Welfare Guarantees from Data

Authored by:

Vasilis Syrgkanis
Darrell Hoy
Denis Nekipelov

Abstract

Analysis of efficiency of outcomes in game theoretic settings has been a main item of study at the intersection of economics and computer science. The notion of the price of anarchy takes a worst-case stance to efficiency analysis, considering instance independent guarantees of efficiency. We propose a data-dependent analog of the price of anarchy that refines this worst-case assuming access to samples of strategic behavior. We focus on auction settings, where the latter is non-trivial due to the private information held by participants. Our approach to bounding the efficiency from data is robust to statistical errors and mis-specification. Unlike traditional econometrics, which seek to learn the private information of players from observed behavior and then analyze properties of the outcome, we directly quantify the inefficiency without going through the private information. We apply our approach to datasets from a sponsored search auction system and find empirical results that are a significant improvement over bounds from worst-case analysis.

1 Paper Body

A major field at the intersection of economics and computer science is the analysis of the efficiency of systems under strategic behavior. The seminal work of [6, 11] triggered a line of work on quantifying the inefficiency of computer systems, ranging from network routing, resource allocation and more recently auction marketplaces [10]. However, the notion of the price of anarchy suffers from the pessimism of worst-case analysis. Many systems can be inefficient in the worst-case over parameters of the model, but might perform very well for the parameters that arise in practice. Due to the large availability of datasets in modern economic systems, we propose a data-dependent analog of the price of anarchy, which assumes access to a sample of strategic behavior from the system. We focus our analysis on auction systems where the latter approach is more interesting due to the private information held by the participants of the system, i.e. their private value for the item at sale. Since efficiency is a function of these private parameters, quantifying the inefficiency of the system from

samples of strategic behavior is non-trivial. The problem of estimation of the inefficiency becomes an econometric problem where we want to estimate a function of hidden variables from observed strategic behavior. The latter is feasible under the assumption that the observed behavior is the outcome of an equilibrium of the strategic setting, which connects observed behavior to unobserved private information. Traditional econometric approaches to auctions [3, 8], address such questions by attempting to exactly pin-point the private parameters from the observed behavior and subsequently measuring the quantities of interest, such as the efficiency of the allocation. The latter approach is problematic in complex auction systems for two main reasons: (i) it leads to statistical inefficiency, (ii) it requires strong conditions on the connection between observed behavior and private information. Even for a single-item first-price auction, uniform estimation of the private value of a player from T samples of observed bids, can only be achieved at O(T 1/3)-rates [3]. Moreover, uniquely identifying the private information from the observed behavior, requires a one-to-one mapping between the two 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

quantities. The latter requires strong assumptions on the distribution of private parameters and can only be applied to simple auction rules. Our approach bridges the gap between worst-case price of anarchy analysis and statistically and modeling-wise brittle econometric analysis. We provide a data-dependent analog of recent techniques for quantifying the worst-case inefficiency in auctions [13, 4, 10], that do not require characterization of the equilibrium structure and which directly quantify the inefficiency through best-response arguments, without the need to pin-point the private information. Our approach makes minimal ? ? T)assumptions on the distribution of private parameters and on the auction rule and achieves O(rates of convergence for many auctions used in practice, such as the Generalized Second Price (GSP) auction [2, 14]. We applied our approach to a real world dataset from a sponsored search auction system and we portray the optimism of the data-dependent guarantees as compared to their worst-case counterparts [1].

2

Preliminaries

We consider the single-dimensional mechanism design setting with n bidders. The mechanism designer wants to allocate a unit of good to the bidders, subject to some feasibility constraint on the vector of allocations $(x_1, \ldots, x_n)$. Let X be the space of feasible allocations. Each bidder i has a private value v_i ? [0, H] per-unit of the good, and her utility when she gets allocation x_i and is asked to make a payment p_i is v_i ? x_i ? p_i . The value of each bidder is drawn independently from distribution with CDF F_i , supported in V_i ? $R+$ and let F = ?i F_i be the joint distribution. An auction A solicits a bid b_i ? B from each bidder i and decides on the allocation vector based on an allocation rule X : B n ? X and a payment rule p : B n ? Rn . For a vector of values and bids, the utility of a bidder is: U_i (b; v_i) = v_i ? X_i (b) ? P_i (b). (1) A strategy ?i : V_i ? B, for each bidder i, maps the value of the bidder to a bid. Given an auction A and distribution of values F, a strategy profile ? is a Bayes-Nash

Equilibrium (BNE) if each bidder i with any value vi ? Vi maximizes her utility in expectation over her opponents bids, by bidding ?i (vi). The welfare of an auction outcome is the expected utility generated for all the bidders, plus the revenue of the auctioneer, which due to the form of bidder utilities boils down to being the total value that the bidders get from the allocation. Thus the expected utility of a strategy profile ? is ? ? X W ELFARE(?; F) = Ev?F ? vi ? Xi (?(v))? (2) i?[n]

We denote with O PT(F) the expected optimal welfare: O PT(F) = Ev?F [maxx?X

P

i?[n]

vi ? xi].

Worst-case Bayes-Nash price of anarchy. The Bayesian price of anarchy of an auction is defined as the worst-case ratio of welfare in the optimal auction to the welfare in a Bayes-Nash equilibrium of the original auction, taken over all value distributions and over all equilibria. Let BN E(A, F) be the set of Bayes-Nash equilibria of an auction A, when values are drawn from distributions F. Then: P OA =

sup F,??BN E(F)

3

O PT(F) W ELFARE(?; F)

(3)

Distributional Price of Anarchy: Refining the P OA with Data

We will assume that we observe T samples b1:T = {b1 , . . . , bT } of bid profiles from running T times an auction A. Each bid profile bt is drawn i.i.d. based on an unknown Bayes-Nash equilibrium ? of the auction, i.e.: let D denote the distribution of the random variable ?(v), when v is drawn from F. Then bt are i.i.d. samples from D. Our goal is to refine our prediction on the efficiency of the auction and compute a bound on the price of anarchy of the auction conditional on the observed data set. More formally, we want to derive statements of the form: conditional on b1:T , with probability at least 1 ? ?: W ELFARE(?; F) ? ?1? O PT(F), where ?? is the empirical analogue of the worst-case price of anarchy ratio. 2

Infinite data limit We will tackle this question in two steps, as is standard in estimation theory. First we will look at the infinite data limit where we know the actual distribution of equilibrium bids D. We define a notion of price of anarchy that is tailored to an equilibrium bid distribution, which we refer to as the distributional price of anarchy. In Section 4 we give a distribution-dependent upper bound on this ratio for any single-dimensional auction. Subsequently, in Section 5, we show how one can estimate this upper bound on the distributional price of anarchy from samples. Given a value distribution F and an equilibrium ?, let D(F, ?) denote the resulting equilibrium bid distribution. We then define the distributional price of anarchy as follows: Definition 1 (Distributional Price of Anarchy). The distributional price of anarchy DP OA(D) of an auction A and a distribution of bid profiles D, is the worst-case ratio of welfare in the optimal

allocation to the welfare in an equilibrium, taken over all distributions of values and all equilibria that could generate the bid distribution D: DP OA(D) =

sup F,??BN E(F) s.t. D(F,?)=D

O PT(F) W ELFARE(?; F)

(4)

This notion has nothing to do with sampled data-sets, but rather is a hypothetical worst-case quantity that we could calculate had we known the true bid generating distribution D. What does the extra information of knowing D give us? To answer this question, we first focus on the optimization problem each bidder faces. At any Bayes-Nash equilibrium each player must be best-responding in expectation over his opponent bids. Observe that if we know the rules of the auction and the equilibrium distribution of bids D, then the expected allocation and payment function of a player as a function of his bid are uniquely determined: xi (b; D) = Eb?i ?D?i [Xi (b, b?i)]

pi (b; D) = Eb?i ?D?i [Pi (b, b?i)] .

(5)

Importantly, these functions do not depend on the distribution of values F, other than through the distribution of bids D. Moreover, the expected revenue of the auction is also uniquely determined: " # X R EV(D) = Eb?D Pi (b) ,

(6) i

Thus when bounding the distributional price of anarchy, we can assume that these functions and the expected revenue are known. The latter is unlike the standard price of anarchy analysis, which essentially needs to take a worst-case approach to these quantities. Shorthand notation Through the rest of the paper we will fix the distribution D. Hence, for brevity we omit it from notation, using xi (b), pi (b) and R EV instead of xi (b; D), pi (b; D) and R EV(D).

4

Bounding the Distributional Price of Anarchy

We first upper bound the distributional price of anarchy via a quantity that is relatively easy to calculate as a function of the bid distribution D and hence will also be rather straightforward to estimate from samples of D, which we defer to the next section. To give intuition about the upper bound, we start with a simple but relevant example of bounding the distributional price of anarchy in the case when the auction A is the single-item first price auction. We then generalize the approach to any auction A. 4.1

Example: Single-Item First Price Auction

In a single item first price auction, the designer wants to auction a single indivisible good. Thus the space of feasible allocations X , are ones where only one player gets allocation xi = 1 and other players get allocation 0. The auctioneer solicits bids bi from each bidder and allocates the good to the highest bidder (breaking ties lexicographically), charging him his bid. Let D be the equilibrium distribution of bids and let Gi be the CDF of the bid of player i. For simplicity we assume that Gi is continuous (i.e. the distribution is atomless). Q Then the expected allocation of a player i from submitting a bid b is equal to xi (b) = G?i (b) = j6=i Gj (b) and his expected payment is pi (b) = b ? xi (b), leading to expected utility: ui (b; vi) = (vi ? b)G?i (b). 3

4

The quantity DP OA is a complex object as it involves the structure of the set of equilibria of the given auction. The set of equilibria of a first price auction when bidders values are drawn from different distributions is an horrific object.1 However, we can upper bound this quantity by a much simpler data-dependent quantity by simply invoking the fact that under any equilibrium bid distribution no player wants to deviate from his equilibrium bid. Moreover, this data-dependent quantity can be much better than its worst-case counterpart used in the existing literature on the price of anarchy. Lemma 1. Let A be the single item first price auction and let D be the equilibrium distribution of maxi?[n] Eb?i ?D?i [maxj6=i bj] ?(D) bids, then DP OA(D) ? 1?e . ??(D) , where ?(D) = Eb?D [maxi?[n] bi] Proof. Let Gi be the CDF of the bid of each player under distribution D. Moreover, let ? denote the equilibrium strategy that leads to distribution D. By the equilibrium condition, we know that for all vi ? Vi and for all b0 ? B, ui (?i (vi); vi) ? ui (b0 ; vi) = (vi ? b0) ? G?i (b0).

(7)

We will give a special deviating strategy used in the literature [13], that will show that either the players equilibrium utility is large or the expected maximum Rother bid is high. Let Ti denote the ? expected maximum other bid which can be expressed as Ti = 0 1 ? G?i (z)dz. We consider the randomized deviation where the player submits a randomized bid in z ? [0, vi (1 ? e??)] with PDF f (z) = ?(vi1?z) . Then the expected utility from this deviation is: E0 [ui (b ; vi)] = b

Adding the quantity 1 ? Ti

vi (1?e??)

Z

0

1 ?

0

R vi (1?e??) 0

1 (vi ? z) ? G?i (z)f (z)dz = ?

(1 ? G?i (z))dz ?

1 ? Ti

Z

vi (1?e??)

G?i (z)dz

on both sides, we get: Eb0 [ui (b0 ; vi)] +

??

? vi 1?e? . Invoking the equilibrium condition we get: ui (?i (vi); vi) +

Subsequently, for any x?i ? [0, 1]: ui (?i (vi); vi) +

(8)

0

1 1 ? e?? Ti ? x?i ? vi ? x?i . ? ?

1 ? Ti

??

? vi 1?e?

5

.

(9)

If x?i is the expected allocation of player i under the efficient allocation rule Xi? (v) ? 1{vi = maxj vj }, then taking expectation of Equation (9) over vi and adding across all players we get: " # X X 1 1 ? e?? E [ui (?i (vi); vi)] + E Ti Xi? (v) ? O PT(F) (10) vi ?v i ? i P The theorem then follows by invoking the fact that for any feasible allocation x: i Ti ? xi ? maxi Ti = ?(D)R EV(D), using the fact that expected total agent utility plus total revenue at equilibrium is equal to expected welfare at equilibrium and setting ? = ?(D). Comparison with worst-case P OA In the worst-case, ?(D) is upper bounded by 1, leading to the well-known worst-case price of anarchy ratio of the single-item first price auction of (1 ? 1/e)?1 , irrespective of the bid distribution D. However, if we know the distribution D then we can explicitly estimate ?, which can lead to a much better ratio (see Figure 1). Moreover, observe that even if we had samples from the bid distribution D, then estimating ?(D) is very easy as it corresponds to the ratio of two expectations, each of which can be estimating to within an O(?1T) error by a simple average and using standard concentration inequalities. Even thought this improvement, when compared to the worst-case bound might not be that drastic in the first price auction, the extension of the analysis in the next section will be applicable even to auctions where the analogue of the quantity ?(D) is not even bounded in the worst-case. In those settings, the empirical version of the price of anarchy analysis is of crucial importance to get any efficiency bound. 1 Even for two bidders with uniformly distributed values U [0, a] and U [0, b], the equilibrium strategy requires solving a complex system of partial differential equations, which took several years of research in economics to solve (see [15, 7])

4

Price of Anarchy

4

3

2

1

0

1

2

3

4

?

Figure 1: The upper bound on the distributional price of anarchy of an auction

?(D) 1?e??(D

as a function of ?(D).

Comparison with value inversion approach Apart from being just a primer to our main general result in the next section, the latter result about the data-dependent efficiency bound for the first price auction, is itself a contribution to the literature. It is notable to compare the latter result with the standard

econometric approach to estimating values in a first price auction pioneered by [3] (see also [8]). Traditional non-parametric auction econometrics use the equilibrium best response condition to pin-point the value of a player from his observed bid, by what is known as value inversion. In particular, if the function: $u_i(b_0; v_i) = (v_i ? b_0) ? G?_i(b_0)$ has a unique maximum for each v_i and this maximum is strictly monotone in v_i, then given the equilibrium bid of a player b_i and given a data distribution D we can reverse engineer the value $v_i(b_i)$ that the player mustPhave. Thus if we know the bid distribution D we can calculate the equilibrium welfare as $E_{b?D}[\sum_i v_i(b_i) ? X_i(b)]$. Moreover, we can calculate the expected optimal welfare as: $E_{b?D}[\max_i v_i(b_i)]$. Thus we can pin-point the distributional price of anarchy. However, the latter approach suffers from two main drawbacks: (i) estimating the value inversion function $v_i(?)$ uniformly over b from samples, can only happen at very slow rates that are at least $O(1/T^{1/3})$ and which require differentiability assumptions from the value and bid distribution as well as strong conditions that the density of the value distribution is bounded away from zero in all the support (with this lower bound constant entering the rates of convergence), (ii) the main assumption of the latter approach is that the optimal bid is an invertible function and that given a bid there is a single value that corresponds to that bid. This assumption might be slightly benign in a single item first price auction, but becomes a harsher assumption when one goes to more complex auction schemes. Our result in Lemma 1 suffers neither of these drawbacks: it admits fast estimation rates from samples, makes no assumption on properties of the value and bid distribution and does not require invertibility of the best-response correspondence. Hence it provides an upper bound on the distributional price of anarchy that is statistically robust to both sampling and mis-specification errors. The robustness of our approach comes with the trade-off that we are now only estimating a bound on the efficiency of the outcome, rather than exactly pinpointing it. 4.2

Generalizing to any Single-Dimensional Auction Setting

Our analysis on DP OA is based on the reformulation of the auction rules as an equivalent pay-yourbid auction and then bounding the price of anarchy as a function of the ratio of how much a player needs to pay in an equivalent pay-your-bid auction, so as to acquire his optimal allocation vs. how much revenue is the auctioneer collecting. For any auction, we can re-write the expected utility of a bid b:

$p_i(b)$ $u_i(b; v_i) = x_i(b) v_i ? $ (11) $x_i(b)$ This can be viewed as the same form of utility if the auction was a pay-your-bid auction and the player $p_i(b)$ submitted a bid of $\frac{x_i(b)}{p_i(b)}$. We refer to this term as the price-per-unit and denote it $ppu(b) = x_i(b)$. Our analysis will be based on the price-per-unit allocation rule $x?(?)$, which determines the expected allocation of a player as a function of his price-per-unit. Given this notation, we can re-write the utility that an agent achieves if he submits a bid that corresponds to a price-per-unit of z as: $u?_i(z; v_i) = x?(z)(v_i ? z)$. The latter is exactly the form of a pay-your-bid auction. Our upper bound on the DP OA, will be based on the inverse of the PPU allocation rule; let $?_i(z) = x?^{?1}_i(z)$ be the price-per-unit

of the cheapest bid that achieves allocation at least z. More formally, 5

$\tau_i(z) = \min_{b-x_i(b)\geq z}\{ppu(b)\}$. For simplicity, we assume that any allocation $z \in [0, 1]$ is achieveable by some high enough bid b.2 Given this we can define the threshold for an allocation: Definition 2 (Average Threshold). The average threshold for agent i is $\int_0^1 T_i = \tau_i(z)\,dz$ (12)

In Figures 3 and 2 we provide a pictorial representation of these quantities. Connecting with the previous section, for a first price auction, the price-per-unit function is $ppu(b) = b$, the price-per-unit allocation function is $x_{\tau_i}(b) = G_{\tau_i}(b)$ and the threshold function is $\tau_i(z) = G^{-1}_{\tau_i}(z)$. The average $\int_0^1 \tau_i(z)dz = \int_0^1 1 - G_{\tau_i}(b)db$ threshold T_i is equal to $\int_0^1 G_{\tau_i}(z)dz = \int_0^1 1 - G_{\tau_i}(b)db$, i.e. the expected maximum other bid. 1

$x_{\tau_i}(ppu) = \tau_i^{-1}(ppu)$

E[Allocation]

E[Allocation]

1

T_i

$x_\tau(ppu)$

$u_i(b)$ $ppu(b)$

v_i

PPU

PPU

Figure 2: For any bid b with PPC ppu(b), the area of a rectangle between $(ppu(b), x_{\tau_i}(ppu(b)))$ and $(v_i, 0)$ on the bid allocation rule is the expected utility $u_i(b)$. The BNE action b^* is chosen to maximize this area.

Figure 3: The average threshold is the area to the left of the price-per-unit allocation rule, integrate from 0 to 1.

We now give our main theorem, which is a distribution-dependent bound on DP OA, that is easy to compute give D and which can be easily estimated from samples of D. This theorem is a generalization of Lemma 1 in the previous section. Theorem 2 (Distributional Price of Anarchy Bound). For any auction A in a single dimensional setting and for any bid distribution P D, the distributional price of anarchy is bounded by $DP\ OA(D) \leq n\max_{x\in X}\sum_{i=1}^{n}T_i \cdot x_i$, where $\mu(D) = \frac{R}{EV(D)}$ $\frac{1}{1-e^{-\mu(D)}}$ Theorem 2 provides our main method for bounding the distributional price of anarchy. All we need is to compute the revenue R EV of the auction and the quantity: $\sum_{i=1}^{n} T = \max_{x\in X}\sum_{i=1}^{n}T_i \cdot x_i$, (13) under the given bid distribution D. Both of these are uniquely defined quantities if we are given D. Moreover, once we compute T_i, the optimization problem in Equation (13) is simply a welfare maximization problem, where each player?s value per-unit of the good is T_i. Thus, the latter can be solved in polynomial time, whenever the welfare maximization problem over the feasible set X is polynomial-time solvable. Theorem 2 can be viewed as a bid distribution-dependent analogue of the revenue covering framework [4] and of the smooth mechanism framework [13]. In particular, the quantity $\mu(D)$ is the datadepenent analogue of the worst-case μ quantity used in the definition of μ-revenue covering in [4] and is roughly related to the λ quantity used in the definition of a (λ, μ)-smooth mechanism in [13].

8

5

Distributional Price of Anarchy Bound from Samples

In the last section, we assumed we were given distribution D and hence we could compute the quantity $? = \text{RTEV}$, which gave an upper bound on the DP OA. We now show how we can estimate this 2

The theory can be easily extended to allow for different maximum achievable allocations by each player, by simply integrating the average threshold only up until the largest such allocation.

6

quantity ? when given access to i.i.d. samples b1:T from the bid distribution D. We will separately estimate T and R EV. The latter is simple expectation and thereby can be easily estimated by an average at ?1T rates. For the former we first need to estimate Ti for each player i, which requires estimation of the allocation and payment functions xi (?; D) and pi (?; D). Since both of these functions are expected values over the equilibrium bids of opponents, we will approximate them by their empirical analogues: T 1X xbi (b) = Xi (b, bt?i) T t=1

T 1X pbi (b) = Pi (b, bt?i). T t=1

(14)

To bound the estimation error of the quantities T?i produced by using the latter empirical estimates of the allocation and payment function, we need to provide a uniform convergence property for the error of these functions over the bid b. Since b takes values in a continuous interval, we cannot simply apply a union bound. We need to make assumptions on the structure of the class of functions FXi = {Xi (b, ?) : b ? B} and FPi = {Pi (b, ?) : b ? B}, so as uniformly bound their estimation error. For this we resort to the technology of Rademacher complexity. For a generic class of functions F and a sequence of random variables Z 1:T , the Rademacher complexity is defined as: " # T 1X t 1:T t RT (F, Z) = E sup ? f (Z) . (15) ? 1:T f ?F T t=1 where each ? t ? {?1/2} is an i.i.d. Rademacher random variable, which takes each of those values with equal probabilities. The following well known theorem will be useful in our derivations: Theorem 3 ([12]). Suppose that for any sample Z 1:T of size T , RT (F, Z 1:T) ? RT and suppose that functions in F take values in [0, H]. Then with probability 1 ? ?:

r T

1 X 2 log(4/?)

f (Zt) ? E[f (Z)] ? 2RT + H (16) sup

T f ?F T t=1 This Theorem reduces our uniform error problem to bounding the Rademacher complexity of classes FXi and FPi , since we immediately have the following corollary (where we also use that the allocation functions lie in [0, 1] and the payment functions lie in [0, H]): Corollary 4. Suppose that for any sample b1:T of size T , the Rademacher complexity of classes FXi and FPi is at most RT . Then with probability 1??/2, both sup —xbi (b)?xi (b)— and sup —pbi (b)?pi (b)— b?B b?B p are at most 2RT + H 2 log(4/?) / T . ? ? We now provide conditions under which the Rademacher complexity of these classes is O(1/ T). Lemma 5. Suppose that B = [0, B] and for each

bidder i and each $b_i \in B$, the functions $X_i(b_i, \cdot) : [0, B]^{n-1} \mapsto [0, 1]$ and $P_i(b, \cdot) : [0, B]^{n-1} \mapsto [0, H]$ can be computed as finite superposition of (i) coordinate-wise multiplication of bid vectors b_{-i} with constants; (ii) comparison indicators $1\{? > ?\}$ of coordinates or constants; (iii) pairwise addition $? + ?$ of coordinates or constants. The Rademacher p complexity for both classes on a sample of size T is $O\left(\sqrt{\log(T)/T}\right)$. The proof of this Lemma follows by standard arguments of Rademacher calculus, together with VC arguments on the class of pairwise comparisons. Those arguments can be found in [5, Lemma 9.9] and [9, Lemma 11.6.28]. Thereby, we omit its proof. The assumptions of Lemma 5 can be directly verified, for instance, for the sponsored search auctions where the constants that multiply each bid correspond to quality factors of the bidders, e.g. as in [2] and [14] and then the allocation and the payment is a function of the rank of the weighted bid of a player. In that case the price and the allocation rule are determined solely by the ranks and the values of the score-weighted bids $?_i b_i$, as well as the position specific quality factors $?_j$, for each position j in the auction. Next we turn to the analysis of the estimation errors on quantities T_i. We consider the following plug-in $?_{b_i}(b)$ in estimator for T_i: We consider the empirical analog of function $?_i(\cdot)$ by $?_{b_i}(z) = \inf x_{b_i}(b) . b?[0,B], x_{b_i}(b)?z$

7

Then the empirical analog of T_i is obtained by: $\int_0^1 T_{b_i} =$

$?_{b_i}(z) dz.$

(17)

0

To bound the estimation error of T_{b_i}, we need to impose an additional condition that ensures that any non-zero allocation requires the payment from the bidder at least proportional to that allocation. Assumption 6. We assume that $p_i(x^{-1}_i(\cdot))$ is Lipschitz-continuous and that the mechanism is worst-case interim individually rational, i.e. $p_i(b) \geq H ? x_i(b)$. $? ? T$) rates of convergence of T_{b_i} to T_i and of the Under this assumption we can establish that $O(? = \max_{x?X} \sum_{i=1}^{n} T?_i ? x_i$ of the optimized threshold to T as well as the empirical empirical analog T i=1 ? T analog Rd EV of the revenue to R EV. Thus the quantity $? ?= d$, will also converge to $? = RTEV$ at R EV that rate. This implies the following final conclusion of this section. Theorem 7. Under Assumption 6 and the premises of Lemma 5, with probability $1 ? ?$:
! r O PT(F) H log(n/?) ? b ? n max{L, H} ? +O (18) W ELFARE(?; F) 1 ? e?b? T

6

Sponsored Search Auction: Model, Methodology and Data Analysis

We consider a position auction setting where k ordered positions are assigned to n bidders. An outcome m in a position auction is an allocation of positions to bidders. $m(j)$ denotes the bidder who is allocated position j; $m^{-1}(i)$ refers to the position assigned to bidder i. When bidder i is assigned to slot j, the probability of click $c_{i,j}$ is the product of the click-through-rate of the slot $?_j$ and the quality score of the bidder, $?_i$, so $c_{i,j} = ?_j ?_i$ (in the data the quality scores for each bidder are varying across different auctions and we used the average score as a proxy for the score of a bidder). Each advertiser has a value-per-click (VPC) v_i

, which is not observed in the data and which we assume is drawn from some distribution F_i. Our benchmark for welfare will be the welfare of the auction that chooses a feasible allocation to maximize the welfare generated, thus $OPT = E_v [\max_m \sum_i \sum_{m=1} (i) v_i]$. We consider data generated by advertisers repeatedly participating in a sponsored search auction. The mechanism 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: Each advertiser i is associated with a click probability $?_i$ and a scoring coefficient s_i and is asked to submit a bid-per-click b_i. Advertisers are ranked by their rank-score $q_i = s_i ? b_i$ and allocated positions in decreasing order of rank-score as long as they pass a rank-score reserve r. All the mentioned sets of parameters $? = (s, ?, ?, r)$ and the bids b are observable in the data. We will denote with $?_{b,?} (j)$ the bidder allocated in slot j under a bid profile b and parameter profile $?1 ?$. We denote with $?_{b,?} (i)$ the slot allocated to bidder i. If advertiser i is allocated position j, then he pays only when he is clicked and his payment, i.e. his ncost-per-click is the minimal bid o max s?

(j+1) ?b?

(j+1) ,r

b,? b,? he had to place to keep his position, which is: $cpc_{ij} (b; ?) = $. Mapping s_i this setting to our general model, the allocation function of the auction is $X_i (b) = ???1 (i) ? ?, b,?$

the payment function is $P_i (b) = ???1 (i) ? ? ? cpci??1 (i) (b; ?)$ and the utility function is: b,? b,?

$U_i (b; v_i) = ???1 (i) ? ?_i ? v_i ? cpci??1 (i) (b; ?) . $ b,?

b,?

Data Analysis We applied our analysis to the BingAds sponsored search auction system. We analyzed eleven phrases from multiple thematic categories. For each phrase we retrieved data of auctions for the phrase for the period of a week. For each phrase and bidder that participated in the auctions for the phrase we computed the allocation curve by simulating the auctions for the week under any alternative bid an advertiser could submit (bids are multiples of cents). See Figure 4 for the price-per-unit allocation curves $x ?_i (?) = ?_i?1 (?)$ for a subset of the advertisers for a specific search phrase. We estimated the average threshold $T?_i$ for each bidder by numerically 8

?

phrase1 phrase2 phrase3 phrase4 phrase5 phrase6 phrase7 phrase8 phrase9 phrase10 phrase11

T ? ?= d R EV .511 .509 2.966 1.556 .386 .488 .459 .419 .441 .377 .502

1 DP OA

?? ?

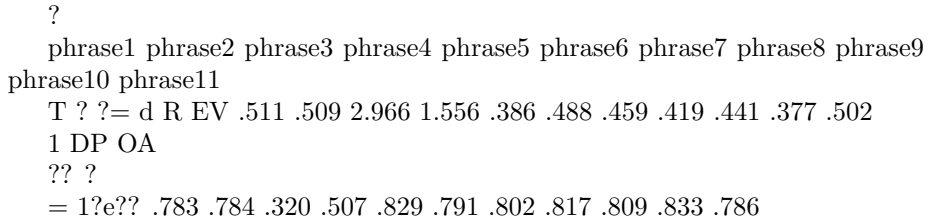= 1?e?? .783 .784 .320 .507 .829 .791 .802 .817 .809 .833 .786

Figure 4: (left) Examples of price-per-unit allocation curves for a subset of six advertisers for a specific keyword during the period of a week. All axes are normalized to 1 for privacy reasons. (right) Distributional Price of Anarchy analysis for a set of eleven search phrases on the BingAds system.

integrating these allocation curves along the y axis. We then applied the approach described in ? = maxx?X P ? Section 3 for each of the search phrases, computing the quantity T i?[n] Ti ? xi = P ? maxm(?) i Ti ? ?i ? ?m?1 (i) . The latter optimization is simply the optimal assignment problem where each player?s value-per-click is T?i and can be performed by greedily assigning players to slots in decreasing order of T?i . We then estimate the expected revenue by the empirical revenue Rd EV . ?

T We portray our results on the estimate ? ?= d and the implied bound on the distributional price R EV of anarchy for each of the eleven search phrases in Table 4. Phrases are grouped based on thematic category. Even though the worst-case price of anarchy of this auction is unbounded (since scores si are not equal to qualities ?i , which is required in worst-case P OA proofs [1]), we observe that empirically the price of anarchy is very good and on average the guarantee is approximately 80% of the optimal. Even if si = ?i the worst-case bound on the P OA implies guarantees of approx. 34% [1], while the DP OA we estimated implies significantly higher percentages, portraying the value of the empirical approach we propose.

2   References

[1] Ioannis Caragiannis, Christos Kaklamanis, Maria Kyropoulou, Brendan Lucier, Renato Paes Leme, and ?va Tardos. Bounding the inefficiency of outcomes in generalized second price auctions. pages 1?45, 2014. [2] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. The American economic review, 97(1):242?259, 2007. [3] Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. Econometrica, 68(3):525?574, 2000. [4] Jason Hartline, Darrell Hoy, and Sam Taggart. Price of Anarchy for Auction Revenue. In ACM Conference on Economics and Computation, pages 693?710, New York, New York, USA, 2014. ACM Press. [5] Michael R Kosorok. Introduction to empirical processes and semiparametric inference. Springer Science & Business Media, 2007. [6] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In STACS 99, pages 404?413. Springer, 1999. [7] Vijay Krishna. Auction Theory. Academic Press, March 2002. [8] H. J. Paarsch and H. Hong. An Introduction to the Structural Econometrics of Auction Data. MIT Press, 2006. 9

[9] D. Pollard. Convergence of Stochastic Processes. Springer-Verlag, 1984. [10] Tim Roughgarden, Vasilis Syrgkanis, and ?va Tardos. The price of anarchy in auctions. CoRR, abs/1607.07684, 2016. [11] Tim Roughgarden and Eva Tardos. How bad is selfish routing? J. ACM, 49(2):236?259, March 2002. [12] S. Shalev-Shwartz and S. Ben-David. Understanding Machine Learning: From Theory to Algorithms. Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014. [13] Vasilis Syrgkanis and Eva Tardos. Composable and efficient mechanisms. In ACM Symposium on Theory of Computing, pages 211?220, 2013. [14] Hal R Varian. Online ad auctions.

The American Economic Review, pages 430?434, 2009. [15] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. The Journal of Finance, 16(1):8?37, 1961.

10