Online Appendix: Adverse Selection and Auction Design for Internet Display Advertising

By Nick Arnosti and Marissa Beck and Paul Milgrom

Note that for any allocation rule z, the quantities $V_B(z)$, $V_P(z)$ and $V(z)$ depend implicitly on the number of bidders and the distributions of advertiser values. At times, we study the variation in the performance of an allocation rule z, as a function of an underlying parameter θ (such as the number of bidders or the distribution from which their values are drawn). In some cases, we make the dependence on θ explicit by writing $V(z;\theta)$. 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:

We must show that second price auctions cannot guarantee $(\frac{1}{2}+\epsilon)V(\text{OMN})$, for any $\epsilon > 0$. Fix $N \geq 2$ and $\epsilon > 0$. To do so, we assume that M_i are iid draws from a power law distribution with parameter a, and maintain the identity $\gamma\mu = \gamma E[M_0] = (1+\epsilon)E[M_{(1)}]$. 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,

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

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

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 $\gamma\mu = \gamma E[M_0] > E[M_{(1)}]$. Suppose that $C \in [1,\infty)$ has density $g(c) = a'c^{-a'-1}$. Then there exists $\delta > 0$ such that if $a' < 1+\delta$,*

$$\sup_b V(\text{SP}_b) = \gamma\mu E[C].$$

PROOF OF LEMMA 1:

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

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

From this, if follows that

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

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 $\int_0^b (\gamma\mu - u)f(u)u^{a'-1}du$ is (weakly) increasing on $[0, \gamma\mu]$ and (weakly) decreasing thereafter. Thus, it is enough to show that when a' is sufficiently small, $\int_0^\infty (\gamma\mu - u)f(u)u^{a'-1}du > 0$.

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

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

PROOF OF PROPOSITION 2:

For a reminder of the theory of sufficient statistics, see *Theory of Point Estimation* by Lehmann and Casella. We begin by establishing that $(N, X_{(N)})$ is a sufficient statistic for C (and thus for X_0). Given C and N, the conditional density of X is

$$f(X; C, N) = a^N C^{aN} \mathbf{1}_{X_{(N)} \ge C} \cdot \left(\prod_{i=1}^N X_i \right)^{-a-1},$$

so $X_{(N)}$ is a sufficient statistic for C.

We now show that $E[C|N, X_{(N)}]$ is non-decreasing in $X_{(N)}$. The notation that follows assumes that C follows an atomless distribution with density g, although the argument can be extended to the case where the distribution of C has atoms.

The conditional density of C, given $(N, X_{(N)})$, is proportional to $g(c)c^{aN+1}$ on

$[1, X_{(N)}]$. Thus,

$$E[C|N, X_{(N)} = x] = \frac{\int_0^x g(c)c^{aN+2}dc}{\int_0^x g(c)c^{aN+1}dc}. \tag{1}$$

We wish to prove that this expression is non-decreasing in x. Its derivative with respect to x is

$$\frac{g(x)x^{aN+2}\int_0^x g(c)c^{aN+1}dc - g(x)x^{aN+1}\int_0^x g(c)c^{aN+2}dc}{\left(\int_0^x g(c)c^{aN+1}dc\right)^2},$$

which can be rewritten as

$$\frac{g(x)x^{aN+1}}{\int_0^x g(c)c^{aN+1}dc}\left(x - E[C|N, X_{(N)} = x]\right).$$

The first term is clearly non-negative, as is the second term (because $M_{(N)} \geq 1$ and $X_{(N)} = CM_{(N)} \geq C$).

We now turn to the second point in the proposition, which states that under OPT, $E[z_0(X)|C, N]$ is decreasing in C whenever C follows a power law distribution.

We begin with a series of claims that hold whenever the match values follow a power law distribution (regardless of the distribution of C). First, we claim that

$$E[z_0(X)|N, X_{(N)}] = E[z_0(X)|C, N, X_{(N)}], \tag{2}$$

that is, given the values of N and $X_{(N)}$, the value of C does not affect the probability that the impression is awarded to the brand advertiser. To see this, recall that under OPT,

$$z_0(X) = \mathbf{1}_{X_{(1)} \leq \gamma E[X_0|X]} = \mathbf{1}_{\frac{X_{(1)}}{X_{(N)}} \leq \gamma E[M_0] \cdot \frac{E[C|N, X_{(N)}]}{X_{(N)}}}, \tag{3}$$

where the first equality holds by definition and the second makes use of the fact that $(N, X_{(N)})$ is a sufficient statistic for C. Clearly, the distribution of $X_{(1)}/X_{(N)} = M_{(1)}/M_{(N)}$ is independent from C, implying that (2) holds. From this and the definition of conditional expectation, it follows that

$$\begin{aligned} E[z_0(X)|C, N] &= E[E[z_0(X)|C, N, X_{(N)}]|C, N]. \\ &= E[E[z_0(X)|N, X_{(N)}]|C, N]. \end{aligned}$$

Because the conditional distribution of $X_{(N)}$ given C, N is stochastically increasing in C, to show that $E[z_0(X)|C, N]$ is decreasing in C, it suffices to show that

$E[z_0(X)|N, X_{(N)}]$ is decreasing in $X_{(N)}$. From (3), we see that

$$E[z_0(X)|N, X_{(N)}] = P\left(\frac{X_{(1)}}{X_{(N)}} \leq \gamma E[M_0] \cdot \frac{E[C|N, X_{(N)}]}{X_{(N)}}\middle| N, X_{(N)}\right).$$

Because $X_{(1)}/X_{(N)}$ is ancillary for C, and $X_{(N)}$ is sufficient for C, Basu's theorem implies that $X_{(1)}/X_{(N)}$ is conditionally independent from $X_{(N)}$, given N. Therefore, in order to show that the quantity $E[z_0(X)|N, X_{(N)}]$ is decreasing in $X_{(N)}$, it is enough to show that the ratio $E[C|N, X_{(N)} = x]/x$ is decreasing in x.

Here, for the first time, we invoke the assumption that C follows a power law distribution – that is, that $g(x) = bx^{-b-1}$ on $[1, \infty)$ for some $b > 1$. Define $\beta = aN - b + 2$. By (1), we have that

$$\frac{1}{x}E[C|N, X_{(N)} = x] = \begin{cases} \frac{\log(x)}{x-1} & \beta = 0 \\ \frac{x-1}{x\log(x)} & \beta = 1 \\ \frac{\beta-1}{\beta} \cdot \frac{x^\beta - 1}{x^\beta - x} & \beta \notin \{0, 1\} \end{cases}.$$

In what follows, we assume $\beta \notin \{0, 1\}$; similar arguments establish monotonicity of the expressions corresponding to $\beta = 0$ and $\beta = 1$. Differentiating with respect to x, we see that

$$\begin{aligned} \frac{d}{dx}\frac{1}{x}E[C|N, X_{(N)} = x] &= \frac{\beta-1}{\beta}(x^\beta - x)^{-2}\left((x^\beta - x)(\beta x^{\beta-1}) - (x^\beta - 1)(\beta x^{\beta-1} - 1)\right). \\ &= \frac{\beta-1}{\beta}(x^\beta - x)^{-2}\left(\beta x^{\beta-1} + x^\beta - \beta x^\beta - 1\right). \end{aligned}$$

We must show that for $x > 1$, this expression is negative. Its sign is determined by the sign of

$$\frac{\beta-1}{\beta}\left(\beta x^{\beta-1} + x^\beta - \beta x^\beta - 1\right),$$

which takes the value zero at $x = 1$. Thus, it is enough to show that this quantity is non-increasing in x for $x \geq 1$. Differentiating, we get

$$\frac{\beta-1}{\beta}\frac{d}{dx}\left(\beta x^{\beta-1} + x^\beta - \beta x^\beta - 1\right) = (\beta-1)^2\, x^{\beta-2}\,(1-x) \leq 0.$$

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 claim that if the mechanism is false-name proof, then $h(x_{-i}) = h(\max\{x_{-i}\})$. Suppose that 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 reduce its price by submitting the remaining bids in the profile x_{-i}, so the mechanism is not winner false-name proof. If $h(x_{-i}) > h(\max\{x_{-i}\})$, then the losing bidder can raise the winner's price by submitting the remaining bids in the profile x_{-i}, so the mechanism is not loser false-name proof.

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 3:

Note that $\alpha = 1$ and $\alpha = \infty$ describe the cases where the brand advertiser never wins or always wins, so $V(\mathrm{MSB}_1) = E[X_{(1)}]$ and $V(\mathrm{MSB}_\infty) = \gamma E[X_0]$. Furthermore,

$$\begin{aligned} V(\mathrm{OMN}) &= E[\max(\gamma C E[M_0], C M_{(1)})] \\ &\leq \gamma E[X_0] + E[X_{(1)}] \\ &\leq 2 \cdot \max(\gamma E[X_0], E[X_{(1)}]), \end{aligned}$$

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 $\max(X_{(1)}, 1)$. It follows that

$$E[X_0|X] = E[\mu C|X] = \frac{\beta + b}{\beta + b - 1}\mu \max(X_{(1)}, 1).$$

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

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

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

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

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

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

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

$$\lambda = F(\mu(\lambda, n))^n, \tag{4}$$

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). For notational simplicity, we treat the case where $\gamma = 1$. Other values of γ follow identically, as changing γ is effectively equivalent to rescaling the brand advertiser's average match value μ.

We begin with a technical lemma, which allows us to compute $V_P(\text{OMN})$, 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 (4). Then*

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

PROOF OF LEMMA 2:

Differentiating the identity $F(\mu(\lambda,n))^n = \lambda$, we obtain

$$nF(\mu(\lambda,n))^{n-1}f(\mu(\lambda,n))\frac{d}{d\lambda}\mu(\lambda,n) = 1. \tag{5}$$

Therefore,

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

where the final line follows from application of (5). The Lemma follows immediately.

Our proof of Theorem 2 references the gamma function Γ, defined by

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx.$$

We make use of the following facts.

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. $P(M_i \le x) = 1 - x^{-a} = F(x)$ for $x \in [1,\infty)$. Let $M_{(j)}$ be the j^{th} order statistic of the M_i. Then*

1) *For any $r \ge 1$, $E[M_i|M_i > r] = rE[M_i]$.*
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 $s > 0$, $\lim\limits_{n\to\infty}\left(\frac{n^{-s}\,\Gamma(n+1)}{\Gamma(n+1-s)}\right) = 1$.*

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

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

PROOF OF LEMMA 3:

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

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:

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

$$\begin{aligned}
V(\mathrm{OMN}) &= V_B(\mathrm{OMN}) + V_P(\mathrm{OMN}) \\
&= \lambda\mu(\lambda, n) + \int_\lambda^1 \mu(x, n) dx. \qquad (6)
\end{aligned}$$

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

$$\begin{aligned} V(\mathrm{MSB}_\alpha) &= V_B(\mathrm{MSB}_\alpha) + V_P(\mathrm{MSB}_\alpha) \\ &= P(M_{(1)} \leq \alpha M_{(2)})\mu(\lambda, n) + V_P(\mathrm{MSB}_\alpha) \\ &= (1-\alpha^{-a})\mu(\lambda, n) + \alpha^{1-a} E[M_{(1)}], \end{aligned} \tag{7}$$

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 $V_B(\mathrm{OMN}) = V_B(\mathrm{MSB}_\alpha)$.

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 $V_P(\mathrm{MSB}_\alpha)/V_P(\mathrm{OMN})$ is decreasing in n. Applying Lemma 2, we see that

$$\begin{aligned} \lim_{n\to\infty} n^{-1/a}\, V_P(\mathrm{OMN}; n, \mu(\lambda, n)) &= \lim_{n\to\infty} n^{-1/a} \int_\lambda^1 (1-x^{1/n})^{-1/a} dx \\ &= \int_\lambda^1 \log(1/x)^{-1/a} dx. \end{aligned} \tag{8}$$

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

$$\lim_{n\to\infty} n^{-1/a} V_P(\mathrm{MSB}_\alpha; n, \mu(\lambda, n)) = \lim_{n\to\infty} n^{-1/a}\alpha^{1-a} E[M_{(1)}] = \alpha^{1-a}\Gamma(1-1/a). \tag{9}$$

Lemma 5 establishes that the ratio $V_P(\mathrm{MSB}_\alpha)/V_P(\mathrm{OMN})$ 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 $V(\mathrm{MSB}_\alpha)/V(\mathrm{OMN})$ is decreasing in n. We compute that

$$\lim_{n\to\infty} n^{-1/a}\mu(\lambda, n) = \log(1/\lambda)^{-1/a}. \tag{10}$$

Combining this with (6), (7), (8) and (9), we conclude that

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

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

Thus, the ratio of these expressions is a lower bound on $V(\mathrm{MSB}_\alpha)/V(\mathrm{OMN})$.

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_α *and* OMN *sell the impression to the brand advertiser with equal probability. Then* $\frac{V_P(\mathrm{MSB}_\alpha; n, \mu(\lambda,n))}{V_P(\mathrm{OMN}); n, \mu(\lambda, n)}$ *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,

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

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 $n \geq 2$,

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

We do this by considering an allocation rule z such that

$$V_P(z; n+1) = E_{n+1}[M_{(n+1)}] V_P(\mathrm{OMN}; n).$$

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 $P_{n+1}(R_1 \leq \mu(\lambda, n)) = P_n(M_{(1)} \leq$

$\mu(\lambda, n))$, so this auction allocates the impression to the brand advertiser with the same probability λ as under OMN. It follows that

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

completing the proof.

LEMMA 5: *Fix $n \geq 2$ 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 $\frac{V_P(\text{MSB}_\alpha; \mu(\lambda,n))}{V_P(\text{OMN};\ \mu(\lambda,n))}$ is decreasing in λ.*

PROOF OF LEMMA 5:

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

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

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 $\frac{V(\text{MSB}_\alpha; \mu(\lambda,n), n)}{V(\text{OMN}; \mu(\lambda,n), n)}$ is decreasing in n.*

PROOF OF LEMMA 6:

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:

$$\frac{V(A)}{V(A')} = \frac{V_B(A')}{V(A')} \cdot \frac{V_B(A)}{V_B(A')} + \frac{V_P(A')}{V(A')} \cdot \frac{V_P(A)}{V_P(A')}, \tag{11}$$

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_α and OMN. Letting $A = \text{MSB}_\alpha$ and $A' = \text{OMN}$ above, we must show that for fixed λ and a, the relative performance of MSB, as given in (11), is decreasing in n.

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

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

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

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 OF LEMMA 7:

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 its 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.