Stochastic Variability in Sponsored Search Auctions: Observations and Models *

Furcy Pin†
École Normale Supérieure
Paris, France
furcy.pin@ens.fr

Peter Key
Microsoft Research
Cambridge, CB3 0FB, UK
peter.key@microsoft.com

April 12, 2011

Abstract

Sponsored search advertisement slots are currently sold via Generalized Second Price (GSP) auctions. Despite the simplicity of their rules, these auctions are far from being fully understood. Our observations on real ad-auction data show that advertisers usually enter many distinct auctions with different opponents and with varying parameters. We describe some of our findings from these observations and propose a simple probabilistic model taking them into account. This model can be used to predict the number of clicks received by the advertisers and the total price they can expect to pay depending on their bid, or even to estimate the players valuations, all at a very low computational cost.

1 Introduction

Each time someone types a query in a search engine such as Bing or Google, an auction is run and used to determine which sponsored links will appear and what price will be paid by the advertiser if their ad is clicked. The *Generalized Second Price (GSP)* auction is the current mechanism used. Although these auctions are very simple to describe and have been in use since at least 2002, their behavior in practice is far from being fully understood. In this paper we develop a new analytical model for such a position auction, motivated and informed by measurements from actual ad-auctions, which attempts to account for some of the real-world variability in such auctions.

In a simplified version of a GSP, each *advertiser* (or *player*) places a bid for a keyword. The *auctioneer* (namely the ad-platform) then displays the ads with the highest bids, allocating the highest bidder to the best slot, and so on, until all the allowed slots are allocated. Payment is then organized on a Pay-Per-Click (PPC) basis: an advertiser only pays when her ad is clicked, and pays the price equal to the bid of the player just below her.

A prerequisite to improving auction design is the ability to predict accurately the behavior of bidders. In particular, one challenge is to estimate a player's *valuation* knowing her bid, and conversely to predict the bids when the valuations are given.

We describe a new analytical model, which allows us to predict a player's valuation given her bid and her opponents' bids. This model relies on the observation that the players participate with the same bid in a great number of auctions with varying conditions. For

*This is a full version of the paper in 12th ACM Conference on Electronic Commerce (EC2011). The authors wish to thank Anton Schwaighofer and Alexandre Proutiere for their help and comments on this paper.

†Work performed during the author's internship at Microsoft Research Cambridge.

instance, we noticed that the number and the identity of the opponents of a given player usually vary across successive auctions. Moreover, the players' bids are in practice multiplied by weights which depend on many criteria, and vary from one auction to another. We also observed that most players do not change their bid often.

For these reasons, the classical model of a single auction repeated an infinite number of times is inappropriate. Because of the stochastic noise in our model, in general players have a unique optimal response to a given opponent's bidding profile, whereas in the classical model where only one single auction is considered (or one auction repeated with the same weights and players), there typically is a complete interval of optimal bids.

Contributions. The main contributions of this paper are first, insights into the inherent randomness of ad-auctions, and a description of the type of randomness encountered, namely high variability in the weights associated with bids, and variability in both the number and identity of players participating in different instances of the same keyword-auction. Second, we describe a model which captures this inherent variability, and which allows us, under additional simplifying assumptions on the distribution of bids, to derive closed form expressions for the optimal bid for a given player assuming they know their own valuation. The simplifying assumptions are that the bids an opponent faces in any particular instance of an auction are well approximated by sampling independently a random number of players with a common bid distribution. There are both empirical and theoretical reasons why this is good approximation for heavily contested keywords, and this assumption also allows players to work with observable-in-principle quantities (bids) of other opponents rather than unobservable valuations, offering the potential for advertisers to find their optimal bid easily. Our model produces utility functions that are smooth functions of a player's own value, in contrast to the discontinuous step functions of classical GSP models. Third, for the ad-platform, our model allows fast computation of players' valuations and click-through estimates to be derived from bids by assuming bids are optimal and finding the appropriate value, with the computation being two orders of magnitude faster than an Incremental Cost Per Click (ICC) model [3].

Related Work. GSP auctions have been studied by Varian in [11], and by Edelman, Ostrovsky and Schwarz (EOS) in [6]. Both examine Nash equilibria in a single auction with complete information. Most subsequent papers [1, 2, 7] have followed this simple model of a single auction. Some works [4, 5, 12] have looked at the dynamics of these auctions by considering an auction repeated multiple times with a fixed set of participants and most parameters fixed. Some others [10] even studied the dynamics of ad auctions with incoming and departing participants using tools from queueing theory.

The beneficial effects of noise on bids was emphasized in [4], where the authors noticed that such noise made some low-revenue equilibria unstable, and suggested adding some artificial noise to the bids. However they did not mention that such variability is naturally present in practice.

The other type of variability we observed in real auctions was a randomness in the identity of players participating in the auctions. The results on equilibria obtained for a single auction are hard to generalize to multiple auctions with varying opponents, and there is a lack of research in this area.

One paper which does consider multiple auctions with varying opponents and a random noise on the weights, is Athey and Nekipelov [3]. It is also probably the closest to our work in spirit. The main difference between our model and theirs is that they assume that players still have full information on their opponent's identity in each auction they enter, while in our model players ignore this information, and are given instead the bid distribution of their opponents. The advantage of our model compared to Athey-Nekipelov's is its simplicity, inducing a low computational cost of the estimation. However, our model being asymmetrically focused on a single player, cannot be used to directly manipulate Nash Equilibria, which Athey-Nekipelov's allows.

Gomes and Sweeney[8] consider a Bayes Nash equilibria of a GSP. Using the standard setting of Bayes-Nash, assuming a common prior distribution $F_v(v)$ of players values, they derive an integral equation which a bidding function $\beta(v)$ has to satisfy in an efficient symmetric equilibrium, and provide a way to calculate the solution recursively. They also discuss conditions under which efficient equilibria do not exist. It is possible to connect their main equation to ours in the case of a fixed number of players, by transforming the value distribution $F_v(v)$ into a distribution across bids $F_b(\cdot)$, using the optimal bid function $\beta(v)$. We comment on this more fully in section 6.2.

Organization. In Section 2, we give a quick introduction to the GSP auction mechanism and our model. We then present in Section 3 the results of studies done on real auctions in the Bing database, and show with these observations that the classical model of a single repeated auction is far from reality. We propose in Section 4 our new analytical model which is based on more realistic assumptions inspired from our observations on real data. Section 5 presents the results of estimations made using that model, Section 6 explores in detail similarities and differences between our model and some other works [3, 8], and we conclude in Section 7.

2 Notations and Assumptions

An overview of some of the sponsored search and some of the technology is given in Jansen and Mullen [9]. In the GSP mechanism currently used, the auctioneer (the ad-platform of a search engine, e.g. Google or Bing) sells to *advertisers* a finite set of *slots* in which ads are displayed. Each time a search query is returned, an auction is run to decide which ad will appear in each slot. The payment is then organized on a *price-per-click (PPC)* basis: an advertiser pays only when someone clicks on her ad.

All slots are sold simultaneously in one single auction, although the slots are not all equivalent. For example, the slots at the top of the result page (called the *mainline*) offer a greater visibility than the ones displayed in the *sidebar*. The *click-through-rate (CTR)* measures the probability of an ad being clicked on. Although CTRs may depend on multiple factors, a classical assumption is that they are *separable* as a product of two independent factors depending respectively on the slot position (*slot-CTR*) and the ad displayed in it (*ad-CTR*).

Notation Consider a set of N *players* (or *bidders*) competing for K slots. Each player i has a private *value* (or *valuation*) v_i she attributes to a click. To each player i is attributed a *weight*, which corresponds to the *ad-CTR* of the player's ad, and each slot j has a specific *slot-CTR* s_j, such that when player i is assigned to slot j she receives a CTR equal to $w_i s_j$.

Without loss of generality, we assume the slots to be ordered in decreasing order of slot-CTR: $s_0 \geq s_1 \geq \cdots \geq s_K$, and we assume $K > N$, by adding fictitious undisplayed slots with slot-CTR $s_j = 0$ if needed.

The auctioneer fixes a *reserve price* r corresponding to the lowest price at which it accepts to display ads.

GSP The GSP mechanism consists as follows. Each player i submits a bid b_i. The players whose weighted bid is lower than the reserve price r are rejected from the auction. The remaining k players are then ordered in decreasing order of their weighted bids, such that $b_1 w_1 \geq b_2 w_2 \geq \cdots \geq b_k w_k \geq r$, breaking ties randomly. The first player according to that order is then displayed in the first slot, the second player in the second slot, and so on, until all players are placed. Each player then pays per click a price equal to the minimal bid necessary to maintain their position. That is, player $i < k$ must bid at least $b_i \geq b_{i+1}w_{i+1}/w_i$ in order to stay above player $i+1$, and thus she has to pay $p_i(b) = b_{i+1}w_{i+1}/w_i$ per click. Similarly, the last player above the reserve price has to pay $p_k(b) = r/w_k$ per click. The *utility* (or *payoff*) of player i is then given by $U_i(v_i, b) = (v_i - p_i(b))w_i\varphi_i(b)$ where the *allocation function* $\varphi_i(b) \in \{s_0, s_1, \ldots, s_K\}$ denotes the CTR of the slot obtained by player i.

However, this simple description is far from being sufficient for a full understanding of the real GSP auctions used in practice. Indeed, we described here the mechanism of one single auction, while in practice, a player enters many distinct auctions with the same single bid, while the other parameters (opponents, weights) vary from one auction to the other, as we will show in the next section. In addition, the weights attributed to each advertiser are only estimates computed by the ad-platform and are not directly communicated to the advertisers.

3 Real Data Analysis

We now present some observations derived from the analysis of full records of one week of auctions run on Microsoft's search engine Bing.

We focused on several aspect of these auctions, such as the evolution of ad qualities or players' behavior during these auctions. In particular, several observations indicate that the classical model of one single auction is inaccurate to describe real auctions, in particular for ads shown frequently.

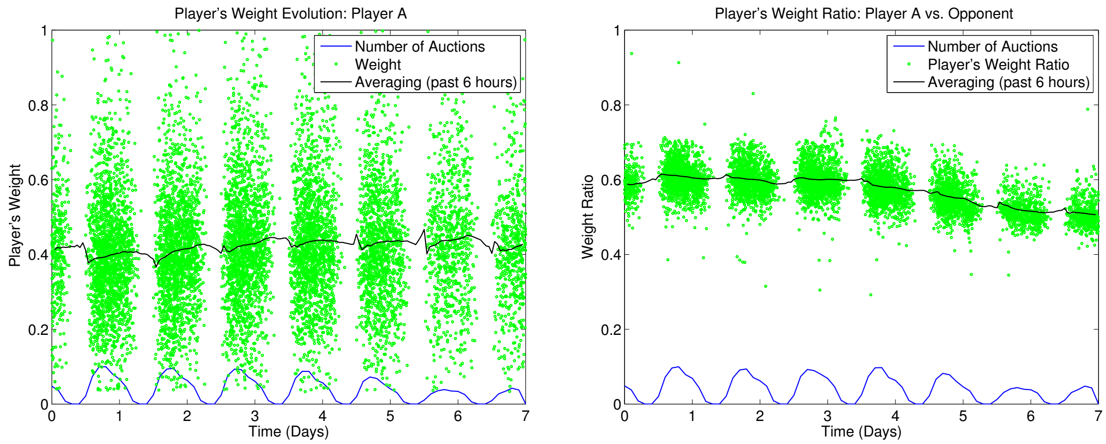
3.1 Stochastic Variability of the Weights

We study the behavior of two typical and distinct ads chosen from the set of ads that participated in the most auctions during the one week period. We will refer to these two ads as Player A and B. Each one of these two ads was associated with one single keyword, and could only participate in an auction when this keyword matched a user query. The two keywords associated to Player A and B were distinct and unrelated.

One of the most important features in ad auctions is the fact that the weights attributed to players are in practice estimations of their ad-CTR computed by the ad-platform itself, and are thus significantly variable.

Figure 1(a) shows the evolution of Player A's weight. As we can see, it is extremely variable (the standard deviation is 40% of the mean). An explanation for this could be the fact that the estimation of ad-CTR depends on multiple factors. Some of them are specific to the ad, but do not evolve fast enough to explain this variability. Some other factors are specific to the auction entered by the player, like the profile of the user performing the query and his propensity to click on ads. However, such factors are common to all ads entering the auction, and should not affect the weight ratio between two given players. To test this, we selected one of Player's A most frequent opponents and measured the weight ratio between these two players. The evolution of this ratio across time is displayed in Figure 1(b). We see that this ratio is also highly variable (the standard deviation is 9.5% of the mean). This suggests that the weights of the players are subject to some additional noise. We were unable to explain where this noise exactly comes from. We believe that the auction-specific parameters may simply have a non-multiplicative effect on the estimations, and may thus alter the weight ratio of two players even if shared by them.

Additionally, we find diurnal variation in real CTRs, and that real players' weighted bids and allocations can vary dramatically between players. When we look at the bids a particular player "faces", we see even if the opponents' bids are concentrating around a few values (for example, a few round numbers) the relative weighted bid distribution is smoothed by the randomness of the player's weights.



(a) Player A's weight

(b) Weight ratio between Player A and an opponent

Figure 1: Estimated probability of click (weight) of Player A's ad across the week and weight ratio against one of her opponents. Each green dot correspond to an auction in which Player A's participated (we display each 10th auction only). The solid black line indicates the average of these weights for the past 6 hours. The solid blue line at the bottom of each figure indicates the frequency of auctions in which Player A participated depending on the time of the day. The vertical axes have been renormalized to 1 for confidentiality reasons.

3.2 Repercussion on the Weighted Bids

Figure 2 shows the evolution of player A and B's bids, weighted bids and allocation. As we can see, the weighted bids are subject to an important noise, and the two players' allocation appears to be very irregular. It is interesting to notice periodicity with days in the average real slot-CTR received by the player, which suggests that the slot-CTR might in practice depend on the time of the day. In particular, the 6th and 7th day, Saturday and Sunday, received less auctions than during the week, although Player B's slot-CTR seems to increase on week-ends.

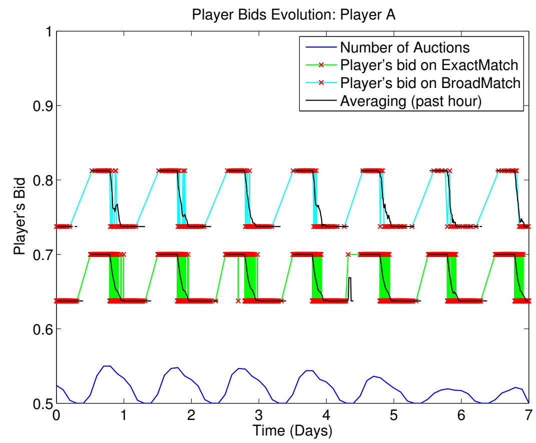
Figure 3 shows the histograms of opponents' bids and relative weighted bids $b_i w_i / w_0$ (where w_0 is Player A or B's weight) faced by the two players. As we can see, even if the bids are concentrating around a few values, mostly round numbers, the relative weighted bid distribution is smoothed by the randomness of the players' weights.

3.3 Insights from our observation

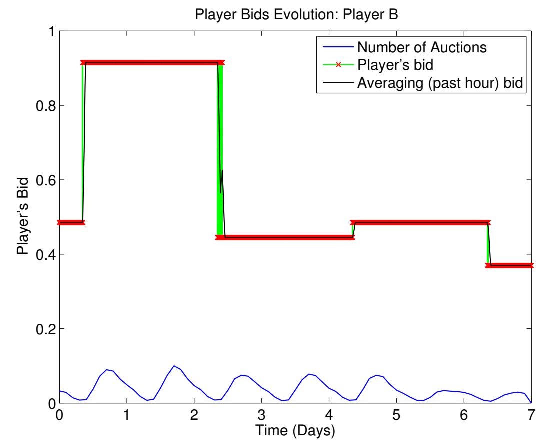
These observations reveal that real ad auctions are far from the classical model of one single auction with known players and complete information. In practice, a player enters many distinct auctions, with distinct opponents, varying weights and reserve prices. For this reason, a player cannot learn her opponents' bids or valuations, since they change from one auction to the other. But she could learn the distribution of the number of opponents and the distribution of their bids, assuming she had access to such information[1].

This inspires the analytical model presented in the next section, where we focus on one single player to whom these distributions would be given.

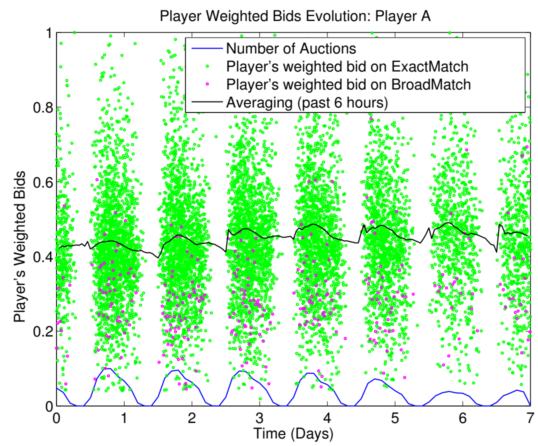
[1] Of course, in practice players do not have access to the full auction records and they probably never will. We emphasize however that it would be easy for the ad-platforms to provide advertisers with the distribution of the number of opponents and the distribution of their bids, or even with real-time estimates of their total number of clicks received and total price paid depending on their bid, by using our model.



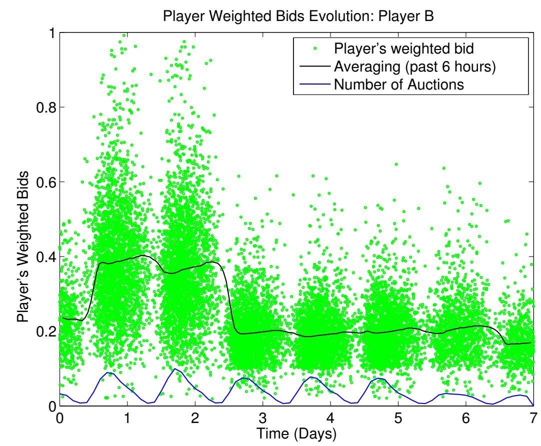
(a) Player A's bids



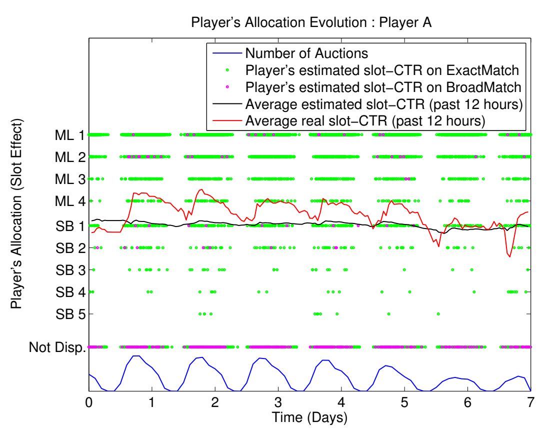
(b) Player B's bids



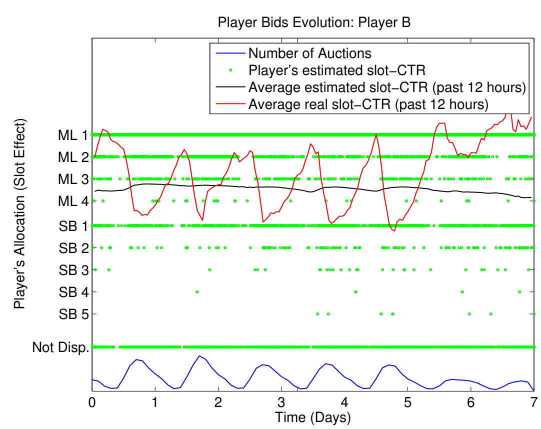
(c) Player A's weighted bids



(d) Player B's weighted bids

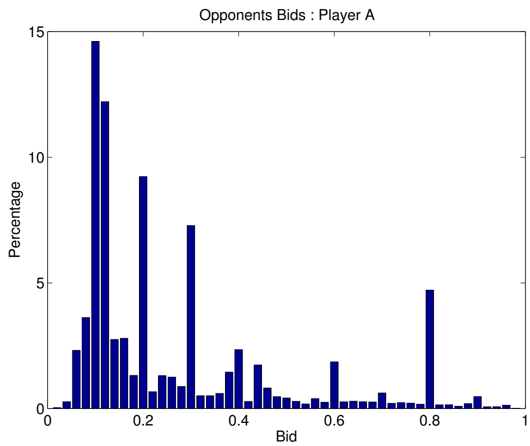


(e) Player A's allocation

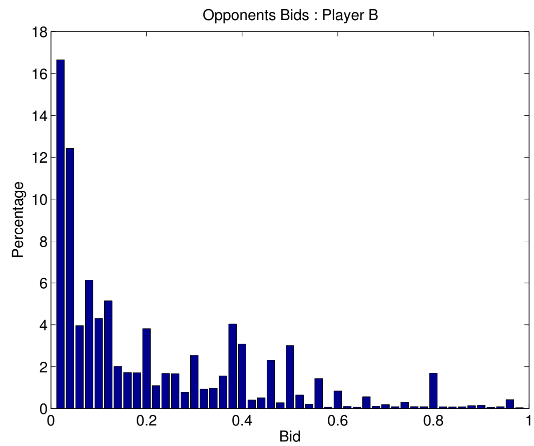


(f) Player B's allocation

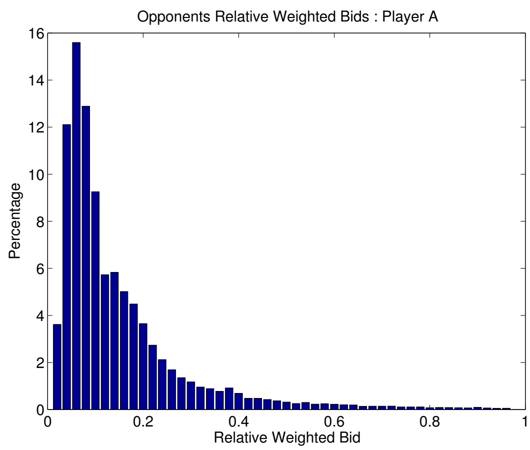
Figure 2: Evolution across the week of two players' bids, weighted bids and allocation. The two players are not related and were advertising on distinct, unrelated keywords. Player A placed distinct bids on ExactMatch and BroadMatch, so we considered the two separately. Player B placed the same bids on all Match Types. All the vertical axes were renormalized. In particular, the position of the various slots on the vertical axes of Figures (e) and (f) do not correspond to the estimates used to compute the average estimated slot-CTR.



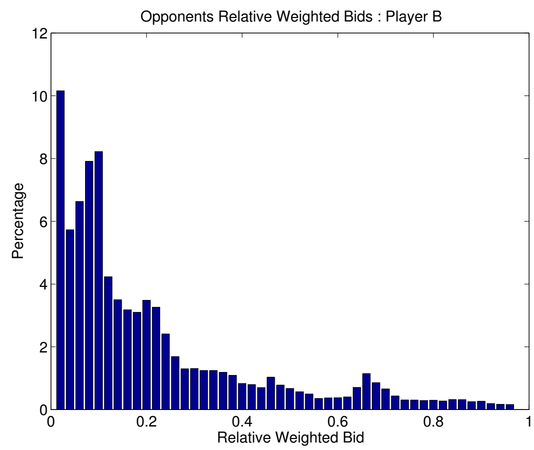
(a) Player A's opponents bids



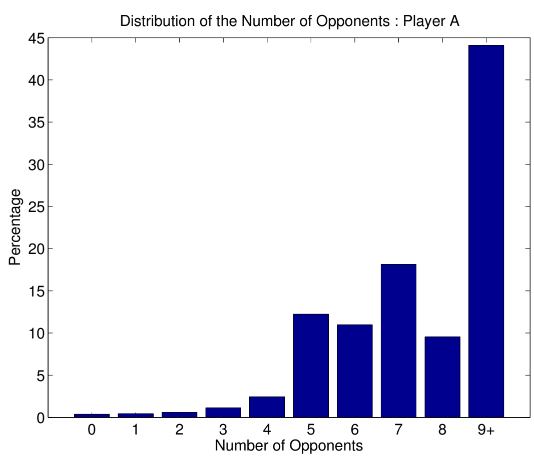
(b) Player B's opponents bids



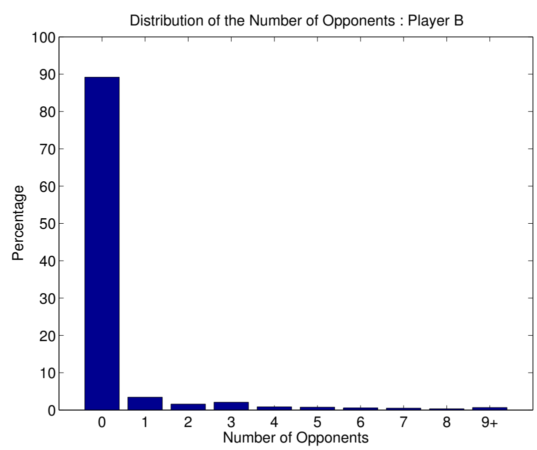
(c) Player A's opponents relative weighted bids



(d) Player B's opponents relative weighted bids



(e) Player A's number of opponents



(f) Player B's number of opponents

Figure 3: Distributions of Player's A and B's opponents bids, relative weighted bid and number of opponents. Horizontal axes have been renormalized on figures (a), (b), (c) and (d).

4 Analytical model

Our analytical model basically consists in considering a player with known value competing in a GSP auction against a given number of unknown opponents submitting random i.i.d. weighted bids with known probability distributions, and studying the player incentives in that auction.

We first emphasize that the assumptions needed for this model are realistic and fit our previous empirical observations.

4.1 Discussion on the relevance of the model

This model may seem far from reality at first sight, since we would expect the player to learn the bids of her adversaries with time. But there are a few effects that make real-life auctions closer to the model than it seems. In practice the player places a single bid on a keyword, and that bid is used for a great number of auctions with different configurations. We observed in particular that most players did not change their bid very often; unfortunately we cannot give any precise data on this for confidentiality reasons.

First the identity of the opponents can change over time, because the players enter auctions when the user query matches their ad. For instance, if a player places a bid on the keyword "insurance", she may enter an auction for the query "house insurance" and another for "car insurance", facing distinct opponents in the two auctions while submitting the same bid to both. We observed from the data set that frequently appearing ads met a wide range of opponents, unfortunately qualitative results are confidential.

Second, the profile of the user performing the query, such as his age or his geographical location, can change the relevance of ads and thus change the opponents met in an auction.

Third, bidders have in practice a monthly (or even daily) budget to spend, and when a bidder runs out of budget, she simply stops bidding until next month. This can also change the identity of a player's opponents across time. Lastly, advertisers may have several ads in the same campaign, and the ad-platform performs a rotation among these ads, so that if the same user performs the same query twice, he does not necessarily see the same ads twice.

In addition to the identity of the opponents, their weighted bids change as well across time, for several reasons. First the opponents can change their bid to explore and to optimize their revenue. Second, their weights, as well as the player's weight, change across time since they depend on several non-constant factors like the number of clicks received in the past, the profile of the user, and the keyword-query matching process.

Because of all these effects and others, the adversary bids faced by a player can be quite irregular, and the relative weighted bids are even more irregular (see Figures 3(c) and 3(d). We also observed that the number of opponents faced in one auction was not fixed, as shown by Figures 3(e) and 3(f). We later generalize our model to an uncertain number of opponents.

4.2 Model

We consider the following model: A *player* called Player 0 with known value v_0 and weight w_0 competes in a GSP auction with *reserve price* r against n *opponents*. These opponents submit random i.i.d. weighted bids with a known probability density function (*p.d.f.*) f, and we denote by F the corresponding cumulative distribution function. The slots are numbered from 0 to n, where we add dummy slots with a CTR of zero if necessary.

Without loss of generality, we assume that all players (including Player 0) have the same weights $w_i = 1$, since we can interpret the relative weighted bids $b_i w_i / w_0$ seen by Player 0 as simple bids. Thus the p.d.f. f corresponds to the distribution of opponents' relative weighted bids. We also denote by $\varphi(b)$ the CTR of the slot attributed to Player 0 when she bids b.

Under these assumptions, when Player 0 bids $b > r$ (we only consider Player 0's bids above the reserve price, but opponents may bid under), the probability that Player 0 gets the k-th slot for $k \leq n$ is then: $\mathbb{P}(\varphi(b) = s_k) = \binom{n}{k} F(b)^{n-k}(1 - F(b))^k$.

When in k-th position, Player 0 has $n - k$ players below her, and the maximum of their bids is a random variable with cumulative distribution $\mathbb{P}(b_1 < t, \ldots, b_{n-k} < t) = \mathbb{P}(b_1 < t)^{n-k} = F(t)^{n-k}$. Therefore, Player 0's expected price per click in that position is:

$$\begin{aligned} \mathbb{E}\big[p(b) \mid \varphi(b) = s_k\big] &= \frac{1}{F(b)^{n-k}} \left(rF(r)^{n-k} + \int_r^b tf(t)(n-k)F(t)^{n-k-1} dt \right) \\ &= b - \int_r^b \frac{F(t)^{n-k}}{F(b)^{n-k}} dt \end{aligned} \tag{1}$$

by integrating by parts.

Thus, her expected price per click is:

$$\begin{aligned} \mathbb{E}[p(b)] &= \sum_{k=0}^{n} \mathbb{E}[p(b) \mid \varphi(b) = s_k] \mathbb{P}(\varphi(b) = s_k) \\ &= \sum_{k=0}^{n} \binom{n}{k} F(b)^{n-k}(1 - F(b))^k \Big(b - \int_r^b \frac{F(t)^{n-k}}{F(b)^{n-k}} dt \Big) \\ &= b - \int_r^b (F(t) + (1 - F(b)))^n dt \ . \end{aligned} \tag{2}$$

It is interesting to notice that the expected price of our player is equal to her bid minus some discount depending on her bid, the reserve price, the number of opponents and the distribution function of their bids. In particular, we see that when the reserve price or the number of opponents increase, this discount becomes smaller and tends toward zero when r tends to b or n tends toward infinity. That is, for a high reserve price or with a high number of players, the GSP mechanism becomes closer to a First Price mechanism (where players pay their bids for each received click).

Using (1), we then see that the player expected utility is:

$$\begin{aligned} \mathbb{E}[U(v, b)] &= \sum_{k=0}^{n} \mathbb{P}(\varphi(b) = s_k) s_k (v - \mathbb{E}[p(b) | \varphi(b) = s_k]) \\ &= \sum_{k=0}^{n} \binom{n}{k} F(b)^{n-k}(1 - F(b))^k s_k \Big(v - b + \int_r^b \frac{F(t)^{n-k}}{F(b)^{n-k}} dt \Big) \end{aligned}$$

which can be rewritten with a more compact notation as:

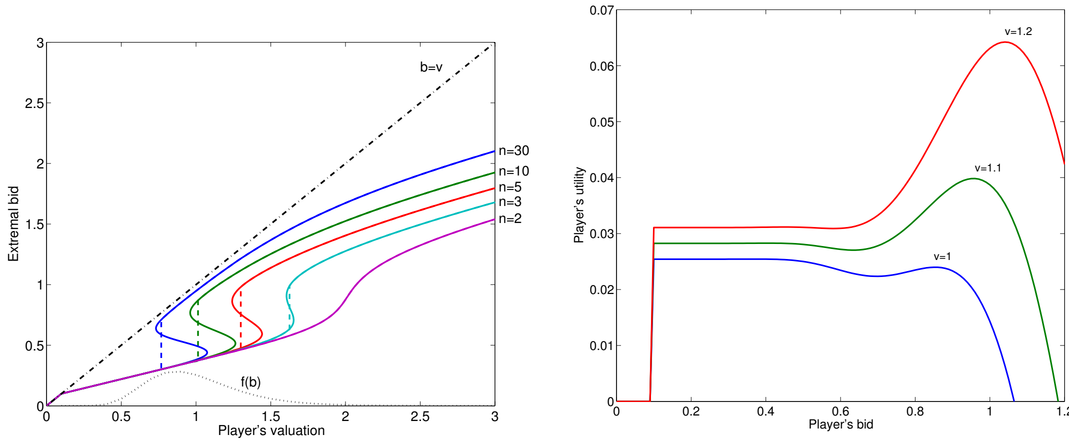
$$\mathbb{E}[U(v, b)] = \sum_{k=0}^{n} \binom{n}{k} s_k \int_r^v (1 - F(b))^k F(t \wedge b)^{n-k} dt \tag{3}$$

where $t \wedge b := \min(t, b)$.

The optimal bid b^{opt} of our player will maximize their expected utility, at which point the partial derivative of the expected utility with respect to the bid is zero. This lead to the following necessary condition, whose proof is given in Appendix A:

Theorem 4.1. *The optimal bid b^{opt} of Player 0 (as well as any local extrema of $\mathbb{E}[U(v, b)]$) must satisfy the following equation in b^*:*

$$v = b^* + \frac{\sum_{k=0}^{n-1} \binom{n-1}{k} s_{k+1} (1 - F(b^*))^k \int_r^{b^*} F(t)^{n-k-1} dt}{\sum_{k=0}^{n-1} \binom{n-1}{k} (s_k - s_{k+1})(1 - F(b^*))^k F(b^*)^{n-k-1}} \ . \tag{4}$$



(a) Solutions of equation (5).

(b) Expected utility for $v = 1, 1.1$ and 1.2.

Figure 4: (a) Plot of the solutions of equation (5) for several numbers of opponents ($n = 2, 3, 5, 10, 30$). When the equation has multiple solution, the dashed lines indicate where the optimal bid jumps from one solution to the other. The dash-dotted line indicates the main diagonal and the dotted line gives the opponent's bid distribution f ($\times 0.2$).
(b) Player 0's expected utility depending of her bid, against $n = 10$ opponents and for several valuations of the player.

Notice that this is a necessary but not sufficient condition: the equation does not characterize entirely the optimal bid which may have several solutions in b^*, namely all the local extrema of $\mathbb{E}[U(v, b)]$. However, we see that this equation determines entirely the value v as a function of b^*, allowing us to retrieve the player's valuation if we know her optimal bid.

It is also immediate that $b^{opt} \leq v$, since the slot-CTRs are decreasing. The player thus never has any incentive to overbid.

4.2.1 Study of the solutions of the equation

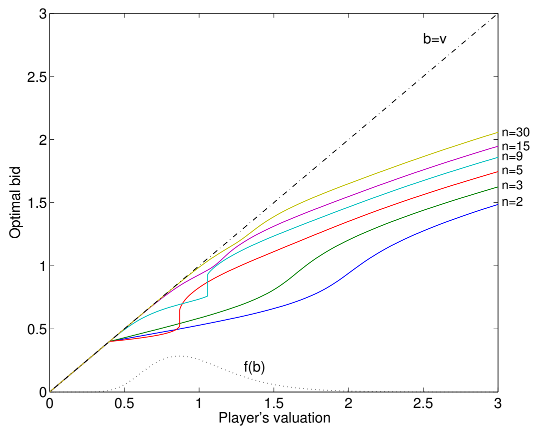
We first study the solutions of equation (4) in a simpler special case: A common approximation in the literature [5] consists in assuming that the slot-CTRs are geometrically decreasing. In the circumstances, the previous equation admits a much simpler form, whose proof is straightforward:

Corollary 4.2. *Assuming the slot-CTRs are geometrically decreasing, i.e. $s_j = \gamma^j$ for some $\gamma \in (0, 1)$, then equation (4) can be reduced to:*

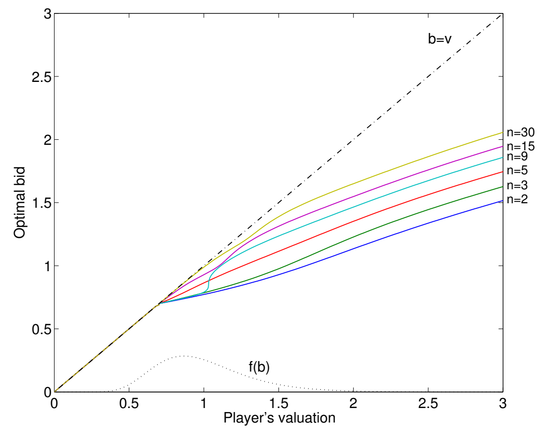
$$v = b^* + \frac{\gamma}{1-\gamma} \int_r^{b^*} \left(\frac{F(t) + \gamma(1 - F(b^*))}{F(b^*) + \gamma(1 - F(b^*))} \right)^{n-1} . \tag{5}$$

Figure 4(a) shows a plot of the extremal bids which are solutions of Equation (5) depending on the player's valuation v when the number of opponents varies[2]. We notice that for some valuations, there are 3 solutions to Equation (5). These solutions correspond to local extrema of the player's utility, but only one of them corresponds to the optimal bid. The dashed lines show were the optimal bid jumps from a solution to the other.

[2]We used the following parameters for all figures in that paragraph: reserve price $r = 0.1$ and decaying rate $\gamma = 0.7$. The considered opponents' bids distribution was a lognormal distribution with mean 1 and variance 0.1.



(a) Reserve price $r = 0.4$



(b) Reserve price $r = 0.7$

Figure 5: Optimal bid depending on the player's valuation for several numbers of opponents ($n = 2, 3, 5, 9, 15, 30$) in an auction with reserve price $r = 0.4$ and $r = 0.7$ and 9 allocated slots with slot-CTRs corresponding to those we observed in practice. The bid distribution is a lognormal with mean 1 and variance 0.1.

To better understand that effect, Figure 4(b) represents the Player's expected utility as a function of her bid when she plays against $n = 10$ opponents and when her value is $v = 1$, 1.1 and 1.2 (values for which Equation (5) has several solutions, according to Figure 4(a)). As we can see in Figure 4(b), the Player's utility curve has 3 local extrema: 2 local maxima and 1 local minimum in-between. When the player's value increases, the global maximum suddenly jumps from one local maximum to the other.

When the player's value is low, her utility curve becomes mostly flat, and the maximum given by the analytic solution of (5) becomes less meaningful, since it belongs to a sort of "plateau" (see curve $v = 1$ on Figure 4(b)).

4.2.2 Results for realistic slot-CTRs

We now show some results obtained with the general equation (4). In particular we can study the impact of the reserve price and the number of players.

Figure 5 shows the optimal bids against various number of opponents when the reserve price is $r = 0.4$ and $r = 0.7$ and the slots-CTRs s_i are taken similar to those observed in practice[3]. As we can see, we still have a discontinuity in the optimal bid for some values of n, but the general shape is much harder to understand and seems to depend highly on the number of opponents, particularly when this is smaller than the number of available slots. Observe, however, that when the reserve price or the number of players increases, this discontinuity tends to disappear.

Our main observation, from both real observation and synthetic models, is that the utility curve of a player is in general not quasi-concave, which causes a discontinuity in the optimal bid response of our player when her value changes. This highlights the pitfalls of an advertiser trying to find their optimal bid by local search, which may only find a local optimum.

[3]In practice on Bing, the first four slots in the mainline have a high CTR with high discrepancy between slots, while the 5 slots on the sidebar have a much lower CTR, which does not decrease much from one slot to the next.

4.3 Generalization to an Uncertain Number of Opponents

In practice, the number of opponents of a considered player may vary from one auction to another. We now enhance our model by considering the number of opponents as a random variable $\mathcal{N}$ in $\{0,\ldots,N\}$, independent of the bids, with $\mathbb{P}(\mathcal{N}=n)=p_n$.

In that case, the Player's expected utility becomes:

$$\mathbb{E}[U(v,b)] = \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k}p_n s_k \int_r^v (1-F(b))^k F(t\wedge b)^{n-k}dt, \tag{6}$$

and the optimal bid must thus verify the following equation:

$$v = b^* + \frac{\sum\limits_{n=1}^{N}\sum\limits_{k=0}^{n-1}\binom{n-1}{k} n p_n s_{k+1}(1-F(b^*))^k \int_r^{b^*} F(t)^{n-k-1}dt}{\sum\limits_{n=1}^{N}\sum\limits_{k=0}^{n-1}\binom{n-1}{k} n p_n (s_k - s_{k+1})(1-F(b^*))^k F(b^*)^{n-k-1}} . \tag{7}$$

We also have the expected allocation:

$$\mathbb{E}[\varphi(b)] = \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k}p_n s_k (1-F(b))^k F(b)^{n-k} , \tag{8}$$

and the expected total price paid by the player:

$$\mathbb{E}[p(b)\varphi(b)] = \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k}p_n s_k (1-F(b))^k F(b)^{n-k}\left(b - \int_r^b \frac{F(t)^{n-k}}{F(b)^{n-k}}dt\right) . \tag{9}$$

4.4 Adding a Mainline Reserve Price

In practice, Bing uses two distinct reserve prices for the mainline, which includes the first four slots, and the sidebar. We will now enhance our model to incorporate a mainline reserve price. We will refer to the sidebar reserve price as r, while the mainline reserve price will be R.

4.4.1 Fixed Reserve Prices

We first assume that the reserve prices are fixed. In that configuration, the expression of the expected utility differs depending on the player's bid compared to the reserve prices:

$$\mathbb{E}[U(v,b,r,R)] = \begin{cases} 0 & \text{if } b<r \\ \xi_1(v,b,r,R) & \text{if } r<b<R \\ \xi_2(v,b,r,R) & \text{if } b>R \end{cases} . \tag{10}$$

As we will see, the value when $b=r$ or $b=R$ is of no importance.

When $b>R$, the expected utility is similar to what we had before:

$$\xi_2(v,b,r,R) = \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k}p_n s_k \int_R^v (1-F(b))^k F(t\wedge b)^{n-k}dt ,$$

while when $r<b<R$, the expected utility become slightly more complicated to express:

$$\xi_1(v,b,r,R) = \sum_{n=0}^{N}\sum_{\substack{j,k=0\\ j+k\le n}}^{n}\binom{n}{j,k}p_n s_{k+j\vee 4}\int_r^v (1-F(R))^j (F(R)-F(b))^k F(t\wedge b)^{n-j-k}dt .$$

Indeed, the probability of having n opponents, k of them bidding between b and R and j bidding above R is equal to $p_n\binom{n}{j,k}(1-F(R))^j(F(R)-F(b))^k F(b)^{n-j-k}$, The player will then be in position $k+j\vee 4$ and will pay an expected price of $b-\displaystyle\int_r^b \frac{F(t)^{n-j-k}}{F(b)^{n-j-k}}dt$.

Because of this case disjunction, the player's utility is not continuous anymore, and thus it is not possible to express the optimal bid with a single equation. Moreover, since we considered the opponents' relative weighted bids, *i.e.* their weighted bids normalized by Player 0's weight, the relative reserve prices seen by Player 0 are in fact r/w_0 and R/w_0, and should then not be fixed but rather random.

We can address both issues simply by adding randomized reserve prices to our model and taking the expectation of the utility over these.

4.4.2 Random Reserve Prices

We now consider that the player is facing random reserve prices of the form $r\varepsilon, R\varepsilon$ where $\varepsilon = 1/w_0$ is some random variable with p.d.f. f_ε and independent from the number of opponents and their bids in the auction. With that additional assumption, we now get that :

$$\mathbb{E}[U(v,b,r,R)] = \int_0^\infty \frac{1}{\varepsilon}\Big(\mathbb{1}_{\{r\varepsilon<b<R\varepsilon\}}\xi_1(v,b,r\varepsilon,R\varepsilon) + \mathbb{1}_{\{b>R\varepsilon\}}\xi_2(v,b,r\varepsilon,R\varepsilon)\Big) f_\varepsilon(\varepsilon)d\varepsilon \quad ,$$

and thus :

$$\begin{aligned}\frac{\partial\mathbb{E}[U(v,b,r,R)]}{\partial b} &= \int_0^\infty \frac{1}{\varepsilon}(\mathbb{1}_{\{r\varepsilon<b<R\varepsilon\}}\frac{\partial\xi_1(v,b,r\varepsilon,R\varepsilon)}{\partial b} + \mathbb{1}_{\{b>R\varepsilon\}}\frac{\partial\xi_2(v,b,r\varepsilon,R\varepsilon)}{\partial b})f_\varepsilon(\varepsilon)d\varepsilon \\ &\quad + \frac{r}{b}f_\varepsilon(\frac{b}{r})\xi_1(v,b,b,\frac{R}{r}b) + \frac{R}{b}f_\varepsilon(\frac{b}{R})(\xi_2(v,b,\frac{r}{R}b,b) - \xi_1(v,b,\frac{r}{R}b,b)) \ ,\end{aligned}$$

and after a few calculations, we find that the optimal bid must solve:

$$v = b^* + \frac{\int_0^\infty \frac{1}{\varepsilon}\mathrm{num}(b^*,r\varepsilon,R\varepsilon)f_\varepsilon(\varepsilon)d\varepsilon + \frac{R}{b^*}f_\varepsilon(\frac{b^*}{R})\mathrm{numA}(b^*,r,R)}{\int_0^\infty \frac{1}{\varepsilon}\mathrm{den}(b^*,r\varepsilon,R\varepsilon)f_\varepsilon(\varepsilon)d\varepsilon + \frac{R}{b^*}f_\varepsilon(\frac{b^*}{R})\mathrm{denA}(b^*,r,R) + \frac{r}{b^*}f_\varepsilon(\frac{b^*}{r})\mathrm{denB}(b^*,r,R)} \ , \tag{11}$$

where :

$$\mathrm{num}(b,r,R) =$$

$$\begin{cases} 0 & \text{if } b<r \\ f(b)\displaystyle\sum_{n=1}^{N}\sum_{\substack{j,k=0\\ j+k\le n-1}}^{n-1}\binom{n-1}{j,k}np_n s_{k+1+j\vee 4}(1-F(R))^j(F(R)-F(b))^k\int_r^b F(t)^{n-j-k-1}dt & \text{if } r<b<R \\ f(b)\displaystyle\sum_{n=1}^{N}\sum_{k=0}^{n-1}\binom{n-1}{k}np_n s_{k+1}(1-F(b))^k\int_R^b F(t)^{n-k-1}dt & \text{if } b>R \end{cases} ,$$

$$\mathrm{den}(b,r,R) =$$

$$\begin{cases} 0 & \text{if } b<r \\ f(b)\displaystyle\sum_{n=1}^{N}\sum_{\substack{j,k=0\\ j+k\le n-1}}^{n-1}\binom{n-1}{j,k}np_n(s_{k+j\vee 4}-s_{k+1+j\vee 4})(1-F(R))^j(F(R)-F(b))^kF(b)^{n-j-k-1} & \text{if } r<b<R \\ f(b)\displaystyle\sum_{n=1}^{N}\sum_{k=0}^{n-1}\binom{n-1}{k}np_n(s_k-s_{k+1})(1-F(b))^kF(b)^{n-k-1} & \text{if } b>R \end{cases} ,$$

and :

$$\mathrm{numA}(b,r,R) = \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k} p_n s_{k\vee 4}(1-F(b))^k \int_{\frac{r}{R}b}^{b} F(t)^{n-k}dt \ ,$$

$$\mathrm{denA}(b,r,R) = \sum_{n=0}^{N}\sum_{k=0}^{n}\binom{n}{k} p_n (s_{k\vee 4})(1-F(b))^k F(b)^{n-k}dt \ ,$$

$$\mathrm{denB}(b,r,R) = \sum_{n=0}^{N}\sum_{\substack{j,k=0\\ j+k\leq n}}^{n}\binom{n}{j,k} p_n s_{k+j\vee 4}(1-F(\frac{R}{r}b))^j (F(\frac{R}{r}b)-F(b))^k F(b)^{n-j-k} \ .$$

Using this relation we can then retrieve the players' valuation knowing their optimal bid if we know the distribution of the number of opponents, the distribution of their relative weighted bids, and the distribution of the players' weights.

4.5 Results

We selected an ad from the auction database, and by looking at all the auctions it participated in we estimated both the distribution of the number of opponents and the distribution of their relative weighted bids. With these distributions, we computed the optimal bid curve of the player using the generalized analytical model with a mainline reserve price, and compared it to the optimal bid curve computed greedily by replaying all the auctions for all possible bids (with a low resolution on bids, since it is a heavy computation) and then choosing the optimal bid.

Figure 6 shows the comparison of the two optimal bid curves. The Optimal Bid Replay curve is not continuous because the bid resolution used to compute it was low. However both curves are quite close and present a similar discontinuity in the optimal bid.

This shows that this discontinuity is not an artifact of our model, and still happens when the parameters are measured from real auctions, or even when we compute explicitly the optimal bid with real auction data.

5 Experiments

One of the main benefits of our model is that it provides a simple method, easy to implement and with low computational cost, for estimating the expected number of clicks a player receives, (8), their expected total price (9), and even their valuations via (7). In this section we present the results of click estimation and value inference using generalized versions of these equation which incorporate a mainline reserve price as sketched in Section 4.4.

We used the full records of 132 millions auctions played over a given time period on the Bing ad-platform. We considered a randomly chosen set of 15777 ads, each associated with one specific keyword (or set of keywords) and with one specific bid. More precisely, each triple formed by an ad and its associated keyword and bid were considered as one specific player (Player 0). We then looked at the record of all auctions in which this player participated and estimated from these the distributions of the number of opponents[4] and of their relative weighted bids $b_i w_i / w_0$.

Using this information only, we then estimated the expected number of clicks received per auction by each player and inferred their valuations, assuming their bids were chosen optimally.

[4]Because the maximal number of slots displayed on Bing was 9 at the of this study, we considered only the 9 opponents with the highest weighted bids in each auction, since the other opponents have no impact on the result of the auction. The number of opponents was also taken to be at most 9.

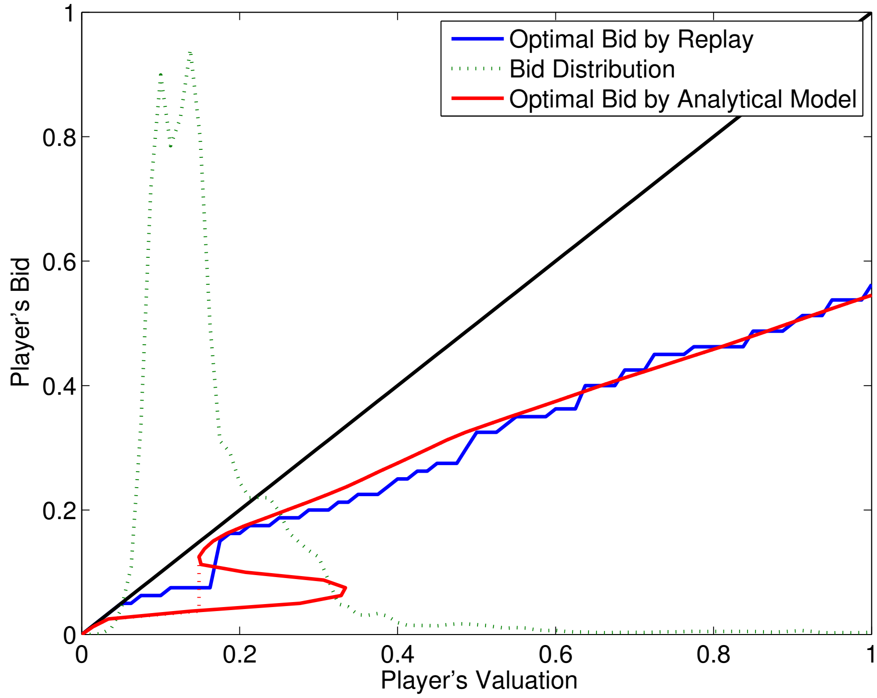


Figure 6: Optimal bid curves computed for an ad from real auction data. The Optimal Bid by analytical model was computed using the optimal bid equation derived from the analytical model, while the Optimal Bid by Replay was computed by replaying all the auctions with all possible bids and selecting the bid with the best utility. The Bid Distribution line represents the distribution of the opponents' relative weighted bid observed by the considered player. Both axes have been arbitrarily renormalized to 1.

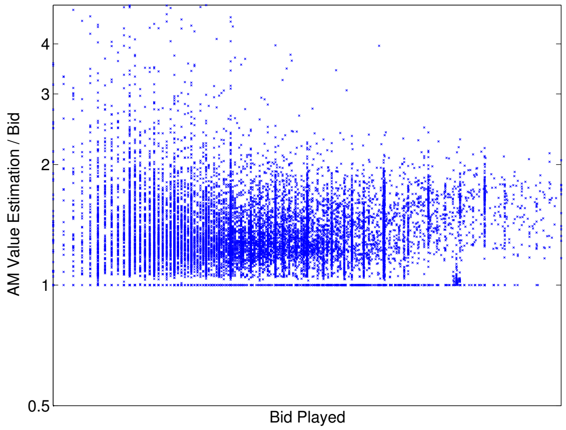
5.1 Valuation Inference Results

We compared our valuation estimations with the ones obtained from the ICC (Incremental Cost per Click) method designed by Athey and Nekipelov in [3][5].

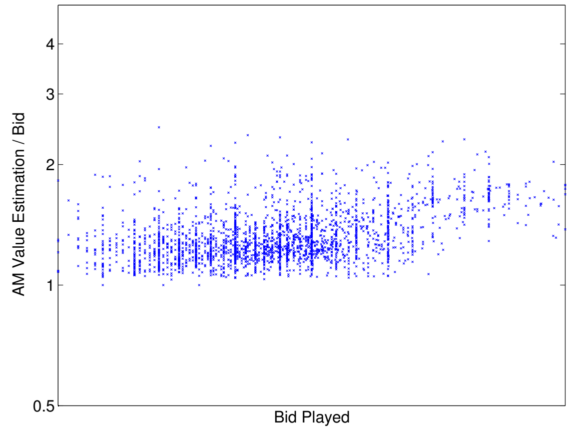
Figure 7 shows the comparison of our analytical model against Athey-Nekipelov's ICC model. A key parameter for the accuracy of the estimations is the slot-CTR : In real auctions, players receive a click with some probability, which is estimated by the CTR $w_i s_j$ for player i in slot j. When replaying the auctions, each displayed player is assumed to receive a fraction of a click corresponding to this probability. That is, player i in slot j will receive $w_i s_j < 1$ click in that auction. For a fair comparison, the same slot-effect estimations were used in both models.

It is interesting to notice in the figure that under both models, the estimated value is significantly higher for the advertisers who submitted a high bid. This suggest that players with a high valuation have a strong incentive to shade their bids, as already suggested by Figures 5 and 6. Also, since the key assumption for both models to infer valuations is that players are playing optimally, this may also suggest that players with high valuations do not play optimally in practice and do not shade their bids enough, which results in an over-approximation of their values by both models.

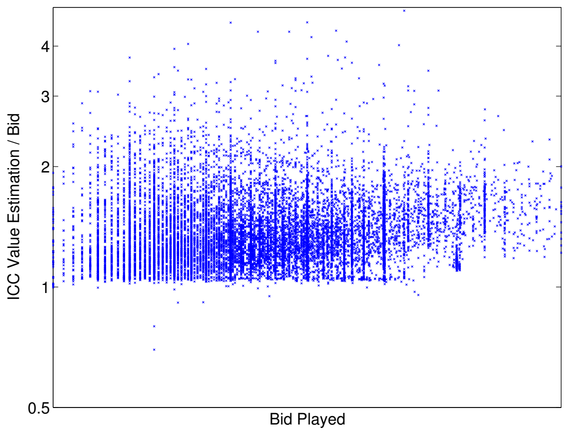
[5]The ICC method consists basically in replaying all the auctions in which a given player placed a given bid and estimating the value as being equal to the incremental cost per click at the bid placed. When the number of auctions is too low to have a smooth cost per click curve, virtual auctions are added by replaying the real one and adding an artificial noise on weights. This method also requires the assumption that advertisers are playing optimally.



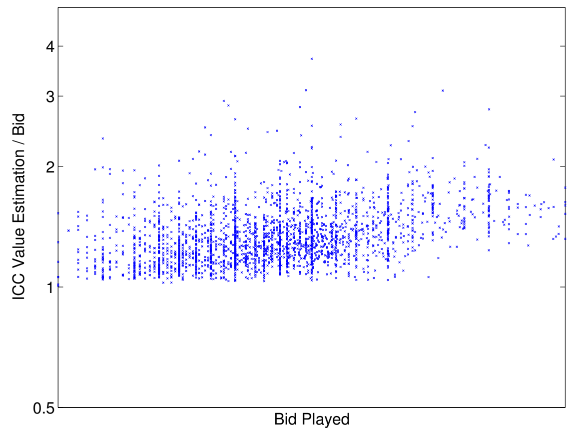
(a) Analytical model estimations



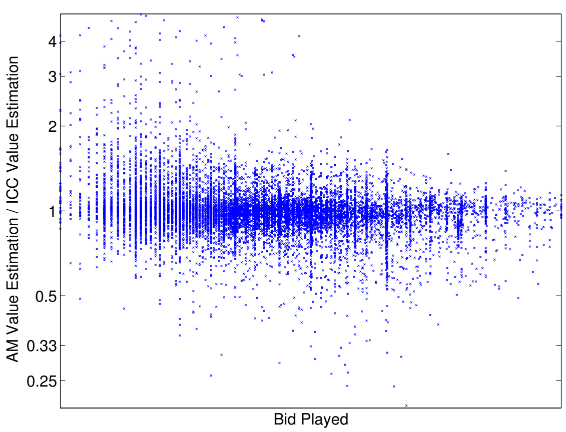
(b) Analytical model estimations



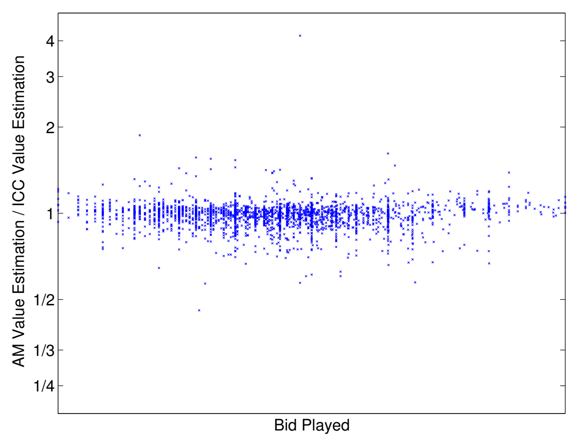
(c) ICC model estimations



(d) ICC model estimations

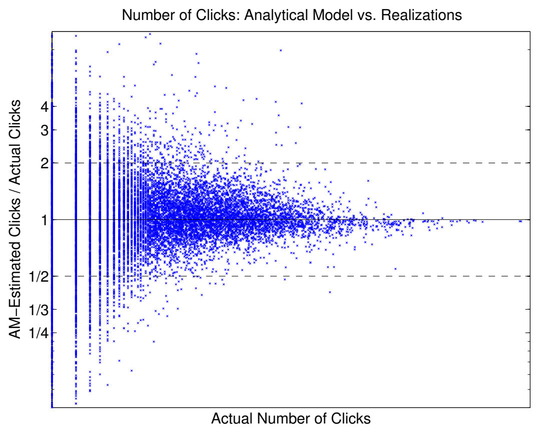


(e) Analytical model against ICC model

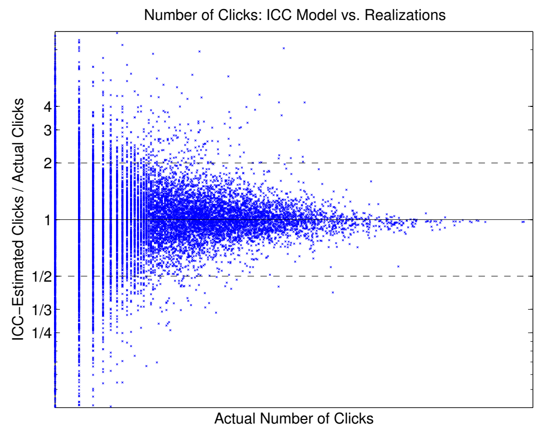


(f) Analytical model against ICC model

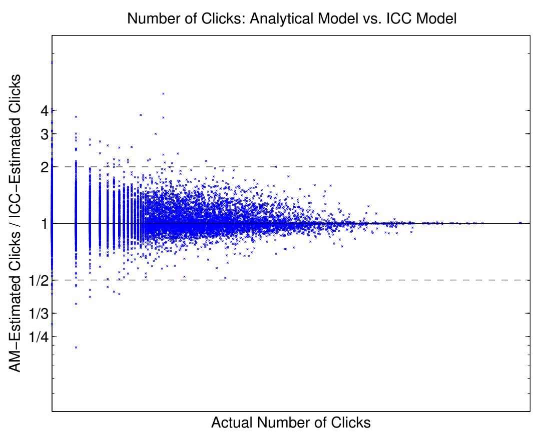
Figure 7: Estimations of the valuations compared to the original bids. Each point corresponds to one ad. The figures on the left show the estimations for all of the 15777 considered ads, while those on the right show the results for the few of them which received more than 100 clicks. The horizontal axes have been removed for confidentiality reasons. They use a logarithmic scale.



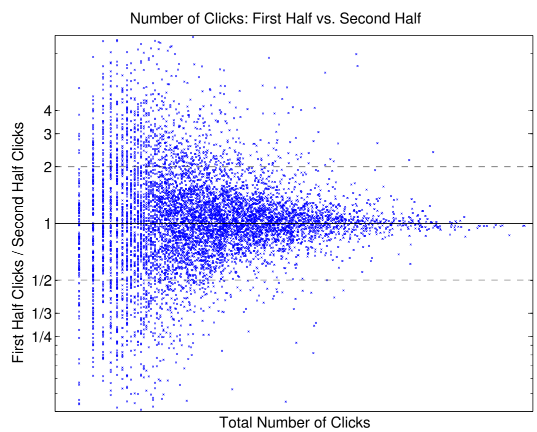
(a) Analytical model against Realizations



(b) Estimations by replay against Realizations



(c) Analytical model against Estimations by replay



(d) First half against Second half (Realizations)

Figure 8: Comparison of the click predictions of our analytical model against the actual number of received clicks (Realizations) and the Estimations by replaying the auctions. Each point corresponds to one ad. We did the estimations for 15777 distinct ads. The horizontal axes have been removed for confidentiality reasons. They use a logarithmic scale.

5.2 Estimations of the Number of Clicks

Since the real valuations are not known in practice, it is impossible to test directly the accuracy of the valuation estimations. We can however test the accuracy of both models by comparing the estimations of the expected number of clicks received by each player to the real number of clicks they actually received.

Figure 8 gives the results of the estimations of the number of clicks with these two models and compare them to their realizations, while Figure 9 shows a summary histogram of the results. They show the comparisons of the estimations of each model against the realizations (actual number of clicks received), and against each other. To test the predictive power of these estimations, we also compared the actual number of click received per auction in the first half of the considered time period against the second half.

As we can see, the estimations of our model and the replays are much closer than they are from the realizations, and as Figure 8(d) shows, the realizations are much more volatile than our estimations are inaccurate. This underline the difficulty of obtaining good estimates of the number of clicks received by a given ad, and shows that both models provide satisfying estimations, particularly for the ads frequently shown.

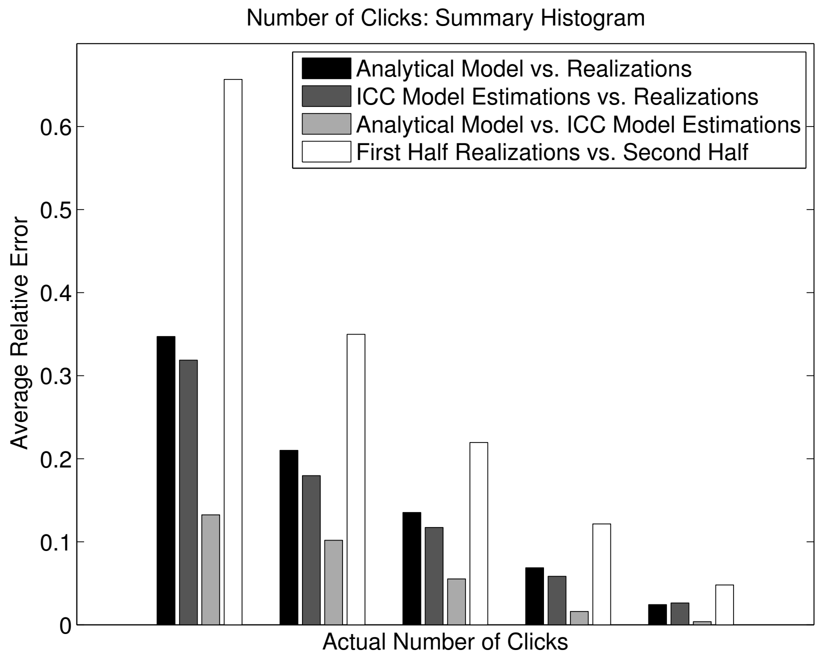


Figure 9: Summary histogram of the click estimations results for various ranges of received click (the number of click increasing exponentially from the left to the right). We show the average relative error of our model against the realizations, of the ICC model against the realizations, of our model against ICC, and of the realizations of the click obtained in the first half of the considered time period against the second half.

5.3 Implementation and Performances

While our estimation results confirm that the accuracy of our estimation method is comparable to Athey-Nekipelov's, our analytical model has a significant advantage over their ICC method: its speed.

The analytical model was implemented in F#. Most of the computing time was spent loading the auction data, the processing of these data was done in parallel of the loading. The total time spent loading and processing 132 million auctions was around 1 hour and a half with a single 8-core computer. Once the data processed, estimating the valuations, total number of clicks received and total price paid using the analytical model took less than 15 minutes, which is less than 0.1 second per ad. As a comparison, the replays of auctions and estimations using Athey-Nekipelov's ICC method was implemented in F# as well and took around half a day running on several HPC nodes.

The great advantage of our model over Athey-Nekipelov's is that ours is tractable on a large scale, and could even be implemented to give predictions for all of Bing's or Google's ads. Indeed, it would suffice to keep records for each ad of the distributions of the number of opponents and their relative weighted bids - which can be done online by updating them at each auction - to give in less than a second the prediction of the number of clicks and the total price depending on the chosen bid, while replaying all past auctions would take several minutes.

Using this, it would be possible to give the advertisers a precise feedback on their ad campaign, by providing them estimates of their click and price curves, so that they could chose their bid optimally, and the search engine would know as an extra their valuation using the analytical equation (7).

6 Comparison with other models

In this section we give theoretical comparisons between our model with Athey and Nekipelov's [3] first, and with Gomes and Sweeney's model [8].

6.1 Comparison with Athey and Nekipelov

Under some assumptions, it is possible to give a theoretical bound on the difference between the expected allocation, price and utility of players predicted by our analytical model and the Athey-Nekipelov model.

Athey and Nekipelov consider a set of $N+1$ players, that we will number from 0 to N, participating in several auctions, with fixed private valuations v_i and submitting public bids b_i. When an auction is run, a random set C of players is chosen to participate, and each player receives a random weight w_i according to a given distribution. All auctions are assumed to be independent from each other. Athey and Nekipelov assume that every player has full knowledge of the distribution of the set C and the weights w_i.

From a player's point of view, say Player 0, the set C^0 of his opponents in a given auction in which she participates is a random variable with known distribution, and thus the set of relative weighted bids to which Player 0's bid will be confronted is a random variable $B^{AN} = (b_i w_i / w_0)_{i \in C^0}$ with known distribution.

In the version of our model generalized to an uncertain number of opponents (Section 4.3), Player 0 is facing a random opponent bid vector with known distribution obtained as $B^{PK} = (b_1, \dots, b_{\mathcal{N}})$ where $\mathcal{N}$ is a random variable with distribution p and $b_1, \dots, b_{\mathcal{N}}$ i.i.d. random variables with c.d.f. F.

Thus, both our analytical model and Athey-Nekipelov's fall under the same general framework, were Player 0 is facing a random vector of opponent bids B taking values in $E = \bigcup_{n=0}^{N} \mathbb{R}^n$, with known probability distribution.

Using this we obtain the following result, whose proof is detailed in Appendix B:

Proposition 6.1. *Let us assume that the set C^0 is drawn uniformly across the subsets of $\{1, \dots, N\}$ of size n, and that the weights w_i are drawn independently. We define the bid c.d.f. F as: $\forall b, F(b) = \frac{1}{n} \sum_{i=1}^{N} \mathbb{P}_w(b_i w_i < b_0 w_0)$, and take B^{AN} and B^{PK} as defined previously from these parameters. Let $\mathbb{E}_B(\varphi(b))$ denote the expected utility of player 0 when bidding b against the random vector of bids B. Similarly, let $\mathbb{E}_B(p(b)\varphi(b))$ denote her expected total price and $\mathbb{E}_B(U(v,b)) = \mathbb{E}_B((v - p(b))\varphi(b))$ denote her expected utility. Then,*

$$|\mathbb{E}_{B^{AN}}(\varphi(b)) - \mathbb{E}_{B^{PK}}(\varphi(b))| \leq \frac{n(n-1)}{2N}$$

$$|\mathbb{E}_{B^{AN}}(p(b)\varphi(b)) - \mathbb{E}_{B^{PK}}(p(b)\varphi(b))| \leq \frac{n(n-1)}{2N} b$$

$$|\mathbb{E}_{B^{AN}}(U(v,b)) - \mathbb{E}_{B^{PK}}(U(v,b))| \leq \frac{n(n-1)}{2N} v$$

In particular, when the total number of players N grows to infinity and the number of players per auction n remains bounded, the difference between the estimations of the two models tends to zero.

As we showed previously, the utility curve of a player may not be quasi-concave in her bid (even with realistic parameter choices). The proposition thus implies that the utility curves are not always quasi-concave under Athey-Nekipelov's model either. Therefore their theorem showing existence and uniqueness of a Nash equilibrium under certain conditions (one of them being that the utility curves are quasi-concave) do not apply in general.

6.2 Comparison with Gomes and Sweeney

In [8], Gomes and Sweeney consider a Bayesian model where N players with random i.i.d. valuations with density $f_v(v)$ and corresponding c.d.f. $F_v(v)$ compete in a single GSP auction with no reserve price (we therefore assume $r = 0$ in this section).

They study the existence of symmetric efficient equilibrium where each player chooses her bid according to a strictly increasing optimal bid function $b = \beta(v)$. An equilibrium is said to be *efficient* if a player with a higher valuation than another player always obtains a higher slot than her. This implies that the bid function $\beta(v)$, if it exists, has to be strictly monotone. They show that, when it exists, the optimal bid function $\beta(v)$ must verify a closed form equation.

The following proposition, whose proof is given in Appendix C, states that under such circumstances, their equation can be derived directly from our equation (4):

Proposition 6.2. *If the optimal bid function $\beta(v)$ exists and is continuous and strictly monotone (and thus invertible), then the opponent bid vector for a given player is a vector of $N-1$ i.i.d. bids with c.d.f. $F(b) = F_v(\beta^{-1}(b))$. It must thus verify (4):*

$$v = \beta(v) + \frac{\sum_{k=0}^{N-2} \binom{N-2}{k} s_{k+1}(1 - F(\beta(v)))^k \int_0^{\beta(v)} F(t)^{N-k-2} dt}{\sum_{k=0}^{N-2} \binom{N-2}{k} (s_k - s_{k+1})(1 - F(\beta(v)))^k F(\beta(v))^{N-k-2}} , \tag{12}$$

which can be turned into:

$$\beta(v) = v - \frac{\sum\limits_{k=2}^{N} \binom{N-2}{k-1} s_{k-1}(1 - F_v(v))^{k-2} \int\limits_0^v (v - \beta(x)) F_v(x)^{N-k-1} f_v(x) dx}{\sum_{k=1}^{N-1} \binom{N-2}{k-1} s_{k-1}(1 - F_v(v))^{k-1} F_v(v)^{N-k-1}} \tag{13}$$

which corresponds to equation (1) of [8], up to notations (Gomes and Sweeney number their slots from 1 to $S < N$, while we number ours from 0 to $K > N$, but we can take $s_k = 0$ for $k \geq S$).

Although our equation (4) and Gomes-Sweeney's related equation (12) are close typographically, the two models have many differences, ranging from the assumptions used in them to the nature of the solutions obtained and their interpretation.

Assumptions. First, our model focuses on the incentives of a single player against a fixed randomized profile of opponents, while Gomes and Sweeney consider the behavior of several players in a symmetrical equilibrium. Given the highly dynamic aspect of these auctions and the very important number of advertiser participating in them, it is unclear how such equilibrium could happen in practice, while a best response approach seems more suitable.

Second, randomness does not arise from the same assumptions and do not apply to the same variables in the two models. In Gomes and Sweeney, randomness applies to the valuations and comes from an uncertainty in the players beliefs, while in our model, randomness is applied directly to the weighted bids and arises from the inherent variability of the weights. In particular, the Gomes-Sweeney model does not take into account the weights attributed to each player, and although their model is easy to generalize to fixed weights (a renormalization would allow to take the weights equal to 1), it cannot be generalized to randomized weights[6].

[6] For instance, consider the following simple example on the Gomes-Sweeney model. 2 players numbered 0 and 1 with i.i.d. values drawn from the c.d.f. F_v compete for one slot in an auction. In that simple case where only one slot is sold, GSP corresponds to a simple VCG auction and is truthful. Thus $\beta(v) = v$ and the Bayes-Nash equilibrium clearly exists. In that case, if Player 0 value is fixed as v_0, she will obtain the slot with probability $F_v(v_0)$. If we now assume that their bids are multiplied by some random i.i.d. weights w_0 and w_1 respectively, then we see easily that in general Player 0 will not win the auction with probability $F_v(v_0)$. For instance if the values are uniformly drawn in $[0, 1]$ and $v_0 = 1$, and the weights are equal to 1 with probability half and 2 with probability half, then Player 0 may loose the auction with probability 1/8 (if $w_0 = 1, w_1 = 2$ and $v_1 > 0, 5$) and thus wins with probability 7/8, which is different from $F_v(v_0) = 1$.

Solutions of the equations. The two models also differ in the nature of the solutions of their equations. The unknown of equation (13) is the whole function $\beta(v)$, while the unknown of (4) is just a tuple (v, b^*). Therefore, our equation is much simpler to solve, since v is uniquely determined by b^*, while Gomes and Sweeney show that the solutions of their equation can be computed as the limit of an recursively defined sequence of functions $\beta^{(n)}(v)$.

Moreover, when the distributions of the bids or values (depending on the model) are smooth, the solution curve of (13) is a continuous mapping from the values to the bids while the solution curve of (4) is a continuous mapping from the values to the bids. As observed earlier, the optimal bid function $b^{opt}(v)$ may thus present a discontinuity in our model, while it is continuous in the Gomes-Sweeney model.

Interpretations. Because of the different nature of the solutions, each model leads to a different interpretation. While Gomes and Sweeney shows that no symmetric efficient equilibrium is possible if the slot-CTRs are too close to each other, we give several examples where the utility curve of a player is not quasi-concave in her bid, resulting in a discontinuity of the optimal bid curve. Because of this, we believe that even asymmetric and inefficient Nash equilibria may not always exist in practice.

7 Conclusion and further works

We have developed a new analytical model of GSP auctions centered on the behavior of one single player facing multiple unknown opponents who submit random bids. This approach has been inspired by a conclusion drawn from real auction data, which showed that most ads enter many auctions with the same bid, facing multiple opponents whose identities differ from one auction to the other, and with highly irregular weights attributed to the bids.

Using this model, we were able to derive simple, yet very general, closed form expressions for the expected number of clicks, total price and utility of the considered player, using only information on the distribution of the number of opponents and of their relative weighted bids. We also expressed exactly the relation between the optimal bid and the player's valuation, highlighting the fact that in some cases encountered in practice, the optimal bid jumps discontinuously from one value to the other when the valuation increases. However, we noticed that this discontinuity tends to disappear when the number of opponents or the reserve price (i.e. the competition) increases.

The main advantage of this model is its very low computational and memory cost, compared to replaying all auctions. Because of this, it will be practical to give better feedback to advertisers by providing them with personalized estimates of their click and price curves, thus helping them to chose their bids more efficiently.

In our analysis, we have assumed that the weights come from independent stochastic processes. In practice, as we alluded to, the actual CTR and its estimated weight is affected by the "type" of the searcher, and advertisers can take this into account when bidding. A simple example of this is demographic targeting, where certain demographics are valued more highly by certain advertisers, and such advertisers can automatically or manually adjust their bids for search queries coming from favored locations. In addition, keywords associated with the same ad or campaign are not independent. The extension of our model to cater for these interesting and practical scenarios is for future work.

References

[1] Abrams, Z., and Ghosh, A. Auctions with revenue guarantees for sponsored search. In *Proceedings of the 3rd International Workshop on Internet and Network Economics* (2007), WINE'07.

[2] Aggarwal, G., Goel, A., and Motwani, R. Truthful auctions for pricing search keywords. In *Proceedings of the 7th ACM conference on Electronic Commerce* (2006), EC '06.

[3] Athey, S., and Nekipelov, D. A structural model of sponsored search advertising auctions. In *Sixth Ad Auctions Workshop* (2010), adauctions2010.

[4] Borgs, C., Chayes, J., Immorlica, N., Jain, K., Etesami, O., and Mahdian, M. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th international conference on World Wide Web* (2007), WWW '07.

[5] Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A. R., Mathieu, C., and Schwarz, M. Greedy bidding strategies for keyword auctions. In *Proceedings of the 8th ACM conference on Electronic commerce* (2007), EC '07.

[6] Edelman, B., Ostrovsky, M., and Schwarz, M. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *American Economic Review* (2007), 242–259.

[7] Edelman, B., and Schwarz, M. Optimal auction design and equilibrium selection in sponsored search auctions. *American Economic Review* (2010).

[8] Gomes, R. D., and Sweeney, K. S. Bayes-nash equilibria of the generalized second price auction. In *Proceedings of the 10th ACM conference on Electronic Commerce* (New York, NY, USA, 2009), EC '09.

[9] Jansen, B. J., and Mullen, T. Sponsored search: an overview of the concept, history, and technology. *Int. J. Electronic Business 6*, 2 (2008), 114–131.

[10] Menache, I., Ozdaglar, A., Srikant, R., and Acemoglu, D. Dynamic online-advertising auctions as stochastic scheduling. In *Workshop on the Economics of Networks, Systems and Computation* (2009), NetEcon '09.

[11] Varian, H. R. Position auctions. *International Journal of Industrial Organization* (2007), 1163–1178.

[12] Vorobeychik, Y., and Reeves, D. M. Equilibrium analysis of dynamic bidding in sponsored search auctions. In *Proceedings of the 3rd international Workshop on Internet and Network Economics* (2007), WINE'07, Springer-Verlag.

Appendix

A Proof of theorem 4.1

The expected utility $\mathbb{E}[U(v,b)]$ of Player 0 is given by (3), and is clearly differentiable in b. The optimal bid, and any local extrema, must thus be a zero of its partial derivative in b.

To compute more easily this partial derivative, we use the following lemma, whose proof is straightforward:

Lemma A.1. *Let f and g be two differentiable real functions on some non-empty interval $[a,b]$, and consider the function I defined on $[a,b]$ by $I(x) = \int_a^b f(x)g(t \wedge x)dt$. Then I is differentiable and:*

$$I'(x) = \int_a^b \Big(\mathbb{1}_{\{t>x\}} f(x) g'(t \wedge x) + f'(x) g(t \wedge x) \Big) dt \ ,$$

where $\mathbb{1}$ is the indicator function.

We thus have:

$$\begin{aligned}
&\frac{\partial \mathbb{E}[U(v,b)]}{\partial b} \\
&= \sum_{k=0}^{n} \binom{n}{k} f(b) s_k \int_r^v \Big(\mathbb{1}_{\{t>x\}} (n-k)(1-F(b))^k F(t\wedge b)^{n-k-1} - k(1-F(b))^{k-1} F(t \wedge b)^{n-k} \Big) dt \\
&= n f(b) \sum_{k=0}^{n-1} \binom{n-1}{k} \int_r^v (\mathbb{1}_{\{t>x\}} s_k - s_{k+1})(1-F(b))^k F(t\wedge b)^{n-k-1} dt \\
&= n f(b) \sum_{k=0}^{n-1} \binom{n-1}{k} (1-F(b))^k \left((v-b)(s_k - s_{k+1}) F(b)^{n-k-1} - s_{k+1} \int_r^b F(t)^{n-k-1} dt \right) \ .
\end{aligned}$$

Finally, we obtain equation (4) with one last manipulation.

B Proof of proposition 6.1

The key idea is that if to compare the probability distributions of B^{AN} and B^{PK} induced by each model. More precisely we define the *total variation distance* of two random variable X and Y on a measurable space (E,Σ) as: $d_{TV}(X,Y) = \sup_{A\in\Sigma} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|$.

Then the following two lemmas hold:

Lemma B.1. *For any coupling of two random variables X and Y, $d_{TV}(X,Y) \leq \mathbb{P}(X \neq Y)$.*

Proof. For any $A \in \Sigma$, we have:

$$\begin{aligned}
&\mathbb{P}(X \in A) - \mathbb{P}(Y \in A) \\
&= \mathbb{P}(X \in A \ \&\ Y \notin A) + \mathbb{P}(X \in A \ \&\ Y \in A) - \mathbb{P}(Y \in A) \\
&\leq \mathbb{P}(X \in A \ \&\ Y \notin A) \leq \mathbb{P}(X \neq Y) \ .
\end{aligned}$$

□

Lemma B.2. *Let X, Y be two random variables taking values in a measurable space (E,Σ), and let $\psi : (E,\Sigma) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a measurable positive real function.*

If ψ is bounded, i.e. if $||\psi||_\infty := \max_{x\in E} |\psi(x)| < \infty$, then we have $|\mathbb{E}_X(\psi(X)) - \mathbb{E}_Y(\psi(Y))| \leq ||\psi||_\infty d(X,Y)$.

Proof. We first assume that ψ is a step function and thus takes a finite number of values $\{a_1, \ldots, a_n\}$, and let $A^+ := \{a_i | \mathbb{P}_X(\psi(X) = a_i) \geq \mathbb{P}_Y(\psi(Y) = a_i)\}$ and $A^- := \{a_i | \mathbb{P}_X(\psi(X) = a_i) < \mathbb{P}_Y(\psi(Y) = a_i)\}$.

We have then:

$$\begin{aligned}
\mathbb{E}_X(\psi(X)) - \mathbb{E}_Y(\psi(Y)) &= \sum_{i=1}^{n} a_i(\mathbb{P}_X(X = a_i) - \mathbb{P}_Y(Y = a_i)) \\
&\leq ||\psi||_\infty(\mathbb{P}_X(X \in A^+) - \mathbb{P}_Y(Y \in A^+)) \\
&\leq ||\psi||_\infty d_{TV}(X, Y) \ .
\end{aligned}$$

The above inequality is obtained by using the fact that

$$a_i(\mathbb{P}_X(X = a_i) - \mathbb{P}_Y(Y = a_i)) \leq ||\psi||_\infty(\mathbb{P}_X(X = a_i) - \mathbb{P}_Y(Y = a_i))$$

when $a_i \in A^+$ and $a_i(\mathbb{P}_X(X = a_i) - \mathbb{P}_Y(Y = a_i)) \leq 0$ when $a_i \in A^-$.

By symmetry between X and Y, we also have $\mathbb{E}_Y(\psi(X)) - \mathbb{E}_X(\psi(Y)) \leq ||\psi||_\infty d_{TV}(X, Y)$ which proves the proposition when ψ is a step function.

The general case follows, since any measurable function can be approximated by a sequence of step functions. □

Since $||\varphi(b)||_\infty \leq 1$, $||p(b)||_\infty \leq p$ and $||U(v, b)||_\infty \leq v$, proposition 6.1, will follow from lemma B.2 if we show that under the assumptions stated, $d_{TV}(B^{AN}, B^{PK}) \leq \frac{n(n-1)}{2N}$. For more convenient notations, for any integers $i \leq j$, we will write for $[i, j] := \{i, i+1, \ldots, j\}$.

To bound the total variation distance of the two variables, we build the following coupling between them:

Let $C = (c_1, \ldots, c_N)$ be a set of n i.i.d. players drawn uniformly in $[1, N]$, containing possibly several copies of the same player, and let $\widehat{C}$ be a set of n players in $\{1, \ldots, N\}$ taken at random uniformly amongst the sets of n distinct players. We then construct the random set C' as $C' = C$ if C do not contain a player twice, $C' = \widehat{C}$ otherwise. It is easy to see that C' always contains n distinct players, and has the same distribution law than $\widehat{C}$. Moreover using the independency of the components of C, we get:

$$\begin{aligned}
\mathbb{P}(C' \neq C) &\leq \mathbb{P}(\exists i \in [1, n], \exists j \in [i+1, n], c_i = c_j) \\
&\leq \sum_{i=1}^{n} \mathbb{P}(c_i \in \{c_j \mid j \in [i+1, n]\}) \\
&\leq \sum_{i=1}^{n} \frac{n-i}{N} = \frac{n(n-1)}{2N} \ .
\end{aligned}$$

Indeed, since the C_i's are i.i.d. drawn uniformly in $[1, N]$, we have that $\mathbb{P}(c_i = j) = 1/N$ for any $i \in [1, n]$ and $j \in [1, N]$. Thus for any set $A \subset [1, N]$ of size k, $\mathbb{P}(c_i \in A) \leq k/N$.

We can then construct B^{AN} and B^{PK} as follow. Take $B^{AN} = (b_i w_i / w_0)_{i \in C'}$ and $B^{PK} = (b_i w_i / w_0)_{i \in C}$. It is then immediate to check that B^{AN} and B^{PK} have the wanted distributions, and that $\mathbb{P}(B^{AN} \neq B^{PK}) \leq \mathbb{P}(C' \neq C) \leq \frac{n(n-1)}{2N}$.

Thus, by lemma B.1, $d_{TV}(B^{AN}, B^{PK}) \leq \frac{n(n-1)}{2N}$, which concludes the proof.

More generally, this proof shows that whenever the total variation distance between B^{AN} and B^{PK} can bounded by some ε, the bounds given in the proposition stay true by replacing $\frac{n(n-1)}{2N}$ with ε.

C Proof of Proposition 6.2

With an integration by part and variable change ($t = \beta(x)$), and since $\beta(0) = 0$ (the only possible bid for a player with valuation zero is zero), we get:

$$\begin{aligned}\int_0^{\beta(v)} F(t)^{N-k-2}dt &= \int_0^{\beta(v)} (\beta(v) - t)(N-k-2)F(t)^{N-k-3}f(t)dt \\ &= \int_0^{v} (\beta(v) - \beta(x))(N-k-2)F_v(x)^{N-k-3}f_v(x)dx \ .\end{aligned}$$

We thus have from (12):

$$\begin{aligned}&\sum_{k=0}^{N-2} \binom{N-2}{k} s_k(1-F_v(v))^k F_v(v)^{N-k-2}(v-\beta(v)) \\ &= \sum_{k=0}^{N-2} \binom{N-2}{k} s_{k+1}(1-F_v(v))^k \\ &\qquad \left[(v-\beta(v))F_v(v)^{N-k-2} + \int_0^v (\beta(v)-\beta(x))(N-k-2)F_v(x)^{N-k-3}f_v(x)dx\right] \\ &= \sum_{k=0}^{N-2} \binom{N-2}{k} (N-k-2)s_{k+1}(1-F_v(v))^k \int_0^v (v-\beta(x))F_v(x)^{N-k-3}f_v(x)dx \\ &= \sum_{k=2}^{N} \binom{N-2}{k-1} (k-1)s_{k-1}(1-F_v(v))^{k-2} \int_0^v (v-\beta(x))F_v(x)^{N-k-1}f_v(x)dx \ ,\end{aligned}$$

and finally

$$\beta(v) = v - \frac{\sum\limits_{k=2}^{N} \binom{N-2}{k-1} s_{k-1}(1-F_v(v))^{k-2} \int\limits_0^v (v-\beta(x))F_v(x)^{N-k-1}f_v(x)dx}{\sum_{k=1}^{N-1} \binom{N-2}{k-1} s_{k-1}(1-F_v(v))^{k-1}F_v(v)^{N-k-1}}$$

which is (13).