Hiding Publisher Identities in Ad Exchanges: Designs and Policies

(The bulk of the work was done by a student.)

Abstract

We study information revelation designs and policies in online ad exchanges that use a second-price auction mechanism. Two auction designs are proposed with varying degrees of control offered to the auction site to choose the timing and the extent of information released to bidders about an impression. Design I offers less control but is easier to implement. Design II offers greater control but requires more sophisticated bidding systems and integration between bidders and the ad exchange. We find that Design II can earn an arbitrarily large multiple of the revenue earned by Design I. We also study simple information revelation policies (that do not require detailed bidder valuation information) under each design and establish performance guarantees on the revenues earned by these policies with respect to the optimal revenue.

Keywords: Ad Exchanges, Information Revelation Policies and Designs, Approximation Guarantees

1 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, 2012). 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. 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, 2012).

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 (known as *cost-per-click pricing*) 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/index.php) 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/).

1.1 Ad Exchange: A Marketplace for Online Advertising

This paper studies online 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 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 rest 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.

[Insert Figure 1 here]

The operational details of an ad exchange are 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 the 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. 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 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 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., 2011). 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?

Information about an impression includes details such as the user's browser, search string (if one exists), ad size and format, location of the ad, and the identity of the publisher. Of these, 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. 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 careful hide/reveal policy for impressions 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[1].

1.2 Our Contribution and Related Literature

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. 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. 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.

Our paper examines two potential ad exchange designs and analyzes information revelation

[1]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. We consider publisher identity information in our analysis for two reasons: One, publisher identity is indeed one of the major pieces of information associated with an impression. Two, there are examples of real-world online exchanges (e.g., OpenX) where publisher identity information is often hidden from bidders.

policies under each of these designs. In Design I, bidders are required to submit bids together with the associated ad, to be displayed if the bid wins. Thus, in Design I, the ad must be chosen at bid time using the information that the exchange reveals to bidders at this time. Design II mimics physical good auctions: while the identity of the publisher may or may not be revealed at bid time, this information is always revealed to the winner of the auction. Design I differs from Design II in an important way. In Design I, only one call is made to a bidder who provides both the bid and the associated ad. In contrast, Designs II can be implemented only if time constraints allow a second call to be made for obtaining an ad from the winner. An example of a one-call auction can be found at AppNexus, whereas OpenX uses a two-call auction. A detailed description of the two designs is provided in Figure 2.

[Insert Figure 2 here]

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 of the second highest bid. We also examine this trade-off to obtain insights that are specific to each design we study. Moreover, our paper distinguishes itself from this literature by considering its timing 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 possible design where the exchange which can postpone the revealing of information to the point in time when the winning bidder has been determined. We find that the impact of revealing information to the winning bidder but hiding information from other bidders can significantly affect an ad exchange's expected revenue.

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. (2012) study a dynamic game between a 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. (2011) study the impact of cookie information on publisher revenue in a common value auction. Different from this literature 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. While doing so, we shed light on how simple information revelation policies within each design perform against the optimal revenue that can be achieved in the design.

2 The Model

In this section, we first formalize the setting for the ad exchange and then define the two potential designs of information revelation that we will investigate in the remainder of the paper.

2.1 Preliminaries

Consider an ad exchange running second-price, sealed-bid auctions for selling impressions from a set of publishers, Θ. When an impression from a publisher $\theta \in \Theta$ is reported to the ad exchange, the ad exchange starts an auction among bidders (who are typically ad networks) from a set C. Bidder $c \in C$ serves a set of advertisers, A^c. The valuation of bidder c (i.e., her expected revenue) from serving an impression of publisher θ to advertiser a is denoted by $v^c(\theta, a)$. The bidder can calculate her expected revenue based on her contract with an advertiser. For example, under a pay-per-click contract, the expected revenue for the bidder is the product of the advertiser's payment for a click on the ad and the click probability.

We assume that the publisher-IDs (i.e., θs) of the impressions are independently and identically distributed realizations from a distribution $\{p_\theta\}$. We further assume that this distribution is public information. Our basic notation is summarized in Table 1.

[Insert Table 1 here]

As we will soon see, in each of the two designs we investigate, the ad exchange chooses publisher-specific probabilities with which publisher-IDs are revealed to bidders. *We will* *assume that these probabilities are known to the bidders.* This is a reasonable assumption because of the high volume of impressions sold at an ad exchange, with individual bidders participating in millions of such auctions each day. For an impression that is hidden (i.e., its publisher-ID θ is not known to the bidders at the time of the auction), these probabilities can be used by bidders to determine their valuations (i.e., expected revenues). On the other hand, if an impression is revealed, then its valuation for a bidder $c \in C$ is $\max_{a \in A^c} v^c(\theta, a)$. Moreover, under a second-price auction, it is well known that self-interested bidders truthfully submit their valuations as their bids.

For each auction, the ad exchange generates revenue by taking a fixed percentage of the payment to the publisher from the winning bidder. Thus, the total revenue of the ad exchange is a common fixed percentage of the total payment made by the bidders over all the auctions. Consequently, the objective of the ad exchange is to maximize the total payment by the bidders.

2.2 Design I

In Design I, a bidder is required to submit both a bid and an ad together to the ad exchange. For a publisher $\theta \in \Theta$, the ad exchange chooses a probability q_{R}^{θ} with which to reveal the impressions from this publisher[2]. Thus, the probability that an impression with publisher θ is hidden is $q_{\mathrm{H}}^{\theta} = 1 - q_{\mathrm{R}}^{\theta}$. Recall from Section 2.1 that p_θ is the probability that an impression is from publisher θ. Thus, the conditional probability that a hidden impression is from θ is

$$\frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}}.$$

We now proceed to calculate the total expected payment by the bidders under Design I. Consider a specific auction. Recall from Section 2.1 that if θ is revealed in an auction, then bidder c submits a bid[3] of $\max_{a \in A^c} v^c(\theta, a)$. The winning bidder is $\arg\max_c \left\{ \max_{a \in A^c} v^c(\theta, a) \right\}$, and her payment equals the second-highest bid, which we denote by $\mathrm{SH}_c \left\{ \max_{a \in A^c} v^c(\theta, a) \right\}$. If θ is hidden by the ad exchange, then the winning bidder is $\arg\max_c \left[\sum_\theta \frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} v^c(\theta, a) \right]$, and her payment equals $\mathrm{SH}_c \left\{ \max_{a \in A^c} \left[\sum_\theta \frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} v^c(\theta, a) \right] \right\}$.

[2]More generally, the ad exchange could consider different classes of impressions for the same publisher and implement a different q_{R}^{θ} for each class of impressions. This generalization does not affect our analysis.

[3]The ad submitted along with this bid is $\arg\max_c v^c(\theta, a)$.

The expected revenue from an auction is formulated below. We use $\boldsymbol{q}_{\mathrm{R}}$ to denote the vector of probabilities $\left(q_{\mathrm{R}}^{\theta},\, \theta \in \Theta\right)$. Let $w^c(\theta) = \max_{a \in A^c} v^c(\theta, a)$. Thus, the expected revenue from an auction under Design I is

$$\mathbb{E}\left[Rev^{\mathrm{I}}(\boldsymbol{q}_{\mathrm{R}})\right] = \sum_{\theta} p_\theta q_{\mathrm{R}}^{\theta} \mathrm{SH}_c \left\{w^c(\theta)\right\} + \left(1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}\right) \mathrm{SH}_c \left\{ \max_{a \in A^c} \sum_{\theta} \frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} v^c(\theta, a) \right\}$$

$$= \sum_{\theta} p_\theta q_{\mathrm{R}}^{\theta} \mathrm{SH}_c \left\{w^c(\theta)\right\} + \mathrm{SH}_c \left\{ \max_{a \in A^c} \sum_{\theta} p_\theta q_{\mathrm{H}}^{\theta} v^c(\theta, a) \right\}, \tag{1}$$

where $\boldsymbol{q}_{\mathrm{R}} \in [0,1]^{|\Theta|}$.

The optimization problem of the ad exchange under Design I is

$$\max_{\boldsymbol{q}_{\mathrm{R}}} \mathbb{E}\left[Rev^{\mathrm{I}}(\boldsymbol{q}_{\mathrm{R}})\right], \text{ subject to} : \boldsymbol{q}_{\mathrm{R}} \in [0,1]^{|\Theta|}.$$

Let OPT_1 denote the optimum value of this problem.

2.3 Design II

Design II works as follows. As in Design I, the ad exchange chooses, for a publisher $\theta \in \Theta$, a probability q_{R}^{θ} with which to reveal the impressions from this publisher. Let $q_{\mathrm{H}}^{\theta} = 1 - q_{\mathrm{R}}^{\theta}$. If θ is revealed in an auction, then bidders are required to simultaneously submit their bids and ads, as before. However, if θ is hidden, then the bidders are only required to submit their bids initially. Then, θ is revealed to the winning bidder, who chooses which ad to serve for that impression.

We now proceed to calculate the payment by the bidders. Consider a specific auction. If θ is revealed, then the following actions are triggered: bidder c submits a bid of $w^c(\theta) = \max_{a \in A^c} v^c(\theta, a)$, the highest bidder wins, and her payment is the second-highest bid $\mathrm{SH}_c\left\{w^c(\theta)\right\}$. If θ is hidden, then the bid computation is more involved. As before, the conditional probability that the hidden impression is from publisher θ is $\frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}}$. The value that bidder c obtains from this impression, if she wins it, is $\frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} w^c(\theta)$ (since the publisher-ID is revealed before ad selection). Therefore, her bid, which equals her expected revenue from the hidden impression, is $\sum_{\theta} \frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} w^c(\theta)$. The winning bidder is $\arg\max_{c} \left\{ \sum_{\theta} \frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} w^c(\theta) \right\}$ and her payment is $\mathrm{SH}_c \left\{ \sum_{\theta} \frac{p_\theta q_{\mathrm{H}}^{\theta}}{1 - \sum_{\hat{\theta} \in \Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} w^c(\theta) \right\}$.

As before, denote the decision vector of the ad exchange by $\boldsymbol{q}_{\mathrm{R}}$. Under Design II, the expected revenue from an auction is as follows:

$$\mathbb{E}\left[Rev^{\mathrm{II}}(\boldsymbol{q}_{\mathrm{R}})\right] = \sum_{\theta} p_\theta q_{\mathrm{R}}^\theta \mathrm{SH}_c\left\{w^c(\theta)\right\} + \left(1 - \sum_{\hat{\theta}\in\Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}\right) \mathrm{SH}_c\left\{\sum_{\theta} \frac{p_\theta q_{\mathrm{H}}^\theta}{1 - \sum\limits_{\hat{\theta}\in\Theta} p_{\hat{\theta}} q_{\mathrm{R}}^{\hat{\theta}}} w^c(\theta)\right\}$$

$$= \sum_{\theta} p_\theta q_{\mathrm{R}}^\theta \mathrm{SH}_c\left\{w^c(\theta)\right\} + \mathrm{SH}_c\left\{\sum_{\theta} p_\theta q_{\mathrm{H}}^\theta w^c(\theta)\right\}, \tag{2}$$

where $\boldsymbol{q}_{\mathrm{R}} \in [0,1]^{|\Theta|}$.

The revenue maximization problem of the ad exchange is as follows:

$$\max_{\boldsymbol{q}_{\mathrm{R}}} \mathbb{E}\left[Rev^{\mathrm{II}}(\boldsymbol{q}_{\mathrm{R}})\right], \text{ subject to : } \boldsymbol{q}_{\mathrm{R}} \in [0,1]^{|\Theta|}.$$

Let OPT_2 denote the optimum value of this problem.

2.4 Optimization within Each Design

To solve the optimization problem for each of the above two designs, the ad exchange needs to know the value of $v^c(\theta, a)$ for all θ, c and a. Typically, it is impractical for the ad exchange to obtain the complete valuation matrix $(v^c(\theta, a))$. Moreover, even if this matrix were known, the optimization problems are intractable. This is because their objective functions – since each involves the second-highest of a set of numbers – do not possess any nice structural property to aid optimization, e.g., concavity. Therefore, we use the optimum values – OPT_1, and OPT_2 – only as theoretical benchmarks to assess the performances of some simple policies.

3 Analysis of Design I

Recall from Section 2.2 that, under Design I, a bidder submits both a bid and an ad, and the ad exchange chooses a probability q_{R}^θ with which to reveal the impressions from publisher θ. The optimization problem of the ad exchange under this design is

$$\max_{\boldsymbol{q}_{\mathrm{R}}} \mathbb{E}\left[Rev^{\mathrm{I}}(\boldsymbol{q}_{\mathrm{R}})\right] = \max_{\boldsymbol{q}_{\mathrm{R}}} \left\{\sum_{\theta} p_\theta q_{\mathrm{R}}^\theta \mathrm{SH}_c\left\{w^c(\theta)\right\} + \mathrm{SH}_c\left\{\max_{a\in A^c} \sum_{\theta} p_\theta q_{\mathrm{H}}^\theta v^c(\theta, a)\right\}\right\},$$

where $\boldsymbol{q}_{\mathrm{R}} \in [0,1]^{|\Theta|}$. Recall that OPT_1 denote the optimum value of this problem.

A natural simple policy to assess first is that of always revealing publisher-IDs to all bidders (complete revealing policy). In Section 3.1, we show in Theorem 1 that this policy is good when the number of bidders is large. However, in general, this policy can be arbitrarily bad (Theorem 2). This motivates the consideration of hiding publisher-IDs (i.e., θ). In Section 3.2, we analyze the extreme policy of always hiding θ. Again, this policy can be arbitrarily bad in general (Theorem 4), but is optimal under a special case (Theorem 3). Finally, in Section 3.3, we show in Theorem 7 that, the policy of revealing with probability $\frac{1}{2}$ and hiding with probability $\frac{1}{2}$ achieves at least 50% of the optimum revenue.

3.1 Performance of the Complete Revealing Policy

This policy chooses $\boldsymbol{q}_{\mathrm{R}} = \mathbf{1}$. Then, the expected revenue generated from an auction is

$$\mathbb{E}[Rev^{\mathrm{I}}(\mathbf{1})] = \sum_{\theta} p_\theta SH_c \left\{ w^c(\theta) \right\}. \tag{3}$$

Theorem 1[4]**.** *Let F denote the joint distribution of a $|\Theta|$-dimensional random variable. Let F_θ denote the marginal distribution for component $\theta \in \Theta$. Assume that for every $\theta \in \Theta$, F_θ has support $[0, u_\theta]$ for some $u_\theta < \infty$. For every c, let $\boldsymbol{w}^c$ be a realization from F, where $\boldsymbol{w}^c = (w^c(\theta) : \theta \in \Theta)$. Then, the difference between the optimal expected revenue and the expected revenue under the policy of complete revealing converges in probability to 0 as $|C|$ approaches infinity. That is, the policy of complete revealing is asymptotically optimal.*

The intuition is simple. When the number of bidders is high, it is highly likely that, for every impression, there will be at least two bidders whose valuations are close to the maximum possible valuation for the corresponding publisher.

Given the result in Theorem 1, a natural question that arises is whether complete revealing would typically achieve the optimal expected revenue under any design. In practice, the answer to this question is negative since the number of bidders is not very high. This is because a bidder (typically an ad network) is an aggregate entity representing many individual advertisers.

Next, we show that, the policy of complete revealing can sometimes produce significantly lower expected revenue relative to the optimum.

Theorem 2. *In general, the percentage loss (relative to the optimum) from the policy of complete revealing can be as high as 100%.*

The intuition behind Theorem 2 is as follows. Consider a situation in which, for each impression, there is an idiosyncratic bidder who has a high valuation while all other bidders have relatively low valuations. Then, the policy of complete revealing will result in low revenue because, for every impression, the second-highest valuation is low. In contrast, the policy of complete hiding can result in higher revenue.

Theorem 2 motivates the consideration of hiding publisher-IDs. There are two opposing forces as a result of hiding publisher-IDs:

- (Positive effect) As in our intuitive explanation above, hiding may reduce the spread between the first- and second-highest bids, since bidders submit their expected valuations.

[4] Proofs of all technical results are available on request.

This helps avoid the situation in the proof of Theorem 2, where the second-highest bid was far away from the highest bid.

- (Negative effect) Hiding reduces the ability of bidders to serve the most appropriate ad (i.e., the ad with the highest $v^c(\theta, a)$). This, in turn, reduces their bids and negatively influences the revenue generated by the auction. As we will see soon, this effect drives the result in Theorem 4.

3.2 Performance of the Complete Hiding Policy

This policy chooses $\boldsymbol{q}_{\mathrm{R}} = \mathbf{0}$. Then, the expected revenue generated from an auction is

$$\mathbb{E}[Rev^{\mathrm{I}}(\mathbf{0})] = SH_c \left\{ \max_{a \in A^c} \sum_{\theta} p_\theta v^c(\theta, a) \right\}. \tag{4}$$

Theorem 3 below identifies a special case in which this policy is optimal. However, Theorem 4 shows that, in general, this policy can be arbitrarily bad, relative to the optimum.

Theorem 3. *The policy of complete hiding is optimal for the special case in which there are two bidders and each bidder works with one advertiser.*

Here, we construct a simple example to illustrate this theorem. Assume that there are two bidders and two publishers, and that each bidder works with one advertiser. The matrix of valuations, where the rows correspond to bidders and the columns correspond to publishers, is as follows:

$$\begin{array}{ccc} & & \theta \\ & & \begin{array}{cc} 1 & 2 \end{array} \\ c & \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{array}$$

Let each publisher-ID be equally likely. In this example, the ad exchange will get zero revenue in every impression if the publisher-ID is revealed: This is because, for each of the two publishers, the second-highest valuation is zero. However, when publisher-IDs are always hidden, every impression will generate a strictly positive revenue (specifically, equal to $\frac{1}{2}$). Therefore, the policy of complete hiding is optimal.

Theorem 4. *In general, the percentage loss (relative to the optimum) from the policy of complete hiding can be arbitrarily close to 100%.*

To illustrate Theorem 4, we construct a simple example with two publishers and four bidders. In the example, each bidder works with one advertiser and each publisher-ID is equally likely. The matrix of valuations, where the rows correspond to bidders and the columns correspond to publishers, is as follows:

$$
\begin{array}{cc}
 & \theta \\
 & \begin{array}{cc} 1 & 2 \end{array} \\
c \quad \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}
\end{array}
$$

In this example, the optimal policy is complete revealing, which generates revenue equal to 1. However, the policy of complete hiding generates revenue equal to $\frac{1}{2}$. Thus, complete hiding achieves 50% of the optimum revenue.

This phenomenon can be exploited by constructing an example with k publishers and $2k$ bidders ($k \geq 2$), each working with one advertiser. The expected revenue from complete hiding is $\frac{1}{k}$, while the optimal expected revenue is 1. The result then follows by choosing an arbitrarily large value for k.

Together, our analysis in Theorems 1 through 4 leads to the following two observations. (i) Individually, both the extremes of complete hiding and complete revealing can be provably good in some situations and provably bad in some other situations. (ii) The effects of hiding and revealing are complementary to each other, in the sense that in the examples used in the proofs of Theorems 2 and 4, when one policy generates poor revenue, the other policy is optimal. Intuitively, hiding helps when the difference between the first and second highest bid is high. Otherwise, if this difference is small, revealing may be better. These observations motivate us to consider a policy that *mixes* hiding and revealing.

3.3 Mixing Hiding and Revealing

We start this section by establishing a useful fact. The expected revenue from any policy, say $\boldsymbol{q}_{\mathrm{R}}$, is at most the sum of the expected revenues from the polices of complete hiding and complete revealing. More formally, for any $\boldsymbol{q}_{\mathrm{R}}$, we have

$$
\begin{aligned}
\mathbb{E}\left[Rev^{\mathrm{I}}(\boldsymbol{q}_{\mathrm{R}})\right] &= \mathbb{E}_\theta\left[q_{\mathrm{R}}^\theta SH_c\left\{w^c(\theta)\right\}\right] + SH_c\left\{\max_a \mathbb{E}_\theta\left[\left(1-q_{\mathrm{R}}^\theta\right)v^c(\theta,a)\right]\right\} \quad \text{(from (1))} \\
&\leq \mathbb{E}_\theta\left[SH_c\left\{w^c(\theta)\right\}\right] + SH_c\left\{\max_a \mathbb{E}_\theta\left[v^c(\theta,a)\right]\right\} \qquad \text{(since } q_{\mathrm{R}}^\theta \in [0,1]\text{)} \qquad (5) \\
&= \mathbb{E}[Rev^{\mathrm{I}}(\mathbf{1})] + \mathbb{E}[Rev^{\mathrm{I}}(\mathbf{0})] \qquad \text{(from (3) and (4))}.
\end{aligned}
$$

This fact, i.e., inequality (5), along with the fact that the maximum of two numbers is larger than their average, immediately implies the first part of the following result.

Theorem 5. *The better of complete hiding and complete revealing provides at least 50% of* *the optimum revenue, i.e.,* $\max\left\{\mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{0})\right], \mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{1})\right]\right\} \geq \frac{1}{2}OPT_1$. *Furthermore, there exist instances for which this bound is tight.*

While Theorem 5 provides a guarantee on the expected revenue of the better of complete hiding and complete revealing, the exercise of determining which of these two polices is better is not easy. It requires both the knowledge of the matrix $(v^c(\theta, a))$, which we recognized earlier as being an impractical requirement, and some amount of sophisticated computation. This motivates the consideration of policies which do not require such knowledge or sophistication. The most natural class is that of uniform policies, which we now define.

Definition. A policy $\boldsymbol{q}_{\mathrm{R}}$ is a uniform policy if $q_{\mathrm{R}}^{\theta} = q$ for all $\theta \in \Theta$, for some $q \in [0, 1]$.

Observe that the complete revealing and complete hiding policies are both uniform. In fact, the better of these two policies is optimal over this class.

Theorem 6. *Either complete hiding or complete revealing is optimal over the class of uniform policies.*

The proof of this result is simple. For a uniform policy, we have $q_{\mathrm{R}}^{\theta} = q$ for every $\theta \in \Theta$. Then, the expected revenue for such a policy is

$$\begin{aligned}
\mathbb{E}[Rev^{\mathrm{I}}(\boldsymbol{q}_R)] &= \sum_{\theta} p_\theta q \mathrm{SH}_c\left\{w^c(\theta)\right\} + SH_c\left\{\max_a \sum_\theta p_\theta (1-q) v^c(\theta, a)\right\} \quad \text{(from (1))} \\
&= q\sum_\theta p_\theta SH_c\left\{w^c(\theta)\}\right) + (1-q)SH_c\left\{\max_a \sum_\theta p_\theta v^c(\theta, a)\right\} \\
&= q\mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{1})\right] + (1-q)\mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{0})\right] \quad \text{(from (3) and (4))} \qquad (6) \\
&\leq \max\left\{\mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{1})\right], \mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{0})\right]\right\}.
\end{aligned}$$

This result, however, is of limited consequence because of the above mentioned difficulty of determining which of these two policies is better. Next, we examine the performance guarantee offered by the uniform policy corresponding to $q \in (0, 1)$. We know from inequality (5) that

$$\mathrm{OPT}_1 \quad \leq \quad \mathbb{E}[Rev^{\mathrm{I}}(\mathbf{1})] + \mathbb{E}[Rev^{\mathrm{I}}(\mathbf{0})]. \qquad (7)$$

Thus, for a uniform policy with $q_{\mathrm{R}}^{\theta} = q$, where $q \in (0, 1)$, we have

$$\begin{aligned}
\mathbb{E}[Rev^{\mathrm{I}}(\boldsymbol{q}_{\mathrm{R}})] \quad &= \quad q\mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{1})\right] + (1-q)\mathbb{E}\left[Rev^{\mathrm{I}}(\mathbf{0})\right] \quad \text{(from (6))} \\
&\geq \quad \min\{q, 1-q\}\,\mathbb{E}[Rev^{\mathrm{I}}(\mathbf{1})] + \min\{q, 1-q\}\,\mathbb{E}[Rev^{\mathrm{I}}(\mathbf{0})] \\
&\geq \quad \min\{q, 1-q\}\,\mathrm{OPT}_1 \quad \text{(from (7))}. \qquad (8)
\end{aligned}$$

Inequality (8) immediately implies the following result.

Theorem 7. *A uniform policy with $q_{\mathrm{R}}^{\theta} = q$, where $q \in (0,1)$, achieves at least* $\min\{q, 1-q\}$ *of the optimal expected revenue.*

Since the choice $q = \frac{1}{2}$ maximizes the guarantee $\min\{q, 1-q\}$ in the above result, the policy with $q = \frac{1}{2}$ is of special interest. For this policy, we have

Corollary 1. *The policy of revealing each publisher-ID with probability $\frac{1}{2}$ achieves at least 50% of the optimum revenue. Furthermore, there exist instances for which this bound is tight.*

4 Analysis of Design II

Recall from Section 2.3 that, under Design II, the ad exchange chooses a probability $\boldsymbol{q}_{\mathrm{R}}$ with which to reveal the impressions from publisher θ. If θ is revealed, then bidders submit bids and ads simultaneously. If θ is hidden, then bidders submit only their bids. The publisher-ID θ is revealed to the winning bidder, who then chooses the ad for that impression. The optimization problem of the ad exchange under this design is

$$\max_{\boldsymbol{q}_{\mathrm{R}}} \mathbb{E}[Rev^{\mathrm{II}}(\boldsymbol{q}_{\mathrm{R}})] = \max_{\boldsymbol{q}_{\mathrm{R}}} \left\{ \sum_{\theta} p_{\theta} q_{\mathrm{R}}^{\theta} \mathrm{SH}_c \left\{ w^c(\theta) \right\} + \mathrm{SH}_c \left\{ \sum_{\theta} p_{\theta} q_{\mathrm{H}}^{\theta} w^c(\theta) \right\} \right\},$$

where $\boldsymbol{q}_{\mathrm{R}} \in [0,1]^{|\Theta|}$. Recall that OPT_2 denotes the optimum value of this problem. We start by comparing the performance of Designs I and II.

4.1 Performance Comparison between Design I and Design II

We show in Theorem 8 below that the optimal revenue under Design II is at least as high as that under Design I. The intuition behind this result is simple: In Design I, a bidder has to commit to the ad at bid time, when the identity of the publisher may not be known. In contrast, Design II allows the bidder to make an optimal ad choice at win time. More formally, consider the same probability vector $\boldsymbol{q}_{\mathrm{R}}$ for Designs I and II. Then, the corresponding expected revenues are as in (1) and (2). Observe that

$$\begin{aligned}
\mathbb{E}\left[Rev^{\mathrm{I}}(\boldsymbol{q}_{\mathrm{R}})\right] &= \sum_{\theta} p_{\theta} q_{\mathrm{R}}^{\theta} \mathrm{SH}_c\left(w^c(\theta)\right) + \mathrm{SH}_c[\max_a \sum_{\theta} p_{\theta} q_{\mathrm{H}}^{\theta} v^c(\theta, a)] \quad \text{(from (1))} \\
&\le \sum_{\theta} p_{\theta} q_{\mathrm{R}}^{\theta} \mathrm{SH}_c\left(w^c(\theta)\right) + \mathrm{SH}_c[\sum_{\theta} p_{\theta} q_{\mathrm{H}}^{\theta} w^c(\theta)] \quad (\text{since } v^c(\theta, a) \le w^c(\theta, a) \text{for all } a) \\
&= \mathbb{E}[Rev^{\mathrm{II}}(\boldsymbol{q}_{\mathrm{R}})] \quad \text{(from (2))},
\end{aligned}$$

Thus, we have

Theorem 8. *For any probability vector $\boldsymbol{q}_{R}$ of revealing publisher-IDs to bidders, the expected revenue under Design II is at least as high as that under Design I. Consequently, the optimal expected revenue under Design II is at least as high as that under Design I.*

The dominance of Design II over Design I, established above, can be strong in some situations. For example, let $C = \{1, 2\}$ and $\Theta = \{1, 2, 3, 4\}$. Let each bidder work with two advertisers. More specifically, let $A^1 = \{1, 2\}$ and $A^2 = \{3, 4\}$. The valuations for the two bidders are as follows:

$$\text{Valuations for Bidder 1 } (c=1) \qquad \text{Valuations for Bidder 2 } (c=2)$$

$$a \begin{matrix} 1 \\ 2 \end{matrix} \overset{\displaystyle\theta}{\overset{\begin{matrix} 1 & 2 & 3 & 4 \end{matrix}}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}} \qquad a \begin{matrix} 3 \\ 4 \end{matrix} \overset{\displaystyle\theta}{\overset{\begin{matrix} 1 & 2 & 3 & 4 \end{matrix}}{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}}$$

Let $p_\theta = \frac{1}{4}$ for all $\theta \in \Theta$. In this example, it can be shown that the optimal policy under both designs is one of complete hiding. Furthermore, the optimal expected revenues under Designs I and II can be verified to be $\frac{1}{4}$ and $\frac{1}{2}$, respectively. Thus, the ratio of the optimal expected revenue under Design II to that under Design I is 2. In fact, the conditions of this example can be aggravated to show that this ratio can exceed any arbitrary number. This establishes the following result:

Theorem 9. *The ratio of the optimal expected revenue under Design II to that under Design I can be arbitrarily large.*

The main takeaways from Theorems 8 and 9 are that (i) the expected revenue of the ad exchange in Design II improves by removing the winning bidder's uncertainty about the impression and (ii) this improvement is significant when the difference between bidders' valuations across publishers is large. These results justify the practice, in some existing ad exchanges (e.g., OpenX), of calling out the winning bidder for an ad.

4.2 Performance of Simple Policies under Design II

In this section, we analyze the performance of four simple policies: (i) the policy of complete revealing publisher-IDs, (ii) the policy of complete hiding publisher-IDs, 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-IDs with probability $\frac{1}{2}$, and (iv) the better of policies (i) and (ii). For the same reasons as those discussed at the beginning of Section 2.2, we will use the optimal expected revenue under Design II (i.e., OPT_2) as a theoretical benchmark to assess the performance of these simple policies.

We now state our results on the performance of these policies.

Theorem 10. *Under Design II, the following statements hold:*

1. *The expected revenue of the policy of revealing publisher-IDs with probability* $\frac{1}{2}$ *is at least 50% of the optimum revenue, i.e.,* $\mathbb{E}\left[Rev^{\mathrm{II}}(\tfrac{\mathbf{1}}{\mathbf{2}})\right] \geq \frac{1}{2}OPT_2$. *Furthermore, there exist instances for which this bound is tight.*
2. *The expected revenue of the better of complete revealing and complete hiding is the highest among the class of uniform policies. Meanwhile, it is at least 50% of the optimum revenue, i.e.,* $\max\left\{\mathbb{E}\left[Rev^{\mathrm{II}}(\mathbf{1})\right], \mathbb{E}\left[Rev^{\mathrm{II}}(\mathbf{0})\right]\right\} \geq \frac{1}{2}OPT_2$.
3. *Assume that* $w^c(\theta)$ *for every* $\theta \in \Theta$, $c \in C$, *are independently and identically distributed with a distribution* F_θ *with support* $[0, u_\theta]$ *for some* $u_\theta < \infty$. *Then, the policy of complete revealing is asymptotically optimal.*
4. *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.

Statement 4 above is more general than Theorem 3 under Design I: In Design II, the winning bidders are able to submit the most appropriate ad. Therefore, complete hiding remains optimal regardless of the number of advertisers each of the two bidders works with.

5 Future Research Directions

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 decision. The hide or reveal question is also relevant to explore in the context of a "chained" auction. In such an auction, some bidders act as resellers of the impression in another auction. The hide or reveal question clearly takes on a higher level of complexity in this setting. 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. One example is OpenX, that runs a modified second-price auction to account for the presence of common valuations among bidders. More detailed analysis of variants of the pure second-price mechanism may therefore need to be studied to reflect the nuances of ad auctions in real-world settings.

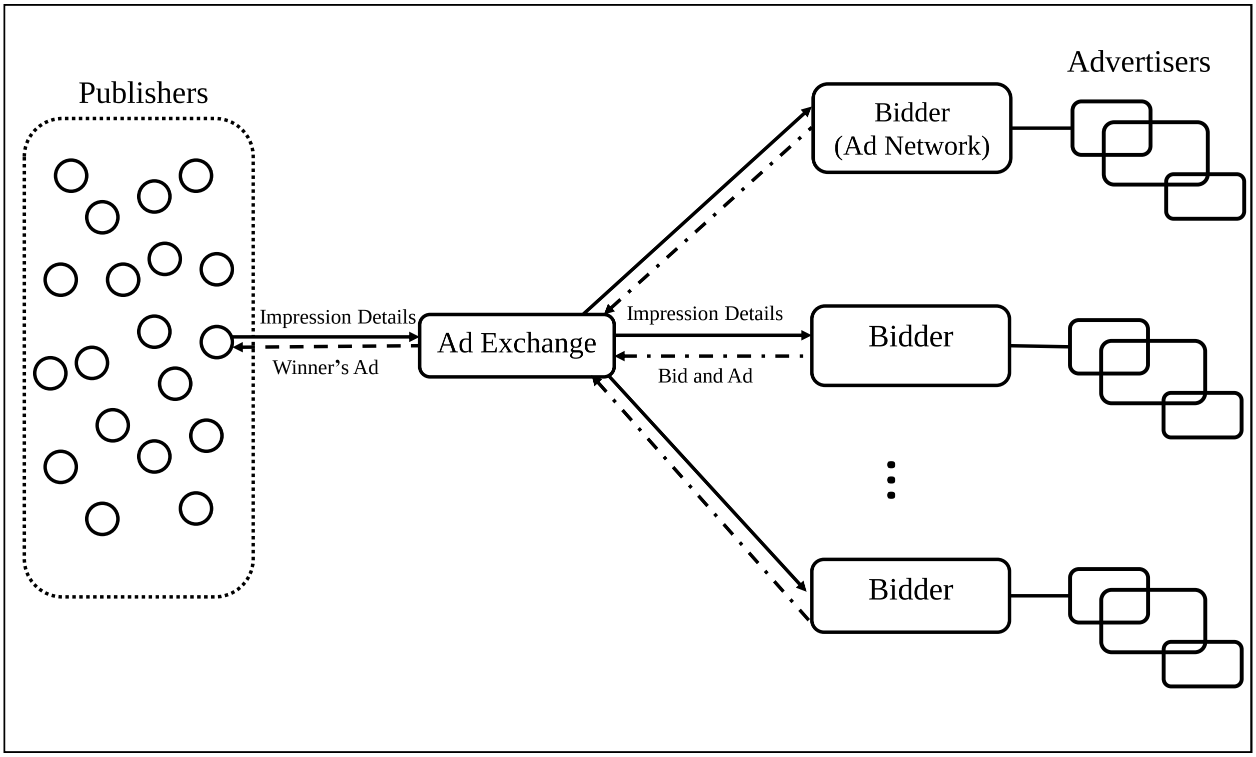

Figure 1: Real-time bidding in online ad exchange

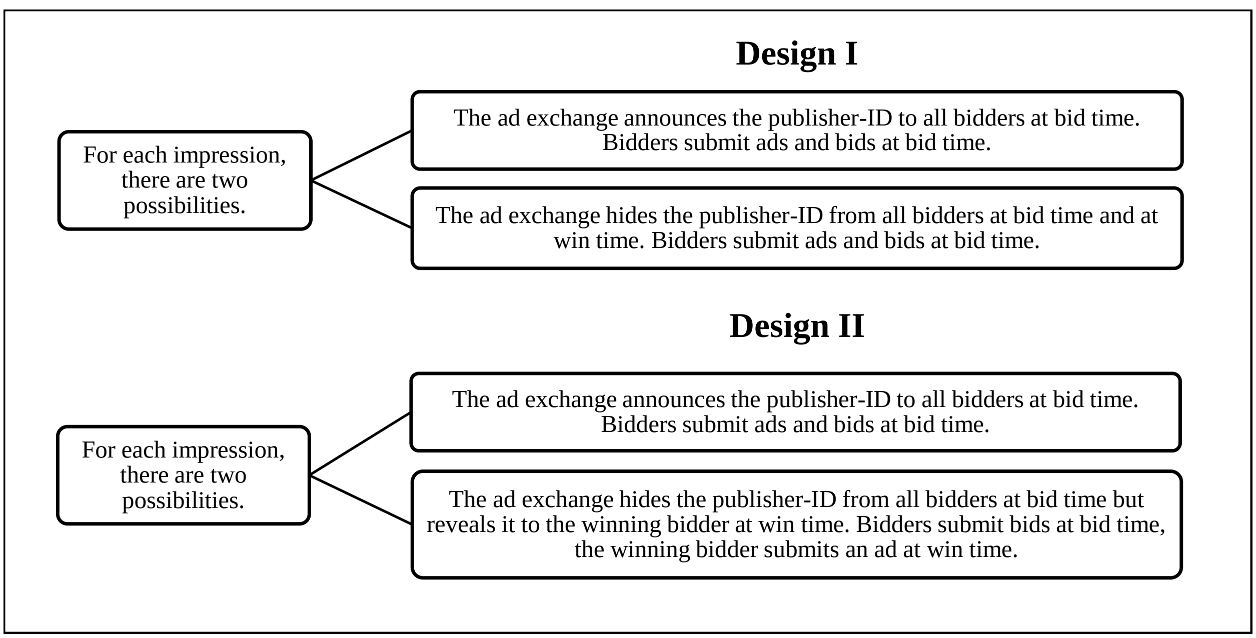

Figure 2: Description of Designs I and II (bid time refers to the time before the bids are collected; win time refers to the time after the bids are collected.)

Table 1: Basic notation

Notation	*Definition*
C	The set of bidders in the auction.
c	A bidder in the auction, $c \in C$.
A^c	The set of advertisers of bidder c.
a	An advertiser, $a \in A^c$.
Θ	The set of publisher-IDs.
θ	A publisher-ID, $\theta \in \Theta$.
$v^c(\theta, a)$	Bidder c 's valuation of advertiser a towards impressions from publisher-ID θ.
p_θ	The probability that an impression is from publisher θ.

References

Abraham, I., S. Athey, M. Babaioff, and M. Grubb. 2011. Peaches, Lemons, and Cookies: Designing Auction Markets with Dispersed Information. *Working Paper*. Harvard University, Cambridge, MA.

Balseiro, S., O. Besbes, and G. Weintraub. 2012. Auctions for Online Display Advertising Exchanges: Approximations and Design. *Working Paper*. Columbia University, New York, NY.

Bergemann, D., and M. Pesendorfer. 2007. Information Structures in Optimal Auctions. *Journal of Economic Theory* 137 (1): 580–609.

Board, S. 2009. Revealing Information in Auctions: The Allocation Effect. *Economic Theory* 38 (1): 125–135.

Cavallo, R., P. McAfee, and S. Vassilvitskii. 2012. Display Advertising Auctions with Arbitrage. In *Proceedings of the Workshop on Ad Auctions, EC-12*.

Central Market Research 2012. *Online Advertising One of the Fastest Growing Advertising Segments*. Available at http://finance.yahoo.com/news/Online-Advertising-One-iw-991038561.html. Accessed April 11, 2013.

eMarketer 2012. *US Online Advertising Spending to Surpass Print in 2012*. Available at http://www.emarketer.com/Article/US-Online-Advertising-Spending-Surpass-Print-2012/1008783. Accessed April 11, 2013.

Eső, P., and B. Szentes. 2007. Optimal Information Disclosure in Auctions and the Handicap Auction. *The Review of Economic Studies* 74 (3): 705–731.

Feldman, J., V. Mirrokni, S. Muthukrishnan, and M. M. Pai. 2010. Auctions with Intermediaries. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, 23–32.

Fu, H., P. Jordan, M. Mahdian, U. Nadav, I. Talgam-Cohen, and S. Vassilvitskii. 2012. Ad Auctions with Data. In *Algorithmic Game Theory*, 168–179. Springer.

Ganuza, J. 2007. Ignorance Promotes Competition: An Auction Model of Endogenous Private Valuations. *RAND Journal of Economics* 35 (3): 583–598.

Ganuza, J., and J. S. Penalva. 2010. Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions. *Econometrica* 78 (3): 1007–1030.

Hof, R. 2011. *Online Ad Spend to Overtake TV by 2016*. Available at http://www.forbes.com/sites/roberthof/2011/08/26/online-ad-spend-to-overtake-tv/. Accessed April 11, 2013.

Interactive Advertising Bureau 2012. *Internet Advertising Revenue Report*. Available at http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_Report_FY_2012_rev.pdf. Accessed April 17, 2013.

Lewis, T. R., and D. Sappington. 1994. Supplying Information to Facilitate Price Discrimination. *International Economic Review* 35 (2): 309–327.

Lieberman, D. 2013. *TV Advertising Is "Surprisingly Weak" Due To Internet And Economy: Analyst*. Available at http://www.deadline.com/2013/03/tv-advertising-weakness-economy-internet-analyst-report/. Accessed April 11, 2013.

Milgrom, P. R., and R. J. Weber. 1982. A Theory of Auctions and Competitive Bidding. *Econometrica* 50 (5): 1089–1122.

Muthukrishnan, S. 2009. Ad Exchanges: Research Issues. In *Internet and Network Economics*, 1–12. Springer.

Ottaviani, M., and A. Prat. 2003. The Value of Public Information in Monopoly. *Econometrica* 69 (6): 1673–1683.

Sweney, M. 2012. *UK Web Advertising Spend to Hit £5bn.* Available at http://www.guardian.co.uk/media/2012/apr/03/uk-web-advertising-spend. Accessed April 11, 2013.