

Selling Impressions: Efficiency vs. Competition*

Dirk Bergemann[†] Tibor Heumann[‡] Stephen Morris[§]
Constantine Sorokin[¶] Eyal Winter^{||}

August 19, 2021

Abstract

In digital advertising, a publisher selling impressions faces a trade-off in deciding how precisely to match advertisers with viewers. A more precise match generates efficiency gains that the publisher can hope to exploit. A coarser match will generate a thicker market and thus more competition. The publisher can control the precision of the match by controlling the amount of information that advertisers have about viewers. We characterize the optimal trade-off when impressions are sold by auction. The publisher pools premium matches for advertisers (when there will be less competition on average) but gives advertisers full information about lower quality matches.

JEL CLASSIFICATION: D44, D47, D83, D84.

KEYWORDS: Second Price Auction, Conflation, Targeted Advertising, Impressions, Two-Sided Private Information, Bayesian Persuasion, Information Design.

*We acknowledge financial support from NSF grants SES-2001208, SES-1948336 and SES-2049744, and ANID Fondecyt Iniciacion 11200365. We thank Steve Berry, Ozan Candogan, Roberto Corrao, Phil Haile, Sergiu Hart, Jason Hartline, Nicole Immorlica, Maxim Ivanov, Anton Kolotilin, Renato Paes Leme, Denis Shishkin, Rann Smorodinsky, Philipp Strack, Alex Wolitzky and Song Zuo for informative discussions. We thank seminar audiences at Informs AMD, UCSD, PUC Chile, the Hebrew University's Center for the Study of Rationality, the China Meetings of the Econometric Society and the Stony Brook International Conference on Game Theory for many helpful comments. This manuscript combines and subsumes the results in Bergemann, Heumann and Morris (2021) and Sorokin and Winter (2021). These two papers shared the current Theorem 1, obtained via different proof strategies.

[†]Yale University, dirk.bergemann@yale.edu.

[‡]Pontificia Universidad Católica de Chile, tiber.heumann@uc.cl.

[§]Massachusetts Institute of Technology, semorris@mit.edu

[¶]Glasgow University and Higher School of Economics, constantine.sorokin@glasgow.ac.uk

^{||}The Hebrew University of Jerusalem; and Lancaster University, eyal.winter@mail.huji.ac.il

1 Introduction

In the internet advertising market, it has become technologically feasible to match advertisers to viewers with ever-greater precision. But do publishers selling impressions have an incentive to do so? Finer matching generates efficiency gains that sellers can hope to exploit. But coarser matching generates market thickness, and so more market competition and less information rent for advertisers. Impressions are typically sold by auction and publishers can control the precision of the match by controlling the information that bidders have access to. We characterize the trade-off between efficiency and competition for the publisher.

We do this in two steps. First, we characterize what information a seller would choose to give buyers about their values in a classic auction in a standard independent private values setting. Second, we provide a model of the market for impressions in digital advertising markets, and show how our characterization applies in this setting. We now discuss these two steps in turn.

Consider a second price auction where bidders' valuations are independently and symmetrically distributed, but initially unknown to the bidders. The seller can choose what information each bidder can learn about their own value. If the seller did not allow them to learn anything, then all bidders would bid their (common) expected value and the good would be randomly (and inefficiently) allocated among them. If the seller allowed bidders to learn their true value, then they would have a dominant strategy (under private values in a second price auction) to bid their values. The good would be allocated efficiently to the bidder with the highest value. The revenue of the seller would equal the value of the efficient allocation minus the bidders' information rent. By permitting bidders to learn something but not everything about their values, the seller can trade off efficiency loss with information rent reduction. Our main result is a characterization of the optimal (among symmetric) information policies for the seller.

Conditional on having a low value, a bidder is likely to be competing with other bidders and earn low information rents. But conditional on having a high value, a bidder is likely to win (facing weak competition) and thus can expect to win at a price significantly below his value, thus earning high information rents. Thus the gains from concealing information will be highest when valuations are high. In the optimal policy, high values are pooled and low values are revealed. There is a critical threshold described by a quantile above which all valuations are bundled together (Theorem 1). The threshold is given by a quantile of the distribution that depends only on the number of bidders and not the distribution of valuations. The optimal quantile until which disclosure occurs is increasing in the number of participating bidders and goes towards 1 (i.e., full disclosure) as the number of

bidders grows arbitrarily large. Thus, the information policy is influencing the distribution of bids, holding fixed the distribution of preferences among the bidders.

The optimal cutoff is designed to keep a moderate level of competition at the top of the bid distribution. If the cutoff is too low, more bidder surplus will be extracted but the expected value conditional on being above the cutoff will be too low. If the cutoff is too high, the expected value conditional on being on being above the cutoff will be high but too little bidder surplus will be extracted. The quantile keeps “around” two bidders at the top of the bid distribution – i.e., with the same expected value and competing with each other. This gives just enough competition to extract the bidder surplus as low efficiency cost. The gains from the optimal information structure will depend on the distribution of values, ranging from no gains to a revenue that can be arbitrarily larger than the expected revenue generated under complete information. As the number of bidders grows large, the gains will remain substantive as long as the distribution of values has fat tails.

Our paper’s motivation is the market for impressions in digital advertising. A large share of digital advertising, whether in search, display advertising or social networks, is allocated by auction mechanisms. The second price auction is commonly used in digital advertising to form a match between competing advertisers (the *bidders*) and a viewer. A match between viewer and advertiser creates an *impression* (or search result) on the publisher’s website. The seller (a *publisher* or *intermediary*) uses an auction platform to sell the attention (“eyeball”) of the viewer to competing advertisers. The viewer is thus the object of the auction. The viewers are typically heterogeneous in many attributes: their demographic characteristics, their preferences, their (past) shopping behavior,; their browsing history and other aspects, observable and unobservable. The advertisers therefore display a corresponding degree of heterogeneity in their willingness to pay for a match between their advertisement and a specific viewer. The private (and the social) value of any particular match is then determined jointly by the viewer’s attributes and the advertiser’s preferences for those attributes. In the presence of this heterogeneity on both sides of the match, viewer and advertiser, internet advertising has moved towards targeted advertising to join the information. The auction can therefore support highly targeted advertising that may increase the social efficiency in the match formation between viewer and advertiser. But - as discussed earlier - allowing for finely targeted bidding may also thin the market among the advertisers, and hence reduce the competition between advertisers. Publishers distinguish two schemes, or algorithms, for mapping preferences and attributes into bids, *automated bidding* and *manual bidding*. In automated bidding, *autobidding* for short, the seller offers a bidding algorithm that generates optimal bids for

the advertisers given the disclosed information.¹ In manual bidding, the seller offers a disclosure algorithm that generates information about the attributes, and in particular a bid recommendation, which each bidder then manually adopts or modifies into a bid for the impression, also referred to as dashboard mechanism in Hartline et al. (2019). Autobidding has become increasingly prevalent in digital advertising to convert the high-dimensional information across millions of impressions into bids with minimal latency, see Aggarwal et al. (2019) and Deng et al. (2021).

In this market, publishers can control the information that advertisers have about their values. While there are many reasons why advertisers' values might be correlated, they will not be if variation in viewers' attributes is horizontal. A second contribution of the paper is to develop a stylized model of the market for impressions, establish that our earlier model applies to this market, and show how the results apply under reasonable assumptions on the market.

In our model of this market, a viewer is characterized by a (perhaps high-dimensional) attribute. The publisher (but not the advertisers) knows the attribute of the viewer. The advertiser (but not the publisher) knows the characteristics of the good being advertised and hence has preference over viewer attributes. The viewer attribute can be combined with an advertiser's preference to generate a match quality, and the value of the viewer to the advertiser is an increasing function of the match quality. A key feature of this model is that the advertiser's private information is not informative about his value of a viewer unless it is combined with information held by the publisher. The publisher sells impressions (the allocation of the viewer to an advertiser) in a second price auction. Because only the publisher knows the attributes of the viewer, the publisher can control the information that advertisers have about their match quality with the viewer.

We present two results in our model. We first show that this model of two-sided information gives rise to the setting of our main result: an independent private values setting, where the advertiser has no information about his value of the viewer, but the combination of the publisher's information about attributes and advertiser's information about preferences fully reveals the state. Now we can identify the optimal information structure from our earlier analysis. But how can this be implemented in practise? We consider *autobidding*. The advertiser reports his preferences to the publisher and the publisher commits to bid optimally for the bidder as a function of the optimal information structure. We show that the bidder will have an incentive to truthfully report his preferences. Thus autobidding implements the optimal outcome in this market.

¹See Google Ads Help Center (2021a), (2021b), or Facebook Business Help Center (2021) for summary descriptions of automated bidding mechanisms.

Literature Levin & Milgrom (2010) suggested that the idea of conflation (central in many commodity markets) by which similar but distinct products are treated as identical in order to make markets thicker or reduce cherry-picking, may be relevant for the design on online advertising markets. The information structure (4) we derive determines exactly when conflation should occur, in the upper interval, and when not, in the lower interval. Ghosh et al. (2009) and Emek et al. (2012) investigate the computational complexity of the determining the optimal information structure when there many preference profiles and many attribute profiles that might be sold in an auction.

The paper relates to the literature studying optimal information disclosure in selling mechanisms. Ganuza (2004) studies the optimal information disclosure in a second-price auction, where bidder's valuation are determined by the quality of a match between a bidder's taste and the good's characteristic represented by Hotelling model on a circle. The seller chooses the optimal public signal about the good's characteristic. He shows that the equilibrium information provision is less than the surplus-maximizing one. His result is about *costly* public signals and so it does not address the trade-off between efficiency and market thickness that is central in our paper.

Bergemann & Pesendorfer (2007) analyze the joint optimal design of auction and information structure. In particular, they allow for asymmetric information structures and personalized reserve prices. Here, we fix the selling mechanism to be a second-price auction as it is a better fit for the markets for selling impressions.

2 Model

There are N agents who bid for an indivisible good in an auction. Bidder i 's valuation is denoted by $v_i \in \mathbb{R}_+$. We assume that the valuations are independently and identically distributed across agents according to an absolutely continuous distribution, denoted by F . The assumptions that F is absolutely continuous helps simplify some of the expressions but all results go through unchanged if we relax this assumption.

The seller can choose how much information each bidder will have about his own valuation. An information structure is denoted by:

$$s_i : \mathbb{R}_+ \rightarrow \Delta\mathbb{R}_+,$$

where $s_i(v_i)$ is the signal observed by bidder i when his valuation is v_i . After observing s_i , the

bidder forms his beliefs about his valuation. An agent's expected valuation is denoted by:

$$w_i \triangleq \mathbb{E}[v_i \mid s_i].$$

We denote by G_i the distribution of expected valuations. Note that we are making two assumptions about the information structure. First, each bidder only observes information about his own valuation, which is reflected by the fact that s_i takes as an argument v_i only (instead of (v_1, \dots, v_N)). Additionally, there is no common source of randomization in the signals. Hence, the signals will be independently distributed across agents. Finally, we assume that the seller is restricted to symmetric information structures, i.e., $s_i(\cdot) = s_j(\cdot)$.

The objective of the seller is to maximize revenue. Since agents are bidding in a second-price auction it is a dominant strategy to bid their expected valuation.² Hence, revenue is equal to the second-highest expected valuation across bidders. We denote the k -th highest valuation by $w_{(k)}$. The objective of the seller is to solve:

$$R \triangleq \max_{\{s: \mathbb{R} \rightarrow \Delta \mathbb{R}\}} \mathbb{E}[w_{(2)}]. \quad (1)$$

3 Optimal Information Structure

Since the expected revenue is equal to the expectation of second-highest valuation, the distribution of expected valuations generated by the signal is a sufficient statistic to compute the seller's expected revenue. Hence, instead of studying explicitly the signal chosen by the seller, we frequently refer to the distribution of expected valuations generated by a signal (recall that this is denoted by G).

The second-order statistic of N symmetrically and independently distributed random variables is distributed according to

$$\Pr(w_{(2)} \leq t) = NG^{N-1}(t)(1 - G(t)) + G^N(t).$$

The expected revenue of the auctioneer is therefore:

$$\mathbb{E}[w_{(2)}] = \int_0^\infty t d(NG^{N-1}(t)(1 - G(t)) + G^N(t)).$$

We now characterize the set of feasible distributions G .

²See Krishna (2009) for a textbook discussion.

By Blackwell (1951), Theorem 5, there exists a signal s that induces a distribution of expected valuations if and only if F is a mean preserving spread of G . F is defined to be a mean preserving spread of G if

$$\int_v^\infty F(t)dt \leq \int_v^\infty G(t)dt, \forall v \in \mathbb{R}_+,$$

with equality for $v = 0$. If F is a mean preserving spread of G we write $F \prec G$.

We can now express the seller's problem as finding an optimization over a distribution G subject to a mean-preserving restriction. The choice of the optimal information structure can be written as the following maximization problem:

$$R = \max_G \int_0^\infty td(NG^{N-1}(t)(1 - G(t)) + G^N(t)) \quad (2)$$

subject to $F \prec G$.

This problem consists of maximizing over feasible distributions of expected valuations. However, the objective function is non-linear in the probability (or density) of the optimization variable G . Moreover, the non-linearity cannot be confined to be either concave or convex on G .

The key step in our argument comes from a change of variables, re-writing the above in terms of the quantile q of the second order statistic. We denote by $S_N(q)$ the cumulative distribution function of the quantile of the second-highest valuation:

$$S_N(q) \triangleq \Pr(G(w_{(2)}) \leq q).$$

We index by N to highlight the dependence on the number of buyers. We observe that $S_N(q)$ is given by:

$$S_N(q) = Nq^{N-1}(1 - q) + q^N.$$

The quantile distribution S_N is independent of the underlying distribution F . Just as the quantile of any random variable is uniformly distributed, the quantile of the second-order statistic of N symmetric independent random variables is distributed according to S_N for any underlying distribution. Hence, the revenue can be computed by taking the expectation over quantiles using measure $S_N(q)$: the revenue given the quantile of the second-order statistic is G^{-1} . So maximization problem (2) can be transformed into:

$$\max_{G^{-1}} \int_0^1 S'_N(q)G^{-1}(q)dq \quad (3)$$

subject to $G^{-1} \prec F^{-1}$

The corresponding constraint states that the seller can choose any distribution of expected valuations whose quantile function G^{-1} is a mean-preserving spread of the quantile function F^{-1} of the initial distribution of valuations. This uses a well-known property of the distribution function, see Shaked & Shanthikumar (2007), Chapter 3, stating that $F \prec G$ if and only if $G^{-1} \prec F^{-1}$. Hence, we have a linear (in G^{-1}) maximization problem subject to a majorization constraint, which will allow us to solve the problem with known methods.

Theorem 1 (Optimal Information Structure)

1. *The unique optimal symmetric information structure is given by:*

$$s(v_i) = \begin{cases} v_j & \text{if } F(v_i) < q_N^* \\ \mathbb{E}[v_j \mid F(v_j) \geq q] & \text{if } F(v_i) \geq q_N^* \end{cases} \quad (4)$$

where the critical quantile $q_N^* \in [0, 1)$ is independent of F .

2. *The critical quantile satisfies $q_2^* = 0$; q_N^* is increasing in N ; $q_N^* \rightarrow 1$ as $N \rightarrow \infty$; and for each $N \geq 3$, q_N^* is the unique solution in $(0, 1)$ to the following polynomial equation of degree N :*

$$S'_N(q)(1 - q) = 1 - S_N(q). \quad (5)$$

Thus, the optimal information structure is to reveal the valuation of all those bidders who have a valuation lower than some threshold determined by a fixed quantile q_N^* and otherwise reveal no information beyond the fact that the valuation is above the threshold. The threshold in terms of the valuation is given by $F^{-1}(q_N^*)$, but the quantile q_N^* is independent of the distribution F of valuations.

The optimal information structures thus supports more competition at the top of the distribution at the expense of an efficient allocation. The information structure fails to distinguish in the allocation between any two valuations that are in the upper tail of the distribution $[F^{-1}(q_N^*), \infty)$. The benefit accrues through more competitive bids among the high value bidders. Namely, if the second highest bid is in the interval, then its competitive bid matches exactly the bid of the winning bid, and thus the information rent of the winning bidder is depressed considerably with a corresponding gain in the revenue for the seller.

Before we prove the result, we will provide some intuition for the critical quantile by confirming that it must be given by equation (5) as long as information takes the form given in the proposition.

Suppose that we fix a quantile threshold q and write $v = F^{-1}(q)$ for the corresponding value. Now what happens to revenue if we decrease the threshold by dq ?

With probability $S'_N(q) dq$ the second-highest bid was not in the pooling zone before the decrease and is after the decrease; and revenue increases by $(\mathbb{E}[v_i | v_i \geq v] - v)$. With probability $1 - S_N(q)$, the second-highest bid was in the pooling zone before the decrease and there is a loss of revenue of

$$\frac{d\mathbb{E}[v_i | v_i \geq v]}{dq} = \frac{1}{1 - q}(\mathbb{E}[v_i | v_i \geq v] - v).$$

Hence, the revenue gain is proportional to $S'_N(q)$ and revenue loss is proportional to $(1 - S_N(q))/(1 - q)$. Equating gain and loss we get (5) (the term $(\mathbb{E}[v_i | v_i \geq v] - v)dq$ appears on both sides so it cancels out). Furthermore, among the class of upper-pooling information structure, revenue is quasi-concave and single-peaked in the cutoff q (cf. Sorokin & Winter (2021)).

To prove Theorem 1, we state a result of Kleiner et al. (2021) in terms of our maximization problem (3).³

Proposition 1 (Kleiner et al. (2021), Proposition 2)

Let G^{-1} be such that for some countable collection of intervals $\{[\underline{x}_i, \bar{x}_i] \mid i \in I\}$,

$$G^{-1}(q) = \begin{cases} F^{-1}(q) & q \notin \cup_{i \in I} [\underline{x}_i, \bar{x}_i] \\ \frac{\int_{\underline{x}}^{\bar{x}} F^{-1}(t) dt}{\bar{x}_i - \underline{x}_i} & q \in [\underline{x}_i, \bar{x}_i] \end{cases}$$

If $\text{conv}S_N$ is affine on $[\underline{x}_i, \bar{x}_i]$ for each $i \in I$ and if $\text{conv}S_N = S_N$ otherwise, then G^{-1} solves problem (3). Moreover, if F^{-1} is strictly increasing the converse holds.

Here, $\text{conv}S_N$ is the convexification of S_N , i.e., the largest convex function that is smaller than S_N . With this result we can prove our main result.

Proof of Theorem 1. The second derivative of the distribution S_N of the quantile of the second order statistic is given by:

$$S''_N(q) = q^{N-3}(N-1)N(N-2-q(N-1)).$$

Hence, $S_N(q)$ is concave if and only if

$$q \geq (N-2)/(N-1),$$

³The proof of Theorem 1 in Sorokin & Winter (2021) was self-contained and did not refer to the verification result of Kleiner et al. (2021).

and convex otherwise. Thus, the convex hull of S_N for $N \geq 3$ is:

$$\text{conv}S_N(q) = \begin{cases} S_N(q) & \text{if } q \leq q_N^*; \\ S'_N(q_N^*)(q - q_N^*) + S(q_N^*) & \text{otherwise.} \end{cases}$$

where q_N^* is defined as in (5) for $N \geq 3$. For $N = 2$, we have $\text{conv}S_2(q) = q$ and define $q_2^* = 0$. Now let G^{-1} be given by:

$$G^{-1}(q) = \begin{cases} F^{-1}(q), & q < q_N^*; \\ \frac{\int_{q_N^*}^1 F^{-1}(t)dt}{1 - q_N^*}, & q \in [q_N^*, 1); \\ 1, & q = 1. \end{cases} \quad (6)$$

Then, G^{-1} satisfies all the assumptions of Proposition 1, so it is the unique optimal solution to (3). For all valuations below $F^{-1}(q_N^*)$ the distribution over expected valuations is the same as that of the real valuations. Hence, types below $F^{-1}(q_N^*)$ know their own values. On the other hand, for valuations above $F^{-1}(q_N^*)$ the distribution over expected valuations is a mass point at the expected valuation conditional on being above $F^{-1}(q_N^*)$. Hence it is clear that this distribution is induced by information structure (4).

To check that q_N^* is strictly increasing in N we define:

$$\psi(q, N) \triangleq S'_N(q)(1 - q) - (1 - S_N(q)).$$

By definition, $\psi(q_N^*, N) = 0$. We now note that:

$$\psi(q, N + 1) - \psi(q, N) = N(q - 1)^2(N(q - 1) + 1)q^{N-2}.$$

so $\psi(q, N + 1) - \psi(q, N) \geq 0$ if and only if $q \geq (N - 1)/N$. As previously argued, $q_N^* < (N - 2)/(N - 1)$ so $q_N^* < (N - 1)/N$, which implies that:

$$\psi(q_N^*, N + 1) < 0. \quad (7)$$

We also have that $\psi(0, N) = -1$ and $\psi(1 - \varepsilon, N) > 0$ for ε small enough, where the last part can be verified by noting that

$$\psi(1, N) = \frac{\partial \psi(1, N)}{\partial q} = 0 \quad \text{and} \quad \frac{\partial^2 \psi(1, N)}{\partial q^2} = N(N - 1) > 0.$$

As previously argued $\psi(q, N + 1)$ has a unique root in $(0, 1)$, so (7) implies that $q_N^* < q_{N+1}^*$.

Finally, if N diverges to infinity and $\lim_{N \rightarrow \infty} q_N^* < 1$, then in the limit we would have that $S_N(q_N^*), S'_N(q_N^*) \rightarrow 0$. So (5) would not be satisfied. We thus must have that $\lim_{N \rightarrow \infty} q_N^* = 1$. ■

It is useful to compare our problem to a Bayesian persuasion model where the objective function permits a nonlinear evaluation $u(x)$ of an outcome but linear in the probability, thus

$$\begin{aligned} & \max_G \int_0^1 u(t) dG(t), \\ & \text{subject to } F \prec G, \end{aligned}$$

as, for example in Dworzak & Martini (2019). Our original maximization problem (2) did not take this form as it was non-linear in probabilities. However, we reformulated the problem to one that is linear in the new optimization variable G^{-1} , changing the direction of the constraint.⁴ For this problem, the convexification of S_N was key to identifying the optimal information structure. The information structure (4) that emerges here is sometimes referred to as "upper censorship" in the Bayesian persuasion literature, as it pools all the states above a cutoff and reveals all the states below the cutoff (Alonso & Camara (2016); Kolotilin et al. (2021)).

Competition at the Top We can gain further intuition about the optimal policy, by examining the degree of competition at the top of the bid distribution. The number of bidders that have the highest expected valuation follows a binomial distribution $B(N, 1 - q_N^*)$. The expected number of bidders at the top of the bid distribution is then given by:⁵

$$\rho_N^* \triangleq N(1 - q_N^*).$$

Now $\rho_2^* = 2$, since $q_2^* = 0$. To characterize ρ_N^* for $N \geq 3$, we can substitute $q_N^* = 1 - \frac{\rho_N^*}{N}$ into (5), to get an N th degree polynomial characterizing ρ_N^* :

$$\frac{\rho_N^*}{N} S'_N \left(1 - \frac{\rho_N^*}{N} \right) = 1 - S_N \left(1 - \frac{\rho_N^*}{N} \right). \quad (8)$$

One can numerically verify that ρ_N^* is decreasing in $N \geq 3$, with $\rho_3^* = 2.25$ and $\rho_N^* \downarrow \rho_\infty^* \approx 1.793$ as $N \rightarrow \infty$. The limit value ρ_∞^* can be solved analytically, as equation (8) reduces to

$$(\rho_\infty^*)^2 + \rho_\infty^* + 1 = e^{\rho_\infty^*}$$

⁴Relatedly, Proposition 12 in Kolotilin & Wolitsky (2020) solves a problem in gerrymandering that is isomorphic to a Bayesian persuasion problem that is non-linear in the probabilities. A linearization of the problem is implicit in their proof.

⁵This is the expression for the mean of any Binomial distribution (see, for example, Papoulis & Saunders (1989)).

as $N \rightarrow \infty$, whose solution is $\rho_\infty^* \approx 1.793$.

We can also say more about the distribution of the number of bidders at the top of the distribution as $N \rightarrow \infty$: the binomial distribution $B(N, 1 - q_N^*)$ converges to a Poisson distribution with parameter ρ_∞^* , by the Law of Rare Events or the Poisson Limit Theorem (see Papoulis & Saunders (1989)).

The bottom line is that the optimal information policy always keeps “about” 2 bidders above the threshold, which is the key to extracting bidder surplus.

Gains from the Optimal Information Policy We now consider what we can say about the revenue gain from the optimal information policy. So let us define the revenue gain to be the ratio of revenue under the optimal policy to revenue under complete information. We will report three examples illustrating the magnitude and variation in revenue gains.

First, consider the case where the distribution of values has a binary support.⁶ This example will illustrate the wide variation in revenue gain that can arise. In particular, suppose the distribution of values has support $\{0, 1/(2(1 - p))\}$, with $\Pr(v_i = 0) = p$. So the mean value is $1/2$ for every p .

When $q_N^* < p$, the optimal information structure is the same as complete information, so there is no benefit from hiding information from the bidders. When $q_N^* > p$, the optimal information structure generates expected valuations with support in $\{0, 1/(2(1 - q_N^*))\}$, with $\Pr(w_i = 0) = q_N^*$. In the limit $p \rightarrow 1$, complete information generates zero revenue while the optimal information generates positive revenue, so the optimal information structure generates infinitely times more revenue than complete information. This is because the probability of two bidders having high valuation in the limit $p \rightarrow 1$ converges to zero, while the optimal information structure keeps a moderate level competition at the top of the bid distribution. Intuitively, when the distribution is heavily skewed, the benefits from manipulating bidders’ information is larger.

Second, consider the case where values are uniformly distribution on $[0, 1]$. This example is a leading benchmark in analysis of auctions. Sorokin & Winter (2021) report detailed numerical analysis of this case as N varies. For example, if $N = 3$ and F is uniform on $[0, 1]$, the revenue gain is approximately 1.1. Given the bounded support of values, it is necessarily the case that as $N \rightarrow \infty$, the revenue gain goes to zero simply because the complete-information revenue converges to the highest value on the support.

Third, consider the case where the distribution of values has fat upper tails. This example will

⁶We stated our main result for continuous distributions of values, but the form of the optimal information structure remains the same (with only minor modifications to describe the optimal signal (4)).

illustrate the possibility of large revenue gain even as the number of bidders becomes large. Arnosti et al. (2016), argues that the Pareto distribution, with its fat tail provides a good fit for demand in the digital advertising market. This case of fat upper tails is analyzed in detail in Bergemann et al. (2021). It is shown that as the number of bidders grows, the revenue gain converges to a constant that depends on the fatness of the tail. In the case of a Pareto tail with distribution $F(v) = 1 - v^{-\alpha}$, the limiting revenue gain (as $N \rightarrow \infty$) is above 1 for $\alpha > 1$ and converges to ∞ as $\alpha \downarrow 1$. The intuition is that despite there being many bidders, there is a high probability of the highest valuation significantly exceeding other valuations. Formally, when the distribution has fat tails:

$$\mathbb{E}[v_{(1)}] - \mathbb{E}[v_{(2)}] \rightarrow \infty, \text{ as } N \rightarrow \infty, \quad (9)$$

so under complete information bidder surplus is not competed away as the number of bidders grows. By contrast, the optimal information structure generates more competition at the tail of the bid distribution and thus reduces bidder surplus substantially.

4 Market for Impressions

We now translate the earlier results into a market for impressions with two-sided information. This translation recasts the optimal information design as bidding mechanisms in the world of digital advertising. The choice of the optimal information structure can then be interpreted in terms of the information policy of the publisher who matches the viewer with the advertisers.

The viewer has attribute $x \in X$ distributed according to F_x . Each advertiser i is represented by a vector $y_i \in Y$, distributed according to F_y , identically and independently distributed across advertisers. An impression is a match between an advertiser and a viewer. The value v_i of advertiser i from attracting a viewer is determined by a function u :

$$u : X \times Y \rightarrow \mathbb{R}_+,$$

such that:

$$v_i \triangleq u(x, y),$$

and we refer to u as the *valuation function*. Since advertiser i 's value for every viewer is determined by y_i , we refer to the vector y_i as the advertiser's *preferences* (with a slight abuse of terminology). The distribution of attributes and preferences (x, y) and the valuation function u induce a distribution of the bidder i 's value v_i , denoted by F .

We assume that the unconditional distribution of values (v_1, \dots, v_N) generated by (x, y) are independent across bidders. We further assume that the unconditional distribution of values are the same as the conditional distribution of values (v_1, \dots, v_N) conditional on either x or y . That is, (x, v_1, \dots, v_N) and (y, v_1, \dots, v_N) are random vectors consisting of independently distributed random variables. Of course, (x, y_i, v_i) are *not* independently distributed. In other words, the preferences y_i provide only information about the value v_i when combined with information about the attributes x of the viewer. Moreover, each advertiser i 's belief about the valuation of their competitors is unaffected by the attributes x or the preferences y_i . Thus, the present model of attributes and preferences generates a model of independent private values.

These assumptions implicitly impose restrictions on the set of valuation functions and distributions of attributes and preferences that we consider. In other words, not every valuation function u and distributions F_x, F_y will generate joint distributions of values that satisfy these assumptions. The attributes of viewer therefore reflect an aspect of horizontal differentiation with value implications that depend on the preferences of the advertiser. Conversely, an aspect of vertical differentiation that has similar implications across all preferences would fail the equivalence between unconditional and conditional value distribution. We now briefly describe two classes of models that satisfy the above conditions, one high-dimensional and one low-dimensional set of models.

Our leading example is given by the following specification. The set of possible attributes and preferences are given by $X = Y = \{-1, 1\}^J$. Each component of the attributes and preferences is distributed uniformly and independently across components. The valuation function is given by:

$$u(x, y_i) = u\left(\frac{1}{\sqrt{J}} \sum_{j=1}^J x_j y_{ij}\right), \quad (10)$$

for some strictly increasing function u . Hence, an advertiser's valuation is increasing in the number of components in which the attributes and preferences are aligned.

A second class of models is given by the following Hotelling location model that has only a one-dimensional space of uncertainty. Suppose that $X = Y = [0, 1]$, where x, y_i are positions in a circle of perimeter 1 uniformly and independently distributed. Let $d(x, y_i)$ be the shortest distance between x and y_i on the circle and

$$u(x, y_i) = u(d(x, y_i)) \quad (11)$$

for some strictly decreasing function u .

Both of these classes of models satisfy the independence conditions above. Of course, independence is (as always) a knife-edge assumption, so we require strong assumptions on the distributions

of preferences and attributes to generate independent values.⁷ However, these assumptions have a clear economic interpretation, namely, we are modeling a two-sided market with (and only with) horizontal differentiation.

We now analyze automated bidding in the second-price auction based on a signal of the value of the impression.⁸ Thus, the publisher commits: (i) to complement the advertiser’s information with a signal regarding the match quality; and (ii) to set the advertiser-optimal bid. In turn, the advertiser submits his preferences y (and thus a description of the attributes he cares about). The central aspect of automated bidding is that the publisher complements the advertiser’s private information y with information about the viewer’s attribute x that is unknown to the advertiser.

Formally, the publisher chooses a signal $s_i : \mathbb{R} \rightarrow \Delta\mathbb{R}$ as a function of the advertiser’s reported value $v_i(x, y_i)$. While the publisher cannot directly observe v_i , she elicits the advertiser’s preferences and knows the viewer attributes, so she can infer v_i . The publisher submits a bid $b_i : Y \times \mathbb{R} \rightarrow \mathbb{R}$ satisfying that:

$$b_i(y_i, s_i) = \mathbb{E}[v_i \mid y_i, s_i(v_i)] \tag{12}$$

That is, the publisher submits a bid on behalf of advertiser i equal to the advertiser’s expected valuation given his preferences y_i and the additional information s_i provided by the publisher. This is an advertiser’s dominant strategy given the available information $(y_i, s(v_i))$. Because (v_i, y_i) are independently distributed, we have that:

$$\mathbb{E}[v_i \mid y_i, s_i(v_i)] = \mathbb{E}[v_i \mid s_i(v_i)].$$

Hence, the publisher’s problem reduces to find an information structure s that solves:

$$R \triangleq \max_{s: \mathbb{R} \rightarrow \Delta\mathbb{R}} \mathbb{E}[b_{(2)}],$$

which is the original problem stated earlier in (1).

We now verify that under the automated bidding in the second price auction, it is optimal for each advertiser to truthfully report their preferences to the publisher. A reporting strategy for bidder i is denoted by:

$$\tilde{y}_i : Y \rightarrow \Delta Y.$$

⁷Finding other models that generate independent values is an interesting direction for future work.

⁸We discussed two alternative bidding procedures - automated bidding and manual bidding - in the introduction, and briefly return to discuss manual bidding in our Section 5 discussion.

Given the reported preferences, the seller discloses to the bidder a signal $s(\tilde{v}_i)$, where \tilde{v}_i is the valuation of an advertiser with preferences \tilde{y}_i . The induced bid is denoted by \tilde{b}_i , while b_i denotes the bid when preferences are reported truthfully.

Proposition 2 (Truthful Reporting in Automated Bidding)

Under the optimal information structure, it is a dominant strategy for an advertiser to report truthfully his preferences to the publisher.

Proof of Proposition 2. By assumption, (v_i, y_i) are independently distributed. Thus the distribution of the bid \tilde{b}_i is the same for every reported preferences \tilde{y}_i . Of course, the joint distribution of (v_i, \tilde{b}_i) does depend on the reporting strategy. In fact, we note that for all $v', b' \in \mathbb{R}$:

$$\Pr(v_i \leq v', \tilde{b}_i \leq b') \leq \min\{\Pr(v_i \leq v'), \Pr(\tilde{b}_i \leq b')\} = \min\{\Pr(v_i \leq v'), \Pr(b_i \leq b')\} = \Pr(v_i \leq v', b_i \leq b'),$$

The first inequality is true for any bivariate distribution, the first equality is because both distributions have the same marginals, and the second equality follows from the fact that:

$$\Pr(v_i \leq v', b_i \leq b') = \begin{cases} F(v') & \text{if } v' \leq b' \text{ or } b' = \mathbb{E}[v_i \mid v_i \geq F^{-1}(q_N^*)], \\ F(b') & \text{otherwise.} \end{cases}$$

By definition, this means that (v_i, b_i) is greater than (v_i, \tilde{b}_i) in the positive quadrant dependent order (see Shaked & Shanthikumar (2007)).

We now write the difference in the expected bidder's surplus under truthful reporting and misreporting as follows:

$$\begin{aligned} \Delta V &\triangleq \mathbb{E}[(v_i - b_i) \Pr(b_i \geq \max_{j \neq i} b_j)] - \mathbb{E}[(v_i - \tilde{b}_i) \Pr(\tilde{b}_i \geq \max_{j \neq i} b_j)] \\ &= \mathbb{E}[(v_i \Pr(b_i \geq \max_{j \neq i} b_j)] - \mathbb{E}[v_i \Pr(\tilde{b}_i \geq \max_{j \neq i} b_j)], \end{aligned}$$

where $\Pr(b \geq \max_{j \neq i} b_j)$ is the probability that the $N - 1$ competing bids are less than b and we use that the distribution of b_i is the same as \tilde{b}_i to cancel out two terms. Clearly $\Pr(b \geq \max_{j \neq i} b_j)$ is an increasing function of b , so $v \Pr(b \geq \max_{j \neq i} b_j)$ is supermodular in (v, b) , which implies that

$$\mathbb{E}[v_i \Pr(\tilde{b}_i \geq \max_{j \neq i} b_j)] \leq \mathbb{E}[v_i \Pr(b_i \geq \max_{j \neq i} b_j)],$$

see (9.A.18.) in Shaked & Shanthikumar (2007). ■

The proposition states that advertisers are willing to submit their preferences honestly to the publisher. The intuition for the proof is that misreporting would not change the distribution of

bids, but instead, only decrease the correlation between an advertiser’s bid and his valuation. The automated bidding algorithm can alternatively be interpreted as a restriction on the bidding language imposed by a publisher. The publisher then runs a second price auction conditional on the realized attributes and the bidding rules.

We can illustrate this in our leading example of binary characteristics described above in (10). The optimal information structure can be implemented in a straightforward manner in the model of characteristics. Namely, the seller informs each bidder about the number of matched characteristics as long as this number is smaller than a threshold number n^* implied by the optimal quantile q_N^* . If the number of matched characteristics exceeds n^* , then the seller only reports that the realized matches exceed the threshold number n^* . Thus, a restriction in terms of the bidding language would allow the bidder to place bids as a function of the matched characteristics up to n^* , but not beyond that. This restriction in terms of the bidding language would then lead to the optimal bids as described by (12).

By interpreting Theorem 1 through the model of binary characteristics (10) we get that the optimal information structure supports a match process between advertiser and viewer that is often referred to as "broad match" in digital advertising, see Dar et al. (2009) and Eliaz & Spiegler (2016). By broad matching, the seller provides matches not only on the exact and narrow matches for specific keywords and characteristics but also for a larger, hence broad set of matches, see Google Ads Help Center (2021c). In our model, these are matches of quality at least n^* . Our paper’s main insight in this context is that "broad matching" should be used only to pool high quality matches, while low quality matches should bid according to their values.

5 Discussion and Conclusion

We studied the revenue-maximizing information structure in a second-price auction without reserve price and provided a model of two-sided private information that generated a model of independent private values. We focused on the market of impression as our main application, but we believe that our analysis is valuable to other markets where the seller’s information is superior to that of the buyers, e.g. markets of second hand goods, and certain markets of experience goods.

The analysis can be extended in a variety of different directions and there are also several interesting questions for future work.

Auction Format In the current environment with independent private values, the revenue equivalence result holds. Thus all classic auction formats, e.g., first-price auction, generate the same expected revenue. Hence, while we formally study the second-price auction, the results extend to all classic auction formats.

Reserve Price Our main analysis focuses on the second-price auction without reserve price. However, the analysis of the optimal information structure can be extended to auctions with a reserve price r with minor modifications (Bergemann et al. (2021)). With a reserve price r , the optimal information structure is as follows. At the top, there is a pooling region followed by a region of full disclosure (bidders learn their valuation if it is in this region). However, the full disclosure region is followed by a second pooling region that generates an expected valuation exactly equal to the reserve price. The reason for the second pooling region is that the seller’s expected revenue is discontinuous in the expected values around the cutoff, so bidders who are marginally below the reserve price r are made more optimistic so that they bid the reserve price.

Asymmetric Information Structures We focus on describing the optimal symmetric information structure. While we do not have a general result showing that the optimal information structure is always symmetric, we have some partial results that point to this direction (cf. Bergemann et al. (2021)). The symmetric information structure that we derive is indeed the unique optimal information structure when there are two or three bidders, thus $N = 2$ or $N = 3$. We can also show that, if the information structure is the optimal symmetric information for $N - 1$ bidders, then it is optimal for the remaining bidder to also observe the optimal symmetric information. Hence, there is no improvement that involves changing the information structure of only one bidder.

Bidder-Optimal Information Structure The bidder-optimal information structure can be analyzed using the same tools we used in our paper. In this case, the optimal information structure exhibits pooling at the bottom and complete information at the top of the bid distribution (cf. Sorokin & Winter (2021)).

Manual Bidding and Obedience We discussed how the optimal information structure can be implemented by automated bidding algorithms. These algorithms generate bids for the advertisers as function of the preferences elicited from the advertisers and the attributes of the viewer. We established in Proposition 2 that automated bidding supports truthful revelation of the private

information by the advertisers. A different class of algorithms is frequently described as manual bidding algorithms. Here, the bidders are asked to reveal their preferences first, and then are invited to bid on the basis of bid recommendations that take into account both the preferences and the attribute information. These algorithms implicitly require both truthtelling and obedience constraints to be satisfied. The additional restrictions imposed by the obedience constraints, namely that the bidder wishes to follow the bid recommendation, may sometimes prevent the implementation of the optimal information structure for a given number N of bidders. Yet, we can show that in the presence of a large number of bidders we can approximate the revenue of the optimal information structure with a two-sided pooling information structure that is only a small variation of the optimal information structure presented here (cf. Bergemann et al. (2021)). Finding the optimal information structure when we impose the dual incentive constraints of truthtelling and obedience on the information design remains an interesting open question.

Horizontal vs Vertical Differentiation An interpretation of the characteristics model is that it generates horizontal differentiation among bidders. By contrast, other models of characteristics would generate vertical differentiation and correlated values. For example, if each dimension of attributes and tastes were represented by the positive orthant, then a higher realization of any attribute would represent an element of vertical differentiation, and the ex-ante distribution of values would be correlated across bidders, thus introducing an element of common values. If valuations are correlated among bidders, then bidders would be subject to the winner’s curse. This would introduce new trade-offs that relate the bidders’ information to the winner’s curse. From a technical perspective, the second price auction would no longer have a dominant-strategy equilibrium, so the characterization of the revenue in terms of the second-order statistic of expected valuations would no longer hold. While extending the analysis formally is not trivial, the linkage principle suggests that the seller would publicly disclose any shocks about common shocks affecting correlations, thus making the valuations conditionally independent which would allow us to go back to our analysis, see Bergemann et al. (2017), (2019).

References

- Aggarwal, G., Badanidiyuru, A. & Mehta, A. (2019), Autobidding with constraints, *in* ‘International Conference on Web and Internet Economics’, Springer, pp. 17–30.
- Alonso, R. & Camara, O. (2016), ‘Persuading voters’, *American Economic Review* **106**, 3590–3605.
- Arnosti, N., Beck, M. & Milgrom, P. (2016), ‘Adverse selection and auction design for internet display advertising’, *American Economic Review* **106**, 2852–2866.
- Bergemann, D., Brooks, B. & Morris, S. (2017), ‘First price auctions with general information structures: Implications for bidding and revenue’, *Econometrica* **85**, 107–143.
- Bergemann, D., Brooks, B. & Morris, S. (2019), ‘Revenue guarantee equivalence’, *American Economic Review* **109**, 1911–1929.
- Bergemann, D., Heumann, T. & Morris, S. (2021), Selling impressions: Efficiency vs market thickness, Technical report.
- Bergemann, D. & Pesendorfer, M. (2007), ‘Information structures in optimal auctions’, *Journal of Economic Theory* **137**, 580–609.
- Blackwell, D. (1951), Comparison of experiments, *in* ‘Proc. Second Berkeley Symp. Math. Statist. Probab.’, University of California Press, Berkeley, pp. 93–102.
- Dar, E. E., Mirrokni, V., Muthukrishnan, S., Mansour, Y. & Nadav, U. (2009), ‘Bid optimization for broad match auctions’, *WWW 2009 Proceedings of the 18th International Conference on World Wide Web* .
- Deng, Y., Mao, J., Mirrokni, V. & Zuo, S. (2021), Towards efficient auctions in an auto-bidding world, *in* ‘Proceedings of the International World Wide Web Conference’.
- Dworczak, P. & Martini, G. (2019), ‘The simple economics of optimal persuasion’, *Journal of Political Economy* **127**, 1993–2048.
- Eliasz, K. & Spiegler, R. (2016), ‘Search design and broad matching’, *American Economic Review* **106**, 563–586.

- Emek, Y., Feldman, M., Gamzu, I., Leme, R. P. & Tennenholtz, M. (2012), Signaling schemes for revenue maximization, *in* ‘ACM EC 12’.
- Facebook Business Help Center (2021), ‘About automated app ads’, <https://www.facebook.com/business/help/309994246788275?id=1858550721111595>, Accessed 16th May 2021 .
- Ganuzza, J.-J. (2004), ‘Ignorance promotes competition: An auction model of endogeneous private valuation’, *RAND Journal of Economics* **35**.
- Ghosh, A., Nazerzadeh, H. & Sundararajan, M. (2009), Computing optimal bundles for sponsored search, Technical report, Yahoo Research.
- Google Ads Help Center (2021a), ‘About app campaigns’, <https://support.google.com/google-ads/answer/6247380?hl=en>, accessed May 25th, 2021 .
- Google Ads Help Center (2021b), ‘About automated bidding’, <https://support.google.com/google-ads/answer/2979071?hl=en>, Accessed 15 May 2021 .
- Google Ads Help Center (2021c), ‘Broad match’, <https://support.google.com/google-ads/answer/2407779?hl=en> accessed May 31st 2021.
- Hartline, J., Johnsen, A., Nekipelov, D. & Zoeter, O. (2019), Dashboard mechanisms for online marketplaces, *in* ‘Proceedings of the 20th ACM Conference on Economics and Computation’.
- Kleiner, A., Moldovanu, B. & Strack, P. (2021), ‘Extreme points and majorization: Economic applications’, *Econometrica* **forthcoming**’.
- Kolotilin, A., Mylovanov, T. & Zapechelnuyk, A. (2021), ‘Censorship as optimal persuasion’, *Theoretical Economics* **forthcoming**.
- Kolotilin, A. & Wolitsky, A. (2020), The economics of partisan gerrymandering.
- Krishna, V. (2009), *Auction Theory*, Academic press.
- Levin, J. & Milgrom, P. (2010), ‘Online advertising: Heterogeneity and conflation in market design’, *American Economic Review Papers and Proceedings* **100**, 603–607.

Papoulis, A. & Saunders, H. (1989), *Probability, Random Variables and Stochastic Processes*, McGraw-Hill.

Shaked, M. & Shanthikumar, J. (2007), *Stochastic Orders*, Springer Verlag.

Sorokin, C. & Winter, E. (2021), Pure information design in classic auctions.