Abstract

We study information revelation policies for ad exchanges that use a second-price auction and the Two-Call design, which corresponds to the Ad Served on the Win Notice OpenRTB specification in the ad industry. Here, the exchange makes a call to all bidders at the beginning of an auction, and makes a second call to the winning bidder at the end of the auction; this second call enables the winning bidder to match the right advertiser for the impression.

While valuations are private to bidders, there are two possibilities as far as the information available to the ad exchange is concerned: When the exchange has no reliable knowledge about bidder valuations, we develop simple information revelation policies that do not use any knowledge of the valuations and establish their performance guarantees. When the exchange has distributional knowledge about bidder valuations, we develop an effective informed heuristic that exploits this information.

Keywords: Ad Exchanges, Information Revelation Designs, Heuristics and Performance Guarantees

Introduction

In the last decade or so, the use of the internet to deliver promotional material to prospective customers has attracted a considerable amount of attention (Central Market Research 2012, Lieberman 2013). Internet advertising revenue in the United States (U.S.) totaled $36.6 billion in 2012, up 15% from that in 2011 (Interactive Advertising Bureau 2012a). This trend is expected to continue: eMarketer (2012) estimates internet ad revenue to exceed $50 billion by 2014, whereas Hof (2011) estimates that internet advertising will reach $76 billion in 2016, and will comprise about 35% of all advertising, overtaking television advertising. Finally, internet advertising is not purely a U.S. phenomenon. For instance, in the United Kingdom, internet advertising has grown 14.4% annually since 2007 (Sweney 2012).

Internet advertising provides advertisers the ability to precisely measure the success of an ad campaign (The Economist 2006). When an ad is shown to a web visitor on a particular website (or the ad publisher’s site), it is possible to track the visitor’s click behavior or in some cases, even whether the click converts to a desired action, such as a sale or a signup. In many cases, however, a click is usually a good
signal of a visitor’s interest in the ad material. It is therefore not surprising to find that many ad revenue models used on the internet are based on the cost-per-click model. In this model, the advertiser pays only for a click, and not for an impression (the event where a visitor is shown an ad). In 2012, 66% of the total online advertising spending in the U.S. was based on performance-based pricing models (such as the cost-per-click model), whereas the impression-based model accounted for only 32% (Interactive Advertising Bureau, 2012a).

Given the success of internet advertising, publishers strive to find better and more effective ways to monetize the web traffic to their websites. One revenue model is to share the click revenue with the agent of the advertiser (such as a demand-side platform, an ad network, etc). Sharing click revenue provides higher returns to the publisher but also entails higher risk (if there is no click, then there is no revenue). A lower revenue (but less risky) option is for the publisher to sell the impression to another party that acquires the right to show an ad for that impression. Such a revenue model (known as cost-per-impression pricing) is becoming an increasingly attractive option for publishers with the advent of the so-called ad exchange. An ad exchange is a supply–demand matching platform where an impression is sold to the highest bidder, often using a second-price auction. In an ad exchange, impressions are auctioned off one by one in real time. Yahoo!’s RightMedia (http://www.rightmedia.com) and Google’s AdX (http://www.google.com/doubleclick/) are examples of prominent ad exchanges. Other lesser known ones are AppNexus (http://www.appnexus.com/) and OpenX (http://www.openx.com/).

Ad Exchange: A Marketplace for Online Advertising

This paper studies internet advertising from the perspective of an ad exchange. Similar to keyword auctions (but arguably, a layer deeper in the marketing funnel), an ad exchange allows advertisers (or their agents) to use the information contained in an impression to infer the intent of the web visitor and match it with an appropriate ad. In a keyword auction, a search engine monetizes search by auctioning keywords (found in search strings) to advertisers. In a similar way, the event of a user’s visit to a website (namely, an impression) is auctioned. The ad exchange is an attractive place for publishers to sell their inventory because it gives them access to a large number of potential demand sources (advertisers). Intuitively, as more bidders (which are typically ad networks) participate in the exchange, publishers should get better prices for their impressions. In turn, as more publishers come to sell, more buyers join in search for a wider advertising audience. The role of an ad exchange is to provide a platform for buyers and sellers to meet. Most ad exchanges earn money by keeping a portion of the amount from the sale of an impression. The remainder of the sale price goes to the seller of the impression, namely, the publisher. Present day ad exchanges are most closely aligned with publisher goals: if publishers make more money, the ad exchange also makes more money.

The operational details of an ad exchange are depicted in Figure 1 (Muthukrishnan 2009). When a visitor comes to a website (e.g., http://www.weather.com/), the publisher sends a request to the ad exchange to auction the impression. Upon receiving this request, the ad exchange immediately starts an auction to sell the impression. Based on the information revealed about the impression, every bidder sends a bid, and in some situations, a link to an ad server that can be called to supply the ad to be displayed. For a bidder to enter a legal bid, its response must be received in real-time, typically in less than 50 milliseconds. Once all legal bids are in, the ad exchange decides the winner of the auction and displays the winner’s ad for the impression. The entire process must be completed fast enough, so that the web user does not experience any perceptible delay in the rendering of the ad. Most advertisers do not directly participate in the exchange; rather they usually work with ad networks that bid on their behalf. Thus, bidders in online ad exchanges are often ad networks and one bidder usually works with multiple advertisers.

As seen above, buyers in an ad exchange respond with a bid based on the information they receive (or possess) about an impression. Previous research has studied the possibility of private information that buyers may have and how the presence of such information could affect the equilibrium price of impressions in the market (Abraham et al. 2013). Our study also examines the impact of impression-related information, but from a different perspective: How much impression-related information should an ad exchange reveal? Or the related question: Can an ad exchange make more money by hiding some details of an impression?
When an auction of an impression starts, the ad exchange sends out bid requests to bidders. A bid request contains multiple attributes describing the impression context. Examples of attributes are publisher identity, user’s browser, search string (if one exists), user’s geographical location, ad format (banner or video), ad size (width and height of the impression) and ad position (header, above or below the fold, etc.) (Interactive Advertising Bureau 2012b). The ad exchange can make hide/reveal decisions on any of the available attributes.

A crucial piece of information contained in an impression is the publisher’s identity. Bidders (i.e., ad networks) can use this information to better predict the probability of a click and, hence, obtain a more accurate estimate of the advertiser’s value for the impression. Usually, the ad network is a re-seller: The ad network buys the impression but gets paid only if the ad it delivers leads to a click. Thus, in deciding how much to bid for an impression, the bidder must estimate the probability of a click for the ad it chooses to deliver for the impression. Our study examines whether, and under what circumstances, should an ad exchange reveal (hide) the identity of the publisher. For the ad exchange, the basic idea behind a revelation policy is to influence the predictive ability of bidders. While some amount of uncertainty about an impression always exists for a bidder, hiding serves to increase the extent of the uncertainty and, thereby, reduce the bidder’s predictive ability.

The analysis in our paper remains valid for any piece of information that has predictive value to the bidder and whose revelation can be controlled by the ad exchange (e.g., any of the attributes mentioned above). We use publisher identity information as an example of the attribute to be hidden/revealed throughout our analysis in this paper for two reasons. Publisher identity could indeed significantly influence the chances of an ad being clicked upon. For example, an outdoor clothing ad will likely work better if it is displayed on a website that contains trekking and camping information. In addition, while publisher identity is always known to the ad exchange, this information is not always made available to bidders. Another example of an attribute that is often missing is the position (on the web page) of the ad slot that is being auctioned. This is an important attribute because if the location is below the fold (implying that the web user has to scroll down to view the slot contents), the chances of a click on such an ad can become drastically lower. The analysis in this paper concerns any attribute that, if hidden, could change the revenue earned by the ad exchange. Of key interest is whether, and under what circumstances, could the ad exchange earn more revenue by controlling the revelation of

---

1 In the OpenX and the Zenovia exchanges, randomly collected samples of impression requests showed that publisher identity information is not always made available to bidders (Chitika, Inc. 2014). Chitika, Inc. is a bidder on both these exchanges. Out of 100000 randomly collected impression requests from OpenX, publisher identity information was missing in 2594 requests, while in Zenovia, out of 13775 randomly collected impression requests, publisher identity information was missing in 988 requests.
impression attribute information.

**The Two-Call Design**

The basic idea of information hiding in auctions is not new. For example, in storage-unit auctions (http://www.storagebattles.com/?do=index), bidders may not be provided full information about the assortment of items contained in the unit. However, once the winner is announced, the contents reveal themselves to the winner. Therefore, the winner has the information to extract her highest value from the items won. Thus, while bidders in physical auctions may not have full information while bidding, they naturally obtain full information after winning the auction. In ad exchanges, however, unless the exchange decides to reveal a publisher’s identity, this information may not be available to the bidder both at Bid-Time and at Win-Time.

In ad industry parlance, the *Ad Served on the Win Notice* specification in the OpenRTB API (Open Real-Time Bidding Application Programming Interface) is one way of serving ads to the ad exchange by bidders (Interactive Advertising Bureau 2012b). Under this specification, the bidder only returns a bid in response to a request. Then, once the winner is identified, the ad is served in response to a request made by the ad exchange to the winner. Notice that when the bid and ad are served separately, a *second* call needs to be made to the winner to obtain the ad. Thus, we refer to this design as the *Two-Call* design.

In the Two-Call design, we require that the impression attribute of interest (here, the publisher identity) that was potentially hidden at Bid-Time, is revealed to the winner at Win-Time. Not only does the Two-Call design allow flexibility for the bidder (since the bidder could make an informed decision about the choice of the ad), it also allows the ad exchange to be strategic about *when* to reveal attribute information to bidders. Figure 2 depicts the Two-Call design pictorially.

![Two-Call Design Diagram](image)

*Figure 2. Description of the Two-Call design (Bid-Time refers to the time before the bids are collected; Win-Time refers to the time after the bids are collected and a winner is decided.)*

**Related Literature**

Our paper stands at the interface of online ad exchange research and the literature on information revelation in physical auctions. In the following paragraphs, we summarize two streams of literature that are related to our work.

Studies on information revelation in physical auctions include Milgrom and Weber (1982), Lewis and Sappington (1994), Ottaviani and Prat (2003), Ganuza (2007), Bergemann and Pesendorfer (2007), Eső and Szentes (2007), Board (2009), and Ganuza and Penalva (2010). The key insight in this stream of work is that the auctioneer’s decision to reveal information to bidders depends on the trade-off between the impact of revealing more information on the total surplus and its impact on bidders’ information rents. In the second-price auction setup in our paper, this trade-off is the same as the impact of revealing more information on the change in the second-highest bid. We also examine this trade-off to obtain insights that are specific to the Two-Call design we study. Moreover, our paper distinguishes itself from this literature by considering the timing of revelation in online ad exchanges. Instead of only
considering the auctioneer’s revelation decision at Bid-Time (i.e., the beginning of the auction), we consider a design where the exchange can postpone the revelation of information to Win-Time.

The second stream of related literature is that on auctions in online ad exchanges. This new online demand-supply matching channel has spawned a host of rich research questions. Feldman et al. (2010) study auction mechanisms with intermediaries where, the ad exchange sells impressions to ad networks in an auction, and ad networks resell them to advertisers in another auction. Cavallo et al. (2012) propose a truth-telling auction mechanism between pay-per-click advertisers and pay-per-impression ad networks. Balseiro et al. (2013) study a dynamic game between one publisher and several advertisers in an ad exchange. Fu et al. (2012) compare the impact of additional information on the auctioneer’s expected revenue under simple mechanisms and under an optimal mechanism. Abraham et al. (2013) study the impact of cookie information on publisher revenue in a common value auction. Different from this literature on online ad exchanges that considers the ad exchange as a passive intermediary, our study considers an active ad exchange that takes information revelation decisions to maximize its revenue.

The rest of the paper is organized as follows. In the next section, we will define and study the Two-Call design. While valuations of an impression are private to bidders, we consider two possibilities as far as the availability of this information to the ad exchange is concerned: For the case when the ad exchange has no reliable knowledge of bidder valuations, we develop simple policies that do not use any knowledge of the valuations and establish their performance guarantees. For the case when the ad exchange has distributional knowledge about bidder valuations, we develop an informed heuristic that exploits this information and demonstrate its near-optimal performance. The last section concludes.

Information Revelation under the Two-Call Design: Model and Simple Policies

As mentioned earlier, the Two-Call design necessitates two contacts between the auctioneer and the bidders. The first contact occurs between the auctioneer and all bidders at the beginning of an auction (i.e., Bid-Time). The second contact occurs between the auctioneer and the winning bidder after the winner has been identified (i.e., Win-Time).

We first discuss the preliminaries and the notation, and then formulate the optimization problem.

Consider an ad exchange running second-price, sealed-bid auctions for selling impressions from a set of publishers, \( \Theta \). When an impression is available, the ad exchange starts an auction among bidders from a set \( C \). Let bidder \( c \in C \) serve ads from the set of advertisers, \( A^c \). We use \( v^c(\theta, a) \) to denote the (private) value that accrues to bidder \( c \) from serving the ad of advertiser \( a \) for an impression from publisher \( \theta \). In a second-price auction, \( v^c(\theta, a) \) is also the bid received from bidder \( c \) if she decides to show the ad of advertiser \( a \) and the publisher is revealed to be \( \theta \). The bidder can calculate her value as the expected revenue based on her contract with an advertiser. For example, under a pay-per-click contract, the expected revenue \( v^c(\theta, a) \) to bidder \( c \) is the product of the advertiser’s payment for a click on the ad and the click probability.

Information on Bidders’ Valuations: Let \( V \) be the matrix of bidder valuations. Each column of this matrix corresponds to a particular publisher \( \theta \), and each row corresponds to a particular combination of a bidder \( c \) and an advertiser \( a \). While the valuation \( v^c(\theta, a) \) is likely to be private to bidder \( c \), the ad exchange may have distributional knowledge about bidder valuations. In this section, however, we will assume that the ad exchange has no reliable knowledge about the valuation matrix \( V \). Our purpose here is to propose some simple policies that do not depend on any knowledge about the valuation matrix. Such policies may be useful to consider when the economic incentives of the bidders (or even the bidders) change over time and hence, it is difficult to collect any reliable knowledge of the valuations. In the next section, we will provide specific examples of what the ad exchange might know about \( V \), and how it can acquire and exploit such knowledge.

In the optimization problem that we will soon define, we will assume that the objective of the ad exchange is to maximize the expected payment, per auction, made by the bidders. Here, an expectation is taken with respect to a valuation distribution, that for our current purposes, is assumed to be un-
known to the ad exchange. Maximizing the expected payment per auction makes sense because an ad exchange usually keeps a fraction of the payment received from the winning bidder and passes the rest over to the publisher.

**Reserve Prices:** Publisher \( \theta \) sets \( r_\theta \) as the reserve price for those of its impressions which are sold with its identity revealed. We assume that reserve prices are not imposed for impressions where publisher identities are hidden. This is because, from a bidder’s perspective, it seems unreasonable to impose a reserve price for impressions with publisher identities hidden. Meanwhile, this assumption lends greater tractability to the analysis.

**Prior Probabilities:** The probability that an impression is from publisher \( \theta \) is \( p_\theta \). We assume that the distribution \( \{p_\theta\} \) is public information. This is because each bidder typically participates in millions of these auctions every day and any “non-trivial” bidder is likely to win in many (thousands) of these auctions in a short period of time. Thus, the bidders have sufficiently large samples of data to be able to estimate these probabilities. Moreover, bidders can make use of various online traffic estimation services that report statistics of page views and visitors of different publishers (e.g., https://www.compete.com/).

**Hide/Reveal Decisions:** For each publisher identity \( \theta \), the ad exchange has to decide \( q^\theta \in [0, 1] \), which is the probability with which \( \theta \) is revealed when an impression is from this publisher. We assume that, once decided by the ad exchange, these probabilities are known to the bidders. This is a reasonable assumption because of the high volume of impressions sold at an ad exchange. Thus, “non-trivial” bidders have sufficiently large samples of data to do the estimation. To learn the hide/reveal decisions, one idea is for a bidder to bid relatively high in all auctions for a short period of time to win almost all the auctions and observe the hide/reveal policy adopted by the ad exchange.

**Bidders’ Valuations of Impressions with Revealed Publisher Identities:** If publisher identity \( \theta \) is revealed, then bidder \( c \in C \) picks the advertiser with the highest valuation within its advertiser base, i.e., \( \arg \max_{a \in A^c} v^c (\theta, a) \). Consequently, bidder \( c \)’s valuation of this impression is \( w^c (\theta) = \max_{a \in A^c} v^c (\theta, a) \).

Since the ad exchange uses a second-price auction, each bidder truthfully submits this valuation as her bid.

**Bidders’ Valuations of Impressions with Hidden Publisher Identities:** If publisher identity \( \theta \) is hidden, then the conditional probability that an impression with hidden publisher identity is from publisher \( \theta \) is \( \frac{p_\theta q^\theta_R}{1 - \sum_{\theta' \in \Theta} p_{\theta'} q_{\theta'}}, \) where \( q^\theta_H = 1 - q^\theta_R \) is the probability of hiding \( \theta \). Because the winning bidder chooses an ad after knowing the publisher identity, bidder \( c \)’s valuation of this impression is \( \sum_{\theta} \frac{p_\theta q^\theta_R}{1 - \sum_{\theta' \in \Theta} p_{\theta'} q_{\theta'}} w^c (\theta) \), which will be reported as her bid.

Our notation is summarized in Table 1. We now proceed to calculate the expected payment per auction under the Two-Call design. An auction that reveals publisher identity \( \theta \) results in a sale only if the highest bid is at least the reserve price \( r_\theta \), and the payment to the ad exchange in this case is \( \max \left\{ SH_c \left\{ w^c (\theta) \right\}, p_\theta \right\} \). Since reserve prices are not applicable for impressions with a hidden publisher identity, every auction that hides the publisher identity at Bid-Time results in a payment of \( \sum_{\theta} \frac{p_\theta q^\theta_R}{1 - \sum_{\theta' \in \Theta} p_{\theta'} q_{\theta'}} w^c (\theta) \) to the ad exchange.

Let \( \tilde{q}_R \) denote the vector of revealing probabilities \( \{q^\theta_R, \theta \in \Theta\} \), implying that the publisher identity of an impression from publisher \( \theta \) is revealed with the probability \( q^\theta_R \) and hidden with the probability

---

2This exercise is a test of the ad exchange’s ability to make hide/reveal decisions where the bidder valuations follow a distribution known to an oracle, but hidden from the ad exchange.

3More generally, the ad exchange could consider different classes of impressions for the same publisher and implement a different \( q^\theta_R \) for each class of impressions. This generalization does not affect our analysis.
Table 1. Notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C )</td>
<td>The set of bidders in the auction.</td>
</tr>
<tr>
<td>( c )</td>
<td>A bidder in the auction, ( c \in C ).</td>
</tr>
<tr>
<td>( A' )</td>
<td>The set of advertisers of bidder ( c ).</td>
</tr>
<tr>
<td>( a )</td>
<td>An advertiser, ( a \in A' ).</td>
</tr>
<tr>
<td>( \Theta )</td>
<td>The set of publisher identities.</td>
</tr>
<tr>
<td>( \theta )</td>
<td>A publisher identity, ( \theta \in \Theta ).</td>
</tr>
<tr>
<td>( v^c(\theta, a) )</td>
<td>The (private) value that accrues to bidder ( c ) from serving the ad ( a ) for an impression from publisher ( \theta ).</td>
</tr>
<tr>
<td>( w^c(\theta) = \max_a v^c(\theta, a) )</td>
<td>The maximum value that accrues to bidder ( c ) from serving an ad for an impression from publisher ( \theta ).</td>
</tr>
<tr>
<td>( p_\theta )</td>
<td>The probability that an impression is from publisher ( \theta ).</td>
</tr>
<tr>
<td>( r_\theta )</td>
<td>The reserve price of an impression set by publisher ( \theta ).</td>
</tr>
<tr>
<td>( q_\theta^R )</td>
<td>The probability of revealing publisher identity ( \theta ). The probability of hiding publisher identity ( \theta ) is ( q_\theta^H = 1 - q_\theta^R ).</td>
</tr>
<tr>
<td>( q_\theta^R )</td>
<td>The vector of revealing probabilities.</td>
</tr>
<tr>
<td>( \text{SH} {x_1, x_2, \ldots, x_n} )</td>
<td>The second-highest value among the values ( x_1, x_2, \ldots, x_n ).</td>
</tr>
</tbody>
</table>

\[ q_\theta^H = 1 - q_\theta^R. \]

Then, the total expected payment, per auction, by the bidders is

\[
\begin{align*}
\text{Rev}^{II}(\vec{q}_R) &= \sum_\theta p_\theta q_\theta^R \mathbb{E} \left[ \max \left\{ \text{SH}_c \{ w^c(\theta) \}, r_\theta \right\} 1 \left( \max_c \{ w^c(\theta) \} \geq r_\theta \right) \right] +
\left( 1 - \sum_{\theta \in \Theta} p_\theta q_\theta^R \right) \mathbb{E} \left[ \text{SH}_c \left\{ \sum_{\theta \in \Theta} \frac{p_\theta q_\theta^H}{1 - \sum_{\theta \in \Theta} p_\theta q_\theta^R} w^c(\theta) \right\} \right] \\
&= \sum_\theta p_\theta q_\theta^R \mathbb{E} \left[ \max \left\{ \text{SH}_c \{ w^c(\theta) \}, r_\theta \right\} 1 \left( \max_c \{ w^c(\theta) \} \geq r_\theta \right) \right] +
\mathbb{E} \left[ \text{SH}_c \left\{ \sum_{\theta} p_\theta q_\theta^H w^c(\theta) \right\} \right].
\end{align*}
\]

(1)

where \( 1 (\cdot) \) is the standard indicator function. The expectation operator \( \mathbb{E} \) above is over the distribution of the bidder valuation matrix \( V \), that is assumed to be unknown (in this section) to the ad exchange. The optimization problem of the ad exchange under the Two-Call design is

\[
\max_{\vec{q}_R \in [0,1]} \text{Rev}^{II}(\vec{q}_R).
\]

Let \( \text{OPT}_2 \) denote the optimum value of this problem.

**Performance of Simple Policies under the Two-Call Design**

The optimization problem (2) is intractable because the second-highest of a set of numbers involved in its objective function does not possess any nice structural property to aid optimization, e.g., concavity. Therefore, we use the optimum value, \( \text{OPT}_2 \), only as a theoretical benchmark to assess the performance of some simple policies, namely, (i) the policy of completely revealing publisher identities, (ii) the policy of completely hiding publisher identities, but revealing them to winning bidders (for simplicity, we refer to this policy as the complete hiding policy in this section), (iii) the policy of revealing publisher identities with probability \( \frac{1}{2} \), and (iv) the better of (i) and (ii).

The results of Theorem 1 below report the performance of these simple policies under the Two-Call
design.\footnote{Proofs of all technical results are available on request.}

**Theorem 1** Under the Two-Call design, the following statements hold:

1. The expected revenue of the policy of revealing publisher identities with probability $\frac{1}{2}$ is at least 50% of the optimum revenue, i.e., $\text{Rev}^{\text{II}}(\frac{1}{2}) \geq \frac{1}{2}\text{OPT}_2$. Furthermore, there exist instances for which this bound is tight.

2. The expected revenue of the better of complete revealing and complete hiding provides at least 50% of the optimum revenue, i.e., $\max \left\{ \text{Rev}^{\text{II}}(1), \text{Rev}^{\text{II}}(0) \right\} \geq \frac{1}{2}\text{OPT}_2$.

3. Assume that $w^c(\theta)$ for every $\theta \in \Theta$, $c \in C$, are independently and identically distributed with a distribution $F_\theta$ with support $[0, u_\theta]$ for some $u_\theta < \infty$ and that the reserve price $r_\theta < u_\theta$. Then, the policy of complete revealing is asymptotically optimal.

4. Without reserve prices, the policy of complete hiding is optimal when there are two bidders, regardless of the number of advertisers that each bidder works with.

5. In general, both the complete hiding policy and the complete revealing policy can be arbitrarily bad, relative to OPT$_2$.

The simple policies of complete hiding, complete revealing, and mixing hiding and revealing equally, do not use any information about bidders’ valuations. Among these, the policy that equally mixes hiding and revealing guarantees an expected revenue that is at least half of the optimum. In the next section, we develop a heuristic that uses knowledge of bidders’ valuations to improve the revenue earned by the ad exchange. While the heuristic continues to offer the theoretical guarantee of 50% of the optimum, we will soon see that its performance on a comprehensive test bed is in fact near-optimal.

**An Informed Heuristic for the Two-Call Design**

Because the optimization problem formulated for the Two-Call design (i.e., Problem (2)) is difficult to solve to optimality, it is worthwhile to develop a heuristic to solve this problem. We refer to the heuristic developed in this section as *informed* because it exploits knowledge of the bidder valuation matrix.

**Development of the Heuristic**

It makes intuitive sense that if the ad exchange is able to acquire some knowledge of bidder valuations, it should be able to perform better. Such knowledge may be acquired, for instance, from repeated interactions with bidders. We consider two kinds of knowledge. First, the ad exchange may have acquired enough experience to assume an analytical distributional form of the valuations. This, of course, implies that the bidder valuations can be assumed to be relatively stationary. On the other hand, it may sometimes be difficult to estimate a reliable, analytical distribution of bidder valuations. Thus, the second kind of knowledge the ad exchange might possess is a data set of valuation matrices (collected across numerous bidder–publisher encounters). We develop a heuristic that can adapt to both the situations described above.

Recall that, under the Two-Call design, the expected revenue that the ad exchange is trying to maximize is

\[
\text{Rev}^{\text{II}}(q_k) = \mathbb{E} \left[ \sum_\theta p_\theta q_k^\theta \max \{ \text{SH}_c \{ w^c(\theta) \} , r_\theta \} 1 \left( \max \{ w^c(\theta) \} \geq r_\theta \right) + \text{SH}_c \left\{ \sum_\theta p_\theta q_k^\theta w^c(\theta) \right\} \right]
\]  

\[
= \sum_\theta p_\theta q_k^\theta \mathbb{E} \left[ \max \{ \text{SH}_c \{ w^c(\theta) \} , r_\theta \} 1 \left( \max \{ w^c(\theta) \} \geq r_\theta \right) \right] + \mathbb{E} \left[ \sum_\theta p_\theta \text{SH}_c \left\{ \sum_\theta p_\theta q_k^\theta w^c(\theta) \right\} \right]
\]  

\[
\text{(since } \sum_\theta p_\theta = 1)\text{ }
\]

\[
= \sum_\theta p_\theta \left\{ q_k^\theta \mathbb{E} \left[ \max \{ \text{SH}_c \{ w^c(\theta) \} , r_\theta \} 1 \left( \max \{ w^c(\theta) \} \geq r_\theta \right) \right] + \mathbb{E} \left[ \text{SH}_c \left\{ \sum_\theta p_\theta q_k^\theta w^c(\theta) \right\} \right] \right\}
\]  

\text{(combining terms for each } \theta)
by hiding a publisher only if the gap above is maybe less attractive to bidders. This balance between the merits and demerits of hiding is achieved by choosing the reserve price for this publisher. Thus, revealing is bad if this gap is large. On the other hand, by hiding a publisher, it is revealed that this publisher gets the higher of these second-highest valuation and its reserve price; this results in a reduced expected value. Put differently, by revealing a publisher identity, the decision of hiding/revealing this publisher identity in twoways. First, by revealing a publisher identity, when the maximum valuation exceeds the reserve price, the ad exchange gets the higher of the second-highest valuation and the reserve price; this results in a reduced expected revenue if the maximum valuation is much higher than the second-highest valuation and the reserve price. Put differently, by revealing a publisher identity, the ad exchange forgoes the revenue corresponding to the gap between the first-highest valuation and the higher of the second-highest valuation and the reserve price for this publisher. Thus, revealing is bad if this gap is large. On the other hand, by hiding a publisher identity, the ad exchange mixes this publisher in a pool of other hidden publishers that may be less attractive to bidders. This balance between the merits and demerits of hiding is achieved by hiding a publisher only if the gap above is sufficiently high.

The notion of what constitutes as sufficiently high is captured in the heuristic, where we set

\[ q^\theta_R = \begin{cases} 
1 & \text{if } \mathbb{E} \left[ \max_c \{ \text{SH}_c \{ w^c(\theta) \} , r_\theta \} 1 \left( \max_c \{ w^c(\theta) \} \geq r_\theta \right) \right] > \alpha \mathbb{E} \left[ \max_c \{ w^c(\theta) \} \right], \\
0 & \text{otherwise.} 
\end{cases} \]  

(5)

The heuristic process takes as input a set \( \Lambda = \{ \alpha_1 = 0, \alpha_2, \ldots, \alpha_{m-1}, \alpha_m = 1 \} \), where \( 0 < \alpha_i < 1 \) for all \( i \in \{2, 3, \ldots, m-1\} \), of the values of the threshold \( \alpha \) in (5). We generate \( |\Lambda| \) hide/reveal policies, one corresponding to each value of \( \alpha_i \in \Lambda \), and then pick the best one.

As indicated above, the purpose of the threshold \( \alpha \) in our heuristic criterion is to balance local versus global goals. Observe that \( \alpha = 0 \) (resp., \( \alpha = 1 \)) corresponds to the policy of complete revealing (resp., complete hiding). The heuristic experimentally sets \( \alpha \) to a value that balances the goals of (a) not going the gap between the maximum valuation and the higher of the second-highest valuation and the reserve price (a local goal, since it only concerns the characteristics of a particular publisher) and (b) the effect of mixing this publisher with a pool of other (better or worse) hidden publishers (a global goal, since introducing a particular publisher in a pool affects the whole pool).
Under the Two-Call design, the comparison in (5) requires knowledge of the distribution of \( w^c(\theta) = \max_a v^c(\theta, a) \), for each bidder \( c \) and each publisher identity \( \theta \). Note that when publisher identity \( \theta \) is revealed, bidder \( c \) submits \( w^c(\theta) \) as her bid. Thus, the nature of the knowledge possessed by the ad exchange about valuations, based on historical interactions with bidders, can be of the following two types. In each case, we discuss the evaluation of the comparison in (5), the subsequent selection of a heuristic policy, and the evaluation of this policy (i.e., the comparison of the heuristic policy with an optimal policy).

- When the ad exchange has analytical distributions of \( w^c(\theta) \), it may be possible to get closed-form expressions of the quantities being compared in (5). Thus, for each \( \alpha_i \in \Lambda \), the ad exchange can obtain a particular hide/reveal policy. To pick the best policy among these \(|\Lambda|\) policies, the ad exchange can simulate data drawn from the valuation distribution and numerically compute the average revenue corresponding to each of the \(|\Lambda|\) policies using the expected revenue expression (3). Specifically, for a reasonably large integer \( N \), the ad exchange can simulate \( N \) samples of the matrix \( \{w^c(\theta)\} \) to compute these values. The heuristic policy we pick is the one that gives the highest average revenue among the \( |\Lambda| \) policies. To find an optimal policy and compare its performance with that of the heuristic policy, we consider the following convenient discretization of the space of the decision vector \( (q^\theta_{kR}, \theta \in \Theta) \): \( q^\theta_{kR} \in \{ \frac{k}{K}, k = 0, 1, 2, \ldots, K \} \) for some integer \( K \), for each \( \theta \in \Theta \). We enumerate over all possible choices of \( q^\theta_{kR} \) over all publishers. Thus, the total number of possible policies is \((K + 1)^{|\Theta|}\). Using (3), we compute the average revenue for each of these policies on the same simulated data that is used to pick the heuristic policy. The policy that generates the highest average revenue is treated as the optimal policy.

- When the ad exchange has no analytical closed-form distributions about valuations, the comparison in (5) is numerically performed using historical observations (of bids submitted by bidders from previous auctions) of the matrix \( \{w^c(\theta)\} \), to get a hide/reveal policy for each \( \alpha_i \in \Lambda \). Again, the heuristic policy is chosen to be the one that generates the highest average revenue among the \(|\Lambda|\) policies. The optimal policy is computed as before, except that historical observations are used instead of simulated data.

Note that both the policy of complete hiding and that of complete revealing are included in the set of \(|\Lambda|\) policies that determine the heuristic policy. Therefore, from the second statement of Theorem 1, we have:

**Corollary 2** The expected revenue of the heuristic policy is at least 50% of the optimum revenue.

**Simplification under Uniform Bidder Valuations**

We observe that the comparison in (5) requires the calculation of two expectations. This calculation is difficult in general, but the computational burden can be greatly eased if closed-form expressions for the expectations are available. We illustrate this simplification below.

Assume that the valuations of bidder \( c \in C \) towards publisher \( \theta \in \Theta \) follow a uniform distribution with the support \([0, b_{c\theta}]\). For a given \( \theta \), let \( b_{(c),\theta} \) be the \( c^{th} \) smallest \( b_{c\theta} \) and let \( b_{(0),\theta} = 0 \). Then, we can derive the following closed-form expressions\(^5\) for the expectations needed to make the hide/reveal comparison in (5):

\[
\mathbb{E} \left[ \max_c \{w^c(\theta)\} \right] = \sum_{c=1}^{|C|} \left\{ \sum_{i=1}^{c} \frac{1}{(i-1,\theta) b_{(i),\theta} - b_{(i-1),\theta}} b_{(i),\theta} \prod_{j>i} b_{(j),\theta} \right\}. 
\]

\(^5\)The detailed derivations of the above expressions are available on request.
• If $r_\theta \in [b_{(s-1),\theta}, b_{(s),\theta}]$ and $s \in \{1, 2, \ldots, |C|\}$, then

$\mathbb{E}\left[\max \{S_{\theta'} \{w^c(\theta)\}, r_\theta\} \cdot 1 \left(\max_c \{w^c(\theta)\} > r_\theta\right)\right]$

$= \frac{1}{\prod_{l=1}^{[C]} b_{(l),\theta}} \sum_{l=1}^{[C]} \frac{b_{(l),\theta} r_{\theta}^{(l) + 1} - r_{\theta}^{(l) + 1 - s}}{1 - s} \sum_{k,l \geq s} b_{(k),\theta} - r_{\theta}^{(k) + 1 - s} \frac{b_{(l),\theta} r_{\theta}^{(l) + 1 - s}}{1 - s} \sum_{m=0}^{[C] - 2} \prod_{s=1}^{m} b_{(s),\theta} \sum_{l=m+1}^{[C]} \frac{b_{(l),\theta} - b_{(l-m),\theta}}{[C] - m} \sum_{u \neq 1, u > m} b_{(u),\theta} - \left(\sum_{m=0}^{[C]} b_{(m+1),\theta} - b_{(m),\theta}\right) \frac{b_{(l),\theta} r_{\theta}^{(l) + 1 - m}}{[C] + 1 - m}$. (7)

Using these expressions, we can easily obtain the heuristic hide/reveal decisions corresponding to any specific value of the threshold $\alpha$. Then, as before, the best value of $\alpha$ can be chosen to arrive at the heuristic policy. In one of our experiments that follow, we use the above closed-form expressions to get the heuristic policy and evaluate its performance by comparing with an optimal policy.

We are now ready to describe our experimental setup and discuss the results obtained.

**Numerical Simulations**

The numerical simulations conducted in this section are designed to evaluate the performance of the various policies developed in this study. Specifically, we are interested in examining the benefit of an informed policy over those that use no knowledge of the bidder valuations. The benchmark for the performance of these policies is, of course, an optimal policy. Since finding an optimal policy requires an exhaustive enumeration over all possible policies, we only consider problems of modest size. In our experiments, we set $|C| = 5$ (number of bidders) and $|\Theta| = 8$ (number of publishers). Thus, we have a $5 \times 8$ valuation matrix, $(w^c(\theta))$. We set $\Lambda = \{0, 0.1, 0.2, \ldots, 0.9, 1\}$.

We consider two settings of the bidder valuation matrix. In the first setting (S1), bidder valuations are drawn from different uniform distributions that depend on (bidder, publisher) pairs (Abraham et al. 2013). In this setting, closed-forms (6) and (7) can be used to find a heuristic hide/reveal policy for a given value of the threshold $\alpha$. In the second setting (S2), we use contextual knowledge to obtain the matrix of bidder valuations. These settings are summarized below.

• (S1) Each $w^c(\theta)$ follows a (distinct) uniform distribution with support $[0, b_{c\theta}]$, where $c$ denotes the bidder $c \in \{1, 2, 3, 4, 5\}$ and $\theta$ denotes the publisher $\theta \in \{1, 2, 3, 4, 5, 6, 7, 8\}$.

• (S2) Here, to generate a different kind of bidder valuation matrix, we use some important factors that affect the value of an impression to a particular bidder. Most bidders (or ad networks) work on pay-per-click contracts with advertisers. In such contracts, the expected value accrued to a bidder from an impression is the revenue-per-click times the probability of a click. To obtain a realistic bidder valuation matrix, we randomly draw click probabilities from a Beta distribution (with chosen shape parameters) and multiply these probabilities by the revenue-per-click that is randomly drawn from a uniform distribution over a certain interval. The Beta distribution is known to be a reasonable model of the click-probability distribution for impressions coming from a particular publisher (Mookerjee et al. 2012). While the ad exchange may have direct access to bid-data, our purpose here is to simulate realistic conditions with a data set (one where the valuations follow no well-known, analytical distribution) to see if the heuristic is able to perform well under these conditions.

We next describe the details of the experiment conducted under each of the settings described above.

**Experiment S1: Uniform Bidder Valuations**

Under this setting, $w^c(\theta)$ follows a uniform distribution with support $[0, b_{c\theta}]$; $c \in C, \theta \in \Theta$. We consider the cases with and without reserve prices. To incorporate reserve prices, for each distribution setup (i.e., for each set of $b_{c\theta}$ values), a randomly drawn value from the uniform distribution over
\[ \left[ 0, \frac{\sum c b_c}{5} \right] \] is set as the reserve price for publisher \( \theta \). Thus, the reserve price for a publisher is likely to be less than the maximum bid for this publisher.

To evaluate the performance of the heuristic policy we need to calculate the optimal expected revenue, which is the highest expected revenue over all possible policies. To compute this optimum, we consider the following discretization of the space of the decision variable \( q^\theta \): \( 0, 0.25, 0.5, 0.75, 1 \), and enumerate over all possible choices of \( q^\theta \) over all publishers. Thus, the total number of possible policies is \( 5^b = 390625 \). The average revenue of each policy is numerically evaluated using (3), and the policy that generates the highest average revenue is chosen as the optimal policy.

We test the heuristic by considering different factors in the experiments:

- **Publisher Distribution**: We consider two different setups of the probabilities \( (p_\theta) \): (1) all publishers are equally likely, and (2) publishers are distributed following the Pareto Principle (a large portion of impressions come from a few publishers), as in

  \[
  (p_\theta) = [0.35, 0.2, 0.12, 0.1, 0.08, 0.06, 0.05, 0.04].
  \]

- **Bidder Heterogeneity**: Both homogeneous and heterogeneous bidder populations are simulated. For homogenous bidders, the values of \( b_c, \theta \) (the uniform bound of the normal distribution from which valuations are drawn) across the bidders \( c \in C \) are set to be similar. For heterogeneous bidders, two bidder types are generated. Bidders of the same type have similar values of \( b_c, \theta \), but the values of \( b_c, \theta \) are significantly different from one type to another. We consider two ways of generating bidder valuations:

  1. In the first approach, specific values of \( b_c, \theta \) are used to generate the valuation matrices. We consider the following scenarios:

     (a) **One Different Bidder**: For any given publisher, randomly pick one bidder and use the uniform distribution over \([0, x]\), where \( x \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\} \), to draw valuations for that bidder. The valuations of the other four bidders towards this publisher are uniformly distributed in \([0, 5]\). Thus, there are ten parameter settings. Each parameter setting corresponds to specific values of \( b_c, \theta \), \( c \in C \), \( \theta \in \Theta \), and thus specific distributions for each \((c, \theta)\) cell of the valuation matrix.

     (b) **Two Different Bidders**: This is similar to the above, except that there are two bidders who are different from the rest. For any given publisher, randomly pick two bidders and use the uniform distribution \([0, x]\), where \( x \in \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\} \), to draw valuations for these bidders. The valuations of the other three bidders towards this publisher are drawn from the uniform distribution over \([0, 5]\). Once again, there are a total of ten parameter settings.

     (c) **Homogenous Bidders**: For any given publisher, the valuations of all five bidders come from the uniform distribution over \([0, 5]\).

  2. In the second approach, there is more bidder heterogeneity since we only set a range from which the upper bound \( b_c, \theta \) for a bidder’s valuation distribution is randomly picked. Three sub-cases (similar to (a), (b), and (c) above) also apply here. Since the heuristic uses an average-based criterion to make hide/reveal decisions, increasing the variation in the valuations could hurt its performance. We consider the following scenarios:

     (d) **One Different Bidder**: For any given publisher, randomly pick one bidder and use the uniform distribution over \([0, x]\), where \( x \) is a random integer between 50 and 150, to draw valuations for that bidder. For each of the other four bidders, randomly pick an integer \( y \) between 1 and 10, and then use the uniform distribution over \([0, y]\) to draw her valuations. We randomly pick 100 such parameter settings.

     (e) **Two Different Bidders**: This is similar to the above, except that there are two bidders who are different from the rest. For any given publisher, randomly pick two bidders. For each of these two bidders, randomly pick an integer \( x \) between 50 and 150, and then use
the uniform distribution $[0, x]$ to draw her valuations. For each of the other three bidders, randomly draw an integer $y$ between 1 and 10, and use the uniform distribution over $[0, y]$ to draw her valuations. Again, we randomly pick 100 such parameter settings.

(f) **Homogenous Bidders:** For any given publisher, for each of the five bidders, randomly pick an integer $x$ between 50 and 150, and use the uniform distribution over $[0, x]$ to draw her valuations. As before, we randomly pick 100 such parameter settings.

Over the two approaches of generating bidder valuations, we have a total of 321 different parameter settings (21 for the first approach, 300 for the second approach). To obtain our heuristic policies, we use the closed-form expressions (6) and (7) for the components of the heuristic criterion to make hide/reveal decisions. For each parameter setting, we randomly generate 1000 valuation matrices. This data is used to obtain the heuristic policy as well as an optimal policy for each parameter setting.

**Results: Experiment S1**

Experimental results for S1 are reported in Table 2. In this table, columns (1)–(4) report the average ratio of the expected revenue generated from the heuristic to that of, respectively, the optimal policy, the policy of mixing hiding and revealing equally, the complete revealing policy, and the complete hiding policy. Note that the maximum and minimum ratios in each column are highlighted in bold font. We make the following observations.

- The heuristic is near-optimal—observe in columns (1) of Table 2, that the expected revenue of the heuristic policy always exceeds 94% of the optimal expected revenue. Moreover, this near-optimal performance is robust across publisher setups (i.e., equally likely publishers or otherwise) as well as with or without reserve prices.
- Observe that the values in columns (2)–(4) are all greater than or equal to 1. Thus the heuristic policy performs better than the three uninformed policies, namely the policy of mixing hiding and revealing equally, the complete revealing policy and the complete hiding policy. Moreover, the values in columns (3) show that the performance of the complete revealing policy changes significantly from one scenario to another. When the gap between the first-highest valuation and the higher of the second-highest valuation and the reserve price of a publisher is low (e.g., scenarios (b) and (c)), complete revealing is near-optimal. Yet, it performs much worse in other scenarios. Similarly, although complete hiding generates a revenue that is near-optimal when the gap is high (scenarios (a) and (d)), its performance in other scenarios is much worse. Taken as a whole, the above findings provide evidence that the heuristic successfully chooses a good policy across the spectrum of available policies—from complete revealing to complete hiding.
- Examining the values in columns (3)–with and without reserve prices—we see that setting a reserve price does typically increase the revenue from the complete revealing policy. Concerning reserve prices, we have two observations. First, while reserve prices (together with complete revealing) can be beneficial for the ad exchange, these prices are typically set by publishers and are not in the control of the ad exchange. Thus, these prices may not be set optimally from the perspective of the ad exchange. Second, the heuristic policy does a much better job of extracting revenue for the ad exchange than publisher-set reserve prices.

**Experiment S2: Bidder Valuations from Empirical Knowledge**

Under this setting, we randomly draw the revenue-per-click of bidder $c$ for publisher $\theta$, $y_{c\theta}$, as an integer between 1 and 100. For click probabilities, we use Beta distributions with the same mean but different variances to see how the dispersion of the Beta distribution affects the heuristic. Three click probability settings are considered:

(i) The click probability of the ad served by bidder $c$ for publisher $\theta$ follows a Beta distribution with shape parameters $\alpha = 0.005$ and $\beta = 2$. The mean of this Beta distribution is $\frac{\alpha}{\alpha + \beta} = \frac{0.005}{0.005+2} \approx 0.0025$, and the variance is $\frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \approx 0.00083$ (or a standard deviation of 0.0288, implying that the variance is high relative to the mean). This high variance setting should favor hiding.

(ii) We next introduce some heterogeneity in publisher click-probabilities. The click probability of the ad served by bidder $c$ for publisher $\theta \in \{1, 2, \ldots, 8\}$ follows a Beta distribution with shape...
parameters $\alpha = 0.005 \times 3^{\theta-1}$ and $\beta = 2 \times 3^{\theta-1}$. These Beta distributions have the same mean as in the previous case (0.0025), but the variance is typically lower, implying that revealing should be favored (relative to (i) above).

(iii) The publishers are once again homogeneous, but the Beta distribution from which click-probabilities are drawn has a much lower variance. This situation should also favor revealing, but more strongly than the previous situation. The click probability of the ad served by bidder $c$ for publisher $\theta$ follows a Beta distribution with shape parameters $\alpha = 0.005 \times 3^7$ and $\beta = 2 \times 3^7$. The mean and the variance of this Beta distribution are approximately 0.0025 and 5.67 $\times$ 10$^{-7}$, respectively.

Similarly, we consider situations with and without reserve prices. When reserve prices are present, $r_{\theta}$ is a random draw from the uniform distribution over $[0, \eta_{\theta}]$, where $\eta_{\theta} = \frac{\alpha}{\alpha + \beta} \times \sum_{y \leq \theta} \frac{y \cdot \theta}{5}$. Note that $\frac{\alpha}{\alpha + \beta}$ is the mean of the Beta distribution and the multiplier is the average revenue-per-click. Thus, the reserve price should typically be lower than the maximum bidder valuation. Also, we consider the same two setups as in Experiment S1 of the publisher probabilities $(p_{\theta})$.

With two publisher setups, two reserve price setups, and three click-probability situations, there are $2 \times 2 \times 3 = 12$ scenarios in this experiment. For each of these scenarios, we randomly pick 100 parameter settings: One parameter setting is composed of 40 integers (i.e., the values of revenue-per-click) randomly drawn between 1 and 100 and a click probability setting. Therefore, we have a total of $12 \times 100 = 1200$ parameter settings.

Since there are no closed-form expressions to evaluate the comparison (10) under this experiment, for each of the 1200 parameter settings, 1000 valuation matrices are simulated to find the heuristic policy as well as an optimal policy.

Results: Experiment S2

Experimental results for S2 are reported in Table 3. Note that the maximum and minimum ratios in each column are highlighted in bold font. Our observations are similar as those under S1. In addition, we observe that, when the variance of the Beta distribution is relatively high (e.g., click probability setting (i)), hiding is beneficial because of the relatively high gap between the first-highest valuation and the higher of the second-highest valuation and the reserve price. This is confirmed from the values corresponding to the intersections of rows (i) and columns (4): The heuristic policy here is identical to the complete hiding policy, which is near-optimal (see columns (1)). On the contrary, when the variance of the Beta distribution is relatively low (i.e., click probability setting (iii)), revealing is more beneficial. Thus, the values corresponding to the intersections of rows (iii) and columns (3) are close to 1.
We also considered instances with a larger size – 20 bidders and 20 publishers – under the following setting: equally likely publishers, without reserve price and click probability setting (iii). The results are similar to the corresponding results in Table 3.

### Table 3. Numerical Experiments Results (click probabilities follow Beta distributions)

<table>
<thead>
<tr>
<th>Publisher Setup</th>
<th>Click Probability Setting</th>
<th>Without Reserve Prices</th>
<th>With Reserve Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publishers are equally likely</td>
<td>(i) 1</td>
<td>1.7416</td>
<td>0.9999 1.5425 3.4699 1</td>
</tr>
<tr>
<td></td>
<td>(ii) 0.9996</td>
<td>1.575</td>
<td>0.9998 1.1248 1.1972 1.0615</td>
</tr>
<tr>
<td></td>
<td>(iii) <strong>0.9876</strong></td>
<td><strong>1.0895</strong></td>
<td><strong>0.9881 1.0876 1.0043 1.1875</strong></td>
</tr>
<tr>
<td>Publishers are not equally likely</td>
<td>(i) 1</td>
<td>1.6767</td>
<td>0.9991 1.4959 3.0778 1</td>
</tr>
<tr>
<td></td>
<td>(ii) 0.9998</td>
<td>1.2262</td>
<td>0.9966 1.1513 1.2932 1.0392</td>
</tr>
<tr>
<td></td>
<td>(iii) <strong>0.9898</strong></td>
<td><strong>1.0724</strong></td>
<td><strong>0.9902 1.0712 1.0054 1.1481</strong></td>
</tr>
</tbody>
</table>

(1) Average ratio of the expected revenue of the heuristic to the optimal expected revenue.
(2) Average ratio of the expected revenue of the heuristic to that of the policy of mixing revealing and hiding equally.
(3) Average ratio of the expected revenue of the heuristic to that of the complete revealing policy.
(4) Average ratio of the expected revenue generated from the heuristic to that of the complete hiding policy.

### Is the Heuristic Policy Typically an Extreme Policy?

From Tables 2 and 3, observe that the better of complete hiding and revealing (the smaller of columns (3) and (4) for the same row) is near-optimal (close to 1). This observation indicates that even though the heuristic policy outperforms both the extreme policies, the better of complete revealing and complete hiding is very close to the heuristic policy. Thus the natural question arises: Does the heuristic typically provide an extreme policy? We construct the following scenario to show that the heuristic policy can hide only some publishers (i.e., rather than hiding all or revealing all) and perform significantly better than the better of complete hiding and complete revealing. In this scenario, publishers are equally likely, and there are no reserve prices. The bidder valuation setup is as follows:

- For publisher \( \theta \in \{1, 2\} \), for bidder \( c = \theta \), use the uniform distribution over \([0, x]\), where \( x \) is a random integer between 100 and 120, to draw her valuations. For each of the other four bidders, randomly pick an integer \( y \) between 1 and 5, and then use the uniform distribution over \([0, y]\) to draw her valuations.
- For publisher \( \theta \in \{3, 4, 5\} \), pick the last two bidders (i.e., \( c = 4 \) and 5), and, for each of these two bidders, use the uniform distribution over \([0, x]\), where \( x \) is a random integer between 50 and 60, to draw her valuations. For each of the other three bidders, randomly pick an integer \( y \) between 1 and 5, and then use the uniform distribution over \([0, y]\) to draw her valuations.
- For publisher \( \theta \in \{6, 7, 8\} \), for each of the five bidders, randomly pick an integer \( y \) between 1 and 5, and then use the uniform distribution over \([0, y]\) to draw her valuations.

We randomly pick 100 parameter settings under this scenario. To evaluate the comparison (10) under this experiment, for each of the 100 parameter settings, 1000 valuation matrices are simulated to find the heuristic policy as well as an optimal policy. The results in this experiment are reported in Table 4 below. Observe that the heuristic policy significantly outperforms each of the three uninformed policies. The smallest value in columns (2)–(4) is 1.6818, which corresponds to a 68% improvement. Therefore, in general, the best policy may not be extreme, but lie somewhere between complete hiding and complete revealing.

### Table 4. Numerical Experiments Results (a special scenario)

<table>
<thead>
<tr>
<th>Ratio Value</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.9579</td>
<td>1.6818</td>
<td>2.3432</td>
<td>1.6818</td>
<td></td>
</tr>
</tbody>
</table>

(1) Average ratio of the expected revenue of the heuristic to the optimal expected revenue.
(2) Average ratio of the expected revenue of the heuristic to that of the policy of mixing revealing and hiding equally.
(3) Average ratio of the expected revenue of the heuristic to that of the complete revealing policy.
(4) Average ratio of the expected revenue of the heuristic to that of the complete hiding policy.
(5) Average ratio of the expected revenue of the heuristic to that of the better of the policies of complete hiding and complete revealing.
Key Insights from Numerical Simulations

We highlight below, the key insights from the numerical simulations conducted in this section.

• The heuristic policy is near-optimal and achieves its excellent performance by appropriately balancing the forces of hiding and revealing. For a publisher, a high gap between the first and the higher of the reserve price and the second highest valuation favors hiding. At the same time, hiding a publisher with relatively large valuations can hurt revenue because bidders cannot discriminate this publisher from other hidden publishers with much lower valuations. Judging by its near-optimal performance, the heuristic policy does a good job of balancing the forces of hiding and revealing. In our experiments, the heuristic policy was, on average, about 98% of the optimal.

• An interesting finding from the experiments is that the heuristic policy can sometimes lie between the two extremes of complete hiding and complete revealing. However, how badly would the ad exchange do if it simply were to implement the best extreme policy (the better of complete hiding and complete revealing)? While the better of complete hiding and complete revealing does quite well, we demonstrate that the heuristic policy can do significantly better. For example, by introducing heterogeneity across publishers, it is easy to construct instances where the heuristic policy shows a 68% improvement over the best extreme policy.

• Are reserve prices (which are set by publishers) an effective way to address loss of revenue from complete revealing? Our finding here is that the introduction of reserve prices does boost the revenue for the complete revealing policy. However, for a given set of (publisher set) reserve prices, the ad exchange is still better off using the heuristic policy over one that reveals all publishers; specifically, an improvement of about 50%, on average.

Conclusions and Future Research

We study the implications of hiding pertinent impression attribute information (specifically, the identity of the publisher) to bidders in an ad auction. The hide/reveal decision is implemented by the auction site with a view to maximize the total expected revenue that can be extracted from the bidders in a second-price auction under the Two-Call design (corresponding to the “Ad Served on the Win Notice” OpenRTB specification).

We study two kinds of information revelation policies. First, we consider simple policies that can operate in conditions where the ad exchange has no knowledge of the bidder valuations for the impressions. While both the policies of complete hiding and complete revealing can individually be arbitrarily bad relative to the optimal policy, it is interesting that the policy of revealing each publisher identity with probability \( \frac{1}{2} \) earns at least half of the optimal revenue. This policy is a mixed-revelation strategy in that it randomizes between the policies of complete hiding and complete revealing. Second, we consider informed policies to make hide/reveal decisions. These policies exploit the information possessed by the ad exchange about bidder valuations: either captured in a valuation distribution, or at least a data set that records bids extracted from numerous publisher–bidder encounters. Informed hide/reveal decisions are guaranteed to be at least as good as uninformed ones (i.e., they both provide the same worst case performance guarantee). However, numerical comparisons show that not only do informed decisions provide much higher revenues than uninformed ones, the performance of informed decisions is near-optimal. The key takeaway from this part of the analysis is that information revelation matters in an ad auction and that hide/reveal decisions must be carefully made for the monetization of web traffic to publishers.

There are several avenues for future work that suggest themselves. We allowed the decision rights to hide or reveal information to lie with the ad exchange. It would be interesting to consider the decision rights to lie with the publishers and study how publishers would make (self) hide or reveal decisions in equilibrium. If such equilibrium decisions can be found, the other natural issue to explore would be the study of incentives to achieve the revenue of the centralized hide or reveal decisions. A related area of study would be to allow the bidder to choose to buy pertinent information that has predictive value (at a price set by the ad exchange or perhaps, more interestingly, in another information auction). That is, items of information in one set could be made available to all publishers for free, whereas other items of useful information could be obtained by the bidder for a set price. Finally, we assumed a second-price auction in all the analysis carried out in the current study. While second-price auctions appear
to be the norm, there are a few ad exchanges that operate some variant of the second price auction; for example, a modified second-price auction to account for the presence of common valuations among bidders. A more-detailed analysis of variants of the pure second-price mechanism may therefore be studied to reflect the nuances of ad auctions in some real-world settings. Our analysis in this paper assumed that reserve prices are not imposed on impressions with hidden publisher identities. While this seems reasonable from the viewpoint of bidders, the publishers may want to impose reserve prices for impressions with hidden identities. The analysis of optimal revelation policies under reserve prices for all impressions could be a challenging generalization of our work.

References


Chitika, Inc. 2014. Private Communications.


(50:5), pp. 1089–1122.