# A Comparative Analysis of a Modified Second-Price Auction

*Author:*
Oskar Stigland
Bachelor Thesis
Fall 2018

*Supervisors:*
Tommy Andersson
Morten Arngren
Vlad Sandulescu

Lund University
School of Economics and Management
Department of Economics

# Contents

# Chapter 1

# Introduction

Online display advertisement is increasingly bought and sold through a process known as *real-time bidding* (RTB), in which bidders, or *advertisers*, receive requests to bid on display advertisement slots for specific users as they log onto a website, often referred to as a *publisher*. The whole process of auctioning out the ad slots from the moment the user starts loading the website takes less than 100 milliseconds. First, the publisher hosts an auction through a so-called *advertisement exchange* (AdX), which sends out bid requests to a number of so-called *demand-side platforms* (DSP). The role of a DSP is to use algorithmic know-how to participate in RTB auctions on behalf of advertisers. After the DSP:s have received the bid requests, they submit bids to the AdX, and whoever posts the highest bid wins the auction and gets to display their advertisement on the website. This somewhat simplified example of the RTB ecosystem is illustrated in figure 1.1.

There are two main reasons why an advertiser enlists a DSP in order to buy online ad slots. First, the DSP has the statistical tools and data-management skills necessary to valuate different impressions. This is often done by different performance-related metrics, of which one of the most common is the so-called *click-through rate* (CTR), i.e. the probability that a certain user, given their demographic information and other characteristics, will click on a given online display advertisement. Secondly, a normal DSP can participate in *billions* of RTB auctions *per day*. There are thousands of websites selling ad slots every second. Hence, simply participating in auctions and finding valuable impressions is a task in its own right. All of these procedures are, of course, strictly algorithmic.

There are a number of interesting research questions within the RTB ecosystem; we will primarily be concerned with the AdX. In terms of economics, the most interesting research question is, perhaps, the question of how to design the auction with respect to revenue maximization, allocative efficiency, and stability. Today, the dominating auction mechanism in RTB is the Vickrey auction, i.e. a sealed-bid second-price auction. However, we will discuss recent suggestions for improving the outcome of RTB auctions by using a sort of modified Vickrey auction. The main reason why we want to discuss modified second-price auctions is that we now have access to enormous amounts of data on bidder behavior in recurrent auctions with very similar items. This means that we can actually get a good estimate of the value distributions for different bidders.

In the dataset analyzed in one of the articles featured in this thesis, hundreds of thousands of auctions are used. In this thesis, a dataset with millions of auctions will be used, with more or less the same advertisers. This is different from a traditional auction environment where we may have only a single auction. That is, RTB auctions may allow the auctioneer to increase revenue by understanding fundamental properties of the bidders' different behaviors and strategies, in a way that is not possible when there's only a single auction.

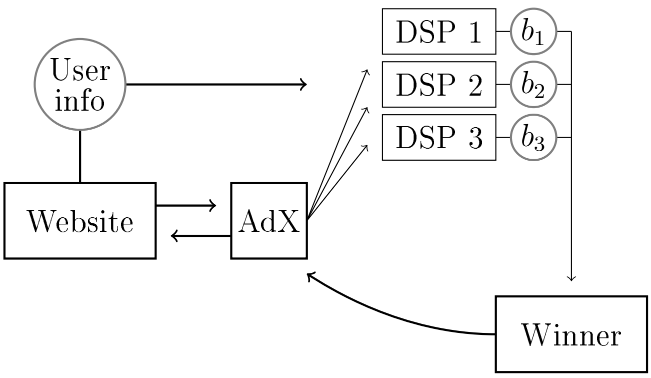



Figure 1.1: A simplified RTB ecosystem

# Chapter 2

# The second-price auction(s)

The so-called *Vickrey* auction, which is a sealed-bid second-price (SP) auction, is the dominating auction mechanism in real-time auctions for online display advertisement. One of the reasons for this, as well as for the general popularity of the Vickrey auction mechanism, is its nice theoretical properties, which incentivizes bidders to always bid their true valuations for whatever item is up for auction. This thesis is mainly concerned with a variant of the Vickrey auction, a so-called *modified second-bid* auction. In this thesis, we will use "second price", rather than "second bid", and hence refer to this mechanism as MSP.

In this chapter, we will start by discussing the theoretical properties of the Vickrey auction. Then, we will consider a general framework for MSP auctions and to what extent the theoretical properties of the original Vickrey auction are retained in an MSP auction. This framework will rely on the recent work by Arnosti, Beck and Milgrom (2016). We will also consider a specific MSP-type mechanism, called the *boosted second-price* (BSP) auction and discuss how it relates to the framework developed by Arnosti, Beck and Milgrom (2016). The BSP auction was introduced in 2017 by Golrezaei et al. The main focus of this thesis is to implement the BSP auction on a new dataset.

Finally, the chapter will also briefly consider the sealed-bid first-price (FP) auction and its historical role in online display advertisement, as well as why its revenue and allocative efficiency is not being compared along with the Vickrey auction and the BSP auction.

## 2.1 Vickrey auctions

Let's consider an auction where we have $n$ bidders, $i = 1, 2, \dots, n$, who are all submitting sealed bids for one item. The bidders' valuations, $v_i$, are independent of each other and drawn from continuous distributions, $v_i \sim V_i(\cdot)$, $i = 1, 2, \dots, n$. Each bidder has some bidding strategy, such that their bid is formulated as a function of their valuation, i.e. $b_i = B_i(v_i)$.

Assuming that all bids are ordered from largest to smallest, $b_1$ will be the winning bid with the winner paying $b_2$. Then, it is a *weakly dominant strategy* for each bidder to always bid their valuation, i.e. such that $B_i(v_i) = v_i$ for $i = 1, 2, \dots, n$. This means that the strategy earns the bidders a *payoff* at least as high as for any other strategy, regardless of what the other bidders do. The payoff is defined as the difference between the valuation, $v_i$, and the payment, $p_i$, i.e. $v_i - p_i$.

Why is it a weakly dominant strategy for any bidder, $i$, to bid their true valuation, $v_i$? Basically, bidders cannot affect their payoff positively by *not* bidding their true valuation. Let's consider bidder $i$ employing some bidding strategy, $B_i(\cdot)$, such that $B_i(v_i) > v_i$. We denote the highest bid for all of the other bidders by $b_*$, i.e. $\max_{j \neq i} b_j = b_*$. Then, there are three possible cases:

(i) $b_* > b_i, v_i$

(ii) $b_i > b_* > v_i$

(iii) $b_i, v_i > b_*$

In case (i), bidder $i$ will not win the item regardless. In case (ii), bidder $i$ will win the item, but at the cost of a negative payoff, since $b_* > v_i$. Finally, in case (iii), bidder $i$ will win the item with a positive payoff, since $b_* < v_i$. However, the payoff would be equally positive if $b_i = v_i$, since bidder $i$ would still win the auction and pay $b_*$. Hence, in all three cases the bidder would have been at least equally well off, with the notable exception of (ii), where the bidder would actually have been better off not bidding such that $b_i > v_i$. We can show that the same result holds for a bidder employing a bidding strategy such that $B_i(v_i) < v_i$ by following the same logic.

This is an important result to bear in mind since the data we will be using is taken from Vickrey auctions, meaning that the empirical distribution of bids for each advertiser, $i$, reveals their actual value distribution, $V_i(\cdot)$. However, it should be noted that it's difficult to accurately assess the value of an impression in RTB. Hence, even if the advertisers have no reason to bid anything else than what they perceive to be the value of the impression, it doesn't mean that this perception is correct. On the other hand, we're working with a large dataset, often with millions, and in some cases tens of millions, of historical bids for different advertisers. Consequently, we will assume that the historical bid distributions are an accurate estimation of the bidders individual value distributions.

## 2.2 Modified second-price auctions

Arnosti, Beck and Milgrom (2016) discuss the problem of an AdX where there are two types of advertisers, *performance advertisers* and *brand advertisers*, who have positively correlated valuations for impressions, but where performance advertisers can estimate the value of individual impressions accurately while brand advertisers can not. A performance advertiser could be, e.g., an online store looking to sell items to specific customer segments, in which case the advertiser has some information on the CTR, which is often, in practice, used as a proxy for the true value. CTR estimations are not accessible to brand advertisers, who are concerned with, e.g., advertising an event or spreading a campaign message.

This environment might lead to brand advertisers being exposed to adverse selection, i.e. winning disproportionately many low-value impressions, and thus leading to an inefficient allocation in the AdX. The authors suggest that the value of any impression can be characterized by a *common value* component and a *match value* component. The common value represents generally desirably qualities in a user, such as high income, high susceptibility to online display advertisement, and so on. The match value, on the other hand, might represent information that is specific to a certain advertiser, such as if the user recently visited the advertiser's website.

The authors present the MSP auctions as the group of auction mechanisms that overcome the disadvantages of using Vickrey auctions in RTB, i.e. being free of adverse selection, while still maintaining full strategy-proofness. The MSP auction is formally defined by the authors as the mechanism parameterized by $\alpha$, $\boldsymbol{z}$, and $\boldsymbol{p}$, where

- $\alpha$ is a constant, $\alpha \geq 1$,
- $z_i(b)$ is the probability of advertiser $i$ winning the auction given some bid, $b$ and
- $p_i(b)$ is the expected payment of advertiser $i$ given some bid, $b$,

such that for performance advertisers, $i = 1, 2, \ldots, n$, and a brand advertiser, $i = 0$, we have that

(i)

$$z_i(b) = P\left(b \geq \alpha \max_{j \neq i} b_j\right)$$

(ii)

$$p_i(b) = z_i(b) \times \alpha \mathbb{E}\left[\max_{j \neq i} b_j\right]$$

(iii)

$$z_0(b) = \prod_{i=1}^{n} P\left(b_i < \alpha \max_{j \neq i} b_j\right)$$

The first conditions say that the winner has to bid higher than and pay the second-highest bid multiplied by the constant $\alpha$. The third condition says that the brand advertiser only wins the auction when no performance advertiser wins the auction. This can happen if no performance advertiser bids in the auction, hence making it more or less an FP auction for the brand advertiser. It can also happen if all of the performance advertisers bid lower than some floor price, at which point the impression is awarded to the brand advertiser at the cost of the floor price. The authors describe this mechanism as being *deterministic*, *anonymous*, *fully strategy-proof*, and *adverse-selection free*.

The MSP auction mechanism is compared to the so-called *omniscient* (OMN) mechanism, which always achieves the optimal allocation by allocating the impression to the bidder with the highest match value, i.e. disregarding the common value component. In practice, it's obviously not feasible, or even possible, for an AdX - and even less so for an individual advertiser - to know the individual match value components for all advertisers. However, the authors argue that the MSP auction, as opposed to the Vickrey auction, comes very close to the upper bound of allocative efficiency posed by the OMN auction, by choosing some optimal $\alpha$. Further, they argue that the MSP auction outperforms the Vickrey auction in terms of revenue.

Next, we turn to the BSP auction, in which we will drop the anonymity of the general MSP auction and instead assign an individual multiplier to each advertiser. For discussion purposes and conceptual understanding, the distinction between brand and performance advertisers, as well as common value and match value, and the idea of an OMN auction as a benchmark will be retained throughout this thesis.

## 2.3 The boosted second-price auction

We consider the same auction environment as before, where we have $n$ bidders who all submit sealed bids for one item. All of the bidders have valuations drawn from continuous distributions and each has some bidding strategy. However, each bidder is also assigned an individual *boost value*, $\beta_i$, corresponding to an individual $\alpha$ in the general MSP auction. Similarly to the MSP auction, the winner of the auction is not the bidder with the highest bid, but the bidder with the highest bid multiplied by their individual boost value, i.e.

$$\text{winner} = \arg\max_i \beta_i b_i, \quad i = 1, 2, \dots, n$$

The payment of the winner is the second-highest boosted bid, scaled by the boost value of the winner. If the bidders are ordered by the size of their boosted bids, the payment is

$$p = \max\left(b_2, \frac{\beta_2 b_2}{\beta_1}\right)$$

However, bidders can never pay more than their initial bid, such that $p \leq b_1$. Hence, for any BSP auction, the revenue is

$$R_{\text{BSP}} = \min\left(b_1, \max\left(b_2, \frac{\beta_2 b_2}{\beta_1}\right)\right)$$

Golrezaei et al. (2017) also consider a set of reserve prices, $\boldsymbol{r}$, such that the auction is parametrized by $\boldsymbol{\beta}$ and $\boldsymbol{r}$, but we will set $\boldsymbol{r} = \mathbf{0}$ and hence only be concerned with $\boldsymbol{\beta}$. Let's consider the key distinctions with the general BSP presented in the previous section. We assume that we have $n$ performance advertisers and one brand advertiser. Then, the BSP auction is parametrized by $\boldsymbol{\beta}$, $\boldsymbol{z}$, and $\boldsymbol{p}$, such that

(i)

$$z_i(b) = P\left(\beta_i b \geq \max_{j \neq i} \beta_j b_j\right)$$

(ii)

$$p_i(b) = z_i(b) \times \mathbb{E}\left[\max_{j \neq i} \frac{\beta_j b_j}{\beta_i}\right]$$

(iii)

$$z_0(b) = P\left(\beta_0 b \geq \max_{j \neq 0} \beta_j b_j\right)$$

The authors also distinguish between brand advertisers and performance advertisers, calling the latter *retargeting advertisers* rather than performance advertisers. However, the BSP auction only differentiates between them by their boost values. Rather than choosing an $\alpha$ such that a proportionate amount of valuable impressions will be awarded to the brand advertiser by more or less making the performance advertisers forfeit the impression, the BSP auction can favor brand advertisers by assigning them higher boost values. This means that whenever a brand advertiser posts a relatively high bid, the AdX tries to capture the match value by assigning more weight to the brand advertiser's bid.

At this point, the natural question is how the boost values, $\boldsymbol{\beta}$, are actually calculated. In contrast to Arnosti, Beck and Milgrom (2016), who, at best, give a very vague description of how they imagine the $\alpha$ should be calculated, Golrezaei et al. (2017) give a more detailed account of how they calibrate $\boldsymbol{\beta}$. They employ a data-driven algorithm they call *BSP alternating minimizer* (`BSP-AM`). First of all, their description of `BSP-AM` is somewhat lacking in practical details. Secondly, they admit themselves that `BSP-AM` aims to solve an NP-complete, non-convex optimization problem, where the algorithm does converge to a coordinate maximum but without guaranteeing that this is in fact an optimal solution.

The approach is simple: iterate through a large number of randomized, historical auctions and compute the optimal boost value for each advertiser with respect to overall revenue in each auction, until the algorithm converges to a set of boost values. While this approach is intuitive and simple to implement, it can become complicated when working with a larger dataset. In the method section, I will discuss this further and introduce another algorithm which seems to be more efficient and which captures more interesting elements of the large amounts of data available from these auctions. My algorithm also seems to perform better with respect to revenue. However, it should be mentioned that I use a much larger dataset and a longer horizon with auctions for several different ad slots, calibrating the boost values over a whole month and testing for a whole week, while Golrezaei et al. (2017) only calibrate boost values for one day (and test them for the same day) with only one ad slot. On the other hand, given the nonstationarity of the RTB environment, this could be argued to be in their favor.

In the previous section, we noted that Arnosti, Beck and Milgrom (2016) argue that MSP auctions are deterministic, anonymous, fully strategy-proof and adverse-selection free. The BSP auction is straightforward in the first two characteristics: it's certainly deterministic and definitely not anonymous. The interesting question is how we can characterize the BSP auction in terms of strategy-proofness and adverse selection. Golrezeai et al. (2017) mention the paper by Arnosti, Beck and Milgrom (2016) briefly:

> "Arnosti et al. (2016) study the adverse selection in online ad markets for the impressions that are sold via auctions vs. guaranteed-delivery contracts, where the valuations of the buyers are correlated via a common value component. They show that to address the 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". **p.6**

In addition to failing to mention that their own approach is very much similar to that of Arnosti, Beck and Milgrom (2016), the authors vaguely dismiss the notion of adverse selection with the argument that it "does not appear" in their private-value setting. However, this is not in any way addressing the issue since the general MSP auction is also employed in a private-value setting. The point is not that the common value component is generally known, but that it represents a portion of any bidder's private valuation, such that it is reasonable to assume that the valuations of different bidders are positively correlated. This section is a disappointment in an otherwise interesting paper.

Intuitively, it seems reasonable that the BSP auction should also decrease adverse selection by assigning higher boost values to brand advertisers. Golrezaei et al. (2017) do not show or discuss what happens with the allocation of impressions in their data when the boost values are introduced. That is, while they report revenue increases, they do not show where this increased revenue actually comes from. In our dataset, we will see that unconstrained boosting may lead to significant re-allocation of impressions, in which brand advertisers do seem to be heavily favored. However, it's hard to make conclusive statements on adverse selection as in Arnosti, Beck and Milgrom (2016), since they suggest to explicitly pick an $\alpha$ such that the brand advertiser(s) does get a proportionate share of valuable impressions. The BSP auction, on the other hand, does not consider the allocative efficiency in such detail when calibrating boost values. Hence, while it certainly decreases adverse selection by the same logic as the general MSP decreases adverse selection, we can not say that it is adverse-selection free.

This leaves us with, perhaps, the most important characteristic: strategy-proofness. If the BSP auction is not strategy-proof and if truthfulness is not a weakly dominant strategy, our results are less legitimate as we would have to consider changes in bidding strategies when testing. This is also why the lack of discussion from Golrezaei et al. (2017) is somewhat disappointing; the results and the discussion suffer from lack of dimensionality since potentially severe effects of introducing a boosting mechanism are not discussed properly. While Golrezaei et al (2017) formally propose at the outset that the BSP auction mechanism is truthful, they never attempt to prove this proposition and more or less dismiss it at the end of the final discussion, where they suggest that the BSP auction mechanism may incentivize bidders to change their behavior in an attempt to increase their payoff. However, it's not clear that this is the case; again, there's a lack of scope and dimensionality.

Going forward, lacking the capacity to complete a more rigorous theoretical exposition, we will follow the proposition of Golrezaei et al. (2017) and assume that the BSP auction is, in fact, a truthful auction mechanism. The authors show that there's a correlation between boost values and certain parameters describing bidder behavior. Specifically, they show a significant relationship between volatile bidding behavior and low boost values. In other words, the BSP auction favors stable bidders, such as brand advertisers, while giving less favor to performance advertisers who are more prone to volatile bidding behavior because of, e.g., retargeting advertisement, more accurate estimation of CTR, and so on. They also show how this relates so the optimal mechanism presented by Myerson (1981).

These results do seem to lend credence to the proposition of truthfulness. That is, we're not considering truthfulness with respect to a single bid, but with respect to entire bidding strategies. If the boost values are directly related to the long-term behavior of bidders, broadly characterized by brand advertisers and performance advertisers, any bidder would have to change their long-term behavior in order to affect their boost value, which seems like an unlikely scenario. However, it does seem likely that a bidder would change their short-term participation rates in response to sudden changes in the allocation of impressions, e.g. because one advertiser might end up with more impressions than accounted for in the short-term budget. These are the types of issues that are not discussed by Golrezaei et al. (2017). Hence, we will attempt to improve upon this by making a distinction between *constrained* and *unconstrained* boosting. The former will be simulated by factoring in more realistic, short-term behavioral responses from the advertisers with respect to, e.g., budgets and winning rates.

In conclusion, we consider the BSP to be a non-anonymous, deterministic, and truthful form of MSP. It does seem, both from the results reported by Golrezaei et al. (2017) and from the results presented later in this paper, that the BSP mechanism impedes the occurrence of adverse selection in the AdX. However, we cannot state that it's adverse-selection free. The paper by Golrezaei et al. (2017) has a number of drawbacks, specifically with respect to the lack of dimensionality and scope when discussing the actual outcome of implementing the BSP. They report the average revenue gains from a number of simulations, which will be reported later in the results section, but they fail to show exactly where this revenue comes from, i.e. what happens with the allocation of impressions. We will attempt to improve upon this by being more transparent about the resulting allocation, as well as by trying to account for plausible behavioral responses.

## Example

Let's consider a simple example. We have three bidders, of which two, 1 and 2, are performance advertisers and one, 3, is a brand advertiser. They all post bids for an impression, such that

$$b_1 = 6, \quad b_2 = 4, \quad b_3 = 2.5$$

In a normal Vickrey auction, these would be truthful bids and advertiser 1 would win the auction, paying $b_2$, such that $R_{\text{SP}} = 4$. Now, let's consider a BSP setting with boost values $\beta_1 = 1$, $\beta_2 = 1$, and $\beta_3 = 2$. The winner is again advertiser 1, since

$$\max_i \beta_i b_i = \max\{1 \times 6, 1 \times 4, 2 \times 2.5\} = 6$$

and the payment is

$$R_{\text{BSP}} = \min\left(6, \frac{2 \times 2.5}{1}\right) = 5$$

which means that $R_{\text{SP}} < R_{\text{BSP}}$. This is, of course, a silly, albeit not entirely unrealistic, example. However, it can be useful for illustrating another important point. Let's instead assume that $\beta_3 = 3$. Then, the winner of the auction is bidder 3, since $\beta_3 b_3 = 7.5$ and the payment is

$$R_{\text{BSP}} = \min\left(2.5, \frac{1 \times 6}{3}\right) = 2$$

which means that $R_{\text{SP}} > R_{\text{BSP}}$. Hence, the BSP generates less revenue that the Vickrey auction. However, while we might have less revenue, it might also mean a better allocation. More interestingly, it illustrates how a more "optimal" allocation is actually a spill-over effect from trying to increase revenues. As we're attempting to close the gap between the highest bids, we will necessarily assign higher boost values to lower bidders, e.g. brand advertisers, meaning that they will often win impressions for which they post relatively high bids. Essentially, this is exactly the function of the anonymous MSP, where an $\alpha$ is chosen such that the gap between the highest bidders is narrowed, while simultaneously leading to the brand advertiser getting a larger portion of valuable impressions.

## 2.4 The shift from first price to second price

Edelman, Ostrovsky and Schwarz (2005) detail the institutional history of markets for online display advertisement. The FP auction had a brief stint of popularity from 1997 to 2002 due to its ease of use, low entry costs and transparency. However, both publishers and advertisers realized that it was prone to instabilities. The FP auction encouraged excessive gaming and bid shading since the bidder that was fast in reacting to other bidders' bidding strategies had an advantage. This also caused allocative inefficiencies. To remedy these problems, Google was the first publisher to turn to the Vickrey auction, which, as we know, is a truthful mechanism. Since then, the Vickrey auction has been the dominant auction format in RTB.

Golrezaei et al. (2016) also comment on the historical lack of popularity for first-price auctions in RTB. They focus on the heterogeneity of bidders in RTB and how this leads to bid shading in FP auctions, while it's hard for the AdX to understand the bidders' respective behaviors and strategies. Thus, instabilities and insecurities have led to the Vickrey auction dominating RTB for almost two decades. These are also the reasons why we will not be focusing on the FP auction when considering revenue and allocation of the Vickrey auction and the BSP auction. The highest bid in a Vickrey auction doesn't really say anything about the revenue from an FP auction with the same bidders, since we have to consider the interplay of bidding strategies and bid shading that would've occurred in an FP auction.

While it would be interesting to consider how an FP auction would play out in terms of revenue and allocation given the data we have, it is simply too demanding and outside the scope of this thesis to attempt such simulations. Even if we know all of the bidders' true valuations, it's difficult, if not impossible, to accurately simulate how a number of intelligent bidders would've behaved to maximize their respective payoffs simultaneously.

## 2.5 Summary

We have considered how both revenue and allocative efficiency can be improved in an RTB setting where there are two types of advertisers, brand advertisers and performance advertisers, who have positively correlated valuations of impressions. We have looked at general MSP auctions, which can potentially increase both revenue and allocative efficiency while still retaining the nice theoretical properties of the Vickrey auction, especially with respect to bidder behavior. We have also considered a specific application of MSP, the BSP auction, for which the retainment of certain properties is not clear. Both the general MSP and the BSP auction increase revenue and allocative efficiency by, directly or indirectly, favoring brand advertisers.

Going forward, we want to investigate the performance of the BSP auction in relation to the Vickrey auction. We will start with an unrealistic scenario in which we assume that all bidders want to win all impressions, i.e. essentially no budget limits, and that the auction is strategy-proof, like the MSP and the Vickrey auction. Then, we will try to conduct a more realistic test and factor in plausible behavioral changes as a response to the changes in impression allocations. We will also look at a number of interesting aspects of the overall problem, such as the bidders' value distributions and if the valuations (i.e. the bids) are positively correlated.

In conclusion, we want to explore other auction mechanisms than the Vickrey or FP auction, which seem to result in suboptimal outcomes in terms of revenue, allocation and stability, especially since RTB is a dynamic auction environment with heterogeneous bidders. MSP and BSP are attempts to amend the drawbacks of the traditional auction mechanisms by employing data-driven methods that aim to exploit the heterogeneity of bidders by understanding and adapting to the behavior and value distribution of each bidder.

# Chapter 3

# Method

As mentioned in the previous chapter, I haven't followed the approach by Golrezaei et al. (2017) when attempting to simulate the effects of introducing a BSP auction mechanism in an AdX. This is due to the fact that I have used a much larger dataset, spanning a longer time period and including more items (i.e. different ad slots). This not only means that the algorithm they proposed is more cumbersome to use due to runtime, but also that the data preprocessing necessary for their algorithm becomes more complicated. This chapter is broadly dedicated to describing this problem and how I've attempted to deal with it.

The first section will describe the data used and how it differs from the data used by Golrezaei et al. (2017). Then, I will describe some of the bidders in the dataset in terms of their participation rate, bid average, and bid variance, as well as their estimated value distributions. Finally, the last sections will be devoted to describing the method I've used to calibrate the boost values, as well as how I've attempted to calibrate boost values for different advertisers when incorporating a constraint on their budget spending.

## 3.1 Data

The data consists of auctions for ad slots held by Adform on behalf of a single website in Denmark. The website is one of the largest in Denmark and has approximately 2.1 million unique visits per week, i.e. something like 300 thousand unique visits per day on average. There are a large number of ad slots on the website. These are characterized by height, width, and position. Summary statistics across the training dataset for some of the most popular ad slots are displayed in table 3.1. The dataset is split up into six different periods, of which the first four are used for calibrating the boost values while the last period is used for testing the revenue when using the boost values. There is a total of 116 million bids in the raw training data. When I have filtered out all auctions with only one bidder and with multiple ad slots, there are 8 million bids in 2.9 million auctions. When the testing data has been filtered for single-bid auctions and auctions with several ad slots, there are 634201 auctions left with 2 million bids. Summary statistics for the different periods are displayed in table 3.2.

Many of the auctions in the dataset only have one bidder. I have filtered out all of these auctions from the testing dataset, since applying boost values doesn't make any difference in terms of revenue when there's only one bidder. There are also many cases of the same bidder posting multiple bids. This is likely due to the fact that DSP:s often run multiple advertisement campaigns at the same time, posting bids on behalf of several different advertisers depending on their specific budgets, targeting criteria, and so on. In the testing dataset, I've only considered the highest bid from an advertiser that has posted several bids.

Table 3.1: **SOMETHING LIKE THIS...**

| Slot ID | Width | Height | Position ID | Number of bids | Bid mean | Bid var |
|---|---|---|---|---|---|---|
| 177545 | ... | ... | . | 5033814 | ... | ... |
| 177547 | ... | ... | . | 4896650 | ... | ... |
| 177546 | ... | ... | . | 4839611 | ... | ... |
| 177544 | ... | ... | . | 4498701 | ... | ... |
| 177885 | ... | ... | . | 3260809 | ... | ... |
| 339967 | 320 | 320 | 2 | 2984862 | ... | ... |
| 494451 | 320 | 160 | 2 | 2566887 | ... | ... |
| 177550 | ... | ... | . | 2485026 | ... | ... |

Table 3.2: Data periods

| Period | Type | Start | End | Number of bids | Number of auctions |
|---|---|---|---|---|---|
| 1 | Training | 2018-10-10 | 2018-10-17 | 25879377 | 10025222 |
| 2 | Training | 2018-10-17 | 2018-10-24 | 22878984 | 9117532 |
| 3 | Training | 2018-10-25 | 2018-11-01 | 27834936 | 10240755 |
| 4 | Training | 2018-11-02 | 2018-11-09 | 15843723 | 4533186 |
| 5 | Training | 2018-11-12 | 2018-11-19 | 23985352 | 6355267 |
| 6 | Testing | 2018-11-27 | 2018-12-04 | 24439817 | 7852427 |

The main reason why the approach suggested by Golrezaei et al. (2017), i.e. iterating through a number of historical auctions until the boost values converge, is that the large amount of data makes it difficult to randomize and shuffle the training data. Ideally, we want to shuffle the training data since RTB is a highly nonstationary environment, meaning that the "market price" of an ad slot can change drastically during a week, or even during a day. Hence, if we calibrate boost values on a chronologically ordered set of auctions, the boost values may be fitted specifically to the last share of auctions, and may not be representative of the earlier auctions. This also highlights the major differences between the dataset used by Golrezaei et al. (2017) and the one I have used. First of all, I'm training and testing over much longer periods (several days, rather than just one day). Secondly, I'm considering bids for an number of different ad slots rather than for just one ad slot. As shown in table 3.1, the ad slots are in fact distinguishable. Hence, there is less granularity and more generality in my approach.

Obviously, I could focus on just one day and come closer to replicating Golrezaei et al. (2017). However, there is another important feature to consider which separates the datasets. They use data from Google's AdX, which is the largest in the world. Hence, it has a large number of frequent bidders with high participation rates. Adform does not have as many auctions with as many bidders during one day. This is not necessarily a problem, but conducting a thorough analysis of the seasonality in the participation of different bidders is definitely a problem in terms of time and effort (for a Bachelor's thesis). Thus, I've chosen to use the entire dataset from Adform and create more general estimates of the boost values. I will show that this also leads to converging boost values and increased revenue, by the same evaluation method as Golrezaei et al. (2017).

## 3.2 Bidders

I have followed the approach of Golrezaei et al. (2017) in focusing on the largest bidders. This is sensible, since applying boost values to smaller bidders will have a negligible effect on the revenue. Some bidders participate in less than 0.1 % of the auctions. Essentially, among the 8 bidders with the highest spending, of which the three smallest have a participation rate of approximately 0.1 %, there is one brand advertiser and seven performance advertisers. This section will provide statistics on these bidders, including estimations of their value distributions. Summary statistics on the different bidders are presented in table 3.3.

There are some interesting points to be made from looking at the bidders. It's clear that bidder 2 is a typical brand advertiser, with a very high participation rate and relatively low bid mean and variance. When considering all auctions (i.e. including all the auctions with only one bidder), bidder 2 has an even higher participation rate relative to the other bidders. On the other end of the spectrum, bidders 6 and 7 seem to be typical retargeting advertisers, with extremely low participation rates but considerable spending due to a high bid mean and (especially for bidder 7) a very high bid variance. Obviously, this structure of bidders' characteristics will determine the possible revenue increase. For example, if we only would've had brand advertisers, with low bid variances and similar bid means, the possible revenue increase would be limited by the similarity of bidding strategies. In this sense, my results are not experiments are not entirely comparable to those by Golrezaei et al. (2017). That is, even if I replicated their algorithm completely and limited the experiment to only one ad slot during one day, the results would still not be comparable since there is a fundamental difference in the bidders participating in the auctions. While Golrezaei et al. (2017) are not as transparent about their bidders, it seems like they have a more "complete" spectrum of bidders, including more brand advertisers.

Table 3.3: The eight largest bidders, ordered by number of bids

| Bidder | Type | Participation rate | Number of bids | Bid mean | Bid var |
|---|---|---|---|---|---|
| 1 | Performance | 0.71 | 4017328 | 1.89 | 2.10 |
| 2 | Brand | 0.90 | 2594512 | 0.60 | 1.06 |
| 3 | Performance | 0.22 | 631380 | 1.35 | 1.92 |
| 4 | Performance | 0.11 | 318965 | 1.66 | 0.91 |
| 5 | Performance | 0.08 | 229228 | 1.40 | 1.21 |
| 6 | Performance | 0.04 | 124821 | 2.22 | 4.61 |
| 7 | Performance | 0.02 | 63483 | 4.45 | 85.11 |
| 8 | Performance | 0.02 | 52920 | 2.07 | 0.57 |

The kernel density estimates in figure 3.1-3.8 are purely for illustrative purposes. I have used a less granular bandwidth to give them a smoother appearance. The estimated cumulative distribution functions (CDFs) are, however, used in calibrating the boost values. However, when plotting the inverse CDFs I have used fewer bins in the underlying histograms. Hence, the figures should be understood primarily as a tool to illustrate the behavior and strategies of the different bidders.

Figure 3.1: Bidder 1

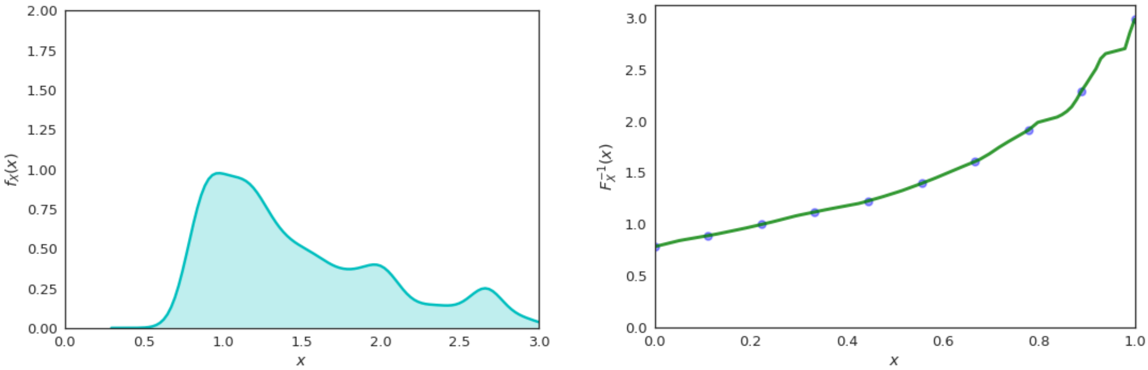



Figure 3.2: Bidder 2

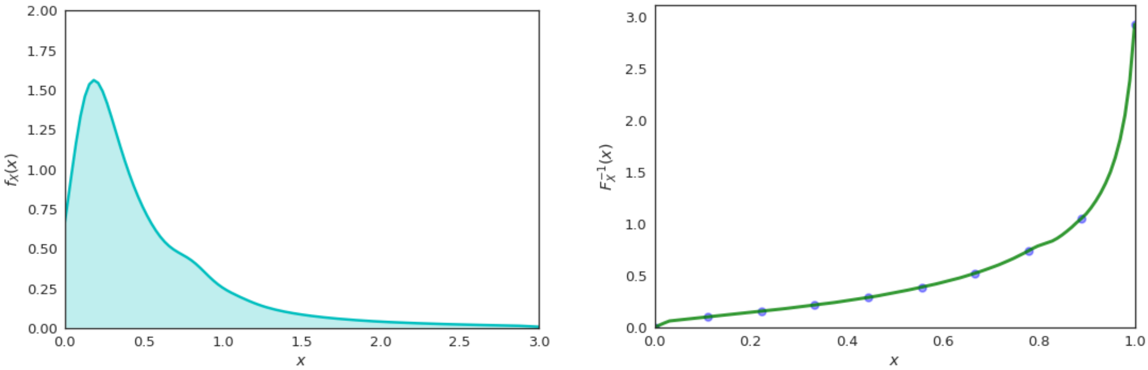



Figure 3.3: Bidder 3

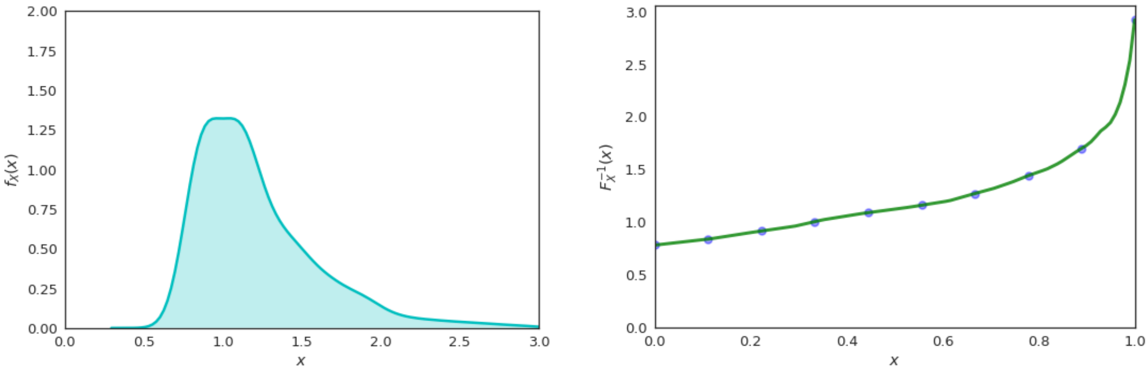


Figure 3.4: Bidder 4

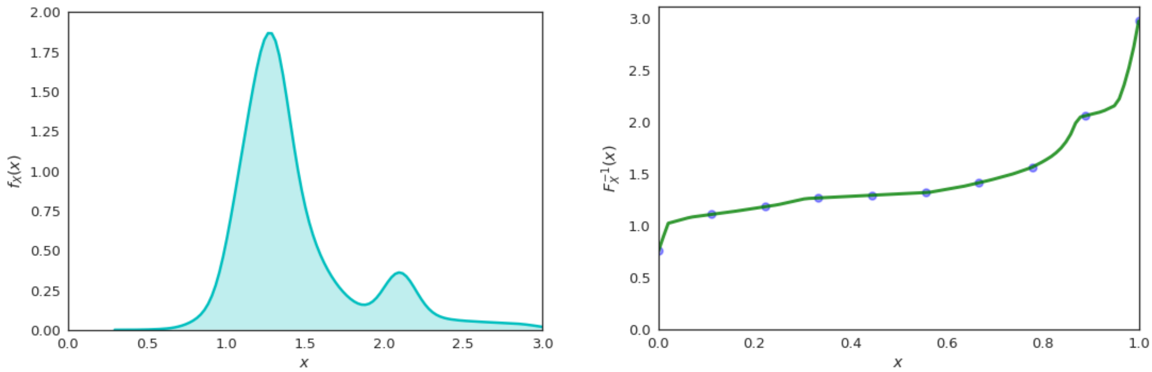



Figure 3.5: Bidder 5

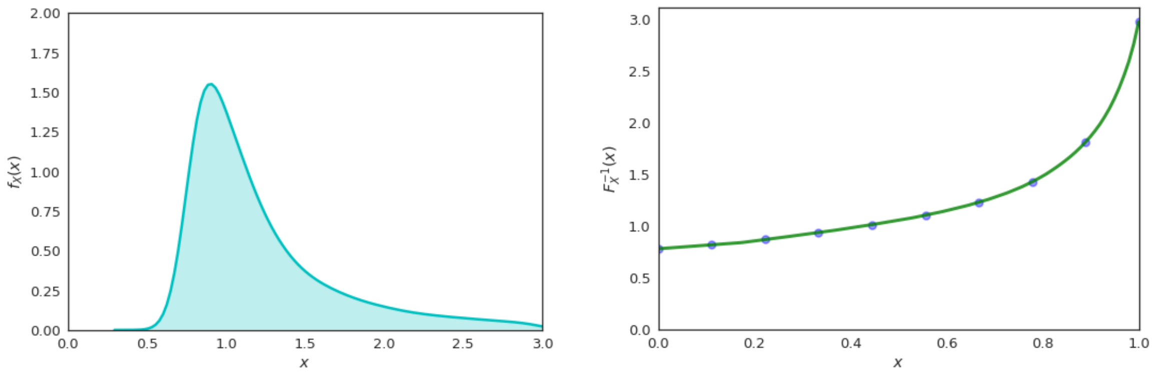
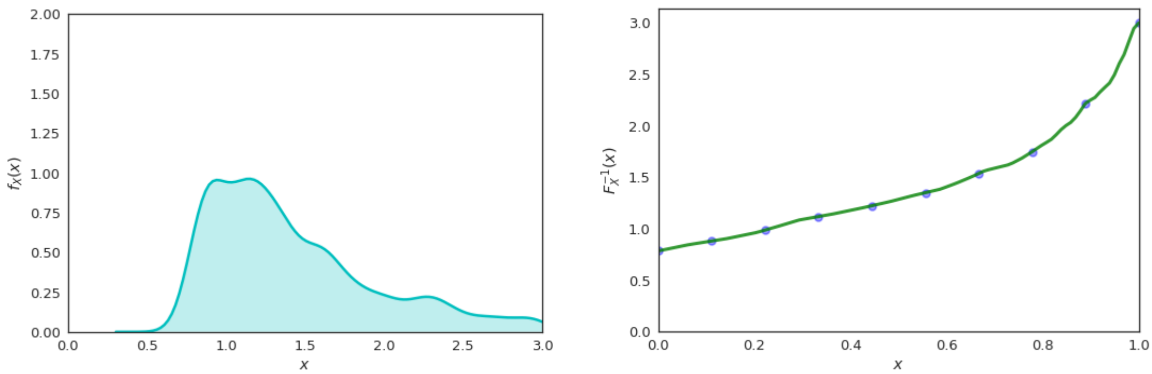



Figure 3.6: Bidder 6



## 3.3 Boosting by sampling

As described in the previous sections, I've been working with dataset which is much larger than the one used by Golarezaei et al. (2017), spanning a longer time period and several different ad slots. Hence, a more general and dynamic approach is required to compute boost values which are representative of bidding behavior across all auctions. The large amounts of data also make it difficult to shuffle and iterate through the historical data. I've solved this by instead creating a stylized, randomized auction which is iterated a large number of times.

First, I extract all of the bidding data for each included advertiser. Then, I estimate all of the participation rates and bid distributions. This makes for a representation of any advertiser's "average" bidding behavior across all historical auctions. Hence, we can consider a simulated, "average" auction, more or less representative of all of the historical data. For each iteration, I sample from a uniform distribution for each advertiser to determine if they will participate in the auction. Then, if more than two advertisers are participating, I take a random sample from each of their bid distributions and hold a BSP auction. Finally, the same optimization procedure as in Golrezeai et al. (2017) is followed, updating each boost value incrementally with respect to maximizing the revenue from the auction. We will see in the results section that the boost values converge relatively fast. The whole procedure is described in more detail in the following subsections, while pseudocode is presented in the two following sections, 3.4 and 3.5, which also discuss the problem of calibrating boost values without any constraints.

### 3.3.1 Estimating bid distributions

First, we need to create an *empirical distribution function* for each bidder. Let's consider bidder $i$. If we define $\boldsymbol{b}_i = (b_i^{(1)}, b_i^{(2)}, \ldots, b_i^{(n)})$ to be all bids by bidder $i$, we define the empirical *cumulative distribution function* (CDF) as

$$\bar{F}_i(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{b_i^{(j)} \leq x}$$

such that $\bar{F}_i : \mathbb{R} \to [0, 1]$. We create a more or less granular histogram and calculate the probability of some value $x$ by considering the cumulative value of all the bins up to and including the one containing $x$. Specifically, I've used 500 bins in the range $[0, 5]$. This excludes some extreme bids for some of the performance advertisers. However, the algorithm converges quickly as a result, and it can be argued that there is no great loss with respect to revenue in using this restriction since the bidders with some extremely high bids also have very low participation rates.

### 3.3.2 Participation sampling

For each advertiser, we have a historical participation rate. Again, let's consider bidder $i$. We denote this participation rate as $\rho_i$. Then, in each auction we sample from the uniform distribution between 0 and 1, such that we get $u \sim \text{Uniform}(0, 1)$. If $u \leq \rho_i$, bidder $i$ will participate in the auction, and if $u > \rho_i$, bidder $i$ will not participate. There are cases where is only one bidder or no bidders. In this case, there is no auction and hence no change to the boost values.

### 3.3.3 Bidding by inverse transform sampling

Whenever we have two or more bidders participating, we use so-called *inverse transform sampling* to get the bid from each bidder. This means that we take the *inverse* of the aforementioned empirical CDF, i.e. $\bar{F}_i^{-1}(x)$. I've done this by interpolating a curve over the cumulative values and the bin edges, i.e. for $\{(\bar{F}_i(x_j), x_j) \mid j = 1, 2, \ldots, n\}$. Thus, we get a continuous approximation of the inverse CDF for each bidder. Then, we sample $u$ from the same uniform distribution as before, i.e. $u \sim \text{Uniform}(0, 1)$. Since $\bar{F}_i^{-1} : [0, 1] \to \mathbb{R}$, this means that we can get random samples from bidder $i$, $b_i$, by $b_i = \bar{F}_i^{-1}(u)$.

### 3.3.4 Boost calibration

For each advertiser, we update the boost values by a *gradient descent method.* We consider the optimal boost value for bidder $i$, $\beta_i$, with respect to maximizing the revenue in a given auction. We have two possibilities: either $i$ is the winner of the auction or $i$ is not the winner of the auction. In any case, bidder $i$ will never pay more than the given bid, i.e. $b_i$. Let's consider another bidder, $j$. We have case (i), in which $i$ is the winner of the auction and $j$ is the second-highest (boosted) bidder, and case (ii), in which $j$ is the winner of the auction.

(i) If $i$ is the winner of the auction, we want a low $\beta_i$, since the payment is scaled in inverse proportion to $\beta_i$. In any case, $R_{\max} = b_i$, such that the optimal beta, $\beta_i^*$, for that particular auction can be found by

$$R_{\max} = \frac{\beta_j b_j}{\beta_i^*} \iff b_i = \frac{\beta_j b_j}{\beta_i^*} \iff \beta_i^* = \frac{\beta_j b_j}{b_i}$$

(ii) If $j$ is the winner of the auction, we want a high $\beta_i$, since the payment is now scaled in proportion to $\beta_i$. Now, we have that $R_{\max} = b_j$, such that the optimal beta $\beta_i^*$, for this particular auction is given by

$$R_{\max} = \frac{\beta_i^* b_i}{\beta_j} \iff b_j = \frac{\beta_i^* b_i}{\beta_j} \iff \beta_i^* = \frac{\beta_j b_j}{b_i}$$

Then, we update $\beta_i$ by

$$\beta_i^{(k+1)} = \beta_i^{(k)} - \alpha \times \left(\beta_i^{(k)} - \beta_i^*\right) = \beta_i^{(k)} - \alpha \times \left(\beta_i^{(k)} - \frac{\beta_j^{(k)} b_j^{(k)}}{b_i^{(k)}}\right)$$

where $\alpha$ is the *learning rate* and where $j$ is either the highest or second-highest bidder, depending on whether $i$ won the auction. After each auction and subsequent calibration, all the boost values are divided by the smallest boost value, such that the smallest boost value is 1.0. In figure 3.7, I have plotted the tuning of the boost values using $\alpha = 1 \cdot 10^{-4}$ for the largest advertisers by spending and iterating 500000 times. It is worth noting that the result is consistent with Golrezaei et al. (2017), since they report that typical brand advertisers are heavily favored by the BSP auction. This is clearly the case here as well.

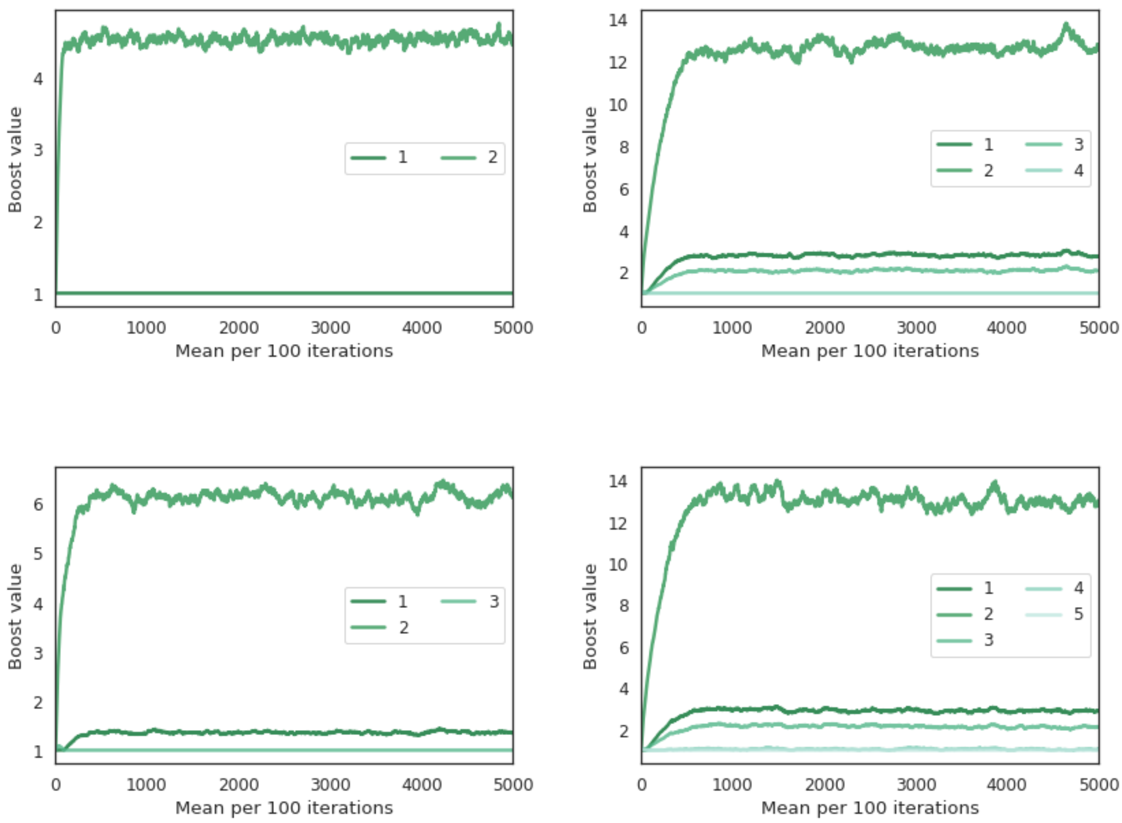



Figure 3.7: Calibration of boost values over 500000 iterations.

Rather than letting the boost values converge to some exact value, I let the simulated auctions iterate for a fixed number of times. In order to get the final boost values, I take the average of the last 400000 iterations, i.e. after the boost values settle in some interval. The reason the boost values do not converge to exact values is probably due to the nonstationarity of the auction environment and the variations in bidding behavior, including the differences in participations rates. For example, a performance advertiser that has a typical retargeting behavior will not participate very often, but will likely have a large impact on the boost values when it does participate. As we're increasing the number of bidders, the calibration of the boost values gets increasingly unstable. One way of decreasing instability is to have a smaller $\alpha$, but this will also mean longer times for calibration. For example, in the case of 4 advertisers with $\alpha = 1 \cdot 10^{-5}$, it takes something like 1.5 million iterations before the boost values settles in some interval. Hence, we will be using $\alpha = 1 \cdot 10^{-4}$.

## 3.4 Unconstrained boosting

Since the algorithm is inspired by Monte Carlo methods and is relying heavily on random sampling, I've chosen to call the procedure `BSP-MC`. The total runtime is approximately $XXX$ seconds. I've included the pseudocode in figure 3.8. to give a better overview of the whole procedure. However, this procedure (and, seemingly, the method employed by Golrezaei et al. (2017)), is not entirely unproblematic. As we will see in the next chapter, it's important to consider exactly what happens when we start applying boost values and where the additional revenue actually comes from. We saw a brief example of this in the theory section. When some boost value is sufficiently high, specifically for the brand advertiser, we might see changes in the distribution of impressions. While this is desirable to some extent, since we want a better allocative efficiency and less adverse selection for the brand advertiser, we might run into problems.

Currently, we assume that all bidders want to win all impressions that they bid for and that they do not have any budget constraints. This is, of course, an unrealistic scenario. For example, if the boost value for the brand advertiser becomes large enough, the brand advertiser could potentially swamp up almost all impressions. Hence, we would like to calibrate boost values with some constraints on the possible changes in how the impressions are allocated. This should, of course, hamper the theoretical revenue increase, but it also makes for a more realistic simulation. A method for constrained boosting is suggested in the next section.

**Algorithm** - `BSP-MC`

```
Bla bla
Bla bla bla
for k = 1 to K do
    Bla bla
    if x = 1 then
        Bla bla
    else
        Bla bla
    end if
    Bla bla
    Bla bla
end for
Bla bla
```

Figure 3.8: Pseudocode...

## 3.5 Constrained boosting

- How?
- How?
- How?

**Algorithm - Constrained** `BSP-MC`

Bla bla
Bla bla bla
**for** $k = 1$ to $K$ **do**
    Bla bla
    **if** $x = 1$ **then**
        Bla bla
    **else**
        Bla bla
    **end if**
    Bla bla
    Bla bla
**end for**
Bla bla

Figure 3.9: Pseudocode...

# Chapter 4

# Results

In this chapter, we want to address the discussions from the theory chapter using the proposed method. As mentioned previously, the results found by Golrezaei et al. (2017) and those presented here might be more or less incomparable. Hence, such comparisons will be held to a minimum. We will start by looking at the correlation between bids from different advertisers. Then, we will look at the possible revenue improvements from unconstrained boosting, as well as the resulting changes in the allocation of winning impressions. Finally, we will look at results from some cases of calibrating boost values while incorporating constraints.

## 4.1 Bid correlation

In proposing the MSP, Arnosti, Beck and Milgrom (2016), discuss the positive correlation of bids from brand and performance advertisers as the main reason for the supposed existence of adverse selection in an AdX. Hence, this seems like an interesting aspect to consider when proposing to implement an auction mechanism which clearly changes the allocation of impressions. I've used the testing dataset and compared the brand advertiser, bidder 2, to two performance advertisers, bidder 1 and bidder 3. I looked at all auctions in which both bidders participate. If a bidder has posted multiple bids, I have takes the average of all those bids. When comparing bidder 2 with bidder 1, there are 462 thousand auctions, but when comparing bidder 2 with bidder 3 there's only a thousand auctions.

In the first case, the correlation coefficient is $-0.05$ and in the second case the correlation coefficient is $-0.06$. Hence, there's very little correlation, and no positive correlation, in the valuations of impressions in the auctions where the bidders participate. One could argue that it's also interesting to consider auctions in which one bidder participates and the other doesn't, in which case the bid of the bidder not participating could be zero. This would of course affect the correlation and possibly make it even smaller (due to the high participation rate of bidder 2 relative to the other bidders). However, as clarified by the discussion by Arnosti, Beck and Milgrom (2016), what we're interested in is the possible existence of a common value component, which should be captured by the current analysis. In contrast to their discussion, the small correlation coefficients suggest that there is little evidence of a common value component.

This result is of course not representative of RTB in general, only of auctions held by Adform for a specific website in Denmark. Nevertheless, the results are interesting and suggest that the proposed existence of a common value component is not entirely uncontroversial. One could also argue that due to the inability of brand advertisers to estimate the value of different impressions, they are less willing to post relatively high bids. However, looking at table 3.3, bidder 2 has significant variance, meaning there is variation in the valuations of bidder 2. Hence, if a common value component did exist, any variations in the common value component should be captured by the bidding behavior and the correlation between the different bidders.

It's important to note that this result is underpinned by the truthfulness of the Vickrey auction. That is, attempting to capture the existence of a common value component only makes sense if the bids are wholly representative of the bidders' actual valuations. Given the truthfulness of the Vickrey auction, there doesn't seem to exist a significant common value component or a positive correlation between bids submitted by brand and performance advertisers.

## 4.2 Results without constraints

We want to consider the distribution of revenue increases for a number of simulations with unconstrained. I've run 100 simulations each when assigning boost values for 2, 3, 4, 5, and 6 bidders, which have been chosen randomly from the 8 top-spending bidders. In total, I ran 500 simulations. For each simulation, I've used 1000000 iterations and taken the average of the last 500000 iterations as the final boost values, with $\alpha = 1 \cdot 10^{-4}$. This is due to the fact that if we choose bidders randomly, we will sometimes end up only with bidders with low participation rates, meaning that we will have relatively few simulated auctions to calibrate boost values from. Each of the 500 sets of boost values are evaluated on the first 100000 auctions in the test data. The results are shown in figure 4.1 and table 4.1. The average revenue increase across all simulations is 84 %, while the maximum revenue increase is 220 %. As a comparison, Golrezaei et al. (2017) achieve revenue increases from 16.55 % to 29.28 %. However, it should be noted that even though the evaluation methods are comparable, the results are not necessarily comparable since the auction environments analyzed seem to be very different.

We get the largest average revenue increase, 118 % and 127 %, when assigning boost values to 4 and 5 bidders, respectively. The standard deviation for the results from assigning boost values to 4 bidders is slightly lower. We run a simulation with the 4 top-spending bidders, using $\alpha = 1 \cdot 10^{-4}$ for 500000 iterations, evaluating on the entire test dataset, in order to get a better understanding of where the revenue comes from. In figure 4.2, I've plotted the change in the distribution of payments (i.e. what the bidder with the winning bid actually pays) for the auctions in the test data, using the original Vickrey auction format and the BSP mechanism. There is a heavy shift to the right, which can be interpreted as the boost values "closing the gap" between the highest and second-highest bids.

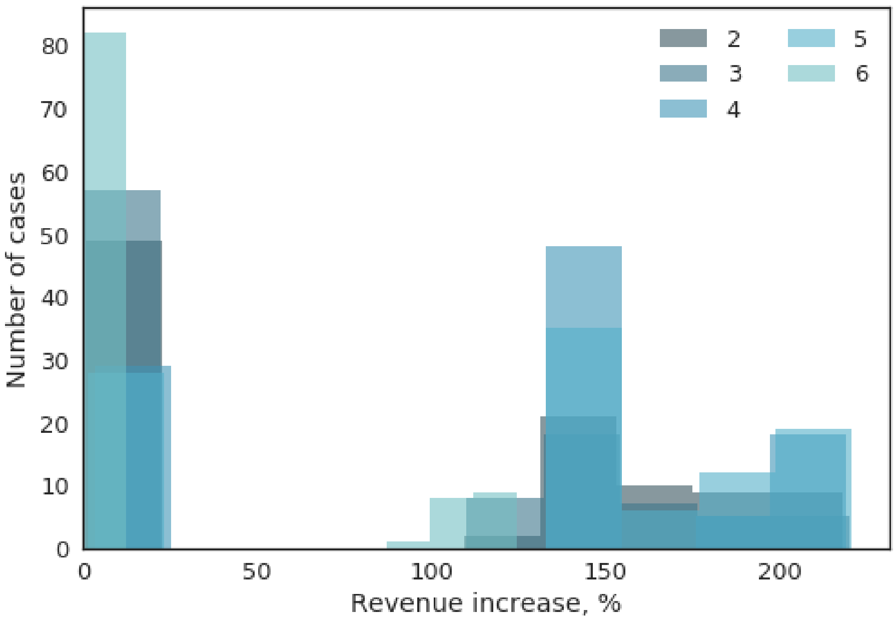



Figure 4.1: Histogram with revenue increases from unconstrained boosting

We also want to look at the changes in the allocation of impressions when implementing the boost values. As mentioned before, it's possible that these are significant since the boost values are calibrated without considering more or less realistic budget constraints on behalf of the bidders. That is, even if the auction is strategy proof, we should consider the fact that bidders can't be expected to be entirely flexible with respect to the amount of impressions they gain or loose as a result of using the boost values. In figure 4.3, I've plotted the changes in the allocation of impressions.

We're clearly dealing with an unrealistic scenario. The increase in revenue is 198 %, but the brand advertiser, 2, absorbs a lot more impressions, while bidders 1, 3 and 4 likely pays a lot more for fewer impressions. It seems unlikely that bidder 2 would maintain a maintain a high participation rate when implementing the boost values. The number of impressions won increases from 34939 to 297401, which clearly has serious budget implications. It's important to understand why this shift in allocation doesn't necessarily decrease the revenue. At first sight, it might seem like a bidder with a lower average bid winning more auctions would lead to a lower average revenue. However, the point is that in the normal Vickrey auction, bidder 2 often has the second-highest bid such that the highest-bidder will pay $b_2$, while bidder 2 will often be the winner in the BSP auction and pay its own bid, i.e. such that the revenue is still $b_2$, due to its high boost value.

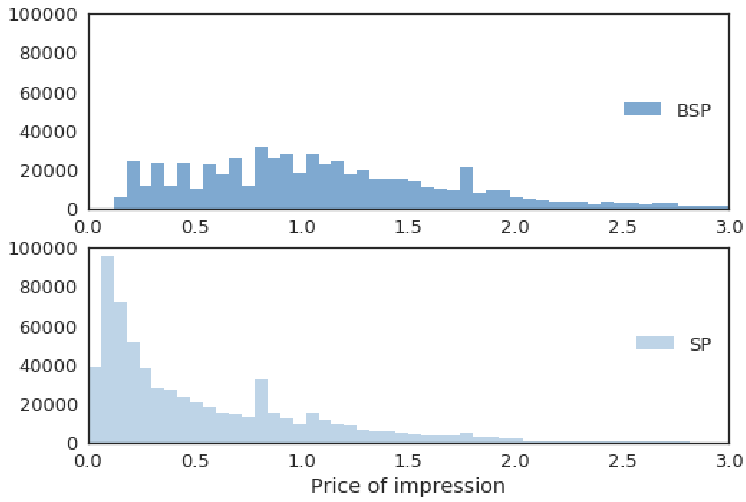



Figure 4.2: Change in price of impressions

While the revenue increase is impressive, the change in the allocation of impressions shows a more problematic aspect of the boosting approach. It would be interesting too see the change in allocation for Golrezaei et al. (2017). Unfortunately, this information is not available. However, given that they also reported significant increases in revenue, it's likely that their simulations also resulted in significant changes in allocation. At this point, it's not really a question of the auction being strategy-proof. Rather, it's a question of the bidders even being able to employ a strategy or even being willing to participate due to the significant changes in budget spending. Hence, we will be turning to the case of constrained boosting in the next section.

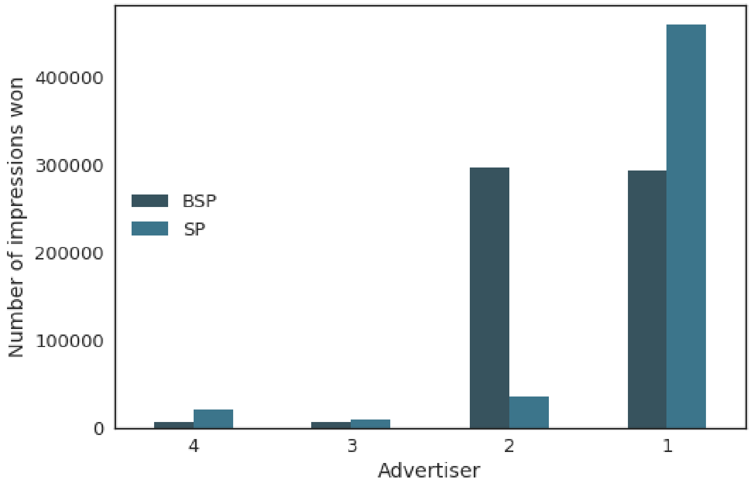



Figure 4.3: Change in allocation of impressions

## 4.3 Results with constraints

- Revenue with different levels of tolerance for exceeding expected budget

## 4.4 Summary

- Mention that results are heavily contingent on the "nature" of the auction, i.e. the disposition of brand advertisers, performance advertisers etc. Potential revenue increases are dependent on the possible improvement available.

# Chapter 5

# Discussion and conclusion

# References