

# Econometrics for Learning Agents

DENIS NEKIPELOV, University of Virginia, denis@virginia.edu  
 VASILIS SYRGKANIS, Microsoft Research, vasy@microsoft.com  
 EVA TARDOS, Cornell University, eva.tardos@cornell.edu

The main goal of this paper is to develop a theory of inference of player valuations from observed data in the generalized second price auction without relying on the Nash equilibrium assumption. Existing work in Economics on inferring agent values from data relies on the assumption that all participant strategies are best responses of the observed play of other players, i.e. they constitute a Nash equilibrium. In this paper, we show how to perform inference relying on a weaker assumption instead: assuming that players are using some form of no-regret learning. Learning outcomes emerged in recent years as an attractive alternative to Nash equilibrium in analyzing game outcomes, modeling players who haven't reached a stable equilibrium, but rather use algorithmic learning, aiming to learn the best way to play from previous observations. In this paper we show how to infer values of players who use algorithmic learning strategies. Such inference is an important first step before we move to testing any learning theoretic behavioral model on auction data. We apply our techniques to a dataset from Microsoft's sponsored search ad auction system.

Categories and Subject Descriptors: J.4 [Social and Behavioral Sciences]: Economics; K.4.4 [Computers and Society]: Electronic Commerce; F.1.2 [Modes of Computation]: Online computation

General Terms: Theory, Economics, Econometrics, Experimental

Additional Key Words and Phrases: Econometrics, no-regret learning, sponsored search, value inference

## 1. INTRODUCTION

The standard approach in the econometric analysis of strategic interactions, starts with assuming that the participating agents, conditional on their private parameters, such as their valuation in an auction, can fully optimize their utility given their opponents actions and that the system has arrived at a stable state of mutual such best-responses, aka a Nash equilibrium.

In recent years, learning outcomes have emerged as an important alternative to Nash equilibria. This solution concept is especially compelling in online environments, such as Internet auctions, as many such environments are best thought of as repeated strategic interactions in a dynamic and complex environment, where participants need to constantly update their strategies to learn how to play optimally. The strategies of agents in such environments evolve over time, as they learn to improve their strategies, and react to changes in the environment. Such learning behavior is reinforced even more by the increasing use of sophisticated algorithmic bidding tools by most high revenue/volume advertisers. With auctions having emerged as the main source of revenue on the Internet, there are multitudes of interesting data sets for strategic agent behavior in repeated auctions.

---

Author's addresses: denis@virginia.edu, vasilis@cs.cornell.edu, eva.tardos@cornell.edu. Éva Tardos was supported in part by NSF grants CCF-0910940; and CCF-1215994, ONR grant N00014-08-1-0031, a Yahoo! Research Alliance Grant, and a Google Research Grant. Denis Nekipelov was supported in part by NSF grants CCF-1449239, CCF-0910940 and a Google Research Grant.

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

*EC'15*, June 15–19, 2015, Portland, OR, USA. Copyright © 2015 ACM 978-1-4503-3410-5/15/06 ...\$15.00.  
<http://dx.doi.org/10.1145/2764468.2764522>

To be able to use the data on agent behavior to empirically test the prediction of the theory based on learning agents, one first needs to infer the agents' types, or valuations of the items, from the observed behavior without relying on the stable state best-response assumption. The idea behind inference based on the stable best response assumption is straightforward: the distribution of actions of players is observed in the data. If we assume that each player best responds to the distribution of opponents' actions, this best response can be recovered from the data. The best response function effectively describes the preferences of the players, and can typically be inverted to recover each player's private type. This is the idea used in all existing work in economics on this inference problem, including [Athey and Nekipelov 2010], [Bajari et al. 2013], and [Xin Jiang and Leyton-Brown 2007].

There are several caveats in this approach. First of all, the assumption that an equilibrium has been reached is unrealistic in many practical cases, either because equilibrium best response functions are hard to compute or the amount of information needed to be exchanged between the players to reach an equilibrium outcome is unrealistically large. Second, the equilibrium is rarely unique especially in dynamic settings. In that case the simple "inversion" of the best responses to obtain the underlying valuations becomes a complicated computational problem because the set of equilibria frequently has a complicated topological structure. The typical approach to avoid this complication is to assume some equilibrium refinement, e.g. the *Markov Perfect Equilibrium* in the case of dynamic games.

In spite of a common understanding in the Economics community that many practical environments cannot be modeled using the traditional equilibrium framework (and that the assumption of a particular equilibrium refinement holds), this point has been overshadowed by the simplicity of using the equilibrium based models.

In this paper we consider a dynamic game where players learn how to play over time. We focus on a model of *sponsored search auctions* where bidders compete for advertising spaces alongside the organic search results. This environment is inherently dynamic where the auction is run for every individual consumer search and thus each bidder with an active bid participates in a sequence of auctions.

Learning models agents that are new to the game, or participants as they adjust to a constantly evolving environment, where they have to constantly learn what may be best play. In these cases, strict best response to the observed past may not be a good strategy, it is both computationally expensive, and can be far from optimal in a dynamically changing environment. Instead of strict best response, players may want to use a learning algorithm to choose their strategies.

No-regret learning has emerged as an attractive alternative to Nash equilibrium. The notion of a no-regret learning outcome, generalizes Nash equilibrium by requiring the no-regret property of a Nash equilibrium, in an approximate sense, but more importantly without the assumption that player strategies are stable and independent. When considering a sequence of repeated plays, having no-regret means that the total value for the player over a sequence of plays is not much worse than the value he/she would have obtained had he played the best single strategy throughout the time, where the best possible single strategy is determined with hindsight based on the environment and the actions of the other agents. There are many well-known, natural no-regret learning algorithms, such as the weighted majority algorithm [Arora et al. 2012], [Littlestone and Warmuth 1994] (also known as Hedge [Freund and Schapire 1999]) and regret matching [S. Hart and Mas-Colell 2000] just to name a few simple ones. We propose a theory of inference of agent valuations just based on the assumption that the agent's learning strategies are smart enough that they have minimal regret, without making any assumptions on the particular no-regret algorithms they employ.

There is a growing body of results in the algorithmic game theory literature characterizing properties of no-regret learning outcomes in games, such as approximate efficiency with respect to the welfare optimal outcome (see e.g. [Roughgarden 2012; Syrgkanis and Tardos 2013]). For instance, [Caragiannis et al. 2015] consider the generalized second price auction in this framework, the auction also considered in our paper and showed that the average welfare of any no-regret learning outcome is always at least 30% of the optimal welfare. To be able to apply such theoretical results on real data and to quantify the true inefficiency of GSP in the real world under learning behavior, we first need a way to infer player valuations without relying on the equilibrium assumption.

*Our contribution.* The main goal of this paper is to develop a theory of value inference from observed data of repeated strategic interactions without relying on a Nash equilibrium assumption. Rather than relying on the stability of outcomes, we make the weaker assumption that players are using some form of no-regret learning. In a stable Nash equilibrium outcome, players choose strategies independently, and their choices are best response to the environment and the strategies of other players. Choosing a best response implies that the players have no regret for alternate strategy options. We make the analogous assumption, that in a sequence of play, players have small regret for any fixed strategy.

Our results do not rely on the assumption that the participating players have correct beliefs regarding the future actions of their opponents or that they even can correctly compute expectations over the future values of state variables. Our only assumption is that the players understand the rules of the game. This is significantly different from most of the current results in Industrial Organization where estimation of dynamic games requires that the players have correct beliefs regarding the actions of their opponents. This is especially important to the analysis of bidder behavior in sponsored search auctions, which are the core application of this paper. The sponsored search marketplace is highly dynamic and volatile where the popularity of different search terms is changing, and the auction platform continuously runs experiments. In this environment advertisers continuously create new ads (corresponding to new bids), remove under-performing ads, learning what is the best way to bid while participating in the game. In this setting the assumption of players who have correct beliefs regarding their opponents and whose bids constitute an equilibrium may not be realistic.

When inferring player values from data, one needs to always accommodate small errors. In the context of players who employ learning strategies, a small error  $\epsilon > 0$  would mean that the player can have up to  $\epsilon$  regret, i.e., the utility of the player from the strategies used needs to be at most  $\epsilon$  worse than any fixed strategy with hindsight. Indeed, the guarantees provided by the standard learning algorithms is that the total utility for the player is not much worse than any single strategy with hindsight, with this error parameter decreasing over time, as the player spends more time learning. In aiming to infer the player's value from the observed data, we define the *rationalizable set*  $\mathcal{NR}$ , consisting of the set of values and error parameters  $(v, \epsilon)$  such that with value  $v$  the sequence of bid by the player would have at most  $\epsilon$  regret. We show that  $\mathcal{NR}$  is a closed convex set. Clearly, allowing for a larger error  $\epsilon$  would result in a larger set of possible values  $v$ . The most reasonable error parameter  $\epsilon$  for a player depends on the quality of his/her learning algorithm. We think of a rationalizable value  $v$  for a player as a value that is rationalizable with a small enough  $\epsilon$ .

Our main result provides a characterization of the rationalizable set  $\mathcal{NR}$  for the dynamic sponsored search auction game. We also provide a simple approach to compute this set. We demonstrate that the evaluation of  $\mathcal{NR}$  is equivalent to the evaluation of a one-dimensional function which, in turn, can be computed from the auction data

directly by sampling. This result also allows us to show how much data is needed to correctly estimate the rationalizable set for the dynamic sponsored search auction. We show that when  $N$  auction samples are observed in the data, the Hausdorff distance between the estimated and the true sets  $\mathcal{NR}$  is  $O((N^{-1} \log N)^{\gamma/(2\gamma+1)})$ , where  $\gamma$  is the sum of the number of derivatives of the allocation and pricing functions and the degree of Hölder continuity of the oldest derivative. In particular, when the allocation and pricing functions are only Lipschitz-continuous, then the total number of derivatives is zero and the constant of Hölder-continuity is 1, leading to the rate  $O((N^{-1} \log N)^{1/3})$ .

This favorably compares our result to the result in [Athey and Nekipelov 2010] where under the assumption of *static Nash equilibrium* in the sponsored search auction, the valuations of bidders can be evaluated with error margin of order  $O(N^{-1/3})$  when the empirical pricing and allocation function are Lipschitz-continuous in the bid. This means that our approach, which does not rely on the equilibrium properties, provides convergence rate which is only a  $(\log N)^{1/3}$  factor away.

In Section 6 we test our methods on a Microsoft Sponsored Search Auction data set. We show that our methods can be used to infer values on this real data, and study the empirical consequences of our value estimation. We find that typical advertisers bid a significantly shaded version of their value, shading it commonly by as much as 40%. We also observe that each advertiser’s account consists of a constant fraction of listings (e.g. bided keywords and ad pairs) that have tiny error and hence seem to satisfy the best-response assumption, whilst the remaining large fraction has an error which is spread out far from zero, thereby implying more that bidding on these listings is still in a learning transient phase. Further, we find that, among listings that appear to be in the learning phase, the relative error (the error in regret relative to the player’s value) is slightly positively correlated with the amount of shading. A higher error is suggestive of an exploration phase of learning, and is consistent with attempting larger shading of the bidder’s value, while advertisers with smaller relative regret appear to shade their value less.

*Further Related Work.* There is a rich literature in Economics that is devoted to inference in auctions based on the equilibrium assumptions. [Guerre et al. 2000] studies the estimation of values in static first-price auctions, with the extension to the cases of risk-averse bidders in [Guerre et al. 2009] and [Campo et al. 2011]. The equilibrium settings also allow inference in the dynamic settings where the players participate in a sequence of auctions, such as in [Jofre-Bonet and Pesendorfer 2003]. A survey of approaches to inference is surveyed in [Athey and Haile 2007]. These approaches have been applied to the GSP and, more generally, to the sponsored search auctions in the empirical settings in [Varian 2007] and [Athey and Nekipelov 2010].

The deviation from the “equilibrium best response” paradigm is much less common in any empirical studies. A notable exception is [Haile and Tamer 2003] where the authors use two assumptions to bound the distribution of values in the first-price auction. The first assumption is that the bid should not exceed the value. That allows to bound the order statistics of the value distribution from below by the corresponding order statistics of the observed bid distribution. The second assumption is that there is a minimum bid increment and if a bidder was outbid by another bidder then her value is below the bid that made her drop out by at least the bid increment. That allows to provide the bound from above. The issue with these bounds is that in many practical instances they are very large and they do not directly translate to the dynamic settings. The computation of the set of values compatible with the model may be complicated even in the equilibrium settings. For instance, in [Aradillas-López et al. 2013] the authors consider estimation of the distribution of values in the ascending auction when the distribution of values may be correlated. It turns out that even if we

assume that the bidders have perfect beliefs in the static model with correlated values, the inference problem becomes computationally challenging.

## 2. NO-REGRET LEARNING AND SET INFERENCE

Consider a game  $G$  with a set  $N$  of  $n$  players. Each player  $i$  has a strategy space  $B_i$ . The utility of a player depends on the strategy profile  $\mathbf{b} \in B_1 \times \dots \times B_n$ , on a set of parameters  $\theta \in \Theta$  that are observable in the data and on private parameters  $v_i \in \mathcal{V}_i$  observable only by each player  $i$ . We denote with  $U_i(\mathbf{b}; \theta, v_i)$  the utility of a player  $i$ .

We consider a setting where game  $G$  is played repeatedly. At each iteration  $t$ , each player picks a strategy  $b_i^t$  and nature picks a set of parameters  $\theta^t$ . The private parameter  $v_i$  of each player  $i$  remains fixed throughout the sequence. We will denote with  $\{\mathbf{b}\}_t$  the sequence of strategy profiles and with  $\{\theta\}_t$  the sequence of nature's parameters. We assume that the sequence of strategy profiles  $\{\mathbf{b}\}_t$  and the sequence of nature's parameters  $\{\theta\}_t$  are observable in the data. However, the private parameter  $v_i$  for each player  $i$  is not observable. The inference problem we consider in this paper is the problem of inferring these private values from the observable data.

We will refer to the two observable sequences as the *sequence of play*. In order to be able to infer anything about agent's private values, we need to make some rationality assumption about the way agents choose their strategies. Classical work of inference assumes that each agents best response to the statistical properties of the observable environment and the strategies of other players, in other words assumes that game play is at a stable Nash equilibrium. In this paper, we replace this equilibrium assumption with the weaker *no-regret* assumption, stating that the utility obtained throughout the sequence is at least as high as any single strategy  $b_i$  would have yielded, if played at every time step. If the play is stable throughout the sequence, no-regret is exactly the property required for the play to be at Nash equilibrium. However, no-regret can be reached by many natural learning algorithms without prior information, which makes this a natural rationality assumption for agents who are learning how to best play in a game while participating. More formally, we make no assumption of what agents do to learn, but rather will assume that agents learn well enough to satisfy the following no-regret property with a small error.

A sequence of play that we observe has  $\epsilon_i$ -regret for advertiser  $i$  if:

$$\forall b' \in B_i : \frac{1}{T} \sum_{t=1}^T U_i(\mathbf{b}^t; \theta^t, v_i) \geq \frac{1}{T} \sum_{t=1}^T U_i(b', \mathbf{b}_{-i}^t; \theta^t, v_i) - \epsilon_i \quad (1)$$

This leads to the following definition of a *rationalizable set under no-regret learning*.

**Definition 2.1 (Rationalizable Set).** A pair  $(\epsilon_i, v_i)$  of a value  $v_i$  and error  $\epsilon_i$  is a rationalizable pair for player  $i$  if it satisfies Equation (1). We refer to the set of such pairs as the *rationalizable set* and denote it with  $\mathcal{NR}$ .

The rationality assumption of the inequality (1) models players who may be learning from the experience while participating in the game. We assume that the strategies  $b_i^t$  and nature's parameters  $\theta^t$  are picked simultaneously, so agent  $i$  cannot pick his strategy dependent on the state of nature  $\theta^t$  or the strategies of other agents  $b_{i-1}^t$ . This makes the standard of a single best strategy  $b_i$  natural, as chosen strategies cannot depend on  $\theta^t$  or  $b_{i-1}^t$ . Beyond this, we do not make any assumption on what information is available for the agents, and how they choose their strategies. Some learning algorithms achieve this standard of learning with very minimal feedback (only the value experienced by the agent as he/she participates). If the agents know a distribution of possible nature parameters  $\theta$  or is able to observe the past values of the parameters  $\theta^t$  or the strategies  $b_{i-1}^t$  (or both), and then they can use this observed past information

to select their strategy at time step  $t$ . Such additional information is clearly useful in speeding up the learning process for agents. We will make no assumption on what information is available for agents for learning, or what algorithms they may be using to update their strategies. We will simply assume that they use algorithms that achieve the no-regret (small regret) standard expressed in inequality (1).

For general games and general private parameter spaces, the *rationalizable set* can be an arbitrary set with no good statistical or convexity properties. Our main result is to show that for the game of sponsored search auction we are studying in this paper, the set is convex and has good convergence properties in terms of estimation error.

### 3. SPONSORED SEARCH AUCTIONS MODEL

We consider data generated by advertisers repeatedly participating in sponsored search auction. The game  $G$  that is being repeated at each stage is an instance of a generalized second price auction triggered by a search query.

The rules of each auction are as follows<sup>1</sup>: Each advertiser  $i$  is associated with a click probability  $\gamma_i$  and a scoring coefficient  $s_i$  and is asked to submit a bid-per-click  $b_i$ . Advertisers are ranked by their rank-score  $q_i = s_i \cdot b_i$  and allocated positions in decreasing order of rank-score as long as they pass a rank-score reserve  $r$ . If advertisers also pass a higher mainline reserve  $m$ , then they may be allocated in the positions that appear in the mainline part of the page, but at most  $k$  advertisers are placed on the mainline.

If advertiser  $i$  is allocated position  $j$ , then he is clicked with some probability  $p_{ij}$ , which we will assume to be separable into a part  $\alpha_j$  depending on the position and a part  $\gamma_i$  depending on the advertiser, and that the position related effect is the same in all the participating auctions:

$$p_{ij} = \alpha_j \cdot \gamma_i \quad (2)$$

We denote with  $\alpha = (\alpha_1, \dots, \alpha_m)$  the vector of position coefficients. All the mentioned sets of parameters  $\theta = (s, \mathbf{b}, \gamma, r, m, k, \alpha)$  are observable in the data.

If advertiser  $i$  is allocated position  $j$ , then he pays only when he is clicked and his payment, i.e. his cost-per-click (CPC) is the minimal bid he had to place to keep his position, which is:

$$c_{ij}(\mathbf{b}; \theta) = \max \left\{ \frac{s_{\pi(j+1)} \cdot b_{\pi(j+1)}}{s_i}, \frac{r}{s_i}, \frac{m}{s_i} \cdot \mathbf{1}\{j \in M\} \right\} \quad (3)$$

where by  $\pi(j)$  we denote the advertiser that is allocated position  $j$  and with  $M$  we denote the set of mainline positions.

We also assume that each advertiser has a value-per-click (VPC)  $v_i$ , which is not observed in the data. If under a bid profile  $\mathbf{b}$ , advertiser  $i$  is allocated slot  $\sigma_i(\mathbf{b})$ , his expected utility is:

$$U_i(\mathbf{b}; \theta, v_i) = \alpha_{\sigma_i(\mathbf{b})} \cdot \gamma_i \cdot (v_i - c_{i\sigma_i(\mathbf{b})}(\mathbf{b}; \theta)) \quad (4)$$

We will denote with:

$$P_i(\mathbf{b}; \theta) = \alpha_{\sigma_i(\mathbf{b})} \cdot \gamma_i \quad (5)$$

the probability of a click as a function of the bid profile and with:

$$C_i(\mathbf{b}; \theta) = \alpha_{\sigma_i(\mathbf{b})} \cdot \gamma_i \cdot c_{i\sigma_i(\mathbf{b})}(\mathbf{b}; \theta) \quad (6)$$

the expected payment as a function of the bid profile. Then the utility of advertiser  $i$  at each auction is:

$$U_i(\mathbf{b}; \theta, v_i) = v_i \cdot P_i(\mathbf{b}) - C_i(\mathbf{b}) \quad (7)$$

<sup>1</sup>ignoring small details that we will ignore and are rarely in effect

The latter fits in the general repeated game framework, where the strategy space  $B_i = \mathbb{R}_+$  of each player  $i$  is simply any non-negative real number. The private parameter of a player  $v_i$  is an advertiser’s value-per-click (VPC) and the set of parameters that affect a player’s utility at each auction and are observable in the data is  $\theta$ . At each auction  $t$  in the sequence the observable parameters  $\theta^t$  can take arbitrary values that depend on the specific auction. However, we assume that the per-click value of the advertiser remains fixed throughout the sequence.

*Batch Auctions.* Rather than viewing a single auction as a game that is being repeated, we will view a batch of many auctions as the game that is repeated in each stage. This choice is reasonable, as it is impossible for advertisers to update their bid after each auction. Thus the utility of a single stage of the game is simply the average of the utilities of all the auctions that the player participated in during this time period. Another way to view this setting is that the parameter  $\theta$  at each iteration  $t$  is not deterministic but rather is drawn from some distribution  $D^t$  and a player’s utility at each iteration is the expected utility over  $D^t$ . In effect, the distribution  $D^t$  is the observable parameter, and utility depends on the distribution, and not only on a single draw from the distribution. With this in mind, the per-period probability of click is

$$P_i^t(\mathbf{b}^t) = \mathbb{E}_{\theta \sim D^t}[P_i(\mathbf{b}^t; \theta)] \quad (8)$$

and the per-period cost is

$$C_i^t(\mathbf{b}^t) = \mathbb{E}_{\theta \sim D^t}[C_i(\mathbf{b}^t; \theta)]. \quad (9)$$

We will further assume that the volume of traffic is the same in each batch, so the average utility of an agent over a sequence of stages is expressed by

$$\frac{1}{T} \sum_{t=1}^T (v_i \cdot P_i^t(\mathbf{b}^t) - C_i^t(\mathbf{b}^t)). \quad (10)$$

Due to the large volume of auctions that can take place in between these time-periods of bid changes, it is sometimes impossible to consider all the auctions and calculate the true empirical distribution  $D^t$ . Instead it is more reasonable to approximate  $D^t$ , by taking a sub-sample of the auctions. This would lead only to statistical estimates of both of these quantities. We will denote these estimates by  $\hat{P}_i^t(\mathbf{b})$  and  $\hat{C}_i^t(\mathbf{b})$  respectively. We will analyze the statistical properties of the estimated rationalizable set under such subsampling in Section 5.

#### 4. PROPERTIES OF RATIONALIZABLE SET FOR SPONSORED SEARCH AUCTIONS

For the auction model that we are interested in, Equation (1) that determines whether a pair  $(\epsilon, v)$  is rationalizable boils down to:

$$\forall b' \in \mathbb{R}_+ : v \cdot \frac{1}{T} \sum_{t=1}^T (P_i^t(b', \mathbf{b}_{-i}^t) - P_i^t(\mathbf{b}^t)) \leq \frac{1}{T} \sum_{t=1}^T (C_i^t(b', \mathbf{b}_{-i}^t) - C_i^t(\mathbf{b}^t)) + \epsilon \quad (11)$$

If we denote with

$$\Delta P(b') = \frac{1}{T} \sum_{t=1}^T (P_i^t(b', \mathbf{b}_{-i}^t) - P_i^t(\mathbf{b}^t)), \quad (12)$$

the increase in the average probability of click from player  $i$  switching to a fixed alternate bid  $b'$  and with

$$\Delta C(b') = \frac{1}{T} \sum_{t=1}^T (C_i^t(b', \mathbf{b}_{-i}^t) - C_i^t(\mathbf{b}^t)), \quad (13)$$

the increase in the average payment from player  $i$  switching to a fixed alternate bid  $b'$ , then the condition simply states:

$$\forall b' \in \mathbb{R}_+ : v \cdot \Delta P(b') \leq \Delta C(b') + \epsilon \quad (14)$$

Hence, the rationalizable set  $\mathcal{NR}$  is an envelope of the family of half planes obtained by varying  $b \in \mathbb{R}_+$  in Equation (14).

We conclude this section by characterizing some basic properties of this set, which will be useful both in analyzing its statistical estimation properties and in computing the empirical estimate of  $\mathcal{NR}$  from the data.

LEMMA 4.1. *The set  $\mathcal{NR}$  is a closed convex set.*

PROOF. The Lemma follows by the fact that  $\mathcal{NR}$  is defined by a set of linear constraints. Any convex combination of two points in the set will also satisfy these constraints, by taking the convex combination of the constraints that the two points have to satisfy. The closedness follows from the fact that points that satisfy the constraints with equality are included in the set.  $\square$

LEMMA 4.2. *For any error level  $\epsilon$ , the set of values that belong to the rationalizable set is characterized by the interval:*

$$v \in \left[ \max_{b': \Delta P(b') < 0} \frac{\Delta C(b') + \epsilon}{\Delta P(b')}, \min_{b': \Delta P(b') > 0} \frac{\Delta C(b') + \epsilon}{\Delta P(b')} \right] \quad (15)$$

*In particular the data are not rationalizable for error level  $\epsilon$  if the latter interval is empty.*

PROOF. Follows from Equation (14), by dividing with  $\delta P(b')$  and taking cases on whether  $\Delta P(b')$  is positive or negative.  $\square$

To be able to estimate the rationalizable set  $\mathcal{NR}$  we will need to make a few additional assumptions. All assumptions are natural, and likely satisfied by any data. Further, the assumptions are easy to test for in the data, and the parameters needed can be estimated.

Since  $\mathcal{NR}$  is a closed convex set, it is conveniently represented by the set of support hyperplanes defined by the functions  $P_i^t(\cdot, \cdot)$  and  $C_i^t(\cdot, \cdot)$ . Our first assumption is on the functions  $P_i^t(\cdot, \cdot)$  and  $C_i^t(\cdot, \cdot)$ .

ASSUMPTION 1. *The support of bids is a compact set  $B = [0, \bar{b}]$ . For each bidvector  $\mathbf{b}_{-i}^t$  the functions  $P_i^t(\cdot, \mathbf{b}_{-i}^t)$  and  $C_i^t(\cdot, \mathbf{b}_{-i}^t)$  are continuous, monotone increasing and bounded on  $B$ .*

*Remark.* These assumptions are natural and are satisfied by our data. In most applications, it is easy to have a simple upper bound  $M$  on the maximum conceivable value an advertiser can have for a click, and with this maximum, we can assume that the bids are also limited to this maximum, so can use  $B = M$  as a bound for the maximum bid. The probability of getting a click as well as the cost per-click are clearly increasing function of the agent's bid in the sponsored search auctions considered in this paper, and the continuity of these functions is a good model in large or uncertain environments.

We note that properties in Assumption 1 implies that the same is also satisfied by the linear combinations of functions, as a linear combination of monotone functions is monotone, implying that the functions  $\Delta P(b)$  and  $\Delta C(b)$  are also monotone and continuous.



The assumption further implies that for any value  $v$ , there exists at least one element of  $B$  that maximizes  $v\Delta P(b) - \Delta C(b)$ , as  $v\Delta P(b) - \Delta C(b)$  is a continuous function on a compact set.

#### 4.1. Properties of the Boundary

Now we can study the properties of the set  $\mathcal{NR}$  by characterizing its boundary, denoted  $\partial\mathcal{NR}$ .

**LEMMA 4.3.** *Under Assumption 1,  $\partial\mathcal{NR} = \{(v, \epsilon) : \sup_b (v\Delta P(b) - \Delta C(b)) = \epsilon\}$*

**PROOF.** (1) Suppose that  $(v, \epsilon)$  solves  $\sup_b (v\Delta P(b) - \Delta C(b)) = \epsilon$ . Provided the continuousness of functions  $\Delta P(b)$  and  $\Delta C(b)$  the supremum exists and it is attained at some  $b^*$  in the support of bids. Take some  $\delta > 0$ . Taking point  $(v', \epsilon')$  such that  $v' = v$  and  $\epsilon' = \epsilon - \delta$  ensures that  $v'\Delta P(b^*) - \Delta C(b^*) > \epsilon'$  and thus  $(v', \epsilon') \notin \mathcal{NR}$ . Taking the point  $(v'', \epsilon'')$  such that  $v'' = v$  and  $\epsilon'' = \epsilon + \delta$  ensures that

$$\sup_b (v''\Delta P(b) - \Delta C(b)) < \epsilon'',$$

and thus for any bid  $b$ :  $v''\Delta P(b) - \Delta C(b) < \epsilon''$ . Therefore  $(v'', \epsilon'') \in \mathcal{NR}$ . Provided that  $\delta$  was arbitrary, this ensures that in any any neighborhood of  $(v, \epsilon)$  there are points that both belong to set  $\mathcal{NR}$  and that do not belong to it. This means that this point belongs to  $\partial\mathcal{NR}$ .

(2) We prove the sufficiency part by contradiction. Suppose that  $(v, \epsilon) \in \partial\mathcal{NR}$  and  $\sup_b (v\Delta P(b) - \Delta C(b)) \neq \epsilon$ . If  $\sup_b (v\Delta P(b) - \Delta C(b)) > \epsilon$ , provided that the objective function is bounded and has a compact support, there exists a point  $b^*$  where the supremum is attained. For this point  $v\Delta P(b^*) - \Delta C(b^*) > \epsilon$ . This means that for this bid the imposed inequality constraint is not satisfied and this point does not belong to  $\mathcal{NR}$ , which is the contradiction to our assumption.

Suppose then that  $\sup_b (v\Delta P(b) - \Delta C(b)) < \epsilon$ . Let  $\Delta\epsilon = \epsilon - \sup_b (v\Delta P(b) - \Delta C(b))$ . Take some  $\delta < \Delta\epsilon$  and take an arbitrary  $\delta_1, \delta_2$  be such that  $\max\left\{|\delta_1|, \frac{|\delta_2|}{\sup_b \Delta P(b)}\right\} < \frac{\delta}{2}$ . Let  $v' = v + \delta_1$  and  $\epsilon' = \epsilon + \delta_2$ . Provided that

$$\sup_b (v\Delta P(b) - \Delta C(b)) \geq \sup_b ((v + \delta_1)\Delta P(b) - \Delta C(b)) - \delta_1 \sup_b \Delta P(b)\delta_1,$$

for any  $\tau \in (-\infty, \Delta\epsilon)$  we have

$$\sup_b (v'\Delta P(b) - \Delta C(b)) < \epsilon - \tau + \delta_1 \sup_b \Delta P(b).$$

Provided that  $|\delta_1 \sup_b \Delta P(b)| < \delta/2 < \Delta\epsilon/2$ , for any  $|\delta_2| < \delta/2$ , we can find  $\tau$  with  $-\tau + \delta_1 \sup_b \Delta P(b) = \delta_2$ . Therefore

$$\sup_b (v'\Delta P(b) - \Delta C(b)) < \epsilon',$$

and thus for any  $b$ :  $v'\Delta P(b) - \Delta C(b) < \epsilon'$ . Therefore  $(v', \epsilon') \in \mathcal{NR}$ . This means that for  $(v, \epsilon)$  we constructed a neighborhood of size  $\delta$  such that each element of that neighborhood is an element of  $\mathcal{NR}$ . Thus  $(v, \epsilon)$  cannot belong to  $\partial\mathcal{NR}$ .  $\square$

The next step will be to establish the properties of the boundary by characterizing the solutions of the optimization problem of selecting the optimal bid single  $b$  with for a given value  $v$  and the sequence of bids by other agents, defined by  $\sup_b (v\Delta P(b) - \Delta C(b))$ .

**LEMMA 4.4.** *Let  $b^*(v) = \arg \sup_b (v\Delta P(b) - \Delta C(b))$ . Then  $b^*(\cdot)$  is upper-hemicontinuous and monotone.*

PROOF. By Assumption (1) the function  $v\Delta P(b) - \Delta C(b)$  is continuous in  $v$  and  $b$ , then the upper hemicontinuity of  $b^*(\cdot)$  follows directly from the Berge's maximum theorem.

To show that  $b^*(\cdot)$  is monotone consider the function  $q(b; v) = v\Delta P(b) - \Delta C(b)$ . We'll show that this function is *supermodular* in  $(b; v)$ , that is, for  $b' > b$  and  $v' > v$  we have

$$q(b'; v') + q(b; v) \geq q(b'; v) + q(b; v').$$

To see this observe that if we take  $v' > v$ , then

$$q(b; v') - q(b; v) = (v' - v)\Delta P(b),$$

which is non-decreasing in  $b$  due to the monotonicity of  $\Delta P(\cdot)$ , implying that  $q$  is supermodular. Now we can apply the Topkis' ([Topkis 1998], [Vives 2001]) monotonicity theorem from which we immediately conclude that  $b^*(\cdot)$  is non-decreasing.  $\square$

Lemma 4.4 provides us a powerful result of the monotonicity of the optimal response function  $b^*(v)$  which only relies on the boundedness, compact support and monotonicity of functions  $\Delta P(\cdot)$  and  $\Delta C(\cdot)$  *without relying on their differentiability*. The downside of this is that  $b^*(\cdot)$  can be set valued or discontinuously change in the value. To avoid this we impose the following additional assumption.

ASSUMPTION 2. For each  $b_1$  and  $b_2 > 0$  the incremental cost per click function

$$ICC(b_2, b_1) = \frac{\Delta C(b_2) - \Delta C(b_1)}{\Delta P(b_2) - \Delta P(b_1)}$$

is continuous in  $b_1$  for each  $b_2 \neq b_1$  and it is continuous in  $b_2$  for each  $b_1 \neq b_2$ . Moreover for any  $b_4 > b_3 > b_2 > b_1$  on  $B$ :  $ICC(b_4, b_3) > ICC(b_2, b_1)$ .

*Remark.* Assumption 2 requires that there are no discounts on buying clicks “in bulk”: for any average position of the bidder in the sponsored search auction, an incremental increase in the click yield requires the corresponding increase in the payment. In other words, “there is no free clicks”. In [Athey and Nekipelov 2010] it was shown that in sponsored auctions with uncertainty and a reserve price, the condition in Assumption 2 is satisfied with the lower bound on the incremental cost per click being the maximum of the bid and the reserve price (conditional on the bidder's participation in the auction).

LEMMA 4.5. Under Assumptions 1 and 2  $b^*(v) = \arg \sup_b (v\Delta P(b) - \Delta C(b))$  is a continuous monotone function.

PROOF. Under Assumption 1 in Lemma 4.4 we established the monotonicity and upper-hemicontinuity of the mapping  $b^*(\cdot)$ . We now strengthen this result to monotonicity and continuity.

Suppose that for value  $v$ , function  $v\Delta P(b) - \Delta C(b)$  has an interior maximum and let  $B^*$  be its set of maximizers. First, we show that under Assumption 2  $B^*$  is singleton. In fact, suppose that  $b_1^*, b_2^* \in B^*$  and wlog  $b_1^* > b_2^*$ . Suppose that  $b \in (b_1^*, b_2^*)$ . First, note that  $b$  cannot belong to  $B^*$ . If it does, then

$$v\Delta P(b_1^*) - \Delta C(b_1^*) = v\Delta P(b) - \Delta C(b) = v\Delta P(b_2^*) - \Delta C(b_2^*),$$

and thus  $ICC(b, b_1^*) = ICC(b_2^*, b)$  for  $b_1^* < b < b_2^*$  which violates Assumption 2. Second if  $b \notin B^*$ , then

$$v\Delta P(b_1^*) - \Delta C(b_1^*) \geq v\Delta P(b) - \Delta C(b),$$

and thus  $v \leq ICC(b, b_1^*)$ . At the same time,

$$v\Delta P(b_2^*) - \Delta C(b_2^*) \geq v\Delta P(b) - \Delta C(b),$$

and thus  $v \geq \text{ICC}(b_2^*, b)$ . This is impossible under Assumption 2 since it requires that  $\text{ICC}(b_2^*, b) > \text{ICC}(b, b_1^*)$ . Therefore, this means that  $B^*$  is singleton.

Now consider  $v$  and  $v' = v + \delta$  for a sufficiently small  $\delta > 0$  and  $v$  and  $v'$  leading to the interior maximum of the of the objective function. By the result of Lemma 4.4,  $b^*(v') \geq b^*(v)$  and by Assumption 2 the inequality is strict. Next note that for any  $b > b^*(v)$   $\text{ICC}(b, b^*(v)) > v$  and for any  $b < b^*(v)$   $\text{ICC}(b, b^*(v)) < v$ . By continuity,  $b^*(v)$  solves  $\text{ICC}(b, b^*(v)) = v$ . Now let  $b'$  solve  $\text{ICC}(b', b^*(v)) = v'$ . By Assumption 2,  $b' > b^*(v')$  since  $b^*(v')$  solves  $\text{ICC}(b, b^*(v')) = v'$  and  $b^*(v') > b^*(v)$ . Then by Assumption 2,  $\text{ICC}(b', b^*(v)) - v > v - \text{ICC}(0, b^*(v))$ . This means that  $b' - b^*(v) < (v' - v)/(v - \text{ICC}(0, b^*(v)))$ . This means that  $|b^*(v') - b^*(v)| < |b' - b^*(v)| < \delta/(v - \text{ICC}(0, b^*(v)))$ . This means that  $b^*(v)$  is continuous in  $v$ .  $\square$

Lemma 4.5, the interior solutions maximizing  $v\Delta P(b) - \Delta C(b)$  are continuous in  $v$ . We proved that each boundary point of  $\mathcal{NR}$  corresponds to the maximant of this function. As a result, whenever the maximum is interior, then the boundary shares a point with the support hyperplane  $v\Delta P(b^*(v)) - \Delta C(b^*(v)) = \epsilon$ . Therefore, the corresponding normal vector at that boundary point is  $(P(b^*(v)), -1)$ .

#### 4.2. Support Function Representation

Our next step would be to use the derived properties of the boundary of the set  $\mathcal{NR}$  to compute it. The basic idea will be based on varying  $v$  and computing the bid  $b^*(v)$  that maximizes regret. The corresponding maximum value will deliver the value of  $\epsilon$  corresponding to the boundary point. Provided that we are eventually interested in the properties of the set  $\mathcal{NR}$  (and the quality of its approximation by the data), the characterization via the support hyperplanes will not be convenient because this characterization requires to solve the computational problem of finding an envelope function for the set of support hyperplanes. Since closed convex bounded sets are fully characterized by their boundaries, we can use the notion of the *support* function to represent the boundary of the set  $\mathcal{NR}$ .

*Definition 4.6.* The support function of a closed convex set  $X$  is defined as:

$$h(X, u) = \sup_{x \in X} \langle x, u \rangle,$$

where in our case  $X = \mathcal{NR}$  is a subset of  $\mathbb{R}^2$  or value and error pairs  $(v, \epsilon)$ , and then  $u$  is also an element of  $\mathbb{R}^2$ .

An important property of the support function is the way it characterizes closed convex bounded sets. Recall that the Hausdorff norm for subsets  $A$  and  $B$  of the metric space  $E$  with metric  $\rho(\cdot, \cdot)$  is defined as

$$d_H(A, B) = \max\{\sup_{a \in A} \inf_{b \in B} \rho(a, b), \sup_{b \in B} \inf_{a \in A} \rho(a, b)\}.$$

It turns out that  $d_H(A, B) = \sup_u |h(A, u) - h(B, u)|$ . Therefore, if we find  $h(\mathcal{NR}, u)$ , this will be equivalent to characterizing  $\mathcal{NR}$  itself.

We note that the set  $\mathcal{NR}$  is not bounded: the larger is the error  $\epsilon$  that the player can make, the wider is the range of values that will rationalize the observed bids. We consider the restriction of the set  $\mathcal{NR}$  to the set where we bound the error for the players. Moreover, we explicitly assume that the values per click of bidders have to be non-negative.

ASSUMPTION 3.

- (i) *The rationalizable values per click are non-negative:  $v \geq 0$ .*
- (ii) *There exists a global upper bound  $\bar{\epsilon}$  for the error of all players.*

*Remark.* While non-negativity of the value may be straightforward in many auction context, the assumption of an upper bound for the error may be less straightforward. One way to get such an error bound is if we assume that the values come from a bounded range  $[0, M]$  (see Remark after Assumption (1)), and then an error  $\epsilon > M$  would correspond to negative utility by the player, and the players may choose to exit from the market if their advertising results in negative value.

**THEOREM 4.7.** *Under Assumption 1, 2 the support function of  $\mathcal{NR}$  is*

$$h(\mathcal{NR}, u) = \begin{cases} +\infty, & \text{if } u_2 \geq 0, \text{ or } \frac{u_1}{|u_2|} \notin [\inf_b \Delta P(b), \sup_b \Delta P(b)], \\ |u_2| \Delta C \left( \Delta P^{-1} \left( \frac{u_1}{|u_2|} \right) \right), & \text{if } u_2 < 0 \text{ and } \frac{u_1}{|u_2|} \in [\inf_b \Delta P(b), \sup_b \Delta P(b)]. \end{cases}$$

**PROOF.** Provided that the support function is positive homogenous, without loss of generality we can set  $u = (u_1, u_2)$  with  $\|u\| = 1$ . To find the support function, we take  $u_1$  to be dual to  $v_i$  and  $u_2$  to be dual to  $\epsilon_i$ . We re-write the inequality of the half-plane as:  $v_i \cdot \Delta P(b_i) - \epsilon_i \leq \Delta C(b_i)$ . We need to evaluate the inner-product

$$u_1 v_i + u_2 \epsilon_i$$

from above. We note that to evaluate the support function for  $u_2 \geq 0$ , the corresponding inequality for the half-plane needs to “flip” to  $\geq -\Delta C(b_i)$ . This means that for  $u_2 \geq 0$ ,  $h(\mathcal{NR}, u) = +\infty$ . Next, for any  $u_2 < 0$  we note that the inequality for the half-plane can be re-written as:  $v_i |u_2| \Delta P(b_i) + u_2 \epsilon_i \leq |u_2| \Delta C(b_i)$ .

As a result if there exists  $b_i$  such that for a given  $u_1$ :  $\frac{u_1}{|u_2|} = \Delta P(b_i)$ , then  $|u_2| \Delta C(b_i)$  corresponds to the support function for this  $b_i$ .

Now suppose that  $\sup_b \Delta P(b) > 0$  and  $u_1 > |u_2| \sup_b \Delta P(b)$ . In this case we can evaluate

$$u_1 v_i + u_2 \epsilon_i = (u_1 - |u_2| \sup_b \Delta P(b)) v_i + |u_2| \sup_b \Delta P(b) v_i + u_2 \epsilon_i.$$

Function  $|u_2| \sup_b \Delta P(b) v_i + u_2 \epsilon_i$  is bounded by  $|u_2| \sup_b \Delta C(b)$  for each  $(v_i, \epsilon_i) \in \mathcal{NR}$ . At the same time, function  $(u_1 - |u_2| \sup_b \Delta P(b)) v_i$  is strictly increasing in  $v_i$  on  $\mathcal{NR}$ . As a result, the support function evaluated at any vector  $u$  with  $u_1/|u_2| > \sup_b \Delta P(b)$  is  $h(\mathcal{NR}, u) = +\infty$ .  $\square$

This behavior of the support function can be explained intuitively. The set  $\mathcal{NR}$  is a convex set in  $\epsilon > 0$  half-space. The unit-length  $u$  corresponds to the normal vector to the boundary of  $\mathcal{NR}$  and  $-h(\mathcal{NR}, u)$  is the point of intersection of the  $\epsilon$  axis by the tangent line. Asymptotically, the boundaries of the set will approach to the lines  $v_i \sup_b \Delta P(b) - \epsilon \leq \sup_b \Delta C(b)$  and  $v_i \inf_b \Delta P(b) - \epsilon \leq \inf_b \Delta C(b)$ . First of all, this means that the projection of  $u_2$  coordinate of the normal vector of the line that can be tangent to  $\mathcal{NR}$  has to be negative. If that projection is positive, that line can only intersect the set  $\mathcal{NR}$ . Second, the maximum slope of the normal vector to the tangent line is  $\sup_b \Delta P(b)$  and the minimum is  $\inf_b \Delta P(b)$ . Any line with a steeper slope will intersect the set  $\mathcal{NR}$ .

Now we consider the restriction of the set  $\mathcal{NR}$  via Assumption 3. The additional restriction on what is rationalizable does not change its convexity, but it makes the resulting set bounded. Denote this bounded subset of  $\mathcal{NR}$  by  $\mathcal{NR}_B$ . The following theorem characterizes the structure of the support function of this set.

**THEOREM 4.8.** *Let  $\bar{v}$  correspond to the furthest point of the set  $\mathcal{NR}_B$  from the  $\epsilon$  axis. Suppose that  $\bar{b}$  is such that  $\bar{v} \Delta P(\bar{b}) - \bar{\epsilon} = \Delta C(\bar{b})$ . Also suppose that  $\underline{b}$  is the point of intersection of the boundary of the set  $\mathcal{NR}$  with the vertical axis  $v = 0$ . Under Assumptions*

1, 2, and 3 the support function of  $\mathcal{NR}_B$  is

$$h(\mathcal{NR}_B, u) = \begin{cases} h(\mathcal{NR}, u), & \text{if } u_2 < 0, \Delta P(\underline{b}) < u_1/|u_2| < \Delta P(\bar{b}), \\ u_1\bar{v} + u_2\bar{\epsilon}, & \text{if } u_2 < 0, u_1/|u_2| > \Delta P(\bar{b}), \\ u_2\underline{\epsilon}, & \text{if } u_2 < 0, \Delta P(\underline{b}) > u_1/|u_2|, \\ u_1\bar{v} + u_2\bar{\epsilon}, & \text{if } u_1, u_2 \geq 0, \\ u_2\bar{\epsilon}, & \text{if } u_1 < 0, u_2 \geq 0, \end{cases}$$

PROOF. We begin with  $u_2 < 0$ . In this case when  $\Delta P(\underline{b}) < u_1/|u_2| < \Delta P(\bar{b})$ , then the corresponding normal vector  $u$  is on the part of the boundary of the set  $\mathcal{NR}_B$  that coincides with the boundary of  $\mathcal{NR}$ . And thus for this set of normal vectors  $h(\mathcal{NR}_B, u) = h(\mathcal{NR}, u)$ .

Suppose that  $u_2 < 0$  while  $u_1/|u_2| \geq \Delta P(\bar{b})$ . In this case the support hyperplane is centered at point  $(\bar{v}, \bar{\epsilon})$  for the entire range of angles of the normal vectors. Provided that the equation for each such a support hyperplane corresponds to  $\frac{u_1}{|u_2|}v - \epsilon = c$  and each needs pass through  $(\bar{v}, \bar{\epsilon})$ , then the support function is  $h(\mathcal{NR}_B, u) = u_1\bar{v} + u_2\bar{\epsilon}$ .

Suppose that  $u_2 < 0$  and  $u_1/|u_2| < \Delta P(\underline{b})$ . Then the support hyperplanes will be centered at point  $(0, \underline{v})$ . The corresponding support function can be expressed as  $h(\mathcal{NR}_B, u) = u_2\underline{\epsilon}$ .

Now suppose that  $u_2 \geq 0$  and  $u_1 > 0$ . Then the support hyperplanes are centered at  $(\bar{v}, \bar{\epsilon})$  and again the support function will be  $h(\mathcal{NR}_B, u) = u_1\bar{v} + u_2\bar{\epsilon}$ .

Finally, suppose that  $u_2 \geq 0$  and  $u_1 \leq 0$ . Then the support hyperplanes are centered at  $(0, \bar{\epsilon})$ . The corresponding support function is  $h(\mathcal{NR}_B, u) = u_2\bar{\epsilon}$ .  $\square$

## 5. INFERENCE FOR RATIONALIZABLE SET

Note that to construct the support function (and thus fully characterize the set  $\mathcal{NR}_B$ ) we only need to evaluate the function  $\Delta C(\Delta P^{-1}(\cdot))$ . It is one-dimensional function and can be estimated from the data via direct simulation. We have previously noted that our goal is to characterize the distance between the true set  $\mathcal{NR}_B$  and the set  $\widehat{\mathcal{NR}}$  that is obtained from subsampling the data. Denote

$$f(\cdot) = \Delta C(\Delta P^{-1}(\cdot))$$

and let  $\hat{f}(\cdot)$  be its estimate from the data. The set  $\mathcal{NR}_B$  is characterized by its support function  $h(\mathcal{NR}_B, u)$  which is determined by the exogenous upper bound on the error  $\bar{\epsilon}$ , the intersection point of the boundary of  $\mathcal{NR}$  with the vertical axis  $(0, \underline{\epsilon})$ , and the highest rationalizable value  $\bar{v}$ .

We note that the set  $\mathcal{NR}$  lies inside the shifted cone defined by half-spaces  $v \sup_b \Delta P(b) - \epsilon \leq \sup_b \Delta C(b)$  and  $v \inf_b \Delta P(b) - \epsilon \leq \inf_b \Delta C(b)$ . Thus the value  $\bar{v}$  can be upper bounded by the intersection of the line  $v \sup_b \Delta P(b) - \epsilon = \sup_b \Delta C(b)$  with  $\epsilon = \bar{\epsilon}$ . The support function corresponding to this point can therefore be upper bounded by  $|u_2| \sup_b \Delta C(b)$ .

Then we notice that

$$d_H(\widehat{\mathcal{NR}}_B, \mathcal{NR}_B) = \sup_{\|u\|=1} |h(\widehat{\mathcal{NR}}_B, u) - h(\mathcal{NR}_B, u)|.$$

For the evaluation of the sup-norm of the difference between the support function, we split the circle  $\|u\| = 1$  to the areas where the support function is linear and non-linear determined by the function  $f(\cdot)$ . For the non-linear segment the sup-norm can be upper

bounded by the global sup-norm

$$\begin{aligned} \sup_{u_2 < 0, \Delta P(\bar{b}) < u_1/|u_2| < \Delta P(\bar{b})} |h(\widehat{\mathcal{N}R}_B, u) - h(\mathcal{N}R_B, u)| &\leq \sup_{\|u\|=1} \left| |u_2| \widehat{f}\left(\frac{u_1}{|u_2|}\right) - |u_2| f\left(\frac{u_1}{|u_2|}\right) \right| \\ &\leq \sup_z \left| \widehat{f}(z) - f(z) \right|. \end{aligned}$$

For the linear part, provided that the value  $\bar{\epsilon}$  is fixed, we can evaluate the norm from above by

$$\sup_{\|u\|=1} |u_1 \sup_b \Delta \widehat{C}(b) - u_1 \Delta C(b)| \leq |\sup_b \Delta \widehat{C}(b) - \Delta C(b)| \leq \sup_z \left| \widehat{f}(z) - f(z) \right|.$$

Thus we can evaluate

$$d_H(\widehat{\mathcal{N}R}_B, \mathcal{N}R_B) \leq \sup_z \left| \widehat{f}(z) - f(z) \right|. \quad (16)$$

Thus, the bounds that can be established for estimation of function  $\widehat{f}$  directly imply the bounds for estimation of the Hausdorff distance between the estimated and the true sets  $\mathcal{N}R_B$ . We assume that a sample of size  $N = n \times T$  is available (where  $n$  is the number of auctions sampled per period and  $T$  is the number of periods). We now establish the general rate result for the estimation of the set  $\mathcal{N}R_B$ .

**THEOREM 5.1.** *Suppose that function  $f$  has derivative up to order  $k \geq 0$  and for some  $L \geq 0$*

$$|f^{(k)}(z_1) - f^{(k)}(z_2)| \leq L|z_1 - z_2|^\alpha.$$

*Under Assumptions 1, 2, and 3 we have*

$$d_H(\widehat{\mathcal{N}R}_B, \mathcal{N}R_B) \leq O((N^{-1} \log N)^{\gamma/(2\gamma+1)}), \quad \gamma = k + \alpha.$$

*Remark.* The theorem makes a further assumption the function  $f$  is  $k$  times differentiable, and satisfies a Lipschitz style bound with parameter  $L \geq 0$  and exponent  $\alpha$ . We note that this theorem if we take the special case of  $k = 0$ , the theorem does not require differentiability of functions  $\Delta P(\cdot)$  and  $\Delta C(\cdot)$ . If these functions are Lipschitz, the condition of the theorem is satisfied with  $k = 0$  and  $\alpha = 1$ , and the theorem provides a  $O((N^{-1} \log N)^{1/3})$  convergence rate for the estimated set  $\mathcal{N}R$ .

**PROOF.** By (16) the error in the estimation of the set  $\mathcal{N}R_B$  is fully characterized by the uniform error in the estimation of the single-dimensional function  $f(\cdot)$ . In part (i) of the Theorem our assumption is that we estimate function  $f(\cdot)$  from the class of convex functions. In part (ii) of the Theorem our assumption is that we estimate the function from the class of  $k$  times differentiable single-dimensional functions. We now use the results for *optimal global convergence rates* for estimation of functions from these respective classes. We note that these rates do not depend on the particular chosen estimation method and thus provide a global bound on the convergence of the estimated set  $\widehat{\mathcal{N}R}_B$  to the true set. By [Stone 1982], the global uniform convergence rate for estimation of the unknown function with  $k$  derivatives where  $k$ -th derivative is Hölder-continuous with constant  $\alpha$  is  $(N^{-1} \log N)^{\gamma/(2\gamma+1)}$  with  $\gamma = k + \alpha$ . That delivers our statement.  $\square$

We note that this theorem does not require differentiability of functions  $\Delta P(\cdot)$  and  $\Delta C(\cdot)$ . For instance, if these functions are Lipschitz, the theorem provides a  $O((N^{-1} \log N)^{1/3})$  convergence rate for the estimated set  $\mathcal{N}R$ .

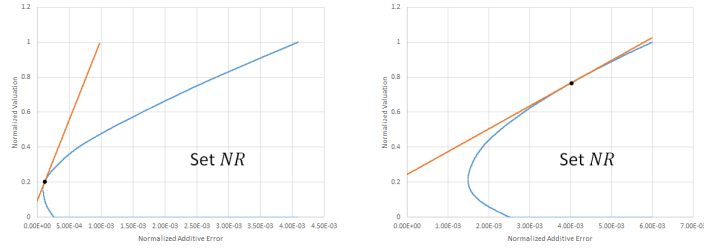


Fig. 1.  $\mathcal{NR}$  set for two listings of a high-frequency bid changing advertiser. Values are normalized to 1. The tangent line selects our point prediction.

## 6. DATA ANALYSIS

We apply our approach to infer the valuations and regret constants of a set of advertisers from the search Ads system of Bing. Our focus is on advertisers who change their bids frequently (up to multiple bid changes per hour) and thus are more prone to use algorithms for their bid adjustment instead of changing those bids manually. Each advertiser corresponds to an “account”. Accounts have multiple listings corresponding to individual ads. The advertisers can set the bids for each listing within the account separately. We examine nine high frequency accounts from the same industry sector for a period of a week and apply our techniques to all the listings within each account. The considered market is highly dynamic where new advertising campaigns are launched on the daily basis (while there is also a substantial amount of experimentation on the auction platform side, that has a smaller contribution to the uncertainty regarding the market over the uncertainty of competitors’ actions.) We focus on analyzing the bid dynamics across the listings within the same account as they are most probably instances of the same learning algorithm. Hence the only thing that should be varying for these listings is the bidding environment and the valuations. Therefore, statistics across such listings, capture in some sense the statistical behavior of the learning algorithm when the environment and the valuation per bid is drawn from some distribution.

*Computation of the Empirical Rationalizable Set.* We first start by briefly describing the procedure that we constructed to compute the set  $\mathcal{NR}$  for a single listing. We assumed that bids and values have a hard absolute upper bound and since they are constrained to be multiples of pennies, the strategy space of each advertiser is now a finite set, rather than the set  $\mathbb{R}_+$ . Thus for each possible deviating bid  $b'$  in this finite set we compute the  $\Delta P(b')$  and  $\Delta C(b')$  for each listing. We then discretize the space of additive errors. For each additive error  $\epsilon$  in this discrete space we use the characterization in Lemma 4.2 to compute the upper and lower bound on the value for this additive error. This involves taking a maximum and a minimum over a finite set of alternative bids  $b'$ . We then look at the smallest epsilon  $\epsilon_0$ , where the lower bound is smaller than the upper bound and this corresponds to the smallest rationalizable epsilon. For every epsilon larger than  $\epsilon_0$  we plot the upper bound function and the lower bound function.

An example of the resulting set  $\mathcal{NR}$  for a high frequency listing of one of the advertisers we analyzed is depicted in Figure 1. From the figure, we observe that the right listing has a higher regret than the left one. Specifically, the smallest rationalizable additive error is further from zero. Upon the examination of the bid change pattern, the right listing in the Figure was in a more transient period where the bid of the advertiser was increasing, hence this could have been a period where the advertiser was experimenting. The bid pattern of the first listing was more stable.

*Point Prediction: Smallest Multiplicative Error.* Since the two dimensional rationalizable set  $\mathcal{NR}$  is hard to depict as summary statistic from a large number of listings, we instead opt for a point prediction and specifically we compute the point of the  $\mathcal{NR}$  set that has the smallest regret error.

Since the smallest possible additive error is hard to compare across listings, we instead pick the smallest multiplicative error that is rationalizable, i.e. a pair  $(\delta, v)$  such that:

$$\forall b' \in B_i : \frac{1}{T} \sum_{t=1}^T U_i(\mathbf{b}^t; \theta^t, v_i) \geq (1 - \delta) \frac{1}{T} \sum_{t=1}^T U_i(b', \mathbf{b}_{-i}^t; \theta^t, v_i) \quad (17)$$

and such that  $\delta$  is the smallest possible value that admits such a value per click  $v$ . Denote  $P^0 = \frac{1}{T} \sum_{t=1}^T P_i^t(\mathbf{b}^t)$  the observed average probability of click of the advertiser and with  $C^0 = \frac{1}{T} \sum_{t=1}^T C_i^t(\mathbf{b}^t)$  the observed average cost, then by simple re-arrangements the latter constraint is equivalent to:

$$\forall b' \in B_i : v \Delta P(b) \leq \Delta C(b) + \frac{\delta}{1-\delta} (vP^0 - C^0) \quad (18)$$

By comparing this result to Equation (14), a multiplicative error of  $\delta$  corresponds to an additive error of  $\epsilon = \frac{\delta}{1-\delta} (vP_i^0 - C_i^0)$ .

Hence, one way to compute the feasible region of values for a multiplicative error  $\delta$  from the  $\mathcal{NR}$  set that we estimated is to draw a line of  $\epsilon = \frac{\delta}{1-\delta} (vP_i^0 - C_i^0)$ . The two points where this line intersects the  $\mathcal{NR}$  set correspond to the upper and lower bound on the valuation. Then the smallest multiplicative error  $\delta$ , corresponds to the line that simply touches the  $\mathcal{NR}$  set, i.e. the smallest  $\delta$  for which the intersection of the line  $\epsilon = \frac{\delta}{1-\delta} (vP_i^0 - C_i^0)$  is non-empty. This line is depicted in orange in Figure 1 and the point appears with a black mark.

Computationally, we estimate this point by binary search on the values of  $\delta \in [0, 1]$ , until the upper and lower point of  $\delta$  in the binary search is smaller than some pre-defined precision. Specifically, for each given  $\delta$ , the upper and lower bound of the values implied by the constraints in Equation 18 is:

$$b': (1-\delta)\Delta P(b') - \delta P^0 > 0 \quad \frac{(1-\delta)\Delta C(b') - \delta C^0}{(1-\delta)\Delta P(b') - \delta P^0} \leq v \leq \frac{(1-\delta)\Delta C(b') - \delta C^0}{(1-\delta)\Delta P(b') - \delta P^0} \quad b': (1-\delta)\Delta P(b') - \delta P^0 > 0$$

If these inequalities have a non-empty intersection for some value of  $\delta$ , then they have a non-empty intersection for any larger value of  $\delta$  (as is implied by our graphical interpretation in the previous paragraph).

Thus we can do a binary search over the values of  $\delta \in (0, 1)$ . At each iteration we have an upper bound  $H$  and lower bound  $L$  on  $\delta$ . We test whether  $\delta = (H + L)/2$  gives a non-empty intersection in the above inequalities. If it does, then we decrease the upper bound  $H$  to  $(H + L)/2$ , if it doesn't then we increase the lower bound  $L$  to  $(H + L)/2$ . We stop whenever  $H - L$  is smaller than some precision, or when the implied upper and lower bounds on  $v$  from the feasible region for  $\delta = H$ , are smaller than some precision.

The value that corresponds to the smallest rationalizable multiplicative error  $\delta$  can be viewed as a point prediction for the value of the player. It is exactly the value that corresponds to the black dot in Figure 1. Since this is a point of the  $\mathcal{NR}$  set the estimation error of this point from data has at least the same estimation error convergence properties as the whole  $\mathcal{NR}$  set that we derived in Section 5.

*Experimental Findings.* We compute the pair of the smallest rationalizable multiplicative error  $\delta^*$  and the corresponding predicted value  $v^*$  for every listing of each of the nine accounts that we analyzed. In Figure 2, on the right we plot the distribution of the smallest non-negative rationalizable error over the listings of a single account. Different listings of a single account are driven by the same algorithm, hence



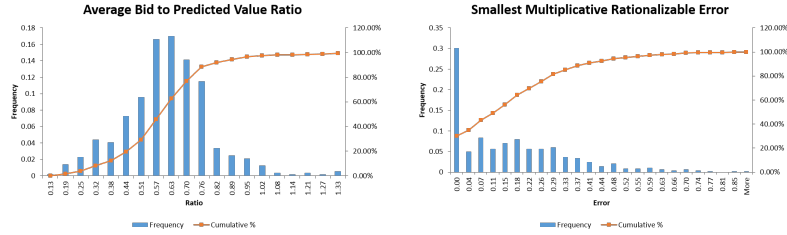


Fig. 2. Histogram of the ratio of predicted value over the average bid in the sequence and the histogram of the smallest non-negative rationalizable multiplicative error  $\delta$  (the smallest bucket contains all listings with a non-positive smallest rationalizable error).

we view this plot as the “statistical footprint” of this algorithm. We observe that all accounts have a similar pattern in terms of the smallest rationalizable error: almost 30% of the listings within the account can be rationalized with an almost zero error, i.e.  $\delta^* < 0.0001$ . We note that regret can also be negative, and in the figure we group together all listings with a negative smallest possible error. This contains 30% of the listings. The empirical distribution of the regret constant  $\delta^*$  for the remaining 70% of the listings is close to the uniform distribution on  $[0, .4]$ . Such a pattern suggests that half of the listings of an advertiser have reached a stable state, while a large fraction of them are still in a learning state.

We also analyze the amount of relative bid shading. It has been previously observed that bidding in the GSP auctions is not truthful. Thus we can empirically measure the difference between the bids and estimated values of advertisers associated with different listings. Since the bid of a listing is not constant in the period of the week that we analyzed, we use the average bid as proxy for the bid of the advertiser. Then we compute the ratio of the average bid over the predicted value for each listing. We plot the distribution of this ratio over the listings of a typical account in the left plot of Figure 2. We observe that based on our prediction, advertisers consistently bid on average around 60% of their value. Though the amount of bid shading does have some variance, the empirical distribution of the ratio of average bid and the estimated value is close to normal distribution with mean around 60%.

Interestingly, we observe that based on our prediction, advertisers consistently bid on average around 60% of their value. Though the amount of bid shading does have some variance, the empirical distribution of the ratio of average bid and the estimated value is close to normal distribution with mean around 60%.<sup>2</sup>

Figure 2 depicts the ratio distribution for one account. We give similar plots for all other accounts that we analyzed in the full version of the paper.

Last we also analyze whether there is any correlation between the smallest rationalizable error and the amount of underbidding for listings that seem to be in a learning phase (i.e. have  $\delta^*$  greater than some tiny amount). We present a scatter plot of the pairs of  $(\delta^*, v^*)$  in Figure 3 for a single

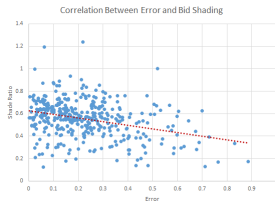


Fig. 3. Scatter plot of pairs  $(v^*, \delta^*)$  for all listings of a single advertiser.

<sup>2</sup>We do observe that for a very small number of listings within each account our prediction implies a small amount of overbidding on average. These outliers could potentially be because our main assumption that the value remains fixed throughout the period of the week doesn’t hold for these listings due to some market shock.

account and for listings that have  $\delta^* > 0.0001$ . Though there does not seem to be a significant correlation we consistently observe a small correlation: listings with higher error shade their bid more.

## REFERENCES

- ARADILLAS-LÓPEZ, A., GANDHI, A., AND QUINT, D. 2013. Identification and inference in ascending auctions with correlated private values. *Econometrica* 81, 2, 489–534.
- ARORA, S., HAZAN, E., AND KALE, S. 2012. The multiplicative weights update method: a meta-algorithm and some applications. *Theory of Computing*.
- ATHEY, S. AND HAILE, P. A. 2007. Nonparametric approaches to auctions. *Handbook of econometrics* 6, 3847–3965.
- ATHEY, S. AND NEKIPELOV, D. 2010. A structural model of sponsored search advertising auctions. In *Sixth ad auctions workshop*.
- BAJARI, P., HONG, H., AND NEKIPELOV, D. 2013. Game theory and econometrics: a survey of some recent results. *Advances in Economics and Econometrics* 3.
- CAMPO, S., GUERRE, E., PERRIGNE, I., AND VUONG, Q. 2011. Semiparametric estimation of first-price auctions with risk-averse bidders. *The Review of Economic Studies* 78, 1, 112–147.
- CARAGIANNIS, I., KAKLAMANIS, C., KANELLOPOULOS, P., KYROPOULOU, M., LUCIER, B., PAES LEME, R., AND TARDOS, E. 2015. On the efficiency of equilibria in generalized second price auctions. *Journal of Economic Theory*.
- FREUND, Y. AND SCHAPIRE, R. 1999. Adaptive game playing using multiplicative weights. *Games and Economic Behavior* 29, 79–103.
- GUERRE, E., PERRIGNE, I., AND VUONG, Q. 2000. Optimal nonparametric estimation of first-price auctions. *Econometrica* 68, 3, 525–574.
- GUERRE, E., PERRIGNE, I., AND VUONG, Q. 2009. Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica* 77, 4, 1193–1227.
- HAILE, P. A. AND TAMER, E. 2003. Inference with an incomplete model of english auctions. *Journal of Political Economy* 111, 1, 1–51.
- JOFRE-BONET, M. AND PESENDORFER, M. 2003. Estimation of a dynamic auction game. *Econometrica* 71, 5, 1443–1489.
- LITTLESTONE, N. AND WARMUTH, M. 1994. The weighted majority algorithm. *Information and Computation* 108, 212–260.
- ROUGHGARDEN, T. 2012. The price of anarchy in games of incomplete information. *SIGecom Exch.* 11, 1, 18–20.
- S. HART, S. AND MAS-COLELL, A. 2000. A simple adaptive procedure leading to correlated equilibrium. *1127-1150*, 68.
- STONE, C. J. 1982. Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics*, 1040–1053.
- SYRGKANIS, V. AND TARDOS, E. 2013. Composable and efficient mechanisms. In *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing*. STOC '13. ACM, New York, NY, USA, 211–220.
- TOPKIS, D. M. 1998. *Supermodularity and complementarity*. Princeton university press.
- VARIAN, H. R. 2007. Position auctions. *International Journal of Industrial Organization* 25, 6, 1163–1178.
- VIVES, X. 2001. *Oligopoly pricing: old ideas and new tools*. MIT press.
- XIN JIANG, A. AND LEYTON-BROWN, K. 2007. Bidding agents for online auctions with hidden bids. *Machine Learning* 67.