, asserts a condition called "Łojasiewicz inequality" that guarantees the convergence of the steepest descent equation to critical points. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real analytic function and let a be a critical point of f. Then, the Łojasiewicz inequality states that there exists some $\theta \in [\frac{1}{2}, 1)$, constant $c > 0$, and an open neighborhood of W of a such that for any $z \in W$, we have

$$|f(z) - f(a)|^\theta \leq c \cdot |\nabla f(z)| \,. \tag{8}$$

Roughly speaking, this inequality ensures that function f is steep enough around its critical points. The Łojasiewicz inequality was later generalized by [25].

Smooth BSP-AM Algorithm: So far, we have investigated the properties of the smooth revenue function. Next, we characterize the convergence properties of the smooth BSP-AM algorithm, presented below.

Smooth BSP-AM Algorithm

Input: $a \in \mathbb{R}^+$.

Choose $\beta_i = 1$, $i \in [n]$, as our initial values.

Until convergence is reached or for some fixed number of iterations:

For $i \in [n]$, update β_i to

$$\beta_i \;=\; \arg\min_y \; \mathrm{R}_{a,\phi}(y, \boldsymbol{\beta}_{-i}) + \lambda(y-\beta_i)^2 \,, \tag{9}$$

where $\mathrm{R}_{a,\phi}(\boldsymbol{\beta})$ and the probability density function ϕ_a are defined in Equations (5) and (7), respectively.

Observe that the smooth BSP-AM algorithm is very similar to the BSP-AM algorithm. The difference is that to update the boost value of one bidder in Equation (2), the revenue function R is replaced by the smooth revenue function $\mathrm{R}_{a,\phi}$. The following theorem presents the convergence properties of the smooth BSP-AM algorithm.

[16]The set $\limsup_{k\to\infty}\{\nabla \mathrm{R}_{a_k,\phi}(\boldsymbol{\beta}_k) \mid \boldsymbol{\beta}_k \to \boldsymbol{\beta}\}$, which is called ϕ-subgradient, is defined in [16]. Note that this set depends on the family of density functions $\{\phi_a(\cdot) : \mathbb{R}^n \to \mathbb{R}^+,\, a \in \mathbb{R}^+\}$. For more details, please see [16].

Theorem 6.3. *Let* $\boldsymbol{\beta}^k = (\beta_1^k, \ldots, \beta_n^k)$ *be the boost values in iteration* k *of the smooth BSP-AM algorithm. Then,* $\boldsymbol{\beta}^k$ *converges to a critical point of the smooth revenue function. Further, there exists* $c > 0$ *and* $\tau \in [0, 1)$ *such that* $||\boldsymbol{\beta}^k - \boldsymbol{\beta}^\infty|| \leq c\tau^k$ *.*

To show Theorem 6.3, we show that the smooth revenue function is analytic and, as a result, satisfies the KL inequality. Then, we show that the Łojasiewicz inequality is satisfied with $\theta = \frac{1}{2}$; see Equation (8). This allows us to invoke Theorem 11 of [4] or Theorem 2.9 of [45] to show that the convergence rate of the smooth BSP-AM algorithm is exponential.[17]

6.1 BSP-AM versus Smooth BSP-AM

To compare the performance of the smooth algorithm with that of the original algorithm, we consider the setting in Section 4. In particular, we only optimize the boost values of the top bidders, and we use the monopoly reserve prices.

To run the smooth BSP-AM algorithm, we set a to 0.01. In this algorithm, we convolve the revenue function with the probability density $\phi_a(\cdot)$, defined in Equation (7), where a determines the support of the distribution. Given this choice of a, we approximately solve the optimization problem in Equation (9). To do so, we use the fact that as a gets smaller, the infima of the smooth revenue function converge to those of the revenue function; see Proposition 6.2. Specifically, to solve (9), we only evaluate the smooth revenue function at the discontinuity and stationary points of the revenue function and choose the one that maximizes the smooth revenue function. Recall that the optimal solution of the optimization problem in the BSP-AM algorithm is one of $O(T)$ discontinuity and stationary points of the revenue function. Although we only consider these $O(T)$ points, this does not necessarily mean that the optimization problems in the BSP-AM and its smooth version lead to the same solution. Due to the impact of smoothing, the best of these $O(T)$ points may not be the same for the revenue function and its smooth version. To compute the smooth revenue function at a point, we sample 100 times from distribution $\phi_a(\cdot)$ and take the average over these samples. Table 3 shows the performance of the smooth BSP-AM algorithm when the number of the selected bidders is 2, 4, 6, or 8. We observe that the convergence speed and revenue gain of the smooth BSP-AM are very close to those of the BSP-AM algorithm.

# of selected bidders	Err	# of iterations	Avg Rev gain & its Std Err in test (in %)
2	2e-6	3	(2.42, 0.01)
4	0	4	(6.33, 0.05)
6	1e-6	3	(6.53, 0.04)
8	1e-03	3	(6.66, 0.04)

Table 3: BSP versus SP auction. Here, the boost values are optimized using the smooth BSP-AM algorithm, and reserve prices are equal to the monopoly prices.

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.* 1169–1177.

[17]The results of [4] is for two-blocks alternating minimizer. However, [45] showed that these results can be extended to the case where we have multiple blocks. In our setting, each block corresponds to one boost value that we optimize over.

[2] Kareem Amin, Afshin Rostamizadeh, and Umar Syed. 2014. Repeated contextual auctions with strategic buyers. In *Advances in Neural Information Processing Systems*. 622–630.

[3] Nick Arnosti, Marissa Beck, and Paul Milgrom. 2016. Adverse selection and auction design for internet display advertising. *The American Economic Review* 106, 10 (2016), 2852–2866.

[4] Hédy Attouch, Jérôme Bolte, Patrick Redont, and Antoine Soubeyran. 2010. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality. *Mathematics of Operations Research* 35, 2 (2010), 438–457.

[5] Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S Matthew Weinberg. 2014. A simple and approximately optimal mechanism for an additive buyer. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*. IEEE, 21–30.

[6] Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. 2017. Sample Complexity of Multi-Item Profit Maximization. *arXiv preprint arXiv:1705.00243* (2017).

[7] Maria-Florina F Balcan, Tuomas Sandholm, and Ellen Vitercik. 2016. Sample complexity of automated mechanism design. In *Advances in Neural Information Processing Systems*. 2083–2091.

[8] Amir Beck. 2017. *First-Order Methods in Optimization*. Vol. 25. SIAM.

[9] Dirk Bergemann and Martin Pesendorfer. 2007. Information structures in optimal auctions. *Journal of economic theory* 137, 1 (2007), 580–609.

[10] L Elisa Celis, Gregory Lewis, Markus Mobius, and Hamid Nazerzadeh. 2014. Buy-it-now or Take-a-chance: Price Discrimination through Randomized Auctions. *Management Science* 60, 12 (2014), 2927–2948.

[11] Vincent Conitzer and Tuomas Sandholm. 2002. Complexity of mechanism design. *arXiv preprint cs/0205075* (2002).

[12] Vincent Conitzer and Tuomas Sandholm. 2003. Applications of automated mechanism design. (2003).

[13] Mahsa Derakhshan, Negin Golrezaei, and Renato Paes Leme. 2019. Lp-based approximation for personalized reserve prices. In *Proceedings of the 2019 ACM Conference on Economics and Computation*. 589–589.

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

[15] 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.

[16] Yuri M Ermoliev, Vladimir I Norkin, and Roger JB Wets. 1995. The minimization of semicontinuous functions: mollifier subgradients. *SIAM Journal on Control and Optimization* 33, 1 (1995), 149–167.

[17] Péter Eső and Balazs Szentes. 2007. Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies* 74, 3 (2007), 705–731.

[18] Anindya Ghose and Sha Yang. 2009. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management science* 55, 10 (2009), 1605–1622.

[19] Negin Golrezaei, Adel Javanmard, Vahab Mirrokni, et al. 2021. Dynamic Incentive-Aware Learning: Robust Pricing in Contextual Auctions. *Operations Research* 69, 1 (2021), 297–314.

[20] Negin Golrezaei and Hamid Nazerzadeh. 2017. Auctions with dynamic costly information acquisition. *Operations Research* 65, 1 (2017), 130–144.

[21] Jiong Guo, Jens Gramm, Falk Hüffner, Rolf Niedermeier, and Sebastian Wernicke. 2006. Compression-based fixed-parameter algorithms for feedback vertex set and edge bipartization. *J. Comput. System Sci.* 72, 8 (2006), 1386–1396.

[22] Trevor Hastie, Robert Tibshirani, and Martin Wainwright. 2015. *Statistical learning with sparsity: the lasso and generalizations*. CRC press.

[23] 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.

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

[25] Krzysztof Kurdyka. 1998. On gradients of functions definable in o-minimal structures. In *Annales de l'institut Fourier*, Vol. 48. Chartres: L'Institut, 1950-, 769–784.

[26] Sebastien Lahaie, Andrés Muñoz Medina, Balasubramanian Sivan, and Sergei Vassilvitskii. 2017. Testing Incentive Compatibility in Display Ad Auctions. (2017).

[27] Anton Likhodedov and Tuomas Sandholm. 2004. Methods for boosting revenue in combinatorial auctions. In *AAAI*. 232–237.

[28] Stanislaw Lojasiewicz. 1963. Une propriété topologique des sous-ensembles analytiques réels. *Les équations aux dérivées partielles* 117 (1963), 87–89.

[29] Stanislas Łojasiewicz. 1993. Sur la géométrie semi-et sous-analytique. *Ann. Inst. Fourier* 43, 5 (1993), 1575–1595.

[30] Mohammad Mahdian, Vahab Mirrokni, and Song Zuo. 2017. Incentive-Aware learning for large markets. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee.

[31] Roger B Myerson. 1981. Optimal auction design. *Mathematics of operations research* 6, 1 (1981), 58–73.

[32] Hamid Nazerzadeh, Renato Paes Leme, Afshin Rostamizadeh, and Umar Syed. 2016. Where to sell: Simulating auctions from learning algorithms. In *Proceedings of the 2016 ACM Conference on Economics and Computation*. ACM, 597–598.

[33] Thomas Nedelec, Noureddine El Karoui, and Vianney Perchet. 2019. Learning to bid in revenue-maximizing auctions. *arXiv preprint arXiv:1902.10427* (2019).

[34] Yu Nesterov. 2012. Efficiency of coordinate descent methods on huge-scale optimization problems. *SIAM Journal on Optimization* 22, 2 (2012), 341–362.

[35] Michael Ostrovsky and Michael Schwarz. 2011. Reserve prices in internet advertising auctions: a field experiment. In *Proceedings of the 12th ACM conference on Electronic commerce*. ACM, 59–60.

[36] Renato Paes Leme, Martin Pál, and Sergei Vassilvitskii. 2016. A Field Guide to Personalized Reserve Prices. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1093–1102.

[37] Tim Roughgarden and Okke Schrijvers. 2016. Ironing in the dark. In *Proceedings of the 2016 ACM Conference on Economics and Computation*. ACM, 1–18.

[38] Tim Roughgarden and Joshua R Wang. 2019. Minimizing regret with multiple reserves. *ACM Transactions on Economics and Computation (TEAC)* 7, 3 (2019), 1–18.

[39] Tuomas Sandholm. 2003. Automated mechanism design: A new application area for search algorithms. In *International Conference on Principles and Practice of Constraint Programming*. Springer, 19–36.

[40] Bernard W Silverman. 2018. *Density estimation for statistics and data analysis*. Routledge.

[41] Sergej L Sobolev. 2008. *Some applications of functional analysis in mathematical physics*. Vol. 90. American Mathematical Soc.

[42] VA Steklov. 1983. Basic problems of mathematical physics.

[43] Paul Tseng. 2001. Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of optimization theory and applications* 109, 3 (2001), 475–494.

[44] Paul Tseng et al. 1988. Coordinate ascent for maximizing nondifferentiable concave functions. (1988).

[45] Yangyang Xu and Wotao Yin. 2013. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM Journal on imaging sciences* 6, 3 (2013), 1758–1789.

A APPENDICES: PROOFS OF MAIN RESULTS

A.1 Proof of Theorem 3.1

To show the result, we consider an instance of this problem in which the feasible region, $\mathcal{B}$, is $\{1, \beta_H\}$ and the reserve prices are set to zero. Here, we would like to find a subset of buyers to whom we assign a higher boost values in order to maximize the revenue of the BSP auction. That is, we would like to partition the buyers in two groups and favor one group over the other one. We reduce this instance of the problem to an edge bipartization problem, which is NP-complete [21]. The edge bipartization problem is the problem of deleting as few edges as possible to make a graph bipartite.

Let H and L be constants such that $H > L$. Given a graph $G = (V, E)$, we map the edge bipartization problem on this graph to the following instance of Problem (Data-Driven BSP) with $|V|$ bidders and $2|E|$ auctions. Each vertex is a bidder, and each edge $e = (u, v)$ represents two auctions. In the first auction, the bids of bidders u and v are respectively H and L, and the bids of other bidders are zero. In the second auction, the bids of bidders u and v are respectively L and H, and the bids of other bidders are zero. Furthermore, we set β_H to $\frac{H}{L}$.

In the following, we describe the structure of the optimal solution of Problem (Data-Driven BSP) for the described instance. Consider an edge $e = (u, v)$. If bidders u and v have the same boosts, then the expected revenue from the auctions associated with this edge will be $2L$. Now, if one of the bidders has a boost of $\frac{H}{L}$ and the other one has a boost of 1, the expected revenue from the auctions associated with this edge is $H + \frac{L^2}{H}$, which is always greater than $2L$. Now let us assign the boosts to the bidders arbitrarily. Let S_H be the set of bidders/vertices with a boost of $\frac{H}{L}$, and let S_L be the set of bidders/vertices with a boost of 1. Assume that there is no edge within set S_L and that there is no edge within set S_H. That is, every edge connects a vertex in S_L to a vertex in S_H. Then, graph G is a bipartite graph whose partition has the parts S_H and S_L. In that case, the total revenue is $|E| \times (H + \frac{L^2}{H})$, which is the maximum possible revenue. When there are k edges within S_L and S_H, i.e., there are k edges that do not connect S_H to S_L, then we have k edges (u, v) such that the boosts of u and v are the same. In that case, the revenue is $(|E| - k) \times (H + \frac{L^2}{H}) + k \times 2L$. Since $H + \frac{L^2}{H} < 2L$, we prefer to minimize the number of edges that do not connect S_H and S_L. In other words, the best solution of Problem (Data-Driven BSP) is obtained by finding as few edges as possible to make graph G bipartite, which is an NP-complete problem.

A.2 Proof of Theorem 3.2

We first show that problem (2) can be solved effectively. We then show that the convergence of the BSP-AM algorithm. We start with simplifying the first term of the objective function in Equation (2), which is the negative of the revenue of the BSP auction. Let $S^t = \{i : b_i^t \geq r_i\}$ be the set of bidders who clear their reserve price in auction $t \in [T]$. In addition, define $\tau_i = \{t \in [T] : b_i^t \geq r_i, |S^t| \geq 2\}$ as the set of the auctions in which both $b_i^t \geq r_i$ and the number of bids that clear their reserve price is at least two.[18] Let h_{-i}^t and s_{-i}^t, $t \in \tau_i$, be respectively the highest and second-highest[19] boosted bids of the bidders in set $S^t \backslash \{i\}$. Define r_{-i}^* and β_{-i}^* respectively as the reserve price and boost value associated with h_{-i}^t. Then, the first term of the objective function can be written as

$$\begin{aligned} \mathrm{R}(y, \boldsymbol{\beta}_{-i}) = -\sum_{t \in \tau_i} \Bigg[& \mathbb{I}\{b_i^t y \ \geq \ h_{-i}^t\} \max\left\{\frac{h_{-i}^t}{y}, r_i\right\} \\ & + \mathbb{I}\{h_{-i}^t \ > \ b_i^t y \ \geq \ s_{-i}^t\} \max\left\{\frac{b_i^t y}{\beta_{-i}^*}, r_{-i}^*\right\} \\ & + \mathbb{I}\{s_{-i}^t \ > \ b_i^t y\} \max\left\{\frac{s_{-i}^t}{\beta_{-i}^*}, r_{-i}^*\right\} \Bigg], \end{aligned} \tag{10}$$

Note that in the first term, bidder i wins the auction and he pays $\max\left\{\frac{h_{-i}^t}{y}, r_i\right\}$. In the second and third terms, he does not win. In the second term, bidder i has the second highest boost bid and because of this, his boost value influences the revenue of the BSP auction. Whereas, in the third term, the boost value of bidder i does not have any impact on the revenue of the BSP auction. Note that $\mathrm{R}(\boldsymbol{\beta}_{-i}, y)$ is discontinuous when the second highest boosted bid in an auction is equal to the first highest boosted bid. This is the case because at this point, the winner changes as we change the boost of bidder i. This implies that at $y = \frac{h_{-i}^t}{b_i^t}$, $t \in \tau_i$, the objective function of the optimization problem in Equation (2) can be discontinuous. Consider two consecutive discontinuity points. By Equation (10), the objective function in Equation (2) can be written as $\mathrm{R}(y, \boldsymbol{\beta}_{-i}) + \lambda(\beta_i - y)^2 = -\frac{A}{y} - By - C + \lambda(\beta_i - y)^2$, where A, B, and C are nonnegative numbers that do not depend on y.[20] Observe that the derivative of the objective function, $\mathrm{R}(y, \boldsymbol{\beta}_{-i}) + \lambda(\beta_i - y)^2$, w.r.t. y is $\frac{A}{y^2} - B - 2\lambda(\beta_i - y)$, and as a result, the stationary points of the objective function have a closed form. This and the fact that the number of discontinuity points is linear in the number of auctions, we conclude that the optimization problem in (2) can be solved by evaluating its objective function at $O(T)$ points.

A.3 Proof of Proposition 6.2

We first argue that the revenue function $\mathrm{R}(\boldsymbol{\beta})$ is strongly lower semicontinuous. Then, the result follows from Theorems 3.7 and 4.7 of [16]. Function $f : \mathbb{R}^n \to \mathbb{R}$ is lower semicontinuous if for any $z \in \mathbb{R}^n$, we have $\liminf_{\tilde{z} \to z} f(\tilde{z}) \geq f(z)$. Further, a lower semicontinuous function is strongly lower semicontinuous if for every $z \in \mathbb{R}^n$, there exists a sequence $z_k \to z$ such that function f is continuous for all z_k, for all $k \in \mathbb{N}$, and $f(z_k) \to z$. Roughly speaking, strong lower semicontinuity implies that point $(z, f(z))$ can be reached by following a continuos path. Now it is easy to verify that the revenue function R is strongly lower semicontinuous. To see why observe that the revenue function is discontinuous at finite number of points and by our definition, at any points of discontinuity z, $\mathrm{R}(z) = \min(\mathrm{R}(z^-), \mathrm{R}(z^+))$.

A.4 Proof of Theorem 6.3

We will first show that the smooth revenue function is analytic. This implies that this function satisfies the KL inequality. Further, by Remark 3.13 of [16], the gradient of the smooth revenue function is locally Lipschitz. Establishing these properties, we then invoke Theorem 9 of [4]. By this theorem, either $\boldsymbol{\beta}^k$ goes to infinity or it converges to a critical point of the smooth revenue function. One can argue that $\boldsymbol{\beta}^k$ cannot to infinity. To see why first observe that the revenue function only depend on the ratio of the boost values. That is, if the boost value of two buyers scale at the same rate, the revenue function stays the same. Then, note that when we update boost value of buyer i, we always have $\max_j\{\frac{\beta_i^k}{\beta_j^{k-1}}\} \leq \frac{b_{\max}}{b_{\min}}$, where $b_{\max}$ and $b_{\min}$ are respectively maximum and minimum non-zero bids and β_i^k is the boost value of buyer i in iteration k. To see why note that when we increase β_i^k such that $\max_j\{\frac{\beta_i^k}{\beta_j}\}$ exceeds $\frac{b_{\max}}{b_{\min}}$, buyer i will be winner in all the auctions in which he submitted a positive bid. Thus, by increasing boost of buyer i beyond $\max_j\{\frac{\beta_i}{\beta_j}\} \leq \frac{b_{\max}}{b_{\min}}$, the revenue only decreases, i.e., the revenue function increases. Based on this argument, in the smooth BSP-AM algorithm, the ratio of any two boost values cannot go to infinity.

Next, we will show that the Łojasiewicz inequality satisfies with $\theta = \frac{1}{2}$; see Equation (8). This allows us to invoke Theorem 11 of [4] to show that the rate of convergence of the smooth BSP-AM algorithm is exponential.

We show that for any given $\boldsymbol{\beta}_{-i}$, function $y \mapsto \mathrm{R}_{a,\phi}(y, \boldsymbol{\beta}_{-i})$ is analytic. By definition, $\mathrm{R}_{a,\phi}(y, \boldsymbol{\beta}_{-i}) = \int_{\mathbb{R}^n} \mathrm{R}(y - z_i, \boldsymbol{\beta}_{-i} - z_{-i})\phi_a(z)dz = \int_{\mathbb{R}^n} \mathrm{R}(y - z_i, \boldsymbol{\beta}_{-i} - z_{-i}) \prod_{i \in [n]} \psi_a(z_i)dz$, where $\psi_a(z_i)$ is given in (7). Note that the revenue of the BSP auction only depends on the boost of the bidders with the highest and second highest boosted

[18] When there is only one bidder in an auction, assigning boosts does not change the allocation and payment of the auction.

[19] When there are only two bidders in set S^t, $t \in \tau_i$, then we set s_{-i}^t, to zero.

[20] Let τ_i^h be the set of auctions in τ_i in which buyer i has the highest boosted bid and his payment is governed by the the second highest boosted bid; that is, $\tau_i^h = \{t : t \in \tau_i, b_i^t \geq \frac{h_{-i}^t}{y} \geq r_i\}$. Similarly, let τ_i^s be the set of auctions in τ_i in which buyer i has the second highest boosted bid and payment of the winner is governed by the the second highest boosted bid; that is, $\tau_i^s = \{t : t \in \tau_i, h_{-i}^t > b_i^t y \geq \max(s_{-i}^t, \beta_{-i}^* r_{-i}^*)\}$. Then, $A = \sum_{t \in \tau_i^h} h_{-i}^t$ and $B = \sum_{t \in \tau_i^s} \frac{b_i^t}{\beta_{-i}^*}$.

bids. Thus, to compute the above integral, one can partition the space of $(y, z_1, z_2, \ldots, z_n)$ such that for all the points in a partition, the bidders with the highest and second highest boosted bids in any auction $t \in [T]$ are the same. Observe that the boundaries of each partition can be described by a linear function of $y, z_1, z_2, \ldots, z_n$. This is so because the boosted bids are linear in boosts. Further note that the revenue function R in each of these partitions is equal to the maximum of the second highest boost bid and reserve of the winner divided by the boost of the winner. This and that $\psi_a(\cdot)$ is polynomial of degree two imply $\mathrm{R}_{a,\phi}(y, \boldsymbol{\beta}_{-i})$ can be written as the sum of the polynomial · log(polynomial) and polynomials. Therefore, for any given $\boldsymbol{\beta}_{-i}$, function $\mathrm{R}_{a,\phi}(y, \boldsymbol{\beta}_{-i})$ is analytic in y.

So far, we have showed that the smooth revenue function is analytic. Next we show that the smooth revenue function satisfies the KL inequality with $\theta = \frac{1}{2}$. To do so, we will verify that at a critical point of the smooth revenue function, the Hessian matrix is non-zero. Then, by [28, 29], we can conclude that $\theta = \frac{1}{2}$.

By the proof of Theorem 3.2, $\mathrm{R}(\boldsymbol{\beta})$ can be written as $-\frac{A(\boldsymbol{\beta}_{-i})}{y} - B(\boldsymbol{\beta}_{-i})y - C(\boldsymbol{\beta}_{-i})$ where A, B, and C are positive and only depend on $\boldsymbol{\beta}_{-i}$. This leads to

$$\frac{\partial \mathrm{R}_{a,\phi}(\boldsymbol{\beta})}{\partial \beta_i} = \int_{\mathbb{R}^n} \left(\frac{A(\boldsymbol{\beta}_{-i} - \boldsymbol{z}_{-i})}{(\beta_i - z_i)^2} - B(\boldsymbol{\beta}_{-i} - \boldsymbol{z}_{-i}) \right) d\boldsymbol{z}\,, \qquad (11)$$

where the second equation holds because of the dominated convergence theorem. [21] This implies that $\frac{\partial^2 \mathrm{R}_{a,\phi}(\boldsymbol{\beta})}{\partial^2 \beta_i} = \int_{\mathbb{R}^n} -\frac{A(\boldsymbol{\beta}_{-i} - \boldsymbol{z}_{-i})}{(\beta_i - z_i)^3} d\boldsymbol{z} < 0$. The above equation shows that at a critical point, the Hessian matrix is not zero.

B ROBUSTNESS TO STRATEGIC BEHAVIORS

So far, we have shown that our data-driven approach that learns/tunes the parameters of the BSP auction using historical data performs well. Generally, learning from historical data has proven to be an effective tool to optimize the parameters of other types of auctions, including SP auctions [6, 14, 35, 36]. However, this technique can have a drawback: it can incentivize the bidders to change their bidding behavior to game the learning algorithm of the seller. For instance, in SP auctions, the bidder may have an incentive to shade his bids in order to enjoy lower reserve prices in the future. In fact, the problem of mitigating the negative impact of strategic behavior of bidders even in the SP auction is still an open problem. Empirical work that studied reserve price optimization in the SP auction has not considered this problem [35, 36].

Further, only recently, several theoretical papers have made some progress in this direction under certain assumptions; see, for example, [1, 2, 19, 24, 30]. It would be interesting to study how strategic bidders change their behavior to game the learning algorithm of BSP auctions and how this behavior impacts the seller. Since the BSP auction favors bidders with more stable bidding behavior, the strategic bidders may have an incentive to stabilize their bids.

Nonetheless, here, we show that the BSP auctions continue to outperform the SP auctions when bidders act strategically. Precisely, we conduct two counterfactual analyses by considering the following two bidding strategies for the top bidders:

Shading: At any auction t, top bidder $i \in [4]$ submits a bid of γb_i^t, where $\gamma \in \{0.9, 0.8, 0.7\}$. [22]

Compressing: At any auction t, top bidder $i \in [4]$ submits a bid of $\gamma b_i^t + \eta$, where $\gamma \in \{0.9, 0.8, 0.7\}$ and η is chosen in such a way that the average bids remain the same. That is, $\sum_{t \in [T]} b_i^t = \sum_{t \in [T]} (\gamma b_i^t + \eta)$.

In particular, we assume that one of the top four bidders either shades or compresses his bids by a constant factor $\gamma \in \{0.9, 0.8, 0.7\}$ while the rest of the bidders submit the same bids as before. We refer to γ as the shading/compression factor. Table 4 shows the revenue gain of the BSP auction relative to the SP auction with monopoly prices. We observe that the BSP auction maintains its dominance over the SP auction; precisely, it still earns 5%-6% more revenue than the SP auction. This suggests that the BSP auction is more robust to the strategic bidding behavior of bidders. We further observe that the revenue gain of the BSP auction is higher under shading than under compressing. This is in line with our empirical finding finding that stable bidders are favored in BSP auctions.

Shading/Compression factor (γ)	Strategic bidder	Average Rev gain under shading (in%)	Average Rev gain & under compression (in%)
0.9	1	6.35 (0.05)	6.13 (0.05)
	2	6.36 (0.05)	6.10 (0.05)
	3	6.47 (0.05)	6.41 (0.06)
	4	6.10 (0.05)	5.62 (0.05)
0.8	1	6.31 (0.06)	6.01 (0.05)
	2	6.15 (0.05)	5.81 (0.05)
	3	6.58 (0.06)	6.39 (0.05)
	4	6.05 (0.05)	5.61 (0.05)
0.7	1	6.15 (0.05)	5.92 (0.05)
	2	6.13 (0.06)	5.56 (0.05)
	3	6.48 (0.05)	6.40 (0.05)
	4	5.66 (0.04)	4.83 (0.04)

Table 4: BSP auction versus SP auction when one of the top four bidders shades/compresses his bids by a factor of γ.

[21] Note that we can apply the dominated convergence theorem if $\beta_i - z_i$ in $\frac{A(\boldsymbol{\beta}_{-i} - \boldsymbol{z}_{-i})}{\beta_i - z_i}$ is bounded away from zero. In fact, $\beta_i - z_i$ is bounded away from zero because the term $\frac{A(\boldsymbol{\beta}_{-i} - \boldsymbol{z}_{-i})}{\beta_i - z_i}$ is the revenue in all the auctions in which there are at least two cleared bids and bidder i has the highest boosted bid where the highest boosted bid is always positive. This implies that the boost of the winner $\beta_i - z_i > 0$.

[22] In practice, to incentivize the bidders not to adopt this behavior, advertising exchanges throttle the bidders whose bid is very low [26, 32].