Deals or No Deals: Contract Design for Online Advertising

[Extended Abstract]

Vahab Mirrokni
Google Research
New York, NY
mirrokni@google.com

Hamid Nazerzadeh
Marshall School of Business
University of Southern California
nazerzad@usc.edu

ABSTRACT

Billions of dollars worth of display advertising are sold via contracts and deals. This paper presents a formal study of preferred deals, a new generation of contracts for selling online advertisement, that generalize the traditional reservation contracts; these contracts are suitable for advertisers with advanced targeting capabilities. We propose a constant-factor approximation algorithm for maximizing the revenue that can be obtained from these deals. We show, both theoretically and via data analysis, that deals, with appropriately chosen minimum-purchase guarantees, can yield significantly higher revenue than auctions. We evaluate our algorithm using data from Google's ad exchange platform. Our algorithm obtains about 90% of the optimal revenue where the second-price auction, even with personalized reserve, obtains at most 52% of the benchmark.

1. INTRODUCTION

Display advertising is the major source of revenue for online publishers and many Internet companies who sell the space on their webpages to advertisers. In 2014, the revenue of this market in the US alone exceeded $20 billion [36].

Display advertising is sold mainly via two channels: reservation contracts and real-time bidding. In a reservation or guaranteed-delivery contract, the advertiser specifies a targeting criteria and the size and duration of the campaign (e.g., 10 million impressions[1] to male users aged 35-50 from California in November) and the publisher guarantees to deliver impressions that match the criteria; the publisher usually pays a penalty if he fails to deliver [22]. The price per impression of the contracts is usually determined through negotiation. Advertisers can also purchase impressions via auction platforms, such as Google's DoubleClick and Yahoo!'s Right Media. Advertisers bid in real-time for a chance to show their ads on a publisher's website. These platforms sell billions of impressions each day [35].

Despite the growth of real-time bidding, still a large fraction of premium (and more valuable) online ad space is still sold via reservation contracts. In this paper, we study preferred deals, a new generation of contracts for selling display advertising: "Preferred Deals is a feature that allows Ad Exchange sellers to offer inventory to buyers at a fixed, pre-negotiated price before the inventory is made available to other buyers in the general auction." [28]

Under the traditional reservation contracts, the advertiser purchases all the impressions that are sent to him by the publisher. Preferred deals give the advertiser the option to choose a subset of impressions from those sent to him. If the advertiser "accepts" an impression, he pays the pre-negotiated price. If he rejects the impression, it will go back to the publisher.

Preferred deals are appealing for advertisers who can evaluate the value of an impression better than the publisher.[2] Many advertisers collect information about users who have visited their websites. They use this information mainly to *re-target* users by showing them ads tailored to their previous activities on the website; such as searching for a product or vacation packages [30]. Those impressions are of high value to the advertisers.

In this paper, we present a formal study of preferred deals and show how they can be structured in order to significantly increase the revenue of a publisher compared to second-price auctions. We show that finding the optimal sequence of deals is NP-Hard; see Section 3. We then present a greedy algorithm for a stochastic setting where the valuations of the buyers are drawn from a joint probability distribution. More specifically, in Section 4, we present our algorithm, called the Auction-Adjusted Greedy, where each advertiser gets the same number of impressions as they would have won in a second-price auction from the remaining impressions. We prove that our algorithm obtains at least half of the optimal revenue when the distributions of the buyers' valuations are independent of each other. When the valuations are correlated via a common value component, our algorithm yields at least $\frac{1}{3}$ of the optimal revenue.

We then evaluate our algorithm using data from Google ad exchange auctions in Section 5. To our surprise, we observe that our algorithm obtains about 90% of the benchmark revenue (the sum of the highest bids in each auction). In

[1] *Impression* is the unit of inventory and simply refers to the display of an ad to a user.

[2] We note that the information Google has about its users is not shared with the ad exchange.

WWW 2017, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/104

addition, it obtains significantly higher revenue, up to more than three times, compared to second-price auctions.

Related Work

In this section, we briefly discuss three lines of work that are closely related to ours; see [33] for a recent survey on display advertising.

Channel coordination. An important challenge in the display advertising market is coordinating the contract and auctions sale channels.[3] A natural approach is to first try to sell an impression in an auction, and if the impression is not sold in the auction, then allocate it to one of the reservation contracts if it matches its targeting criteria. The reserve price in the auction can be determined so that the contracts can meet their guarantees by the end of the campaigns [10]. However, this approach may give rise to adverse selection if the valuations of the contract buyers and bidders in the auction are correlated. For instance, consider an advertisers who is willing to pay on average \$0.05 for an impression to an Angeleno. But most of the impressions to the users that are from zip codes with higher than average income in Los Angeles are sold at the auction. Therefore, the quality of impressions allocated to the advertiser under the deal would deteriorate, which subsequently may result in paying penalties or losing the advertiser in the future. To address this issue, [24] propose a randomized bidding mechanism to allocate a *representative* set of impressions to the buyers. [4] characterize adverse-selection proof mechanisms. They show that under certain assumptions on the correlation of the valuation of the reservation buyers and the bidders in the auction, a simple variation of the second-price auction is adverse-selection proof. Preferred deals can alleviate, if not eliminate, the adverse selection concerns by allowing the buyers to choose the impressions they want before they are sent to the auction.

In this paper, our main focus in on the problem of designing the contracts. Once the contracts are determined, the delivery of the impressions in order to satisfy the contracts can itself be complicated, particularly when there is uncertainty about the volume of impressions; see

Right-To-Choose Auctions. A line of research closely related to our work is the study of "right-to-choose" auctions where the seller auctions off the right-to-choose an item from the available items. These auctions are commonly used to sell real estate ([5, 27, 12]), antiques and jewelry [19], and water rights [3]. It has been observed empirically and via field and lab experiments that right-to-choose auctions can increase revenue compared to sequential auctions. The theoretical justification for the revenue increase is offered for the case of risk-averse buyers ([27, 12]). Recently, [18] show that a variation of right-to-choose auctions can obtain significantly higher *welfare* compared to sequential auctions when the buyers' valuations are subadditive.

Bundling. The intuition for higher deals' revenue compared to auctions is that the seller can bundle impressions over time and charge a higher price for the bundle. Bundling is a well-known tool for price discrimination and revenue maximizing ([1, 34, 8, 13, 26]). For sponsored search auctions, [25] show that the problem of finding the optimal bundling is NP-Hard and propose approximation algorithms; see also [21, 20].

[7] study the problem of selling n items to one buyer. They show that a simple strategy of either pricing each item separately or selling them all as one bundle obtains a constant approximation ratio of the optimal mechanism; see also [29]. Our results provide a constant approximation ratio for the multi-buyer case under the assumptions that n is large and the items are ex-ante identical. We believe that some of the ideas developed in our work can be used to find approximately optimal pricing schemes for more general settings of this problem.

2. SETTING & PRELIMINARIES

We consider a seller (publisher) of a set of items (impressions) to n buyers (advertisers), over a horizon of length T.

Let $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ denote the vector of valuations of the buyers for an impression. We consider the case where the vector of valuations of the buyers for each impression is drawn independently and identically from joint distribution $F : \mathbb{R}^n \to \mathbb{R}$. To simplify the presentation, we assume that the valuation of buyer i is distributed over $[\underline{v}_i, \overline{v}_i]$, $0 \leq \underline{v}_i < \overline{v}_i$, and the margins are bounded and positive on $[\underline{v}_i, \overline{v}_i]$. Therefore, the probability of having two equal positive bids is zero.

We allow the valuations of the buyers to be correlated via a common value component. Namely, $v_i = \nu_i + \eta$, where ν_i is the private signal of the buyer, which is distributed independently of the other buyers' valuations; η is the common value component.

The buyers are risk-neutral with quasi-linear utility: the utility buyer i obtains from purchasing the impression at price ρ is equal to $v_i - \rho$.

2.1 Preferred Deals

We define a **preferred deal** using three parameters $\mathcal{I}$, ρ, and μ: by accepting the deal, buyer i gains access to the impressions in $\mathcal{I}_i$ and agrees to buy a fraction μ_i of them, each at price ρ_i.

The above definitions follows industry practice [28]. We note that $\mu_i = 1$ corresponds to the traditional reservation contracts where the buyer specifies a set desired impressions, $\mathcal{I}$, using different features (e.g., context of webpage, demographics of users, etc.) and agrees to purchase all the impressions (satisfying those constraints) that the seller sends to him. In practice, the deals can be specified with the number of impressions or the total spending instead of the fraction. We note that because the price-per-impression is fixed and the total number of impressions is very large, all these forms of contracts are essentially equivalent.

We assume that the minimum purchase constraint has to be met (only) in expectation. This is justified by the fact that the number of impressions in practice is quite large. Let F_i denote the distribution of valuations of buyer i. In order to purchase μ_i fraction of the impressions, buyer i will have to purchase all impressions with $v_i \geq \theta_i$ where $\theta_i = F_i^{-1}(1-\mu_i)$ (i.e., $\mu_i = 1-F_i(\theta_i)$). Note that buyer i will purchase an impression if $v_i \geq \rho_i$. Hence, the minimum purchase is binding only if $1 - F_i(\rho_i) < \mu_i$. The per impression expected utility of the buyer from deal $(\mathcal{I}_i, \mu_i, \rho_i)$ is equal to $(1 - F_i(\theta_i))\big(\mathrm{E}_{\mathcal{I}}[v_i|v_i \geq \theta_i] - \rho_i\big)$ if the minimum purchase is binding; otherwise, it is equal to $(1 - F_i(\rho_i))\big(\mathrm{E}_{\mathcal{I}}[v_i|v_i \geq \rho_i] - \rho_i\big)$.

[3]Channel coordination has been studied in other context of operations management; for example, see [14] on online retail.

We study the seller's optimization problem to maximize the revenue using a **sequence of deals**. The seller chooses a priority list $(\pi_1, \pi_2, \cdots, \pi_n)$ where n denotes the number of buyers. The seller offers to each buyer i a preferred deal $(\rho_i, \mu_i, \mathcal{I}_i)$. By accepting the deal, the buyer agrees to purchase fraction μ_i of the impressions in $\mathcal{I}_i$ each at price ρ_i. Each buyer receives impressions that are not purchased by the buyers with higher priorities. For example, all the impressions are sent to π_1, the buyer with the highest priority. The buyer with the second highest priority receives all the impressions that are not purchased by π_1.

We assume that the seller offers each buyer a deal. A more general approach would be to offer deals only to a subset of buyers, and the buyers that are not offered a deal could participate in an auction. To simplify the presentation, we consider the former model, but our results can be extended to the latter.[4]

2.2 Benchmark

We compare the performance of our algorithm with a benchmark mechanism that can extract the whole surplus of the buyers, cf. [11, 31].

Benchmark Mechanism: The following is an optimal direct mechanism for the setting described in Section 2: charge each buyer i an initial payment (or entrance fee) equal to

$$T \times (\mathrm{E}[v_i | v_i = v_{(1)}] - \mathrm{E}[\max_{j \neq i}\{v_j\} | v_i = v_{(1)}]$$

and then allocate each impression via the second-price auction with no reserve. Note that this mechanism charges in advance the expected utility of each buyer from the future second-price auctions. Hence, it obtain the revenue equal to the expected maximum welfare equal to $T \times \max_i\{v_i\}$.

However, the current practice of the online advertising market does **not** allow for such contracts that require an initial payment without clearly specifying the number of allocated impressions. The *market design question* we study in this work is that: *are proffered deals near-optimal?* We remark that these contracts do not require initial-payments and have already been implemented in practice ([28]), but prior to our work, there was no rigorous study of their powers and limitations compared with the optimal mechanism.

As the first step of our analysis, we note that selling each impression individually at auctions may not be able to compete with the benchmark mechanism.[5]

Proposition 1 (Revenue Gap). *The gap between the revenues of the benchmark mechanism and the second price auction (with an optimal reserve) can be arbitrarily larger.*

Proof. We prove the claim using the following example. Suppose there is only one buyer whose valuations is distributed according to a (truncated) revenue-equal distribution, $F(v) = \frac{M}{M-1}\left(1 - \frac{1}{v}\right)$, $v \in [1, M]$, for a large constant M. Observe that the revenue of any posted price p is at most equal to 1.

$$\begin{aligned} p(1 - F(p)) &= p\left(1 - \frac{M}{M-1}\left(1 - \frac{1}{p}\right)\right) \\ &\leq \frac{M}{M-1}\left(1 - \frac{1}{M}\right) = 1. \end{aligned}$$

On the other hand, the benchmark mechanism obtains revenue equal to $E[v] = O(\log M)$.

$$E[v] = \int_1^M x \times \frac{M}{M-1}\frac{1}{x^2}dx = \frac{M}{M-1}\log M.$$

Therefore, the gap between the revenues grows larger as M increases. $\square$

In contrast with second-price auctions, we observe that preferred deal can obtain the optimal revenue when there is only one buyer using $\mu = 1$ and $\rho = \mathrm{E}[v]$. The intuition is that deals can obtain higher revenue because they *bundle* the impressions *over time* and charge the price of the bundle instead of selling each impression individually. For multiple buyers, the preferred deals may not be optimal, but as we show later, they can obtain near-optimal revenue.

In the above example and proposition, we assumed that the seller can extract all the surplus of the buyer. This implicitly implies that the seller is a monopoly and has full bargaining power (no outside option for the buyer). But from a practical perspective, this may not quite be the case. Online publishers have a monopoly on the advertising space on their webpages and to some extent on the users who visit their websites. If an advertiser aims to target the New York Times visitors, he has to buy impressions from the New York Time. Nevertheless, some of those users, or users similar to them, can be reached via other websites. In practice, the terms of these contracts can be negotiated and the surplus is somehow shared between the seller and the buyer, for instance via offering a deal at a lower price. In this paper, we do not consider the bargaining solution and the division of the surplus, and mainly in order to simplify the presentation, we assume that the seller can extract the surplus of the buyer. However, as shown by our theoretical result and later in our empirical analysis in Section 5, we emphasize that the gap between the revenues obtained from the deals and the auction could be dramatic, even if the surplus is shared. Therefore, even a seller with limited bargaining power could benefit from deals and contracts.

3. CHERRY-PICKING AND HARDNESS

One of the main challenges in finding the optimal sequence of preferred deals is *cherry-picking*. A buyer picks the impressions that have the highest value for him, but there could exist other buyers with higher valuations for those impressions. Therefore, there are two sources of inefficiency from a sequence of preferred deals: i) The buyer with the highest valuation does not purchase the impression because the price was too high. ii) An impression goes to a buyer with a lower valuation because that buyer had higher priority. The latter appears to be the bigger challenge. Due to these complexities, it is not obvious how to design even a simple greedy algorithm for this problem. A naive greedy algorithm may allocate all the impressions to the first buyer it picks if the buyer has positive valuations for all the impressions. In

[4]We examined the latter model empirically and observe that the seller would prefer to offer deals to all the buyers instead of a combination of deals and auction.

[5]For correlated valuations, the whole surplus can also be extracted via randomized static mechanisms [17]. But we note that the revenue gap presented in Proposition 1 holds for independent valuations, which is an important special case of our model where surplus extraction is not possible via auctioning off each impression individually.

fact, in this rest of this section, we show that the problem of finding the optimal sequence of deals is computationally hard.

Let $\omega(\mathcal{I}, S)$ denote the maximum expected per impression revenue that can be obtained using a sequence of deals from impressions in $\mathcal{I}$ and the set of buyers S. If S has only one element ($|S| = \{i\}$), then $\omega(\mathcal{I}, S) = \mathrm{E}_{\mathcal{I}}[v_i]$, which corresponds to the preferred deal ($\rho_i = \mathrm{E}_{\mathcal{I}}[v_i], \mu_i = 1, \mathcal{I}$). For $|S| \geq 2$, we can define ω recursively.

$$\omega(\mathcal{I}, S) = \max_{j \in S, \theta_j \in [\underline{v}_j, \overline{v}_j]} \Big\{ \mathrm{Pr}_{\mathcal{I}}[v_j \geq \theta_j] \ \mathrm{E}_{\mathcal{I}}[v_j | v_j \geq \theta_j] + \mathrm{Pr}_{\mathcal{I}}[v_j < \theta_j] \ \omega(\mathcal{I} \setminus \{v_j \geq \theta\}, S \setminus \{j\}) \Big\}. \tag{1}$$

In the equation above, $\mathcal{I} \setminus \{v_j \geq \theta\}$ denotes the set of impressions in $\mathcal{I}$ where $v_j < \theta$. Suppose buyer i, at threshold θ_i, maximizes the r.h.s. The seller offers to i preferred deal ($\rho_i = \mathrm{E}_{\mathcal{I}}[v_i | v_i \geq \theta_i], \mu_i = \Pr[v_i \geq \theta_i], \mathcal{I}$). Note that the parameters of the deals are determined such that they extract the buyer's surplus; see the discussion after Example 1. Unfortunately, solving the recursion above takes exponential time.

Theorem 1 (Computational Hardness). *The problem of finding an optimal (revenue-maximizing) sequence of preferred deals is NP-Hard.*

We present a proof in Appendix A. We use a reduction from a well-known NP-Hard problem, the *maximum acyclic subgraph* [32]. As suggested by its name, the problem is defined as finding the largest acyclic subgraph of a given directed graph. If the graph is weighted, then the goal is to find a subset that maximizes the sum of the weights of edges in the acyclic subgraph. This problem is equivalent to finding an ordering (ranking) of the vertices that maximizes the number (or the total weight) of the *forward edges* (those consistent with the orderings that are directed from nodes with higher rankings to lower rankings).[6]

4. AUCTION-ADJUSTED GREEDY ALGORITHM

We present an algorithm called Auction-Adjusted Greedy ($\mathcal{AAG}$) in Figure 1. The algorithm recursively chooses one of the buyers, denoted by i, and offers him a contract $(\mu_i, \rho_i, \mathcal{I}_i)$. The choice for θ is guided by the auction. Consider the first step. θ_j is chosen such that the buyer wins exactly the same number of impressions as he would have in a second-price auction with no reserve price. Note that the buyer wins the same number of impressions, however, now he can cherry-pick and choose the impressions that maximize his valuations. Therefore, we have

$$\mathrm{E}\left[v_i | v_{(1)}(S) = v_i\right] \leq \mathrm{E}\left[v_i | v_i \geq \theta_i\right], \tag{2}$$

where $v_{(1)}(S) = \max_{i \in S}\{v_i\}$. On the other hand, as v_i may not be the highest bid for impressions $v_i \geq \theta_i$, the algorithm may lose revenue. We approximate the "opportunity cost" of allocating impressions to v_i with the expected highest valuation of the other buyers, $E_{\mathcal{I}}[v_{(1)}(S \setminus \{i\})]$. In other words, we use the expected value of the highest v_i as the proxy for $\omega(\mathcal{I}, S)$.

The algorithm chooses buyer i with the highest "bang for the buck." Buyer i maximizes the ratio of the expected revenue from the impressions allocated to the buyer divided by the opportunity cost.

We show that the algorithm obtains a constant fraction of the optimal revenue. We use $\mathrm{E}[\max_i\{v_i\}]$, which is the trivial upper bound on the revenue as the benchmark.

Figure 1: Auction-Adjusted Greedy ($\mathcal{AAG}$)

Input: Stream of impressions $\mathcal{I}$ and a set of buyers S.
Output: Sequence of Deals $\mathcal{D}$.

If ($|S| = \{i\}$)
  Return ($\rho_i = E_{\mathcal{I}}[v_i], \mu_i = 1$).

For each buyer $j \in S$, choose θ_j such that:

$$\Pr[v_j \geq \theta_j] = \Pr\left[v_i = v_{(1)}(S)\right].$$

Choose buyer

$$i \in \arg\max_j \{\mathrm{E}_{\mathcal{I}}[v_j | v_j \geq \theta_j] / \mathrm{E}_{\mathcal{I}}[v_{(1)}(S \setminus \{j\}) | v_j \geq \theta_j]\}. \tag{3}$$

$\mu_i \leftarrow \mathrm{Pr}_{\mathcal{I}}[v_i \geq \theta_i]$.
$\rho_i \leftarrow \mathrm{E}_{\mathcal{I}}[v_i | v_i \geq \theta_i]$.
$\mathcal{D} \leftarrow (\rho_i, \mu_i, \mathcal{I}) + \mathcal{AAG}(\mathcal{I} \setminus \{v_i \geq \theta_i\}, S \setminus \{i\})$.
Return $\mathcal{D}$

Theorem 2. *The expected per impression revenue of the sequence of deals found by the Auction-Adjusted Greedy algorithm is at least equal to* $\frac{1}{3}\mathrm{E}[\max_i\{v_i\}]$. *For the independent valuation* ($\eta_t = 0, t \in \mathcal{I}$), *the algorithm obtains at least half of the optimal solution,* $\frac{1}{2}\mathrm{E}[\max_i\{v_i\}]$.

In Appendix B, we prove the theorem above using induction. By Eq. (2), the algorithm obtains at least the same revenue as the benchmark from the chosen buyer. We then bound the revenue of the benchmark from the impressions allocated to the chosen buyer.

For independent valuations, we show that there always exists a buyer such that for him the proxy opportunity cost, $\mathrm{E}[v_{(1)}(S \setminus \{j\}) | v_j \geq \theta_j] = \mathrm{E}[v_{(1)}(S \setminus \{j\})]$, is less than or equal to the revenue that can be obtained from that buyer via a deal, $\mathrm{E}[v_j | v_j \geq \theta_j]$. Because the algorithm chooses a buyer i that maximizes the ratio of the revenue to cost, for chosen buyer i we have $\mathrm{E}_{\mathcal{I}}[v_{(1)}(S \setminus \{i\})] \leq \mathrm{E}_{\mathcal{I}}[v_i | v_i \geq \theta_i]$. Similarly, for correlated valuations, we show that there exists a buyer such that $\mathrm{E}[v_{(1)}(S \setminus \{j\}) | v_j \geq \theta_j] \leq 2 \ \mathrm{E}[v_j | v_j \geq \theta_j]$. To do so, we map the set of impressions for which buyer j has the highest valuations to the set of impressions where $v_j \geq \theta_j$. For the independent valuations, the expected valuations of the highest bid among other bidders, $\mathrm{E}[v_{(1)}(S \setminus \{j\})]$, are the same for both sets. For the correlated valuations, the expected value is higher for the impressions in the second set

[6]For the maximum acyclic subgraph problem, choosing a random ordering of nodes obtains a $\frac{1}{2}$ approximation. For our problem, choosing a random order does not immediately lead to a good algorithm because the algorithm has to decide on the price and minimum-purchase requirement for each deal.

(i.e., $\mathrm{E}[v_{(1)}(S\setminus\{j\})|v_j > v_{(1)}(S\setminus\{j\})] \leq \mathrm{E}[v_{(1)}(S\setminus\{j\})|v_j > \theta_j]$ because the valuations are positively correlated), hence the difference in the approximation ratio.

In the next section, we show that on real-data the algorithm may obtain higher revenue than the worst-case bounds provided by the theorem.

5. EMPRICAL EVALUATION

We evaluate our algorithm by simulating it over logs of auction data. The data was collected from Google's Doubleclick ad exchange platform. On the day of data collection, we found the top 4 ad units that generated the highest revenue. For each of those ad units, we focus on the top 10 advertisers with the highest total bid (summed over all the impressions on that day). We then create a dataset for each ad unit by sampling 0.1% of the auctions that day and collecting the bids of the 10 chosen advertisers; if an advertiser does not participate in an auction, we let his bid be equal to 0. The number of impressions (auctions) in our four datasets ranges from $200,000$ to 2 million. Unfortunately, due to the proprietary nature of our data, we are not able to provide statistical information about the bids.

The Auction-Adjusted Greedy algorithm is defined in Figure 1 using the distribution of the valuations. In simulations, we use the actual realizations instead of a distribution. More specifically, we assume that the bids are equal to the valuations because bidding truthfully is a dominant strategy in second-price auctions.[7] Given a subset of impressions $\mathcal{I}$, we use the fraction of auctions where j has the highest bid as $\Pr_{\mathcal{I}}[v_j = v_{(1)}]$. Similarly, for $\mathrm{E}_{\mathcal{I}}[v_j|v_j \geq \theta_j]$, we use the average of v_j calculated over the impressions in $\mathcal{I}$ where the bid of advertiser j is larger than or equal to θ_j.

We compare the auction-adjusted greedy algorithm with four other mechanisms. The first algorithm is the second-price auction with no reserve price. We then optimize the reserve price which will remain constant across the auctions. Namely, we discretize the space of bids and numerically find the revenue-maximizing reserve. We also look at the second-price auction with personal reserve prices where the reserves for buyers may differ and the impression is allocated to the highest bidder whose bid is above his personal reserve. We calculate the optimal reserve for each buyer individually by finding the optimal posted-price for the buyer in the absence of other bidders.

In addition, we consider a *Max-Margin Greedy Algorithm*. This algorithm at each step chooses the buyer with the highest margin, the revenue from the buyer minus the proxy opportunity cost; namely,

$$j \leftarrow \arg\max_{j\in S}\left\{\max_{\theta}\left\{\Pr[v_j \geq \theta] \times \mathrm{E}[v_j - v_{(1)}(S\setminus\{j\})|v_j \geq \theta]\right\}\right\}.$$

The results are presented in Table 1 as the percentage of the benchmark. Recall that the benchmark is the first best solution, the sum of the highest valuation for each impression which is equal to the welfare obtained by the second-price auction with no reserve. Surprisingly, our algorithm obtains about 90% of the optimal revenue. Recall that there is significant potential for loss of welfare (and revenue) for preferred deals because of the imposed ordering. An impression may be allocated to a buyer with a lower valuation because that buyer had higher priority. However, our study suggests that the adverse impact of cherry picking might be limited.

Our algorithm outperforms the other greedy algorithm. As suggested by theory, both preferred deal mechanisms obtain significantly higher revenue than the second-price auctions.

Among the reserve price mechanisms, the one with personal reserve yields the highest revenue, but at the cost of reducing the efficiency (due to high reserve prices). Note that the auction-adjusted algorithm obtains a higher welfare than this auction in 3 out 4 of the datasets. As discussed in Section 2.2, even if the seller shares the surplus equally with the buyer, by reducing the price of the deal, the corresponding deal obtains more revenue, and almost-equal if not higher welfare, than the auction with large reserve prices.

6. CONCLUSION

One of our main goals in this work is to understand the power and limitations of preferred deals. Our numerical results are quite promising and suggest that there is significant potential to increase the revenue of the publishers using the preferred deals, if the minimum purchase requirements are determined appropriately. Our algorithm can serve as a baseline that identifies potential profitable deals and their parameters.

We considered only the strict priority rule in this paper. Another possibility is that two or more buyers can be assigned to the same priority level with a tie-breaking rule when more than one buyers are willing to purchase the impression. One way to break ties is that those buyers bid for the impression. For instance, this corresponds to a second-price auction with personal reserve if no minimum purchase is specified and all the buyers have the same priority. Other solutions could be random allocation or assigning an impression to the contract that is behind in its fulfillment. Other directions for future research include analyzing settings where buyers are budget-constrained (cf. [23, 9]) or arrive over time (cf. [2, 16, 6]).

We believe that further research in this area can shape the future of the display advertising market by determining how the deals and contracts will be implemented in practice.

APPENDIX

A. PROOF OF THEOREM 1

Consider an instance of the maximum acyclic subgraph problem, given by directed graph $G(V,E)$ with n vertices and m edges. Given an instance of the maximum acyclic subgraph problem, we construct the following instance of the preferred deal optimization problem.

The instance has n buyers, each corresponding to a node, and $m+n\bar{d}$ "types" of items where $\bar{d}$ is equal to the maximum degree in the graph plus one. For each edge $e = (i,j) \in E(G)$, we add a type e where the valuation of buyer i for an item of type e is equal to 3, and the valuation of buyer j for an item of type e is equal to 2. Other buyers' valuations for type e items are set equal to 0. Finally, we add $\bar{d}$ "exclusive" types of items for each buyer i with value 1. Now let the number of items be equal to $T = \tau\big(n\bar{d} + m\big)$ where τ is the scaling parameter. The items are drawn uniformly from the

[7]See [15] for a discussion on the pros and cons of this assumption.

Mechanisms		Dataset 1		Dataset 2		Dataset 3		Dataset 4	
Preferred Deals	Auction-Adjusted Greedy	87%		88%		94%		87%	
	Max-Margin Greedy	82%		70%		94%		86%	
2nd-Price Auction	No Reserve	28%	100%	33%	100%	18%	100%	21%	100%
	Optimal Reserve	33%	87%	37%	99%	44%	80%	25%	95%
	Personal Reserve	41%	83%	42%	90%	49%	88%	52%	79%

Table 1: Respectively revenue and welfare as the fraction of the benchmark. For the preferred deals, revenue is equal to welfare.

types, i.e., an item belongs to each type with probability equal to $1/(n\bar{d}+m)$.

To convey some intuition, we start with the following lemma.

LEMMA 1. *For any solution with value k to an instance of the maximum acyclic subgraph problem, there exists a solution with the expected revenue of $\tau\big((2m+k)+n\bar{d}\big)$ to the corresponding instance of the deal optimization problem.*

PROOF. We construct the following sequence of deals. The priorities are the same as the ordering of the nodes. Each buyer i is offered a deal such that they purchase all "available" items, including the $\bar{d}$ exclusive items. In other words, the threshold θ for all buyers will be equal to 1; see (1). From each item that corresponds to a forward edge (edge in the acyclic subgraph), the revenue is equal to 3 because the buyer with the higher valuation comes first in the ordering and has a higher priority. Therefore, we obtain revenue 3 from each forward edge and revenue 2 from backward edges, which adds up to $\big(3k+2(m-k)\big)\tau = \big(k+2m\big)\tau$. The expected revenue from exclusive items is equal to $\tau n\bar{d}$, which completes the proof. □

We now prove the other direction, starting with the following lemma.

LEMMA 2. *For the problem instance defined above, in the (optimal) solution of (1), the thresholds (θ_is) are equal to 1 and all the exclusive items are allocated to the buyers.*

PROOF. We prove the claim via contradiction. We show that the revenue of any solution can be increased by allocating to each buyer all her exclusive items. Consider buyer i. It is easy to see that in an optimal solution all the "available" items with valuation 3 should be allocated to buyer i. We now consider the following two cases:

- If some (more than 0), but not all, exclusive items are allocated to buyer i, then we can simply increase the revenue by allocating all the exclusive items and setting ρ equal to the expected value of the items allocated to her.

- Suppose no exclusive items are allocated to buyer i. This means that $\rho > 1$. Suppose we reduce ρ to 1. Note that we may lose some revenue from those items that buyer i values at 2, but there are other buyers who value them at 3. The expected total loss from those items is at most equal to τd_i where d_i denotes the degree of note i. Because $\bar{d} > d_i$, the revenue will be increased.

□

We are now ready to prove the theorem.

LEMMA 3. *If there exists a sequence of preferred deals with the expected revenue of $\tau\big((2m+k)+n\bar{d}\big)$ for the problem instance described above, then, there exists an ordering with k forward edges for the corresponding maximum acyclic subgraph problem.*

PROOF. By Lemma 2, if there exists a sequence of preferred deals of total revenue $\tau\big((2m+k)+n\bar{d}\big)$ for the problem instance described above, then there exists a sequence of deals of at least the same revenue such that each buyer purchases all the items available to her. Note that if a buyer purchases any exclusive items, she will also purchase all the available items with value 2 or 3. The total revenue from the exclusive items is equal to $\tau\big(n\bar{d}\big)$ plus 3 from the forward edges and 2 from the backward edges. Therefore, the expected revenue $\tau\big((2m+k)+n\bar{d}\big)$ implies that there are k types of value 3 in the solution that correspond to k forward edges for the solution to the acyclic subgraph problem with the same ordering as the deals. □

B. PROOF OF THEOREM 2

We prove Theorem 2 via induction on the number of the buyers. For $|S| = 1$, the claim trivially follows using deal $(\rho = E[v], \mu = 1)$. We now consider $|S| \geq 2$. We drop the dependence on S when it is clear from the context. Let i be the first buyer chosen by the algorithm; see Eq. (3). We start with providing an upper bound on the benchmark.

$$\begin{aligned}
& \mathrm{E}\left[v_{(1)}(S)\right] \\
= \;& \mathrm{E}\left[v_{(1)}\ \mathbf{1}(v_{(1)} = v_i)\right] + \mathrm{E}\left[v_{(1)}\ \mathbf{1}(v_{(1)} \neq v_i, v_i \geq \theta_i)\right] \\
& +\mathrm{E}\left[v_{(1)}\ \mathbf{1}(v_{(1)} \neq v_i, v_i < \theta_i)\right] \\
\leq \;& \mathrm{E}\left[v_{(1)}\ \mathbf{1}(v_{(1)} = v_i)\right] + \mathrm{E}\left[v_{(1)}(S \setminus \{i\})\ \mathbf{1}(v_i \geq \theta_i)\right] \\
& +\mathrm{E}\left[v_{(1)}(S \setminus \{i\})\ \mathbf{1}(v_i < \theta_i)\right] \\
\leq \;& \Pr[v_i \geq \theta_i]\ \left(\mathrm{E}\left[v_i | v_i \geq \theta_i\right] + \mathrm{E}\left[v_{(1)}(S \setminus \{i\}) | v_i \geq \theta_i\right]\right) \\
& + \Pr[v_i < \theta_i]\ \mathrm{E}\left[v_{(1)}(S \setminus \{i\}) | v_i \leq \theta_i\right] \qquad (4)
\end{aligned}$$

The last inequality follows from Eq. (2). The algorithm obtains expected per-impression revenue of

$$\Pr[v_i \geq \theta_i]\mathrm{E}\left[v_i | v_i \geq \theta_i\right].$$

Therefore, to prove the theorem, using the induction hypothesis, it suffices to show that $\mathrm{E}\left[v_i | v_i \geq \theta_i\right] \geq \mathrm{E}\left[v_{(1)}(S \setminus \{i\})\right]$ for independent valuations and

$$2 \times \mathrm{E}\left[v_i | v_i \geq \theta_i\right] \geq \mathrm{E}\left[v_{(1)}(S \setminus \{i\}) | v_i \geq \theta_i\right]$$

for correlated valuations. We prove these inequalities respectively in Lemmas 6 and 7.

Recall that for any buyer j, we have $v_j = \eta + \nu_j$ where η is the common value component and ν_j is distributed independently among the buyers. Let f^η and f_j^ν respectively denote the p.d.f. of distributions of η and ν_j.

Using the fact that ν_i is distributed independently of ν_j, we prove the following lemma.

LEMMA 4. *For any buyer j, we have*

$$\begin{aligned} & E[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j] \qquad (5) \\ = \; & \Pr[v_j \geq \theta_j] \; E[\nu_j|v_j \geq \theta_j] - \Pr[v_j \geq \theta_j] E\left[\max_{k \in S\setminus\{j\}} \{\nu_k\}\right] \end{aligned}$$

PROOF. For any buyer j, we have

$$\begin{aligned} & \mathrm{E}[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j] \\ = \; & \int_{\underline{\eta}_j}^{\overline{\eta}_j} \int_{\underline{\nu}_j}^{\overline{\nu}_j} \left(\mathbf{1}(\eta + \nu \geq \theta_j) \; \mathrm{E}[v - v_{(1)}(S \setminus \{j\})|v_j = \eta + \nu] \right) \\ & \times f^\eta(\eta) f_j^\nu(\nu) d\eta d\nu \\ & \times f^\eta(\eta) f_j^\nu(\nu) d\eta d\nu \\ = \; & \int_{\underline{\eta}_j}^{\overline{\eta}_j} \int_{\underline{\nu}_j}^{\overline{\nu}_j} \left(\mathbf{1}(\eta + \nu \geq \theta_j) \; \left(\nu - \mathrm{E}\left[\max_{k \in S\setminus\{j\}} \{\nu_k\}\right]\right) \right) \\ & \times f^\eta(\eta) f_j^\nu(\nu) d\eta d\nu \\ = \; & \Pr[v_j \geq \theta_j] \; \mathrm{E}[\nu_j|v_j \geq \theta_j] \\ & - \Pr[v_j \geq \theta_j] \mathrm{E}\left[\max_{k \in S\setminus\{j\}} \{\nu_k\}\right] \end{aligned}$$

□

Recall that $v_{(1)}$ is the random variable corresponding to the highest valuation. We now define random variable $\tilde{v}$ as follows. Let $v_\ell = \eta + \nu_\ell$ be the highest valuation. For all $k \neq \ell$, *re-sample* ν_k and let $\tilde{\nu}_k$ denote the new realization. Let $\tilde{v} = \eta + \max_{k \neq \ell}\{\tilde{\nu}_k\}$. We observe that $\mathrm{E}[v_{(1)}] \geq \mathrm{E}[\tilde{v}]$. The reason is that $\tilde{v}$ represents the highest expected value among a smaller number of buyers. Using this observation, we obtain the following result.

LEMMA 5. *We have*

$$E[v_{(1)}] - E[\eta] \;\; \geq \;\; \sum_{j=1}^{n} \Pr\left[v_j = v_{(1)}\right] E\left[\max_{k \neq j}\{\nu_k\}\right]$$

We are now ready to prove the induction step.

LEMMA 6. *Let i be the chosen buyer. For independent valuations, we have $E[v_i|v_i \geq \theta_i] \geq E\left[v_{(1)}(S \setminus \{i\})\right]$.*

PROOF. Summing up Eq. (5), we have

$$\begin{aligned} & \sum_{j=1}^{n} \mathrm{E}\left[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j\right] \\ \geq \; & \sum_{j=1}^{n} \Pr[v_j \geq \theta_j] \; \left(\mathrm{E}[v_j|v_j \geq \theta_j] - \mathrm{E}\left[\max_{i \neq j}\{v_i\}\right]\right) \qquad (6) \\ = \; & \sum_{j=1}^{n} \Pr[v_j = v_{(1)}] \; \left(\mathrm{E}[v_j|v_j \geq \theta_j] - \mathrm{E}\left[\max_{i \neq j}\{v_i\}\right]\right) \qquad (7) \\ \geq \; & \sum_{j=1}^{n} \Pr[v_j = v_{(1)}] \; \left(\mathrm{E}[v_j|v_j = v_{(1)}] - \mathrm{E}\left[\max_{i \neq j}\{v_i\}\right]\right) \qquad (8) \\ = \; & \mathrm{E}[v_{(1)}] - \sum_{j=1}^{n} \Pr[v_j = v_{(1)}] \mathrm{E}\left[\max_{i \neq j}\{v_i\}\right] \\ \geq \; & 0 \end{aligned}$$

Eq. (7) follows from definition of θ_j and (8) follows from (2). We obtain the last inequality using Lemma 5. Note that because the sum of $\mathrm{E}\left[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j\right]$ is non-negative, there exists buyer k for whom $\mathrm{E}\left[v_k - v_{(1)}(S \setminus \{k\})|v_k \geq \theta_k\right] \geq 0$. For that buyer $\mathrm{E}\left[v_k|v_k \geq \theta_k\right]/\mathrm{E}\left[v_{(1)}(S \setminus \{k\})\right] \geq 1$. By definition, (3), for buyer i we have

$$\mathrm{E}\left[v_i|v_i \geq \theta_i\right]/\mathrm{E}\left[v_{(1)}(S \setminus \{i\})\right] \geq 1.$$

□

LEMMA 7. *For correlated valuations (i.e., $\eta > 0$), we have $2 \; E[v_i|v_i \geq \theta_i] \;\; \geq \;\; E\left[v_{(1)}(S \setminus \{i\})|v_i \geq \theta_i\right]$.*

PROOF. Similar to the previous lemma, summing up Eq. (5), we have

$$\begin{aligned} & \sum_{j=1}^{n} \mathrm{E}\left[v_j|v_j \geq \theta_j\right] + \sum_{j=1}^{n} \mathrm{E}\left[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j\right] \\ \geq \; & \mathrm{E}[v_{(1)}] + \sum_{j=1}^{n} \mathrm{E}\left[v_j - v_{(1)}(S \setminus \{j\}|v_j \geq \theta_j\right] \qquad (9) \\ = \; & \mathrm{E}[v_{(1)}] \qquad (10) \\ & + \sum_{j=1}^{n} \Pr[v_j \geq \theta_j] \; \left(\mathrm{E}[\nu_j|v_j \geq \theta_j] - \mathrm{E}\left[\max_{i \in S\setminus\{j\}}\{\nu_i\}\right]\right) \\ \geq \; & \sum_{j=1}^{n} \Pr[v_j \geq \theta_j] \; \mathrm{E}[\nu_j|v_j \geq \theta_j] + \mathrm{E}[\eta] \geq 0 \qquad (11) \end{aligned}$$

Using (2) and (5) respectively, we obtain (9) and (10). Inequality (11) follows from Lemma 5, and the last inequality holds simply because all the terms are non-negative. Note that because the sum of $\mathrm{E}\left[2v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j\right]$ is non-negative, there exists buyer k for whom

$$\mathrm{E}\left[2v_k - v_{(1)}(S \setminus \{k\})|v_k \geq \theta_k\right] \geq 0.$$

For that buyer, we have

$$2\mathrm{E}\left[v_k|v_k \geq \theta_k\right] \geq \mathrm{E}\left[v_{(1)}(S \setminus \{k\})|v_k \geq \theta_k\right].$$

By definition, (3), for buyer i we have

$$2\mathrm{E}\left[v_i|v_i \geq \theta_i\right] \geq \mathrm{E}\left[v_{(1)}(S \setminus \{i\})|v_i \geq \theta_i\right].$$

□

Acknowledgment.

We are grateful to Umar Syed for helping us with data collection. We would like to thank Nitish Korula, Max Lin, Martin Pál, Maher Said, and Christopher Tignor for their valuable comments and suggestions. This work was supported in part by a Google Faculty Research Award.

C. REFERENCES

[1] W. J. Adams and J. L. Yellen. Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics*, pages 475–498, 1976.

[2] S. Alaei, E. Arcaute, S. Khuller, W. Ma, A. Malekian, and J. Tomlin. Online allocation of display advertisements subject to advanced sales contracts. In *Proceedings of the third international workshop on data mining and audience intelligence for advertising*, pages 69–77, 2009.

[3] J. E. Alevy, O. Cristi, O. Melo, et al. Right-to-choose auctions: A field study of water markets in the Limari valley of Chile. *Agricultural and Resource Economics Review*, 39(2):213–226, 2010.

[4] N. Arnosti, M. Beck, and P. Milgrom. Adverse selection and auction design for internet display advertising. *American Economic Review*, 2017.

[5] O. Ashenfelter and D. Genesove. Testing for price anomalies in real-estate auctions. *The American Economic Review*, 82(2):501, 1992.

[6] M. Babaioff, J. D. Hartline, and R. D. Kleinberg. Selling ad campaigns: online algorithms with cancellations. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 61–70. ACM, 2009.

[7] M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A simple and approximately optimal mechanism for an additive buyer. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 21–30, 2014.

[8] Y. Bakos and E. Brynjolfsson. Bundling information goods: Pricing, profits, and efficiency. *Management Science*, 45(12):1613–1630, 1999.

[9] S. R. Balseiro, O. Besbes, and G. Y. Weintraub. Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Science*, 61(4):864–884, 2015.

[10] S. R. Balseiro, J. Feldman, V. Mirrokni, and S. Muthukrishnan. Yield optimization of display advertising with ad exchange. *Management Science*, 60(12):2886–2907, 2014.

[11] D. P. Baron and D. Besanko. Regulation, asymmetric information, and auditing. *The RAND Journal of Economics*, pages 447–470, 1984.

[12] R. Burguet. Right to choose in oral auctions. *Economics Letters*, 95(2):167–173, 2007.

[13] G. P. Cachon and P. Feldman. Pricing services subject to congestion: Charge per-use fees or sell subscriptions? *Manufacturing & Service Operations Management*, 13(2):244–260, 2011.

[14] R. A. Caldentey and G. Vulcano. Online auction and list price revenue management. *Management Science*, 5(53):795–813, 2007.

[15] L. E. Celis, G. Lewis, M. Mobius, and H. Nazerzadeh. Buy-it-now or take-a-chance: Price discrimination through randomized auctions. *Management Science*, 60(12):2927–2848, 2014.

[16] F. Constantin, J. Feldman, S. Muthukrishnan, and M. Pál. An online mechanism for ad slot reservations with cancellations. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1265–1274. Society for Industrial and Applied Mathematics, 2009.

[17] J. Cremer and R. P. McLean. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–57, November 1988.

[18] N. R. Devanur, J. Morgenstern, and V. Syrgkanis. Draft auctions. *Working Paper*, 2013.

[19] K. Eliaz, T. Offerman, and A. Schotter. Creating competition out of thin air: An experimental study of right-to-choose auctions. *Games and Economic Behavior*, 62(2):383–416, 2008.

[20] Y. Emek, M. Feldman, I. Gamzu, R. PaesLeme, and M. Tennenholtz. Signaling schemes for revenue maximization. *ACM Transactions on Economics and Computation*, 2(2):5, 2014.

[21] E. Even-Dar, M. Kearns, and J. Wortman. Sponsored search with contexts. In *Internet and Network Economics*, pages 312–317, 2007.

[22] U. Feige, N. Immorlica, V. S. Mirrokni, and H. Nazerzadeh. A combinatorial allocation mechanism with penalties for banner advertising. In *Proceedings of the 17th International Conference on World Wide Web* (WWW), pages 169–178, 2008.

[23] J. Feldman, N. Korula, V. Mirrokni, S. Muthukrishnan, and M. Pál. Online ad assignment with free disposal. In *Internet and network economics*, pages 374–385. Springer, 2009.

[24] A. Ghosh, P. McAfee, K. Papineni, and S. Vassilvitskii. Bidding for representative allocations for display advertising. In *Internet and Network Economics*, pages 208–219. Springer, 2009.

[25] A. Ghosh, H. Nazerzadeh, and M. Sundararajan. Computing optimal bundles for sponsored search. In *Internet and Network Economics, Third International Workshop (WINE)*, pages 576–583, 2007.

[26] S. M. Gilbert, R. S. Randhawa, and H. Sun. Optimal per-use rentals and sales of durable products and their distinct roles in price discrimination. *Production and Operations Management*, 23(3):393–404, 2014.

[27] J. K. Goeree, J. Wooders, and C. R. Plott. Bidders' choice auctions: Raising revenues through the right to choose. *Journal of the European Economic Association*, pages 504–515, 2004.

[28] Google AdX Documentations. Preferred deals overview. `https://support.google.com/adxbuyer/answer/2604595?hl=en&ref_topic=2813049`, 2015. Accessed: 2015-04-1.

[29] S. Hart and N. Nisan. Approximate revenue maximization with multiple items. In *Proceedings of the 13th ACM Conference on Economics and Computation*, pages 656-656, 2012.

[30] M. Helft and T. Vega. Retargeting ads follow surfers to other sites. *New York Times*, August 30 2010. `http://www.nytimes.com/2010/08/30/technology/30adstalk.html`.

[31] S. M. Kakade, I. Lobel, and H. Nazerzadeh. Optimal dynamic mechanism design and the virtual pivot mechanism. *Operations Research*, 61(4):837–854, 2013.

[32] R. M. Karp. *Reducibility among combinatorial problems*. Springer, 1972.

[33] N. Korula, V. Mirrokni, and H. Nazerzadeh. Optimizing display advertising markets: Challenges and directions. *IEEE Internet Computing*, 2016.

[34] R. P. McAfee, J. McMillan, and M. D. Whinston. Multiproduct monopoly, commodity bundling, and correlation of values. *The Quarterly Journal of Economics*, pages 371–383, 1989.

[35] S. Muthukrishnan. Ad exchanges: Research issues. In *Internet and Network Economics, 5th International Workshop (WINE)*, pages 1–12, 2009.

[36] PwC. IAB internet advertising revenue report, 2015. `http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_FY_2014.pdf`.