Display Advertising Auctions with Arbitrage*

Ruggiero Cavallo
Yahoo
cavallo@yahoo-inc.com

R. Preston McAfee
Microsoft
preston@mcafee.cc

Sergei Vassilvitskii
Google
sergeiv@google.com

Submitted: July 12, 2012
Revised: September 5, 2014

Abstract

Online display advertising exchanges connect web publishers with advertisers seeking to place ads. In many cases the advertiser obtains value from an ad impression (a viewing by a user) only if it is clicked, and frequently advertisers prefer to pay contingent on this occurring. But at the same time many publishers demand payment independent of click. Arbitragers with good estimates of click-probabilities can resolve this conflict by absorbing the risk and acting as intermediary, paying the publisher on allocation and being paid only if a click occurs. This paper examines the incentives of advertisers and arbitragers and contributes an efficient mechanism with truthful bidding by the advertisers and truthful reporting of click predictions by arbitragers as dominant strategies while, given that a hazard rate condition is satisfied, yielding increased revenue to the publisher. We provide empirical evidence based on bid data from Yahoo's Right Media Exchange suggesting that the mechanism would increase revenue in practice.

1 Introduction

Advertising exchanges, such as Yahoo's Right Media, Google's AdX, AppNexus and OpenX, are an increasingly common and important mechanism for allocating display advertising. In these exchanges, an advertiser such as an automobile company will bid to put its ad on web pages, such as those from news outlets or car magazines. The particular mechanisms used by these exchanges vary. In all cases, however, a loading web page calls an exchange for an ad, the exchange then holds an automated auction, chooses the highest eligible bidder, and assigns that bidder's ad to the web page. Confusingly, both the call for the ad and the ad on the page are known as an "impression." Many advertisers pay by the impression (cost per impression or "CPM"), and generally the publisher (web page) would prefer to be paid by the impression. However, advertisers often prefer to pay only when their ad is clicked, known as CPC or cost per click payments.

The preference of advertisers to pay for clicks conflicts with the desire of publishers to be paid by the impression. Advertisers want to pay by the click because they cannot

*This research was done while all three authors were employed at Yahoo Research.

observe the quality of the inventory they are buying and are suspicious that they will be charged for low quality, non-converting impressions; they only want to pay for the quality they receive.[1] Similarly, publishers want to ensure they receive payment for the quality they provide. For uninformed publishers the restriction to CPM payment is fundamental: without having good estimates of click-probabilities for the various advertisers, a publisher cannot even identify a sensible CPC payment scheme (what should the per-click prices be?).

Some advertising exchanges, such as Yahoo's Right Media Exchange (RMX), permit advertisers to bid either on a cost per impression or a cost per click basis. In order to compare a cost per impression bid to a cost per click bid, a forecast of the probability of a click is needed. RMX itself performs such forecasts, but could an exchange practically use a third party to perform the forecasting? Such questions arise when a third party is better at forecasting than other participants. Indeed, third party companies have emerged to effectively bridge the gap between publishers and advertisers, guaranteeing clicks to advertisers while making per impression payments to the publishers. In this work we propose mechanisms for bringing the click-prediction expertise of these companies directly into the exchange by abstracting them as arbitrage agents. The arbitrager forecasts the probability of a click, which allows the construction of an expected or average price per impression, namely the bid times the probability of a click. The arbitrager is then paid when the ad is clicked and pays the publisher some amount per impression.[2] We adopt the perspective of the exchange designer who wishes to implement efficient allocations, and provide mechanisms that establish incentives for all parties—arbitragers, publishers and advertisers—to participate and bid truthfully, leading to *maximal* exchange efficiency.

The presence of arbitragers presents an important opportunity and also a challenge. If an arbitrager is the only party with knowledge of the true click probability for an ad, then achieving an efficient allocation depends critically on his participation. The associated challenge is that, in order to create appropriate incentives for the arbitrager's participation, the mechanism must yield him at least as much expected value as he could achieve by directly buying from the publisher and selling to the CPC advertiser (this is his "outside option"). The case where the arbitrager bypasses the mechanism—which is possible if the arbitrager comes with his own data and does not need any information from the exchange in order to do his prediction—would inevitably lead to some loss, with the publisher and advertiser seeking to set prices and bids in a way that minimizes the surplus extracted by the arbitrager, sometimes leading to inefficient allocation.

Given the Myerson-Satterthwaite Theorem [Myerson and Satterthwaite, 1983], which considers a bilateral exchange problem with private information about cost by

[1] Some advertisers prefer to pay only when they make a sale or some other action is taken, known as cost per action or CPA, as sales are an even more relevant outcome. For convenience we frame the environment in terms of CPC-demanding advertisers, but the same analysis holds for advertisers who are only willing to pay CPA, assuming actions are observable to the publisher.

[2] We describe these third-parties as *arbitragers* since they profit from buying and selling in cases where a trade will yield efficiency but has not taken place. An analogy to arbitrage in financial markets is stretched slightly by the fact that in our setting the trade cannot take place without the arbitrager, since he provides information that is critical to the process.

the seller and about value by the buyer, it would be surprising if efficient mechanisms for handling arbitrage exist. That theorem states that efficient exchange cannot be arranged by any mechanism respecting incentive compatibility and individual rationality. It has widely been interpreted to state that efficient exchange is impossible with two-sided private information. However, we show that when the common hazard rate assumption holds, we can construct a mechanism that provides dominant strategies to both the advertiser and the arbitrager (both of whom hold critical private information), and which yields an expected revenue increase for the publisher over not accepting the arbitrager's bid. This mechanism is straightforward to implement and does not depend on specific details of the distribution of values. As such, it is a natural candidate for use in an exchange. The mechanism's simplicity is important because many billions of auctions are carried out daily.

After providing an outline of related work, we present the main results of the paper in Section 2: we provide an efficient mechanism and analyze the revenue and efficiency it yields in comparison with a CPC-bidder-excluding approach. In Section 3 we present two alternative mechanisms that are inefficient but have the advantage of always (weakly) increasing revenue and efficiency over the CPC-bidder-excluding solution, which is important for settings where the publisher demands *in every instance* the baseline revenue of the CPC-bidder-excluding solution and there is no third-party (such as the exchange itself) willing to take on the risk associated with click uncertainty. In Section 4 we generalize the model and results with respect to the number of CPC-bidders and arbitragers. In Section 5 we use real bid data from Yahoo's Right Media Exchange to evaluate, retrospectively and hypothetically, the revenue and efficiency impact of moving to our efficient mechanism. In Section 6 we conclude. All proofs are provided in an Appendix.

1.1 Related work

There is an extensive theory of arbitrage in asset markets, see e.g., [Merton, 1987] or [Fama and French, 2004] for an overview. Arbitrage in asset markets is generally risk-free, where the present analysis involves risk-shifting. There are also extensive analyses of risk-shifting based on risk aversion, e.g., [Arrow, 1971]. We study risk-shifting based on expertise—a party able to forecast an event more accurately is the natural candidate for taking the risk of the action. Since in our setting the arbitrager's most critical role is to provide private click-probability information, there is also a connection to scoring rules and other belief-elicitation mechanisms (see, e.g., [Savage, 1971], [de Finetti, 1974], and [Karni, 2009]), though our problem has added constraints due to arbitragers' "outside option" of bypassing the mechanism and directly buying and selling instead.

We focus specifically on display advertising on the Internet. An increasing share of such ad impressions are sold via ad exchanges, which facilitate the sale of billions of impressions per day; see the survey by Muthukrishnan [2009] for a complete description of advertising exchanges. Previous work has considered aspects of the mechanism design problem faced by the ad exchange. Chakraborty et al. [2010] consider the problem faced by the central exchange when the bidders have additional constraints

that limit the number of auctions they can participate in. The authors develop a joint optimization framework to decide whom to solicit bids from at each step. Feldman et al. [2010] investigate the setting of auctions with intermediaries, where the individual bidders must purchase the good not from the seller directly, but rather through a self-interested intermediate third party. They design optimal auctions in this setting and analyze the resulting equilibria. Both of these situations are quite different from our setting where the goal is to design mechanisms to incorporate CPC-CPM arbitrager agents into the exchange in an efficient and revenue-positive manner.

In their work on arbitraging in sponsored search markets, Bu et al. [2008] explore a different setting for arbitrage, namely content match publishers (for example, Google's AdSense or Yahoo's Content Match) attracting additional traffic by means of search advertising. In contrast to our work, they are not interested in enabling these kinds of arbitragers, but rather analyze the strategies and equilibria in these settings. Finally, there is other recent work that addresses the impact of asymmetrically informed advertisers in online auctions; e.g., Abraham et al. [2011] consider the case where select advertisers have access to cookie information. This is different from our setting where the informed parties are non-advertisers that we seek to bring within the system.

2 An efficient solution

We proceed directly to the main contribution of the paper: a mechanism that achieves efficiency in dominant strategies, and a proof that for distributions satisfying a hazard rate condition the expected revenue impact of adopting it over the CPC-bidder-excluding option is positive. We consider a simplified model in which there is a single impression to be allocated, a set of CPM-bidders I, a single CPC-bidder who obtains value only if the impression is clicked, and a single arbitrager who gets an unbiased signal of the true probability with which the CPC-bidder's ad would be clicked were it to be shown; we denote this true probability p^*. There is a common prior over p^*, and the arbitrager alone forms an updated posterior as result of his privately observed signal. All parties will thus seek to maximize their expected utility with respect to the *arbitrager's* true beliefs about p^*.[3] All parties are assumed to be risk-neutral, and so there is no meaningful distinction between this setup and one in which the arbitrager simply knows p^*. The single-CPC-bidder and single-arbitrager restrictions serve to make the exposition clearer but are otherwise not very significant, as we will see in Section 4 when we address the general case.

The mechanism elicits reported values (bids) from each advertiser, and predicted probability-of-click from the arbitrager. We assume p^* is strictly positive, and so the mechanism will not admit a prediction that $p^* = 0$. For any vector b of real numbers we let $b^{(k)}$ denote the kth highest value in b.

[3] One instantiation of the model is this. There is a true value p^* drawn from Beta(1,1), the uniform distribution. The realization of p^* is observed by neither party, but both the CPC-bidder and the arbitrager know the process generating p^*. In addition, nature draws a signal k from $B(n, p^*)$, the binomial distribution with n trials and success (click) probability p^*.This signal is privately revealed to the arbitrager. The arbitrager performs the Bayes update, so that the arbitrager's posterior on p^* is Beta$(1 + k, 1 + (n - k))$.

Definition 1. (SECOND PRICE FOR ALL (SP) MECHANISM) *The CPC-bidder announces a value-per-click $\hat{v}$, the arbitrager announces a probability-of-click $\hat{p}$ for the CPC-bidder's ad, and each CPM-bidder $i \in I$ announces a value-per-impression b_i, with b denoting the vector of CPM bids.*

- *If $\hat{p}\hat{v} < b^{(1)}$, the impression is allocated to bidder $\arg\max_{i \in I} b_i$ who pays $\max\{\hat{p}\hat{v}, b^{(2)}\}$ (ties can be broken arbitrarily).*
- *If $\hat{p}\hat{v} \geq b^{(1)}$, the impression is allocated to the CPC-bidder and the arbitrager pays the highest CPM bid, $b^{(1)}$. If the impression is clicked, additionally the arbitrager is paid $\hat{v}$ and the CPC-bidder pays $\frac{b^{(1)}}{\hat{p}}$.*

We illustrate the mechanism on the following simple example: there are two CPM-bidders who announce bids \$0.04 and \$0.06; the CPC-bidder announces value \$0.90; and the arbitrager announces probability of click 0.1. The impression will be allocated to the CPC-bidder (since $0.1 \cdot \$0.90 > \$0.06$) and the arbitrager will pay \$0.06. If the CPC-bidder's advertisement is clicked, then additionally the arbitrager will be paid \$0.90 and the CPC-bidder will pay \$0.60.[4]

An efficient allocation is one in which the bidder with highest expected value receives the impression; the SP mechanism achieves efficient allocations in dominant strategies.

Theorem 1. *The SP mechanism is truthful and efficient in dominant strategies.*

Note that in the truthful dominant strategy equilibrium the expected revenue to the publisher, conditional on allocation to the CPC-bidder, equals:

$$\begin{aligned} & b^{(1)} + p^* \left(\frac{b^{(1)}}{p^*} - v^* \right) \\ = \; & 2b^{(1)} - p^* v^* \end{aligned}$$

Conditional on allocation to a CPM-bidder, equilibrium revenue equals $\max\{p^*v^*, b^{(2)}\}$.

In the SP mechanism the arbitrager's expected net payment equals his expected contribution to efficiency—i.e., if the CPC-bidder obtains the good, the difference between his expected value and the highest CPM-bidder's value. A mechanism that scaled the arbitrager's payments down uniformly (charging him $\epsilon b^{(1)}$ if the CPC-bidder wins, and additionally paying him $\epsilon\hat{v}$ if a click occurs, for some small $\epsilon > 0$) would also be strategyproof;[5] but this would not be an effective solution here because it would create an incentive for the arbitrager to bypass the mechanism and directly buy (from the publisher) and sell (to the CPC-bidder), extracting an amount that approximates his contribution to efficiency. This would inevitably lead to lost efficiency in some cases, and so in the efficient solution we must pay the arbitrager (in expectation) his

[4] For the purposes of this section the net revenue of the mechanism can be allocated arbitrarily between the ad exchange and the publisher. The choice of how this allocation is made may have important practical implications, which we address in Section 2.1; for now it is sufficient to only consider the impact of payments on the privately informed parties (i.e., the advertisers and arbitrager).

[5] We thank an anonymous referee for this suggestion.

entire contribution to efficiency, which is the best possible payment the arbitrager could obtain if directly bargaining with the CPC-bidder outside the mechanism. Remarkably, as we will soon demonstrate, while doing this the SP mechanism still increases overall revenue under a fairly standard assumption about the value distribution.

2.1 Revenue and efficiency implications

Theorem 1 tells us that the SP mechanism is ideal with respect to efficiency, but we need also to consider the revenue impact of folding a CPC-bidder and arbitrager into the system in the way proposed. Assume all aggregate payments taken in by the mechanism are given to the publisher; the publisher would benefit from moving to the SP mechanism from a second-price auction that excludes the CPC-bidder if doing so will lead to increased expected revenue.[6] Without allowing CPC-bidders, revenue would always equal $b^{(2)}$. Then allowing a CPC-bidder, in the case where the CPC-bidder wins (i.e., when $p^*v^* \geq b^{(1)}$), the *change* in resulting revenue equals:

$$2b^{(1)} - p^*v^* - b^{(2)} \tag{1}$$

Whether or not this is positive in expectation depends on the distribution from which values are drawn. But there is another factor: when the CPC-bidder/arbitrager submit the *second highest* bid, this *increases* revenue (with certainty) over what it would have been if they were excluded. Specifically, when $b^{(1)} > p^*v^* > b^{(2)}$, the change in resulting revenue from including them equals:

$$p^*v^* - b^{(2)}$$

Then the total change in expected revenue from allowing the CPC-bidder/arbitrager and moving to the SP mechanism from the second-price CPC-bidder-excluding approach equals:

$$Pr(p^*v^* \geq b^{(1)})\, \mathbb{E}[2b^{(1)} - p^*v^* - b^{(2)} \,|\, p^*v^* \geq b^{(1)}] \tag{2}$$

$$+\, Pr(b^{(1)} > p^*v^* > b^{(2)})\, \mathbb{E}[p^*v^* - b^{(2)} \,|\, b^{(1)} > p^*v^* > b^{(2)}] \tag{3}$$

Now to make things precise, assume there are $n-1$ CPM-bidders and one CPC-bidder, and assume p^*v^* and every CPM bid is drawn independently from the same distribution, with p.d.f. f and c.d.f. F. We assume there is at least one CPM-bidder (i.e., $n > 1$), since otherwise efficiency is obtained trivially by simply always allocating to the CPC-bidder. There is a $1/n$ chance of the CPC-bidder having the highest bid and a $1/n$ chance of him having the second highest bid. Then letting $\mathcal{E}_k$ denote the

[6] The expectation is taken over all bidders' private types. Sufficiency of an *expected* revenue gain is justified by the small size and fast pace of individual transactions in the advertising realm, since these factors suggest that advertisers and publishers should approximately experience the average. As mentioned in the introduction, if the publisher is unwilling to risk obtaining less revenue than he would under the CPC-bidder-excluding solution, a risk-neutral exchange can transfer this amount to the publisher and execute all payments in the SP mechanism; then our requirement completely analogously becomes that the exchange not run at a loss.

expected value of the kth highest draw from the distribution, the expected revenue gain spelled out in Eqs. (2–3) reduces to:

$$\begin{aligned}
&\frac{1}{n}(2\mathcal{E}_2 - \mathcal{E}_1 - \mathcal{E}_3) + \frac{1}{n}(\mathcal{E}_2 - \mathcal{E}_3) \\
&= \frac{1}{n}(3\mathcal{E}_2 - \mathcal{E}_1 - 2\mathcal{E}_3)
\end{aligned}$$

This can be expressed:

$$\begin{aligned}
&\frac{1}{n}\Big[3\int n(n-1)f(x)(1-F(x))F(x)^{n-2}x\,dx - \int nf(x)F(x)^{n-1}x\,dx \\
&\quad - 2\int \frac{n(n-1)(n-2)}{2} f(x)(1-F(x))^2F(x)^{n-3}x\,dx\Big] \\
&= 3(n-1)\int f(x)(1-F(x))F(x)^{n-2}x\,dx - \int f(x)F(x)^{n-1}x\,dx \\
&\quad - (n-1)(n-2)\int f(x)(1-F(x))^2F(x)^{n-3}x\,dx
\end{aligned}$$

In the case where values are distributed $U[0,1]$ this reduces to:

$$\begin{aligned}
&3(n-1)\int_0^1 (1-x)x^{n-1}\,dx - \int_0^1 x^n\,dx - (n-1)(n-2)\int (1-x)^2x^{n-2}\,dx \\
&= \frac{1}{n(n+1)}
\end{aligned}$$

Since expected revenue without the CPC-bidder equals $\frac{n-2}{n}$ in the uniform case, the expected *percentage* revenue gain equals:

$$\frac{1}{n(n+1)} \Big/ \frac{n-2}{n} = \frac{1}{(n-2)(n+1)}$$

So, particularly when n is small, there is a significant revenue gain if bidder values are uniformly distributed. We established through numerical calculations of expected revenue that this is also the case for normally distributed values, with a variety of different standard deviations tested (see Figure 1 for an example). But we can go significantly further and analytically demonstrate that expected revenue will be increased for *any* distribution that has a monotonically increasing hazard rate.

Definition 2 (Hazard rate)**.** *The hazard rate at value x for distribution function F with density f is:*

$$\frac{f(x)}{1-F(x)}$$

Theorem 2. *When all agent values are i.i.d. according to a distribution with monotonically increasing hazard rate, the SP mechanism yields greater expected revenue than the second-price CPC-bidder-excluding mechanism.*

There's also a significant efficiency gain over a mechanism that excludes the CPC-bidder/arbitrager. Taking $\mathbb{E}$[value to bidder receiving the impression] as our efficiency metric, letting E_n represent this quantity for n bidders in the symmetric case, the percentage efficiency gain equals $\frac{E_n - E_{n-1}}{E_{n-1}}$. For uniform values this equals $\frac{1}{n^2-1}$. (See Figure 1.)

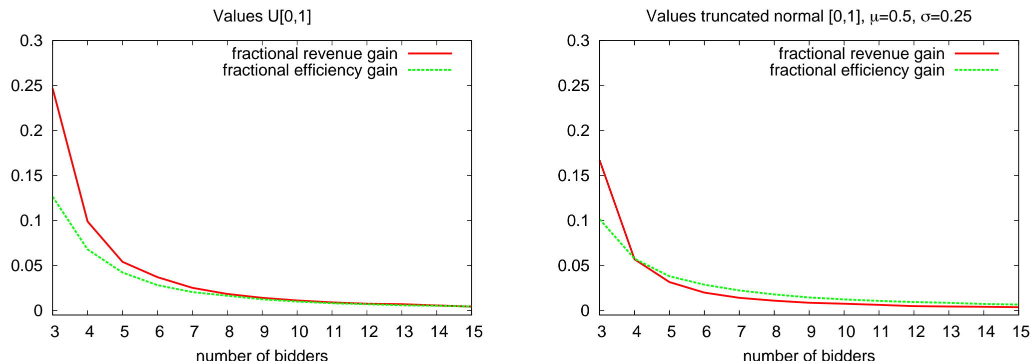

Figure 1: Expected revenue and efficiency gains of allowing a single CPC-bidder and arbitrager to participate under the SP mechanism, for uniform (left) and normal (right) value distributions. Specifically, let X denote the average efficiency (or revenue) of the CPC-bidder-excluding mechanism, and Y be that for the SP mechanism; we plot $(Y - X)/X$.

2.2 Value for an unclicked impression

As the proof of Theorem 1 exposes (see the Appendix), efficiency of the SP mechanism is critically dependent on the assumption that the CPC-bidder obtains non-zero value *only upon click*. In this subsection we show that if this is not the case, and the bidder has some value for an unclicked impression, we can achieve efficiency with a variant of the SP mechanism in an ex post equilibrium.

Imagine that the CPC-bidder has value v^* per click and additional value k^* for the impression *independent of whether there's a click*. Then his expected utility under the SP mechanism, conditional on being allocated the impression, equals:

$$p^* \left(v^* - \frac{b^{(1)}}{\hat{p}} \right) + k^*$$

This will be positive (and thus the CPC-bidder would want to win) whenever $v^* > \frac{b^{(1)}}{\hat{p}} - \frac{k^*}{p^*}$. But if $v^* \in \left(\frac{b^{(1)}}{\hat{p}} - \frac{k^*}{p^*}, \frac{b^{(1)}}{\hat{p}} \right)$, he will not win if he reports truthfully, and thus incentive compatibility no longer holds. There is no longer a feasible dominant strategy because the CPC-bidder's optimal bid would balance his *unconditional* allocation value (k^*) with his *conditional* value (v^*), and the odds of this conditional value being reaped depends on p^*, the best estimate of which is privately known by the arbitrager.

Nonetheless the mechanism can be adapted and we can achieve truthfulness with a weaker equilibrium notion as follows. The adapted mechanism elicits two reports from

the CPC-bidder: his unconditional impression value and his click value. Note that the CPC-bidder, though he now obtains positive value for an unclicked impression, remains unwilling to pay-per-impression.

Definition 3. (SP MECHANISM FOR A CPC-BIDDER WITH IMPRESSION VALUE (SP-IV)) *The CPC-bidder announces a value-per-click $\hat{v}$ and a value-per-impression $\hat{k}$, the arbitrager announces a probability-of-click $\hat{p}$, and each CPM bidder $i \in I$ announces a value-per-impression b_i, with b denoting the vector of CPM bids.*

- *If $\hat{p}\hat{v} + \hat{k} < b^{(1)}$, the impression is allocated to bidder* $\arg\max_{i \in I} b_i$ *who pays* $\max\{\hat{p}\hat{v} + \hat{k}, b^{(2)}\}$ *(ties can be broken arbitrarily).*
- *If $\hat{p}\hat{v} + \hat{k} \geq b^{(1)}$, the impression is allocated to the CPC-bidder and the arbitrager pays $b^{(1)} - \hat{k}$. If the impression is clicked, additionally the arbitrager is paid $\hat{v}$ and the CPC-bidder pays $\frac{b^{(1)}}{\hat{p}}$.*

The SP-IV mechanism modifies the SP mechanism by considering impression value to the CPC-bidder in determining the (efficient) allocation, and by decreasing the arbitrager's charge (when the CPC-bidder is allocated the impression) by the CPC-bidder's reported impression value.

Theorem 3. *The* SP-IV *mechanism is truthful and efficient in ex post Nash equilibrium for the generalized setting where the CPC-bidder may derive non-zero value for an unclicked impression.*

This mechanism does not recover dominant strategy truthfulness. To see this consider the case where the arbitrager over-reports the probability of click ($\hat{p} > p^*$) and:

$$b^{(1)} > \hat{p}v^* + k^* \quad \text{and} \quad \hat{p} > \frac{p^* b^{(1)}}{p^* v^* + k^*}$$

These inequalities are quite compatible, and note that, if truthful, the CPC-bidder will not receive the impression. Now if the CPC-bidder were to over-report (either v^*, k^*, or both) to be allocated the impression his expected utility would equal $p^*v^* - \frac{p^*}{\hat{p}} b^{(1)} + k^*$, and this is strictly positive. Therefore in this case truthfulness is not optimal and so the mechanism is not strategyproof.

In the truthful ex post Nash equilibrium of the SP-IV mechanism, the expected revenue, conditional on allocation to the CPC-bidder, equals:

$$\begin{aligned} & b^{(1)} - k^* - p^* \left(v^* - \frac{b^{(1)}}{p^*} \right) \\ &= 2b^{(1)} - p^* v^* - k^* \end{aligned}$$

Conditional on allocation to a CPM-bidder, it is $\max\{p^*v^* + k^*, b^{(2)}\}$.

If we assume that the CPC-bidder's *total* expected value for the impression (i.e., $p^*v^* + k^*$) is drawn from the same distribution as the CPM bidders' values, then the revenue and efficiency analysis of Section 2.1 applies with equal validity in this enriched setting, though we must settle for analyzing results in an ex post Nash rather than dominant strategy equilibrium.

For the rest of the paper we return to the model wherein CPC-bidders are assumed to possibly obtain non-zero value *only* upon click.

3 Inefficient alternatives

The SP mechanism is a compelling solution: it is efficient in dominant strategies and, for a broad range of natural value distributions, will increase revenue over a solution that excludes CPC-bidders. However, for some distributions it will not meet this revenue criterion, and even when it does, on given problem *instances* the revenue may be lower than would result under the CPC-bidder-excluding solution. In the presence of a risk-neutral third party, such as the advertising exchange itself, this poses no problem: the exchange can provide a guaranteed revenue bound to the publisher while making a profit in expectation. If there is no such third-party, then efficiency is unachievable while guaranteeing the CPC-bidder-excluding revenue benchmark.

Theorem 4. *There exists no mechanism that is truthful and efficient in dominant strategies, ex post no-deficit, ex post individually rational for the CPC-bidder, interim individually rational for the arbitrager, and always yields the publisher revenue at least as great as the second highest CPM bid.*

This motivates a consideration of mechanisms that are not perfectly efficient. Moreover, another concern under the SP mechanism is collusion: since the arbitrager is paid the CPC-bidder's reported value if a click occurs, if the CPC-bidder has the highest true value then reporting a dramatically *higher* value does not diminish his utility at all, yet it may greatly increase that of the arbitrager. The two parties can potentially collude to extract an unbounded sum from the mechanism.

In this section we consider alternative mechanisms that mitigate these issues at the cost of efficiency. They will achieve *revenue and efficiency dominance* with respect to the CPC-bidder-excluding auction, i.e., they will never yield less equilibrium revenue or efficiency on any instance, and they will also eliminate the collusion problem described above. We propose the ASP and BSP mechanisms, which differ from the SP mechanism only in how the click-contingent payments for the CPC-bidder and arbitrager are defined, should they submit the highest joint bid.

Definition 4. (Arbitrager second-price (ASP) and Bidder second-price (BSP) mechanisms) *The CPC-bidder announces a value-per-click $\hat{v}$, the arbitrager announces a probability-of-click $\hat{p}$, and each CPM-bidder $i \in I$ announces a value-per-impression b_i, with b denoting the vector of CPM bids.*

- *If $\hat{p}\hat{v} < b^{(1)}$, the impression is allocated to bidder $\arg\max_{i \in I} b_i$ who pays $\max\{\hat{p}\hat{v}, b^{(2)}\}$ (ties can be broken arbitrarily).*
- *If $\hat{p}\hat{v} \geq b^{(1)}$, the impression is allocated to the CPC-bidder and the arbitrager pays $b^{(1)}$. If the impression is clicked then additional payments are made, defined in the ASP and BSP mechanisms, respectively, as:*
  - *(ASP) The arbitrager is paid $\hat{v}$ and the CPC-bidder pays $\hat{v}$.*
  - *(BSP) The arbitrager is paid $\frac{b^{(1)}}{\hat{p}}$ and the CPC-bidder pays $\frac{b^{(1)}}{\hat{p}}$.*

The ASP mechanism can be thought of as second-price for the arbitrager but *first-price* for the CPC-bidder; the BSP mechanism can be thought of as second-price for

the CPC-bidder but first-price (in a sense) for the arbitrager. In the case of the ASP mechanism the CPC-bidder will have an incentive to under-report his value, while the arbitrager's best strategy is truth; under the BSP mechanism the arbitrager will have incentive to under-report his predicted probability-of-click, though the CPC-bidder is best off being truthful.

Since in these mechanisms *over*-reporting is a dominated strategy, there is no risk of the CPC-bidder/arbitrager winning when they do not have the highest value. Thus they will only improve efficiency over a mechanism that excludes them; and revenue also will only increase since they can only increase the second-highest bid. The following two propositions follow for the same reasons given in the proof of Theorem 1.

Proposition 1. *Truthful reporting is a dominant strategy for the arbitrager in the ASP mechanism.*

Proposition 2. *Truthful reporting is a dominant strategy for the CPC-bidder in the BSP mechanism.*

We can also show that the ASP and BSP mechanisms, though imperfect, under very weak assumptions will at least weakly increase both revenue and efficiency over the CPC-bidder-excluding mechanism.

Theorem 5. *Assuming CPM-bidders play only undominated strategies, then on every possible value profile the ASP and BSP mechanisms both yield weakly greater revenue than the second-price CPC-bidder-excluding mechanism.*

Theorem 6. *Assuming all agents play only undominated strategies, then on every possible value profile the ASP and BSP mechanisms both yield weakly greater allocation value (efficiency) than the second-price CPC-bidder-excluding mechanism.*

To understand the equilibrium behavior of the potentially (rationally) non-truthful parties in the ASP and BSP mechanisms, we consider the expected value they would achieve as a function of their reports. First consider the ASP mechanism. Let f_p and F_p denote, respectively, the p.d.f. and c.d.f. representing the CPC-bidder's beliefs about the click-probability which the arbitrager will (truthfully, by Proposition 1) report. Recall that F (with no subscript) denotes the c.d.f. for any one of the $n-1$ CPM bidders' bids. The CPC-bidder's expected utility from reporting $\hat{v}$ when his true value is v^* equals:

$$\int f_p(p) F(p\hat{v})^{n-1} p(v^* - \hat{v})\, dp \tag{4}$$

From here, in the special case of uniformly distributed values, we can compute the CPC-bidder's equilibrium bid without assuming anything about his beliefs regarding p^*. We now provide a detailed treatment of this special case. We will consider bidders with values restricted to the interval $[0, 1]$, merely for simplicity.

3.1 Uniformly distributed values

Proposition 3. *If CPM-bidder values are i.i.d. according to $\mathcal{U}(0,1)$ and only undominated strategies are played, the CPC-bidder's optimal bid in the ASP mechanism is $\frac{n-1}{n}v^*$, regardless of his beliefs about p^*.*

Now considering the BSP mechanism, let f_v and F_v denote, respectively, the p.d.f. and c.d.f. representing the arbitrager's beliefs about the value that the CPC-bidder will (truthfully, by Proposition 2) report. Let g and G, respectively, denote the p.d.f. and c.d.f. for the *highest* CPM value. The arbitrager's expected utility from reporting $\hat{p}$ when the true probability is p^* equals:

$$\int_0^1 f_v(v) \int_0^{\hat{p}v} g(c)\Big(\frac{p^*}{\hat{p}}c - c\Big)\, dc\, dv \tag{5}$$

Proposition 4. *If CPM-bidder values are i.i.d. according to $\mathcal{U}(0,1)$ and only undominated strategies are played, the arbitrager's optimal report in the BSP mechanism is $\frac{n-1}{n}p^*$, regardless of his beliefs about v^*.*

This analysis demonstrates that, in the case of uniformly distributed values, the outcomes of the ASP and BSP mechanisms are identical in equilibrium: in each case the equilibrium "joint bid" by the CPC-bidder/arbitrager pair will be $\frac{n-1}{n}p^*v^*$, with CPM-bidders reporting truthfully. However this equivalence will not hold for other distributions and, unfortunately, in general the equilibrium bids will depend on the CPC-bidder's (arbitrager's) belief about the arbitrager's (CPC-bidder's) report.[7] But we can fall back on Theorems 5 and 6 to conclude that, whatever the distributions over v^* and p^*, efficiency and revenue will only be increased over the second-price CPC-bidder-excluding mechanism.

The equilibrium expected value to the highest bidder—which is a measure of efficiency—in the ASP or BSP mechanisms with uniformly distributed values equals:

$$\int_0^1 \Big[\Big(\frac{n-1}{n}x\Big)^{n-1} x + \int_{\frac{n-1}{n}x}^1 (n-1)y^{n-1}\, dy\Big]\, dx$$
$$= \frac{n-1}{n} + \frac{1}{n+1}\Big[\Big(\frac{n-1}{n}\Big)^{n-1} - \Big(\frac{n-1}{n}\Big)^{n+1}\Big]$$

Figure 2 depicts the expected percentage efficiency and revenue gains of moving to either the ASP or BSP mechanism from the arbitrager-excluding approach, in the same manner as did Figure 1 for the SP mechanism, for the uniform values case. The revenue increase is slightly more than that of the SP mechanism, although the efficiency gain is less, which could have been deduced from the fact that the SP mechanism is perfectly efficient for all distributions and the ASP and BSP mechanisms are not, regardless of the distribution.

[7] We might assume that these beliefs correspond to the prior, in the case of the CPC-bidder and p^*, and f in the case of the arbitrager and v^*. One could then derive equilibrium bids for any given instantiation of these distributions. But the point here is that the expectations held by the arbitrager and CPC-bidder about each other's types, whatever they may be, turn out to be *irrelevant* when CPM-bidder values are uniformly distributed.

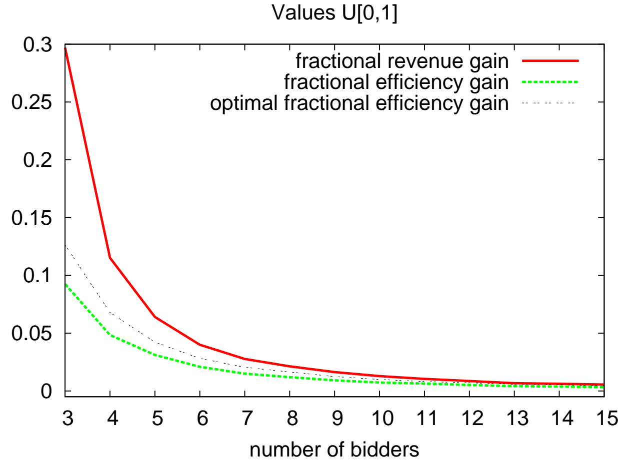

Figure 2: Expected revenue and efficiency gains of allowing a single CPC-bidder and arbitrager to participate under the ASP and BSP mechanisms, for uniformly distributed values. The dashed line represents the efficiency improvement achieved by the SP mechanism, which always allocates the impression efficiently.

4 Generalizations

Perhaps the two most important ways in which the setting we've considered thus far can be generalized are: allowing for multiple CPC-bidders, and allowing for multiple arbitragers. To handle incentives with these generalizations we will have to use somewhat more complex mechanisms, but the basic analysis at play with respect to the SP mechanism and the single CPC-bidder/arbitrager setting continues to hold. We now let C denote the *set* of (CPC) bidders only willing to pay-per-click, and let R denote the set of arbitragers.

4.1 Multiple CPC-bidders, each with one arbitrager

We first consider the case where CPC-bidders and arbitragers come in pairs, so there is no competition amongst arbitragers for any given CPC-bidder, but there is competition between CPC-bidder/arbitrager pairs. We call this the *captive arbitrager* setting.

Definition 5. (SP mechanism for captive arbitrager settings (SP-ca)) *Each CPC-bidder $c \in C$ announces a value-per-click $\hat{v}_c$, each arbitrager $r_c \in R$ announces a probability-of-click $\hat{p}_c$ for his associated CPC-bidder c, and each CPM-bidder $i \in I$ announces a value-per-impression b_i. Let b_c denote $\hat{p}_c\hat{v}_c$ for each $c \in C$, let b denote the set of bids $\{b_j\}_{j\in I\cup C}$, and let $h \in \arg\max_{j\in I\cup C} b_j$, breaking ties arbitrarily.*

- *If $h \in I$: the impression is allocated to bidder h who pays $b^{(2)}$.*
- *Alternatively if $h \in C$: the impression is allocated to bidder h and arbitrager r_h pays $b^{(2)}$. If the impression is clicked, additionally h pays $\frac{b^{(2)}}{\hat{p}_h}$ and r_h is paid $\hat{v}_h$.*

Theorem 7. *The* SP-CA *mechanism is truthful and efficient in dominant strategies for the captive arbitrager setting, for any number of CPC-bidder/arbitrager pairs.*

4.2 Multiple CPC-bidders, one arbitrager for all

We now consider the case where there is a single arbitrager that can potentially pair up with more than one CPC-bidder; we call this the *sole arbitrager* setting. We assume that any given CPC-bidder's click-probability is completely independent of that of any other CPC-bidders (this eliminates the possibility of a "winner's curse" effect that the arbitrager may otherwise need to correct for).

Whereas in the previous setting we could base the arbitrager's payment on bids by CPC-bidders (since his reports had no interaction with any CPC bids except one), here instead we must charge the arbitrager based only on CPM bids, though CPC-bidders can still be charged based on other CPC bids.

Definition 6. (SP MECHANISM FOR SOLE ARBITRAGER SETTINGS (SP-SA)) *Each CPC-bidder $c \in C$ announces a value-per-click $\hat{v}_c$, the arbitrager announces a probability-of-click $\hat{p}_c$ for each CPC-bidder c, and each CPM-bidder $i \in I$ announces a value-per-impression b_i. Let b_c denote $\hat{p}_c \hat{v}_c$ for each $c \in C$, let b denote the set of bids $\{b_j\}_{j \in I \cup C}$, and let $h \in \arg\max_{j \in I \cup C} b_j$, breaking ties arbitrarily.*

- *If $h \in I$: the impression is allocated to bidder h who pays $b^{(2)}$.*
- *Alternatively if $h \in C$: the impression is allocated to bidder h and the arbitrager pays $\max_{i \in I} b_i$. If the impression is clicked, additionally h pays $\frac{b^{(2)}}{\hat{p}_h}$ and the arbitrager is paid $\hat{v}_h$.*

The key difference from the SP-CA mechanism is that in the SP-SA mechanism the arbitrager pays the maximum CPM bid rather than the maximum bid overall.

Theorem 8. *The* SP-SA *mechanism is truthful and efficient in dominant strategies for the sole arbitrager setting, for any number of CPC-bidders.*

We note that in the case where there are no CPM-bidders, the SP-SA mechanism will run a deficit: the unconditional payments reduce to zero and, letting $h \in \arg\max_{j \in C} p_j^* v_j$, in the truthful equilibrium the expected on-click payments equal $\max_{j \in C \setminus \{h\}} p_j^* v_j - p_h^* v_h$, which is less than 0 except in the case of a tie. Since in this scenario the arbitrager plays a role in every bid for the impression, it is somewhat akin to a single-bidder scenario, and it can be show that dominant strategy efficiency is impossible to achieve with a budget balanced mechanism that is ex post individually rational for the bidders and provides incentives to the arbitrager to participate. Thus in a world where CPM-bidders are rare and there is only a single arbitrager, budget concerns may compel one to adopt an extension of one of the inefficient mechanisms described in Section 3.

4.3 Multiple CPC-bidders and multiple arbitragers

We finally consider the most general setting, in which there are multiple CPC-bidders and multiple arbitragers, with each arbitrager potentially predicting the probability-of-click for more than one (or even all) of the CPC-bidders. We call this the *many arbitrager* setting. For any arbitrager $r \in R$, C_r will denote the set of CPC-bidders for whom r submits a prediction bid. In this model where more than one arbitrager may submit a prediction regarding each CPC-bidder's probability-of-click, there may be varying estimates, a result of different information held by the different arbitragers. There is an allocation that is efficient with respect to the totality of signals held across arbitragers about each CPC-bidder, but we do not propose a mechanism that seeks to pool information in this way. Instead, we seek a strategyproof mechanism that selects a single arbitrager for each CPC-bidder, and is efficient with respect to the information held by each arbitrager about the CPC-bidder it has been paired with. The arbitragers are paired with CPC-bidders in a way that is independent of bids and predictions, which makes the presence of other potential arbitragers for each CPC-bidder irrelevant from an incentives perspective; strategyproofness then holds by the same reasoning applied in Theorems 1 and 7.

Definition 7. (SP MECHANISM FOR MANY ARBITRAGER SETTINGS (SP-MA)) *Each CPC-bidder $c \in C$ announces a value-per-click $\hat{v}_c$, each arbitrager $r \in R$ announces a probability-of-click $\hat{p}_{r,c}$ for each CPC-bidder $c \in C_r$ (where C_r can be defined arbitrarily by r), and each CPM bidder $i \in I$ announces a value-per-impression b_i.*

- *For each $c \in C$, let R_c denote $\{r \in R : c \in C_r\}$. For each $c \in C$, an $r_c \in R_c$ is chosen according to a decision rule that is independent of reported values and click probabilities.*

Let b_c denote $\hat{p}_{r_c,c}\hat{v}_c$ for each $c \in C$, let b denote the set of bids $\{b_j\}_{j \in I \cup C}$, and let $h \in \arg\max_{j \in I \cup C} b_j$, breaking ties arbitrarily.

- *If $h \in I$: the impression is allocated to bidder h who pays $b^{(2)}$.*
- *Alternatively if $h \in C$: the impression is allocated to bidder h and arbitrager r_h pays $\max_{i \in I \cup \{c \in C : r_c \neq r_h\}} b_i$. If the impression is clicked, additionally h pays $\frac{b^{(2)}}{\hat{p}_{r_h}}$ and r_h is paid $\hat{v}_h$.*

Theorem 9. *The* SP-MA *mechanism is strategyproof in the many arbitrager setting.*

The SP-MA mechanism specifies selecting an arbitrager for each CPC-bidder independent of the reported predictions, but this need not equate with selection at random. Zooming out from our formal static model, in practice each impression auction takes place in a broader context of the ongoing allocation of millions of impressions per day. Some arbitragers may have access to stronger signals than others, and evidence of this may accumulate over time. One possibility in choosing arbitragers is to discriminate amongst them based on previous predictive accuracy (e.g., if r won with an estimated click probability of 0.7 one thousand times and, of those, seven hundred led to clicks, r has been very accurate). This introduces a new consideration for strategic arbitragers,

but it only presents a bias towards *accurate* prediction, which corresponds perfectly with *truthful* prediction given our model.

Another approach is to not choose arbitragers independent of their bids, but instead to favor those who announce *higher* predictions. Creating a bias for over-estimation in this way would not lead to any decrease in revenue for the publisher, though it would break strategyproofness and may decrease efficiency somewhat. If the maximum prediction is selected, arbitragers now face a tradeoff of wanting to be highest (to be selected) and wanting to be accurate, which will lead the winner to be overoptimistic, reducing efficiency. Another factor that would come into play if arbitragers have imperfect information is the winner's curse: if each arbitrager believes the *average* of all arbitragers' beliefs is a better estimate than their own individual estimate, this provides an incentive to submit a lower bid.

5 Empirical analysis

In this section we perform an empirical evaluation of the mechanisms introduced in the paper, specifically evaluating the monotone hazard rate assumption, and the non-decreasing revenue claim of Theorem 2. We find that while the hazard rate assumption does *not* hold for very low bids, the proposed mechanism is nevertheless revenue positive.

5.1 Data

The data used for the experiments is a random sample from six days of live traffic on the RightMedia Exchange (RMX) collected over two months in 2011. At the time of writing, RMX is the largest ad exchange in the industry with over ten billion transactions daily (see http://rightmedia.com/about). The final dataset consists of approximately 1.5 million events, each corresponding to an auction for an ad placement. For each auction event we recorded all of the valid advertiser bids submitted to the publisher (i.e. those satisfying the supply and demand constraints), as well as the campaign type of the advertiser, i.e., whether they are a CPM or a CPC bidder.

5.2 Hazard rate

To investigate the hazard rate assumption, we computed the empirical distribution of all of the bids across all of the auctions. While the hazard rate is essentially constant for bids above \$0.01 (per-impression bid), it is *decreasing* for the very low bids, as shown in Figure 3. We note that the presence of these low bids cannot be easily discounted, since they represent a sizable portion of all submitted bids. At the same time, in practice publishers often set up reserve prices that exclude exactly the bids that violate the hazard rate assumption.

5.3 Revenue and efficiency impact

The monotone hazard rate assumption forms a sufficient but not necessary condition for non-decreasing expected revenue in the presence of the arbitrager. To check whether

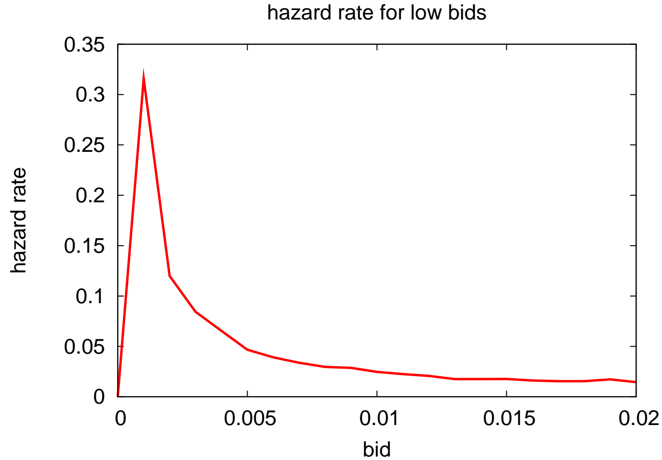

Figure 3: The hazard rate for low bid values.

the SP mechanism yields greater expected revenue than the second-price CPC-bidder-excluding mechanism, we used the collected auction data to calculate the revenues under different mechanisms.

Each auction consists of a set of bids, each in terms of expected value for the impression; some bids are CPM (the greatest of which corresponds to $b^{(1)}$ in our notation), and some are CPC and constituted by a click-value (v_j, for bidder j) multiplied by a click-probability estimate (p_j) provided by RMX.[8] The quantities of interest for each auction are the revenue and efficiency gains from moving to the SP mechanism from a standard second-price auction among only CPM bidders. When there is more than one CPC-bidder, we assume each has a different arbitrager, and thus the relevant comparison mechanism would be SP-CA. In either case, if the high bidder is a CPC-bidder, the revenue *gain* equals two times the second highest overall bid (where a CPC "bid" means $p_j v_j$ here), minus the highest overall bid, minus the second highest CPM bid (see Section 4.1); and the efficiency gain equals the highest overall bid minus the highest CPM bid. Alternatively if a CPM bidder wins, the revenue gain equals the difference between the highest CPC bid and the second highest CPM bid, or 0 if this is negative. Table 1 provides two examples to illustrate these calculations, one with four bidders (two of whom are CPC) and another with three bidders (one of whom is CPC).

We note that a genuine evaluation of the CPC-bidder-excluding mechanism is difficult, for a couple reasons. First, though we would like to evaluate the impact of the mechanism on cases where the publisher refused CPC bids, this is not possible because in such cases no data on the excluded bids is available.[9] So instead we look at cases

[8]RMX provides click probabilities which form the basis for CPC-payment second pricing in cases where publishers are willing to accept payment per click. However, other major exchanges do not provide this service; and even in the case of RMX, bringing better-informed third-parties into the system would improve efficiency.

[9]Further complicating matters is the fact that whether CPC bids were excluded from a given auction is unobservable; if CPC bids appear in the bid log for an auction then we know the publisher accepted

Bidder type	Bid-per-impression
CPC	$0.04 \cdot 40 = 16$
CPC	$0.02 \cdot 60 = 12$
CPM	9
CPM	3

Bidder type	Bid-per-impression
CPM	20
CPC	$0.05 \cdot 30 = 15$
CPM	12

Table 1: Two auction examples illustrating the nature of the dataset and the measures we apply to it. In the first auction (left), the expected revenue gain from the SP mechanism over the CPC-bidder-excluding mechanism equals $2 \cdot 12 - 16 - 3 = 5$ and the efficiency gain equals $16 - 9 = 7$. In the second auction (right), the revenue gain equals $15 - 12 = 3$ and the efficiency gain is 0.

where CPC bids *were* accepted, where we have the full set of submitted CPM and CPC bids, and use results there to drawn an inference about the expected effect in cases where the publisher is unwilling to accept CPC bids and the arbitrager becomes critical.[10] Another difficulty stems from the fact that it is impossible to determine whether "CPC bidders" would have stayed out of the auction or instead submitted CPM bids if such was there only option—faced with an exclusionary mechanism some of the CPC-bidders may switch to being CPM-bidders and take on the CPC-CPM conversion risk, rather than be locked out of the auction completely. Motivated by this ambiguity, we simulated the two mechanisms in two scenarios, each representing an opposite end of the spectrum with respect to whether a CPC-bidder would be willing to bid CPM in the CPC-bidder-excluding mechanism:

1. *All.* In this scenario we assume that all of the CPC-bidders remain CPC, and thus do not appear in the CPC-bidder-excluding mechanism.

2. *Random.* In this scenario we assume that a single randomly chosen CPC-bidder remains CPC, while all others become CPM-bidders.

We present both the average revenue and efficiency gains in Table 2, as a percentage increase of moving to the SP mechanism from the CPC-bidder-excluding mechanism. Due to the methodological challenges described above, the most prudent interpretation of these numbers will focus on the *signs* (which, happily, are mostly positive), rather than taking the magnitude as an indication of the expected revenue and efficiency impact for the exchange as a whole.

Under the *All* scenario, the SP mechanism provided more revenue to the publisher on each day. In the stricter *Random* scenario the results were mixed, with the overall gain fluctuating between $\pm 1\%$.

CPC bidders, otherwise it could be either that CPC bidders were excluded or none showed up at the auction.

[10] An assumption that bids (and bid types) are identically distributed across CPC-bid-accepting and CPC-bid-excluding auctions would validate this inference.

Metric	Scenario	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Avg. over all days
Revenue	*All*	11.34	4.01	19.24	13.50	8.25	1.07	7.62
	Random	1.00	-1.80	0.68	0.47	-0.90	0.08	0.002
Efficiency	*All*	4.40	3.61	3.67	2.48	5.30	0.44	2.42
	Random	0.31	1.32	0.15	0.05	0.93	0.02	0.26

Table 2: Empirical revenue and efficiency changes (in %) after moving to mechanism SP, for 6 days over the course of two months. The last column is the average over all auctions for the 6 days, and since some days had more auction samples than others, this is not identical to the sample average of the numbers in each row.

6 Conclusion

Given a perfect estimator of whether any particular ad will be clicked, including as many bidders as possible in an ad auction is ideal from an efficiency standpoint. Advertisers often prefer to bid on a CPC basis, while publishers can and do refuse CPC bids in favor of CPM bids. The mismatch between preferred payments can be resolved satisfactorily with intermediaries serving an arbitrage role and making appropriate CPC-pricing possible with their private click-probability information. We provided an efficient mechanism in which truthful reporting is a dominant strategy for all parties, including arbitragers. There is reason to be optimistic about the revenue impact to publishers, since for any value distribution with a monotonically increasing hazard rate, expected revenue will increase. We also provided two mechanisms which never reduce publisher revenue, while still increasing efficiency.

The empirical picture is less definitive. We found a mixed revenue-impact with respect to real bid data from Yahoo's Right Media Exchange. This could be because of the nature of the bid distribution (which does not satisfy the hazard condition), but there are other factors to consider. In practice no arbitrager is a perfect predictor, and the prediction engine at work in the RMX data—which estimates click-probabilities for CPC-bidders—is imperfect. There is reason to believe this issue may be especially relevant in our data, because at the time the data was collected the prediction engine in use was relatively new and experimental.

Important directions for future work include, first, verifying that despite inevitably imperfect click-prediction, efficiency is increased by including CPC-bidders and arbitragers; and second, achieving a better understanding of the factors that determine whether moving to the efficient mechanism will increase publisher revenue. Future work should also more thoroughly address how to structure allocation decisions in a context of multiple CPC-bidders and competing arbitragers with imperfect predictions.

References

[Abraham *et al.*, 2011] Ittai Abraham, Susan Athey, Moshe Babaioff, and Michael Grubb. Peaches, lemons, and cookies: Designing auction markets with dispersed information. Harvard University (unpublished), 2011.

[Arrow, 1971] Kenneth J. Arrow. Insurance, risk and resource allocation. *Essays in the Theory of Risk-Bearing*, pages 134–143, 1971.

[Bu *et al.*, 2008] Tian-Ming Bu, Xiaotie Deng, and Qi Qi. Arbitrage opportunities across sponsored search markets. *Theoretical Computer Science*, 407:182–191, November 2008.

[Chakraborty *et al.*, 2010] Tanmoy Chakraborty, Eyal Even-Dar, Sudipto Guha, Yishay Mansour, and S. Muthukrishnan. Selective call out and real time bidding. In Amin Saberi, editor, *WINE*, volume 6484 of *Lecture Notes in Computer Science*, pages 145–157. Springer, 2010.

[de Finetti, 1974] Bruno de Finetti. *Theory of Probability*. Wiley, New York, 1974.

[Fama and French, 2004] Eugene F. Fama and Kenneth R. French. The capital asset pricing model: Theory and evidence. *The Journal of Economic Perspectives*, 18(3):25–46, 2004.

[Feldman *et al.*, 2010] Jon Feldman, Vahab S. Mirrokni, S. Muthukrishnan, and Mallesh M. Pai. Auctions with intermediaries: extended abstract. In David C. Parkes, Chrysanthos Dellarocas, and Moshe Tennenholtz, editors, *Proceedings of the 11th ACM Conference on Electronic Commerce (EC-10)*, pages 23–32. ACM, 2010.

[Karni, 2009] Edi Karni. A mechanism for eliciting probabilities. *Econometrica*, 77(2):603–606, 2009.

[Merton, 1987] Robert C. Merton. A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, 42(3):483–510, 1987.

[Muthukrishnan, 2009] S. Muthukrishnan. Ad exchanges: Research issues. In Stefano Leonardi, editor, *WINE*, volume 5929 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2009.

[Myerson and Satterthwaite, 1983] Roger Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 28:265–281, 1983.

[Savage, 1971] Leonard Jimmie Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66:783–801, 1971.

Appendix

In this Appendix we provide the proofs of all results stated in the main text.

Theorem 1. *The SP mechanism is truthful and efficient in dominant strategies.*

Proof. We will show that no agent can ever gain by bidding other than his true value (strategyproofness), which, given the allocation rule, entails dominant strategy efficiency. For every CPM bidder strategyproofness is immediate because the auction faced is a standard Vickrey auction.

Considering the CPC-bidder and the arbitrager, whether they win the auction depends on the *product* of their reports. First take the CPC-bidder's perspective and consider arbitrary report $\hat{p}$ for the arbitrager and any bids b by the CPM-bidders. The CPC-bidder obtains value v^* on click and 0 otherwise. Conditional on being allocated the impression his expected utility will be:

$$p^* \left(v^* - \frac{b^{(1)}}{\hat{p}} \right) \tag{6}$$

Given an allocation decision, the payment the CPC-bidder makes is independent of his bid, and, given Expression (6), his utility is maximized by being allocated the impression if and only if $v^* \geq b^{(1)}/\hat{p}$. Reporting $\hat{v} = v^*$ achieves this, and thus truthful reporting is a dominant strategy.

Now take the arbitrager's perspective and consider arbitrary bid $\hat{v}$ for the CPC-bidder and bids b by the CPM-bidders. If allocated the impression the arbitrager obtains expected utility $p^*\hat{v} - b^{(1)}$ (he pays $b^{(1)}$ for sure and receives $\hat{v}$ with probability p^*). Given an allocation decision, the payment he receives is independent of his bid, and therefore the arbitrager's utility is maximized if he is allocated the impression if and only if $p^*\hat{v} \geq b^{(1)}$. Reporting $\hat{p} = p^*$ achieves this, and thus truthful reporting is a dominant strategy. □

Theorem 2. *When all agent values are i.i.d. according to a distribution with monotonically increasing hazard rate, the SP mechanism yields greater expected revenue than the second-price CPC-bidder-excluding mechanism.*

Proof. We need to show that for arbitrary distribution F satisfying the hazard rate condition, $3\mathcal{E}_2 - \mathcal{E}_1 - 2\mathcal{E}_3 \geq 0$. Equivalently, $2(\mathcal{E}_2 - \mathcal{E}_3) - (\mathcal{E}_1 - \mathcal{E}_2) \geq 0$. We can express $\mathcal{E}_1$ as:

$$\mathcal{E}_1 = n \int F(x)^{n-1} f(x) x \, dx$$

We can express $\mathcal{E}_2$ as:

$$\begin{aligned}
\mathcal{E}_2 &= n(n-1) \int (1 - F(x)) F(x)^{n-2} f(x) x \, dx \\
&= n \left[F(x)^{n-1} (1 - F(x)) x \Big|_0^\infty - \int F(x)^{n-1} \Big(- x f(x) + (1 - F(x)) \Big) \, dx \right] \\
&= n \int F(x)^{n-1} f(x) x \, dx - n \int F(x)^{n-1} (1 - F(x)) \, dx \\
&= \mathcal{E}_1 - n \int F(x)^{n-1} (1 - F(x)) \, dx
\end{aligned}$$

In a similar manner,

$$\begin{aligned}
\mathcal{E}_3 &= \frac{n(n-1)(n-2)}{2} \int (1-F(x))^2 F(x)^{n-3} f(x) x \, dx \\
&= \frac{n(n-1)}{2} \left[F(x)^{n-2}(1-F(x))^2 x \Big|_0^\infty - \int F(x)^{n-2} \left((1-F(x))^2 - 2x(1-F(x))f(x) \right) dx \right] \\
&= n(n-1) \int xF(x)^{n-2}(1-F(x))f(x)\, dx - \frac{n(n-1)}{2} \int F(x)^{n-2}(1-F(x))^2 dx \\
&= \mathcal{E}_2 - \frac{n(n-1)}{2} \int F(x)^{n-2}(1-F(x))^2 \, dx
\end{aligned}$$

So $2(\mathcal{E}_2 - \mathcal{E}_3) = n(n-1)\int F(x)^{n-2}(1-F(x))^2\,dx$ and $(\mathcal{E}_1 - \mathcal{E}_2) = n\int F(x)^{n-1}(1-F(x))\,dx$. Then the expected revenue gain, $2(\mathcal{E}_2 - \mathcal{E}_3) - (\mathcal{E}_1 - \mathcal{E}_2)$, is

$$\begin{aligned}
& n(n-1) \int F(x)^{n-2}(1-F(x))^2 \, dx - n \int F(x)^{n-1}(1-F(x))\, dx \\
&= n \int F(x)^{n-2}(1-F(x)) \Big((n-1)(1-F(x)) - F(x) \Big) dx \\
&= n \int F(x)^{n-2}(1-F(x)) \Big(n-1-nF(x) \Big) dx \\
&= \frac{n}{n-1} \int (n-1)F(x)^{n-2} f(x) \left[\frac{1-F(x)}{f(x)} \cdot (n-1-nF(x)) \right] dx \qquad (7)
\end{aligned}$$

Note that $(n-1)F(x)^{n-2}f(x)$ is a density, as

$$\int (n-1)F(x)^{n-2} f(x)\, dx = F(x)^{n-1}\big|_0^\infty = 1$$

Let $p_x = (n-1)F(x)^{n-2}f(x)$. Then Eq. (7) can be rewritten:

$$\begin{aligned}
& \frac{n}{n-1} \mathbb{E}_p \left[\frac{1-F(x)}{f(x)} \cdot (n-1-nF(x)) \right] \\
\geq\ & \frac{n}{n-1} \mathbb{E}_p \left[\frac{1-F(x)}{f(x)} \right] \mathbb{E}_p \Big[n-1-nF(x) \Big], \qquad (8)
\end{aligned}$$

where the inequality follows because, using the assumption of increasing hazard rate, $\frac{1-F(x)}{f(x)}$ and $n-1-nF(x)$ are positively correlated (both decreasing in x).

We know that $\mathbb{E}_p\big[\frac{1-F(x)}{f(x)}\big] \geq 0$ since both numerator and denominator are non-negative. To evaluate the second term of Eq. (8),

$$\begin{aligned}
\frac{1}{n-1} \mathbb{E}_p \big[n-1-nF(x) \big] &= \int F(x)^{n-2} f(x) \Big((n-1) - nF(x) \Big) dx \\
&= F(x)^{n-1} \Big|_0^\infty - F(x)^n \Big|_0^\infty = 0
\end{aligned}$$

Therefore, the change in revenue from incorporating the arbitrager/CPC-bidder through the SP mechanism is at least 0. □

Theorem 3. *The* SP-IV *mechanism is truthful and efficient in ex post Nash equilibrium for the generalized setting where the CPC-bidder may derive non-zero value for an unclicked impression.*

Proof. Noting that the CPC-bidder's expected value for the impression equals $p^*v^*+k^*$, the mechanism chooses an allocation that is efficient with respect to agent reports, so the theorem follows if truthtelling is an ex post Nash equilibrium.

For each CPM-bidder, truthfulness is a dominant strategy because each is facing a second-price auction. Now, for arbitrary CPM bids and arbitrary true click probability p^*, assume that the arbitrager truthfully bids $\hat{p} = p^*$. If the CPC-bidder is allocated the impression his expected utility will be $p^*v^* - b^{(1)} + k^*$. Conditional on winning the impression, this is independent of his announced values $\hat{v}$ and $\hat{k}$. His utility is thus maximized if he is allocated the impression if and only if $p^*v^* + k^* \geq b^{(1)}$. Reporting $\hat{v} = v^*$ and $\hat{k} = k^*$ achieves this, and thus truthfulness is a best-response.

Now, taking the arbitrager's perspective, consider arbitrary CPM bids and truthful reports $\hat{v} = v^*$ and $\hat{k} = k^*$ for the CPC-bidder, for arbitrary true click and impression values v^* and k^*. If allocated the impression the arbitrager obtains expected utility $p^*v^* - b^{(1)} + k^*$. This is identical to the CPC-bidder's expected utility and, conditional on winning the impression, is also independent of the arbitrager's announced $\hat{p}$. His utility is thus maximized if he is allocated the impression if and only if $p^*v^* + k^* \geq b^{(1)}$. Reporting $\hat{p} = p^*$ achieves this, and thus truthfulness is a best-response. □

The proof of Theorem 4 will make use of the following simple lemma which was not included in the main text.

Lemma 1. *If a mechanism is truthful in dominant strategies, then conditioned on winning the auction, the arbitrager's expected payment must be independent of his announced click-probability p.*

Proof. Assume otherwise. Then there exist reports p and p' such that both p and p' lead to the arbitrager winning the impression, but where the arbitrager's payoff for reporting p is greater than that if he reports p'. Then when the true probability is p the arbitrager has a beneficial deviation in reporting p'. □

Theorem 4. *There exists no mechanism that is truthful and efficient in dominant strategies, ex post no-deficit, ex post individually rational for the CPC-bidder, interim individually rational for the arbitrager, and always yields the publisher revenue at least as great as the second highest CPM bid.*

Proof. Consider arbitrary mechanism that is truthful and efficient in dominant strategies, ex post no-deficit, ex post individually rational for the advertisers, interim individually rational for the arbitrager, and guarantees $b^{(2)}$ to the publisher (where $b^{(k)}$ denotes the kth-highest CPM bid). Under such a mechanism the CPC-bidder pays nothing unless he is allocated the ad and it is clicked (by ex post individual rationality, combined with the assumption that CPC-bidders obtain no value for unclicked impressions). Then if the CPC-bidder is allocated the ad (i.e., if $p^*v \geq b^{(1)}$), the arbitrager

must unconditionally pay (at least) $b^{(2)}$ to the publisher. If no click occurs then no other payments are made (this follows from ex post no-deficit and individual rationality for the CPC-bidder).

Now, to satisfy interim individual rationality for the arbitrager, if a click does occur, then he must (at least in expectation) be paid a quantity x satisfying: $p^*x \geq b^{(2)}$. Moreover, x must be paid by the CPC-bidder, in order to satisfy ex post no-deficit while ensuring that the publisher receives $b^{(2)}$. By Lemma 1, the magnitude of x must be independent of the arbitrager's report p. Fix arbitrary b with $b^{(1)} = b^{(2)}$, fix arbitrary $v > b^{(1)}$, and assume $x < v$. Then there exists a p^* such that $p^*x < b^{(2)}$ and $p^*v > b^{(1)}$ (this holds for any $p^* \in (b^{(1)}/v, b^{(1)}/x)$). Here the arbitrager has a beneficial deviation in reporting $p' < b^{(1)}/v$ rather than p^*, thus losing the auction and paying nothing. Therefore, interim individual rationality for the arbitrager requires that $x \geq v$. To satisfy ex post individual rationality for the CPC-bidder, $x \leq v$, and so $x = v$, i.e., the CPC-bidder pays the arbitrager his reported value. But this is inconsistent with dominant strategy truthfulness, since when $p^*v > b^{(1)}$, the CPC-bidder has a beneficial deviation in reporting $v - \epsilon$ for some $\epsilon \in (0, p^*v - b^{(1)})$. □

Theorem 5. *Assuming CPM-bidders play only undominated strategies, then on every possible value profile the ASP and BSP mechanisms both yield weakly greater revenue than the second-price CPC-bidder-excluding mechanism.*

Proof. Under ASP, BSP, and the second-price CPC-bidder-excluding mechanism, truthfulness is a dominant strategy for CPM-bidders. So assume all CPM-bidders bid truthfully. The revenue in ASP or BSP will be different from that in the second-price CPC-bidder-excluding mechanism if and only if the CPC-bidder/arbitrager's joint bid ($\hat{p}\hat{v}$) is greater than the second highest CPM-bidder value ($b^{(2)}$). And in this case the total revenue to the mechanism will equal either $b^{(1)}$ or $\hat{p}\hat{v}$ (which is $> b^{(2)}$), whereas in the second-price CPC-bidder-excluding mechanism revenue equals $b^{(2)}$. Thus regardless of the strategies chosen by the CPC-bidder and arbitrager, revenue can only be increased. □

Theorem 6. *Assuming all agents play only undominated strategies, then on every possible value profile the ASP and BSP mechanisms both yield weakly greater allocation value (efficiency) than the second-price CPC-bidder-excluding mechanism.*

Proof. Under ASP, BSP, and the second-price CPC-bidder-excluding mechanism, truthfulness is a dominant strategy for CPM-bidders. Under the ASP and BSP mechanisms, *over*-bidding by the CPC-bidder or arbitrager is dominated by truthtelling: if a particular report leads to allocation when truthtelling would not, then negative expected utility will result, and conditional on being allocated the impression, bidding higher than truth can only *increase* the payment that must be made to the center. So if only undominated strategies are played, the allocation under ASP or BSP will differ from that under the second-price CPC-bidder-excluding mechanism only if the CPC-bidder/arbitrager are allocated the impression with a joint bid that is weakly greater than the highest CPM-bid; but this joint bid will be no greater than the *true*

joint value and thus the impression will have been allocated efficiently and allocation value will have been (weakly) improved. □

Proposition 3. *If CPM-bidder values are i.i.d. according to $\mathcal{U}(0,1)$ and only undominated strategies are played, the CPC-bidder's optimal bid in the ASP mechanism is $\frac{n-1}{n}v^*$, regardless of his beliefs about p^*.*

Proof. All agents besides the CPC-bidder have truthfulness as a dominant strategy, so assume they are truthful. Considering Eq. (4), and noting that $F(x)^{n-1} = x^{n-1}$ in the case of the uniform distribution over $[0,1]$, the CPC-bidder's expected utility for bidding $\hat{v}$ when his true value is v^* equals:

$$\begin{aligned} &\int_0^1 f_p(p)p^{n-1}\hat{v}^{n-1}p(v^* - \hat{v})\,dp \\ &= \hat{v}^{n-1}(v^* - \hat{v})\int_0^1 f_p(p)p^n\,dp \end{aligned} \tag{9}$$

Then taking the partial derivative with respect to the CPC-bidder's bid, we get:

$$\begin{aligned} \frac{\partial}{\partial \hat{v}}(Eq.\ (9)) &= ((n-1)v^*\hat{v}^{n-2} - n\hat{v}^{n-1})\int_0^1 f_p(p)p^n\,dp \\ &= \hat{v}^{n-2}((n-1)v^* - n\hat{v})\int_0^1 f_p(p)p^n\,dp \end{aligned}$$

This has roots at $\hat{v} = 0$ and $\hat{v} = \frac{n-1}{n}v^*$, the latter of which is the maximum. Thus, given truthfulness of the CPM-bidders and the arbitrager (which is a dominant strategy for them), if values are distributed uniformly, the CPC-bidder's expected utility is maximized by bidding $\frac{n-1}{n}$ times his true value, regardless of f_p. □

Proposition 4. *If CPM-bidder values are i.i.d. according to $\mathcal{U}(0,1)$ and only undominated strategies are played, the arbitrager's optimal report in the BSP mechanism is $\frac{n-1}{n}p^*$, regardless of his beliefs about v^*.*

Proof. All agents besides the CPC-bidder have truthfulness as a dominant strategy, so assume they are truthful. Considering Eq. (5), the arbitrager's expected utility for reporting $\hat{p}$ when the true probability is p^* equals:

$$\begin{aligned} &\left(\frac{p^*}{\hat{p}} - 1\right)\int_0^1 f_v(v)\int_0^{\hat{p}v} g(c)c\,dc\,dv \\ &= \left(\frac{p^*}{\hat{p}} - 1\right)\int_0^1 f_v(v)\Big[\hat{p}vG(\hat{p}v) - \int_0^{\hat{p}v} G(c)\,dc\Big]\,dv \end{aligned}$$

Noting that $G(x) = x^{n-1}$ in the case of the uniform distribution over $[0,1]$, this equals:

$$\begin{aligned} &\left(\frac{p^*}{\hat{p}} - 1\right)\int_0^1 f_v(v)\Big[\hat{p}^n v^n - \int_0^{\hat{p}v} c^{n-1}\,dc\Big]\,dv \\ &= \left(\frac{p^*}{\hat{p}} - 1\right)\int_0^1 f_v(v)\Big[\hat{p}^n v^n - \frac{1}{n}\hat{p}^n v^n\Big]\,dv \\ &= (\hat{p}^{n-1}p^* - \hat{p}^n)\int_0^1 f_v(v)v^n\frac{n-1}{n}\,dv \end{aligned} \tag{10}$$

Then taking the partial derivative with respect to the arbitrager's report, we get:

$$\frac{\partial}{\partial \hat{p}}(Eq.\ (10)) = ((n-1)p^*\hat{p}^{n-2} - n\hat{p}^{n-1}) \int_0^1 f_v(v)v^n \frac{n-1}{n}\, dv$$

This has roots at $\hat{p} = 0$ and $\hat{p} = \frac{n-1}{n}p^*$, the latter of which is the maximum. Thus, given truthfulness of the CPM-bidders and the uniform distribution over values, the arbitrager's expected utility is maximized by reporting $\frac{n-1}{n}$ times the true probability-of-click, regardless of f_v. □

Theorem 7. *The* SP-CA *mechanism is truthful and efficient in dominant strategies for the captive arbitrager setting, for any number of CPC-bidder/arbitrager pairs.*

Proof. The mechanism chooses the efficient outcome according to agent reports, and strategyproofness holds by the same argument used in Theorem 1: conditional on winning, the winning bidder (or arbitrager) pays an amount independent of his bid (or probability-of-click prediction); and his expected utility for winning is positive if and only if a truthful bid would win. □

Theorem 8. *The* SP-SA *mechanism is truthful and efficient in dominant strategies for the sole arbitrager setting, for any number of CPC-bidders.*

Proof. Again, strategyproofness for all bidders holds by the same argument used in Theorem 1: conditional on winning, the winning bidder pays an amount independent of his bid; and his expected utility is maximized by winning exactly when a truthful bid would win.

For the arbitrager an analogous argument applies. Conditional on a CPM-bidder winning the arbitrager's utility is 0. Conditional on some CPC-bidder h winning, the arbitrager's expected utility equals:

$$p_h^*\hat{v}_h - \max_{i \in I} b_i$$

Conditional on the allocation, this quantity is independent of the arbitrager's reports. It is maximized if $h \in \arg\max_{c \in C} p_c^*\hat{v}_c$, and this is achieved by truthfully reporting $\hat{p}_c = p_c^*$, $\forall c \in C$. Also, this quantity is negative if and only if $p_h^*\hat{v}_h < \max_{i \in I} b_i$, so the arbitrager would never want to win the auction through a non-truthful over-report or lose it through a non-truthful under-report. □

Theorem 9. *The* SP-MA *mechanism is strategyproof in the many arbitrager setting.*

Proof. Strategyproofness for all bidders holds by the same argument deployed in proving the previous theorems. For the arbitragers, a strategy has two components: which CPC-bidders to submit a prediction for, and what predictions to submit given that some prediction is submitted. First consider whether r could ever gain by submitting a prediction for CPC-bidders other than those for whom he has a predicted probability-of-click. Here we see that our model, which takes arbitragers as risk-neutral Bayesian

optimizers, leads to r "having a prediction" for every CPC-bidder and thus submitting a $\hat{p}_c$ for every $c \in C$.

Now, for arbitrary arbitrager r, given a set C_r of CPC-bidders for whom r will submit some prediction, submitting truthful predictions is a dominant strategy. This holds because, given that r wins the impression with CPC-bidder h, his expected utility equals:

$$p^*_h \hat{v}_h - \max_{i \in I \cup \{c \in C : r_c \neq r_h\}} b_i \tag{11}$$

This is (conditionally) independent of the predictions he reports. Note that it *is* dependent on the set C_r of CPC-bidders for whom r makes a prediction—specifically, it can potentially be made *higher* by increasing the size of C_r—but as argued above C_r will equal the complete set anyway. Then Eq. (11) is maximized if, conditional on r winning with *some* CPC-bidder h, $h \in \arg\max_{c \in C} p^*_c \hat{v}_c$. Given the allocation rule this is achieved by the arbitrager truthfully reporting predictions. Moreover, by nature of the allocation rule, Eq. (11) will be non-negative exactly in the case that truthful bidding leads to r winning the auction (with some CPC-bidder), and so r could never gain by bidding untruthfully in order to win (lose) the auction in circumstances other than those under which he wins (loses) with truthful bidding. □