

Double Marginalization in Performance-Based Advertising: Implications and Solutions

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

First Draft: March 2010. This version: July 2011

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, Yahoo and Microsoft. In the latter settings they also reduce publisher revenues relative to pay-per-exposure methods. 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, improving both consumer surplus as well as the publisher's expected profits.

*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. Although the literature 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 (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, such price and revenue distortions arise in the practically important case of advertising auctions. Auction-based allocation of advertising resources is the norm in sponsored search advertising (see discussion and references in Section 2) and is increasingly being used to sell radio and TV ads (Nisan et al. 2009). I show that in the auction-based variants of PPA advertising currently practiced by Google, Yahoo and Microsoft, the price distortions that form the focus of this work *always* lead to lower publisher revenues relative to pay-per-exposure methods. I propose a number of enhancements to today’s auction-based PPA mechanisms that restore equilibrium pricing of advertised goods to efficient levels, improving both consumer surplus as well as the publisher’s expected profits.

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 shows how these ideas apply to auction-based search advertising. Section 5 proposes ways of restoring efficient product pricing in auction-based search advertising. Section 6 presents some important modeling extensions. Finally, Section 7 concludes.

2 Related Work

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 researchers and practitioners in the fields of marketing, economics, information systems and computer science. 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. (2008) and Yao and Mela (2009).

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 2, i.e. (using the language of my paper) that advertisers price their product at the point that maximizes the joint advertiser-publisher profit.

Feng and Xie (2007) 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.

3 The core intuitions

For pedagogical reasons I first establish the connection between the double marginalization problem in supply chain management and price distortions in performance-based advertising through a simple posted price setting that best exposes the core intuitions. Section 4 then shows that similar phenomena arise in the practically important settings where advertising is sold via auctions.

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 $\pi_m(w|p) = D(p)(w - c_m) =$

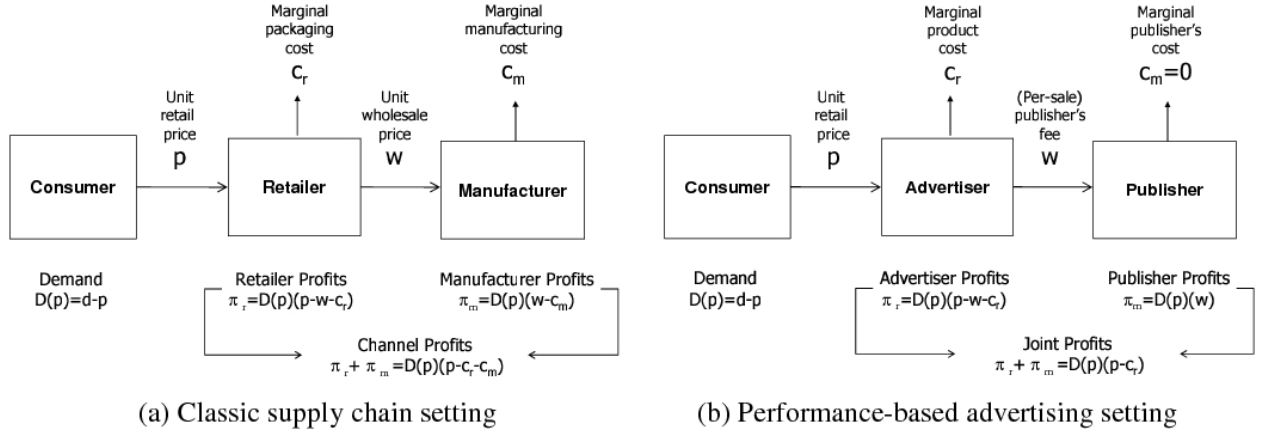


Figure 1: Analogies between a classic supply chain setting and performance-based advertising.

$(d - \frac{d+c_r+w}{2})(w - c_m)$. 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 Analogy with performance-based advertising

I now show that 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.

To keep the exposition as simple as possible my baseline models assume that consumers will not be able to purchase the product unless the product is advertised on the publisher's resource. This assumption implies that the advertising resource acts as the exclusive gateway to the channel where the product is sold. In an online setting this might mean that the only way to purchase the product would be to click on a keyword ad and be taken to the advertiser's online storefront (that is assumed to not be accessible otherwise). In an offline setting this might mean that the only way to purchase

the product is by calling a number that is only available on the ad.

Under these assumptions, gaining access to the publisher’s resource results in demand for the product equal to $D(p)$ (Figure 1b). Let $V(p)$ denote the *ex-ante* value the advertiser expects to obtain from leasing the advertising resource. This value is equal to the 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.¹

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 sales profit. For example, if $D(p) = d - p$ (where $d > c_r$) the retail price would be set to $p^* = \frac{d+c_r}{2}$, 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 .²

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 $\frac{d+c_r+w}{2} = \frac{3d+c_r}{4}$. 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.

¹The preceding assumptions aim to produce a minimally complex model for the purpose of mathematical tractability. A more realistic set of assumptions would be to posit that the advertiser generates demand $D_0(p)$ through other channels and that consumers have a variety of ways in which they can purchase the product. Access to an additional advertising resource then offers the advertiser *incremental* demand $D_1(p)$. Section 6.1 extends my baseline models to such multi-channel settings and shows that all key results qualitatively apply there as well. The magnitude of price and revenue distortions is then positively correlated with the fraction of demand that is generated via performance-based advertising.

²Theory does not give a crisp prediction with respect to the level of fee f , which generally depends on the relative market power of the two parties. One way of deriving an expression for f is to assume that the advertiser has a reservation utility u , in which case the maximum fee that will induce the advertiser to purchase the ad is $f = V(p^*) - u$.

3.3 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. *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.

As before, let $V(p)$ denote the *ex-ante* value (demand times unit profit) the advertiser expects to obtain from leasing the advertising resource. 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 (Figure 2):

- In pay-per-sale contracts $U(p)$ is equal to the 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 *click-through 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.³

³Addressing 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.

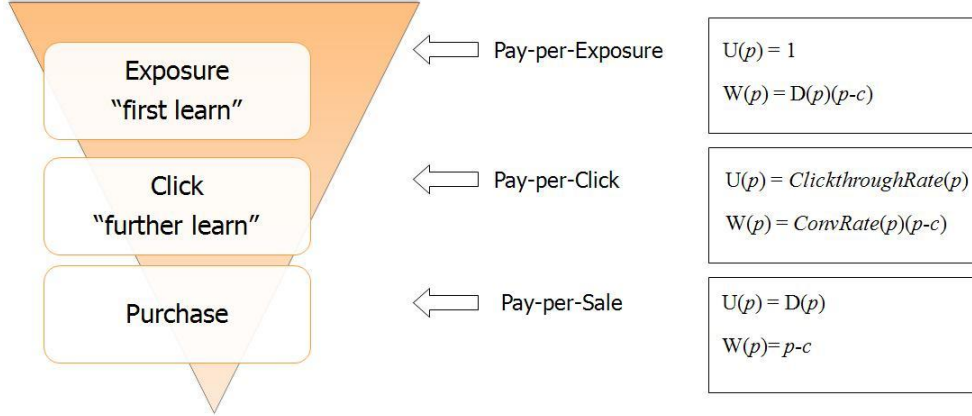


Figure 2: The idea behind pay-per-action is to pay the publisher upon every occurrence of a triggering action further down the purchasing funnel (e.g. click, sale). The total value (sales profits) $V(p)$ advertisers expect from leasing the advertising resource can, thus, be expressed as the product of the triggering action frequency $U(p)$ times the value per action $W(p)$. The definitions of $U(p)$ and $W(p)$ depend on the specific form of triggering action used. In contrast, their product $V(p)$ is independent of the payment method.

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 3.2) equal to the product price p^E that the advertiser would choose under a pay-per-exposure contract.⁴ 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 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.*

⁴I am using superscripts to denote the payment mechanism that quantities of interest are associated with. For example, p^E denotes the equilibrium product price advertisers will choose when advertising is sold using a pay-per-exposure (E) mechanism whereas p^A denotes the equilibrium product price when advertising is sold using a pay-per-action (A) mechanism.

As was the case with the classic double marginalization problem, the above phenomena are fairly general: they occur in any situation where there is a downward sloping triggering action frequency $U(p)$ and persist in settings with competition among advertisers and/or publishers as long as they each retain *some* market power and are able to mark up their respective prices above marginal cost. The case of downstream competition among advertisers is discussed in Section 6.2. When it comes to competing publishers, Proposition 1 holds as long as each publisher j retains enough market power to charge a per-action fee $w_j > 0$ that is above marginal cost. Finally, the key results are orthogonal to the specific mechanism used by the publisher to allocate his resource to advertisers. More concretely, Section 4 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.

The assumption $U'(p) < 0$ (more broadly $U'(p) \neq 0$) is a necessary condition for price distortions to occur. This assumption 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. This assumption obviously holds in the context of pay-per-sale contracts where prices are known to consumers by the time (completion of a sale) that payment to the publisher is triggered. It also holds in pay-per-click or pay-per-call settings when one or more of the following conditions hold:

1. product prices are displayed on the ad
2. at least some consumers have 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.)
3. at least some consumers have some industry knowledge (i.e. “know” something about the advertiser’s value function and the publisher’s price structure) to be able to infer (possibly imperfectly) these prices

Current sponsored search ads typically do not display product prices. Nevertheless, even in such settings there are usually information spillovers (e.g. through consumer reviews, word of mouth, social networks, retail channels, etc.) that result in some consumers being informed about product prices before they click on the ad. By conditions 2 and 3 I expect that there will still be a negative relationship between $U(p)$ and product price. However, this relationship is expected to be milder, and the resulting price distortions smaller, relative to settings where product prices are listed on ads.

4 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 search advertising auctions.

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 uni-dimensional type $q \in [\underline{q}, \bar{q}]$, 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 3.3 value $V(p, q)$ will be equal to the sales profit that the advertiser expects to realize by leasing the resource.

The following are assumed to hold for all $p \in \mathbb{R}^+$ and $q \in [\underline{q}, \bar{q}]$:⁶

A1 $V(p, q)$ is uni-modal in p , attaining its unique maximum at some $p^*(q) > 0$

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

A3 $V_2(p, q) > 0$

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. 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 $[\beta^E]'(y) \geq 0$, 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:

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

where $G(y) = [F(y)]^{N-1}$ is the probability that every other bidder's type is less than or equal to y

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.

⁶ I use the notation $f_i(x_1, \dots, x_n)$, $i = 1, \dots, n$ to denote the partial derivative of $f(\cdot)$ with respect to its i th variable. For example, $V_2(p, q) = \frac{\partial V(p, q)}{\partial q}$.

and $G'(y)$ is the corresponding density.⁷ 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:

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

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 PPE bidding

Proposition 2. *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:*

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

⁷ 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".

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

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 3.3 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.⁸ For that reason they were quickly abandoned in favor of, more sophisticated, *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). Google, Yahoo and Microsoft 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

⁸For example, consider a setting where the advertisers’ per-action value $W(p, q)$ is negatively correlated with their total value $V(p, q) = U(p, q)W(p, q)$ 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 though the total value she expects to generate by leasing the resource is the lowest among all advertisers.

problem is to choose a bid $b^A(q, u)$ and a price $p^A(q, u)$ that maximize:

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

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:

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

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 $\frac{\partial}{\partial q} [V(p^A(q), q)] \geq 0$ 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 3. *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:

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

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 3 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.

As was the case in the fixed price setting, the above results also hold in settings where there are multiple competing advertisers and/or multiple competing publishers. The case of multiple advertisers competing in the product market is discussed in Section 6.2. In the case of multiple publishers the setting reverts to one of simultaneous competing auctions (Bapna et al. 2010). Theory then predicts that, if auction participation costs are sufficiently high, advertisers probabilistically choose to participate in a single auction; within that auction, they bid as if that auction was the only one. Therefore all results qualitatively hold, with appropriately adjusted functions $\Phi(\cdot)$ and $\Psi(\cdot)$.⁹

The rest of the section will show that, in this special context, this price distortion reduces consumer surplus, social welfare and publisher revenues. Just as I did in Section 3.3, I will focus my attention on the practically important case where $U_1(p, q) < 0$ for all p, q .¹⁰

From Proposition 3, 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:

⁹If auction participation costs are low, Bapna et al. (2010) predict an equilibrium where bidders place bids in all auctions. Price distortions persist in such an equilibrium because their fundamental driver, i.e. the advertiser's desire to minimize the amount paid to the publisher by decreasing the triggering action frequency and increasing the value per action, is still there.

¹⁰Proposition 4 and Corollaries 1 and 2 hold in settings where there is no publisher moral hazard. See Section 5.1 for settings with publisher moral hazard.

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.*

To derive the impact on publisher revenues, the key result 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 4. *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 3, in general it will be $p^A(\cdot) \neq p^*(\cdot)$. From Proposition 2 (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.*

5 Restoring efficiency in search advertising auctions

The previous two sections showed that performance-based advertising is vulnerable to price distortions akin to double marginalization; such distortions reduce joint profits and, in the case of auctions, publisher profits. These results raise the reasonable question of whether the industry should revert to selling advertising using PPE. This section begins by establishing that, if there is any amount of publisher moral hazard and if such price distortions can be avoided, PPA induces higher joint *and* publisher profits relative to traditional PPE. This makes it worthwhile to design mechanisms that eliminate price distortions in RBR-PPA settings. The rest of the section proposes two such mechanisms and explains why these mechanisms result in strict joint profit and publisher profit improvements relative to standard RBR-PPA mechanisms that are currently used in practice.

5.1 The case for performance-based advertising

A number of studies provide theoretical justification for the current trend towards performance-based advertising payment schemes and distill conditions under which publishers would opt to use

such schemes over traditional pay-per-exposure. Hu (2006) and Zhao (2005) explain the growing popularity of performance-based advertising from a moral hazard perspective. Their main argument is that publishers can make non-contractible efforts that improve the attractiveness of advertising campaigns (and thus, the willingness of advertisers to pay for such campaigns). Performance-based payment provides the necessary incentives for them to do so. Liu and Viswanathan (2010) examine the publisher's choice between pay-per-performance and pay-per-exposure from an adverse selection framework. They find that when there is competition among publishers and consumers are uncertain about each publisher's quality, higher quality publishers have incentive to offer pay-per-performance as a means to separate themselves from low quality competitors. All of these studies, however, assume that product prices are independent of the payment mechanism used by the publisher.

Below I provide a simplified version of the moral hazard argument that forms the basis of Hu (2006) and Zhao (2005), adapted to the context this paper.¹¹ I augment the setting of Section 4 by assuming that the publisher can engage in non-contractible costly effort ϵ that will improve the effectiveness of the resource he offers to advertisers. Such effort could be related to the quality of the content that is offered in conjunction with advertising, the amount of purchased information (on consumer demographics, interests, etc.) that the publisher uses to better match advertisements to consumers, etc. In this augmented setting I denote the advertiser's expected value from leasing the resource by $V(p, q, \epsilon) = D(p, q, \epsilon)(p - c(q))$ or, more generally, by $V(p, q, \epsilon) = U(p, q, \epsilon)W(p, q)$ where $D_3(p, q, \epsilon) \geq 0$ (more generally $U_3(p, q, \epsilon) \geq 0$). Notice that the assumption I am making is that publisher effort positively affects demand (more generally, triggering action frequency) whereas unit profit (more generally, value per action) only depends on the advertiser's parameters. These assumptions imply $V_3(p, q, \epsilon) \geq 0$.

In such a setting, at the limit where the publisher sets every advertiser's quality weight to her respective expected equilibrium triggering action frequency, by Proposition 4, publisher profits are equal to:

$$\pi_P(p^i(\cdot, \epsilon^i), \epsilon^i) = R^E(p^i(\cdot, \epsilon^i), \epsilon^i) - c_P(\epsilon^i) = N \int_{\underline{q}}^{\bar{q}} \left(\int_{\underline{q}}^q V(p^i(y, \epsilon^i), y, \epsilon^i) G'(y) dy \right) F'(q) dq - c_P(\epsilon^i) \quad (6)$$

where $i \in \{A, E\}$ denotes the payment scheme, $p^i(\cdot), \epsilon^i$ are the corresponding equilibrium product price and publisher effort levels and $c_P(\cdot)$ is the cost of the publisher's effort.

When advertising is sold under PPE, payment is typically exchanged (or contractually agreed upon) before the publisher engages in action ϵ^E . Standard moral hazard arguments then predict that, at any equilibrium, it will be $\epsilon^E = 0$ and publisher profits will be equal to $\pi_P(p^*(\cdot, 0), 0)$, where $p^*(q, \epsilon) = \arg \max_p V(p, q, \epsilon)$. In contrast, when advertising is sold using RBR-PPA, equation (4) augmented by the above set of assumptions imply that:

1. advertisers still bid their expected value per action $W(p, q)$; this is independent of the pub-

¹¹I prefer to motivate pay-per-performance advertising using moral hazard than adverse selection, as the former set of concerns persists even after repeated interactions between the advertiser and the publisher, whereas the latter tends to be resolved once the publisher and advertiser have acquainted themselves with each other sufficiently well.

lisher's efforts

2. publisher revenues are equal to the product of triggering action frequency $U(p, q, \epsilon)$ times the expected payment per action; the latter is a function of bids and, thus, independent of the publisher's efforts.

PPA publishers, therefore, have an incentive to set their effort at the level where the resulting triggering action frequency $U(p, q, \epsilon^A)$ maximizes $\pi_P(p^A(\cdot, \epsilon^A), \epsilon^A) = R^E(p^A(\cdot, \epsilon^A), \epsilon^A) - c_P(\epsilon^A)$. Since $V_3(p, q, \epsilon) \geq 0$ the envelope theorem gives $\frac{\partial}{\partial \epsilon} [V(p^*(q, \epsilon), q, \epsilon)] \geq 0$. Because (see equation (6)) the publisher's revenue $R^E(p^i(\cdot, \epsilon^i), \epsilon^i)$ is an integral of advertiser value functions, this also implies $\frac{\partial}{\partial \epsilon} [R^E(p^*(q, \epsilon), \epsilon)] \geq 0$. If, additionally, it is $\frac{\partial}{\partial \epsilon} [\pi_P(p^*(q, \epsilon), \epsilon)]_{\epsilon=0} \geq 0$ (a condition satisfied as long as $c'_P(0)$ is not very large) the optimum publisher PPA effort level ϵ^A will be above zero. All this means that, if PPA advertisers can be induced to price their products at the efficient level $p^*(q, \epsilon)$, PPA results in:

1. higher publisher effort ($\epsilon^A > 0 = \epsilon^E$),
2. higher joint profit ($V(p^*(q, \epsilon^A), q, \epsilon^A) > V(p^*(q, 0), q, 0)$) and
3. higher publisher profit ($\pi_P(p^*(q, \epsilon^A), \epsilon^A) > \pi_P(p^*(q, 0), 0)$)

relative to PPE. This makes worthwhile to design mechanisms that induce efficient pricing under PPA and (assuming that such mechanisms are in place) rational for publishers to choose PPA over PPE. In the the rest of this Section I propose two such mechanisms.

The reader can appreciate that the above argument requires the additional assumption of some form of publisher moral hazard but, at the same time, is orthogonal to the double marginalization concerns that form the core focus of this paper. For the sake of notational and expositional clarity, in the rest of this paper I will omit any further mention of publisher efforts. I will, however, explicitly note all results that depend on the presence or absence of publisher moral hazard.

5.2 Price-based quality weights

The use of *quantity discount* contracts is a common method for solving the double marginalization problem in supply chain settings (see, for example, Jeuland and Shugan 1983). 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, it can be shown that such a contract provides the retailer with incentives to keep her price at levels that optimize joint profits.

This idea is not directly applicable in auction-based mechanisms where the publisher cannot directly set per-action fees. However, in RBR-PPA 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 3 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 truthful elicitation of each round’s product prices is possible (see more about this point below), 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 5. Consider a RBR-PPA mechanism where:

- (i) Each round advertisers $i = 1, \dots, N$ are asked to disclose their bids b_i as well as to announce (and truthfully commit to) product prices 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 if there is no publisher moral hazard
3. results in higher publisher revenue to that of a PPE mechanism with endogenous prices if there is publisher moral hazard

The above solution is conceptually elegant and works in settings where the triggering action frequency is a deterministic or stochastic function of product price, as well as in the important special case of personalized ad settings where $U(p, q)$ is a probability function. In practice, it requires the advertiser to *truthfully commit* to a product price and the publisher to be able to monitor the extent to which the advertiser adheres to her commitment. It is easy to show that, in the absence of effective monitoring, the advertiser has an incentive to announce a very low price to the publisher (so that the publisher assigns her a high quality weight) while actually charging another, higher, price to consumers. The feasibility of price-based fee schedules is, therefore, limited to settings where product prices are transparent or where the publisher has access to a reliable infrastructure for monitoring actual prices. In such cases the publisher can write a contract that provides sufficiently harsh penalties if the advertiser deviates from her announced price, making truthful price announcements rational.

5.3 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 by committing to reward (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 the triggering action was (not) observed in the current round). The ideas that follow assume that the publisher has interacted with the advertiser sufficiently often to have correctly learned her triggering action frequency function $U(p, q)$.

In auction-based settings 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 M1:

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 provides the details.

Proposition 6. *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 mechanism M1 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:*

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

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 if there is no publisher moral hazard*
3. *results in higher publisher revenue relative to that of an infinite horizon PPE mechanism with endogenous prices if there is publisher moral hazard*

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. 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.

An important property of 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.

5.4 Why existing RBR-PPA implementations are inefficient

Most current implementations of RBR-PPA (e.g. by Google) involve a dynamic process whereby the publisher iteratively learns an advertiser's quality weight from observations of the advertiser's triggering action frequencies in past periods (Pandey and Olston 2006). This section shows that the discounted lifetime joint publisher-advertiser profit - as well as corresponding publisher's profit - that can be achieved by such mechanisms is strictly lower than what can be achieved by Mechanism M1. Therefore, supplementing existing RBR-PPA mechanisms with the schedule of fixed fees and subsidies discussed in the previous section offers tangible improvements to current practices.

I first consider a "pure learning" setting where the learning process eventually converges to a steady state "correct" quality weight U , such that at equilibrium the advertiser chooses her price so that the resulting triggering action frequency is indeed equal to U . In such a setting, the quality

weight remains stable after convergence and the situation becomes equivalent to that of a single round setting. In fact, the setting becomes equivalent to that of Proposition 3. As that proposition shows, price distortions then persist ad infinitum, reducing joint and publisher profits.

One can attempt to induce permanent disincentives to distort prices by extending the above (“pure learning”) quality weight updating mechanism with a sanctioning (“punishment”) component: For example, the publisher can commit to increase an advertiser’s quality weight when triggering actions occur and to decrease it when they don’t, irrespective of how many rounds have elapsed in the game. 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 3) $b_i = W(p^A(q_i), q_i)$. 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. Since the publisher’s (auctioneer’s) revenues are positively related to the auction winner’s resource value, publisher revenues are also reduced. 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 summarizes the result:

Proposition 7. *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 cannot credibly disclose her product price to the publisher, the discounted lifetime joint profit and publisher profit that can be achieved by an RBR-PPA mechanism are strictly lower than what can be achieved by Mechanism M1.*

In conclusion, restoring efficient product pricing in search advertising auctions while also maintaining allocative efficiency 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.

6 Model Extensions

6.1 Multiple channels

To better isolate the double marginalization phenomena that form the focus of this paper, the models of the previous sections made the assumption that the advertising resource is the exclusive gateway to the only channel through which the advertiser sells her products and, therefore, that without this

resource, demand is zero. A more realistic set of assumptions would be to assume that the advertiser generates demand $D_0(p)$ through a variety of standard PPE channels (whose cost is of no interest to us) and sets a single price across all channels. Access to an additional advertising resource then offers the advertiser *incremental* demand $D_1(p)$. This section shows that the key results of this paper qualitatively hold in such multi-channel settings but the magnitude of price and profit distortions are then scaled by a factor that is positively related to the fraction of demand that is generated through PPA advertising.

To more clearly illustrate the implication of having multiple channels I set $D_0(p) = (1 - \alpha)D(p)$ and $D_1(p) = \alpha D(p)$, where $0 < \alpha \leq 1$, that is, I assume that the incremental demand represents a fraction α of total demand $D(p)$. Finally, I denote with $V(p) = D(p)(p - c_r)$ the advertiser's total sales profit (before advertising expenses).

When access to the additional resource is made using PPE contracts the advertiser's profit function is $\pi_r(p|f) = D_0(p)(p - c_r) + D_1(p)(p - c_r) - f = D(p)(p - c_r) - f = V(p) - f$. As in Section 3.2, the publisher's fee f does not impact the advertiser's pricing problem; the advertiser will choose the price p^E that maximizes her total sales profit. That price solves:

$$V'(p^E) = 0 \tag{7}$$

In contrast, when the additional resource is sold using pay-per-sale contracts with a per-sale fee of w the advertiser's profit function becomes $\pi_r(p|w) = D_0(p)(p - c_r) + D_1(p)(p - c_r - w) = V(p) - \alpha D(p)w$. The price p^A that maximizes $\pi_r(p|w)$ now solves:

$$V'(p^A) - \alpha D'(p^A)w = 0 \tag{8}$$

Under the assumption $D'(p) < 0$, comparison of (7) and (8) shows that, as before, it is $p^A > p^E$ and $V(p^A) < V(p^E)$ for all $0 < \alpha \leq 1$. However, the degree to which p^A diverges from p^E is moderated by the fraction α of total demand that is generated using PPA advertising. The smaller that fraction, the lower the product price p^A under PPA. (This is easy to show from monotone comparative statics theory by observing that $\frac{\partial^2 \pi_r(p|w)}{\partial p^A \partial \alpha} = -D'(p^A)w > 0$.) Furthermore, under the assumption $V''(p) < 0$, the smaller the difference $p^A - p^E$, the smaller the difference $V(p^E) - V(p^A)$, i.e. the sales profit losses due to the higher product price.

It is straightforward to show that the above results apply under the more general formulation of Section 3.3 as well as in the auctions setting of Section 4.

In summary, the price and profit distortions that form the focus of this paper are also present in settings where advertisers use multiple advertising channels and derive only part of their demand through PPA advertising. However, the relative magnitude of the distortions is then scaled back and is positively correlated with the fraction of total demand that is generated through PPA methods. The practical significance of these phenomena is likely to be higher in sectors that do a lot of advertising using keyword auctions or other methods that use performance-based pricing (e.g.

referrals). Given the industry’s trend towards performance-based advertising (beyond just keyword auctions) the impact of double marginalization is likely to increase in the future as more firms will generate more of their demand using performance-based methods. It is, therefore, important that practitioners are aware of it and take it into consideration when choosing how to advertise, as well as when developing future advertising payment mechanisms.

6.2 Downstream Competition

The preceding sections assume a monopolist advertiser (even in the auction model, the advertiser who wins the auction acts as a monopolist in the product market). This section shows that upward price distortions due to double marginalization persist in many settings with imperfect downstream competition, i.e. in settings where there are multiple advertisers competing in the same market and where each advertiser has sufficient market power to individually choose her product prices. I convey the key intuition in a fixed price setting where advertising is sold using either pay-per-exposure (PPE) or using pay-per-sale (PPS). The extension to general PPA settings is straightforward.

I label the focal competitor as i and use the notation $-i$ to refer to every other competitor. Each competitor’s demand has the general form $D^i(p_i, p_{-i})$ where I am assuming $D_1^i(p_i, p_{-i}) \leq 0$ and $D_2^i(p_i, p_{-i}) \geq 0$, i.e. that demand has a negative relationship with one’s own product price and a positive relationship with everybody else’s price. The following proposition conveys the key result:

Proposition 8. *In any setting with imperfect competition among advertisers, if advertiser i unilaterally changes the way she pays for advertising from PPE to PPS and the cross-price elasticities of demand $D_2^i(p_i, p_k)$, $k \neq i$ are not very high, the shift leads to higher equilibrium product prices charged by i . This result is independent of how every other advertiser pays for advertising.*

To understand the intuition behind this result, keep in mind that PPS causes upward price distortions if and only if increases to an advertiser’s own price result in decreases to her demand. Price increases can then be attractive from the advertiser’s perspective because demand losses can be compensated by gains in unit profit and reduced payments to the publisher. The negative relationship between own price and demand would certainly be true if an advertiser’s price increase did not affect anyone else’s price. In competitive settings, however, price increases by one advertiser may lead to price increases by her competitors. By $D_2^i(p_i, p_{-i}) \geq 0$ such competitive price increases would moderate (or even reverse) the net impact of an advertiser’s own price increase on her demand. If, however, the cross-price elasticities of demand $D_2^i(p_i, p_{-i})$ are sufficiently low so that the net impact of own price increases on own demand remains negative, the key result of this paper remains valid.

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 (Mahdian and Tomak 2008) and also provide publishers with incentives to make non-contractible efforts that improve the attractiveness of advertising campaigns (Hu 2006, Zhao 2005).

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, Yahoo and Microsoft. 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 as well as contracts that provide deterrents to misreporting.

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). 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, throughout most of the analysis I 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. Section 6.1 shows that, if this assumption is relaxed, for example if the advertiser sells through multiple channels and charges the same price 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 performance-based advertising.

Finally, 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.

References

- 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.
- Athey, S., and Ellison, G. (2008) Position Auctions with Consumer Search. Working paper.
- Bapna, R., Dellarocas, C. and Rice, S. (2010) Vertically-differentiated Simultaneous Vickrey Auctions: Theory and Experimental Evidence. *Management Science* 56 (7), 1074-1092.
- 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. 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>.
- Hu, J. (2004) Performance-based Pricing Models in Online Advertising. Available on SSRN <http://ssrn.com/abstract=501082>.

- Jeuland, A. P. and Shugan, S. M. (1983) Managing Channel Profits. *Marketing Science* 2 (3), 239-272.
- 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. (2008) Current Issues in Keyword Auctions. Available on SSRN <http://ssrn.com/abstract=1008496>.
- Liu, D. and Viswanathan, S. (2010) Information Asymmetry and Payment Schemes in Online Advertising. Available at SSRN: <http://ssrn.com/abstract=1698524>.
- 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 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.
- Nisan, N., J. Bayer, D. Chandra, T. Franji, R. Gardner, Y. Matias, N. Rhodes, M. Seltzer, D. Tom, H. Varian and D. Zigmond (2009). Google's auction for TV ads. *Automata, Languages and Programming*. Lecture Notes in Computer Science, Volume 5556/2009, 309-327.
- Riley, J. G., Samuelson, W. F. (1981) Optimal Auctions. *American Economic Review* 71 (3) 381-392.
- Spengler, J. (1950) Vertical Integration and Antitrust Policy. *Journal of Political Economy* 58, 347-352.
- 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.
- Yao, S. and Mela. C. F. (2009) Sponsored Search Auctions: Research Opportunities in Marketing (March 12, 2009). Available on SSRN: <http://ssrn.com/abstract=1367402>.
- 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 3) $b^A(q, u) = W(p^A(q, u), q) = V(p^A(q, u), q)/U(p^A(q, u), q)$ 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

Parts 1 and 2. Assume that every other bidder bids $\beta^E(y)$ and that (as I will show) $[\beta^E]'(y) > 0$. An advertiser of type q will choose bid $b^E(q)$ and price $p^E(q)$ that maximize:

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

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

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

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

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

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 $[b^E]'(q) > 0$, 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):

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

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 3

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:

$$\frac{\partial \Pi^A}{\partial p^A(q, u)} = V_1(p^A(q, u), q) \Psi(ub^A(q, u)) - U_1(p^A(q, u), q) Z^A(b^A(q, u), u) \quad (9)$$

Assumption A1 implies that:

$$\begin{aligned} V_1(p, q) &> 0 \text{ for all } p < p^*(q) \\ V_1(p, q) &= 0 \text{ for } p = p^*(q) \\ V_1(p, q) &< 0 \text{ for all } p > p^*(q) \end{aligned} \quad (10)$$

- If $U_1(p, q) < 0$ for all p then (9) and (10) 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 (9) and (10) 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 (9) and (10) 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:

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

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:

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

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:

$$V_1(p^A(q), q)\Psi(V(p^A(q), q)) - U_1(p^A(q), q)Z^A(W(p^A(q), q), U(p^A(q), q)) = 0 \quad (11)$$

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 $\frac{\partial}{\partial q} [V(p^A(q), q)] \geq 0$, 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 (11) we obtain:

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

where

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

Proof of Proposition 4

The proof makes use of the fact that, if $\frac{\partial}{\partial q} [V(p^A(q), q)] \geq 0$, 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 3, Part 3).

Part 1. Substituting $u = U(p^A(q), q)$ and $b^A(q) = W(p^A(q), q)$ into (3) I obtain:

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

Part 2. Substituting $u = U(p^A(q), q)$ and $b^A(q) = W(p^A(q), q)$ into (4) I obtain:

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

Proof of Proposition 5

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:

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

Solving we obtain:

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

which in turn implies that $p^A(q) = p^*(q)$ and, by Proposition 4, that the mechanism is allocation and revenue equivalent to a PPE mechanism with endogenous product prices. Part 3 of the Proposition is a direct consequence of the result of Section 5.1.

Proof of Proposition 6

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

$$\begin{aligned}
\Omega_H(q, u) &= -f_H(q) + (-g_H(q) + V(p_H, q)) \Psi(ub_H(q)) - U(p_H(q), q) Z(u, b_H(q)) \\
&\quad + \delta \Psi(ub_H(q)) [U(p_H(q), q) \Omega_H(q, u) + (1 - U(p_H(q), q)) \Omega_L(q, u)] + \delta (1 - \Psi(ub_H(q))) \Omega_H(q, u) \\
\Omega_L(q, u) &= -f_L(q) + (-g_L(q) + V(p_L, q)) \Psi(ub_L(q)) - U(p_L(q), q) Z(u, b_L(q)) \\
&\quad + \delta \Psi(ub_L(q)) [U(p_L(q), q) \Omega_H(q, u) + (1 - U(p_L(q), q)) \Omega_L(q, u)] + \delta (1 - \Psi(ub_L(q))) \Omega_L(q, u)
\end{aligned} \tag{12}$$

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:

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

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:

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

and, in turn, imply the following:

$$\begin{aligned}
g_H(q) &= -\delta(1 - U(p^*(q), q))(\Omega_H(q, u) - \Omega_L(q, u)) \\
g_L(q) &= \delta U(p^*(q), q)(\Omega_H(q, u) - \Omega_L(q, u)) \\
Z(u, b_H(q)) &= \delta\Psi(ub_H(q))(\Omega_H(q, u) - \Omega_L(q, u)) \\
\Omega_H(q, u) &= \frac{V(p^*(q), q)\Psi(ub_H(q)) - U(p^*(q), q)Z(u, b_H(q))}{1 - \delta}
\end{aligned} \tag{13}$$

Substituting (12) into (13) and solving for $f_H(q), f_L(q), g_H(q), g_L(q)$ we obtain:

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

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 3) 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:

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

Furthermore, under the above assumptions, from Proposition 4 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. Part 3 of the Proposition is a direct consequence of the result of Section 5.1.

Proof of Proposition 7

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

Proof of Proposition 8

Advertiser i 's profit functions (inclusive of advertising expenses) under PPE and PPS are, respectively:

$$\pi_i^E(p_i, p_{-i}) = D^i(p_i, p_{-i})(p_i - c_i) - f$$

$$\pi_i^S(p_i, p_{-i}) = D^i(p_i, p_{-i})(p_i - c_i - w)$$

Since equilibrium price p_i is unaffected by the fixed fee f , understanding how product price changes when advertiser i shifts from PPE to PPS is equivalent to understanding how the price p_i^A that maximizes:

$$\pi_i(p_i, p_{-i}(p_i, \cdot), w) = D^i(p_i, p_{-i}(p_i, \cdot))(p_i - c_i - w) - f \quad (14)$$

changes as w grows from zero to a positive value. If $\partial p_i^A / \partial w > 0$ for all positive w then I can assert that the switch from PPE to PPS leads to higher product prices. From monotone comparative statics theory (Milgrom and Shannon 1994) a sufficient condition for $\partial p_i^A / \partial w > 0$ is $\partial^2 \pi_i(p_i, p_{-i}(p_i, \cdot), w) / \partial p_i \partial w > 0$. Differentiating (14) I obtain:

$$\frac{\partial^2 \pi_i(p_i, p_{-i}(p_i, \cdot), w)}{\partial p_i \partial w} = -D_1^i(p_i, p_{-i}(p_i, \cdot)) - \sum_{k \neq i} D_k^i(p_i, p_k(p_i, \cdot)) \frac{\partial p_k(p_i, \cdot)}{\partial p_i}$$

where $D_k^i(p_i, p_k(p_i, \cdot))$ denotes the cross-elasticity of i 's demand with respect to k 's price. Since it is $-D_1^i(p_i, p_{-i}(p_i, \cdot)) > 0$, the above expression is positive as long as the magnitudes of the cross-elasticities $D_k^i(p_i, p_k(p_i, \cdot))$ are sufficiently small.