Mechanism Design for Mixed Bidders

Yoram Bachrach
Microsoft Research,
Cambridge, UK
yobach@microsoft.com

Sofia Ceppi
The University of Edinburgh,
Edinburgh, UK
sceppi@inf.ed.ac.uk

Ian A. Kash
Microsoft Research,
Cambridge, UK
iankash@microsoft.com

Peter Key
Microsoft Research,
Cambridge, UK
peter.key@microsoft.com

Mohammad Reza Khani
Microsoft R&D,
Bellevue, USA
rezak@microsoft.com

ABSTRACT

The Generalized Second Price (GSP) auction has appealing properties when ads are simple (text based and identical in size), but does not generalize to richer ad settings, whereas truthful mechanisms such as VCG do. However, a straight switch from GSP to VCG incurs significant revenue loss for the search engine. We introduce a transitional mechanism which encourages advertisers to update their bids to their valuations, while mitigating revenue loss. In this setting, it is easier to propose first a payment function rather than an allocation function, so we give a general framework which guarantees incentive compatibility by requiring that the payment functions satisfy two specific properties. Finally, we analyze the revenue impacts of our mechanism on a sample of Bing data.

General Terms

Algorithmic Game Theory; Mechanism Design; Revenue Optimization

Keywords

Payment Framework; Incentive Compatibility; Online Advertising;

1. INTRODUCTION

Sponsored search is the main source of revenue for most search engines, such as Google, Yahoo! or Bing. Search engines use the Generalized Second Price (GSP) mechanism to select and price ads. In GSP, advertisers are rank-ordered by decreasing bids (more generally by rank score) and slots are assigned in this order. The price of a slot is the minimum bid an advertiser has to make in order to maintain that position, which equates to the next highest bid in the simplest form of GSP. Payment is made when an ad is clicked. The GSP auction's equilibria, bidding strategies, and other properties are well studied (see, e.g., [11, 21]).

However, online advertising is becoming more complex. There may be different ad formats (*e.g.*, text-ads or image-ads) with different sizes, multiple page-templates for search results, and other constraints on showing ads. For these settings, GSP is not well defined and if generalized can be ill-behaved [8].

While such scenarios may be currently rare enough to make GSP tolerable, this appears unlikely to be the case in the future. Therefore, there is an incentive for migrating from GSP to another mechanism. Truthful mechanisms, such as the famous VCG mechanism, are attractive as they gracefully handle such complex scenarios. Truthful mechanisms also remove the computational burden of calculating the optimum bid from advertisers, making the whole system more transparent. Further, they simplify counterfactual analysis (e.g. of auction parameter changes), and more naturally allow the same bid of an advertiser to be used across multiple auctions. Given these advantages, it is perhaps not surprising that Facebook has decided to use VCG and that Google has adopted VCG for contextual ads [22].

The equilibria of GSP result in bid shading, where advertisers bid below their true valuation. A major obstacle to migrating from GSP to VCG is the requirement that advertisers must update their bids, increasing them to their true valuations.

As Varian and Harris [22] note, Google "thought very seriously about changing the GSP auction to a VCG auction during the summer of 2002." One problem was that "the VCG auction required advertisers to *raise their bids above those they had become accustomed to in the GSP auction*." GSP has become entrenched over the past decade. If a switch to VCG were made today, advertisers might not update their bids quickly and even if they did update, it might not be to their true valuation.

We would expect there to be a continuum of advertiser reactions; at one extreme, advertisers using sophisticated optimization techniques would quickly understand the new system and adapt their bids; at the other extreme, advertisers may not check their accounts for months, leaving a pool of advertisers still using their old GSP bids. Thus, for a long period of time there would be advertisers using outdated bids, leading to significant revenue loss.

In this work we introduce a transitional mechanism that attempts to mitigate these problems. Our mechanism operates in a stylized model where there are two types of bidders corresponding to the two extremes: *adaptive* bidders who play an optimal strategy given the mechanism and *non-adaptive* bidders who are still using their GSP bid and, at least in the short-term, do not react to the change of mechanism. Thus, the mechanism is designed to be truthful for adaptive bidders, but not necessarily for non-adaptive bidders who are, after all, still playing as if they are in a non-truthful mechanism.

Of course this stylized model is insufficient for a real transition: advertisers do not come identified as "adaptive" or "non-adaptive". Perhaps the simplest way to think about deploying our mechanism in practice is as a way of doing a staged transition. Initially, the

WWW 2016, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4143-1/16/04.
http://dx.doi.org/10.1145/2872427.2882983 .

search engine is running GSP. Then it picks some arbitrary set of advertisers and classifies them as adaptive, while classifying all others as non-adaptive. Over time, it classifies more and more advertisers as adaptive, until eventually all are adaptive and the search engine is now using VCG. Thus the search engine can manage the transition without dropping all of its price support at once, as a straight transition to VCG would. Viewed this way, our mechanism simply provides a way of implementing a reasonable auction at an intermediate stage where some advertisers have been transitioned to the new mechanism but not all have.

However, our results and nomenclature also suggest that the search engine can do better than simply picking which advertisers to classify as adaptive arbitrarily. In particular, since our mechanism promises good incentive properties for adaptive bidders but not necessarily for non-adaptive ones, our approach will tend to be more effective if bidders who show a tendency to update their bids frequently are among the first to be classified as adaptive.

An ideal deployment would thus look something like first grouping advertisers based on their historical frequency of bid changes and other such data the search engine may have. Then, classify the most frequently changing group as adaptive and all others as non-adaptive. After giving this group some time to adapt their bids, the second most frequent group's classification can be changed to adaptive. This process can then be repeated until all groups are adaptive, although the slower to adapt groups will need to be given more time to adapt before reclassifying the next group.

Our Contribution: With the goal of designing a hybrid mechanism that earns good revenue and which encourages migration, we propose a transitional mechanism which is designed to behave exactly as the current auction when all advertisers are classified as non-adaptive, exactly as VCG when all advertisers are classified as adaptive, and behave sensibly at times in between.

The key difficulty then is to decide how to do the allocation and pricing when some ads are treated as adaptive but others are not. For example, if the bid of an adaptive ad is higher than the bid of a non-adaptive ad, it does not necessarily mean that the adaptive ad should have a better allocation, since we know the non-adaptive ad is shading its bid downwards. The usual approach for designing truthful mechanisms is to give a monotone allocation function and then derive the unique payments using Myerson's lemma [17]. However, in our setting it is not easy to give first an allocation function, as the bid of a nonadaptive ad implies that *payments* of ads above him should be altered (since we want to treat him like a GSP bidder, and GSP is fundamentally a payment rule). Instead, we start by describing the payment rule we want, and then show how to derive a suitable allocation rule.

In fact, we design a framework more general than simply handling adaptive ads and GSP-style non-adaptive ads, which allows for complexities present in real systems. For example, some ads may arrive with pre-established contracts that result in a fixed payment if their ad is clicked. Or, some participants may represent information about non-strategic entities such as organic results which the auction uses to affect the allocation and payments.

To describe this framework more formally, assume ad slots are indexed from 1, the top slot, to n, the bottom one, and there are n advertisers. The designer has to formulate a payment function $p^{(i)} : \mathbb{R}^{n-i} \mapsto \mathbb{R}$ for each $i \in [n]$ which specifies the payment of the advertiser assigned to position i given types of advertisers assigned to positions below him. Our framework requires the payment function satisfy two simple properties: (i) Minimum Marginal Increase (MMI): the payment has to be high enough so that truthful bidders assigned to lower slots do not envy the winner of a higher slot and (ii) Exact Marginal Increase (EMI): the marginal payment increase of the slot directly above a truthful ad has to be equal to the truthful ad's bid. Given a set of payment functions satisfying MMI and EMI, our framework shows how to construct an allocation rule with truthful payments that are exactly those given by the payment functions applied to the realized allocation.

While we primarily focus on deriving a practical rule for transitioning from GSP to VCG, the set of ad auction mechanisms that fit in our framework is quite general. In fact, we prove that by using our framework one can design any truthful mechanism in which the payment of an ad is derived solely from ads below that ad (subject to a few additional requirements). Equivalently, our framework encapsulates mechanisms where raising the bid of an ad does not affect the allocation of ads that were previously allocated below it. More broadly, while the question of what properties of allocation rules lead to truthful mechanisms has been intensively studied (see, e.g., [4, 5, 13]), the question of what properties of payment rules lead to truthful mechanisms has not, so our approach may be of independent interest.

We design a large class of candidate hybrid mechanisms, and then select a particular representative mechanism, analyzing it both theoretically and via simulations using Bing data. On the theory side, we show that transitional behavior is particularly nice if the system starts in the lowest symmetric Nash equilibrium. Under a stylized model, allocations and prices do not change, leading to a painless transition.

In our simulations, we examine transitions both in our stylized model and under somewhat more reasonable assumptions about the transition process. We find that essentially all the costs of a transition are in terms of revenue—the welfare effects for both advertisers and users are small. We also see that the hybrid mechanism can have significant revenue benefits if bidders directly update to true valuations, but these benefits decrease if the bidders fail to do that for various reasons such as not knowing the true valuations, not being utility maximizers, or not trusting the system.

To summarize, our three main contributions are:

1. a new framework for deriving truthful mechanisms from payment rules that satisfy MMI and EMI (Section 4),
2. a specific hybrid mechanism to enable transitioning from GSP to VCG (Section 5),
3. an evaluation of the mechanism using Bing data (Section 6).

2. RELATED WORK

Sponsored search auctions are arguably the most successful recent application of auction theory to a business environment. As a result, much research has been conducted regarding the influence of the mechanism used for the auction on social welfare and the generated revenue. In the case where the VCG mechanism is used, truthfulness is the dominant bidding behavior. However, the same does not hold for the GSP auction and predicting bids in this case is trickier.

A complete information analysis of GSP auctions is discussed by Edelman et al. [11], Varian [21], and Aggarwal et al. [2]. A common theme in this line of work is the equivalence between the auctioneer's revenue and bidders' utility under a VCG auction and under the lowest symmetric Nash equilibrium of a GSP auction (which is sometimes referred to as the "bidder-optimal locally envy free equilibrium"). Ashlagi et al. [6] generalize this, showing that in many auction types in which the payments are a function of the lowest ranked bids, there exists an equilibrium in which bidders' utility is equivalent to their utility under the VCG auction. Roberts et al. [19] generalize this along a different axis, showing that this result also holds for a variety of rank score functions other than simply ranking by highest bid.

Much of the research on equilibria in GSP auctions has focused on symmetric Nash equilibria. Edelman and Schwarz [12] examined the revenue of different symmetric Nash equilibria, noting that under a certain comparison to optimal revenue possible in the Bayesian setting, the "lowest" equilibrium is the reasonable one. A generalized auction proposed by Aggarwal et al. [1] allows advertisers to specify not only a bid but also the positions they desire, ruling out the bottom positions. They show that this auction has a symmetric Nash equilibrium implementing the same outcome (i.e., allocation and pricing) as the VCG auction.

Complementary to studies on symmetric equilibria, several researchers have studied the inefficiency that can result from asymmetric equilibria [15, 9, 16]. Some studies of auction tuning have also explored the full set of equilibria [20]

Taking a Bayesian perspective, Gomes and Sweeney [14] examined the existence and uniqueness of efficient Bayes-Nash equilibria in a GSP auction. Several models have also been proposed for inferring the valuations of advertisers based on the observed bid data [18, 7]. The model by Pin and Key [18] considers advertisers best responding in an uncertain environment in a repeated auction setting, relating the bidding behavior to scenarios when the Bayes-Nash Equilibria of Gomes and Sweeney [14] are known to exist. The model of Athey and Nekipelov [7] starts directly from the Bayes-Nash Equilibria, but has a different model of the information available to the bidders. Instead, Vorobeychik [23] proposed a framework based on agent simulation to approximate Bayes-Nash equilibria in GSP auctions which relies on restricting the space of bidding strategies.

The dynamics leading to the equilibrium outcomes in GSP auctions are less studied. Cary et al. [10] consider dynamics under a greedy bidding strategy, where each bidder chooses the optimal bid for the next round assuming the other bidders do not change their bids. They show this bidding strategy has a unique fixed point, with payments identical to those of the VCG mechanism.

Closest to our work, Aggarwal et al. [3] propose a framework that can implement both the GSP and VCG mechanisms in terms of an assignment game with appropriate models of bidder utility. In their model each bidder specifies a maximum price she is willing to pay and a value for each item. The auctioneer also sets a reserve price for each item. Aggarwal et al. [3] prove that there exist a unique bidder optimal stable assignment in this setting which can be found in polynomial time. They use this assignment as their proposed allocation. Here VCG can be modeled if each bidder submits her true valuation as the maximum price and value to their mechanism. Similarly, GSP can be modeled if each bidder submits her bid as the maximum price and a large number as her value. Their framework can be used to derive a hybrid auction consisting of truthful bidders and max price bidders. Here we cover more general payment functions and allow arbitrary types of bidders.

3. PRELIMINARIES

We study the standard model of a sponsored search auction. There is a set of $\{1, \ldots, n\}$ of ads, denoted by $[n]$. For notational convenience, and without loss of generality, we assume that there are also n ad positions, also called slots, where position i has CTR (Click-Through-Rate) f_i (again for convenience we assume that there are no ad quality scores—the probability of any ad being clicked in slot i is exactly f_i). Without loss of generality we assume that the first position has the highest CTR and the n-th position has the lowest ($0 < f_n < \ldots < f_1 \leq 1$). From now on we treat each ad and the bidder associated with the ad as synonymous. Ads can have different characteristics. For example they could be:

- adaptive ads which bid optimally given the auction design (which given our truthful design means they update their bids to their valuations);
- Generalized Second Price (GSP) ads which assume that they are participating in a GSP auction; or
- first price ads which pay a fixed amount upon being clicked.

We provide a framework for mechanism designers which only requires a payment rule to be specified, and which places special constraints on the payments of adaptive (ADP) ads. Therefore, we specially differentiate between ADP and non-ADP ads and assume each ad has a type taken from set $\mathcal{T} = \{ADP, \text{non-}ADP\} \times \mathbb{R}^+$ that specifies the awareness attribute and bid. Here, non-ADP ads can be of any nature and our framework does not limit the designer's ability for deciding their allocation and payments. This can be thought of as modeling a situation where we are designing a system for the ADP bidders, and therefore care about their incentives. However, there will be some "legacy" bidders whose behavior does not reflect the new system. Since the right way to treat these legacy bidders will depend on their exact nature, we do not put any constraints on what they are charged (and indeed neglect modeling this entirely in proving the theoretical results of Section 4). In section E, we examine what happens when we look to maintain parity with existing GSP prices for non-ADP bidders, and where the non-ADP bidders behave as GSP bidders in equilibrium.

We denote an assignment of ads to slots by the permutation $\Pi = (\pi_1, \ldots, \pi_n)$ where ad π_i is assigned to position i. The (expected) payment of ad π_i is the cost per click for being in position i multiplied by the CTR of position i. Throughout the paper we work with expected payments $p^{(i)}, \forall i \in [n]$ as opposed to cost per click. Further, we assume that all ads have quasilinear utilities, *i.e.*, if ad π_i is assigned to position i and pays $p^{(i)}$ then its utility will be $u_i(\Pi, p) = f_i * v(\pi_i) - p^{(i)}$, where $v(\pi_i)$ is the (true) valuation of ad π_i. Throughout the paper, we use $b(i)$ for the bid of ad i and, if the ad type is ADP, we also use $v(i) = b(i)$ to emphasize that the bid and valuation are the same given that we seek to design truthful auctions.

As discussed in Section 1, this model makes the stylized assumption that ads come labeled as either ADP ads or non-ADP ads, i.e., the nature of each ad is known. Further, our model assumes away the growing richness of the ad ecosystem that is part of the motivation for a mechanism switch. However, at this point basic text ads still represent the bulk of ad impressions, so a solution that works well for them would be useful to aid a near-term transition in anticipation of this future richness.

4. TOP INTERFERENCE FREE PAYMENT FRAMEWORK

In this section we introduce a framework for designing mechanisms in ad auction-like settings where it is easier to provide a payment rule than to give an allocation rule. In other words, we do not know exactly what we want the final allocation of ads to be, but we do know what we want the payment for ad π_i assigned to position i in the overall assignment Π to be. More formally, we explore the space of payment rules which are a set of n functions $\mathcal{P} = \{p^{(i)}\}_{i \in [n]}$, where the function

$$p^{(i)} : \overbrace{\mathcal{T} \times \ldots \times \mathcal{T}}^{n-i} \longmapsto \mathbb{R}^+$$

gives the expected payment of an ADP ad in slot i as $p^{(i)}(\pi_{i+1}, \ldots, \pi_n)$. (Recall that we allow the payments of non-ADP bidders to be arbitrary.) Note that $p^{(i)}$ is the expected payment of the ad assigned to position i and $p^{(i)}/f_i$ is its cost per click.

This formulation implicitly restricts the set of payment rules we consider. Without loss of generality, the payment of an adaptive ad does not depend on its own bid. However, with loss of generality, we also assume that the payment does not depend on the bids of ads assigned to slots above it. This is a natural restriction in a setting without externalities, and indeed one that is satisfied by both the GSP and VCG payment rules.

Our framework specifies two further intuitive properties which we require the payment rule satisfy. We show that these two properties are necessary in the sense that any anonymous mechanism whose payment rule for an ad depends only on ads below it can be implemented using our framework. By anonymous, we mean that permuting the input to the mechanism simply permutes the output (up to tie breaking among ads with identical bids). In order to specify these properties we need to restrict the domain of the payment rules to exclude nonsensical inputs where the ADP bidders are mis-ordered, e.g., where ADP ads are not assigned to slots monotonically with respect to their bid.

Let $\Pi = (\pi_1, \ldots, \pi_n)$ denote the assignment of ads to positions. We use the notation $\Pi^{(k)} = (\pi_k, \ldots, \pi_n)$ to show the partial assignment of the last $n-k+1$ ads to positions k to n. In the following we define what partial assignments are valid and thus form the domain of payment rule $\{p^{(i)}\}_{i\in[n]}$, using $v(i)$ instead of $b(i)$ to emphasize that ADP bids are the true valuations — an assumption that will be validated by Theorem 1 below.

DEFINITION 1 (VALID ORDERING). *A partial assignment of $n-k+1$ ads $\Pi^{(k)}$ is valid if and only if for any $i, j \in \{k, \ldots, n\}$ such that π_i and π_j are ADP ads and $i < j$, we have $v(\pi_i) \geq v(\pi_j)$.*

Recall that Myerson's characterization of truthful mechanisms asks a designer to give a monotone allocation rule and the payments are then uniquely derived from the area above the curve. A monotone allocation rule, when seen from the payment perspective, implies monotone marginal increases of the payments (see Figure 1). We first define the marginal increases:

DEFINITION 2 (MARGINAL OPERATOR $\nabla^{(i,j)}$). *For two positions $i, j \in [n]$ where $i < j$, the marginal increase of payment rule $\mathcal{P}$ for a valid assignment Π is*

$$\nabla^{(i,j)}\mathcal{P}(\Pi) = \frac{p^{(i)}(\Pi^{(i+1)}) - p^{(j)}(\Pi^{(j+1)})}{f_i - f_j}.$$

Now we are ready to specify the first property the payment rule $\mathcal{P}$ should satisfy.

DEFINITION 3 (EXACT MARGINAL INCREASE (EMI)). *The payment rule $\mathcal{P}$ satisfies EMI, if for any valid assignment Π and position $i \in [n-1]$, if π_{i+1} is an ADP ad then*

$$\nabla^{(i,i+1)}\mathcal{P}(\Pi) = v(\pi_{i+1}).$$

The intuition behind the EMI requirement is that, since ADP ads are shown in the order of their bid, the minimum bid ADP ad π_i needed to get shown above the ADP ad π_{i+1} is exactly $v(\pi_{i+1})$. Thus, the marginal payment he should make for being in slot i as opposed to $i+1$ is exactly this minimum bid.

As there are non-ADP ads that can be placed between ADP ads, we need a second property that generalizes EMI to ensure that the payments for ADP ads remain incentive compatible.

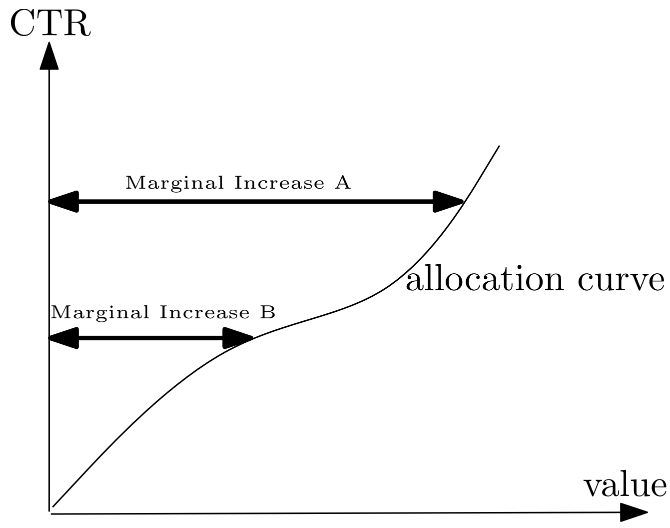

Figure 1: The area above the allocation curve of a winner is his payment. The two arrows show the marginal increase of the payment at different points. If the allocation curve is monotone the marginal increases are also monotone.

DEFINITION 4 (MINIMUM MARGINAL INCREASE (MMI)). *The payment rule $\mathcal{P}$ satisfies MMI if for any valid assignment Π, position $i \in [n]$ such that π_i is a ADP ad, and position $j \in \{1, \ldots, i-1\}$ we have*

$$\nabla^{(j,i)}\mathcal{P}(\Pi) \geq v(\pi_i).$$

Now we give our algorithm to derive the final allocation given payment rule $\mathcal{P}$ which satisfies EMI and MMI. Our algorithm is very simple and intuitive. It starts filling from position n all the way up to position 1. The ADP ads get assigned to positions in increasing order of valuation. Assume that the current ADP ad to be assigned to a position is π. Our algorithm tries to fill all the remaining positions by non-ADP ads which are not yet assigned, choosing the ads sequentially such that the payment of the next position is minimized. Then, our algorithm puts ad π in the position i for which its utility is maximized. The algorithm then takes the next ADP ad and restarts from position $i-1$, fixing all ads in positions i and below. The formal description of our Allocation Algorithm (AA) is given as Algorithm 1 in the appendix.

DEFINITION 5. *A mechanism is* Payment Derived *if it arises from applying the Allocation Algorithm (AA) to a set of payment functions $\mathcal{P} = \{p^{(i)}\}_{i\in[n]}$ that satisfy EMI and MMI.*

The following theorem shows payment derived mechanisms are Incentive Compatible (IC) for ADP ads.

THEOREM 1. *A Payment Derived mechanism is an incentive compatible mechanism for ADP ads.*

The proof is deferred to the appendix.

Having shown that every payment derived mechanism is truthful, it is natural to characterize the mechanisms that are implementable in our framework. We show that this class is characterized by three natural axioms and one technical one.

First note that payment functions $\{p^{(i)}\}_{i\in[n]}$ only use the bid and nature of the ads and do not use the identity (index) of the ads to determine payments[1]. This means that payment derived mechanisms satisfy anonymity, defined formally below.

[1]This would appear to rule out current systems that incorporate

DEFINITION 6. *[Anonymous Mechanism (AM)] A mechanism $\mathcal{M} = (x, p)$ with allocation function x and payment function p is anonymous if the following holds. Let θ and θ' be two type profiles that are permutations of each other (i.e. the set of natures and bids are the same but the identities of ads are permuted) and have no ties. Say $\theta = \sigma(\theta')$, for some fixed permutation σ. Then we have $x(\theta) = \sigma(x(\theta'))$ and $p(\theta) = \sigma(p(\theta'))$. For type profiles with ties we require permutations to permute the output, except that the payments and allocations of tied bidders can be exchanged arbitrarily.*

Secondly, note that the payment of the ad assigned to position i is specified by looking only at the ads that are assigned to positions below i. Therefore, mechanisms in our framework also satisfy the following property.

DEFINITION 7 (TOP INTERFERENCE FREE (TIF)). *A mechanism $\mathcal{M} = (x, p)$ satisfies TIF, if when an ad changes its type and gets a better position then the allocation of ads assigned to lower positions remains unaltered. More formally, let $x(\theta)$ be the allocation given by x on type profile $\theta = \{\theta_1, \ldots, \theta_n\}$ and $x(\theta')$ the allocation given by x on type profile θ' where $\theta'_h = \theta_h, \forall h \in [1, \ldots, k-1, k+1, \ldots, n]$ and $\theta'_k \neq \theta_k$. Assume that ad k is in position i with allocation $x(\theta)$ and in position j with allocation $x(\theta')$ such that $j < i$. Mechanism $\mathcal{M}$ satisfies TIF if the ads assigned to positions $i+1$ to n are the same in both allocations $x(\theta)$ and $x(\theta')$.*

In the following theorem, we prove that our framework can implement all mechanisms that are incentive compatible, anonymous, and top interference free, as well as satisfying an additional technical axiom (2T), which essentially requires that the mechanism is well-behaved when considering the top two slots (and one which seems to be satisfied for reasonable mechanisms). Hence, requiring EMI and MMI does not restrict the designer in ways that current standard designs such as VCG and GSP do not.

THEOREM 2. *A mechanism is Payment Derived if and only if it satisfies IC (for ADP ads), AM, TIF, and 2T.*

See Appendix C for a definition of the technical axiom (2T), a proof of the theorem, and a discussion of 2T.

5. PRICING FUNCTIONS

In the preceding section, we designed a general framework. In this section, we apply it to the desired special case of transitioning from a GSP auction to a VCG auction. Here a GSP mechanism simply assigns ads in the order of their rank score (determined by a function y) and charges each ad the smallest bid for which it maintains its assigned position (see [19, Eqn 2.1]). To do so, we must decide how to price GSP (i.e., non adaptive) bidders, since our framework is silent about how they should be charged. We make the simplest decision, namely to charge them the same amount an adaptive bidder would be charged in the same slot, and show that this has several desirable properties. For a particular mechanism GSP mechanism G^y with rank score function y, this results in a mechanism $\mathcal{M}(G^y)$.

We begin by discussing what happens when all users are utility maximizers but the mechanism classifies some as ADP and others as (non-ADP) GSP bidders. When the hybrid mechanism is first put into use, all bidders are classified as GSP bidders, but as time goes on bidders will be reclassified. Given perfect rationality, this means i's bid changes from $b(i)$ to $v(i)$. We can then show that if GSP bidders begin from the *lowest revenue* Symmetric Nash Equilibrium (SNE) then the revenue and allocations are unaltered, provided a particular hybrid pricing function is employed, irrespective of the order in which users are selected for reclassification. (Since SNE rank bidders in decreasing order of bid, in this section we assume that ad i is in slot i.) Formally, we have the following proposition whose proof is deferred to the appendix.

PROPOSITION 1. *For any GSP mechanism G^y, if GSP bidders bid as in the lowest revenue SNE and ADP bidders bid truthfully, the revenue, allocation and prices paid in $\mathcal{M}(G^y)$ will be independent of the number and identity of ADP and GSP bidders if and only if the payment function satisfies MMI and EMI and further satisfies*

$$p^{(i-1)}(\Pi^{(i)}) = \begin{cases} p^{(i)}(\Pi^{(i+1)}) + v(i)(f_{i-1} - f_i) \\ \qquad\qquad\qquad\qquad \text{if } \theta_i = (ADP, v(i)) \\ b(i) f_{i-1} \qquad\qquad \text{if } \theta_i = (GSP, b(i)) \end{cases} \tag{1}$$

As long as G^y admits an SNE, the proposition applies. The following corollary results from applying a sufficient condition [19] for this.

COROLLARY 1. *Proposition 1 holds for any GSP mechanism G^y that uses a rank score of the form*

$$y(b, i) = \max\big(0, g(i)b - h(i)\big)\,, \tag{2}$$

where g and h are arbitrary non-negative values that can depend on i.

The necessary and sufficient conditions only hold when we start from the *lowest* SNE. For example if we are in another SNE, then moving just one bidder i from GSP to ADP will not change the position or prices paid by i or those below i, but potentially changes prices (and hence positions) of bidder(s) above i (since Equality (10) in the proof of the proposition does not hold for $i-1$ anymore). Hence, we want to construct price functions that satisfy EMI and MMI when other equilibria hold, and for more general non-truthful prices. Specifically, we shall consider two natural examples, where the pricing functions for ad i are the same for adaptive and non-adaptive i. First,

$$\mathbf{A}: \quad p^{(i-1)}(\Pi^{(i)}) = \max\Big(p^{(i)}(\Pi^{(i+1)}) + v_{max}(\Pi^{(i)}) (f_{i-1} - f_i), b_{max}(\Pi^{(i)}) f_{i-1}\Big) \tag{3}$$

where

$$v_{max}(\Pi^{(i)}) \stackrel{def}{=} \max\{v(\theta_j) : j \geq i \cap \theta_j = (ADP, v(j))\} \tag{4}$$

$$b_{max}(\Pi^{(i)}) \stackrel{def}{=} \max\{b(\theta_j) : j \geq i \cap \theta_j = (GSP, b(j))\} \tag{5}$$

are the largest ADP valuation and GSP bid at or below i, respectively, and

$$\mathbf{B}: \quad p^{(i-1)}(\Pi^{(i)}) = \max\Big(p^{(\arg\max v(\Pi^{(i)}))}(\Pi^{(\arg\max v(\Pi^{(i)})+1)}) + v_{max}(\Pi^{(i)})(f_{i-1} - f_{\arg\max v(\Pi^{(i)})}), b_{max}(\Pi^{(i)}) f_{i-1}\Big) \tag{6}$$

where

$$\arg\max v(\Pi^{(i)}) \stackrel{def}{=} \arg\max_j \{v(j) : j \geq i \cap \theta_j = (ADP, v(j)\} \tag{7}$$

is the identity of the largest ADP ad at or below i.

other information such as click probability and other quality measures into a rank score. However, our results still apply in this more general setting as long as the rank scores are treated anonymously.

Either of these hybrid auctions is consistent with Proposition 1, and so in fact they are identical in this case. It is easy to see, however, that they do differ in other scenarios. In any "reasonable" extension of GSP, an advertiser ought to pay at least the bid of a GSP bidder below him, and **B** is the lowest set of prices consistent with this, EMI, and MMI. At the other extreme, **A** may charge GSP bidders prices higher than their bids, which **B** is guaranteed not to. See Appendix E for further discussion. In our simulations, we therefore use rule **B**.

6. SIMULATION RESULTS

We saw in Proposition 1 that if bidders always play the lowest SNE, bidders adapt their bid immediately on reclassification, and that adjustment consists of instantly switching to the advertiser's true value, then there would be no effect on efficiency or revenue from switching to the hybrid auction. Of course, these assumptions are not realistic. In this section, we discuss a variety of simulations that analyze the practical effects of the hybrid auction in more realistic scenarios.

6.1 Simulation Setup

We base our simulations on a *non-random* sample of Bing data on 3984 auctions. It is a filtered subset of a larger random sample that ensures the auctions are "interesting." In particular, we wanted thick auctions (with at least 12 participants), and with other properties such that techniques for inferring true values from GSP bids could give reasonable answers. The metrics have been normalized. Nevertheless, we believe the sample is representative enough to allow a meaningful exploration of our approach.

We restrict each auction to the top 12 participants, and only actually run an auction for the top three slots. In order to run our simulations we need to have an estimate of true valuations of GSP ads. One estimate of true valuations is to assume that GSP ads have played the minimum symmetric Nash equilibrium and invert their bids to their valuations. In this case, we would essentially be baking in the first of the assumptions from Proposition 1, so unsurprisingly the transition would happen without any changes as the allocation and payment of ads remain identical at each point of time.

Instead, we use the stochastic formulation from Pin and Key [18]. This approach derives the valuations under the hypothesis that each advertiser chooses her bid to maximize her expected net utility under the assumption that she faces a stationary bid distribution. In our calculations we assume that the CTRs are known, with the opposing bid distribution estimated from the opponents' empirical bid distributions.

We simulate the following four different mechanisms, re-running the simulations 10 times to derive standard errors.

- **GSP.** The first mechanism is GSP run on the original set of bids when no updates have happened. This represents the current state of the world and serves as a benchmark to which the other approaches can be compared.

- **VCG-V.** The second mechanism is VCG run on the final set of true valuations when all the ads have updated their bids. This represents the ideal end state when all bidders have transitioned to being truthful. It also serves as a sanity check on the reasonableness of our value estimation (i.e. it should display similar performance to GSP).

- **HYBRID.** The third one is a payment derived mechanism, using pricing rule **B** described in (6).

- **VCG-B.** Finally the last mechanism is VCG run on the current set of bids when some ads have updated their bids and some have not. This is the obvious alternative strategy for transitioning: simply transition directly to VCG and wait for bidders to catch up.

6.2 Perfectly Rational Bidders

In our first set of simulations, bidders are chosen in a random order to be "active". When an ad is active it (a) becomes classified as ADP by HYBRID and (b) is perfectly rational and immediately updates its bid to its true value under both VCG-B and HYBRID. Thus, the primary assumption we are relaxing is that bidders are playing the lowest SNE.

Figure 2 shows the normalized average revenue, welfare, and click yield for different mechanisms during the transition. 95% confidence intervals are also plotted. The estimated revenue from ultimately running VCG (i.e. VCG-V) is close to GSP, which is consistent with the reasonableness of our value estimation procedure. Immediately switching to VCG (curve VCG-B) results in a significant revenue drop, which is steadily recovered as more advertisers update their bids. In contrast, there is a more modest revenue drop under the hybrid mechanism (since bidders are not always following the lowest SNE). In particular, revenue always dominates directly switching, substantially so in the initial time steps. In both the estimated welfare and click yield there are no significant differences between VCG-B and Hybrid auction. The observation that welfare and click yield do not differ much in the Hybrid auction and in the VCG-B strengthens the importance of the revenue improvement that the former has over the latter because it is not coming at the cost of other important metrics. Note also that, in the worst case, the drop in welfare is less than 1.5% and the drop in click yield is less than 0.3% from the optimal case. Thus, we focus on revenue in our subsequent simulations.

6.3 Cautious Bidders

The second set of simulations relaxes the idea that bidders are willing to immediately jump to their true value, no matter how large a bid increase this implies. Instead we parameterize them with a triple (p, q, i). At each time step, bidders decide randomly whether to update their bid, doing so with probability p. If their consumer surplus decreased in the last step (i.e. because the bid changes of others changed their slot or increased their price) they update with a higher probability q. This allows us to model advertisers who are attentive only when needed. Finally, when they update, they increase their bid by a percentage i until they reach their true value. Bidders are treated as ADP as soon as they change their bid for the first time.

Different parameterizations lead to somewhat different pictures, but all share the same general trends as in Figure 2. Note however, that the x-axes are on different scales since it now takes significantly more than 12 rounds for all bidders to fully adapt. Figure 3, with parameters (0.3,0.6,0.1) shows that these cautious updates hurt the performance of the hybrid relative to a direct switch to VCG. There is still a benefit for the first 8 rounds, but then essentially all bidders are classified as ADP, so performance is the same as if we had switched directly.

This observation is the basis for the approach we suggest in the introduction, of transitioning bidders in an ordering based on how active they are. Figure 4 shows that the benefits persist longer if we have bidders who are lax about updating (unless something bad happens) with parameters (0.1,0.9,0.1). Larger values of i (not shown) lead to more of the benefits of the hybrid approach being maintained, since the period when a bidder is not GSP but not yet truthful is shortened.

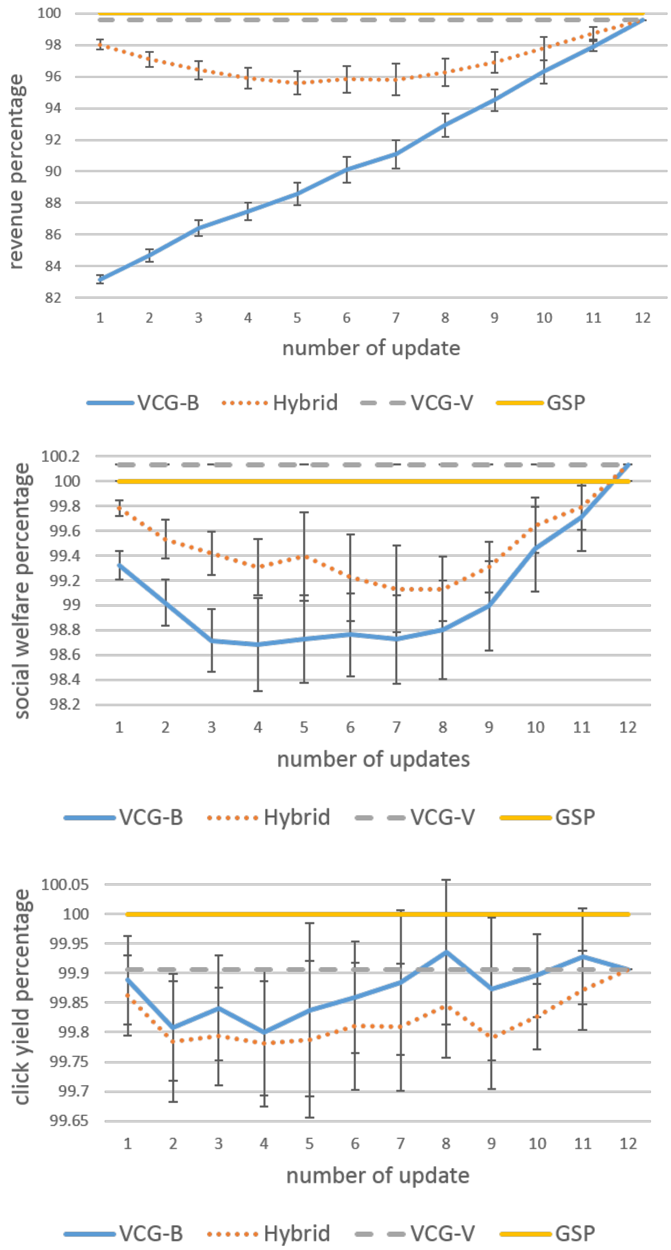

Figure 2: Perfectly Rational Bidders

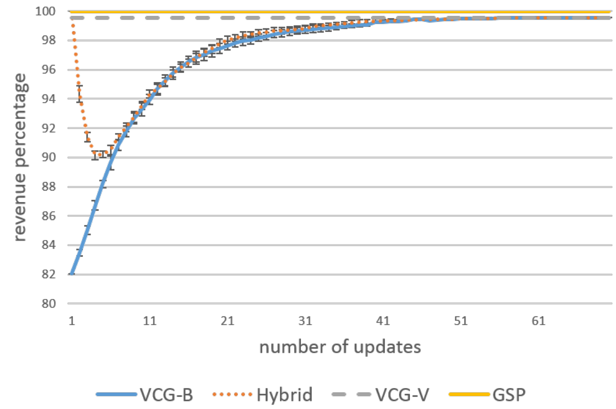

Figure 3: Cautious Bidders

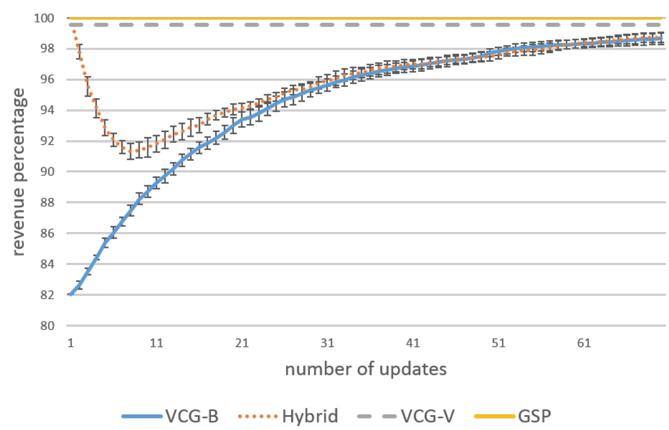

Figure 4: Cautious Bidders, Lax Updates

7. CONCLUDING REMARKS

Our motivation and focus is the sponsored search setting, where ad-slots are auctioned off. We have examined a possible transition from a GSP based auction to a VCG auction, and noted that a simple switch to the VCG mechanism is likely to cause a dramatic loss of revenue, as some advertisers would keep using their old and shaded GSP bid. We aimed to provide a pathway for migrating from one non-truthful mechanism (GSP) to a truthful mechanism whilst mitigating the revenue loss that would occur if there was a switch to a truthful mechanism but bidders did not immediately update (increase) their bids to their true valuation.

We have proposed a hybrid mechanism approach which allows for a staged transition between a GSP auction and a VCG auction in sponsored search settings, which in particular allows maintaining high revenue during the transition period. Our approach allows rolling out the new mechanism to increasingly larger proportions of the advertiser population; when all advertisers are considered non-adaptive the mechanism behaves exactly as the current auction, and when the transition is complete, and all advertisers are adaptive ones, the mechanism behaves exactly as a VCG mechanism; in between these two points, the mechanism behaves sensibly so as to maintain good revenue.

In contrast to the standard approach for designing truthful mechanisms, we do not start with a a monotone allocation function and then derive the unique payments using Myerson's lemma [17]. Instead, our key ingredient is a *payment function*, which maps the bids for lower value goods to the payment for a higher value good, and which needs to satisfy two properties. The two properties, EMI and MMI relate solely to the bids of truthful agents, and place constraints on the discrete derivative of the payment function.

We have given details of a "bottom-up" allocation procedure, which when used with an EMI and MMI payment function gives an Incentive Compatible (IC) mechanism, and hence gives incentive for bidders to change type to truthful. If in addition, the mechanism is "Top Interference Free", TIF , then this characterization is both necessary and sufficient for IC anonymous mechanisms which satisfy an additional technical axiom. Any mechanism derived from a "bottom-up" procedure, such as standard GSP or VCG, are all examples of TIF mechanisms.

We have provided an empirical evaluation based on simulations indicating that our approach does indeed mitigate the revenue loss during the transition.

A number of questions remain open for future research. First, can alternative transition mechanisms achieve a better revenue retention during the transition? There are at least two obvious alternatives to our approach. One option is to directly switch to VCG, which was we have shown requires a much more dramatic initial revenue loss at each stage of the transition. It is possible that it could still lose less money overall, if it causes the transition to take less time, but this seems unlikely, particularly as our approach appears to give stronger incentives for bidders classified as ADP to update their bids. (For example, an immediate switch to VCG keeps allocations unchanged with lower prices, relying on the price signal to encourage advertisers to increase their bid while in our mechanism becoming classified as ADP results in a drop of both price and clicks.) Nevertheless, this is a simple approach that has been successfully used in at least one instance [22]. Another option is to switch to VCG but add a layer that attempts to effectively raise bids on behalf of legacy bidders through some estimate of their value. While this approach sounds appealing, there are a number of difficulties in practice. What should we do if this causes a bidder to end up with a payment larger than her bid? When should we stop this prediction and how does this effect the incentives of bidders to update their bid (particularly with regard to a payment-capping style approach to resolving the first difficulty)?

Second, how easy is it to extend our approach to other auction settings or richer domains? Our approach of deriving mechanisms from payments seems quite general. Are there other natural settings covered by our existing framework? Can it be extended to domains that do not satisfy TIF by adding additional structure to the payment rules or changing the allocation algorithm?

APPENDIX

A. ALLOCATION ALGORITHM (AA)

Description of the algorithm. Set T contains all the ADP ads which are not yet assigned. Similarly set N contains all the non-ADP ads which are not yet assigned. In Line 3 we initialize the value of ℓ which keeps the index of current position to be filled. In Line 5 we select a ADP ad with minimum value in order to assign it to a position. In Line 7 we provisionally fill the next $|N|$ positions with non-ADP ads. In Lines 8 and 9 we find and assign a position with the best profit for ADP ad π. In Line 10 we remove all the non-ADP ads which are assigned permanently (appear after the position i) from N. In Line 13 we fill the remaining positions by the rest of non-ADP ads. Finally at Line 14 we set the payments of allocated ads according to $p^{(i)}$.

Note that at Lines 5 and 19, we might have multiple valid choices, in which case we break the ties by the choosing the ad with the smallest index. The only other instance where a tie can happen is at Line 8, when we select the largest feasible j.

Algorithm 1: Allocation Algorithm (AA)

input : n ads $\{1, \ldots, n\}$ and payment rule $\{p^{(i)}\}_{i\in n}$.
output: assignment $(\pi_1, \ldots, \pi_n)$ of ads to positions and their payments.

1 $T \leftarrow$ `Extract-ADP-ads`$(\{1, \ldots, n\})$;
2 $N \leftarrow$ `Extract-NonADP-ads`$(\{1, \ldots, n\})$;
3 $\ell \leftarrow n$;
4 **while** $|T| > 0$ **do**
5 Let $\pi \in T$ be a ADP ad with minimum value ;
6 Remove π from T;
7 `Fill-With-NonADP-ads` ;
8 $i \leftarrow \arg\max_{j\in\{\ell-|N|,\ell-|N|+1,\ldots,\ell\}} f_j \cdot v(\pi) - p^{(j)}(\pi_{j+1}, \ldots, \pi_n)$;
9 $\pi_i \leftarrow \pi$;
10 $N \leftarrow N - \{\pi_{i+1}, \ldots, \pi_\ell\}$;
11 $\ell \leftarrow i - 1$;
12 **end**
13 `Fill-With-NonADP-ads` ;
14 Set the payment of π_i to be $p^{(i)}(\pi_{i+1}, \ldots, \pi_n)$;

`/* Sub-procedure Fill-With-NonADP-ads provisionally assigns all the remaining non-ADP ads to next available positions. At each step it selects a non-ADP ad which makes the next payment as small as possible. */`

15 `Fill-With-NonADP-ads`:
16 **begin**
17 $N' \leftarrow N$;
18 **for** $i \leftarrow \ell$ **downto** $\ell - |N| - 1$ **do**
19 $\pi_i \leftarrow \arg\min_{\pi\in N'} p^{(i-1)}(\pi, \pi_{i+1} \ldots, \pi_n)$;
20 $N' \leftarrow N' - \{\pi_i\}$
21 **end**
22 **end**

B. PROOF OF THEOREM 1

PROOF. Let $\mathcal{M}$ be the resulting mechanism after applying AA to a set of payment functions. Observe that $\mathcal{M}$ assigns ADP ads to positions in the increasing order of their value, *i.e.*, the larger the value of a ADP ad is, the higher position he receives. This follows from Line 5 of AA.

In order to prove incentive compatibility of mechanism $\mathcal{M}$, we show that an arbitrary ADP ad gets the best utility when he bids his true valuation. Assume that θ is an arbitrary type profile, $\mathcal{M}$ outputs assignment $\Pi = (\pi_1, \ldots, \pi_n)$ for θ, and π_k is a ADP ad. We show that utility of π_k does not increase if he bids v' considering three cases: (1) he is considered in the same iteration of Algorithm 1, (2) he is considered in a later iteration, and (3) he is considered in an earlier one.

Case (1): Since he is considered in the same iteration, all that changes is that Line 8 optimizes with respect to v' rather than v, giving him a weakly worse position. Thus, he does not benefit.

Case (2): Since he is considered in a later iteration, some other ADP with value $v'' \geq v$ is considered in his original iteration and assigned to slot k''. By MMI, $\nabla^{(k',k'')}\mathcal{P}(\Pi') \geq v'' \geq v$. Thus, his marginal payment for all the clicks he gets beyond what he would get in slot k'' is at least his value, and he is no better off than he would have been originally taking slot k'', which is a contradiction.

Case (3): Without loss of generality, let π_k be the bidder in the lowest slot (according to π) who can benefit from lowering his bid. Let k'' be the highest slot below k such that $\pi_{k''}$ (with value v'') is ADP. By the taxation principle, there is a price that $\pi_{k''}$ faced for every slot at or below k'', and at those prices he preferred k''. π_k could have faced those same prices by bidding $v'' - \epsilon$ for sufficiently small ϵ, and as $v \geq v''$ he too prefers slot k'' among all those options. By EMI, $\nabla^{(k''-1,k'')}\mathcal{P}(\Pi) = v'' \leq v$. Thus, he weakly prefers taking slot $k'' - 1$ to taking slot k''. Since slot

$k'' - 1$ was one of his options, he weakly prefers slot k to it, a contradiction. □

C. PROOF OF THEOREM 2

Before we begin, we observe that our requirements ensure that the mechanism allocates adaptive ads in increasing order of value. Recall that in our framework IC only applies to ADP ads.

OBSERVATION 1. *If $\mathcal{M}$ satisfies IC, AM, and TIF then it allocates ADP ads in increasing order.*

PROOF. Suppose for contradiction that $v_1 < v_2$ but ad (ADP, v_1) got a better slot than ad (ADP, v_2). By IC, if we replace (ADP, v_1) by another copy of (ADP, v_2), its slot can only improve. By TIF, this means that the slot of the original copy is unchanged. But by AM, if the original ads had been permuted the ad with value v_1 could raise its bid to v_2 and receive a worse slot, contradicting IC. □

We begin with sufficiency. Let $\mathcal{M}' = (x', p')$ be a mechanism with allocation function x' and payment function p' which satisfies IC, AM, TIF. We use $\mathcal{M}'$ to propose payment rule $\mathcal{P}$ such that a payment derived mechanism $\mathcal{M}$ using $\mathcal{P}$ is equivalent to $\mathcal{M}'$. In order to deal with technicalities of ties, we assume that in the case of ties $\mathcal{M}'$ use the same tie breaking rule as our framework.

Let $\Pi^{(k)} = (\pi_k, \dots, \pi_n)$ for $k \in \{3, \dots, n\}$ be the assignment of $\mathcal{M}'$ for positions k to n. We say that ad i is less than j with regard to $\Pi^{(k)}$ (show by $i \prec_{\Pi^{(k)}} j$), if there exist a type profile θ such that

- The allocation $x'(\theta)$ is the same as $\Pi^{(k)}$ for positions k to n.
- $i, j \in \theta$.
- Ad i is assigned to position $k-1$ and ad j is assigned to a position better than i ($x'(i) > x'(j)$).

Intuitively, $i \prec_{\Pi^{(k)}} j$ means that fixing $\Pi^{(k)}$ allocation x' prefers to assign i to position k over j. In the following lemma we prove that $\prec_{\Pi^{(k)}}$ is a total order over all the ads which can be assigned to position $k-1$ fixing $\Pi^{(k)}$. It turns out that TIF is almost, but not quite strong enough to prove this lemma. The difficulty is that it has no "bite" when applied to the case of $k = 3$ (i.e. the final 2 slots). That is, ad in slot 2 has no (non-fixed) ads below it (so the definition is vacuous), while an ad in slot 1 that changes type and remains in slot 1 forces the other ad to stay in slot 2 (again making the definition vacuous). Thus, we need a property that ensures the mechanism is well-behaved in this case.

DEFINITION 8 (TWO TRANSITIVE(2T)). *$\mathcal{M}$ is two transitive if for all choices of $\Pi^{(3)}$ the relation $\prec_{\Pi^{(3)}}$ is transitive.*

Thus, we further assume for our proof of sufficiency that $\mathcal{M}'$ satisfies 2T.

LEMMA 1. *The relation $\prec_{\Pi(k)}$ is antisymmetric (if $i \prec_{\Pi^{(k)}} j$ then $j \not\prec_{\Pi^{(k)}} i$), total, and transitive.*

PROOF. We prove the lemma by contradiction. To show antisymmetry, let θ be a type profile for which $i \prec_{\Pi^{(k)}} j$ and θ' be a type profile for which $j \prec_{\Pi^{(k)}} i$. Consider a sequence of type profiles that are intermediate between θ' and θ in the sense that the transition from one profile to the next results from changing the type of a single ad from its value in θ' to its value in θ. Let $\theta' = \theta_1, \theta_2, \dots, \theta_{a-1}, \theta_a = \theta$ be the sequence of type profiles.

We show that this sequence maintains the following invariant: the ad in slot $k-1$ is not i and has already changed its type. Clearly this is true for θ', since j is in slot $k-1$ in $x(\theta')$ and its type will never change. Suppose it is true for θ_b, and let ad e be the ad that changes type between θ_b and θ_{b+1}. By our invariant, e is not in slot $k-1$ of $x(\theta_b)$. Thus, by TIF, if the ad in slot $k-1$ changes from $x(\theta_b)$ to $x(\theta_{b+1})$ it must be that in $x(\theta_{b+1})$ the ad in slot $k-1$ is in fact e, which is not i and has already changed its type. This shows that i is not is slot $k-1$ of $x(\theta)$, contradicting our assumption.

To see that the relation is total, take some θ and i and j. Create θ' by replacing all ads other than i and j shown in a slot above k in $x(\theta)$ with a copy of either i or j. Thus, by an inductive argument that shows this does not change the allocation below slot $k-1$, some copy of i or j must be allocated to slot $k-1$.

Finally, transitivity follows via a similar construction. If $k > 3$, simply transform θ to θ' by replacing all ads above slot k with copies of one of the relevant i, j, or ℓ. If $k = 3$, transitivity is by assumption (i.e. 2T). □

Now we are ready to specify how we build set of payment functions $\mathcal{P} = \{p^{(i)}\}_{i \in [n]}$ from $\mathcal{M}'$. Let $\Pi^{(k)}$ be a valid assignment of ads to positions k to n. By Lemma 1 we have a total ordering of all possible candidate ads for position $k-1$, so set $p^{(k-1)}(\Pi^{(k)})$ to be the infimum of ADP ads among those candidates (i.e. the minimum bid an ADP ad could make and be shown in this position).

Now we prove that the payment-derived mechanism using the payment rule $\mathcal{P}$ we have constructed always gives the same allocation and payments as $\mathcal{M}'$ which finishes the proof of this sufficiency.

LEMMA 2. *The payment-derived mechanism $\mathcal{M} = (x, \mathcal{P})$ is the same as $\mathcal{M}' = (x', p')$.*

PROOF. We need only verify that $\mathcal{M}$ always gives the same allocation as $\mathcal{M}'$ as the fact that the payments are the same (at least up to a constant) then follows via revenue equivalence (recall that we only care about the payments of ADP bidders). (The constant can be matched by changing p to p'' to include this constant shift.)

Now we prove by contradiction that the allocation functions x and x' are the same. Let θ be a type profile for which $x(\theta) \neq x(\theta')$. Let $\Pi^{(k)} = (\pi_k, \dots, \pi_n)$ be the largest common suffix of $x(\theta)$ and $x(\theta')$ and assume for now that $p^{(k-1)}(\Pi^{(k)})$ is finite. Let e be the ad assigned to position $k-1$ in $x(\theta)$ and e' be the ad assigned to position $k-1$ in $x'(\theta)$. Note that

$$e' \prec_{\Pi^{(k)}} e \tag{8}$$

since $x'(\theta)$ assigned positions $k-1$ to e'.

Because $x(\theta)$ assigns position $k-1$ to e as opposed to e' this means that

$$p^{(k-2)}(e, \pi_k, \dots, \pi_n) < p^{(k-2)}(e', \pi_k, \dots, \pi_n)$$

(Recall Lines 5 and 19 of algorithm AA). This means that there exists a ADP ad x with valuation $\nabla^{(k-2,k-1)}\mathcal{P}((e, \pi_k, \dots, \pi_n)) < v(x) < \nabla^{(k-2,k-1)}\mathcal{P}((e', \pi_k, \dots, \pi_n))$. Now if we replace the rest of ads with ADP bidders with valuation $v(x)$ then they appear before e but after e'. This contradicts with Equation 8 and the fact that $\prec_{\Pi^{(k)}}$ is a total order for any $\Pi^{(k)}$.

Now we deal with the case where $p^{(k-1)}(\Pi^{(k)})$ is infinite. Intuitively, this is the case where only non-adaptive ads can be shown before the suffix $\Pi^{(k)}$. Since prices are all infinite, we need a way for the algorithm to match the order that $\mathcal{M}'$ chooses. We do this by allowing prices of the form (∞, a), where a is the type of a non-ADP ad. The total order $\prec_{\Pi(k)}$ then gives a well defined notion of the lowest price as the one whose a is lowest according to that ordering. With this enlarged set of prices, the proof proceeds as before. □

Finally, the necessary part is easy to prove. Let $\mathcal{M}$ be a payment derived mechanism. The AM and TIF properties follow by the fact that in algorithm AA (see Algorithm 1) when assigning the next ADP ad, AA uses only its value and neither its index nor the value of higher ADP ads. The IC property of $\mathcal{M}$ is the result of Theorem 1. 2T follows from the greedy nature of the allocation.

C.1 Discussion of Two Transitivity (2T)

Two transitivity is a technical assumption. The intuition is that we require the mechanism to be well-behaved when considering the top two slots, which TIF is not strong enough to enforce. If all non-ADP ads are of the same nature, a sufficient condition is that $\mathcal{M}'$ is monotone for non-ADP ads (MN). That is, if a non-ADP ad raises its bid, it gets a (weakly) better slot.

LEMMA 3. *IF $\mathcal{M}$ satisfies IC, AM, TIF, and MN and all non-ADP ads have the same nature then it satisfies 2T*

PROOF. By AM and IC/MN, $\prec_{\Pi(3)}$ is transitive if all 3 ads are ADP or non-ADP respectively. Thus, WLOG let 2 be ADP and 1 be non-ADP (replace IC by MN below if only 1 is ADP). There are 3 cases.

Case 1: $N \prec V_1 \prec V_2$. By IC, $N \prec V_2$ (otherwise an adaptive ad V_1 could raise its bid and go from slot 1 to slot 2).

Case 2: $V_1 \prec N \prec V_2$. By AM+IC, $V_1 \prec V_2$ (otherwise an adaptive ad V_2 could raise its bid and go from slot 1 to slot 2 when facing N).

Case 3: $V_1 \prec V_2 \prec N$. By IC, $V_1 \prec N$ (otherwise an adaptive ad V_1 could raise its bid and go from slot 1 to slot 2). $\square$

Such a nice sufficient condition is not obvious if there is more than 1 nature of non-ADP ad. Non-degenerate examples still appear to satisfy 2T, but we do not know of a less technical way to explain the way in which they are non-degenerate. To see why, consider an example with 2 slots. If the bidders have at least one adaptive ad, this becomes a form of second price auction. However, when there are two non-adaptive ads of different natures, nothing obviously constrains the rule for determining the order in which they are shown in a way that corresponds to enforcing transitivity. This same example shows why we require MN in Lemma 3. Without it we would be equally at a loss in this situation.

D. PROOF OF PROPOSITION 1

PROOF. Both directions of the proof follow almost directly from the definitions of a lowest SNE, EMI and MMI. In order for payments of ADP and GSP bidders to be identical, the payment functions $p^{(i)}(\Pi^{(i+1)})$ must be the same regardless of whether i is ADP or GSP, and independent of the mix of bidder types in $\Pi^{(i+1)}$. Consequently no GSP or ADP bidder wants to change bid or position, since by the definition of an SNE

$$\left(v(i) - p^{(i)}(\Pi^{i+1})\right) f_i \geq \left(v(i) - p^{(j)}(\Pi^{j+1})\right) f_j \quad \text{for all } i, j. \tag{9}$$

By standard arguments about the lowest SNE (see, e.g., [21, 19]) we in fact have that for all i,

$$b(i) f_{i-1} = b(i+1) f_i + v(i)(f_{i-1} - f_i). \tag{10}$$

Thus, by induction, the two conditions of (1) are in fact equal at the lowest SNE. This gives that the form is necessary and sufficient for prices to coincide, as pricing must be equivalent to the case $\theta_j = (GSP, b(j))$ for all $j \geq i$. As this outcome is equivalent to the outcome of a truthful auction, it follows that EMI and MMI are satisfied as well. $\square$

E. FURTHER DISCUSSION OF PRICING RULES

OBSERVATION 2. *When setting prices according to* (3)*, GSP bidders may pay more than their bid.*

PROOF. Consider the following example. There are three non adaptive advertisers (1, 2, and 3) whose bids are $b(1) = 12$, $b(2) = 11$, and $b(3) = 10$, and one adaptive advertiser (4) whose bid is $b(4) = v(4) = 22$. There are four available positions, with $f_1 = 1$, $f_2 = 0.7$, $f_3 = 0.6$, and $f_4 = 0.5$. The resulting allocation is then $\Pi = (1, 2, 3, 4)$. When setting prices according to (3), the advertiser in the first position pays $p^{(1)} = 13.6$, which is higher than his bid. (The payments for the other positions are $p^{(2)} = 7$, $p^{(3)} = 2.2$, and $p^{(4)} = 0$.) $\square$

We say that an ADP bidder is indifferent between slots $i+1$ and i if his utility is the same for both positions during the course of our allocation procedure.

OBSERVATION 3. *When setting prices according to either* (3) *or* (6)*, GSP bidders pay exactly their bid when an indifferent ADP bidder is put immediately below them.*

PROOF. Let b be the value of a GSP bidder that is assigned to slot i and v be the value of an adaptive bidder assigned to slot $i+1$. We have

$$f_i(v - b) = f_{i+1} v - p^{(i+1)}(\Pi^{(i+2)}).$$

Rewriting shows that the GSP bidder pays exactly his bid. $\square$

OBSERVATION 4. *When setting prices according to* (6)*, GSP bidders never pay more than their bid.*

PROOF. When bidder i is being priced by another GSP bidder (i.e. the second branch of the pricing rule applies), this is trivially true as GSP bidders are allocated in order of bid. Otherwise, let $j = \arg\max v(\Pi^{(i)})$. Let b be the value of the GSP bidder assigned to slot i and v be the value of the ADP bidder assigned to slot j. Then

$$f_i(v - b) \leq f_j v - p^{(j)}(\Pi^{(j+1)}).$$

Rewriting shows that the GSP bidder pays at most his bid. $\square$

Thus, we have seen that rule **B** possesses the desirable property that GSP bidders never pay more than their bids. But at the cost of charging both GSP and ADP bidders less than rule **A** in general. The situations in which rule **A** causes a GSP bidder to pay more than his bid are actually quite specific, and in unreported simulations we found them to be rare. Since our primary goal is to mitigate revenue loss, there is a reasonable case to use rule **A**, perhaps with a cap to ensure no bidder is ever charged more than his bid, although this may have some incentive implications depending on how bidders are classified as ADP vs GSP.

Finally, we note that this distinction between rules **A** and **B** shows off the generality of our framework relative to previous work. It turns out that the natural way of applying the approach of Aggarwal et al. [3] to accommodate both ADP and GSP bidders results in rule **B**, while our approach allows the options of considering rule **A** (and convex combinations of the two), as well as being adaptable to other considerations like first-price bidders or organic results participating in the allocation rule.

References

[1] G. Aggarwal, J. Feldman, and S. Muthukrishnan. Bidding to the top: Vcg and equilibria of position-based auctions. In *Approximation and Online Algorithms*, pages 15–28. Springer, 2007.

[2] G. Aggarwal, A. Goel, and R. Motwani. Truthful auctions for pricing search keywords. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 1–7. ACM, 2006.

[3] G. Aggarwal, S. Muthukrishnan, D. Pál, and M. Pál. General auction mechanism for search advertising. In *Proceedings of the 18th international conference on World wide web (WWW)*, pages 241–250. ACM, 2009.

[4] A. Archer and R. Kleinberg. Truthful germs are contagious: a local to global characterization of truthfulness. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 21—30, 2008.

[5] I. Ashlagi, M. Braverman, A. Hassidim, and D. Monderer. Monotonicity and implementability. *Econometrica*, 78(5):1749—1772, 2010.

[6] I. Ashlagi, D. Monderer, and M. Tennenholtz. Mediators in position auctions. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 279–287. ACM, 2007.

[7] S. Athey and D. Nekipelov. A structural model of sponsored search advertising auctions. In *Sixth ad auctions workshop*, 2010.

[8] Y. Bachrach, S. Ceppi, I. A. Kash, P. Key, and D. Kurokawa. Optimising trade-offs among stakeholders in ad auctions. In *ACM Conference on Economics and Computation, EC '14, Stanford , CA, USA, June 8-12, 2014*, pages 75–92, 2014.

[9] I. Caragiannis, C. Kaklamanis, P. Kanellopoulos, and M. Kyropoulou. On the efficiency of equilibria in generalized second price auctions. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 81–90. ACM, 2011.

[10] M. Cary, A. Das, B. Edelman, I. Giotis, K. Heimerl, A. R. Karlin, C. Mathieu, and M. Schwarz. Greedy bidding strategies for keyword auctions. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 262–271. ACM, 2007.

[11] B. Edelman, M. Ostrovsky, and M. Schwarz. Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords. 2005.

[12] B. Edelman and M. Schwarz. Optimal auction design and equilibrium selection in sponsored search auctions. *The American Economic Review*, pages 597–602, 2010.

[13] R. M. Frongillo and I. A. K. and. General truthfulness characterizations via convex analysis. In *10th Conference on Web and Internet Economics (WINE)*, 2014. To Appear.

[14] R. Gomes and K. Sweeney. Bayes–nash equilibria of the generalized second-price auction. *Games and Economic Behavior*, 86:421–437, 2014.

[15] R. P. Leme and É. Tardos. Sponsored search equilibria for conservative bidders. In *Fifth Workshop on Ad Auctions*, 2009.

[16] B. Lucier, R. Paes Leme, and E. Tardos. On revenue in the generalized second price auction. In *Proceedings of the 21st international conference on World Wide Web*, pages 361–370. ACM, 2012.

[17] R. B. Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6:58–73, 1981.

[18] F. Pin and P. Key. Stochastic variability in sponsored search auctions: observations and models. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 61–70. ACM, 2011.

[19] B. Roberts, D. Gunawardena, I. A. Kash, and P. Key. Ranking and tradeoffs in sponsored search auctions. In *Proc. of the Fourteenth ACM Conf. on Electronic Commerce*, EC '13, pages 751–766, 2013.

[20] D. R. Thompson and K. Leyton-Brown. Revenue optimization in the generalized second-price auction. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 837–852, New York, NY, USA, 2013. ACM.

[21] H. R. Varian. Position auctions. *international Journal of industrial Organization*, 25(6):1163–1178, 2007.

[22] H. R. Varian and C. Harris. The VCG auction in theory and practice. *The American Economic Review*, 104(5):442–445, 2014.

[23] Y. Vorobeychik. Simulation-based game theoretic analysis of keyword auctions with low-dimensional bidding strategies. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 583–590. AUAI Press, 2009.