Double Marginalization in Performance-Based Advertising: Implications and Solutions

Chrysanthos Dellarocas
Boston University
School of Management
dell@bu.edu

First draft: August 2009. Current draft: March 2010

Abstract

An important current trend in advertising is the replacement of traditional pay-per-exposure (pay-per-impression) pricing models with performance-based mechanisms in which advertisers pay only for measurable actions by consumers. Such pay-per-action (PPA) mechanisms are becoming the predominant method of selling advertising on the Internet. Well-known examples include pay-per-click, pay-per-call and pay-per-sale. This work highlights an important, and hitherto unrecognized, side-effect of PPA advertising. I find that, if the prices of advertised goods are endogenously determined by advertisers to maximize profits net of advertising expenses, PPA mechanisms induce firms to distort the prices of their goods (usually upwards) relative to prices that would maximize profits in settings where advertising is sold under pay-per-exposure methods. Upward price distortions reduce both consumer surplus and the joint publisher-advertiser profit, leading to a net reduction in social welfare. They persist in current auction-based PPA mechanisms, such as the ones used by Google and Yahoo. In the latter settings they always reduce publisher revenues relative to pay-per-exposure methods. In extreme cases they also lead to rat-race situations where, in their effort to outbid one another, advertisers raise the prices of their products to the point where demand for them drops to zero. I show that these phenomena constitute a form of double marginalization and discuss a number of enhancements to today's PPA mechanisms that restore equilibrium pricing of advertised goods to efficient levels.

Keywords: performance-based advertising, sponsored search, keyword auctions, double marginalization, mechanism design

Acknowledgments: Many thanks to Jason Hartline, Lorin Hitt, De Liu, Michael Schwarz, Siva Viswanathan and seminar participants at Harvard, Wharton, Tepper, Toulouse IDEI and Boston University for valuable feedback on an earlier version of this work.

Half the money I spend on advertising is wasted;
the trouble is, I don't know which half.
John Wanamaker, owner of America's first department store

1 Introduction

John Wanamaker's famous quote has been haunting the advertising industry for over a century. It now serves as the motivation behind much of the innovation taking place in Internet-based advertising. From Google, Yahoo and Microsoft, to Silicon Valley upstarts, some of the best and brightest technology firms are focusing a significant part of their energies on new mechanisms to reduce advertising waste. These come in many forms but have one thing in common: a desire to replace traditional *pay-per-exposure* (also known as *pay-per-impression*) pricing models, in which advertisers pay a lump sum for the privilege of exposing an audience of uncertain size and interests to their message, with performance-based mechanisms in which advertisers pay only for measurable *actions* by consumers. Pay-per-click sponsored search, invented by Overture and turned into a multi-billion dollar business by Google, Yahoo and other online advertising agencies, is perhaps the best known of these approaches: advertisers bid in an online auction for the right to have their link displayed next to the results for specific search terms and then pay only when a user actually clicks on that link, indicating her likely intent to purchase. Pay-per-call, pioneered by firms such as Ingenio (acquired by AT&T in 2007), is a similar concept: the advertiser pays only when she receives a phone call from the customer, usually initiated through a web form. Pay-per-click and pay-per-call are viewed by many as only an intermediate step towards what some in the industry consider to be the "holy grail of advertising": the pay-per-sale approach where the advertiser pays only when exposure to an advertising message leads to an actual sale. All of these approaches are attempting to reduce all or part of Wanamaker's proverbial waste by tying advertising expenditures to consumer actions that are directly related or, at least, correlated with the generation of sales. In the rest of the paper I will refer to them collectively as *pay-per-action* (PPA) pricing models.

The current surge in pay-per-action advertising methods has generated considerable interest from researchers in a variety of fields including economics, marketing, information systems and computer science. This is important and timely since most of these methods have been invented by practitioners and their properties and consequences are not yet fully understood. Although the literature (surveyed in Section 2) has made significant advances in a number of areas, an important area that, so far, has received almost no attention is the impact of various forms of PPA advertising on the prices of the advertised products. With very few exceptions (also discussed in Section 2), papers in this stream of research have made the assumption that the prices of the goods being advertised are set exogenously and independently of the advertising payment method.

In this paper I make the assumption that the prices of the goods being advertised are an endogenous decision variable of firms bidding for advertising resources. I find that, if such prices are endogenously determined by advertisers to maximize profits net of advertising expenses, PPA advertising mechanisms induce firms to distort the prices of their goods (usually upwards) relative to the prices that would maximize their profits in settings where there is no advertising or where advertising is sold under pay-per-exposure methods. Upward price distortions reduce both consumer surplus and the joint publisher-advertiser profit, leading to a net reduction in social welfare.

My results are driven by the fact that the publisher and the advertisers are independent profit maximizing agents that interact in a "supply chain" type of relationship (i.e. the publisher supplies an advertising slot to an advertiser who then uses it to advertise and sell a product to consumers). When both of these agents independently set the marginal prices they charge to their respective downstream customers to optimize their profit margins, an inefficiency somewhat akin to the well-known concept of double marginalization occurs.

The phenomenon is very general and occurs in any setting where the publisher and advertisers have some degree of market power. Specifically, I show that such price and revenue distortions also arise in search advertising auctions, such as the ones currently used by Google and Yahoo. In the latter settings they *always* lead to lower publisher revenues relative to pay-per-exposure methods. Price distortions are more severe in settings where the set of competing advertisers have similar valuations for the advertising slot. In the extreme case where the set of competing advertisers have identical valuations, they lead to rat race situations where, in their effort to outbid one another, advertisers raise the prices of their products to the point where demand for them drops to zero.

I discuss how today's auction-based search advertising mechanisms can be enhanced to restore efficient pricing. I propose two broad sets of ideas. The first idea requires the advertiser to disclose to the publisher both her bid as well as the product price she intends to charge. The publisher then makes the advertiser's current round quality weight a function of both her past history (which is what the current generation of PPA keyword auctions already do) and her current round product price (something current mechanisms *do not* do). The second idea is to construct mechanisms whereby advertisers are charged some form of penalty, e.g. an auction participation fee, following periods where no triggering actions (i.e. clicks, sales, etc.) occur. I show that my enhanced mechanisms maximize the joint publisher-advertiser profit and have identical revenue and allocation properties to an optimal pay-per-exposure mechanism.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 introduces the paper's core intuitions through a simple posted price example setting. Section 4 illustrates methods of restoring efficient pricing in the context of my example setting. Section 5 shows how the ideas of the previous sections apply to the practically important case of auction-based search advertising. Section 6 proposes ways of restoring efficient product pricing in auction-based search advertising. Finally, Section 7 concludes.

2 Related Work

This work relates to a number of important streams of marketing, economics, information systems and computer science literature. Nevertheless, the phenomenon discussed herein has so far not been addressed by any of these literatures.

Sponsored search advertising

Pay-per-click online advertising, such as sponsored search links, is one of the most successful and highly publicized methods of performance-based advertising. It is the main source of revenue for sites like Google and Yahoo and one of the fastest growing sectors of the advertising industry. Not surprisingly, this field has experienced an explosion of interest by both researchers and practitioners. Important advances have been made on understanding the properties of the *generalized second price* (GSP) auction mechanisms that are currently the prevalent method of allocating advertising resources in such spaces (see, for example, Athey and Ellison 2008; Edelman et al. 2007; Varian 2007). A related stream of research has proposed several extensions to baseline GSP auctions that aim to improve their properties. The following is an illustrative subset: Aggarwal et al. (2006) propose an alternative advertising slot auction mechanism that is revenue-equivalent to GSP but induces truthful bidding (GSP does not); Feng et al. (2007), Lahaie and Pennock (2007) and Liu and Chen (2006) explore the allocative efficiency and publisher revenue implications of alternative methods for ranking bidders, including "rank by bid" and "rank by revenue"; Katona and Sarvary (2008) explore the equilibrium behavior of keyword auctions under more sophisticated assumptions about users' search behavior; Ashlagi et al. (2008) and Liu et al. (2008b) explore auction design in the presence of competing publishers.

Growing attention is also being given to the perspective of advertisers bidding on such auctions; the most important problems here include how to identify appropriate keywords (Abhishek and Hosanagar 2007; Joshi and Motwani 2006; Rutz and Bucklin 2007) and how to dynamically allocate one's budget among such keywords (Borgs et al. 2007; Cary et al. 2007; Feldman et al. 2007; Rusmevichientong et al. 2006). Finally, researchers have paid attention to incentive issues that are inherent in pay-per-action advertising, most important among them being the potential for click fraud, i.e. the situation where a third party maliciously clicks on an advertiser's sponsored link without any intention of purchasing her product (Immorlica et al. 2006; Wilbur and Zhu 2008) as well as the advertiser's incentive to misreport the frequency of her triggering action in order to avoid paying the publisher (Agarwal et al. 2009; Nazeradeh et al. 2008).

The above theoretical and algorithmic contributions are complemented by a growing number of empirical works (e.g. Animesh et al. 2009; Ghose and Yang 2009; Goldfarb and Tucker 2007; Rutz and Bucklin 2008; Yao and Mela 2009b). For comprehensive overviews of current research and open questions in sponsored search auctions the reader is referred to excellent chapters by Feldman et al. (2008), Lahaie et al. (2007), Liu et al. (2008a) and Yao and Mela (2009a).

Interestingly, almost all papers on this burgeoning field assume that an advertiser's value per sale is exogenously given and do not consider how the performance-based nature of advertising affects the advertiser's pricing of the products being sold. The only two exceptions I am aware of are Chen and He (2006) and Feng and Xie (2007). Chen and He (2006) study seller bidding strategies in a paid-placement position auction setting with endogenous prices and explicit consumer search. However, they only assume a pay-per-exposure mechanism and derive results that are essentially identical to my Proposition 7 i.e. (using the language of my paper) that advertisers price their product at the point that maximizes their per-exposure value function. Feng and Xie (2007) focus on the quality signaling aspects of advertising and propose a model that is in many ways orthogonal to mine. I discuss their paper later in this section.

Performance-based contracting

Performance-based advertising is a special case of performance-based contracting. Contract theory has devoted significant attention to such contracts, as they can help balance incentives in principal-agent settings where moral hazard exists or where the sharing of risk between the two parties is a concern (Holmstrom 1979; Holmstrom and Milgrom 1987, 1991). In the context of information goods, Sundararajan (2004) studies optimal pricing under incomplete information about the buyers' utility. He finds that the optimal pricing usually involves a combination of fixed-fee and usage-based pricing. Closer to the context of this work, Hu (2006) and Zhao (2005) study how performance-based advertising contracts that optimally balance the incentives of both the publisher and the advertise can be constructed. They both find that the optimal contract must have both a fixed and a performance-based (i.e. PPA) component. Once again, however, both of these studies consider the prices of advertised products as fixed and not as an endogenous decision variable under the control of advertisers.

Advertising and Product Prices

The relationship between product prices and advertising has received quite a bit of attention in the economics and marketing literature. These literatures have primarily focused on the quality signaling role of prices in conjunction with advertising. The main result is that the simultaneous presence of prices and advertising improves a firm's ability to successfully signal its quality to consumers because firms can partially substitute quality-revealing price distortions with quality-revealing advertising expenditures (see, for example, Fluet and Garella 2002; Hertzendorf and Overgaard 2001; Milgrom and Roberts 1986; Zhao 2000). Almost all works in this stream of literature assume that advertising is sold under a traditional pay-per-exposure model.

The only exception I am aware of is Feng and Xie (2007). They study how the move from exposure-based to performance-based advertising affects the ability of price and advertising to signal product quality. Their main result is that such a move generally reduces the number of situations where advertising expenditures can be used to signal quality and increases the prices charged to consumers, since firms must now rely harder on the price signal to reveal their quality. Their result relies on the assumption that higher quality firms are more likely to have a higher proportion of repeat customers who would be clicking and purchasing the product anyway, but who nevertheless induce incremental advertising charges in a performance-based model. Therefore, performance-based advertising is relatively more wasteful for high quality vs. low quality firms and this moderates a high quality firm's incentive to spend more on advertising.

My results are orthogonal to this work since in my model price distortions are unrelated to the advertisers' desire to signal their quality or to the presence of repeat customers; they occur even in settings where consumers have perfect knowledge of each advertiser's quality or where repeat customers do not exist. As I explain in Section 3, my results are driven by the fact that the publisher and the advertisers are independent profit maximizing agents that interact in a "supply chain" type of relationship (i.e. the publisher supplies an advertising slot to the advertiser who then uses it to advertise and sell a product to consumers). When both of these agents independently set the marginal prices they charge to their respective downstream customers to optimize their profit margins, an inefficiency akin to the well-known concept of double marginalization occurs.

In summary, the traditional literature on advertising has examined various aspects of the relationship between product prices and advertising expenditures in settings that essentially correspond to what I call pay-per-exposure. On the other hand, the rapidly growing literature on performance-based advertising has largely assumed that product prices are exogenous to the choice of advertisement payment mechanism. This work breaks new ground by showing that when one endogenizes product prices, performance-based advertising mechanisms create incentives for price distortions that have negative consequences for most stakeholders.

3 The core intuitions

In this section I establish the connection between the double marginalization problem and price distortions in performance-based advertising through the use of a example posted price setting. I also review some of the classic solutions to the problem.

3.1 The double marginalization problem

Consider a setting where an upstream monopolist manufacturer sells raw materials to a retailer who, in turn, packages them and sells them to consumers. The retailer is a monopolist in the downstream market. The manufacturer's constant marginal production cost is denoted by c_m and the retailer's marginal packaging cost by c_r. The manufacturer sells the product to the retailer at wholesale price w and the retailer sells the product to end consumers at a retail price p. (Figure 1a). Demand is assumed to be linear, taking the form $D(p) = d - p$, where d is assumed to be greater than total marginal cost $c = c_m + c_r$.

If the manufacturer and retailer are vertically integrated, the retail price p would be chosen to maximize the total channel profit $\pi(p) = D(p)(p - c)$. It follows that the profit maximizing retail price equals $\frac{d+c}{2}$, yielding an efficient sales quantity of $\frac{d-c}{2}$ and a channel profit of $\frac{(d-c)^2}{4}$. The manufacturing and retail departments can then negotiate a transfer price w to divide the total channel profit.

If the manufacturer and retailer are not vertically integrated, they independently choose their respective prices to maximize their own profits. In this setting, the manufacturer moves first by offering a wholesale price w. Consequently, the retailer faces a profit of $\pi_r(p|w) = D(p)(p - c_r - w)$ and responds by choosing a profit maximizing retail price $\frac{d+c_r+w}{2}$. Anticipating the retailer's reaction, the manufacturer chooses w to maximize his profit function $\pi_m(w|p) = D(p)(w - c_m) =$

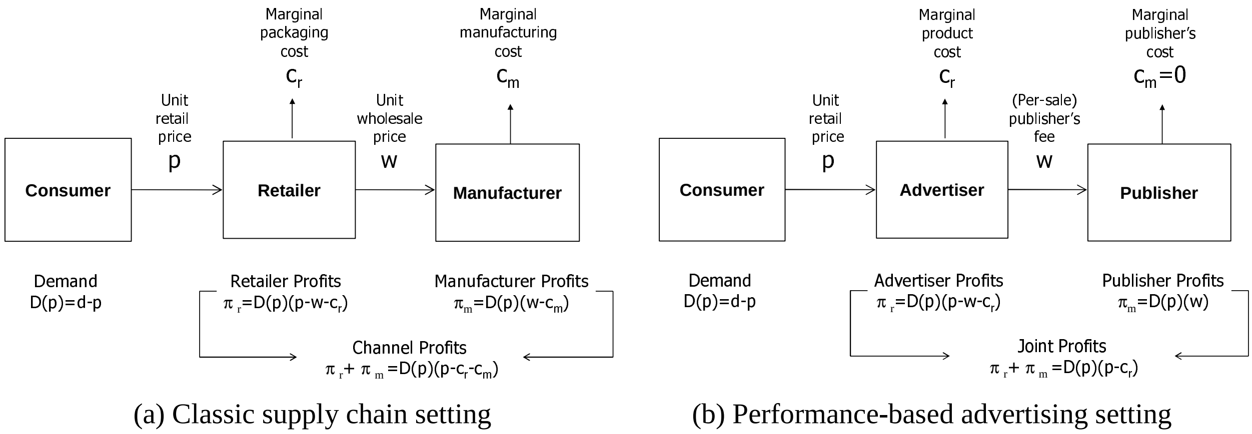

Figure 1: Analogies between a classic supply chain setting and performance-based advertising.

$\left(d - \frac{d+c_r+w}{2}\right)(w - c_m)$. The profit maximizing wholesale price then equals $\frac{d+c_m-c_r}{2} > c_m$ (under the assumption $d > c = c_m + c_r$), inducing a retail price $\frac{3d+c}{4}$ that exceeds the efficient retail price level of $\frac{d+c}{2}$. Consequently, the manufacturer ends up with a profit of $\frac{(d-c)^2}{8}$ and the retailer earns $\frac{(d-c)^2}{16}$. The total profit realized in this channel shrinks to $\frac{3(d-c)^2}{16}$, which corresponds to only 75% of the integrated channel profit. Consumers also end up clearly worse off since they pay a higher retail price.

The above inefficiency stems from the double price distortion which occurs when two independent firms (manufacturer, retailer) stack their price-cost margins, thus the term *double marginalization.* The phenomenon has been identified by Cournot (1838) even though its first concrete analysis is most often attributed to Spengler (1950). It is easy to show that it occurs under any downward sloping demand curve $D(p)$ and that it is also not confined to monopoly settings; it is present in any setting where the manufacturer and retailer each have *some* market power and therefore, mark up their respective prices above marginal cost.

3.2 A survey of classic solutions

An extensive literature has proposed solutions to the double marginalization problem (see, for example, Jeuland and Shugan 1983; Rey and Vergé 2008). Most solutions rely on the manufacturer imposing some vertical restraints to the retailer. For instance, including a *resale price maintenance* provision in a retail contract can directly stipulate the efficient retail price. Alternatively, a *quantity forcing* contract specifies the minimum volume of sales the retailer must achieve. Setting this to the integrated channel efficient quantity forces the retailer to choose the efficient price level. Both such contracts involve the manufacturer directly interfering with the retailer's pricing decision and are often frowned upon, or legally unenforceable in some countries.

A less intrusive solution is to use a *two part tariff.* Specifically, the manufacturer charges the retailer both a fixed franchise fee f and a marginal wholesale price w. The retailer profit function becomes $\pi_r(p|w, f) = D(p)(p - c_r - w) - f$. As before, for a given f the profit maximizing retail price is $\frac{d+c_r+w}{2}$ with the corresponding retailer profit being $\frac{(d-c_r-w)^2}{4} - f$. The manufacturer can set the franchise fee up to $\frac{(d-c_r-w)^2}{4} - u$ where u is the reservation utility of the retailer. The manufacturer thus chooses w to maximize $\pi_m(w|p,f) = D(p)(w-c_m)+f = \left(d - \frac{d+c_r+w}{2}\right)(w-c_m)+\frac{(d-c_r-w)^2}{4} - u$. The profit maximizing wholesale price then is $w = c_m$, inducing the efficient retail price of $p = \frac{d+c}{2}$ and an efficient integrated channel profit equal to $\frac{(d-c)^2}{4}$. The central idea of two part tariffs in this setting is to eliminate one of the margins: the manufacturer sets his price equal to marginal cost and uses the franchise fee to gain his share of the surplus. As a result, the retailer faces a margin that is identical to that of an integrated channel. With the price distortion removed channel efficiency is restored.

The use of *quantity discount* contracts is yet another method for solving the double marginalization problem. For example, consider a contract that specifies a unit wholesale price equal to $w = \frac{f}{D} + g$, where f, g are non-negative constants and D is the total quantity bought by the retailer. With such a scheme, the higher the consumer demand (and, therefore, the quantity ordered by the retailer) the lower the retailer's average cost. Since higher demand implies lower retail price, this provides the retailer with incentives to keep her price low. In fact, it is easy to see that this price contract is equivalent to a two-part tariff: for both contracts the total cost to the retailer for a given quantity D is $f + Dg$. The quantity discount scheme can be viewed as an average-cost version of a two-part tariff. Consequently, the reasoning outlined in the previous paragraph applies here as well. If the manufacturer sets $g = c_m$, the retailer's equilibrium price will equal the retail price at the efficient level. The manufacturer and retailer can then divide the surplus by appropriately setting f.

3.3 Analogy with performance-based advertising

In this section I show how the previous setting has a direct analogy to an advertising context. Instead of a manufacturer, consider a monopolist publisher who owns an advertising resource, such as a billboard located at a busy city square, a time slot in prime time TV, or space at the top of a popular web page. Instead of a retailer, consider a monopolist advertiser who produces a product or service at marginal cost c_r. Since the advertising resource is an information good, I assume that the publisher's marginal cost c_m is zero. I further assume that *at least some* consumers will not be aware of the product unless the product is advertised on the publisher's resource. Therefore, gaining access to the publisher's resource results in *incremental* demand for the product equal to $D(p)$ (Figure 1b).

Consider first a setting where access to the advertising resource is made using *pay-per-exposure* (PPE) contracts. In such contracts the publisher charges the advertiser an upfront flat fee f that is independent of the actual revenue that the advertiser is able to realize thanks to the advertising resource. The advertiser's profit function is then $\pi_r(p|f) = D(p)(p - c_r) - f$. It is easy to see that, in such a setting, the publisher's fee f does not impact the advertiser's pricing problem. The advertiser will choose the price that maximizes her incremental sales profit. For example, if $D(p) = d - p$ (where $d > c_r$) the retail price would be set to $\frac{d+c_r}{2}$, yielding an efficient sales quantity of $\frac{d-c_r}{2}$ and a channel profit of $\frac{(d-c_r)^2}{4}$ that is split between the advertiser and the publisher through the use of the flat fee

f.

Let us now see how the situation changes when the publisher sells access to the advertising resource using performance-based contracts. The most straightforward example of such contracts is a *pay-per-sale* (PPS) contract whereby the advertiser pays the publisher a fee w for every sale realized. In this setting the advertiser's profit function is equal to $\pi_r(p|w) = D(p)(p-c_r-w)$ and the publisher's profit function equal to $\pi_m(w|p) = D(p)w$. The reader can readily verify that the above profit functions are identical to those of the classic manufacturer-retailer double marginalization example of Section 3.1 when $c_m = 0$. With reference to that example, if $D(p) = d-p$, at equilibrium the publisher will set his per-sale fee w at $\frac{d-c_r}{2}$ and the advertiser will price her product at $\frac{3d+c_r}{4}$. Notice that this price is higher than the sales profit maximizing price $\frac{d+c_r}{2}$ that the advertiser would choose for her product if advertising was sold using a PPE contract. Accordingly, it leads to lower demand. The publisher ends up with a profit of $\frac{(d-c_r)^2}{8}$ and the advertiser earns $\frac{(d-c_r)^2}{16}$. The total profit realized in this channel shrinks to $\frac{3(d-c_r)^2}{16}$, which corresponds to only 75% of the channel profit if advertising was sold using PPE contracts. Consumers also end up clearly worse off since they pay a higher retail price.

As was the case with the classic double marginalization problem, the above phenomena are fairly general: they occur for any downward sloping demand curve $D(p)$ and are also present in any setting where the publisher and advertiser each have *some* market power and are, therefore, able to mark up their respective prices above marginal cost. Finally, they are orthogonal to the specific mechanism used by the publisher to allocate his resource to advertisers. More concretely, Section 5 shows that similar price distortions are present in auction-based allocation methods, such as the ones used by search engines to allocate online advertising space.

3.4 Generalization to arbitrary pay-per-action contracts

Pay-per-sale is only one example of a more general class of performance-based advertising contracts that are often referred to by the term pay-per-action. This section shows how the phenomena identified in the previous sections generalize to this broader class of contracts.

Pay-per-action (PPA) approaches make payment to the publisher contingent on a *triggering action* that is either a sale, or some other consumer action (e.g. click, call) that has the following properties:

1. It is linked to the advertising, i.e. only consumers who have been exposed to this particular advertising message can perform the triggering action.

2. It is a *necessary* step of a consumer's purchase decision process. This means that even though not all consumers who perform the triggering action may buy the product, consumers cannot purchase the product without performing the triggering action.

Let $V(p)$ denote the *ex-ante* value the advertiser expects to obtain from leasing the advertising resource. In most real-life settings this value will be equal to the incremental sales profit that the advertiser expects to realize by leasing the resource and can thus be expressed as $V(p) = D(p)(p-c_r)$

where p is the unit price of the advertised product, c_r is the corresponding unit cost and $D(p)$ is the increase in demand due to advertising. Value $V(p)$ will be split between the advertiser and the publisher and will, thus, also be referred to as the joint advertiser-publisher profit. Irrespective of the precise functional form of $V(p)$, assumptions 1-2 above allow us to uniquely express it as a product $V(p) = U(p)W(p)$ where:

$U(p)$ is the expected *triggering action frequency* (TAF) and

$W(p)$ is the expected *value-per-action* (VPA).

The precise meanings of $U(p)$ and $W(p)$ depend on the specifics of the payment contract. For example:

- In pay-per-sale contracts $U(p)$ is equal to the incremental demand due to advertising while $W(p)$ is equal to the unit profit.
- In pay-per-click contracts $U(p)$ is equal to the per-period audience size times the probability of a click (the *clickthrough rate*) and $W(p)$ is equal to the conditional probability of a purchase given a click (the *conversion rate*) times the unit profit.
- Traditional per-per-exposure contracts are a special case of the above framework where $U(p) = 1$ and $W(p) = V(p)$.

A special case of particular importance in online settings is that of personalized ads, such as the ones displayed in response to a keyword search. In such settings the size of the audience is one. The above framework applies to this special case as well, the only difference being that function $U(p)$ would then denote the triggering action *probability*.

In the rest of the paper I am assuming that the publisher has a way of obtaining reliable observations of triggering actions and, therefore, that strategic misreporting of triggering actions from the part of the advertiser is not an issue.[1]

Consider now a general *pay-per-action* (PPA) contract whereby the advertiser pays the publisher a fee w every time a triggering action occurs. In this more general setting the advertiser's profit function is equal to $\pi_r(p|w) = U(p)(W(p) - w) = V(p) - U(p)w$ and the publisher's profit function equal to $\pi_m(w|p) = U(p)w$. Let p^* be the price that maximizes $V(p)$. It is easy to see that this is the price that maximizes $\pi_r(p|w)$ when $w = 0$. It is also (see Section 3.3) equal to the product price p^E that the advertiser would choose under a pay-per-exposure contract.[2] Now, in any PPA setting where the publisher has some market power, he will charge a price $w > 0$ (otherwise, he will make no profits). From monotone comparative statics theory (Topkis 1998) it follows that if the cross-partial

[1] Addressing an advertiser's incentive to misreport the frequency of payment triggering action to the publisher is an important consideration in pay-per-action schemes but orthogonal to the focus of this paper. See Agarwal et al. (2009) and Nazerzadeh et al. (2008) for discussion and proposed solutions.

[2] Throughout the paper I am using superscripts to denote the payment mechanism that quantities of interest are associated with. For example, p^E denotes the product price advertisers will choose when advertising is sold using a pay-per-exposure (E) mechanism whereas p^A denotes the product price when advertising is sold using a pay-per-action (A) mechanism.

derivative $\frac{\partial \pi_r^2}{\partial w \partial p}$ is positive (negative) everywhere, the profit maximizing price $p^A = \arg\max_p \pi_r(p|w)$ is a monotonically increasing (decreasing) function of w. It is $\frac{\partial \pi_r^2}{\partial w \partial p} = -U'(p)$, which is positive if and only if $U'(p) < 0$ for all p. Under these conditions, if $w > 0$ it will be $p^A > p^* = p^E$, leading to joint advertiser-publisher profit $V(p^A) < V(p^*) = V(p^E)$. I have thus shown the following result:

Proposition 1: *In all PPA settings where (i) both the publisher and the advertiser have some market power and (ii) the triggering action frequency satisfies $U'(p) < 0$ for all p, the following hold:*

1. *The advertiser prices her products at a price p^A that is strictly higher than the price p^E she would charge if advertising was sold using pay-per-exposure.*

2. *The above price distortion leads to lower consumer surplus and a lower joint advertiser-publisher profit compared with a setting where advertising was sold using pay-per-exposure.*

The assumption $U'(p) < 0$ implies that *at least some* consumers have some (possibly imperfect) information about the price of the products being advertised before performing the action that triggers payment to the publisher.[3] This assumption obviously holds in the context of pay-per-sale contracts. It also holds in pay-per-click or pay-per-call settings where either (i) product prices are displayed on the ad, or (ii) a section of the consumer population has access to such prices before clicking or calling through separate information channels (e.g. product reviews, price comparison charts, previous clicks or calls to the same company, etc.) or (iii) a section of the population has sufficient industry knowledge (i.e. "knows" something about the advertiser's value function and the publisher's price structure) to be able to infer (possibly imperfectly) these prices.

4 Restoring efficiency in performance-based advertising

In Section 3.2 I discussed two classic solutions to the double marginalization problem: two part tariffs and quantity discounts. I begin this section by arguing that none of these solutions satisfactorily applies to performance-based advertising. The rest of the section then proposes two new ideas that can be used to eliminate price distortions in such settings.

The equivalent of a two part tariff in the performance-based advertising setting would be a contract where the publisher charges the advertiser a fixed fee f and a per-action fee equal to his marginal cost. The assumption of zero publisher marginal cost implies that there would only be a fixed fee. But this would be equivalent to a pure PPE contract. Two part tariffs, are thus, not useful in restoring efficiency if one wants to retain a performance-based compensation component. For zero marginal cost, quantity discount contracts translate to contracts where the per-action fee charged by the publisher would be equal to $\frac{f}{D}$, i.e. inversely proportional to the volume D of incremental sales due to advertising (or, more broadly, equal to $\frac{f}{U}$, i.e. inversely proportional to the triggering action

[3]If no consumer knows anything about the price of an advertised product before performing the action (click, call, sale, etc.) that triggers payment to the publisher then it will be $U'(p) = 0$ for all p. The advertiser's action of raising the price of her product then has no effect in terms of reducing the expected total payment to the publisher. In such a setting the advertiser has no incentive to deviate from the price p^* that maximizes the value she obtains from the advertising resource.

frequency U). The problem here is that sales (or, more broadly, triggering actions) only materialize *after* the advertiser obtains the resource and, sometimes, are spread out over a substantial time period. Effective quantity discount contracts are often impractical in such cases. The problem is even more severe in personalized ad settings where the triggering action occurs either once (in which case the advertiser pays the publisher) or not at all (in which case the advertiser pays nothing). In such cases quantity discount contracts are infeasible.

Although the standard approaches to eliminating inefficiencies due to double marginalization do not readily apply to performance-based advertising, it is possible to construct variants of these approaches that are better suited to this context. The rest of this section proposes two such ideas.

4.1 Price-based fee schedules

The main complication of applying the quantity discount idea in the context of performance-based advertising is that the incremental quantity sold because of advertising (or, more broadly, the triggering action frequency) is only known ex-post. However, if the publisher has sufficiently precise knowledge of the advertiser's *expected* triggering action frequency function $U(p)$ and if the advertiser can credibly disclose her product price p to the publisher, the publisher can construct a contract where the per-action fee charged to the advertiser can be set ex-ante equal to $\frac{f}{U(p)}$, i.e. inversely proportional to the expected triggering action frequency given the advertiser's price. It is then straightforward to prove the following result:

Proposition 2. *In PPA settings where the advertiser discloses her product price p and the publisher sets the per-action fee charged to the advertiser equal to $\frac{f}{U(p)}$, where f is a constant small enough to make the advertiser's participation individually rational and $U(p)$ is an accurate assessment of the advertiser's expected triggering action frequency (or probability) as a function of product price, the advertiser prices her products at the point $p^A = p^*$ that maximizes the joint advertiser-publisher profit.*

The above solution is effective in a wide range of settings. It works well in settings where the triggering action frequency is a stochastic function of product price and also in the important special case of personalized ad settings where $U(p)$ is a probability function. On the minus side it requires the advertiser to truthfully announce her product price and the publisher to be able to monitor the extent to which the advertiser adheres to her announcement.

4.2 History-based fee schedules

In settings where the advertiser and the publisher interact repeatedly, the publisher can use the "shadow of the future" to induce the advertiser to price her products at the joint profit-maximizing level. The applicable solutions differ depending on whether the triggering action frequency is a deterministic or stochastic function of product price.

Deterministic settings. In deterministic settings the publisher can induce efficient product pricing by committing to make next period's per-action fee a declining function of the current period's observed triggering action frequency: the more often the advertiser pays the publisher today, the lower the per-action fee she will have to pay tomorrow. This scheme gives the advertiser a disincentive to increase her current period product price. If $U'(p) < 0$, as product price increases, the current period's triggering action frequency declines. A lower triggering action frequency leads to a lower total advertising expenditure this period but (because of the increase in next period's per-action fee) to a higher advertising expenditure next period.

The following proposition shows that, in settings where the triggering action frequency is a deterministic function of product price, the publisher can always design a per-action fee schedule that balances these two forces and induces the advertiser to price her products at the point that maximizes the joint advertiser-publisher profit.

Proposition 3. *Consider an infinite horizon repeated PPA setting where (i) the triggering action frequency is a deterministic function $U(p)$ of product price and (ii) the per-action fee w_t $(t = 0, 1, ...)$ charged by the publisher to the advertiser at period t has the following expression:*

$$w_t = \begin{cases} w_0 & \text{if } t = 0 \\ w_0 \left[1 + \frac{1}{\delta}\left(1 - \frac{U_{t-1}}{U(p^*)}\right)\right] & \text{if } t > 0 \end{cases}$$

where $0 < w_0 < W(p^*)$, $p^* = \arg\max_p V(p)$, $U(p)$ is the triggering action frequency function as a function of product price p, U_{t-1} is the observed triggering action frequency at period $t-1$ and δ is the advertiser's discount factor.

In such a setting, for all $t = 0, 1, ...,$ the advertiser will price her products at the level $p_t^A = p^*$ that maximizes the joint advertiser-publisher profit.

Stochastic settings. If the relationship between product price and the observed triggering action frequency (or probability) is stochastic, repeated interaction can still induce efficient pricing. However, the solution now always involves the publisher charging advertisers some fixed (i.e. non-performance-based) fees in addition to performance-based fees, at least some of the time.

I will convey the intuition behind these ideas by restricting my attention to personalized ad settings, i.e. settings where the audience size is equal to one and the triggering action probability is $U(p)$. Such settings are practically important in online domains as well as analytically tractable. According to Proposition 2, the publisher can induce the advertiser to price her product at the point p^* that maximizes the joint advertiser-publisher profit by asking her to disclose her current period product price p^A and then setting the current period per-action fee equal to $\frac{f}{U(p^A)}$. If the advertiser's (per-period) reservation utility is u, the publisher can set the constant f as high as $V(p^*) - u$, resulting in discounted lifetime publisher profits $\frac{1}{1-\delta}[V(p^*) - u]$ and discounted lifetime advertiser profits $\frac{u}{1-\delta}$.

If the publisher cannot credibly learn the advertiser's current period product price he can still induce her to price at point p^* by setting up a schedule of fixed and per-action fees that depend on the advertiser's history of past triggering actions. For example, I will prove that the following two-state mechanism can achieve the desired objective:

Mechanism M1:

State H: All advertisers begin the game at state H. While they are in state H advertisers are charged a per-action fee w_H. If a triggering action occurrs during the current period the advertiser remains in state H, otherwise she transitions to state L.

State L: For every period that advertisers are in state L they are charged a fixed fee f and a per-action fee w_L. If a triggering action occurrs during the current period the advertiser transitions to state H otherwise she remains in state L.

The following result provides the details:

Proposition 4. *Consider an infinite horizon PPA setting that is characterized by joint advertiser-publisher profit $V(p)$, triggering action probability $U(p)$ and the above two-state mechanism M1. The advertiser's per-period reservation utility is u. If the publisher sets:*

$$
\begin{aligned}
f &= (V(p^*) - u)\,\tfrac{1}{\delta} \\
w_H = w_L &= (V(p^*) - u)
\end{aligned}
$$

the advertiser always prices her products at $p^A = p^*$ and the publisher extracts the maximum lifetime profits that are consistent with the advertiser's reservation utility.

The gist of the above idea is to give the advertiser an incentive to increase the probability of a triggering action (by lowering her price) by charging her an additional fixed fee on periods that immediately follow periods where the advertiser leased the resource but there was no triggering action. Observe that, compared to a pure PPA solution (e.g. the price-based fee schedule of Section 4.1), this solution involves a lower per-action fee ($V(p^*) - u$ versus $\frac{V(p^*)-u}{U(p^*)}$) coupled with an occasional fixed fee $(V(p^*) - u)\,\frac{1}{\delta}$. The discounted payoffs of both the advertiser and the publisher are identical to those obtained using a pure PPA (e.g. price-based fee) solution.

If advertisers are averse to the idea of paying a non-performance-based fee to the publisher, it is possible to design a variant of this approach where the expected present value of the non-performance-based fees is equal to zero. The idea is to extend Mechanism M1 so that, in addition to charging advertisers a fixed fee f when they are at state L, the publisher also offers them a fixed subsidy s when they are at state H such that the discounted lifetime sum of the fixed fees and subsidies is zero. The following result provides the details:

Proposition 5. *Consider an infinite horizon PPA setting that is characterized by joint advertiser-publisher profit $V(p)$, triggering action probability $U(p)$ and an extension of mechanism M1 that involves a subsidy s while at state H and a fee f while at state L. The advertiser's per-period reservation utility is u. If the publisher sets:*

$$
\begin{aligned}
s &= (V(p^*) - u)\,\tfrac{1-U(p^*)}{U(p^*)} \\
f &= (V(p^*) - u)\,\tfrac{1-\delta(1-U(p^*))}{\delta U(p^*)} \\
w_H = w_L &= \tfrac{V(p^*)-u}{U(p^*)}
\end{aligned}
$$

then (i) the advertiser always prices her products at $p^A = p^*$, (ii) the discounted lifetime sum of the non-performance-based fees and subsidies is equal to zero and (iii) the publisher extracts the maximum lifetime profits that are consistent with the advertiser's reservation utility.

Notice that in the above result the per-action fee is equal to the per-action fee $\frac{V(p^*)-u}{U(p^*)}$ in a pure PPA setting. This is intuitive since the system of non-performance-based fees and subsidies has zero net present value so the publisher obtains his net revenue from the per-action fees only.

All preceding solutions require the introduction of non-performance-based fees in addition to performance-based fees, at least part of the time. It is tempting to think that one could perhaps design a mechanism that induces optimal pricing using performance-based fees alone. The idea would be to charge a higher per-action fee following a period where there was no triggering action and a lower fee following a period where such action occurred. As was the case above, such mechanisms will give the advertiser additional incentives to increase the probability of a triggering action by keeping her price low. The following result shows that such schemes are indeed feasible. However all such mechanisms where the publisher charges positive per-action fees *must* involve a rule whereby the advertiser is expelled from the game after k consecutive periods where no triggering action occurs. For that reason, they can never be as efficient as mechanisms that charge both fixed and per-action fees.

Proposition 6: *All pure PPA mechanisms that induce advertisers to always price their products at the point $p^A = p^*$ that maximizes the per-period joint advertiser-publisher profit $V(p)$ have the following properties:*

1. *Either the publisher's per-action fees are zero for all periods or advertisers are expelled from the game after k consecutive periods where no triggering action occurs, for some integer $k \geq 1$.*

2. *In all such mechanisms the advertiser's discounted lifetime payoff (net of publisher fees) is equal to $\frac{1-\delta^k}{1-\delta} V(p^*)$.*

3. *For all finite k, the discounted lifetime joint publisher-advertiser profit is strictly lower than what can be achieved by an infinite horizon two-state fee schedule that involves both fixed and per-action fees (such as mechanism M1).*

The detailed expressions of the per-action fees imposed by the publisher to the advertiser and of the associated publisher profits as a function of k can be obtained for special cases of such mechanisms.[4] The intuition behind Part 3 of the above result is simple. The probability of k consecutive periods with no triggering action $(1 - U(p^*))^k$ is always positive (though, for large k, potentially very small). Therefore, with virtual certainty, all advertisers will be expelled from the game after a sufficiently large finite number of periods. For that reason, the social efficiency of this mechanism can never be as high as that of a game where the advertiser is never expelled. In practice, however, it can get arbitrarily close.

[4]Appendix A provides a full analysis of such a special case and shows that there exists an optimal k that maximizes the publisher's discounted lifetime profits.

A note about incomplete information settings. The analysis of this and the following sections assumes that the publisher has perfect knowledge of the advertiser's payoff function. In real-life applications this is seldom the case. However, if there is repeated interaction between the advertiser and the publisher, as it is common in real-life advertising settings, the literature on strategic learning (Young 2004; Young 2007) has shown that under a wide range of assumptions it is possible for the publisher to learn the advertiser's payoff function within an arbitrarily small margin of error. The results of this paper then hold at the limit where the publisher has fully learned the advertiser's payoff function.

5 Double marginalization in search advertising auctions

This Section shows that the price distortions that were identified in Section 3 occur in the practically important case of sponsored search keyword auctions. It also analyzes the economic implications of these distortions for the various stakeholders.

Consider, as before, a setting where a monopolist publisher owns an advertising resource and leases it on a per-period basis to a heterogeneous population of N advertisers. Advertisers are characterized by a privately known unidimensional type $q \in [\underline{q}, \overline{q}]$, independently drawn from a distribution with CDF $F(q)$. An advertiser's type relates to the attractiveness of her products or services to consumers; I assume that *ceteris paribus* higher types are, on average, more attractive. In the rest of the paper I will refer to q as the advertiser's quality, even though other interpretations are possible.[5] An advertiser's quality affects the *ex-ante* value $V(p,q)$ she expects to obtain from leasing the advertising resource for one period. As in Section 3.4 value $V(p,q)$ will be equal to the incremental sales profit that the advertiser expects to realize by leasing the resource.

In the rest of the paper I will use the notation $f_i(x_1, ..., x_n)$, $i = 1, ..., n$ to denote the partial derivative of $f(\cdot)$ with respect to its ith variable. For example, $V_2(p,q) = \frac{\partial V(p,q)}{\partial q}$. The following are assumed to hold for all $p \in \mathbb{R}^+$ and $q \in [\underline{q}, \overline{q}]$:

A1 $\quad V(p,q)$ is unimodal in p, attaining its unique maximum at some $p^*(q) > 0$

A2 $\quad \lim_{p \to \infty} V(p,q) = 0$

A3 $\quad V_2(p,q) > 0$

A1 and A2 are common and intuitive consequence of treating $V(p,q)$ as a sales profit function. A3 implies that *some* information about an advertiser's type becomes available to consumers at some point during the advertising-purchasing process, but still allows for a fairly general range of settings (for example: settings where this information might be noisy, where only a subset of consumers are informed, where firms might attempt to obfuscate their true types, etc.).

Each period the publisher allocates the resource to one of the competing firms using a Vickrey auction. Auction-based allocation of advertising resources is the norm in sponsored search advertis-

[5] For example, in settings with network effects (e.g. when the advertisers are social networks) q can be the size of the advertiser's user base.

ing and is also not uncommon in offline settings (e.g. superbowl ads). The double marginalization effects I discuss in this paper are orthogonal to whether the publisher offers one or several (identical or vertically-differentiated) resources. This allows us to ignore the multi-unit mechanism design complications present, say, in sponsored search position auctions (Athey and Ellison 2008; Edelman et al. 2006; Varian 2007) and focus on a single-unit auction.

Traditional *pay-per-exposure* (PPE) methods charge advertisers a fee that is levied upfront and is independent of the ex-post value that advertisers obtain by leasing the resource. In a Vickrey auction setting each advertiser bids the highest such fee she is willing to pay. The auctioneer allocates the resource to the top bidder and charges her the second-highest bid. Assuming that every other bidder of type y bids an amount equal to $\beta^E(y)$ and that, as I will later show, it is $\left[\beta^E\right]'(y) \geq 0$, at a symmetric Bayes-Nash equilibrium an advertiser of type q bids $b^E(q)$ and sets the price of her product at $p^E(q)$ to maximize her net expected profit:

$$\Pi^E(q; b^E(q), p^E(q), \beta^E(\cdot)) = \int_{\underline{q}}^{\left[\beta^E\right]^{-1}(b^E(q))} \left(V(p^E(q), q) - \beta^E(y)\right) G'(y)dy \tag{1}$$

where $G(y) = F^{N-1}(y)$ is the probability that every other bidder's type is less than or equal to y and $G'(y)$ is the corresponding density.[6] At equilibrium it must also be $\beta^E(q) = b^E(q)$. The above specification subsumes the special case where product prices $p(q)$ are given exogenously. In the latter case, a bidder of type q simply chooses a bid $b^E(q; p(\cdot))$ that maximizes $\Pi^E(q; b^E(q; p(\cdot)), p(q), \beta^E(\cdot))$ subject to $\beta^E(q) = b^E(q; p(\cdot))$.

I use the following shorthand notation:

$\Pi^E(q)$	advertisers' PPE equilibrium profit function (endogenous product prices)
$\Pi^E(q; p(\cdot))$	advertisers' PPE equilibrium profit function (exogenous product prices)

According to standard auction theory (e.g. Riley and Samuelson 1981) the expected publisher revenue associated with bids $\beta(y)$ is equal to:

$$R^E(\beta(\cdot)) = N \int_{\underline{q}}^{\overline{q}} \left(\int_{\underline{q}}^{z} \beta(y)G'(y)dy \right) F'(z)dz \tag{2}$$

I use the following shorthand notation:

$R^E = R^E(b^E(\cdot))$	publisher's PPE equilibrium revenue (endogenous product prices)
$R^E(p(\cdot)) = R^E(b^E(\cdot; p(\cdot)))$	publisher's PPE equilibrium revenue (exogenous product prices)

The following result characterizes the most important properties of allocating the resource using

[6] Throughout this paper I restrict my attention to symmetric Bayes-Nash equilibria. Unless specified otherwise, all subsequent references to "equilibrium" thus imply "symmetric Bayes-Nash equilibrium".

PPE bidding

Proposition 7: *In PPE settings where advertisers endogenously set the prices of their advertised products to maximize profits net of advertising expenditures:*

1. *Advertisers bid their expected ex-ante value of acquiring the resource, given their price:*

$$b^E(q) = V(p^E(q), q)$$

2. *Advertisers set the price of their products at the point that maximizes their ex-ante expected value of acquiring the resource:*

$$p^E(q) = \arg\max_p V(p, q) = p^*(q)$$

3. *Equilibrium PPE publisher revenues are equal to or higher to publisher revenues obtained in any PPE setting where the prices of advertised products are set exogenously:*

$$R^E = R^E(p^*(\cdot)) \geq R^E(p(\cdot))$$

with the inequality strict if and only if $p^*(q) \neq p(q)$ *for at least one* $q \in [\underline{q}, \overline{q}]$.

The principal takeaway of the above proposition is that auction-based PPE mechanisms induce all competing advertisers to price their products at the point that maximizes the joint advertiser-publisher profit. This is the point that maximizes the social efficiency, as well as the publisher revenues that are attainable through the use of the auction mechanism.

Pay-per-action (PPA) approaches make payment to the publisher contingent on a *triggering action* that is either a sale, or some other consumer action (e.g. click, call) that (i) is linked to the advertising, i.e. only consumers who have been exposed to this particular advertising message can perform the triggering action and, (ii) is a *necessary* step of a consumer's purchase decision process. As discussed in Section 3.4 the above assumptions allow us to uniquely express the advertiser's ex-ante value function as a product $V(p, q) = U(p, q)W(p, q)$ where $U(p, q)$ is the expected *triggering action frequency* (or, in the special case of personalized ads, the triggering action *probability*) and $W(p, q)$ is the expected *value-per-action*.

Early PPA sponsored search mechanisms used a simple *rank-by-bid* (RBB) allocation rule: Advertisers bid the maximum per-action fee they were willing to pay. The slot was allocated to the highest bidder who paid a per-action fee equal to the second highest bid. RBB mechanisms do not always allocate the resource to the advertiser that values it the most and, thus, do not optimize publisher revenues.[7] For that reason they were quickly abandoned in favor of, more sophisticated,

[7] For example, consider a setting where the per-action value $W(p, q)$ and the total value $V(p, q) = U(p, q)W(p, q)$ are negatively correlated for all q. Then, the advertiser who has the highest per-action value $W(p, q)$ will be able to place the highest bid and win the resource, even thought the total value she expects to generate by leasing the resource is the lowest among all advertisers.

rank-by-revenue mechanisms. The idea behind rank-by-revenue (RBR) is the following: The publisher computes a quality weight u_i for each advertiser. The quality weight is typically based on past performance data and attempts to approximate that advertiser's expected triggering action frequency $U(p,q)$. Once bidders submit their bids b_i the publisher computes a *score* $s_i = u_i b_i$ for each bidder. Assuming that (as we will show) advertisers place bids equal to their expected value-per-action $W(p,q)$, the score is an estimation of the expected total value $V(p,q) = U(p,q)W(p,q)$ that the advertiser will generate by acquiring the resource. The publisher allocates the resource to the bidder with the highest score and charges the winning bidder an amount $u_2 b_2/u_1$ equal to the second highest score divided by the winning bidder's quality weight. It has been shown that RBR methods have better allocative efficiency and auctioneer revenue properties than RBB (Feng et al. 2007; Lahaie and Pennock 2007; Liu and Chen 2006). Both Google and Yahoo use variants of this mechanism in their sponsored link auctions. For these reasons, the rest of the section will focus on RBR-PPA mechanisms.

Let $\Phi(q,s)$ denote an advertiser's beliefs about every other bidder's joint quality (q) and score (s) distribution. At equilibrium these beliefs must be consistent with bidding and publisher behavior. Let $\Phi(q) \equiv F(q)$ and $\Phi(s)$ be the corresponding marginal distributions and let $\Psi(s) = [\Phi(s)]^{N-1}$ denote the advertiser's belief that every other bidder's score will be less than s. Denote the advertiser's current quality weight as u. The single period specification of the advertiser's RBR-PPA bidding problem is to choose a bid $b^A(q,u)$ and a price $p^A(q,u)$ that maximize:

$$\begin{aligned} \Pi^A(q,u;b^A(q,u),p^A(q,u)) &= \int_0^{ub^A(q,u)} U(p^A(q,u),q)\left(W(p^A(q,u),q) - \tfrac{s}{u}\right)\Psi'(s)ds \\ &= V(p^A(q,u),q)\Psi(ub^A(q,u)) - U(p^A(q,u),q)\left(\int_0^{ub^A(q,u)} \tfrac{s}{u}\Psi'(s)ds\right) \end{aligned} \tag{3}$$

where $V(p^A(q,u),q)$ is the expected total value of the resource, $\Psi(ub^A(q,u))$ is the probability of winning the auction, $\int_0^{ub^A(q,u)} \frac{s}{u}\Psi'(s)ds$ is the expected per-action payment to the publisher and $U(p^A(q,u),q)$ is the expected number of times (or probability) that the payment will be made if the advertiser wins the auction.

The corresponding single period RBR-PPA publisher revenue is equal to:

$$R^A(b^A(\cdot,\cdot),p^A(\cdot,\cdot)) = N\int_{\underline{q}}^{\overline{q}} E_{u|q}\left[U(p^A(q,u),q)\left(\int_0^{ub^A(q,u)} \frac{s}{u}\Psi'(s)ds\right)\right]F'(q)dq \tag{4}$$

where $E_{u|q}[\cdot]$ denotes expectation with respect to u conditional on an advertiser's type being q.

The publisher's objective is to use u as an approximation of an advertiser's triggering action frequency. Of particular interest, therefore, is the behavior of the system at the limit where the publisher has obtained "correct" estimates of all quality weights, i.e. where each quality weight is equal to the respective advertiser's expected equilibrium triggering action frequency*:*

$$u_i = U(p^A(q_i,u_i),q_i)$$

I use the following shorthand notation to refer to equilibrium quantities in such "correct quality weight" equilibria:

$p^A(q) = p^A(q, U(p^A(q), q))$	equilibrium product prices
$b^A(q) = b^A(q, U(p^A(q), q))$	equilibrium bids
$\Pi^A(q) = \Pi^A(q, U(p^A(q), q); b^A(q), p^A(q))$	equilibrium advertiser's profits
$R^A = R^A(b^A(q), p^A(q))$	equilibrium publisher's revenue

In the rest of the paper I make the following technical assumption:

A4 $\quad \frac{\partial}{\partial q}\left[V(p^A(q), q)\right] \geq 0$ for all q

The above assumption states that, at any PPA equilibrium, and given the price that each advertiser type charges for her products, higher types derive higher value from obtaining the advertising resource. The following proposition summarizes equilibrium bidding behavior and revenues in the above setting:

Proposition 8: *If advertising is sold on an RBR-PPA basis, the following hold:*

1. *Advertisers set the price of their products at a point $p^A(q, u)$ that has the following properties:*

$$\begin{array}{ll} p^A(q,u) > p^*(q) & \text{if } U_1(p,q) < 0 \text{ for all } p \\ p^A(q,u) < p^*(q) & \text{if } U_1(p,q) > 0 \text{ for all } p \\ p^A(q,u) = p^*(q) & \text{if } U_1(p,q) = 0 \text{ for all } p \end{array}$$

2. *In settings that admit interior solutions advertisers bid their expected ex-ante value per action given their price:*

$$b^A(q, u) = W(p^A(q, u), q)$$

3. *In settings that admit interior solutions and where, additionally, the publisher sets every advertiser's quality weight to her respective equilibrium triggering action frequency, product prices $p^A(q)$ are solutions of the following equation:*

$$V_1(p^A(q), q)G(q) - U_1(p^A(q), q)Z^A(q) = 0 \tag{5}$$

where $Z^A(q) = \left(\int_{\underline{q}}^{q} V(p^A(y), y)G'(y)dy\right)/U(p^A(q), q)$ is the expected per-action payment to the publisher.

Proposition 8 shows that the price distortion due to the double marginalization effect that is induced by performance-based advertising persists in RBR-PPA settings. Note that Parts 1 and 2 of the Proposition hold true for any quality weights u, i.e. not only in settings where the publisher has "correct" assessments of each advertiser's expected triggering action frequency.

The rest of the section will show that, in this special context this price distortion has a series of negative implications: It always reduces consumer surplus, social welfare and publisher revenues.

Furthermore, in settings where the competing advertisers have homogeneous valuations of the advertising resource it induces a rat race situation where all advertisers raise their prices at the point where demand for their product approaches zero. Just as I did in Section 3.4, I will focus my attention on the practically important case where $U_1(p, q) < 0$ for all p, q.

Impact on consumers and social welfare

From Proposition 8, if $U_1(p, q) < 0$ it is $p^A(q) > p^*(q)$ for all q: product prices increase, reducing consumer surplus. Furthermore, the joint advertiser-publisher profit $V(p^A(q), q)$ is then always less than the optimum $V(p^*(q), q)$. If we define social welfare as the sum of consumer surplus plus the joint advertiser-publisher profit the following corollary immediately ensues:

Corollary 1: *Consumer surplus and social welfare are strictly lower in a RBR-PPA mechanism with endogenous product prices and perfectly estimated quality weights than in a corresponding PPE mechanism.*

Impact on publisher revenues

The key result here is that, if every advertiser's quality weight is a correct assessment of her equilibrium triggering action frequency, RBR-PPA is allocation and revenue equivalent to a PPE setting where product prices are exogenously set to $p^A(\cdot)$.

Proposition 9: *If advertising is sold on a RBR-PPA basis and the publisher sets every advertiser's quality weight to her respective expected equilibrium triggering action frequency then:*

1. *Advertiser profits are identical to her equilibrium profits in a PPE setting where prices are exogenously set to $p^A(\cdot)$:*
$$\Pi^A(q) = \Pi^E(q; p^A(\cdot))$$

2. *Publisher revenues are identical to his equilibrium revenues in a PPE setting where every advertiser exogenously prices her products at $p^A(q)$:*
$$R^A = R^E(p^A(\cdot))$$

Once we have this result in place it is easy to show that RBR-PPA results in lower publisher revenues than traditional PPE. By Proposition 8, in general it will be $p^A(\cdot) \neq p^*(\cdot)$. From Proposition 7 (Part 3) it will then be $R^E = R^E(p^*(\cdot)) > R^E(p^A(\cdot)) = R^A$. Therefore:

Corollary 2: *Equilibrium publisher revenues are strictly lower in a RBR-PPA mechanism with endogenous product prices and perfectly estimated quality weights than in a corresponding PPE mechanism.*

Rat race phenomena

I will now show that, for given $U(p, q)$ and $W(p, q)$, the magnitude of the price distortion induced by a RBR-PPA mechanism has a positive relationship with the ratio of a bidder's expected (per-action) payment relative to her value-per-action. When this ratio tends to one, the advertiser tends to raise the price of her products to the point where demand for them drops to zero and the resource becomes useless. This happens when the set of advertisers competing for the resource are homogeneous with respect to the value they expect to generate by leasing the advertising resource.

Proposition 10. *Let:*

$$\zeta^A(q) = \frac{Z^A(q)}{W(p^A(q), q)G(q)} = \frac{\int_{\underline{q}}^{q} V(p^A(y), y)G'(y)dy}{V(p^A(q), q)G(q)} \tag{6}$$

denote the expected (per-action) payment-to-valuation ratio of an advertiser of type q conditional on that advertiser winning the publisher's auction. Fixing $U(p, q)$ and $W(p, q)$, the following statements summarize how the magnitude of $\zeta^A(q)$ impacts equilibrium RBR-PPA prices and the equilibrium value of the advertising resource to the winning advertiser:

1. $\frac{\partial p^A(q;\zeta^A(q))}{\partial \zeta^A(q)} \geq 0$
2. *If $W_1(p, q) > 0$ for all p then it is $\lim_{\zeta^A(q)\to 1} V(p^A(q; \zeta^A(q)), q) = 0$*

The intuition behind this result is the following: The higher the per-action payment to the publisher, the higher the advertisers' marginal gain from increasing $p^A(q)$ and thus reducing the triggering action frequency (i.e. the frequency of paying the publisher). At the limit where the expected per-action payment $Z^A(q)$ approaches a bidder's value-per-action $W(p^A(q), q)$ an advertiser's losses from the reduction in demand that results from price increases are almost exactly compensated by the corresponding reduction in the payment to the publisher. At the same time, if $W_1(p, q) > 0$, higher product prices result in a higher value-per-action, which allows the advertiser to place a higher bid. Competition among bidders for the advertising resource then pushes product prices upwards to the point where the triggering action frequency (and thus the value of the resource to the advertiser) goes to zero. This is a rat-race situation that, clearly, has negative consequences for all parties involved.

Integrating (6) by parts gives:

$$\zeta^A(q) = 1 - \frac{\int_{\underline{q}}^{q} \frac{\partial}{\partial y}\left[V(p^A(y), y)\right] G(y)dy}{V(p^A(q), q)G(q)} \tag{7}$$

From (7) it follows that $\zeta^A(q)$ is inversely related to the variability of the bidder population's *equilibrium* value $V(p^A(y), y)$ as a function of the bidders' type. The more homogeneous the value across bidders, the smaller the distance between the valuations of any two consecutive bidders and thus the higher the expected payment relative to the winning bidder's value. At the limit where $\frac{\partial}{\partial y}\left[V(p^A(y), y)\right] \to 0$, it is $\zeta^A(q) \to 1$. Intuitively, if the bidder population is homogeneous with respect to the value it expects to generate from advertising, the bidding competition for obtaining the resource becomes more intense and drives product prices up to the point where demand for them drops to zero.

6 Restoring efficiency in search advertising auctions

Section 4 discusses several ways of restoring efficient pricing in a posted price performance-based advertising setting. This section adapts some of these ideas to the context of search advertising auctions. The key difference in this context is that the publisher does not have direct control of per-action fees since these are decided by the auction mechanism.

6.1 Price-based quality weights

In Section 4.1 I discussed how the idea of using quantity discounts to induce efficient pricing in settings that are prone to double marginalization can be adapted to the context of performance-based advertising: the publisher asks the advertiser to disclose her current round product price p and charges a per-action fee that is inversely proportional to the advertiser's expected triggering action frequency $U(p)$, given her price. This idea is not directly applicable in auction-based mechanisms where the publisher cannot directly set per-action fees. However, in RBR schemes the publisher has indirect control over both the probability that an advertiser will obtain the resource as well as over the price that she is going to pay through her quality weight u. It is easy to show that:

Lemma 1: *For all q, u it is* $\frac{\partial \Pi^A(q,u)}{\partial u} \geq 0$

Recall that an advertiser's quality weight is meant to be an estimate of her expected triggering action frequency. RBR mechanisms, therefore, offer a "bonus" to advertisers that can increase their triggering action frequency. In that sense, they have a form of "quantity discount" already built into them. The problem, however, is that most current implementations of RBR-PPA treat each advertiser's quality weight as a point estimate that is based on the advertiser's history of observed triggering actions in *past* rounds only. Furthermore, once the estimation process converges the quality weight becomes fixed and independent of the advertiser's current period product price. Proposition 8 shows that price distortions persist at that limit.

If the publisher has an accurate estimate of each advertiser's entire triggering action frequency function $U(p, q_i)$ and if reliable elicitation of each round's product prices is possible, the following result shows that the publisher can induce efficient product pricing by making each advertiser's current round quality weight a function of both her past history and current period product price.

Proposition 11: *Consider a RBR-PPA mechanism where:*

(i) *Each round advertisers $i = 1, ..., N$ are asked to disclose their bid b_i and current round product price p_i*

(ii) *Each advertiser's current round quality weight is set to $u_i = \tilde{U}_i(p_i)$*

At the limit where the publisher has a perfect estimate of each advertiser's triggering action frequency function, i.e. where $\tilde{U}_i(p) = U(p, q_i)$ for all i and p, the above mechanism:

1. *induces all advertiser types to always price their products at $p^*(q)$ and to bid their expected value-per-action $W(p^*(q), q)$*

2. *results in identical allocation, advertiser profits and publisher revenue to those of a PPE mechanism with endogenous prices*

The above solution is effective in a wide range of settings. It works well in settings where the triggering action frequency is a stochastic function of product price and also in the important special case of personalized ad settings where $U(p,q)$ is a probability function. Its main practical drawback is that it requires the advertiser to truthfully announce a product price and the publisher to be able to monitor the extent to which the advertiser adheres to her announcement. It also complicates the publisher's learning task, as it requires him to have full knowledge of the triggering action frequency function for all prices and all advertiser types.

6.2 History-based fee schedules

Section 4.2 shows that, in settings with repeated interaction, the publisher can induce the advertiser to price her products at the efficient level by committing to "punish" her in the next round of interaction if her current round triggering action frequency was not sufficiently high (or, in the special case of personalized ads, if no triggering action was observed in the current round). This punishment consists either of temporarily increasing the per-action fee charged to the advertiser or of temporarily charging a fixed fee on top of the per-action fee. The advantage of such mechanisms is that they do not require the advertiser to disclose product prices to the publisher.

In auction-based settings the publisher cannot directly control the per-action fee charged to advertisers. Therefore, the most straightforward type of intervention would be to add a system of fixed fees and subsidies on top of the per-action fee that is decided by auction. In the important case of personalized ad settings I will prove that, for properly calibrated fees and subsidies, the following two-state mechanism can induce efficient bidding and pricing:

Mechanism M2:

State H: All advertisers begin the game at state H. While they are in state H advertisers pay an auction participation fee $f_H(q)$. If they win the auction they are charged a fixed fee $g_H(q)$ plus the per-action fee that is decided by auction. If a triggering action occurs during the current period the advertiser remains in state H, otherwise she transitions to state L.

State L: For every period that advertisers are in state L they pay an auction participation fee $f_L(q)$. If they win the auction they are charged a fixed fee $g_L(q)$ plus the per-action fee that is decided by auction. If a triggering action occurs during the current period the advertiser transitions to state H otherwise she remains in state L.

The following proposition (analogous to Proposition 5) provides the details.

Proposition 12. *Consider an infinite horizon RBR-PPA setting that is characterized by joint advertiser-publisher profit $V(p,q)$, triggering action probability $U(p,q)$ and the above two-state mech-*

anism M2 involving auction participation fees $f_H(q), f_L(q)$ *and fixed fees* $g_H(q), g_L(q)$*. If the publisher has a correct assessment of every advertiser's triggering action probability function* $U(p,q)$ *and sets the fees as follows:*

$$
\begin{aligned}
f_H(q) &= 0\\
f_L(q) &= \tfrac{1-\delta}{\delta G(q)} Z(q)\\
g_H(q) &= -\tfrac{1-U(p^*(q),q)}{G(q)} Z(q)\\
g_L(q) &= \tfrac{U(p^*(q),q)}{G(q)} Z(q)
\end{aligned}
$$

where $Z(q) = \int_{\underline{q}}^{q} V(p^*(y), y) G'(y) dy / U(p^*(q), q)$ *is the expected per-action payment to the publisher, then the above mechanism:*

1. *induces all advertiser types to always price their products at* $p^*(q)$ *and to bid their expected value-per-action* $W(p^*(q), q)$

2. *results in identical allocation and expected lifetime discounted advertiser profits and publisher revenue to those of an infinite horizon PPE mechanism with endogenous prices*

Observe that the solution involves a positive auction participation fee in state L only. Furthermore, notice that the fixed fee $g_H(q)$ charged to the winner at state H is negative, i.e. a subsidy.

The following is the intuition behind the above result and an explanation of why both auction participation fees (paid by everyone) and fixed fees (paid by the auction winner only) are needed. As was the case in Section 4.2 the threat of lower payoffs (due to the participation fee $f_L(q)$) in state L provides a disincentive to advertisers to raise the price of their products too much: higher prices reduce the probability that a triggering action will be observed in the current round and thus increase the probability of transitioning to (or remaining in) state L. At the same time, the positive probability of transitioning to lower payoff state L if one wins the auction and doesn't generate a triggering action while at state H, *reduces* a state-H advertiser's expected valuation of winning the auction (and, thus, her expected bid). Subsidy $g_H(q)$ is then necessary to correct for this distortion and restore the advertiser's expected valuation of winning the auction to her value-per-action. Similarly, while in state L, the only way to transition to the higher payoff state H is by winning the auction and generating a triggering action. This *increases* a state-L advertiser's expected valuation of winning the auction. Fixed fee $g_L(q)$ is necessary to correct this distortion.

The above schedule of auction participation and fixed fees is calibrated so that their expected lifetime discounted sum is zero. Therefore, they do not change either the advertisers' or the publisher's lifetime discounted payoffs relative to a mechanism (e.g. PPE) where advertisers bid and price in identical ways but where there are no extra fees or subsidies involved.

Just as I did in Section 4.2 it is instructive to ponder whether one can induce efficient pricing by a "pure PPA" mechanism, i.e. without resorting to auction participation or fixed fees. Lemma 1 suggests that it might be possible to construct effective punishment mechanisms by exploiting an advertiser's quality weight: the publisher can commit to increase an advertiser's quality weight when triggering actions occur and to decrease it when they don't. Although the threat to decrease an advertiser's quality weight might succeed in reducing, or eliminating, price distortions, it suffers from another important drawback: to the extent that punishment periods occur with positive probability it would also reduce the auction's allocative efficiency. Recall that RBR mechanisms allocate the slot to the bidder who achieves the maximum score $s_i = u_i b_i$, where (by Proposition 8) $b_i = W(p^A(q_i), q_i)$. Score $s_i = u_i W(p^A(q_i), q_i)$ is perfectly correlated with the advertiser's value $V(p^A(q_i), q_i) = U(p^A(q_i), q_i)W(p^A(q_i), q_i)$ and leads to allocative efficiency if and only if $u_i = U(p^A(q_i), q_i)$ for all i. If there are positive probability punishment states where $u_i \neq U(p^A(q_i), q_i)$ for some i then during such states the resource may not be allocated to the advertiser who values it the most. This, in turn, reduces the mechanism's expected lifetime joint advertiser-publisher profit below the optimal. Positive probability punishment states occur in any setting where the triggering action frequency is a stochastic function of product price or where function $U(p, q)$ denotes a probability. The following proposition ensues:

Proposition 13: *In any setting where (i) there is a stochastic relationship between the prices of advertised products and the observed triggering action frequency/probability, (ii) the advertiser's quality weight is a function of her triggering action history and (iii) the advertiser does not disclose her product price to the publisher, the discounted lifetime joint publisher-advertiser profit that can be achieved by a pure PPA mechanism is strictly lower than what can be achieved by Mechanism M2.*

In conclusion, restoring efficient product pricing in search advertising auctions requires either the ability to truthfully elicit the advertiser's current round product price or payment mechanisms where the publisher charges auction participation fees and fixed fees/subsidies on top of the performance-based fees.

7 Concluding Remarks

Technological advances have made it increasingly feasible to track the impact of individual advertising messages on consumer behavior. Accordingly, pay-per-performance advertising mechanisms, whereby the publisher is only paid when consumers perform certain actions (e.g. clicks, calls, purchases) that are tied to a specific advertising stimulus, have been gaining ground. Such pay-per-action (PPA) mechanisms are proving popular with advertisers because they help limit their risk when investing in new and often untested advertising technologies as well as allow them to better estimate their advertising ROI.

This paper highlights an important, and previously unnoticed, side-effect of PPA advertising. I show that, in settings where at least a subset of consumers has some information about product price before performing the action (click, call, purchase, etc.) that triggers payment to the publisher, PPA mechanisms induce advertisers to distort the prices of their products - usually upwards - as it is more beneficial to them to pay the publisher fewer times but realize a higher net profit per sale. Such equilibria always reduce social welfare and reduce the payoffs of most stakeholders involved: consumers are always left with a lower surplus (because they pay higher prices) and one or both of advertiser profits and publisher revenues decline.

Price distortions persist in the rank-by-revenue auction-based variants of PPA advertising currently practiced by Google and Yahoo. Interestingly, in the latter settings they *always* reduce publisher revenues relative to more traditional pay-per-exposure methods.

Fortunately, it is possible to introduce enhancements to these mechanisms that restore efficient pricing. I propose two broad sets of ideas. The first idea requires the advertiser to disclose to the publisher both her bid as well as the product price she intends to charge. The publisher then makes the advertiser's current round quality weight a function of both her past triggering action history and her current round product price. Although conceptually simple, this idea relies on advertisers truthfully revealing their product prices to the publisher. In practice, this is not a trivial requirement and might require additional infrastructure that can monitor the accuracy of the advertisers' price announcements.

The second idea is to construct mechanisms whereby advertisers are charged some form of non-performance-based penalty, e.g. an auction participation fee, following periods where no triggering actions (clicks, sales, etc.) occur, coupled with subsidies following periods where triggering actions occur. The attractive feature of this solution is that penalties and subsidies can be designed so that their equilibrium lifetime discounted sum nets to zero. Therefore, if advertisers are billed periodically (e.g. monthly or quarterly) by the publisher, the non-performance-based component of their bill would be close to zero (and might even be a small credit), a desirable feature from the perspective of most advertisers since it keeps their expenditures primarily based on actual performance.

To keep my models tractable but also to better highlight the phenomena that form the focus of the paper I made a number of simplifying assumptions. I am arguing that these assumptions do not detract from the essence of the phenomenon.

First, I assumed that both the publisher and the advertisers are risk neutral. This assumption was important to retain tractability. Real-life advertisers tend to be risk-averse; in fact, risk aversion is an important motivator for the introduction of pay-per-performance payment methods (Mahdian and Tomak 2008). However, although risk aversion will introduce some additional forces in favor of performance-based advertising, the fundamental drivers of price distortion, that are due to the publisher and advertisers each independently attempting to maximize their payoffs, will persist.

Second, I have assumed that the advertising resource is the only channel through which advertisers can connect to consumers or alternatively, that if the advertiser also sells her products through additional channels, she sets her prices separately for each channel. If this assumption is not true, for example if the advertiser sells through multiple channels and charges the same prices across all channels, price distortions will still occur, however their magnitude will be smaller and proportional to the relative share of revenues that can be attributed to the advertising resource.

Third, I assumed that the publisher is a monopolist. As argued in Section 3 this is not an essential assumption: as long as the publisher has some market power (i.e. if he can charge the advertisers a per-action fee above marginal cost) the price distortions discussed in this paper will occur.

Internet technologies have spurred a tremendous wave of innovation in advertising methods. New ideas are being continuously invented and tried out by ambitious entrepreneurs, often without being rigorously analyzed. Given the fast pace of competition and innovation in the Internet arena it is only natural that some of these ideas might have shortcomings or side-effects that are not immediately obvious to their inventors. One role for academia in this fast-changing environment is to place these new ideas under a rigorous theoretical lens, identify their shortcomings and propose economically sound improvements. This work is very much in this spirit.

References

Abhishek, V. and Hosanagar, K. (2007) Keyword generation for search engine advertising using semantic similarity between terms. *Proceedings of the 9th International Conference on Electronic Commerce*, Minneapolis, MN, USA, 89-94.

Agarwal, N., Athey, S. and Yang, D. (2009) Skewed Bidding in Pay-per-Action Auctions for Online Advertising. *American Economic Review*, 99(2), 441–47.

Aggarwal, G., Goel, A. and Motwani, R. (2006) Truthful auctions for pricing search keywords. *Proceedings of the 7th ACM Conference on Electronic Commerce*, Ann Arbor, MI, USA, 1-7.

Animesh, A., Ramachandran, V. and Viswananthan, S. (2009) Quality Uncertainty and the Performance of Online Sponsored Search Markets: An Empirical Investigation. *Information Systems Research*, forthcoming.

Ashlagi, I., Mondered, D. and Tennenholtz, M. (2008) Competing Ad Auctions. *Proceedings of the 4th Workshop on Ad Auctions*, Chicago, IL, USA (July 8-9, 2008).

Athey, S., and Ellison, G. (2008) Position Auctions with Consumer Search. Working paper.

Borgs, C., Chayes, J., Immorlica, N., Jain, K., Etesami, O. and Mahdian, M. (2007) Dynamics of Bid Optimization in Online Advertisement Auction. *Proceedings of the 16th International Conference on World Wide Web*, Banff, Alberta, Canada, 531-540.

Cary, M., Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A. R., Mathieu, C. and Schwarz, M. (2007) Greedy bidding strategies for keyword auctions. *Proceedings of the 8th ACM Conference on Electronic Commerce*, San Diego, CA, USA, 262-271.

Chen, Y. and He, C. (2006) Paid placement: Advertising and Search on the Internet. NET Institute Working Paper No. 06-02. http://ssrn.com/abstract=936472.

Cournot, A. (1838). *Recherches sur les Principes Mathématiques de la Théorie des Richesses.* English edition: Research into the Mathematical Principles of the Theory of Wealth, Edited by N. Bacon, New York: MacMillan, 1897.

Edelman, B., Ostrovsky, M., and Schwarz, M. (2007) Internet Advertising and the Generalized Second-price Auction: Selling Billions of Dollars Worth of Keywords. *American Economic Review* 97, 242-259.

Feldman, J., Muthukrishnan, S., Pal, M. and Stein, C. (2007) Budget optimization in search-based advertising auctions. *Proceedings of the 8th ACM Conference on Electronic Commerce*, San Diego, CA, USA, 40-49.

Feldman. J. and Muthukrishnan, S. (2008) Algorithmic Methods for Sponsored Search Advertising. Chapter 4 in Liu, Z. and Xia, C. H., eds. *Performance Modeling and Engineering*, Springer, New York, NY.

Feng, J., Bhargava, H. K. and Pennock. D. M. (2007) Implementing sponsored search in Web search engines: Computational evaluation of alternative mechanisms. *INFORMS Journal on Computing* 19 (1), 137-148.

Feng, J. and Xie, J. (2007) Performance-based Advertising: Price and Advertising and Signals of Product Quality. http://ssrn.com/abstract=1027296.

Fluet, C. and Garella, P.G. (2002) Advertising and Prices as Signals of Quality in a Regime of Price Rivalry. *International Journal of Industrial Organization* 20, 907-930.

Ghose, A. and Yang, S. (2009) An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets. http://ssrn.com/abstract=1022467

Goldfarb, A. and Tucker, C. (2007) Search Engine Advertising: Pricing Ads to Context. Available on SSRN http://ssrn.com/abstract=1021451

Hertzendorf, M. N. and Overgaard, P. B. (2001) Price Competition and Advertising Signals: Signaling by Competing Senders. *Journal of Economics and Management Strategy* 10 (4), 621-662.

Holmstrom, B. (1979) Moral hazard and observability. *Bell Journal of Economics* 10 (1) 74-91.

Holmstrom, B. and Milgrom, P. (1987) Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55 (2) 303-328.

Holmstrom, B. and Milgrom, P. (1991) Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics and Organization* 7 (Sp), 24-52.

Hu, J. (2004) Performance-based Pricing Models in Online Advertising. Available on SSRN http://ssrn.com/abstract=501082.

Immorlica, N., Jain, K., Mahdian, M. and Talwar, K. (2005) Click Fraud Resistant Methods for Learning Click-Through Rates. *Lecture Notes in Computer Science* 3828, 34-45.

Jeuland, A. P. and Shugan, S. M. (1983) Managing Channel Profits. *Marketing Science* 2 (3), 239-272.

Joshi, A. and Motwani, R. (2006) Keyword Generation for Search Engine Advertising. *Proceedings of the 6th IEEE International Conference on Data Mining Workshops* (ICDMW), Hong Kong, PRC, 490-496.

Katona, Z. and Sarvary, M. (2008) The race for sponsored links: A model of competition for search advertising. Working Paper.

Lahaie, S. and Pennock, D. M. (2007) Revenue analysis of a family of ranking rules for keyword auctions. *Proceedings of the 8th ACM Conference on Electronic Commerce*, San Diego, CA, 50-56.

Lahaie, S., Pennock, D. M., Saberi, A. and Vohra, R. V. (2007) Sponsored Search Auctions. Chapter 28 in Nisan, N., Roughgarden, T., Tardos, E. and Vazirani, V. eds., *Algorithmic Game Theory*, Cambridge University Press, New York, NY.

Liu, D. and Chen, J. (2006) Designing Online Auctions with Past Performance Information. *Decision Support Systems* 42 (3), 1307-1320.

Liu, D., Chen, J. and Whinston A. B. (2008a) Current Issues in Keyword Auctions. Available on SSRN http://ssrn.com/abstract=1008496.

Liu, D., Chen, J. and Whinston, A. B. (2008b) Competing Keyword Auctions. *Proceedings of the 4th Workshop on Ad Auctions*, Chicago, IL, USA (July 8-9, 2008).

Mahdian, M. and Tomak, K. (2008) Pay-per-action Model for Online Advertising. *International Journal of Electronic Commerce* 13 (2), 113-128.

Milgrom, P. and Roberts, J. (1986) Price and Advertising Signals of Product Quality. *Journal of Political Economy* 94 (4), 796-821.

Milgrom, P. and Shannon, C. (1994) Monotone Comparative Statics. *Econometrica* 62(1), 157-80.

Nazerzadeh, H., Saberi, A, and Vohra, R. (2008) Dynamic Cost-Per-Action Mechanisms and Applications to Online Advertising. *Proceedings of the 17th international conference on World Wide Web* (WWW'08), Beijing, China, 179-188.

Rey, P. and Vergé, T. (2008) The Economics of Vertical Restraints. Chapter 9 in Buccirossi, P. ed. *Handbook of Antitrust Economics,* MIT Press, Cambridge MA.

Riley, J. G., Samuelson, W. F. (1981) Optimal Auctions. *American Economic Review* 71 (3) 381-392.

Russmevichientong, P. and Williamson, D. (2006) An adaptive algorithm for selecting profitable keywords for search-based advertising services. *Proceedings of the 7th ACM Conference in Electronic Commerce*, Ann Arbor, MI, USA, 260-269.

Rutz, O. J. and Bucklin, R. E. (2007) A Model of Individual Keyword Performance in Paid Search Advertising. http://ssrn.com/abstract=1024765.

Rutz, O. J. and Bucklin, R. E. (2008) From Generic to Branded: A Model of Spillover Dynamics in Paid Search Advertising. http://ssrn.com/abstract=1024766.

Spengler, J. (1950) Vertical Integration and Antitrust Policy. *Journal of Political Economy* 58, 347-352.

Sundararajan, A. (2004) Nonlinear Pricing of Information Goods. *Management Science* 50 (12), 1660-1673.

Topkis, D. M. (1998) *Supermodularity and Complementarity.* Princeton University Press, Princeton, NJ.

Varian, H. (2007) Position Auctions. *International Journal of Industrial Organization* 25, 1163–1178.

Wilbur, K. and Zhu, Y. (2009) Click Fraud. *Marketing Science* 28 (2), 293-308.

Yao, S. and Mela. C. F. (2009a) Sponsored Search Auctions: Research Opportunities in Marketing (March 12, 2009). Available on SSRN: http://ssrn.com/abstract=1367402.

Yao, S. and Mela, C. F. (2009b) A Dynamic Model of Sponsored Search Advertising(March 19, 2009). Available at SSRN: http://ssrn.com/abstract=1285775.

Young, H. P. (2004) *Strategic Learning and Its Limits.* Oxford University Press, Oxford, UK.

Young, H. P. (2007) The possible and the impossible in multi-agent learning. *Artificial Intelligence* 171, 429-433.

Zhao, H. (2000) Raising Awareness and Signaling Quality to Uninformed Consumers: A Price-Advertising Model. *Marketing Science* 19 (4), 390-396.

Zhao, H. (2005) Incentive-based compensation to advertising agencies: A principal-agent approach. *International Journal of Research in Marketing* 22 (3), 255-275.

Proofs

Proof of Lemma 1

It is:

$$\Pi^A(q,u;b^A(q,u),p^A(q,u)) = V(p^A(q,u),q)\Psi(ub^A(q,u)) - U(p^A(q,u),q)Z^A(b^A(q,u),u)$$

where $Z^A(b,u) = \int_0^{ub} \frac{s}{u}\Psi'(s)ds$. Differentiating with respect to u and substituting (see Proposition 8) $b^A(q,u) = W(p^A(q,u),q) = V(p^A(q,u),q)/U(p^A(q,u),q)$ we obtain:

$$\frac{\partial \Pi^A(q,u)}{\partial u} = \frac{Z^A(W(p^A(q,u),q),u)}{u} > 0$$

Proof of Proposition 1

The proof is sketched in the main body of the text.

Proof of Proposition 2

The advertiser's profit function is equal to $\pi_r(p|w) = U(p)(W(p)-w) = V(p) - U(p)w$. If $w = \frac{f}{U(p)}$ the profit function becomes $\pi_r(p|w) = V(p) - f$. The price p^A that maximizes $\pi_r(p|w)$ is then equal to the price p^* that maximizes $V(p)$. The constant f must be less than $V(p^*) - u$ where u is the advertiser's reservation utility.

Proof of Proposition 3

If each period's observed triggering action frequency is a deterministic function of product price it is $U_t = U(p_t)$. The Bellman equation of the infinite horizon PPA setting of interest to this context then is:

$$\Omega(w_t) = \max_p \left[V(p) - U(p)w_t + \delta\Omega(h(U(p)))\right]$$

where

$$h(U) = w_0\left[1 + \frac{1}{\delta}\left(1 - \frac{U}{U(p^*)}\right)\right]$$

and p^* is the price that maximizes $V(p)$, i.e. satisfies $V'(p^*) = 0$ and $V''(p^*) < 0$.

Notice that in the above system product prices $p_t = p(w_t)$ only depend on the current period's per-action fee w_t. Specifically, they must satisfy the first order condition:

$$V'(p(w_t)) - U'(p(w_t))\left[w_t - \delta\Omega'(h(U(p(w_t))))h'(U(p(w_t)))\right] = 0 \tag{8}$$

and the corresponding second order condition. Substituting:

$$\Omega'(w_t) = -U(p(w_t))$$

$$h'(U) = -\frac{w_0}{\delta U(p^*)}$$

into (8) we obtain:

$$V'(p(w_t)) - U'(p(w_t)) \left[w_t - w_0 \frac{U(p(h(U(p(w_t)))))}{U(p^*)} \right] = 0 \tag{9}$$

I will show that, for $w_t = w_0$, $p(w_0) = p^*$ solves (9). Observe that it is $h(U(p^*)) = w_0$. Substituting $w_t = w_0$ and $p(w_t) = p(w_0) = p^*$ into (9) we obtain:

$$V'(p^*) - U'(p^*) \left[w_0 - w_0 \frac{U(p^*)}{U(p^*)} \right] = V'(p^*) = 0$$

For $w_t = w_0$ and $p(w_0) = p^*$ the second-order condition that corresponds to (9) is:

$$\begin{aligned} &V''(p^*) - U''(p^*) \left[w_0 - \delta\Omega'(h(U(p^*)))h'(U(p^*)) \right] \\ &+\delta \left[\Omega''(h(U(p^*)) \left[h'(U(p^*))U'(p^*) \right]^2 + \Omega'(h(U(p^*)))h''(U(p^*)) \left[U'(p^*) \right]^2 \right] < 0 \end{aligned}$$

Because it is $\Omega''(w) = 0$, $h''(U) = 0$ and $w_0 - \delta\Omega'(h(U(p^*)))h'(U(p^*)) = 0$ the above simplifies to $V''(p^*) < 0$ which is true by definition.

In summary, I have shown that if the current period per-action fee is w_0 and the next period fee is equal to $h(U)$, the current period product price that maximizes the advertiser's value function is equal to p^*. Furthermore (because it is $h(U(p^*)) = w_0$) if $w_t = w_0$ then it is also $w_{t+1} = w_0$. Therefore, at equilibrium the publisher will charge the advertiser a per-period fee equal to w_0 at all periods and the advertiser will find it optimal to charge a product price $p(w_0) = p^*$ at all periods.

Proof of Proposition 4

The Bellman equations that describe the system of interest in this proposition are:

$$\begin{aligned} \Omega_H &= V(p_H) - U(p_H)w_H + \delta \left[U(p_H)\Omega_H + (1 - U(p_H))\Omega_L \right] \\ \Omega_L &= -f + V(p_L) - U(p_L)w_L + \delta \left[U(p_L)\Omega_H + (1 - U(p_L))\Omega_L \right] \end{aligned} \tag{10}$$

where p_H, p_L satisfy the first-order conditions:

$$\begin{aligned} V'(p_H) + U'(p_H) \left[-w_H + \delta \left(\Omega_H - \Omega_L \right) \right] = 0 \\ V'(p_L) + U'(p_L) \left[-w_L + \delta \left(\Omega_H - \Omega_L \right) \right] = 0 \end{aligned}$$

The requirement $p_H = p_L = p^*$, such that $V(p^*) = 0$ implies:

$$w_H = w_L = \delta\left(\Omega_H - \Omega_L\right) \tag{11}$$

Substituting (11) into (10) and solving for f, w_H, w_L under the constraint $\Omega_H = \frac{u}{1-\delta}$ that characterizes the situation where the publisher extracts all the surplus subject to leaving the advertiser with an average per-round reservation utility u, gives:

$$\begin{gathered} f = \left(V(p^*) - u\right)\tfrac{1}{\delta} \\ w_H = w_L = \left(V(p^*) - u\right) \end{gathered}$$

Proof of Proposition 5

The Bellman equations that describe the system of interest in this proposition are:

$$\begin{aligned} \Omega_H &= s + V(p_H) - U(p_H)w_H + \delta\left[U(p_H)\Omega_H + (1 - U(p_H))\Omega_L\right] \\ \Omega_L &= -f + V(p_L) - U(p_L)w_L + \delta\left[U(p_L)\Omega_H + (1 - U(p_L))\Omega_L\right] \end{aligned} \tag{12}$$

where p_H, p_L satisfy the first-order conditions:

$$\begin{gathered} V'(p_H) + U'(p_H)\left[-w_H + \delta\left(\Omega_H - \Omega_L\right)\right] = 0 \\ V'(p_L) + U'(p_L)\left[-w_L + \delta\left(\Omega_H - \Omega_L\right)\right] = 0 \end{gathered}$$

The requirement $p_H = p_L = p^*$, such that $V(p^*) = 0$ implies:

$$w_H = w_L = \delta\left(\Omega_H - \Omega_L\right) \tag{13}$$

Let N_H denote the net present value of the system of non-performance-based fees f and subsidies s. It is:

$$\begin{gathered} N_H = s + \delta\left[U(p^*)N_H + (1 - U(p^*))N_L\right] \\ N_L = -f + \delta\left[U(p^*)N_H + (1 - U(p^*))N_L\right] \end{gathered} \tag{14}$$

Solving the above system and setting $N_H = 0$ gives:

$$s = f\frac{\delta(1 - U(p^*))}{1 - \delta(1 - U(p^*))} \tag{15}$$

Substituting (13) and (15) into (12) and solving for f, w_H, w_L under the constraint $\Omega_H = \frac{u}{1-\delta}$ that characterizes the situation where the publisher extracts all the surplus subject to leaving the advertiser with an average per-round reservation utility u, gives:

$$\begin{gathered} s = (V(p^*) - u)\,\frac{1-U(p^*)}{U(p^*)} \\ f = (V(p^*) - u)\,\frac{1-\delta(1-U(p^*))}{\delta U(p^*)} \\ w_H = w_L = \frac{V(p^*)-u}{U(p^*)} \end{gathered}$$

Proof of Proposition 6

Consider a setting where each period the publisher charges the advertiser a per-action fee $w(h)$ where h is the advertiser's history of past triggering actions. The Bellman equations that describe the system of interest in this proposition are:

$$\Omega(h) = V(p(h)) - U(p(h))w(h) + \delta\left[U(p(h))\Omega(h_+(h)) + (1 - U(p(h)))\Omega(h_-(h))\right] \tag{16}$$

where $h_+(h), h_-(h)$ denote the history following the occurrence (non-occurrence) of a triggering action in the current round respectively and where $p(h)$ satisfies the first-order condition:

$$V'(p(h)) + U'(p(h))\left[-w(h) + \delta\left(\Omega(h_+(h)) - \Omega(h_-(h))\right)\right] = 0$$

The requirement $p(h) = p^*$, such that $V(p^*) = 0$ implies:

$$w(h) = \delta\left(\Omega(h_+(h)) - \Omega(h_-(h))\right) \qquad \text{for all } h \tag{17}$$

Substituting (17) into (16) gives:

$$\Omega(h) = V(p^*) + \delta\Omega(h_-(h)) \qquad \text{for all } h \tag{18}$$

If there is no termination state, i.e. no state for which equation (18) does not hold then the above equation has the solution $\Omega(h) = \frac{V(p^*)}{1-\delta}$ for all h, which implies that $w(h) = 0$ for all h.

A termination state is a state where the game would end for the advertiser. For example a state that, once reached, the advertiser would no longer be allowed to obtain the resource. Such a state would have zero payoff. Although such a state can be potentially reachable through multiple paths, repeated application of equation (18) shows that such a state is always reached from an initial state h_0 after k consecutive periods where no triggering action occurs, for some integer $k \geq 1$. Then, by recursive application of (18) the advertiser's lifetime payoff is equal to:

$$\begin{aligned} \Omega(h_0) &= V(p^*) + \delta\Omega(h_-(h_0)) = V(p^*) + \delta\left[V(p^*) + \delta\Omega(h_-(h_-(h_0)))\right] \\ &= (1 + \delta + ... + \delta^{k-1})V(p^*) = \tfrac{1-\delta^k}{1-\delta}V(p^*) \end{aligned}$$

The probability that there will be k consecutive periods without a triggering action is $\tau = (1 - U(p^*))^k$. Therefore, the probability that an advertiser will be expelled at or before state $i + k$ is given by the CDF of a geometric distribution with parameter τ. For any probability P arbitrarily close to 1 there is, thus, a finite i, such that the probability that the advertiser will be expelled at or before state $i+k$ is higher than P. This means that the discounted lifetime joint advertiser-publisher profit in this mechanism is always lower than that achievable by an infinite horizon mechanism where the per-period joint profit was always equal to the maximum $V(p^*)$.

Proof of Proposition 7

Parts 1 and 2

Assume that every other bidder bids $\beta^E(y)$ and that (as I will show) $\left[\beta^E\right]'(y) > 0$. An advertiser of type q will choose bid $b^E(q)$ and price $p^E(q)$ that maximize:

$$\Pi^E(q; b^E(q), p^E(q), \beta^E(\cdot)) = \int\limits_{\underline{q}}^{\left[\beta^E\right]^{-1}(b^E(q))} \left(V(p^E(q), q) - \beta^E(y)\right) G'(y)dy$$

First-order conditions with respect to bid and price give:

$$\frac{\partial\left[\beta^E\right]^{-1}(b^E(q))}{\partial b^E(q)} \left(V(p^E(q), q) - b^E(q)\right) G'(\left[\beta^E\right]^{-1}(b^E(q))) = 0 \quad \text{and} \quad V_1(p^E(q), q)G(\left[\beta^E\right]^{-1}(b^E(q))) = 0$$

At a symmetric equilibrium it must be $b^E(q) = \beta^E(q)$ which implies that $G(\left[\beta^E\right]^{-1}(b^E(q))) = G(q) > 0$, $G'(\left[\beta^E\right]^{-1}(b^E(q))) = G'(q) > 0$ and $\frac{\partial\left[\beta^E\right]^{-1}(b^E(q))}{\partial b^E(q)} = \frac{1}{\left[b^E\right]'(q)} > 0$. The above then reduces to:

$$b^E(q) = V(p^E(q), q) \qquad \text{and} \qquad V_1(p^E(q), q) = 0$$

Assumption A1 implies that $p^E(q) = p^*(q)$ is uniquely defined for all q and also that $V_{11}(p^E(q), q) < 0$. Assumption A3 and the envelope theorem further imply that $\frac{\partial V(p^*(q),q)}{\partial q} > 0$ and hence that $\left[b^E\right]'(q) > 0$, as originally assumed. The corresponding Hessian matrix is:

$$H^E(b^E(q), p^E(q), q) = \begin{bmatrix} -\frac{G'(q)}{V_2(p^E(q),q)} & 0 \\ 0 & V_{11}(p^E(q), q)G(q) \end{bmatrix}$$

It is straightforward to show that H^E is negative definite and, therefore, that the above pair $(b^E(q), p^E(q))$ corresponds to a local maximum of $\Pi^E(q; b^E(q), p^E(q), \beta^E(\cdot))$ for all q.

Part 3

In a PPE setting where prices are exogenously set standard auction theory predicts that each bidder will bid her expected valuation $b^E(q; p(\cdot)) = V(p(q), q)$. Substituting into (2):

$$R^E(p(\cdot)) = N \int_{\underline{q}}^{\overline{q}} \left(\int_{\underline{q}}^{z} V(p(y), y)G'(y)dy \right) F'(z)dz$$

Because $V(p^*(y), y) \geq V(p(y), y)$ for all y (with equality iff $p^*(y) = p(y)$), it is $R^E = R^E(p^*(\cdot)) \geq R^E(p(\cdot))$ with equality iff $p^*(y) = p(y)$ for all y.

Proof of Proposition 8

Part 1

An advertiser of type q will choose bid $b^A(q, u)$ and price $p^A(q, u)$ that maximize (3). The latter can be equivalently rewritten as:

$$\Pi^A(q, u; b^A(q, u), p^A(q, u)) = V(p^A(q, u), q)\Psi(ub^A(q, u)) - U(p^A(q, u), q)Z^A(b^A(q, u), u)$$

where $Z^A(b, u) = \int_0^{ub} \frac{s}{u}\Psi'(s)ds$. Differentiating with respect to $p^A(q, u)$ gives:

$$\frac{\partial \Pi^A}{\partial p^A(q, u)} = V_1(p^A(q, u), q)\Psi(ub^A(q, u)) - U_1(p^A(q, u), q)Z^A(b^A(q, u), u) \tag{19}$$

Assumption A1 implies that:

$$\begin{aligned} &V_1(p, q) > 0 \text{ for all } p < p^*(q) \\ &V_1(p, q) = 0 \text{ for } p = p^*(q) \\ &V_1(p, q) < 0 \text{ for all } p > p^*(q) \end{aligned} \tag{20}$$

- If $U_1(p, q) < 0$ for all p then (19) and (20) imply that $\frac{\partial \Pi^A}{\partial p^A(q,u)} > 0$ for all $p^A(q, u) \leq p^*(q)$ and, therefore, that the advertiser can strictly increase net profits if she raises the price of her products above $p^*(q)$.
- If $U_1(p, q) > 0$ for all p then (19) and (20) imply that $\frac{\partial \Pi^A}{\partial p^A(q,u)} < 0$ (which is equivalent to $\frac{\partial \Pi^A}{\partial(-p^A(q,u))} > 0$) for all $p^A(q, u) \geq p^*(q)$ and, therefore, that the advertiser can strictly increase net profits if she reduces the price of her products below $p^*(q)$.
- Finally, if $U_1(p, q) = 0$ for all p then (19) and (20) imply that $\frac{\partial \Pi^A}{\partial p^A(q,u)} = V_1(p^A(q, u), q)\Psi(ub^A(q, u))$ and, therefore, that the price that maximizes $V(\cdot)$ also maximizes $\Pi^A(\cdot)$

Note that the above hold for all u and for any positive $b^A(q, u)$.

Part 2

For arbitrary u the first-order condition with respect to bid gives:

$$U(p^A(q, u), q)\left(W(p^A(q, u), q) - b^A(q, u)\right)\Psi'(ub^A(q, u))u = 0$$

which implies:

$$b^A(q, u) = W(p^A(q, u), q)$$

Part 3

In the special case where $u = U(p^A(q), q)$, the first-order condition with respect to bid gives:

$$U(p^A(q), q) \left(W(p^A(q), q) - b^A(q)\right) \Psi'(U(p^A(q), q)b^A(q)) = 0$$

which implies:

$$b^A(q) = W(p^A(q), q)$$

Taking the first order condition with respect to price and substituting the expression for $b^A(q)$ we obtain:

$$V_1(p^A(q), q)\Psi(V(p^A(q), q)) - U_1(p^A(q), q)Z^A(W(p^A(q), q), U(p^A(q), q)) = 0 \tag{21}$$

Note that in the special setting where (a) everyone bids their expected value-per-action, and (b) quality weights are equal to everyone's triggering action frequency, every advertiser's score $s = ub$ is simply equal to her expected value function $V(p^A(q), q)$. If we assume that $\frac{\partial}{\partial q}\left[V(p^A(q), q)\right] \geq 0$, i.e. that equilibrium value functions are monotonically increasing with q then advertiser scores will also be monotonically increasing with q. Under these assumptions the probability distribution $\Psi(s)$ of every other bidder's score being less than s is identical to the distribution $G(q)$ of every other bidder's type being less than q, where $s = V(p^A(q), q)$. Substituting into (21) we obtain:

$$V_1(p^A(q), q)G(q) - U_1(p^A(q), q)Z^A(q) = 0$$

where

$$\begin{aligned} Z^A(q) = Z^A(W(p^A(q), q), U(p^A(q), q)) &= \tfrac{1}{U(p^A(q),q)} \int_0^{U(p^A(q),q)W(p^A(q),q)} s\Psi'(s)ds \\ = \tfrac{1}{U(p^A(q),q)} \int_0^{V(p^A(q),q)} s\Psi'(s)ds &= \tfrac{1}{U(p^A(q),q)} \int_{\underline{q}}^{q} V(p^A(y), y)G'(y)dy \end{aligned}$$

Proof of Proposition 9

The proof makes use of the fact that, if $\frac{\partial}{\partial q}\left[V(p^A(q), q)\right] \geq 0$, the probability distribution $\Psi(s)$ of every other bidder's score being less than s is identical to the distribution $G(q)$ of every other bidder's type being less than q, where $s = V(p^A(q), q)$ (see Proof of Proposition 8).

Part 1

Substituting $u = U(p^A(q), q)$ and $b^A(q) = W(p^A(q), q)$ into (3) I obtain:

$$\begin{aligned} \Pi^A(q, U(p^A(q), q); b^A(q), p^A(q)) &= \int_0^{U(p^A(q),q)W(p^A(q),q)} U(p^A(q), q) \left(W(p^A(q), q) - \tfrac{s}{U(p^A(q),q)}\right) \Psi'(s)ds \\ &= \int_0^{V(p^A(q),q)} \left(V(p^A(q), q) - s\right) \Psi'(s)ds = \int_{\underline{q}}^{q} \left(V(p^A(q), q) - V(p^A(y), y)\right) G'(y)dy \\ &= V(p^A(q), q)G(q) - \int_{\underline{q}}^{q} V(p^A(y), y)G'(y)dy = \Pi_E(p^A(\cdot)) \end{aligned}$$

Part 2

Substituting $u = U(p^A(q), q)$ and $b^A(q) = W(p^A(q), q)$ into (4) I obtain:

$$\begin{aligned} R^A(b^A(\cdot,\cdot), p^A(\cdot,\cdot)) &= N \int_{\underline{q}}^{\overline{q}} U(p^A(q), q) \left(\int_0^{U(p^A(q),q)W(p^A(q),q)} \frac{s}{U(p^A(q),q)} \Psi'(s) ds \right) F'(q) dq \\ &= N \int_{\underline{q}}^{\overline{q}} \left(\int_0^{V(p^A(q),q)} s\Psi'(s) ds \right) F'(q) dq \\ &= N \int_{\underline{q}}^{\overline{q}} \left(\int_{\underline{q}}^{q} V(p^A(y), y) G'(y) dy \right) F'(q) dq = R_E(p^A(\cdot)) \end{aligned}$$

Proof of Proposition 10

Part 1

From Proposition 8, at an RBR-PPA equilibrium it is:

$$\Pi^A(q; p^A(q), \zeta^A(q)) = V(p^A(q), q)G(q) - U(p^A(q), q)Z^A(q) = V(p^A(q), q)G(q)(1 - \zeta^A(q)) \quad (22)$$

From Proposition 8 we also know that, if $U_1(p, q) < 0$, it is $p^A(q) > p^*(q)$, which, by Assumption A1 implies that $V_1(p^A(q), q) < 0$. From monotone comparative statics theory (Milgrom and Shannon 1994) a sufficient condition for $\partial p^A(q)/\zeta^A(q) > 0$ is that $\partial^2\Pi^A(q; p(q), \zeta(q))/\partial p(q)\partial\zeta(q) > 0$. Differentiating (22) I obtain:

$$\frac{\partial^2 \Pi^A(q; p^A(q), \zeta^A(q))}{\partial p^A(q) \partial \zeta^A(q)} = -V_1(p^A(q), q)G(q) > 0$$

Part 2

The price $p^A(q)$ that maximizes $\Pi^A(q; p(q), \zeta(q))$ must satisfy the first-order condition:

$$\begin{aligned} &V_1(p^A(q), q)G(q) - U_1(p^A(q), q)Z^A(q) \\ &\quad = U_1(p^A(q), q) \left(W(p^A(q), q)G(q) - Z^A(q) \right) + U(p^A(q), q)W_1(p^A(q), q)G(q) = 0 \end{aligned} \quad (23)$$

At the limit where $\zeta^A(q) \to 1$ it is $W(p^A(q), q)G(q) - Z^A(q) \to 0$ (follows directly from the definition $\zeta^A(q) = \frac{Z^A(q)}{W(p^A(q),q)G(q)}$). Then (23) simplifies to:

$$U(p^A(q), q)W_1(p^A(q), q)G(q) = 0$$

If $W_1(p, q) > 0$ the above implies $U(p^A(q), q) = 0$, which also implies $V(p^A(q), q) = 0$.

Proof of Proposition 11

An advertiser of type q will choose bid $b^A(q, u)$ and price $p^A(q, u)$ that maximize (3). The latter can be equivalently rewritten as:

$$\Pi^A(q,u;b^A(q,u),p^A(q,u)) = V(p^A(q,u),q)\Psi(ub^A(q,u)) - U(p^A(q,u),q)Z^A(b^A(q,u),u)$$

where $Z^A(b,u) = \int_0^{ub} \frac{s}{u}\Psi'(s)ds$. In the special case where $u = U(p^A(q),q)$ first-order conditions with respect to $b^A(q), p^A(q)$ give:

$$\begin{aligned}
&\Psi'(U(p^A(q),q)b^A(q))U(p^A(q),q)\left[V(p^A(q),q) - U(p^A(q),q)b^A(q)\right] = 0 \\
&V_1(p^A(q),q)\Psi(U(p^A(q),q)b^A(q)) \\
&\quad +\Psi'(U(p^A(q),q)b^A(q))U_1(p^A(q),q)b^A(q)\left[V(p^A(q),q) - U(p^A(q),q)b^A(q)\right] = 0
\end{aligned}$$

Solving we obtain:

$$\begin{aligned}
&b^A(q) = \tfrac{V(p^A(q),q)}{U(p^A(q),q)} = W(p^A(q),q) \\
&V_1(p^A(q),q) = 0
\end{aligned}$$

which in turn implies that $p^A(q) = p^*(q)$ and, by Proposition 9, that the mechanism is allocation and revenue equivalent to a PPE mechanism with endogenous product prices.

Proof of Proposition 12

The Bellman equations that describe the system of interest in this proposition are:

$$\begin{aligned}
&\Omega_H(q,u) = -f_H(q) + (-g_H(q) + V(p_H,q))\Psi(ub_H(q)) - U(p_H(q),q)Z(u,b_H(q)) \\
&\quad +\delta\Psi(ub_H(q))\left[U(p_H(q),q)\Omega_H(q,u) + (1-U(p_H(q),q))\Omega_L(q,u)\right] + \delta(1-\Psi(ub_H(q)))\Omega_H(q,u) \\
&\Omega_L(q,u) = -f_L(q) + (-g_L(q) + V(p_L(q),q))\Psi(ub_L(q)) - U(p_L(q),q)Z(u,b_H(q)) \\
&\quad +\delta\Psi(ub_L(q))\left[U(p_L(q),q)\Omega_H(q,u) + (1-U(p_L(q),q))\Omega_L(q,u)\right] + \delta(1-\Psi(ub_L(q)))\Omega_L(q,u)
\end{aligned} \tag{24}$$

where

$$Z(u,b) = \frac{1}{u}\int_0^{ub} s\Psi'(s)ds$$

and $b_H(q), b_L(q), p_H(q), p_L(q)$ satisfy the first-order conditions:

$$\begin{aligned}
&-g_H(q) + U(p_H(q),q)\left[W(p_H(q),q) - b_H(q)\right] - \delta(1-U(p_H(q),q))(\Omega_H(q,u) - \Omega_L(q,u)) = 0 \\
&-g_L(q) + U(p_L(q),q)\left[W(p_L(q),q) - b_L(q)\right] + \delta U(p_H(q),q)(\Omega_H(q,u) - \Omega_L(q,u)) = 0 \\
&V_1(p_H(q),q)\Psi(ub_H(q)) + U_1(p_H(q),q)\left[-Z(u,b_H(q)) + \delta\Psi(ub_H(q))(\Omega_H(q,u) - \Omega_L(q,u))\right] = 0 \\
&V_1(p_L(q),q)\Psi(ub_L(q)) + U_1(p_L(q),q)\left[-Z(u,b_L(q)) + \delta\Psi(ub_L(q))(\Omega_H(q,u) - \Omega_L(q,u))\right] = 0
\end{aligned}$$

We would like to set the schedule of fees and subsidies so that (for all q) advertisers (i) price their products at the socially efficient level $p^*(q)$ defined by $V_1(p^*(q),q) = 0$, (ii) bid their expected value-per-action $W(p^*(q),q)$ and (iii) obtain a lifetime discounted payoff equal to the payoff they would obtain if they bid and priced that way without any fees or subsidies imposed. Mathematically these requirements are equivalent to:

$$\begin{aligned}
&p_H(q) = p_L(q) = p^*(q)\\
&b_H(q) = b_L(q) = W(p^*(q), q)\\
&\Omega_H(q,u) = \tfrac{V(p^*(q),q)\Psi(ub_H(q)) - U(p^*(q),q)Z(u,b_H(q))}{1-\delta}
\end{aligned}$$

and, in turn, imply the following:

$$\begin{aligned}
&g_H(q) = -\delta(1 - U(p^*(q), q))(\Omega_H(q,u) - \Omega_L(q,u))\\
&g_L(q) = \delta U(p^*(q), q)(\Omega_H(q,u) - \Omega_L(q,u))\\
&Z(u, b_H(q)) = \delta\Psi(ub_H(q))(\Omega_H(q,u) - \Omega_L(q,u))\\
&\Omega_H(q,u) = \tfrac{V(p^*(q),q)\Psi(ub_H(q)) - U(p^*(q),q)Z(u,b_H(q))}{1-\delta}
\end{aligned} \tag{25}$$

Substituting (24) into (25) and solving for $f_H(q), f_L(q), g_H(q), g_L(q)$ we obtain:

$$\begin{aligned}
&f_H(q) = 0\\
&f_L(q) = \tfrac{1-\delta}{\delta\Psi(uW(p^*(q),q))} Z(q)\\
&g_H(q) = -\tfrac{1-U(p^*(q),q)}{\Psi(uW(p^*(q),q))} Z(q)\\
&g_L(q) = \tfrac{U(p^*(q),q)}{\Psi(uW(p^*(q),q))} Z(q)
\end{aligned}$$

where $Z(q) = Z(u, W(p^*(q), q)) = \frac{1}{u}\int_0^{uW(p^*(q),q)} s\Psi'(s)ds$. If the publisher sets each advertiser's quality weight equal to her triggering action frequency and if, additionally, it is $\frac{\partial}{\partial q}V(p^A(q), q) \geq 0$ for all q then (see proof of Proposition 8) it is $\Psi(uW(p^*(q), q)) = \Psi(U(p^*(q), q)W(p^*(q), q)) = \Psi(V(p^*(q), q)) = G(q)$ and the above simplify to:

$$\begin{aligned}
&f_H(q) = 0\\
&f_L(q) = \tfrac{1-\delta}{\delta G(q)} Z(q)\\
&g_H(q) = -\tfrac{1-U(p^*(q),q)}{G(q)} Z(q)\\
&g_L(q) = \tfrac{U(p^*(q),q)}{G(q)} Z(q)
\end{aligned}$$

Furthermore, under the above assumptions, from Proposition 9 the numerator of $\Omega_H(q)$ is identical to the advertiser's payoff in a PPE mechanism with optimal endogenous pricing. Therefore, the advertiser's lifetime discounted payoff is identical to the lifetime discounted payoff of an infinite sequence of PPE mechanisms with endogenous prices. It is similarly easy to show that the net lifetime discounted sum of fees $f_H(q), f_L(q), g_H(q), g_L(q)$ is equal to zero. Therefore, the publisher's revenue is also identical to that of an infinite sequence of PPE mechanisms with endogenous prices.

Proof of Proposition 13

The proof is sketched in the main body of the text.

Appendix: Example of a pure PPA mechanism that induces efficient product pricing

Consider a mechanism that has the following properties:

1. All advertisers begin the game at an initial "perfect" state 0.
2. During a period where the advertiser's state is i the publisher charges a per-action fee w_i
3. At the end of a period where the advertiser's state is i: If a triggering action occurs, the advertiser transitions back to state 0. If no triggering action occurs the advertiser transitions to state $i+1$
4. Once the advertiser reaches state k she is expelled from the game (i.e. no longer allowed to obtain the advertising resource)

It is easy to see that the above mechanism is a special case of the general class of mechanisms described by Proposition 6. Therefore, for an appropriately chosen schedule of fees w_i ($i = 0, 1, .., k-1$), it can induce advertisers to always price their products at $p^A = p^*$. In this appendix I provide a full analysis of the mechanism.

Specializing the general results of Proposition 6 to this special case we immediately obtain the following:

1. The Bellman equations that describe the advertiser's payoffs are:

$$\Omega_i = \begin{cases} V(p^*) - U(p^*)w_i + \delta\left[U(p^*)\Omega_0 + (1-U(p^*))\Omega_{i+1}\right] = \frac{1-\delta^{k-i}}{1-\delta}V(p^*) & i < k \\ 0 & i = k \end{cases}$$

2. The publisher's fee schedule is given by:

$$w_i = \delta\left(\Omega_0 - \Omega_{i+1}\right) = \delta\frac{\delta^{k-i-1} - \delta^k}{1-\delta}V(p^*) \quad i = 0, 1, ..., k-1$$

The lifetime discounted joint advertiser-publisher profit is J_0, where J_0 is the solution of:

$$J_i = \begin{cases} V(p^*) + \delta\left[U(p^*)J_0 + (1-U(p^*))J_{i+1}\right] & i < k \\ 0 & i = k \end{cases}$$

Taking successive differences:

$$J_i - J_{i+1} = \begin{cases} \delta(1-U(p^*))\left(J_{i+1} - J_{i+2}\right) & i < k-1 \\ V(p^*) + \delta U(p^*)J_0 & i = k-1 \end{cases}$$

and backwards-substituting into J_0 we obtain:

$$J_0 = \frac{1-\delta^k(1-U(p^*))^k}{1-\delta[1-\delta^k(1-U(p^*))^k U(p^*)]}V(p^*)$$

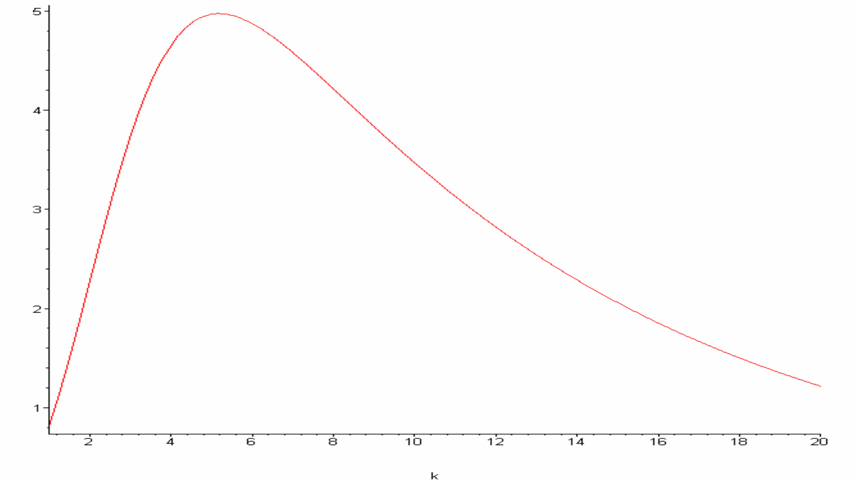

Figure 2: Publisher's revenue R_0 as a function of the threshold k of consecutive rounds without triggering actions after which an advertiser is expelled.

The publisher's discounted lifetime profit is simply:

$$R_0 = J_0 - \Omega_0 = \left[\frac{1 - \delta^k (1 - U(p^*))^k}{1 - \delta[1 - \delta^k (1 - U(p^*))^k U(p^*)]} - \frac{1 - \delta^k}{1 - \delta} \right] V(p^*)$$

It is easy to see that:

1. Both Ω_0 and J_0 are monotonically increasing functions of k.
2. At the limit $k \to \infty$, $\Omega_0 \to J_0 \to \frac{1}{1-\delta} V(p^*)$, $w_i \to 0$ for all i, and $R_0 \to 0$

For finite values of k, as k increases R_0 increases, then decreases. There is, therefore, an optimal value k^* that maximizes the publisher's profits. The intuition behind this result is the following: When k is small (i.e. when advertisers get expelled from the game after a relatively small number of consecutive rounds without an observed triggering action), advertisers leave the game too soon, which hurts the publisher's lifetime profit per advertiser. When k is large then most of the time the fees charged by the publisher are very low, which again hurts his lifetime profit per advertiser.

Figure 2 plots R_0 for $\delta = 0.9$, $V(p^*) = 1$ and $U(p^*) = 1/2$. The integer value that maximizes R_0 is $k^* = 5$. Setting $k = 5$ to the above formulae we obtain:

$$\begin{array}{ll} J_0 = 9.06 & w_0 = 0.59 \\ \Omega_0 = 4.09 & w_1 = 1.25 \\ R_0 = 4.97 & w_2 = 1.97 \\ & w_3 = 2.78 \\ & w_4 = 3.68 \end{array}$$

Notice that the split of the joint profits J_0 between the publisher (R_0) and the advertiser (Ω_0) is relatively equitable (55% goes to the publisher and 45% goes to the advertiser). Also, the per-action publisher fees w_i monotonically increase as the number of consecutive periods with no triggering action increases. Finally, per Proposition 6, the discounted lifetime joint advertiser-publisher profit $J_0 = 9.06$ that is attainable with this mechanism is strictly below the optimal $J^* = \frac{V(p^*)}{1-\delta}=10$ that is attainable through the mechanisms analyzed in Propositions 4 and 5. The latter mechanisms involve the use of both performance-based and non-performance-based fees.