Adverse Selection and Auction Design for Internet Display Advertising

Nick Arnosti, Marissa Beck and Paul Milgrom

"Half the money I spend on advertising is wasted;
the trouble is, I don't know which half."
- John Wanamaker (19th century advertising pioneer)

Abstract

We model an online display advertising environment in which "performance" advertisers can measure the value of individual impressions, whereas "brand" advertisers cannot. If advertiser values for ad opportunities are positively correlated, second-price auctions for impressions can be very inefficient. Bayesian-optimal auctions are complex, introduce incentives for false-name bidding, and disproportionately allocate low-quality impressions to brand advertisers. We introduce "modified second bid" auctions as the unique auctions that overcome these disadvantages.

When advertiser match values are drawn independently from heavy tailed distributions, a modified second bid auction captures at least 94.8% of the first-best expected value. In that setting and similar ones, the benefits of switching from an ordinary second-price auction to the modified second bid auction may be large, and the cost of defending against shill bidding and adverse selection may be low.

1 Introduction

Since the pioneering papers by Edelman, Ostrovsky, and Schwarz (2007) and Varian (2007), there has been a growing body of research focusing on auctions for sponsored search advertising on the internet. These automated auctions, which are initiated when a consumer enters a search query, allow advertisers to bid in real time for the opportunity to post ads near unpaid search results.

Markets for internet display advertising, which determine ad placements on all other kinds of web pages, traditionally operated quite differently. Following practices established for offline media, most impressions were sold through contracts that guaranteed an advertiser a large volume of impressions over a pre-specified time period, for a negotiated price. More recently, the high revenues associated with sponsored search auctions prompted the display advertising industry to introduce opportunities to bid in auctions for individual impressions. Proponents of real-time bidding argue that it alone has the potential to allow each online impression to be routed to the advertiser who values it most highly. Achieving this theoretical ideal, however, requires that advertisers know their values for individual impressions.

What advertisers know about their values depends in part on the types of feedback they collect. Some advertisements are designed to elicit an immediate response from the consumer – most commonly, a click on a link to a website where a sale may occur. By aggregating click and sales data from many impressions, an advertiser might learn how various characteristics of past opportunities have correlated with ad performance. With this knowledge, such a "performance advertiser" could predict the value of future advertising opportunities with reasonable accuracy.

Other advertisements, meanwhile, do not seek an immediate response from the user. Instead, they aim to notify the user of a sales event, movie opening, or other upcoming opportunity, or to raise the user's awareness of an individual brand or product. Recent work by Lewis and Rao (2015) highlights the difficulty of estimating the return from such a campaign; it follows that the difficulty of attributing value to individual ads is even greater. The limited feedback available to these brand advertisers makes it hard for them to reliably estimate values for individual ad opportunities.

For brand advertisers, real-time bidding presents an additional challenge by introducing the possibility of adverse selection. Advertisers' values for different ad opportunities tend to be positively correlated: all advertisers want to avoid showing their ads to automated web crawlers and most prefer to show more ads to consumers with greater disposable incomes and responsiveness to online advertising. As a result, performance advertisers tend to bid higher on opportunities that are more valuable to brand advertisers and brand advertisers may have difficulty measuring the resulting losses and adjusting their bids accordingly.

This work examines how to allocate impressions efficiently when advertiser valuations are positively correlated and advertisers are differentiated in their abilities to estimate the value of individual impressions. We use efficiency as our primary objective for two reasons. First, the rules of the mechanism are often chosen by an ad exchange that is competing with other exchanges to attract both publishers and advertisers. Second, in a world where many ads are sold through negotiated contracts, any revenue analysis must specify a bargaining protocol through which prices are set. By focusing on efficiency, we highlight fundamental informational challenges while remaining agnostic regarding details such as the bargaining power of each participant.

In our model, each performance bidder's value for an ad opportunity is the product of two factors: the overall quality of the opportunity, which affects the value for all advertisers, and an idiosyncratic match quality, which is unique to that particular bidder. If values are primarily determined by overall quality with little variation in idiosyncratic match qualities, then the benefits from using real-time bidding to assign individual impressions are minimal. If, on the other hand, differences in bidders' value estimates arise mainly from differences in match quality with little variation in overall quality, then adverse selection becomes unimportant and real-time bidding using a second-price auction may achieve a nearly efficient allocation. Ideally, one would want a mechanism that performs well in both extreme cases as well as in intermediate ones in which common and idiosyncratic components are both significant determinants of advertiser valuations. Such a robust mechanism might be especially attractive if advertisers disagree about their ability to measure value and the importance of adverse selection.

With these goals in mind, we consider several ad allocation mechanisms:

1. Second-price auctions with reserves, resembling current practice.

2. An optimal auction, which maximizes expected value, given the auctioneer's beliefs.

3. A new class of "modified second bid" (MSB) auctions, which allocate each impression to the highest performance bidder provided its bid exceeds the second highest bid by a factor of more than α (for some $\alpha \geq 1$).

4. The "omniscient" or "first-best" benchmark, which operates as if the auctioneer perfectly observed the overall quality of each impression.

We demonstrate that when there is substantial uncertainty about the common value of each impression, the second-price auction with an optimal reserve may capture as little as half of the value delivered by an optimal mechanism. The Bayesian optimal mechanism, too, has significant drawbacks: it may depend in a very detailed way on assumptions about the environment, assign a disproportionately large number of low quality impressions to the brand advertiser, and be manipulable through shill bidding. The first two drawbacks are especially problematic if advertisers disagree about the correlations among their values. This motivates us to study the class of deterministic, anonymous mechanisms that eliminate both adverse selection and the incentive to create multiple identities. Within our model, we find that this class consists of precisely the set of modified second bid auctions.

Our introduction and analysis of MSB auctions follows a respected tradition in market design, according to which a proposed mechanism is identified by properties that are neither inherited from past practice nor derived from any underlying optimization. Common properties are ones like stability, envy-freeness, and strategy-proofness, which are intuitively appealing and do not rely on detailed distributional assumptions. A goal of this approach is to respect the Wilson doctrine by creating a mechanism whose performance is not too sensitive to detailed assumptions about the economic environment. But can such a mechanism perform well in this setting? By insisting on a mechanism that is immune to shill bids and shields brand advertisers from adverse selection, do we impose large losses in the efficiency of the allocation? These are the questions to which we turn in Section 7.

To achieve an efficient allocation, a mechanism must somehow separate the cases in which some performance advertiser has a high match value (and so should win the impression) from those in which no performance advertiser is a good match (implying that the brand advertiser should win). Informally, modified second bid auctions assume the former case whenever one advertiser significantly outbids the others, and otherwise assumes the latter. Of course, this inference is imperfect: whenever two performance advertisers have similarly high match values, the MSB inefficiently awards the impression to the brand advertiser. If, however, a large fraction of the value is generated by a relatively small number of particularly good matches, one might expect that losses due to this type of error would be small.

We formalize this intuition by considering settings in which performance match values are IID draws from a power law (Pareto) distribution, which is the prototypical distribution used to model heavy tails. We demonstrate that for any power law distribution, any number of bidders, any distribution of the common value, and any mean value for the brand advertiser a modified second bid auction meets a high performance standard: it always captures at least 94.8% of the value generated by our omniscient benchmark. This comparison suggests not only that the MSB mechanism eliminates adverse selection and shill bidding at low cost, but also implies a much stronger statement. For even if the auctioneer could implement the omniscient solution (which would require that it know whether the brand or performance value is higher), that would still permit only modest improvements over MSB.

2 Related Work

There is a large and growing literature on auctions for internet advertising. Various papers focus on the role of risk aversion, budget constraints, dynamic contract formation and fulfillment in the presence of uncertain supply, revealing and/or selling cookies, the role of intermediaries, and many other concerns. In this work, we ignore dynamic considerations and budgets to focus on the information structure and the potential difficulty of even the seemingly simple goal of determining a myopically efficient allocation.

The line of work most closely related to our own is that which asks the question of how to jointly allocate impressions across advance contracts and spot market sales. The primary differences between these papers and our own are the assumptions about the correlation between advertiser valuations. Prior work has focused on extreme cases by assuming either that spot market bids do not provide information about the contract advertiser's value (Chen, 2013; Balseiro et al., 2014) or that they perfectly reveal this value (Ghosh et al., 2009). In either case, the question of determining a (myopically) efficient allocation is trivial, so these papers focus on other topics. By contrast, we formulate a model with a mix of common and private values, so that spot market bids provide information about the brand advertiser's value for the impression, but do not perfectly determine it.

Milgrom and Weber (1982) provide one of the earliest analyses of auction environments with a mix of common and private values. Due to the symmetry in the assumed environment, standard auctions yield an efficient allocation. Some later work has investigated information asymmetry in settings where auctions have a common value component. Most recently, Abraham et al. (2013) study information asymmetry and adverse selection in the context of internet advertising. They focus on the auctioneer's revenue and most of their work considers an environment with pure common values, so that achieving an efficient allocation is trivial.

3 Model

We consider the allocation of a single impression. There are $N+1$ advertisers competing for this impression, with N random and $P(N \geq 2) = 1$. Advertiser i has value X_i for the impression. The value X_i is the product of two terms: a "common value" C and a "match value" M_i. We interpret C as capturing attributes of the user that are valuable to all advertisers, such as the user's income and responsiveness to online advertising. Meanwhile, the value M_i captures idiosyncratic components that contribute to the quality of the match for advertiser i. We assume that, given N, the M_i are drawn independently from (not-necessarily-identical) distributions F_i, that C is drawn from a distribution G, and that $M = (M_0, \ldots, M_N)$ is independent from C. We use $X_{(k)}$ and $M_{(k)}$ to denote the k^{th} highest value and match value factor, respectively.

We assume that advertisers $i \in \{1, \ldots, N\}$ (the "performance advertisers") observe their values X_i, but not its components C and M_i. Meanwhile, advertiser 0 (whom we refer to as a "brand advertiser") **cannot observe** X_0.[1] The assumption that there is a single uninformed advertiser and multiple informed advertisers is made for expositional simplicity and could be easily relaxed as part of a richer model.

Making use of the revelation principle, we consider a *mechanism* to be a mapping from the privately held information $X = (X_1, \ldots, X_N)$ to allocation probabilities z and payments p. For $i \in \{0, \ldots, N\}$, we let $z_i(X)$ be the probability that advertiser i wins, and $p_i(X)$ be advertiser i's expected payments, given X. Given an allocation rule z, we define the total surplus from impressions awarded to the brand advertiser by $V_B(z) = E[X_0 z_0(X)]$. Similarly, we define the surplus from impressions awarded to performance advertisers by $V_P(z) = E\left[\sum_{i=1}^{N} X_i z_i(X)\right]$. We consider objectives that are weighted sums of these two terms. In particular, for $\gamma \in [0, 2]$, we define

$$V_\gamma(z) = \gamma V_B(z) + (2-\gamma) V_P(z).$$

In particular, when $\gamma = 1$, this corresponds to the total efficiency of the allocation.

Throughout the paper, we use as a benchmark the "first best" or "omniscient" mechanism,

[1] If the brand advertiser can measure the value of its overall campaign, but nothing more, then it observes $E[X_0]$, but not the distributions G and F_0 separately. This motivates the axiom of Adverse-Selection-Freeness discussed in Section 6.

which operates as the publisher would if the common value C were observed. In this case, the publisher could deduce the match values M_i from the performance advertiser reports of X_i and condition the allocation directly on those match values. Given a weight γ, the allocation rule of this mechanism is simple: allocate the impression to the performance advertiser with the highest value provided that $(2-\gamma)M_{(1)} > \gamma E[M_0]$ and otherwise allocate the impression to the brand advertiser. We denote this allocation rule by $OMN(\gamma)$. Because we often consider the case $\gamma = 1$ (the publisher weighs value to brand and performance advertisers equally), we use OMN to refer to $OMN(1)$, and $V(\cdot)$ to refer to $V_1(\cdot)$. The omniscient allocation rule is in general not implementable, but it provides an upper bound on the welfare attainable by any mechanism and a way to measure the losses due to the fact that C is unobservable.

Note that for any allocation rule z, the quantities $V_B(z)$, $V_P(z)$ and $V_\gamma(z)$ depend implicitly on the number of bidders and the distributions of advertiser values. At times, we will discuss the ways in which the performance of an allocation rule varies with an underlying parameter θ. In such cases, we make the dependence on θ explicit by writing $V(z;\theta)$.

4 Second-Price auction

If all advertisers could determine their values for the impression, then running a simple second-price auction would yield an efficient allocation. It is natural to wonder whether this might still prove to be an effective solution when the brand advertiser cannot observe its value for the impression.

In our model, the brand advertiser cannot observe its value, so it has no dominant strategy in a second-price auction. Thus, we must make some assumption about how it bids. Naively, it might bid its expected value $E[X_0]$ in each auction. This would be optimal if its value were independent of the values of other advertisers (i.e., if C were constant), but could be a very poor strategy if the correlation between the brand value and the values of other advertisers is high. Instead, we make the rational expectations assumption that the brand advertiser chooses b to maximize its expected profit, as given by

$$\Pi(b) = E[(X_0 - X_{(1)})\mathbf{1}_{X_{(1)} \leq b}]$$

where $X_{(i)}$ denotes the i^{th} highest performance advertiser value. Note that in order to compute the optimal bid, the brand advertiser requires a great deal of information: it must know the distribution of the top performance bid and the correlation between $X_{(1)}$ and X_0 (as determined by the distribution of C). In practice, these could be challenging to learn.

If the publisher has more information about competing bidders, a reasonable alternative is for the publisher to submit a "proxy bid" on behalf of the brand advertiser. This resembles a solution to the allocation problem commonly used in practice: the publisher signs a contract with the brand advertiser and submits each opportunity (along with a corresponding reserve price) to a real-time exchange. If the reserve is met, the impression is awarded to the top bidder. Otherwise, it is allocated to the brand advertiser. We let $SP(b)$ denote the allocation rule of the second-price auction when the brand advertiser submits a bid of b (or the publisher does so on the brand advertiser's behalf). This mechanism is simple and intuitive, and allows performance advertisers to win impressions for which they have very high values. However, for any reserve price, the brand advertiser wins more often when C is low than when C is high, which can cause the resulting allocation to be inefficient.

How inefficient might the second-price auction (with optimally chosen reserve)[2] be? Two extremes are to set the reserve at zero (so that the brand advertiser never wins) or to set it arbitrarily high (so that the brand advertiser always wins). Thus, the second price auction can always deliver a value of $\max(E[X_0], E[X_{(1)}])$. Clearly, the ominiscient mechanism cannot deliver more than $E[X_0] + E[X_{(1)}]$, so for any distribution of match and common values, $\sup_{b \geq 0} V(SP(b)) \geq \frac{1}{2} V(OMN)$. The next result states that this bound is tight, even if performance match values are IID.

Proposition 1.

1. *For any F_i, G, N, $E[M_0]$ there exists $b \in \{0, \infty\}$ such that $V(SP(b)) \geq \frac{1}{2} V(OMN)$.*

2. *For any $\epsilon > 0$, there exists $F, G, N, E[M_0]$ such that if performance match values are drawn IID from F,*

$$\sup_b V(SP(b)) < \left(\frac{1}{2} + \epsilon\right) V(OMN).$$

To show that the bound of 1/2 is tight, we use a sequence of examples in which both match and performance values follow power law distributions: for $x \geq 1$, $Pr(M_i > x) = x^{-a}$ and $Pr(C > x) = x^{-c}$, where $a, c > 1$ are parameters determining the weight of the power law tails.

For the second-price auction to perform this poorly, three things must be true. First, the omniscient mechanism needs to deliver nearly $E[X_0] + E[X_{(1)}]$. Second, the highest bid in the auction must convey virtually no information about the optimal assignment, so that $\sup_b V(SP(b)) \approx \max(E[X_0], E[X_{(1)}])$. Third, it must be the case that $E[X_0] = E[X_{(1)}]$. The first condition holds as $a \downarrow 1$, for in that case nearly all of the performance value can be captured from a vanishingly small fraction of the impressions, while delivering the remainder to the brand advertiser. For fixed a, the second condition holds as $c \downarrow 1$, for then nearly all of the variation in the highest bid comes from variations in C (which are irrelevant when determining the optimal assignment). We can always choose $E[M_0]$ such that the third condition holds.

Although the circumstances above are quite specific and the worst-case bound of 50% is never exactly met, the above reasoning highlights the conditions that are likely to lead to bad performance

[2]Note that in our model, the brand advertiser and publisher have the same information, and thus the second-price auction and "contract with proxy bidding" are mathematically equivalent. In particular, if the publisher uses a reserve of b, the resulting total surplus is $E[X_0 \mathbf{1}_{X_{(1)} \leq b}] + E[X_{(1)} \mathbf{1}_{X_{(1)} > b}] = \Pi(b) + E[X_{(1)}]$. In other words, the brand advertiser's optimal bid in a second-price auction is precisely the bid which maximizes allocative efficiency.

of the second-price auction. If the publisher faces substantial uncertainty about both match and common value factors, the performance of even an optimally selected second-price auction may be unsatisfying. Motivated by this fact, we next consider the properties of the Bayesian-optimal mechanism.

5 Bayesian Optimal Auction

In this section, we consider the allocation rule OPT that maximizes the expected allocative efficiency. Formally, $OPT \in \arg\max_z V(z)$. Unsurprisingly, OPT may be very complex and depend in a detailed way on the distributions F_i and G. Furthermore, we illustrate through examples two other potentially undesirable features of OPT:

1. Performance advertisers may have an incentive to submit shill bids.
2. The brand advertiser in general gets an unrepresentative set of impressions.

The first of these concerns is motivated by the fact that in many web-based environments, bidders reveal only virtual identities, concealing their real identities. This introduces the possibility that a bidder might create several accounts and submit multiple bids for the same impression. If the publisher uses the optimal mechanism, such behavior might prove beneficial. The optimal mechanism uses the complete bid profile to draw inferences about C and guide the allocation: the impression is awarded to the top bidder only if its bid exceeds the expected value of the impression to the brand advertiser (given the bid profile). By submitting several low bids, the top bidder can mislead the mechanism to the conclusion that C must be low and, thereby, win impressions that otherwise would have gone to the brand advertiser. Such behavior is undesirable, as it results in a less efficient allocation.

The fact that the brand advertiser disproportionately wins impressions of low quality may also present challenges. In order to calculate the value that this mechanism awards to brand advertisers, it is necessary to understand the correlation between brand and performance values (i.e., the distribution of C). This is unobserved and in practice the publisher and brand advertiser might be unable to agree upon an appropriate price to compensate the brand advertiser for adverse selection.

The following example illustrates each of the concerns above. In this example, C is binary and the brand advertiser wins with probability 7/9 when C is low and with probability 4/9 when C is high. Furthermore, a bidder with a value of 4 has an incentive to use a second profile to submit a bid of 1.

Example 1. *Suppose that C is uniformly distributed on $\{1, 2\}$ and that given N, $M_1, \ldots, M_N$ are*

IID uniform on $\{1,2,4\}$*, so that* $X_i \in \{1,2,4,8\}^3$*. Suppose that* $3 < E[M_0] < 4$*.*

The optimal mechanism specifies the following allocation for the case $N = 2$*:*

- *If* $X_{(1)} = 8$*, allocate to the top performance advertiser (in this case,* $M_{(1)} = 4 > E[M_0]$*).*
- *If* $X_{(1)} \in \{1,2\}$*, allocate to the brand advertiser (in this case,* $M_{(1)} \leq 2 < E[M_0]$*).*
- *If* $X_{(1)} = 4$*,*
  - *If* $X_{(2)} = 1$*, allocate to the top performance advertiser (because* $M_{(1)} = 4 > E[M_0]$*).*
  - *If* $X_{(2)} = 2$*, allocate the impression to the brand advertiser. In this case, there are two equally likely possibilities:* $C = 1, M_{(1)} = 4, M_{(2)} = 2$ *and* $C = 2, M_{(1)} = 2, M_{(2)} = 1$*. Thus,* $E[M_{(1)}|X] = 3 < E[M_0]$*.*
  - *If* $X_{(2)} = 4$*, allocate to the brand advertiser. Again, there are two equally likely possibilities:* $C = 1, M_1 = M_2 = 4$*, and* $C = 2, M_1 = M_2 = 2$*. Thus,* $E[M_1|X] = E[M_2|X] = 3 < E[M_0]$*.*

Motivated by these observations, in the following section we seek a mechanism which eliminates shill bidding and adverse selection.

6 An Axiomatic Approach

The previous section demonstrated that the optimal mechanism can in general be manipulated by shill bids and may award brand advertisers impressions of disproportionately low value. In this section, we seek to study mechanisms which avoid these concerns.

We consider an extended game in which performance advertisers have the option to create (and bid from) multiple accounts. In such a setting, it may be infeasible to discriminate against individual bidders, so we seek a mechanism which treats bidders symmetrically. Furthermore, we require a mechanism for which it is a dominant strategy for each advertiser to play as a single bidder and to report its value truthfully. This restriction is similar to that imposed by Yokoo et al. (2004). Compared to their work, we introduce the additional requirement that the publisher must not be able to profitably exploit anonymous bidding by submitting false-name bids that are lower than any of the other bids. This is an imperfect stand-in for the restriction the seller should be unable to submit low bids that increase the sales price it receives but have little risk of winning.

Definition 1. *Given a mechanism* (z, p)*, let* z_{-0} *and* p_{-0} *denote the allocations and payments of performance advertisers. The mechanism is* **anonymous** *(among performance advertisers) if, for*

[3] Although we consider a discrete example in which ties may occur, it is easy to see that the logic extends to the case where all distributions are continuous, so we ignore tie breaking in our analysis.

any $n \geq 2$, any permutation σ on $\{1, \ldots, n\}$ and any $x \in \mathbb{R}_+^n$, the following hold:

$$\sigma(z_{-0}(x)) = z_{-0}(\sigma(x)) \qquad \text{and} \qquad \sigma(p_{-0}(x)) = p_{-0}(\sigma(x)),$$

Definition 2. *The mechanism (z,p) is* **strategy-proof** *if, for all $n \geq 2$, all $x \in \mathbb{R}_+^n$, all $i \in \{1, \ldots, n\}$, and all $\hat{x}_i$,*

$$x_i z_i(x) - p_i(x) \geq x_i z_i(\hat{x}_i, x_{-i}) - p_i(\hat{x}_i, x_{-i}).$$

The mechanism (z,p) is **bidder false-name proof** *if no bidder can benefit by submitting multiple bids, meaning that for all $n \geq 2$, $x \in \mathbb{R}_+^n$, all $m \geq 1$, and all $y \in \mathbb{R}_+^m$:*

$$x_i z_i(x) - p_i(x) \geq x_i \left(z_i(x,y) + \sum_{j=1}^{m} z_{n+j}(x,y) \right) - \left(p_i(x,y) + \sum_{j=1}^{m} p_{n+j}(x,y) \right).$$

A mechanism (z,p) is **publisher false-name proof** *if submitting bids below that of the lowest performance bidder cannot raise the price paid by a winning bidder; that is, if for all $n \geq 2$, $x \in \mathbb{R}_+^n$, $m \geq 1$ and $y \in \mathbb{R}_+^m$ satisfying* $\max y \leq \min x$*:*

$$z_i(x) > 0 \implies p_i(x,y) \leq p_i(x).$$

The mechanism (z,p) is **fully strategy-proof** *if it is strategy-proof, bidder false-name proof and publisher false-name proof.*

We also formalize the idea that a mechanism should award brand advertisers a "representative" sample of impressions, for *any* distribution of match and common values. We think of this as a robustness criterion: if the mechanism has this property, then the publisher and brand advertiser do not need to agree on details of the environment. Both agree that the value awarded to the brand advertiser is simply $E[X_0]$ times the number of impressions awarded (regardless of their beliefs about C).

Definition 3. *A mechanism (z,p) is* **adverse-selection free** *if, for every joint distribution of (C, M) such that M is independent from C, $z_0(CM)$ is also independent of C.*

We now turn our attention to a simple class of mechanisms which have all of the above properties: the family of modified second bid (MSB) auctions. These auctions offer the impression to the top bidder, at a price equal to α times the second highest bid (for some $\alpha \geq 1$). If the top bidder is unwilling to pay this price, the impression is awarded to the brand advertiser. It turns out that these are the *only* deterministic mechanisms which satisfy the three properties given above.

Theorem 1. *A deterministic mechanism (z,p) is anonymous, fully strategy-proof, and adverse-selection free if and only if it is a modified second bid auction.*

Theorem 1 characterizes MSB auctions as the unique auctions which are anonymous, deterministic, adverse-selection free, and strategy-proof in a setting where bidders may create multiple identities. We note here that other auctions may have any three of these properties. Relaxing anonymity allows the value of α to depend on the bidder. Relaxing determinism permits mechanisms which randomly set aside opportunities for the brand advertiser (and award the rest via a second-price auction among performance advertisers). A second-price auction with reserve (corresponding to the brand advertiser's bid) is deterministic, anonymous, and fully strategy-proof, but not adverse-selection free. Finally, consider the auction in which the highest bidder wins whenever its bid is at least twice the lowest competing bid, and in that case pays the higher of the second-highest bid and twice the lowest competing bid. This mechanism is deterministic, anonymous, and adverse-selection free, but not bidder false-name proof.

7 Performance Analysis

Although we have characterized MSB auctions as the only mechanisms that are deterministic, anonymous, false-name-proof and adverse-selection-free, it is not clear that these properties are necessary. Perhaps shill bidding can be identified and prevented by other means. Additionally, the publisher may be able to convince the brand advertiser to accept below-average impressions in return for a lower price. If so, it is reasonable to wonder whether these axioms are costly, and whether we could achieve notably better performance by abandoning them.

In this section, we consider the performance of the modified second bid auction in a simple environment where the performance match values are IID (i.e., $F_i = F$ for $i \in \{1, \ldots, N\}$). In particular, we compare the performance of the simple classes of auctions discussed above (the second-price auction and the modified second bid auction), the optimal mechanism, and the omniscient benchmark (which directly observes C).

The following theorem states that it is always possible to choose a (trivial) MSB auction which delivers at least half of the match value delivered by OMN. The reason for this is the same as for the second price auction: choosing $\alpha = 1$ ensures that the impression is never awarded to the brand advertiser, and setting α to be arbitrarily large ensures that the brand advertiser always wins. Furthermore, the theorem states that this bound is tight, and that even the optimal auction may be unable to deliver substantially more than half of the value generated by the first-best solution.

Proposition 2.

1. *For any F_i, G, N, there exists $\alpha \in \{1, \infty\}$ such that $V(MSB(\alpha)) \geq \frac{1}{2}V(OMN)$.*
2. *For any $\epsilon > 0$, there exists $F, G, N, E[M_0]$ such that if performance match values are drawn IID from F,*

$$V(OPT) < \left(\frac{1}{2} + \epsilon\right) V(OMN).$$

How to interpret Proposition 2? In theoretical computer science, constant-factor approximations are often celebrated, and Proposition 2 states that no mechanism provides a better guarantee than that of MSB. However, many mechanisms (including the second price auction) offer the same guarantee, so this argument alone does not justify the adoption of MSB auctions. Furthermore, the possibility of *doubling* match value is economically very meaningful. Thus, one possible conclusion from Proposition 2 is that worrying about the choice of mechanism is "barking up the wrong tree": in some cases, the "real" gains may come from better information about the common value component.

We offer a third interpretation: perhaps worst-case analysis is overly pessimistic, in that the distributions used in the proof of Proposition 2 are unrealistic in some way. In particular, it has been observed that in online advertising, a large fraction of the total value comes from a small number of very valuable impressions. That is, the distribution of advertiser values can be said to be heavy-tailed. Motivated by this thought, in Theorem 2, we consider the case where match values are drawn independently from a power law distribution.[4] Under this assumption, the somewhat pessimistic conclusion of Proposition 2 reverses sharply.

Theorem 2. *Suppose that N is deterministic, and that the performance match values are IID draws from a power law distribution. Then for any γ, there exists a choice of α such that the following hold simultaneously:*

- $V_B(MSB(\alpha)) = V_B(OMN(\gamma))$.
- $V_P(MSB(\alpha)) \geq 0.885 \cdot V_P(OMN(\gamma))$
- $V_\gamma(MSB(\alpha)) \geq 0.948 \cdot V_\gamma(OMN(\gamma))$.

Although the assumptions in Theorem 2 are strong, so are the conclusions. Suppose that the publisher contracts with the brand advertiser and commits to using an allocation rule z such that $V_B(z) \geq v_B$, for some $v_B \in (0, E[X_0])$. Subject to this constraint, the publisher aims to maximize the value of allocations to performance advertisers.[5] Theorem 2 states that it is possible to choose α such that under $MSB(\alpha)$, the contract is fulfilled, and performance advertisers get at least 88.5% of the value that could be delivered to performance advertisers *if the publisher directly observed C*. Furthermore, the third component of the theorem implies that if the value of the contract with the brand advertiser is chosen optimally, then an MSB auction delivers at least 94.8% of the value of the first-best solution. These results hold for any weight of the power law tail, expected brand value $E[X_0]$, distribution of C, and number of performance advertisers N. In particular, this implies that

[4]Recall that the power law distribution has cdf given by $F(x) = 1 - x^{-a}$ on $[1, \infty)$. The parameter a determines the weight of the power law tail; for $a > 1$, the mean of the distribution is finite (in particular, it is $\frac{a}{a-1}$).

[5]Although in Theorem 2, the publisher's goal is to maximize $\gamma V_B(z) + (2 - \gamma) V_P(z)$, there is a well-known equivalence between this approach and that of maximizing $V_P(z)$ subject to a lower-bound on $V_B(z)$.

when match values follow a power law, MSB auctions are "nearly optimal" even in cases where C has a degenerate distribution (so that adverse selection is of no concern). Finally, our result that a second-price auction may attain only 50% of the performance of OMN is directly comparable, because the examples discussed in Section 4 are ones in which match values follow a power law distribution. Taken together, these facts suggest that when match values are heavy tailed,

1. The benefit of moving from a second-price auction to an MSB auction may be significant,

2. Protecting against adverse selection and shill bidding may come at little cost, and

3. The potential gains from more accurate information about C may be minimal.

We now turn to the question of how to choose α. Theorem 2 states that there exists a "good" choice of α, but offers no guidance on how to find this value. Fortunately, this turns out to be a simple problem. For any value v_B guaranteed to the brand advertiser, the fact that MSB is adverse-selection free implies that the brand advertiser must win with probability $v_B/E[X_0]$. So long as the publisher can estimate the distribution of the top two performance bids (for example, using data from previous auctions), it is simple to choose α accordingly. Furthermore, with knowledge of $E[X_0]$ and the joint distribution of $X_{(1)}$ and $X_{(2)}$, the publisher can compute, for any α, the value of the resulting allocation for both brand and performance advertisers (thus making it possible to determine an appropriate guarantee v_B to the brand advertiser). Importantly, these calculations *do not rely on assuming that performance match values are IID or follow a particular distribution*; it is enough that the match and common values are independent. Furthermore, they do not rely on *any* assumptions about the (unobserved) common value C. By contrast, bidding optimally (or setting an optimal reserve) in a second-price auction requires an understanding of the correlation between brand and performance advertiser values (i.e., of the joint distribution of $X_{(1)}$ and C).

What about the revenue of MSB auctions? Although our primary focus in this paper is allocative efficiency, the publisher naturally cares about how this surplus is divided. In order to address this topic, we must posit a division of the surplus associated with impressions allocated to the brand advertiser (a point on which we have so far remained agnostic). We suppose that the publisher and brand advertiser split this surplus proportionately; that is, the revenue to the publisher is δV_B, for some $\delta \in (0,1)$. It turns out that when match values follow a power law distribution with parameter a, for any strategy-proof incentive compatible mechanism z, publisher revenues from performance advertisers are at most $(1-a^{-1})V_P(z)$ (see the Appendix for details). Thus, in this setting, maximizing revenue is effectively equivalent to maximizing $\delta V_B + (1-a^{-1})V_P$. It follows that there exists an MSB auction which delivers revenue at least 94.8% of maximum revenue that could be achieved by *any* such mechanism, even if C were observed. We state this formally below.

Corollary 1. *Suppose that N is deterministic, that performance match values are IID draws from a power law distribution, and that the publisher and brand advertiser split surplus in any fixed* *proportion. Then there exists α such that $MSB(\alpha)$ delivers at least 94.8% of the revenue of the optimal dominant strategy incentive compatible auction.*

8 Conclusion

In this paper, we consider an auction setting in which values are correlated, and some agents may be uncertain of their own values. We introduce the modified second bid auction as the unique deterministic mechanism that is anonymous, false-name-proof, and free of adverse selection. We demonstrate that when value distributions are heavy tailed, MSB auctions may significantly improve upon more traditional reserve-price approaches, and capture nearly all of the value obtained by an omniscient benchmark.

Our work makes two contributions: one methodological, and one practical. On the methodological side, we believe that the "not-quite-optimal" approach used in this paper is a general and appealing way to derive and analyze new mechanisms. Following a long tradition in economics, we derive a candidate mechanism axiomatically, by characterizing the class of mechanisms with certain desirable properties. Mechanism in hand, we then evaluate the costs of these properties by providing performance guarantees that hold across a range of environments.

On the practical side, we demonstrate that when allocating impressions jointly to advance contracts and spot market bidders, traditional reserve-based approaches may leave significant match value "on the table." We posit that there may be significant gains to using a mechanism which adaptively sets a reserve based on submitted bids. Furthermore, most of these gains may be realized by a mechanism which is appealingly simple to understand and implement.

Of course, our theoretical findings are conditional on assumptions, and when those are varied, the conclusions will vary as well. In practice, it may happen that performance advertisers selling similar products participate in an auction to target the same users, and thus their match values may be positively correlated. In such cases, the signature the MSB mechanism utilizes - a single very high bid by a performance advertiser and lower bids by others - would miss more of the very good matches, degrading the average performance compared to the case of independent match values. Still, the improvements in our examples are so large that, even if some of the high value matches are missed in this way, substantial performance gains may still be possible.

When performance match values are correlated with the match value of the brand advertiser, then the set of impressions delivered to the brand advertiser by MSB will not be "representative." However, so long as rare transitory factors are a major determinant of extreme match values, the dependence is likely to be small and adverse selection may be largely neutralized by an MSB auction.

References

Ittai Abraham, Susan Athey, Moshe Babaioff, and Michael Grubb. Peaches, lemons, and cookies: designing auction markets with dispersed information. In *ACM Conference on Electronic Commerce*, pages 7–8, 2013.

Santiago R Balseiro, Jon Feldman, Vahab Mirrokni, and S Muthukrishnan. Yield optimization of display advertising with ad exchange. *Management Science*, 60(12), 2014.

Ying-Ju Chen. Optimal dynamic auctions for display advertising. *Available at SSRN 2216361*, 2013.

Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, 2007.

Arpita Ghosh, Preston McAfee, Kishore Papineni, and Sergei Vassilvitskii. Bidding for representative allocations for display advertising. *Internet and Network Economics (WINE)*, 2009.

Randall A Lewis and Justin M Rao. The unfavorable economics of measuring the returns to advertising. *QJE*, 2015.

Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5), 1982.

Hal R. Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163 – 1178, 2007. ISSN 0167-7187.

Makoto Yokoo, Yuko Sakurai, and Shigeo Matsubara. The effect of false-name bids in combinatorial auctions: new fraud in internet auctions. *Games and Economic Behavior*, 46:174–188, 2004.

9 Appendix: Proofs

We begin by completing the proof that second price auctions cannot guarantee $(\frac{1}{2}+\epsilon)V(OMN)$, for any $\epsilon > 0$. Throughout the appendix, we use the letter μ to refer to the brand advertiser's expected match value $E[M_0]$.

Proof of Proposition 1. Fix $N \geq 2$ and $\epsilon > 0$. We assume that M_i are iid draws from a power law distribution with parameter a, and maintain the identity $\mu = E[M_0] = (1+\epsilon)E[M_{(1)}]$. As $a \downarrow 1$, it becomes possible to capture nearly all of the value from performance advertisers by allocating to them a vanishingly small fraction of impressions. Thus,

$$\frac{V(OMN)}{E[C]E[M_{(1)}]} = \frac{E[\max(\mu, M_{(1)})]}{E[M_{(1)}]} \to \frac{\mu + E[M_{(1)}]}{E[M_{(1)}]} = 2+\epsilon.$$

Choose a sufficiently close to one such that $V(OMN) > 2E[M_{(1)}]E[C]$. Let C be drawn from a power law distribution with parameter a'. By Lemma 1, if a' is sufficiently close to one, then $\sup_b V(SP(b)) = \mu E[C] = (1+\epsilon)E[M_{(1)}]E[C]$. It follows that $\sup_b V(SP(b))/V(OMN) < \frac{1+\epsilon}{2}$. □

Lemma 1. *Suppose that $M_{(1)} \sim F$, which has density f on $[0,\infty)$, and that $E[M_{(1)}^{1+\epsilon}] < \infty$ for some $\epsilon > 0$. Fix $\mu = E[M_0] > E[M_{(1)}]$. Suppose that $C \in [1,\infty)$ has density $g(c) = a'c^{-a'-1}$. Then there exists $\delta > 0$ such that if $a' < 1+\delta$,*

$$\sup_b V(SP(b)) = \mu E[C].$$

Proof. Note that this is equivalent to showing that if a is sufficiently small, then the brand advertiser wants to increase its bid without bound (i.e. always win).

We see that

$$\begin{aligned} V(SP(b)) = \quad & E[C\mu \mathbf{1}_{CM_{(1)} \leq b}] \quad + \quad E[CM_{(1)}\mathbf{1}_{CM_{(1)}>b}] \\ &= \mu \int_1^\infty cF(b/c)g(c)dc + \int_1^\infty \int_{m=b/c}^\infty cmf(m)g(c)dmdc. \end{aligned}$$

From this, if follows that

$$\begin{aligned} \frac{d}{db}V(SP(b)) &= \int_1^\infty (\mu - b/c)f(b/c)g(c)dc. \\ &= a'b^{-a'} \int_0^b (\mu - u)f(u)u^{a'-1}du, \end{aligned}$$

where we have performed the change of variables $u = b/c$ and used the fact that $g(c) = a'c^{-a'-1}$.

We will show that for all a' sufficiently close to one, the above expression is non-negative for all b, implying that it is optimal for the brand advertiser to win all impressions. Viewed as a function of b, the integral $\int_0^b (\mu - u) f(u) u^{a'-1} du$ is (weakly) increasing on $[0, \mu]$ and (weakly) decreasing thereafter. Thus, it is enough to show that when a' is sufficiently small, $\int_0^\infty (\mu - u) f(u) u^{a'-1} du > 0$.

Because $E[M_{(1)}^{1+\epsilon}] < \infty$ for some $\epsilon > 0$, we may apply the dominated convergence theorem to see that as $a' \downarrow 1$,

$$\int_0^\infty (\mu - u) f(u) u^{a'-1} du \to \int_0^\infty (\mu - u) f(u) du = \mu - E[M_{(1)}] > 0.$$

□

Proof of Theorem 1. By inspection, any MSB auction is strategy-proof, deterministic, anonymous, false name proof and adverse selection free. Conversely, it is well-known that any strategy-proof deterministic and anonymous mechanism is characterized by a "threshold price" function h such, for any competing bids x_{-i}, bidder i wins if and only if its bid exceeds its threshold price $h(x_{-i})$ and conditional on winning, i pays this threshold price. Any such mechanism also has the property that only the top performance bidder can win, which requires that $h(x_{-i}) \geq \max\{x_{-i}\}$.

We now show that if the mechanism is false name proof, then $h(x_{-i}) = h(\max\{x_{-i}\})$. For suppose there exists exists x_{-i} such that $h(x_{-i}) \neq h(\max\{x_{-i}\})$, and examine the incentives when there are two bidders, one with value exceeding $h(x_{-i})$ and the other with value $\max\{x_{-i}\}$. If $h(x_{-i}) < h(\max\{x_{-i}\})$, then the first bidder can benefit by submitting the remaining bids in the profile x_{-i} to reduce the price, contradicting bidder false-name proofness. If $h(x_{-i}) > h(\max\{x_{-i}\})$, then the seller can benefit by submitting the remaining bids in the profile x_{-i} to raise the price, contradicting seller false-name proofness.

Next, we show that if the mechanism is adverse selection free, h must be homogeneous of degree one. For suppose not. Then there exists $c \in \mathbb{R}_+, n \geq 2$, and $x_{-i} \in \mathbb{R}_+^{n-1}$ such that (without loss of generality) $h(m_{-i}) < h(cm_{-i})/c$. Fix $m_i \in (h(m_{-i}), h(cm_{-i})/c)$. Suppose that $C \in \{1, c\}$ with $P(C = 1) \in (0, 1)$, that $P(M_{-i} = m_{-i}) = 1$, and that $P(M_i = m_i) = 1$. We show that $z_0(CM) = \mathbf{1}_{\{C=c\}}$, proving that the auction associated with h is not adverse-selection free.

When $C = 1$, $z_i(CM) = z_i(m) = \mathbf{1}_{\{m_i > h(m_{-i})\}} = 1$, so $z_0(CM) = 0$. When $C = c$, $z_i(CM) = z_i(cm) = \mathbf{1}_{\{cm_i > h(cm_{-i})\}} = 0$. Because only the top performance bidder (bidder i) can win the auction, this implies that $z_0(CM) = 1$.

Thus, for any mechanism that is deterministic, strategy proof, false-name proof, and adverse-selection free, there is a threshold price function h that is homogeneous of degree one and depends only on its maximum argument: $h(\max\{x_{-i}\}) = \alpha \max\{x_{-i}\}$ for some α. The fact that $h(x_{-i}) \geq \max\{x_{-i}\}$ implies $\alpha \geq 1$.

□

Proof of Proposition 2. Note that $\alpha = 1$ and $\alpha = \infty$ describe the cases where the brand advertiser never wins or always wins, so $V(MSB(1)) = E[X_{(1)}]$ and $V(MSB(\infty)) = E[X_0]$. Furthermore,

$$V(OMN) = E[\max(CE[M_0], CM_{(1)})] \leq E[X_0] + E[X_{(1)}] \leq 2 \cdot \max(E[X_0], E[X_{(1)}]),$$

which proves the first claim.

For the second claim, we let C be distributed according to $G(x) = 1 - x^{-b}$ on $[1, \infty)$. Fix $N \geq 2$, and suppose the M_i are iid draws from $F(x) = x^{\beta/N}$ on $[0, 1]$. Straightforward calculations reveal that if we define $\hat{g}$ to be the conditional density of C given performance values X, then

$$\hat{g}(c) = \frac{(\beta + b)}{\max(X_{(1)}, 1)} \left(\frac{c}{\max(X_{(1)}, 1)} \right)^{-\beta - b - 1} \quad \text{on } [\max(X_{(1)}, 1), \infty).$$

In other words, given X, C is distributed as a power law random variable with parameter $(b+\beta)$, conditioned on being greater than $\max(X_{(1)}, 1)$. Let $\mu = E[X_0]$. It follows that

$$E[\mu C | X] = \frac{\beta + b}{\beta + b - 1} \mu \max(X_{(1)}, 1).$$

If $\mu = \frac{\beta + b - 1}{\beta + b}$, then $E[\mu C | X] = \max(X_{(1)}, 1)$, so it is optimal to always award the impression to the brand advertiser. This generates a total value of

$$V(OPT) = E[\mu C] = \frac{b}{b-1} \frac{\beta + b - 1}{\beta + b}.$$

Meanwhile, straightforward calculations reveal that the first-best solution generates value

$$V(OMN) = E[C] E[\max(M_{(1)}, \mu)] = \frac{b}{b-1} \left(\frac{\beta}{1+\beta} + \frac{\mu^{1+\beta}}{1+\beta} \right).$$

As $b \to 1$, $\mu \to \frac{\beta}{\beta+1}$. From this, we see that

$$\lim_{b \to 1} V(OMN)/V(OPT) = 1 + \frac{\beta^\beta}{(1+\beta)^{1+\beta}}.$$

As $\beta \to 0$, this tends to 2, implying that even the optimal mechanism cannot guarantee more than 1/2 of the value generated by the first-best solution.

□

We now turn to the proof of Theorem 2. Throughout this section, we assume that N is deterministically equal to $n \geq 2$, and that the M_i are iid draws from a distribution with density f and cdf F. We use the letter μ to represent the brand advertisers expected match value $E[M_0]$. Rather than specifying μ directly, we choose an alternative parameterization by letting $\lambda \in [0, 1]$

be the probability that the brand advertiser receives the impression under the first-best solution, and defining $\mu(\lambda, n)$ by

$$\lambda = F(\mu(\lambda, n))^n, \tag{1}$$

Thus, $\mu(\lambda, n)$ gives the brand advertiser's expected value, as a function of the number of bidders n and the fraction of impressions λ won by the brand advertiser under the first-best allocation (throughout, we fix the distribution F of each performance match value).

We begin with a technical lemma, which allows us to compute $V_P(OMN)$, given λ, n, and the function $\mu(\lambda, n)$.

Lemma 2. *Suppose that $P(N = n) = 1$ and that the M_i are iid draws from a distribution with density f and cdf F. Let $\lambda = P(M_{(1)} \leq E[M_0])$, and define the function μ as in (1). Then*

$$V_P(OMN) = E[C] \int_{\lambda}^{1} \mu(x, n) dx.$$

Proof of Lemma 2. Differentiating the identity $F(\mu(\lambda, n))^n = \lambda$, we obtain

$$nF(\mu(\lambda, n))^{n-1} f(\mu(\lambda, n)) \frac{d}{d\lambda} \mu(\lambda, n) = 1. \tag{2}$$

Therefore,

$$\begin{aligned}
\frac{d}{d\lambda} V_P(OMN) &= \frac{d}{d\lambda} E[C] E[M_{(1)} \mathbf{1}_{M_{(1)} > \mu(\lambda, n)}] \\
&= \frac{d}{d\lambda} \int_{\mu(\lambda, n)}^{\infty} x n F(x)^{n-1} f(x) dx \\
&= -\mu(\lambda, n) n F(\mu(\lambda, n))^{n-1} f(\mu(\lambda, n)) \frac{d}{d\lambda} \mu(\lambda, n) \\
&= -\mu(\lambda, n),
\end{aligned}$$

where the final line follows from application of (2). The Lemma follows immediately. □

Our proof of Theorem 2 makes use of the following facts about the power law distribution and the gamma function.

Fact 1 (Power Law Distribution)**.** *Suppose that $\{M_i\}_{i=1}^n$ are IID draws from a power law distribution with parameter a, i.e. $P(M_i \leq x) = 1 - x^{-a} = F(x)$ for $x \in [1, \infty)$. Let $M_{(j)}$ be the j^{th} order statistic of the M_i. Then*

1. *For any $r \geq 1$, $E[M_i | M_i > r] = rE[M_i]$.*

2. *$M_{(1)}/M_{(n)}, M_{(2)}/M_{(n)}, \ldots, M_{(n-1)}/M_{(n)}$ are independent from $M_{(n)}$ and are distributed as the order statistics of $M_1, \ldots, M_{n-1}$.*

3. $E[M_{(1)}] = \Gamma(1-1/a)\Gamma(n+1)/\Gamma(n+1-1/a)$.

Fact 2 (Gamma Function)**.**

1. *For any* $s > 0$, $\Gamma(s+1) = s\Gamma(s)$.
2. *For any* $s > 0$, $\lim_{n\to\infty} \left(\frac{n^{-s}\,\Gamma(n+1)}{\Gamma(n+1-s)}\right) = 1$.

Lemma 3. *If match values are independent draws from a power law distribution with parameter* a*, then for any* $\alpha \geq 1$,

$$V_P(MSB(\alpha)) = \alpha^{1-a} E[X_{(1)}].$$

Proof.

$$\begin{aligned}
V_P(MSB(\alpha)) &= E\left[X_{(2)}\frac{M_{(1)}}{M_{(2)}}\mathbf{1}_{\frac{M_{(1)}}{M_{(2)}}>\alpha}\right] \\
&= E[X_{(2)}]E\left[\frac{M_{(1)}}{M_{(2)}}\mathbf{1}_{\frac{M_{(1)}}{M_{(2)}}>\alpha}\right] \\
&= E[X_{(2)}]E\left[\frac{M_{(1)}}{M_{(2)}}\right]\alpha P\left(\frac{M_{(1)}}{M_{(2)}} > \alpha\right) \\
&= E\left[X_{(2)}\frac{M_{(1)}}{M_{(2)}}\right]\alpha P\left(\frac{M_{(1)}}{M_{(2)}} > \alpha\right) \\
&= E[X_{(1)}]\alpha^{1-a}
\end{aligned}$$

The first line uses the fact that $X_{(1)}/X_{(2)} = M_{(1)}/M_{(2)}$. The second and fourth lines use the independence of $M_{(2)}$ and $M_{(1)}/M_{(2)}$ established by Fact 1.2. The third and final lines use the fact that $M_{(1)}/M_{(2)}$ follows a power law distribution; the third line also applies Fact 1.1.

□

Proof of Theorem 2. For notational simplicity, we treat the case where $\gamma = 1$. Other values of γ follow by an identical argument, as changing γ is effectively equivalent to rescaling the brand advertiser's average match value μ.

Both OMN and MSB have allocation rules that are independent of C, so it is clear that the distribution of C does not matter. For simplicity, in this proof we take C to be identically one. This leaves us with three parameters of interest: the number of performance bidders n, the average value of the brand advertiser $\mu = E[M_0]$, and the weight of the power law tail, a. As above, we define λ to be probability that the brand advertiser wins the impression under OMN, and use $\mu(\lambda, n)$ to refer to the brand value implied by the given values of λ and n (for fixed a), so $\lambda = P(M_{(1)} \leq \mu(\lambda, n))$.

The omniscient benchmark achieves total surplus given by

$$\begin{aligned} V(OMN) &= V_B(OMN) + V_P(OMN) \\ &= \lambda\mu(\lambda, n) + \int_\lambda^1 \mu(x, n)dx. \end{aligned} \tag{3}$$

Meanwhile, for any $\alpha \geq 1$,

$$\begin{aligned} V(MSB(\alpha)) &= V_B(MSB(\alpha)) + V_P(MSB(\alpha)) \\ &= P(M_{(1)} \leq \alpha M_{(2)})\mu(\lambda, n) + V_P(MSB(\alpha)) \\ &= (1 - \alpha^{-a})\mu(\lambda, n) + \alpha^{1-a}E[M_{(1)}], \end{aligned} \tag{4}$$

where the final line follows from Fact 1.2 and Lemma 3.

We choose the MSB parameter α such that the brand advertiser is awarded the impression with probability λ. In other words, we select α such that $1 - \alpha^{-a} = \lambda$. Because both allocation rules deliver a representative sample of impressions to the brand advertiser, the first statement in Theorem 2 follows immediately. In other words, our choice of α ensures that $V_B(OMN) = V_B(MSB(\alpha))$.

Of course, the value of impressions allocated to performance advertisers will be lower under MSB than under OMN. We establish in Lemma 4 that for fixed λ and a, the ratio $V_P(MSB(\alpha))/V_P(OMN)$ is decreasing in n. Applying Lemma 2, we see that

$$\lim_{n\to\infty} n^{-1/a}\, V_P(OMN; n, \mu(\lambda, n)) = \lim_{n\to\infty} n^{-1/a} \int_\lambda^1 (1 - x^{1/n})^{-1/a}dx = \int_\lambda^1 \log(1/x)^{-1/a}dx. \tag{5}$$

By Lemma 3 and Facts 1.3 and 2.2, we see that

$$\lim_{n\to\infty} n^{-1/a}V_P(MSB(\alpha); n, \mu(\lambda, n)) = \lim_{n\to\infty} n^{-1/a}\alpha^{1-a}E[M_{(1)}] = \alpha^{1-a}\Gamma(1 - 1/a). \tag{6}$$

Lemma 5 establishes that the ratio $V_P(MSB(\alpha))/V_P(OMN)$ worsens as $\lambda \to 1$. Taking $\lambda \to 1$ and applying L'Hospital's rule, we see that

$$\begin{aligned} \lim_{\lambda\to 1} \frac{(1-\lambda)^{1-1/a}\Gamma(1 - 1/a)}{\int_\lambda^1 \log(1/x)^{-1/a}dx} &= (1 - 1/a)\Gamma(1 - 1/a) \lim_{\lambda\to 1} \frac{(1-\lambda)^{-1/a}}{\log(1/\lambda)^{-1/a}} \\ &= \Gamma(2 - 1/a). \end{aligned}$$

where the final line follows from the identity $\Gamma(s + 1) = s\Gamma(s)$ and the fact that $\lim_{\lambda\to 1}(1 - \lambda)/\log(1/\lambda) = 1$. Because $a > 1$, we have $2 - 1/a \in (1, 2)$. The minimum of the gamma function over the interval $(1, 2)$ exceeds 0.885, completing the proof of the second claim in Theorem 2.

We now turn our attention to the third claim. We show in Lemma 6 that $V(MSB(\alpha))/V(OMN)$

is decreasing in n. We compute that

$$\lim_{n\to\infty} n^{-1/a}\mu(\lambda, n) = \log(1/\lambda)^{-1/a}. \tag{7}$$

Combining this with (3), (4), (5) and (6), we conclude that

$$\lim_{n\to\infty} n^{-1/a}\ V(MSB(\alpha); n, \mu(\lambda, n)) = \lambda\log(1/\lambda)^{-1/a} + \Gamma(1-1/a)(1-\lambda)^{1-1/a}.$$

$$\lim_{n\to\infty} n^{-1/a}\ V(OMN; n, \mu(\lambda, n)) = \lambda\log(1/\lambda)^{-1/a} + \int_\lambda^1 \log(1/x)^{-1/a}dx.$$

It follows that the ratio between these expressions is a lower bound on $V(MSB(\alpha))/V(OMN)$.

The minimum of this lower bound for $\lambda \in (0,1)$ and $1/a \in (0,1)$ exceeds 0.948, completing the proof of the third claim. □

Lemma 4. *Suppose that the M_i are IID draws from a power law distribution with parameter a. Fix $\lambda \in (0,1)$ and let $\alpha = (1-\lambda)^{-1/a}$, so that $MSB(\alpha)$ and OMN sell the impression to the brand advertiser with equal probability. Then $\frac{V_P(MSB(\alpha);n,\mu(\lambda,n))}{V_P(OMN);n,\mu(\lambda,n)}$ is decreasing in n.*

Proof of Lemma 4. For this proof only, we adopt additional notation to indicate the number of bidders. We fix the match value distribution, let $E_n[\cdot]$ denote the expectation of its argument conditioned on $N = n$, and let $P_n(\cdot)$ denote the probability of the argument given $N = n$.

Fix $\lambda \in (0,1)$ and $a > 1$, and let $\alpha = (1-\lambda)^{-1/a}$. Note that Fact 1.2 implies that when $N = n+1$, the values $R_i = M_{(i)}/M_{(n+1)}$ for $i = 1, \ldots, n$ are distributed as the order statistics of n iid draws from a power law distribution with parameter a, and are independent from $M_{(n+1)}$. Thus,

$$\begin{aligned} V_P(MSB(\alpha); n+1) &= E_{n+1}\left[M_{(n+1)}R_1\mathbf{1}_{\frac{R_1}{R_2}>\alpha}\right] \\ &= E_{n+1}\left[M_{(n+1)}\right] E_{n+1}\left[R_1\mathbf{1}_{\frac{R_1}{R_2}>\alpha}\right] \\ &= E_{n+1}\left[M_{(n+1)}\right] V_P(MSB(\alpha); n). \end{aligned}$$

The second line follows from the independence of $M_{(1)}/M_{(2)}$ from $M_{(n+1)}$ and the fact that $R_1/R_2 = M_{(1)}/M_{(2)}$, while the final line follows from Fact 1.2. Thus, to prove the lemma, it suffices to show that for any $n \geq 2$,

$$V_P(OMN; n+1) \geq E_{n+1}[M_{(n+1)}]V_P(OMN; n),$$

We do this by considering an allocation rule z such that $V_P(z; n+1) = E_{n+1}[M_{(n+1)}]V_P(OMN; n)$. When $N = n+1$, this rule uses the ratio $R_1 = M_{(1)}/M_{(n+1)}$ to determine how to allocate the impression: it goes to the top performance advertiser whenever R_1 exceeds $\mu(\lambda, n)$. Note that Fact

1.2 implies that $P_{n+1}(R_1 \le \mu(\lambda, n)) = P_n(M_{(1)} \le \mu(\lambda, n))$, so this auction allocates the impression to the brand advertiser with the same probability λ as under OMN. It follows that

$$\begin{aligned} V_P(OMN; n+1) &\ge E_{n+1}[M_{(n+1)} R_1 \mathbf{1}_{R_1 > \mu(\lambda,n)}] \\ &= E_{n+1}[M_{(n+1)}] E_{n+1}[R_1 \mathbf{1}_{R_1 > \mu(\lambda,n)}] \\ &= E_{n+1}[M_{(n+1)}] V_P(OMN; n), \end{aligned}$$

completing the proof. □

Lemma 5. *Fix $n \ge 2$ and suppose that $P(N = n) = 1$ and match values are drawn independently from a power law distribution with parameter a. If $\alpha = (1-\lambda)^{-1/a}$, then $\frac{V_P(MSB(\alpha);\mu(\lambda,n))}{V_P(OMN;\ \mu(\lambda,n))}$ is decreasing in λ.*

Proof. We will prove the equivalent statement that the log of this ratio is decreasing. Lemmas 2 and 3 establish that $V_P(OMN) = E[C] \int_\lambda^1 \mu(x, n)dx$ and $V_P(MSB(\alpha)) = (1-\lambda)^{1-1/a} E[X_{(1)}]$. It follows that

$$\frac{d}{d\lambda} \log(V_P(MSB(\alpha))) - \frac{d}{d\lambda} \log V_P(OMN) = \frac{-1}{1-\lambda} + \frac{\mu(\lambda, n)}{\int_\lambda^1 \mu(x, n)dx.}.$$

Because $\mu(x, n)$ is increasing in x, $\int_\lambda^1 \mu(x, n)dx > (1-\lambda)\mu(\lambda, n)$, proving that the expression above is negative. □

Lemma 6. *Fix $\lambda \in (0, 1)$ and $a > 1$, and let $\alpha = (1-\lambda)^{-1/a}$. Suppose that $N = n$, and that match values are drawn iid from a power law distribution with parameter a. Then the ratio $\frac{V(MSB(\alpha);\mu(\lambda,n),n)}{V(OMN;\mu(\lambda,n),n)}$ is decreasing in n.*

Proof. Note that for any allocation rules A and A', it is possible to express the ratio of total value as a convex combination of the ratio of brand value and the ratio of performance value:

$$\frac{V(A)}{V(A')} = \frac{V_B(A')}{V(A')} \cdot \frac{V_B(A)}{V_B(A')} + \frac{V_P(A')}{V(A')} \cdot \frac{V_P(A)}{V_P(A')}, \tag{8}$$

Fix λ and a, and let $\alpha = (1-\lambda)^{-1/a}$, so that the brand advertiser is equally likely to win the impression under $MSB(\alpha)$ and OMN. Letting $A = MSB(\alpha)$ and $A' = OMN$ above, we must show that for fixed λ and a, the relative performance of MSB, as given in (8), is decreasing in n.

We know that $V_B(MSB(\alpha)) = V_B(OMN)$, and the first part of this Lemma establishes that the ratio $V_P(MSB(\alpha))/V_P(OMN)$ is less than one and decreasing in n. Thus, it suffices to show that the ratio $V_P(OMN)/V(OMN)$ is increasing in n (fixing λ and allowing $\mu = \mu(\lambda, n)$ to vary), or equivalently that $V_P(OMN; n, \mu(\lambda, n))/V_B(OMN; n, \mu(\lambda, n))$ is increasing in n.

Lemma 2 states that $V_P(OMN) = \int_\lambda^1 \mu(x,n)dx$, and $V_B(OMN) = \lambda\mu(\lambda,n)$. It follows that

$$\frac{V_P(OMN)}{V_B(OMN)} = \frac{1}{\lambda}\int_\lambda^1 \frac{\mu(x,n)}{\mu(\lambda,n)}dx.$$

Suppose that $n' > n$. We claim that $\frac{\mu(x,n')}{\mu(x,n)}$ is increasing in x. From this, it follows that

$$\int_\lambda^1 \frac{\mu(x,n')}{\mu(\lambda,n')}dx = \int_\lambda^1 \frac{\mu(x,n')}{\mu(x,n)}\frac{\mu(x,n)}{\mu(\lambda,n')}dx \geq \int_\lambda^1 \frac{\mu(\lambda,n')}{\mu(\lambda,n)}\frac{\mu(x,n)}{\mu(\lambda,n')}dx = \int_\lambda^1 \frac{\mu(x,n)}{\mu(\lambda,n)}dx.$$

All that remains is to prove our claim that $\frac{\mu(x,n')}{\mu(x,n)}$ is increasing in x. Note that

$$\frac{d}{dx}\frac{\mu(x,n')}{\mu(x,n)} > 0 \Leftrightarrow \frac{\frac{d}{dx}\mu(x,n')}{\mu(x,n')} - \frac{\frac{d}{dx}\mu(x,n)}{\mu(x,n)} > 0.$$

Thus, it suffices to show that $\frac{d}{dx}\log(\mu(x,n))$ is increasing in n. We compute

$$\frac{\frac{d}{dx}\mu(x,n)}{\mu(x,n)} = \frac{1}{ax}\frac{\frac{1}{n}x^{1/n}}{(1-x^{1/n})} = \frac{1}{axn(x^{-1/n}-1)}.$$

Making the substitution $z = 1/n$, we see that the above expression is increasing in n if and only if $(x^{-z}-1)/z$ is increasing in z. But

$$\begin{aligned}\frac{d}{dz}\frac{x^{-z}-1}{z} &= \frac{1}{z^2}\left(-z\log(x)x^{-z} - (x^{-z}-1)\right)\\ &= \frac{1}{z^2}(x^{-z}(\log x^{-z} - 1) + 1).\end{aligned}$$

To see that this is non-negative, let $y = x^{-z}$. The minimum of $y(\log y - 1) + 1$ is at $y = 1$, when the value of the expression is zero. □

We now turn our attention to Corollary 1. Lemma 7 establishes that for any dominant strategy incentive compatible mechanism (z,p), revenue from performance advertisers is at most $(1-a^{-1})V_P(z)$, and that MSB auctions achieve this bound. If the publisher gets a fraction δ of the surplus from ads assigned to the brand advertiser, it follows that the revenue from the optimal mechanism is at most $\sup_z \delta V_B(z) + (1-a^{-1})V_P(z)$; by Theorem 2, a suitably-chosen MSB auction gets at least 94.8% of this benchmark.

Lemma 7. *Suppose that (z,p) is a dominant-strategy mechanism in which $x_i = 0$ implies $p_i(x) = z_i(x) = 0$. If M_i is drawn from a power law distribution with parameter a, then $E[p_i(X)] \leq (1-a^{-1})E[X_i z_i(X)]$, with equality if (z,p) corresponds to an MSB auction.*

Proof. It is well-known that if the mechanism is dominant strategy incentive compatible for bidder i, then from this bidder's perspective, the mechanism makes a single take-it-or-leave-it offer. For any offer price $\hat{p}$ (which may depend arbitrarily on others' bids), we consider two cases:

1. $C = c > \hat{p}$. In this case, because $X_i > C$, bidder i wins the impression, receives an expected value of $E[cM_i] = c/(1 - a^{-1})$, and pays $\hat{p}$, which is less than $(1 - a^{-1})$ times their expected value.

2. $C = c < \hat{p}$. In this case, bidder i wins the impression whenever $M_i > \hat{p}/c$, and conditional on winning, has an expected value of $c\,(\hat{p}/c)\frac{a}{a-1} = \hat{p}/(1 - a^{-1})$ (by Fact 1.1). Bidder i pays exactly $\hat{p}$ upon winning, implying that in this case, expected publisher revenues (from bidder i) are exactly $(1 - a^{-1})$ times expected total surplus (from bidder i).

Under an MSB auction, the threshold $\hat{p} = \alpha \max X_{-i} > \alpha C$ (since each $X_j > C$), so the first case above never occurs, and thus revenue and surplus from bidder i are proportional.

□