Deals or No Deals: Contract Design for Online Advertising

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

Hamid Nazerzadeh
Marshall School of Business, University of Southern California, Los Angeles, CA, hamidnz@marshall.usc.edu

Billions of dollars worth of display advertising are sold via contracts and deals. This paper is the first 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 an approximation algorithm for maximizing the revenue that can be obtained from these deals. We evaluate our algorithm using data from Google’s ad exchange platform. Our algorithm obtains about 90% of the optimal revenue. Furthermore, we show, both theoretically and via data analysis, that deals, with appropriately chosen minimum-purchase guarantees, can yield significantly higher revenue than auctions.

Area of review: Optimization.

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 [PwC (2015)].

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 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 [Feige et al. (2008)]. 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 [Muthukrishnan (2009)].

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

---

1 Impression is the unit of inventory and simply refers to the display of an ad to a user.
eneration 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.” (Google AdX Documentations 2015)

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. 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 (Helft and Vega 2010). Those impressions are of high value for advertisers.

Contributions and Organization

In this paper, we initiate the first 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.

First, in Section 2, we consider a special case with one buyer called first-look deals, where the publisher agrees to send the impressions to one buyer (advertiser or an ad network). If the buyer does not purchase an impression, it will be sent to an auction. We show that if the buyer is risk-neutral and has the standard additive valuation, then offering a “price-only” first-look deal is suboptimal. On the other hand, a first-look deal, in addition to the price-per-impression, can enforce a minimum purchase requirement under which the buyer agrees to purchase a certain number of impressions during the campaign. We show that using a combination of price-per-impression and minimum purchase requirement, deals/contracts can significantly increase the revenue of the publisher compared to second-price auctions. We also discuss how publishers can incentivize the buyers to sign such deals.

In Section 3, we consider the multi-buyer case where buyers are assigned different priorities and an impression is offered to a buyer if it is not accepted by the buyers with higher priorities. We show that finding the optimal sequence of deals is NP-complete. We then present a greedy algorithm for a stochastic setting where the valuations of the buyers are drawn from a joint probability distribution. In our algorithm, called the Auction-Adjusted Greedy, 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

\[2\] We note that the information Google has about its users is not shared with the ad exchange.
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. We observe that our algorithm obtains about 90% of the optimal revenue. In addition, it obtains significantly higher revenue, up to more than three times, compared to second-price auctions. Furthermore, we observe that a small number of deals can capture a large fraction of boost in the revenue.

In Section 6, we further extend our analysis to the case where the seller has incomplete information with regards to the distribution of the buyers’ valuations. We design the optimal dynamic mechanism for this setting and show that first-look deals (one-buyer) can obtain the optimal revenue.

**Related Work**

In this section, we briefly discuss three lines of work that are closely related to ours; see Korula et al. (2016) 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 (Balseiro et al. 2014). 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 advertiser 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, Ghosh et al. (2009) propose a randomized bidding mechanism to allocate a representative set of impressions to the buyers. Arnosti et al. (2014) 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.

\(^3\) Channel coordination has been studied in other context of operations management; for example, see Weng (1995), Caldentey and Vulcano (2007), Huang et al. (2014) on online retail.
In this paper, our main focus is 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 Feldman et al. (2009), Turner (2012), Ciocan and Farias (2012), Chen (2013), Agrawal et al. (2014), Shamsi et al. (2014).

**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 state (Ashenfelter and Genesove 1992, Goeree et al. 2004, Burguet 2007), antiques and jewelry (Eliaz et al. 2008), and water rights (Alevy et al. 2010). 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 (Goeree et al. 2004, Burguet 2007). Devanur et al. (2013) 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 (Stigler 1963, Adams and Yellen 1976, McAfee et al. 1989, Bakos and Brynjolfsson 1999, Cachon and Feldman 2011, Gilbert et al. 2014). For sponsored search auctions, Ghosh et al. (2007) show that the problem of finding the optimal bundling is NP-complete and propose approximation algorithms; see also Even-Dar et al. (2007), Emek et al. (2014).

Babaioff et al. (2014) 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 Hart and Nisan (2012). 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. **First-Look Deals**

In this section, we present a special case of our model where the publisher (she) offers a deal only to one buyer (he) and discuss some of our main ideas and assumptions. We will use the terms seller and publisher interchangeably.

Consider a seller of a sequence of impressions, denoted by $\mathcal{I}$. The seller can sell the impressions in an open second-price auction \[4\] In addition, there is a buyer asking for first-look access to the

---

4 The format of the auction is not important for our results. The second-price format is used by Google and many other exchanges.
impressions before they are sent to the auction. Such an arrangement can be made using a preferred deal.

**Preferred Deal:** A preferred deal for the sequence of impressions $\mathcal{I}$ is defined by two parameters, $\rho$ and $\mu$. By accepting the contract, the buyer gains first-look access to the impressions in $\mathcal{I}$ and agrees to buy a fraction $\mu$ of them, each at price $\rho$. The buyer will not be allowed to participate in the auction.\(^5\)

We note that $\mu = 1$ corresponds to the traditional reservation contracts. 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 equivalent.

Let $v$ denote the valuation of the buyer for an impression. The buyer is risk-neutral with quasi-linear utility: the utility from purchasing the impression at price $\rho$ is equal to $v - \rho$.

### 2.1. Minimum Purchase Requirement and Improvement Over Auction

As our first step of analysis, we observe that first-look deals with no minimum purchase, $\mu = 0$, are suboptimal.

**Proposition 1** Consider a first-look deal $(\rho, \mu)$ where $\mu = 0$. The seller can obtain a (weakly) higher revenue per impression by allowing the first-look buyer to participate in the open second-price auction and setting the personal reserve price for the buyer equal to $\rho$.

In the second-price auction, the impression is allocated to the highest bidder among those whose bid is above their personal reserves. Note that the potential first-look buyer will purchase an impression, both in the auction and under the deal, only if his valuation is larger than $\rho$. For those impressions, the revenue in the aforementioned auction is at least equal to $\rho$. The revenue is strictly larger if there is another bidder in the auction with a bid higher than $\rho$.

In the rest of this section, we consider a stochastic setting where $v$ is drawn independently and identically (i.i.d.) from distribution $F$. To simplify the presentation, we assume that the support of $F$ is over $[v, \overline{v}]$ and its p.d.f. $f$ exists and is positive in $[v, \overline{v}]$.

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. In order to purchase $\mu$ fraction of the impressions, the buyer will have to purchase all impressions with $v \geq \theta$ where $\theta = F^{-1}(1 - \mu)$ (i.e., $\mu = 1 - F(\theta)$). Note that under the first-look deal $(\rho, \mu)$, the buyer

\(^5\)Google AdX Documentations (2015): “If the buyer does not bid, or if the bid is below the fixed price, the inventory goes to the general auction and the buyers that are part of a Preferred Deal are excluded from bidding on that impression in the general auction.”
will always purchase an impression if \( v \geq \rho \). Hence, the minimum fraction purchase is binding only if \( 1 - F(\rho) < \mu \). The per impression expected utility of the buyer from contract \((\mu, \rho)\) is equal to \((1 - F(\theta))(E[v|v \geq \theta] - \rho)\) if the minimum purchase is binding; otherwise, it is equal to \((1 - F(\rho))(E[v|v \geq \rho] - \rho)\).

Earlier, we argued that the deals with no minimum purchase are suboptimal. On the flip side, we observe that imposing an optimal minimum purchase can significantly increase the revenue.

**Proposition 2** An optimal first-look deal can obtain arbitrarily larger revenue than the second-price auction.

The proof is given in Appendix A.1. To gain the intuition, let us start with a simple example.

**Example 1** Suppose the distribution of valuations of the buyer, \( F \), is the uniform distribution over \([0, 1]\) and the revenue from the auction (excluding the first-look candidate) is normalized to be equal to 0. In this case, allowing the buyer to participate in the auction can generate revenue at most equal to \(\frac{1}{4} \) using reserve price \(\frac{1}{2}\). However, the first-look deal \((\rho = \frac{1}{3}, \mu = 1)\) will yield revenue equal to \(\frac{1}{2}\). Furthermore, the welfare obtained via the deal, \(\frac{1}{2}\), is larger than \(\frac{3}{8}\), the welfare of the auction with reserve price \(\frac{1}{2}\). Therefore, the deal outperforms both in the terms of revenue and welfare.

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. Another advantage of this approach is that estimating the expected value of an advertiser’s valuation can be significantly easier than estimating the exact value of each impression, cf. Bakos and Brynjolfsson (1999).

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). However, 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 Proposition 2 and later in our empirical analysis (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. Note that in Example 1, any price $\rho \in (\frac{1}{4}, \frac{1}{2})$ increases the seller’s revenue over the auction while leaving a positive surplus for the buyer. Furthermore, we observe that deals can obtain higher welfare compared to auctions with high reserve prices. Therefore, even a seller with limited bargaining power could benefit from deals and contracts.

2.2. Benefits of a Deal for the Buyer

Thus far, we have argued that deals can significantly benefit the seller. Now let us look at deals from a buyer’s perspective. An important advantage of signing a first-look deal for a buyer is the opportunity to purchase high-quality impressions that otherwise could go to another buyer. Compared to bidding in auctions, deals can simplify the process of purchasing impressions. This is the main reasons that the reservation contracts are widely used in practice since they reduce (or eliminate) the chance of under-delivery of a campaign as ad campaigns usually aim to reach a wide audience via a large number of impression.

The following example shows that competition among buyers could incentivize them to sign a first-look deal, which boosts the seller’s revenue.

Example 2 There are two buyers. The valuation of each buyers for each impression is drawn i.i.d. from the uniform distribution over $[0, 1]$. Consider the follow scenarios.

- The seller auctions off the impressions via the second-price auction with the (optimal) reserve $\frac{1}{2}$. In this case, the seller obtains expected per impression revenue equal to $\frac{5}{12}$.
- The seller offers a first-look contract $(\rho, \mu = 1)$, $\rho \in (\frac{5}{12}, 1)$ to only one of the buyers. If the buyer rejects the deal, his utility will be equal to 0. However, he will obtain a positive utility from the deal. The seller’s per impression revenue from the deal is equal to $\rho > \frac{5}{12}$.

3. Preferred Deals

In this section, we consider a general model where the seller offers deals to multiple buyers. Of course, the first-look privilege to the whole inventory can be granted to one buyer only. We show that similar contracts, offered to several buyers, can further increase the revenue of the seller.

More precisely, we study the problem of maximizing the revenue of the seller 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 $(\rho_i, \mu_i, I_i)$. By accepting the deal, the buyer agrees to purchase fraction $\mu_i$ of the impressions in $I_i$ each at price $\rho_i$. Each buyer receives impressions that are not purchased by the buyers with higher priorities. For example, all the impressions are sent

---

6 Reservation contracts often specify a penalty for under-delivery of impressions; see Feige et al. (2008), Kong et al. (2013), Chen et al. (2015), Hojjat et al. (2014).
to $\pi_1$, the buyer with the highest priority. The buyer with the second highest priority receives all the impressions that are not purchased by $\pi_1$.

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.\(^7\)

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.

Due to these complexities, it is not obvious at the first glance how to design even a simple greedy algorithm for this problem. In fact, we show that the problem of finding the optimal sequence of deals is computationally hard.

**Theorem 1 (Computational Hardness)** Suppose the vector of valuations of the buyers for each impression $t \in I$, denoted by $(v_{1,t}, v_{2,t}, \cdots, v_{n,t})$, is known in advance to the seller and the buyers. The problem of finding the optimal (revenue-maximizing) sequence of preferred deals is NP-Hard.

In Appendix A.2 we prove the theorem above using a reduction from a well-known NP-Hard problem, the *maximum acyclic subgraph* (Karp 1972). 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). Guruswami et al. (2008) show that it is even hard to approximate the weighted version of the problem by a factor better than $\frac{1}{2}$. In the next section, we present an approximation algorithm for a setting where the valuations are drawn from a known distribution.

\(^7\) 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.

\(^8\) 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.

\(^9\) 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 of each deal.
4. Auction-Adjusted Greedy Algorithm

In this section, we present a greedy algorithm for finding an approximately optimal sequence of deals. Let $v = (v_1, v_2, \cdots, v_n)$ 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 $F : \mathbb{R}^n \rightarrow \mathbb{R}$. We assume that the valuation of buyer $i$ is distributed over $[v_i, \overline{v}_i]$, $0 \leq v_i < \overline{v}_i$, and the margins are bounded and positive on $[v_i, \overline{v}_i]$. 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, $v_i = \nu_i + \eta_i$, where $\nu_i$ is the private signal of the buyer, which is distributed independently of the other buyers’ valuations; $\eta_i$ is the common value component.

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

$$
\omega(I, S) = \max_{j \in S, \theta_j \in [v_j, \overline{v}_j]} \{\Pr_{I}[v_j \geq \theta_j] \mathbb{E}_I[v_j|v_j \geq \theta_j] + \Pr_{I}[v_j < \theta_j] \omega(I \setminus \{v_j \geq \theta\}, S \setminus \{j\})\}. (1)
$$

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

In Figure 1 we present an algorithm called Auction-Adjusted Greedy (AAG). The algorithm repeatedly (recursively) chooses one of the buyers, denoted by $i$, and offers him a contract $(\mu_i, \rho_i, I_i)$. The choice for $\theta$ is guided by the auction. Consider the first step. $\theta_j$ 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

$$
\mathbb{E}[v_i|v_{(1)}(S) = v_i] \leq \mathbb{E}[v_i|v_i \geq \theta_i],
$$

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, $\mathbb{E}_I[v_{(1)}(S \setminus \{i\})]$. In other words, we use the expected value of the highest $v_i$ as the proxy for $\omega(I, S)$. 

Figure 1  

**Auction-Adjusted Greedy (AAG) Algorithm**

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

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

For each buyer $j \in S$
  Choose $\theta_j$ such that: $\Pr_I[v_j \geq \theta_j] = \Pr_I[v_i = v_{(1)}(S)]$.

Choose buyer $i \in \arg \max_j \{E_I[v_j | v_j \geq \theta_j] / E_I[v_{(1)}(S \setminus \{j\}) | v_j \geq \theta_j]\}$. (3)

$\mu_i \leftarrow \Pr_I[v_i \geq \theta_i]$.  
$\rho_i \leftarrow E_I[v_i | v_i \geq \theta_i]$.  
$D \leftarrow (\rho_i, \mu_i, I) + AAG(I \setminus \{v_i \geq \theta_i\}, S \setminus \{i\})$.  
Return $D$

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 $E[\max_i \{v_i\}]$, which is the trivial upper bound on the revenue as the benchmark.

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

In Appendix A.3 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, $E[v_{(1)}(S \setminus \{j\}) | v_j \geq \theta_j] = E[v_{(1)}(S \setminus \{j\})]$, is less than or equal to the revenue that can be obtained from that buyer via a deal, $E[v_j | v_j \geq \theta_j]$. Because the algorithm chooses a buyer $i$ that maximizes the ratio of the revenue to cost, for chosen buyer $i$ we have $E_I[v_i | v_i \geq \theta_i] \leq E_I[v_i | v_i \geq \theta_i]$. Similarly, for correlated valuations, we show that there exists a buyer such that $E[v_{(1)}(S \setminus \{j\}) | v_j \geq \theta_j] \leq 2 E[v_j | v_j \geq \theta_j]$. 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, $E[v_{(1)}(S \setminus \{j\})]$, are the same for
both sets. For the correlated valuations, the expected value is higher for the impressions in the second set (i.e., \(E[v_{j1}(S \setminus \{j\}) | v_j > v_{j1}(S \setminus \{j\})] \leq E[v_{j1}(S \setminus \{j\}) | v_j > \theta_j]\) because the valuations are positively correlated), hence the difference in the approximation ratio.

In the next section, we show that the algorithm obtains higher revenue than the worst-case bounds provided by the theorem.

5. Empirical Evaluation

We evaluate our algorithm by simulating it over logs of auction data. The data was collected from Google’s Doubleclick ad exchange platform one day in the Winter of 2014. On that day, we found the top 4 ad units that generated the highest revenue. For 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. Given a subset of impressions \(I\), we use the fraction of auctions where \(j\) has the highest bid as \(\Pr_{I}[v_j = v_{j1}]\). Similarly, for \(E_{I}[v_j | v_j \geq \theta_j]\), we use the average of \(v_j\) calculated over the impressions in \(I\) where the bid of advertiser \(j\) is larger than or equal to \(\theta_j\).

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,

\[
j \leftarrow \arg \max_{j \in S} \left\{ \max_{\theta} \left\{ \Pr[v_j \geq \theta] \times E[v_j - v_{j1}(S \setminus \{j\}) | v_j \geq \theta] \right\} \right\}.
\]

See Celis et al. (2014) for a discussion on the pros and cons of this assumption.
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. 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.

In Table 2, we present more details with regards to the solution of the different mechanisms on Dataset 1. Advertisers are labeled with the ordering that they are chosen by the auction-adjusted greedy algorithm. For each mechanisms, in the table we present the revenue obtained from each advertisers as a percentage of the benchmark as well as the percentage of the impressions allocated. For the preferred deal mechanisms, we also present the price $\rho$. The prices are normalized such that the highest $\rho$ is equal to 1.

We note that the contribution of the buyers to both greedy algorithms is somewhat similar. The main reason that the max-margin greedy obtains a lower revenue than the auction-adjusted greedy is that it allocates a smaller number of impressions to advertisers with higher valuations. Although the price-per-impression could be higher, the total revenue goes down.

Finally, observe that a large fraction of the revenue is generated by a small number of deals. This is an important insight since the process of making a deal may require effort and negotiations. Our algorithm can serve as a baseline that identifies potential profitable deals and their parameters.

\[11^{\text{th}}\] The corresponding ad unit generated the highest revenue in the ad exchange on the day of data collection and contains about two million impressions.
In this section, we consider a model where buyers may have more information about the distribution of the valuation than the seller. We will focus on first-look deals, the case when there is one buyer and the remaining impressions are sold in an auction. We show that an optimal first-look deal can maximize the revenue among all selling mechanisms. At the end of this section, we discuss the multiple buyers case.

Modeling asymmetry of information between the seller and the buyer is delicate. The private type of the buyer can be the information about the distribution of his valuations. To avoid the well-known difficulties of multi-dimensional mechanism design, we consider a single-dimensional setting where the set of possible distributions of the valuations is parameterized by a scalar. This allows us to identify an optimal mechanism.

More specifically, we assume that the valuation of the buyer for each impression is distributed according to the following: with probability $1 - \alpha$, the valuation is equal to 0 and with probability $\alpha$, the valuation is drawn from a known distribution $F$. Distribution $F$ is defined over $[\underline{v}, \overline{v}]$ with positive margins.

We think of $\alpha$ as the match probability. In other words, $\alpha$ represents the probability that the user matches the targeting criteria of the buyer. For instance, an advertiser may want to target users who have previously visited the advertiser’s website. The valuation of the advertiser for those users may change depending on other factors, including the users’ behaviors and actions on the website (e.g., the valuation would be higher if the user has an abandoned item in her shopping cart). $\alpha$ is the private information of the buyer, but its distribution, denoted by $G$, is known to the seller. Distribution $G$ is defined over $[\underline{\alpha}, \overline{\alpha}]$ with p.d.f. $g$. As in the previous sections, the realizations of the valuations of the buyer are only known to the buyer.

### Table 2

The contribution of each advertiser to the revenue as a percentage of the benchmark, the percentage of the impressions allocated, and the deal’s price for each algorithm.

<table>
<thead>
<tr>
<th>Advertisers</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Revenue</strong></td>
<td>7.57%</td>
<td>16.90%</td>
<td>28.32%</td>
<td>19.84%</td>
<td>9.88%</td>
<td>2.98%</td>
<td>0.02%</td>
<td>0.01%</td>
<td>0.51%</td>
<td>0.57%</td>
</tr>
<tr>
<td><strong># impressions</strong></td>
<td>1.46%</td>
<td>4.14%</td>
<td>21.10%</td>
<td>20.15%</td>
<td>17.36%</td>
<td>16.33%</td>
<td>0.15%</td>
<td>0.09%</td>
<td>3.12%</td>
<td>7.13%</td>
</tr>
<tr>
<td><strong>Deal Price</strong></td>
<td>1.00</td>
<td>0.78</td>
<td>0.25</td>
<td>0.19</td>
<td>0.11</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Max-Margin Greedy Algorithm</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Revenue</strong></td>
</tr>
<tr>
<td><strong># impressions</strong></td>
</tr>
<tr>
<td><strong>Deal Price</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Second-Price Auction</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Revenue</strong></td>
</tr>
<tr>
<td><strong># impressions</strong></td>
</tr>
</tbody>
</table>
Let us start with the optimal first-look deal with full distributional information. Suppose \( \alpha \) is known and we normalized it to be equal to 1. If the buyer does not purchase an impression, the impression is sent to an auction. Let random variable \( c(v), c : \mathbb{R} \to \mathbb{R} \), denote the expected value of the revenue from the auction. Note that the value of the impression for the buyer can be correlated with the bids from the auction. The expected net profit of the optimal first-look deal is equal to

\[
\max_{\theta} \Pr[v \geq \theta] E[v - c(v)|v \geq \theta].
\]

(4)

We now consider the case where \( \alpha \) is drawn from distribution \( G \). We define the virtual value function \( \psi \) as follows.

\[
\psi(\alpha, v) = v - c(v) - \frac{1 - G(\alpha)}{f(v)} \int_{v}^{r(\alpha)} (1 - F(\theta))d\theta.
\]

(5)

**Assumption 1** Function \( \psi \) is differentiable and increasing in \( \alpha \in [\alpha, \bar{\alpha}] \) and \( v \in [\underline{v}, \overline{v}] \).

The monotonicity in \( \alpha \) holds if \( G \) satisfies the standard monotone hazard-rate property. Similarly, \( \psi \) is monotone in \( v \) if \( F \) has a monotone hazard-rate and \( c'(v) \leq 1 \). Note that if the bids in the auction are independent of the first-look buyer’s valuations, then \( c'(v) = 0 \). In addition, for the correlated valuation model described in Section 4, the mixture of common and private values, we have \( c'(v) \leq 1 \).

Now let \( r_\alpha \) be the solution of \( \psi(\alpha, v) = 0 \). If \( \psi(\alpha, v) > 0 \) for all \( v \in [\underline{v}, \overline{v}] \), let \( r_\alpha = v \). If \( \psi(\alpha, \overline{v}) < 0 \), let \( r_\alpha = \overline{v} \).

**Theorem 3** If Assumption 1 holds, then the following mechanism is optimal.\(^{12}\) The seller offers the buyer a menu of first-look deals \( (\rho_\alpha = p(\alpha)/(\alpha(1 - F(r_\alpha))), \mu_\alpha = \alpha(1 - F(r_\alpha))) \), \( \alpha \in [\alpha, \bar{\alpha}] \), where

\[
p(\alpha) = \alpha \int_{\underline{v}}^{\overline{v}} v f(v)dv - \int_{\underline{v}}^{\overline{v}} \int_{r_\alpha}^{v} (1 - F(\theta))d\theta dz.
\]

(6)

The parameters of the contracts are chosen such that the buyer of type \( \alpha \) will choose first-look deal \( (\rho_\alpha, \mu_\alpha) \) and will purchase all impressions with valuation above \( r_\alpha \).

To gain some intuition, suppose \( F \) and \( G \) have monotone hazard rates. The optimal posted-price for the buyer is then equal to the solution of \( r - \frac{1 - F(r)}{f(r)} = c(r) \). For large \( \alpha \) such that \( \frac{1 - G(\alpha)}{\alpha f(\alpha)} < 1 \), we have \( r_\alpha < r \). Therefore, a buyer with higher match probability will purchase a larger number of impressions compared to the optimal posted-price setting. Note that for \( \alpha = \bar{\alpha} \), all the impression where \( v \geq c(v) \) are allocated to the first-look buyer. On the other hand, the buyer will purchase a smaller number of impressions for small values of \( \alpha \) such that \( \frac{1 - G(\alpha)}{\alpha f(\alpha)} > 1 \) because \( r_\alpha > r \).

\(^{12}\) An optimal mechanism maximizes the revenue (more precisely the net profit) among all mechanisms subject to the participation constraint; see Appendix A.4 for more details.
We prove the theorem above in Appendix A.4. We first show that the following direct mechanism is optimal: at time 0, the buyer reports $\alpha$ to the mechanism. The buyer of type $\alpha$ makes an initial payment equal to $T \times (p(\alpha) - \alpha(1 - F(r_\alpha))r_\alpha)$. Afterwards, the seller sends each impression to the buyer at a price equal to $r_\alpha$. Observe that at price $r_\alpha$, the buyer will purchase $\mu_\alpha = \alpha(1 - F(r_\alpha))$ fraction of the impressions. Note that the total payment under the direct mechanism and the optimal deals is equal to $T \times p(\alpha)$.

When there are multiple buyers, the optimal direct mechanism would consist of initial payments followed by a sequence of auctions where the allocation rule of the auction favors buyers with higher initial payments. 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. For multiple buyers, the preferred deals may not be optimal, but we believe that they can obtain near-optimal solutions; see the discussion in the next section.

7. 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 theoretical results show that preferred deals can obtain a constant fraction of the optimal revenue when the distributions of the valuations are known to the seller. We also showed that under certain assumptions, preferred deals can implement the optimal mechanism for a setting with one buyer where the buyer has more information than the seller. We conjecture that preferred deals can obtain near-optimal revenue for multiple buyers and more general settings. We leave this as an important direction for future research. Other directions include analyzing settings where buyers are budget-constrained (cf. [13]) or arrive over time (cf. [15], [20], [21]).

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.

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.

---

13 See [16], [17], [18], [19], [20], [21] for results in similar settings.
Appendix A: Proofs

A.1. Proof of Proposition 2

We prove the claim using the following example. Suppose the revenue from the second-price auction, excluding the first-look candidate, is always equal to 1. Consider (truncated) revenue-equal distribution \( F(v) = \frac{M}{M - 1} \left( 1 - \frac{1}{v} \right), \ v \in [1, M] \). Note that the revenue of any posted price \( p \) is at most equal to 1.

\[
p(1 - F(p)) = p \left( 1 - \frac{M}{M - 1} \left( 1 - \frac{1}{p} \right) \right) = \frac{M}{M - 1} \left( 1 - \frac{p}{M} \right) \leq \frac{M}{M - 1} \left( 1 - \frac{1}{M} \right) = 1.
\]

On the other hand, the revenue of the first-look deal \((\rho = E[v], \mu = 1)\) is \( O(\log M) \).

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

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

A.2. Proof of Theorem 1

Consider a 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 \( n\bar{d} + m \) items where \( \bar{d} \) denotes the maximum degree in the graph. For each edge \( e = (i, j) \in E(G) \), we add an item \( e \) where the valuation of buyer \( i \) for item \( e \) is equal to 3, and the valuation of buyer \( j \) for item \( e \) is equal to 2. Other buyers’ valuations for item \( e \) are set equal to 0. Finally, we add \( \bar{d} \) “exclusive” items for each buyer \( i \) with value 1. 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 revenue of \( (2m + k) + n\bar{d} \) 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 the remaining items, including the \( \bar{d} \) exclusive items. 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 \( 3k + 2(m - k) = k + 2m \). The total revenue from exclusive items is equal to \( n\bar{d} \), which completes the proof. \( \square \)

We now prove the other direction.

Lemma 2 If there exists a sequence of preferred deals of total revenue \( (2m + k) + n\bar{d} \) for the problem instance described above, then, there exists an ordering with \( k \) forward edges for the corresponding maximum acyclic subgraph problem.
Proof: We first claim that if there exists a sequence of preferred deals of total revenue \((2m + k) + n\bar{d}\) for the problem instance described above, then there exists a sequence of deals with at least the same revenue such that each buyer purchases all the items “available” to her (that are not purchased by the buyers with higher priorities).

To prove the claim, we show that the revenue 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:

- Suppose the number of exclusive items that are allocated to buyer \(i\) is greater than 0 and less than \(\bar{d}\). In this case, we can simply increase the revenue by allocating all the exclusive items and setting \(\rho\) equal to the total value of all items allocated to her divided by the number of the items.

- Suppose no exclusive items are allocated to buyer \(i\). This means that \(\rho\) is large. Suppose we reduce \(\rho\) such that she purchases all the items available to her. 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 total loss from those items is at most equal to \(d_i\), where \(d_i\) denotes the degree of note \(i\). Because \(\bar{d} \geq d_i\), the revenue can be (weakly) increased.

Therefore, if there exists a sequence of preferred deals of total revenue \((2m + k) + n\bar{d}\) 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 \(n\bar{d}\) plus 3 from the forward edges and 2 from the backward edges. Therefore, revenue \((2m + k) + n\bar{d}\) implies that there are \(k\) items 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.

\[\square\]

A.3. 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.

\[
E[v_{(1)}(S)] = E[v_{(1)} 1(v_{(1)} = v_i)] + E[v_{(1)} 1(v_{(1)} \neq v_i, v_i \geq \theta_i)] + E[v_{(1)} 1(v_{(1)} \neq v_i, v_i < \theta_i)]
\leq E[v_{(1)} 1(v_{(1)} = v_i)] + E[v_{(1)}(S \setminus \{i\}) 1(v_i \geq \theta_i)] + E[v_{(1)}(S \setminus \{i\}) 1(v_i < \theta_i)]
\leq \Pr[v_i \geq \theta_i] (E[v_i|v_i \geq \theta_i] + E[v_{(1)}(S \setminus \{i\})|v_i \geq \theta_i]) + \Pr[v_i < \theta_i] E[v_{(1)}(S \setminus \{i\})|v_i \leq \theta_i].
\tag{7}
\]

The last inequality follows from Eq. (2). The algorithm obtains expected per-impression revenue of \(\Pr[v_i \geq \theta_i]E[v_i|v_i \geq \theta_i]\). Therefore, to prove the theorem, using the induction hypothesis, it suffices to show that \(E[v_i|v_i \geq \theta_i] \geq E[v_{(1)}(S \setminus \{i\})]\) for independent valuations and \(2 \times E[v_i|v_i \geq \theta_i] \geq E[v_{(1)}(S \setminus \{i\})|v_i \geq \theta_i]\) for correlated valuations. We prove these inequalities respectively in Lemmas 4 and 5.

Recall that for any buyer \(j\), we have \(v_j = \eta + \nu_j\), where \(\eta\) is the common value component and \(\nu_j\) is distributed independently among the buyers. Let \(f^\eta\) and \(f_j^\nu\) respectively denote the p.d.f. of distributions of \(\eta\) and \(\nu_j\).
For any buyer \( j \), we have
\[
E[v_j - v(1)(S \setminus \{ j \})|v_j \geq \theta_j] = \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} (1 + \nu) E[v - v(1)(S \setminus \{ j \})|v_j = \eta + \nu] f^*(\eta)f^*_j(\nu)d\eta d\nu
\]
\[
= \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} (1 + \nu) E[v_j - \max_{\nu_k \in S \setminus \{ j \}} \nu_k|v_j = \nu] f^*(\eta)f^*_j(\nu)d\eta d\nu
\]
\[
= \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} (1 + \nu) (\nu - \max_{\nu_k \in S \setminus \{ j \}} \nu_k) f^*(\eta)f^*_j(\nu)d\eta d\nu
\]
\[
= \Pr[v_j \geq \theta_j] E[v_j|v_j \geq \theta_j] - \Pr[v_j \geq \theta_j] E\left[\max_{\nu_k \in S \setminus \{ j \}} \nu_k\right]
\]
(8)

The last expressions are obtained because \( \nu_i \) is distributed independently of \( \nu_j \). On the other hand, we have:

**Lemma 3**
\[
E[v(1)] - E[\ell] \geq \sum_{j=1}^{n} \Pr[v_j = v(1)] E\left[\max_{\nu_k \neq j} \nu_k\right]
\]

**Proof:** Recall that \( v(1) \) is the random variable corresponding to the highest valuation. We now define random variable \( \tilde{\nu} \) as follows. Let \( \nu_j = \eta + \nu_i \) be the highest valuation. For all \( k \neq \ell \), re-sample \( \nu_k \) and let \( \tilde{\nu_k} \) denote the new realization. Let \( \tilde{v} = \eta + \max_{k \neq \ell} \{\tilde{\nu_k}\} \). We observe that \( E[v(1)] \geq E[\tilde{v}] \). The reason is that \( \tilde{v} \) represents the highest expected value among a smaller number of buyers. Therefore, we obtain the following.
\[
0 \leq E[v(1)] - E[\tilde{v}]
\]
\[
= \sum_{j=1}^{n} \left( \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} \Pr[v_j \geq v(1)(S \setminus \{ j \})|v_j = \eta + \nu] E[v - v(1)(S \setminus \{ j \})|v_j = \eta + \nu] f^*(\eta)f^*_j(\nu)d\eta d\nu \right)
\]
\[
= \sum_{j=1}^{n} \left( \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} \Pr[\nu \geq \max_{k \neq \ell} \nu_k] \left[ v_j - \max_{k \neq j} \nu_k \right] f^*(\eta)f^*_j(\nu)d\eta d\nu \right)
\]
\[
= \sum_{j=1}^{n} \left( \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} \Pr[\nu \geq \max_{k \neq \ell} \nu_k] \left( v_j - \max_{k \neq j} \nu_k \right) f^*(\eta)f^*_j(\nu)d\eta d\nu \right)
\]
\[
= \sum_{j=1}^{n} \left( \int_{\eta_j}^{\nu_j} \int_{\nu_j}^{\nu_j + \nu \geq \theta_j} \nu f^*(\eta)f^*_j(\nu)d\eta d\nu - \Pr[v_j \geq \theta_j] E\left[\max_{k \neq j} \nu_k\right] \right)
\]
\[
= \sum_{j=1}^{n} \left( \Pr[v_j = v(1)] \left( E[v_j|v_j = v(1)] - E[\ell] \right) - \Pr[v_j = v(1)] E\left[\max_{k \neq j} \nu_k\right] \right)
\]
\[
= E[v(1)] - E[\ell] - \sum_{j=1}^{n} \Pr[v_j = v(1)] E\left[\max_{k \neq j} \nu_k\right]
\]

We are now ready to prove the induction step.

**Lemma 4** Let \( i \) be the chosen buyer. For independent valuations (i.e., \( \nu_j = v_j \)), we have
\[
E[v_i|v_i \geq \theta_i] \geq E\left[v(1)(S \setminus \{ i \})\right].
\]
Proof: Summing up Eq. (8), we have
\[
\sum_{j=1}^{n} E [v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j] \geq \sum_{j=1}^{n} \Pr[v_j \geq \theta_j] \left( E[v_j|v_j \geq \theta_j] - E \left[ \max_{\theta \neq j} \{v_i\} \right] \right)
\]
\[
= \sum_{j=1}^{n} \Pr[v_j = v_{(1)}] \left( E[v_j|v_j \geq \theta_j] - E \left[ \max_{\theta \neq j} v_i \right] \right)
\]
\[
\geq \sum_{j=1}^{n} \Pr[v_j = v_{(1)}] \left( E[v_j|v_j = v_{(1)}] - E \left[ \max_{\theta \neq j} v_i \right] \right)
\]
\[
= E[v_{(1)}] - \sum_{j=1}^{n} \Pr[v_j = v_{(1)}] E \left[ \max_{\theta \neq j} \{v_i\} \right]
\]
\[
\geq 0
\]
Eq. (9) follows from definition of \( \theta_j \) and (10) follows from (2). We obtain the last inequality using Lemma 3. Note that because the sum of \( E \left[ v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j \right] \) is non-negative, there exists buyer \( k \) for whom \( E \left[ v_k - v_{(1)}(S \setminus \{k\})|v_k \geq \theta_k \right] \geq 0 \). For that buyer \( E[v_k|v_k \geq \theta_k]/E \left[ v_{(1)}(S \setminus \{k\}) \right] \geq 1 \). By definition, (3), for buyer \( i \) we have \( E[v_i|v_i \geq \theta_i]/E \left[ v_{(1)}(S \setminus \{i\}) \right] \geq 1 \). □

Lemma 5 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. (8), we have
\[
\sum_{j=1}^{n} E[v_j|v_j \geq \theta_j] + \sum_{j=1}^{n} E[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j]
\]
\[
\geq E[v_{(1)}] + \sum_{j=1}^{n} E[v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j]
\]
\[
= E[v_{(1)}] + \sum_{j=1}^{n} \Pr[v_j \geq \theta_j] \left( E[v_j|v_j \geq \theta_j] - E \left[ \max_{\theta \neq j} \{v_i\} \right] \right)
\]
\[
\geq \sum_{j=1}^{n} \Pr[v_j \geq \theta_j] E[v_j|v_j \geq \theta_j] + E[\eta]
\]
\[
\geq 0
\]
Using (2) and (8) respectively, we obtain (11) and (12). Inequality (13) follows from Lemma 3 and the last inequality holds simply because all the terms are non-negative. Note that because the sum of \( E \left[ 2v_j - v_{(1)}(S \setminus \{j\})|v_j \geq \theta_j \right] \) is non-negative, there exists buyer \( k \) for whom \( E \left[ 2v_k - v_{(1)}(S \setminus \{k\})|v_k \geq \theta_k \right] \geq 0 \). For that buyer, we have \( 2E[v_k|v_k \geq \theta_k] \geq E \left[ v_{(1)}(S \setminus \{k\})|v_k \geq \theta_k \right] \). By definition, (3), for buyer \( i \) we have \( 2E[v_i|v_i \geq \theta_i] \geq E \left[ v_{(1)}(S \setminus \{i\})|v_i \geq \theta_i \right] \). □

A.4. Proof of Theorem 3

Let us start with the outline of the proof. First, we formally present the revenue optimization problem for the seller, (14). We then provide upper bounds on the seller revenue, respectively, via (15) and (18). In Proposition 3 we present an incentive-compatible mechanism that obtains a revenue equal to the upper bounds and is therefore optimal. Finally, we discuss how first-look deals can implement the optimal mechanism.
Using the revelation principle (Myerson 1986), without loss of generality, we focus on direct mechanisms. Let \( h_t \) be the history of the game up to time \( t \). Namely, \( h_0 = \emptyset, h_1 = \langle \hat{\alpha} \rangle, \ldots, h_t = \langle \hat{\alpha}, \hat{v}_1, \ldots, \hat{v}_{t-1} \rangle \), where \( \hat{\alpha} \) and \( \hat{v}_i \) respectively denote the report of the buyer for \( \alpha \) and \( v_t \). Let \( H \) denote the space of all histories. A direct mechanism can be defined using functions \( \hat{q} : H \times R \rightarrow R^+ \) and \( \hat{p} : H \times R \rightarrow R \), which respectively map the history and the report of the buyer to the allocation and payment. Using the one-stage deviation principle (Blackwell 1965; Fudenberg and Tirole 1991), we can write the seller’s profit optimization problem as follows.

\[
\max_{\alpha, H \rightarrow R^+, \hat{p} : H \rightarrow R} \int_{\alpha}^{\pi} E \left[ \sum_{t=1}^{T} \left( \int_{\alpha}^{\pi} (\hat{p}(h_t, x) - \hat{q}(h_t, x)c(x))\alpha f(x)dx \right) \right] g(\alpha)d\alpha
\]

subject to:

\[
\hat{q}(h_t, v)v - \hat{p}(h_t, v) \geq \hat{q}(h_t, v')v - \hat{p}(h_t, v') \quad \forall v, v' \in [\alpha, \pi], h_t \in H
\]

\[
E \left[ \sum_{t=1}^{T} \int_{\alpha}^{\pi} \left( \hat{q}(h_{t,\alpha}, x) - \hat{p}(h_{t,\alpha}, x) \right)f(x)dx \right] \geq 0, \quad \forall \alpha, \alpha' \in [\alpha, \pi]
\]

where \( h_{t,\alpha} = \langle \alpha, \{v_r \}_{r=1}^{t-1} \rangle \) corresponds to the truthful history and \( h_{t,\alpha'} = \langle \alpha', \{v_r \}_{r=1}^{t-1} \rangle \) denotes the history where the buyer misreports \( \alpha \) but is truthful afterwards. The objective corresponds to the expected profit of the seller, the payment from the buyers minus the opportunity cost of selling an impression in the auction. The expectations are taken with respect to the evolution of history and realizations of the valuations of the buyer. The first set of constraints corresponds to the incentive compatibility constraints at each time \( t \). The second set of constraints represents the incentive compatibility for reporting \( \alpha \). The last set is the individual rationality constraints.

**Lemma 6** The optimal solution of the optimization problem below provides an upper bound on the profit of the seller, the solution of (14).

\[
\max_{q : R^2 \rightarrow R^+, p : R^2 \rightarrow R} \int_{\alpha}^{\pi} \left( \int_{\alpha}^{\pi} (p(\alpha, x) - q(\alpha, x)c(x))\alpha f(x)dx \right) g(\alpha)d\alpha
\]

subject to:

\[
q(\alpha, v)v - p(\alpha, v) \geq q(\alpha, v')v - p(\alpha, v') \quad \forall v, v' \in [\alpha, \pi], \alpha \in [\alpha, \pi]
\]

\[
\int_{\alpha}^{\pi} (q(\alpha, x) - p(\alpha, x))f(x)dx \geq \int_{\alpha}^{\pi} (q(\alpha', x) - p(\alpha', x))f(x)dx \quad \forall \alpha, \alpha' \in [\alpha, \pi]
\]

\[
\int_{\alpha}^{\pi} (q(\alpha, x) - p(\alpha, x))f(x)dx \geq 0 \quad \forall \alpha \in [\alpha, \pi]
\]

**Proof:** Note that the first and second IC constraints in the optimization problem (15) are relaxations of the corresponding constraints in (14) because instead of the whole history, they are specified respectively in terms of only the reports of \( v_t \) and \( \alpha \). The rest follows from the linearity of expectation because the valuations are i.i.d. over time. \( \square \)

Observe that (15) represents the optimal mechanism for a two-stage mechanism where in the first round the buyer reports \( \alpha \), and in the second round the buyer learns her valuation and then reports it to the seller,
Lemma 7 If \(q \) and \(p \) are feasible solutions of (16), then \( \frac{du(u)}{d\alpha} = \int_0^\pi (1 - F(v))q(\alpha, v)dv \).

Proof: Note that we can rewrite the second constraint in (16) as follows:

\[
U(\alpha) - U(\alpha') \geq \int_0^\pi (\alpha - \alpha')f(v)u(\alpha', v)dv.
\]

Similarly,

\[
U(\alpha) - U(\alpha') \leq \int_0^\pi (\alpha - \alpha')f(v)u(\alpha, v)dv.
\]

W.l.o.g., assume \(\alpha \geq \alpha' \). Dividing both sides by \(\alpha - \alpha'\), as \(\alpha'\) converges to \(\alpha\), we obtain

\[
\frac{dU}{d\alpha} = \int_0^\pi f(v)u(\alpha, v)dv.
\]

Integrating by parts, we have

\[
\frac{dU}{d\alpha} = (1 - F(\pi))u(\alpha, \pi) - (1 - F(\pi))u(\alpha, \pi) - \int_0^\pi (1 - F(v))\frac{\partial u(\alpha, v)}{\partial v}dv = \int_0^\pi (1 - F(v))q(\alpha, v)dv.
\]

The last expression is obtained by substituting \(\frac{\partial u(\alpha, v)}{\partial v}\) with \(q(\alpha, v)\). This follows from standard arguments using the second period incentive compatibility in (16).

By the lemma above, setting \(U(\alpha) = 0\), we have,

\[
U(\alpha) = \int_0^\pi \int_0^\pi (1 - F(v))q(\alpha, v)dvd\alpha.
\]

With algebraic manipulation we get

\[
\int_0^\pi U(\alpha)g(\alpha)d\alpha = \int_0^\pi \left( \int_0^\pi q(z, v)(1 - F(v))dvdz \right) g(\alpha)d\alpha = \int_0^\pi \int_0^\pi q(\alpha, v)(1 - G(\alpha))(1 - F(v))dvd\alpha.
\]

Plugging into the objective of (16), we have

\[
\int_0^\pi \left( \int_0^\pi q(\alpha, v)(v - c(v))af(v)dv - U(\alpha) \right) g(\alpha)d\alpha = \int_0^\pi \int_0^\pi q(\alpha, v) \left( v - c(v) - \frac{1 - G(\alpha)}{af(\alpha)} \frac{1 - F(v)}{f(v)} \right) af(v)dvdg(\alpha)d\alpha.
\]

Therefore, the expected profit of the seller is at most equal to

\[
\max_{q:R^2 \rightarrow R^+} T \times \int_0^\pi \int_0^\pi q(\alpha, v) \left( v - c(v) - \frac{1 - G(\alpha)}{af(\alpha)} \frac{1 - F(v)}{f(v)} \right) af(v)dvdg(\alpha)d\alpha.
\]
The objective of (18) can be maximized by allocating the item if and only if
\[ v - c(v) - \frac{1 - G(\alpha) 1 - F(v)}{\alpha g(\alpha) f(v)} \geq 0. \]
Therefore, the optimal solution of (18) is equal to
\[ T \times \int_{\alpha}^{\infty} \int_{r_{\alpha}}^{\infty} \left( v - c(v) - \frac{1 - G(\alpha) 1 - F(v)}{\alpha g(\alpha) f(v)} \right) \alpha f(v) \alpha g(\alpha) d\alpha. \] 

\( (19) \)

We now present an optimal direct mechanism.

**Proposition 3** Suppose Assumption 1 holds. The following mechanism is incentive compatible and obtains profit equal to the optimal solution of (19): The buyer reports \( \alpha \) to the seller. Let \( \hat{\alpha} \) denote the report. The buyer pays \( T \times \left( p(\hat{\alpha}) - \hat{\alpha}(1 - F(r_{\hat{\alpha}}))r_{\hat{\alpha}} \right) \), where \( p(\cdot) \) is defined in (6). After that, for each impression, the seller posts a reserve price of \( r_{\hat{\alpha}} \).

**Proof:** Note that the buyer will purchase an impression every time his valuation is equal to or above \( r_{\hat{\alpha}} \). Hence, \( \frac{\partial u(\alpha, v)}{\partial \alpha} = q(\alpha, v) \), as implied by the incentive compatibility constraints on \( v \). Incentive compatibility on \( \alpha \) follows from Kakade et al. (2013), Lemma 3.6.\(^{14}\) The interval dominance property is satisfied because Assumption 1 implies that \( r_{\alpha} \) is decreasing in \( \alpha \). By (17), the revenue is equal to the optimal solution of (19). \( \square \)

To complete the proof, note that the revenue of the first-look contract described in Theorem 3 is equal to the revenue of the optimal direct mechanism above.

**References**


\(^{14}\)Kakade et al. (2013) present an optimal dynamic mechanism for a class of separable environments. We note that the model we consider here is not separable; but Lemma 3.6, which provides necessary and sufficient conditions for incentive compatibility, applies to more general and non-separable settings, including ours.


