Information Asymmetry And Payment Schemes in Online Advertising

De Liu and Siva Viswanathan¹

Abstract

Pay-for-performance advertising schemes such as pay-per-click (PPC) and pay-per-sale (PPS) have grown in popularity with the recent advances in digital technologies for targeting advertising and measuring outcomes. However, several publishers still continue to use well-established pay-per-impression (PPI) payment schemes. Given the choice of multiple payment schemes - PPI, PPC, and PPS - our study examines the role of information asymmetry between advertisers and publishers in determining the optimal choices for publishers. We highlight the role of payment schemes as a means of leveraging private information available to publishers and advertisers. Our study identifies the conditions under which different types of publishers finds it optimal to offer PPI, and the conditions under which pay-for-performance schemes are optimal. The strategy of offering multiple payment schemes simultaneously by a publisher is also analyzed. Our results provide insights into a number of commonly observed publisher strategies. We discuss the implications of our findings for advertisers, publishers and technology providers.

Keywords: Online Advertising, Pay-for-Performance, Information Asymmetry, Adverse Selection

¹Corresponding Author: De Liu is an Associate Professor at the Gatton College of Business and Economics, University of Kentucky, Lexington. Contact: de.liu@uky.edu. Siva Viswanathan is an Associate Professor at the Robert H Smith School of Business, University of Maryland, College Park. Contact: sviswana@rhsmith.umd.edu.
1. Introduction and motivation

The growth of digitization has led to the emergence of new business models and pricing mechanisms in a wide range of sectors including financial services, consumer retailing, and advertising among others. The advertising sector, in particular, has undergone significant changes in terms of both the technologies used, as well as the payment mechanisms. As noted by Esteban et al 2006, one the key elements of a firm’s marketing mix that have undergone significant technological changes is advertising. Digital technologies have significantly improved the targetability (the ability to target individual consumers) as well as the measurability (the ability to measure the outcomes) of advertising. This has led to a number of new payment schemes that try to cater to the differing needs of publishers as well as advertisers. The three most commonly used payment schemes are (i) Pay-per-impression (PPI), (ii) Pay-per-Click (PPC), and (iii) Pay-per-Sale (PPS). In the case of PPI, advertisers are charged for each impression or exposure, and in the case of PPC the advertiser pays only when a user clicks on the link. PPS, requires that the advertisers pay only when a sale, originating from the advertisement, is consummated. While traditional media is largely dominated by PPI, the growth of sponsored search advertising (for e.g., Google, Yahoo!, and Amazon) has led to the popularity of PPC as well as PPS, driven largely by the ability to measure outcomes at a granular level. As we move from PPI to PPS, the payments (i.e., the advertising expenses) are more closely tied to the outcomes (i.e., the advertiser’s sales and revenue). It is not surprising then that, PPS is often considered the “holy grail of advertising”\(^2\). However, the predictions that these new payment mechanisms (PPC and PPS, in particular) would supplant the more common PPI mechanism have not materialized. On the contrary, these different payment mechanisms coexist, with several publishers choosing PPI over PPC and PPS. Different payment schemes arise because as technology progresses more metrics become available and practical. The significance of different payment schemes lies in that they allow the payment transfer to be more or less sensitive to one party’s information about the other party. Different payment schemes also reveal different amount of

information about each party. Given the choice of these three payment schemes, this paper examines the optimal choice of payment scheme for a publisher. More specifically, our study seeks to understand how the uncertainty faced by advertisers as well as publishers affect their outcomes.

Our main findings are as follows. We find that the choice of a pay-for-performance payment scheme can serve as an effective differentiation strategy for a publisher. Our results show that while a low-quality publisher generally benefits from choosing the less transparent payment scheme (for e.g., PPI compared to PPS), pay-for-performance schemes (PPS or PPC) are more likely to be offered by a high-quality publisher. However, pay-for-performance schemes typically result in allocational inefficiencies, and in some cases may tip the balance in favor of PPI even for a high-quality publisher. A high-quality publisher has to balance the losses from pooling with low-quality publishers in a PPI scheme with losses from allocational inefficiencies in a PPS/PPC payment scheme. Pay-for-performance (PPS/PPC) is optimal in the presence of a larger number of low-quality publishers as high-quality publishers seek to avoid the losses that arise from pooling with low-quality publishers. As technologies for estimating advertiser performance improve, high-quality publishers are more likely to adopt pay-for-performance payment models and separate from low-types. However, as the market matures and publisher types are revealed, PPI is more likely. Our analyses also show that PPS/PPC might not be the best choice for a publisher when there is significant variation in the quality of the advertisers; PPI becomes optimal for a publisher as the potential allocational inefficiencies from PPS/PPC increases with dispersion in advertiser quality.

A few online advertising papers have compared different pay-for-performance models with a focus on their role in reducing the risk for advertisers and publishers. Hu (2004, 2010) studies advertising payment mechanisms in a principal-agent setting with moral hazard - along the lines of earlier work by Holmstrom (1979), Holmstrom and Milgrom (1987, 1991), Lal and Staelin (1985) - where each party’s efforts are private information. Hu (2004, 2010) finds that an online advertising contract that includes appropriate performance-based elements gives the publisher and the advertiser proper incentives to make their efforts, and helps align
the interests of the publisher and the advertiser. In contrast, we focus on a principal-agent characterized by adverse selection, where the publisher’s type and the advertisers’ types are private information.

The growth of online sponsored search markets has given rise to a nascent but growing stream of research examining various aspects of these markets. Edelman et al (2007) and Varian (2007) analyze the equilibria in sponsored search auctions under the “generalized second-price” auction rules. Several papers examine various aspects of advertising auction design (Feng, 2007; Weber and Zheng, 2007; Liu et al, 2010; Chen et al, 2009, Wilbur and Zhu 2009). There is also a growing body of empirical research on online advertising, including the examination of bidding patterns (Zhang and Feng, 2005, Edelman and Ostrovsky, 2007) and bidding strategies (Ghose and Yang, 2009; Yang and Ghose, 2009, Animesh et al, 2010). Most of these papers take the pay-per-click payment scheme for granted.

In a recent study analyzing the choice of payment schemes, Zhu and Wilbur (2010) examine a “hybrid” advertising auction where brand and direct advertisers choose between PPI and PPC bids. They demonstrate that the publisher suffers a revenue loss as a result of not knowing the advertiser’s type. Our study extends the analyses by Zhu and Wilbur (2010) in the sense that we study the broader issue of whether the “hybrid” scheme is an optimal choice for publishers among all the payment schemes.

Our work is closest in spirit to Sundararajan (2003) who studies optimal pricing of advertising services in a principal-agent framework where the pricing contracts can be based on impressions as well as performance. Sundararajan (2003) shows that with performance uncertainty and risk-averse advertisers, performance-based pricing is always profit-improving for publishers. He also observes that publishers cannot screen out lower-quality advertisers with performance-based pricing but they can use it to credibly signal their superior marketing effectiveness. Our work builds upon and extends Sundararajan (2003) in a number of ways. First, while Sundararajan (2003) examines the role of information asymmetry regarding publisher’s effectiveness and advertiser’s quality separately, our study simultaneously takes into account the two-sided information asymmetry. Second, rather than focusing on
contracting with a single advertiser, we model the competition between multiple advertisers in an auction framework. This framework highlights the impact of payment schemes on resource allocation efficiency and explains why performance-for-performance may not always be the optimal choice for publishers in our model settings. Finally, our paper provide insights on the case where the publisher simultaneously offering two payment schemes for advertisers to choose.

Our research is also related to the literature on usage-based pricing where prices are tied to measurable outcomes. There is a vast literature on usage-based pricing and non-linear pricing in a variety of contexts such as networks and service industries - with a large focus on the role of congestion externalities, capacity constraints, and delays (Mendelson 1985; Dewan and Mendelson 1990; Gupta et al 1997; Masuda and Whang 2006). Recent work on non-linear pricing (see Sundararajan 2004, Choudhary 2008, Jain and Kannan, 2002, Bashyam 2000) focuses on the unique properties of information goods. While the pay-for-performance schemes for advertising and usage-based pricing schemes for information goods are similar in some respects, pay-for-performance schemes in advertising differ from usage-based pricing in several ways. While usage-based pricing for information goods is usually discussed in a context without any capacity or provisioning constraints, capacity constraints play a pivotal role in pay-for-performance models—particularly in the context of advertising. Consequently, the problem of buyer selection, which is rarely the focus of usage-based pricing models, becomes paramount in the case of pay-for-performance schemes. Secondly, while the value of the service depends on some ex-post outcomes (for instance, actual usage or clicks), in the context of usage-based pricing the outcome (actual usage) is a property of the buyer. However, in the context of performance-based advertising payment schemes, the outcomes (clickthrough rates or sales) depend on both the advertiser (buyer) as well as the publisher (seller). Another important feature of pricing models of information goods as well as pricing in network and service industries is the role of consumer characteristics and preferences in determining the optimal pricing strategies. Heterogeneity in consumer valuations, consumer expertise, among others, makes price-discrimination strategies profitable for firms.
- with usage based pricing and/or fixed-rate tariffs serving as effective price-discrimination mechanisms.

Finally, our work is also related to research on the role of pay-for-performance pricing schemes as risk sharing mechanisms. Gallini and Wright (1990) study the design of the optimal licensing contract for an innovator, given the potential for opportunism by both parties- the licensor as well as the licensee. Gallini and Wright (1990) find that while fixed fee contracts are more likely for low value innovations, output-related royalties are optimal in the case of high-value innovations. In addition, their study also examines the role of exclusivity on the choice of the optimal contract. Chemmanur et al (2010) find that leasing with metering provisions (corresponding to output-based pricing) emerges as an equilibrium solution to a double-sided information asymmetry problem between a capital-goods manufacturer who has private information about quality, and an entrepreneur who has private information about maintenance costs and usage intensities. They also find that in equilibrium, leasing co-exists with a sales contract (corresponds to a flat-rate pricing) in some cases, while it is the sole financing contract in other cases. Our model is in the spirit of these models of two-sided information asymmetry; however, the issues that we focus on are distinct and the mechanisms are specific to existing markets for online advertising.

The rest of the paper is organized as follows. Section 2 outlines the modeling primitives and Section 3 details the equilibrium analyses. Section 4 analyzes the case where the publisher has a choice of multiple payment schemes, and Section 5 discusses extensions to the basic model. Section 6 discusses the implications of the findings and concludes.

2. Modeling Primitives

A risk-neutral publisher has a single advertising slot to sell to one of the \( n \) risk neutral advertisers. The publisher can choose between two payment schemes, PPI and PPS. Our results can be easily re-interpreted for the choice between PPI and PPC or between PPC and PPS. Extensions to multiple slots and to PPC are discussed in section 5.
Performance. We start by assuming that the probability of a sale at any given impression is jointly determined by advertiser quality \((a)\) and publisher quality \((b)\):

\[
\text{Probability of sale} = a \times b
\]

A publisher’s quality \((b)\) is interpreted as the publisher’s ability to deliver advertisements to the right audience. By this interpretation, a high-quality publisher has a better ability to match advertisers with its subscriber base, leading to a higher probability of sales for all advertisements\(^3\). An advertiser’s quality \((a)\) is interpreted as the attractiveness/persuasiveness of the advertiser’s message. A higher quality advertiser has higher-quality product and better-crafted message, leading to higher probability of sale among its intended audience.

For simplicity, we assume that there are two levels of advertiser quality, \(a_l\) and \(a_h\) \((a_l < a_h)\) and use \(x \in \{l, h\}\) to denote an advertiser’s (quality) type. Similarly, there are two levels of publisher quality, \(b_l\) and \(b_h\) \((b_l < b_h)\) and we use \(y \in \{l, h\}\) to denote a publisher’s (quality) type. The probability of being a high-quality advertiser is \(\alpha\) and the probability of being a high-quality publisher is \(\beta\). The quality types \(x\) and \(y\) are private information and independently distributed (we discuss the interdependent case in section 5) but \(\alpha\) and \(\beta\) are common knowledge.

Because a publisher may learn about an advertiser’s quality through repeated interactions or historical performance data purchased from a third party, the publisher receives a signal \(\hat{x}\) on an advertiser’s quality. By the same argument, an advertiser receives a signal \(\hat{y}\) on a publisher’s quality (all advertisers receive the same signal about a publisher). Once again, for simplicity, we assume that there are two types of signals, \(\hat{x}, \hat{y} \in \{l, h\}\). The probability of misclassifying an advertiser (publisher) is \(\gamma_A\) \((\gamma_P)\), that is, \(P(\hat{x} = l|x = h) = P(\hat{x} = h|x = l) = \gamma_A\) and \(P(\hat{y} = l|y = h) = P(\hat{y} = h|y = l) = \gamma_P\), for some \(0 \leq \gamma_A, \gamma_P \leq 0.5\). \(\gamma_A\) and \(\gamma_P\) are interpreted as misclassification rates or imprecision of quality signals. \(\gamma_A\) \((\gamma_P)\) can be affected by the quality of the performance metric, by the

\(^3\)The publisher can also be interpreted as an advertising network (for instance, Microsoft Media Network, AOL Advertising, 24/7 Real Media, admob, etc.) that serves as an intermediary between publishers and advertisers. The performance of an advertising network can be interpreted along the same lines as that of the publisher: its ability to match advertisers with the right publishers.
amount of historical data available, and by the estimation technique. For example, when there is a lot of heterogeneity among audience, we expect more volatile performance outcomes and poorer performance metric. We also expect precision of quality signals to be low for new advertisers/publishers and for parties that have poor estimating technologies.

**Valuation per sale.** An advertiser’s valuation per sale (valuation for short) is $v \in [0, 1]$. Advertisers’ valuations are private information, independent, and identically distributed according to the distribution $F(v)$ with a strictly positive and differentiable density function $f(v)$. We assume that the increasing hazard rate (IHR) condition holds, i.e., $f(v) / [1 - F(v)]$ is increasing.

**Payment schemes.** Payment schemes affect not only the way of advertisers are charged but also the way advertising slots are auctioned. Under the PPI scheme, advertisers submit sealed bids on their willingness-to-pay per impression. The advertising slot will be assigned to the highest bidder, who pays the second highest bid. Under the PPS scheme, advertisers submits sealed bid on their willingness-to-pay per sale. Because a publisher receives signals about advertiser quality, the publisher runs a scoring auction using the signals on advertiser quality. Specifically,

**Assumption 1.** Under the PPS auction, each advertiser is assigned a score $z$ that is a product of her bid $t$ and her expected quality, i.e.,

\begin{equation}
    \text{score } z = t \times w, \text{ where } w = E[a|x]
\end{equation}

The advertiser with the highest score is the winner and pays the lowest price to stay above the second highest score, $z^{(2)}$. In other words, the winner with weighting factor $w$ pays $z^{(2)}/w$ per sale. This style of auctions are first used by search engines, such as Overture and Google to sell their sponsored link advertisements (Edelman and Ostrovsky, 2007). Over time, advertising auctions have achieved broader applications as search engines become brokers for many other types of Internet advertising, such as display advertising, mobile advertising, and advertising on interactive TV.

\footnote{Our results will not change if the upper bound is an arbitrary positive number.}
The game timeline. The game proceeds as follows. At first, Nature determines publisher and advertiser quality. All parties learn their own quality. The publisher then chooses a payment scheme and associated auction format. The two parties then each learn a signal about the other party's type. In the PPS auction, the publisher assigns a weighting factor for each advertiser based on their quality signals. Each advertiser, after learning the payment scheme, her weighting factor, and the signal about the publisher's type, decides how much to bid at the same time as other advertisers make their decisions. The publisher chooses the winner according to the announced auction rule and the winner pays accordingly.

We assume that the publisher offers a single payment scheme. An extension where the publisher offers multiple payment schemes for advertisers to choose from is discussed in Section 5. Let $\rho^m_y$ denote the probability that a $y$-type publisher chooses payment scheme $m \in \{I, S\}$. We say a publisher plays a pure strategy when he chooses a payment scheme with probability 1. Otherwise, we denote his play as a mixed strategy. When both publisher types play pure strategies, we may use a payment scheme vector to represent their strategies. For example, $(I, I)$ represents that both types choose PPI and $(I, S)$ represents that low-quality publisher chooses PPI while high-quality chooses PPS.

3 Equilibrium Analysis

To analyze this game, we use backward induction. We first characterize the advertisers' bidding equilibrium under each circumstance. We then derive the publisher's expected revenue under each payment scheme given the overall publisher strategy profile $\{\rho^m_y\}$. Finally, we find the publisher strategy profiles that are in equilibrium.

3.1 Quality Estimates
<table>
<thead>
<tr>
<th>Notation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v$</td>
<td>advertiser’s valuation per sale</td>
</tr>
<tr>
<td>$a \in {a_l, a_h}$</td>
<td>quality of an advertiser</td>
</tr>
<tr>
<td>$b \in {b_l, b_h}$</td>
<td>quality of a publisher</td>
</tr>
<tr>
<td>$x, y \in {l, h}$</td>
<td>advertiser’s (quality) type and publisher’s (quality) type</td>
</tr>
<tr>
<td>$\alpha, \beta$</td>
<td>probability of a high-quality advertiser and that of a high-quality publisher</td>
</tr>
<tr>
<td>$\hat{x}, \hat{y} \in {l, h}$</td>
<td>signal of an advertiser’s quality and that of a publisher’s quality</td>
</tr>
<tr>
<td>$\hat{a}, \hat{b}$</td>
<td>estimated quality of an advertiser and a publisher</td>
</tr>
<tr>
<td>$\hat{b}$</td>
<td>expected advertiser’s estimates of a publisher’s quality</td>
</tr>
<tr>
<td>$\gamma_A, \gamma_P$</td>
<td>the probability of misclassifying an advertiser and a publisher</td>
</tr>
<tr>
<td>$w_{\hat{x}}$</td>
<td>the weighting factor for advertiser with $\hat{x}$-type signal</td>
</tr>
<tr>
<td>$m \in {I, S}$</td>
<td>a publisher’s payment scheme choice (PPI or PPS)</td>
</tr>
<tr>
<td>$\rho^m_y$</td>
<td>probability for a $y$-type publisher to choose payment scheme $m$.</td>
</tr>
<tr>
<td>$\phi^m(\cdot)$</td>
<td>the equilibrium probability of winning under payment scheme $m$.</td>
</tr>
<tr>
<td>$\pi^m$</td>
<td>publisher’s expected revenue under payment scheme $m$</td>
</tr>
</tbody>
</table>

**Table 1. Notations**

Under the PPS auction, the publisher needs to estimate advertisers’ expected quality and use it as weighting factor. From the publisher’s point of view, an advertiser receives an $\hat{x}$-type signal with probability $P(\hat{x}) = \sum_x P(x) P(\hat{x}|x)$. The expected quality of an advertiser with signal $\hat{x}$ is given by

$$\hat{a}_{\hat{x}} \equiv E [a|\hat{x}] = \frac{\sum_x P(x) P(\hat{x}|x) a_x}{\sum_x P(x) P(\hat{x}|x)}$$

We can easily verify that (see Online Appendix)

$$a_l \leq \hat{a}_l \leq \hat{a}_h \leq a_h$$

Similarly, advertisers must estimate a publisher’s quality under the PPI auction. Recall that advertisers observe a signal about the publisher’s type. Furthermore, when a high-quality publisher and a low-quality publisher chooses a payment scheme with different probabilities, the publisher’s payment scheme choice provides additional information about his type. A rational advertiser will anticipate the publisher’s equilibrium strategy and update her belief about the publisher’s type based on his payment scheme choice and quality signal.
We denote $\mu(\cdot|\hat{y}, m)$ as the advertisers’ updated belief about a publisher’s type given his quality signal $\hat{y}$ and payment scheme $m$. The probability of a publisher with signal $\hat{y}$ choosing payment scheme $m$ is $P(m|\hat{y}) \equiv \sum_y P(y) \rho^m_y P(\hat{y}|y)$. Provided that a payment scheme is chosen with a positive probability in equilibrium, the advertiser’s updated belief can be calculated using the Bayes rule as $\mu(y|\hat{y}, m) = \frac{P(y)\rho^m_y P(\hat{y}|y)}{\sum_y P(y)\rho^m_y P(\hat{y}|y)}$. When a payment scheme $m$ is never chosen, the Bayes rule does not apply and advertisers can form arbitrary beliefs as long as $\sum_y \mu(y|\hat{y}, m) = 1$. We term such beliefs off-equilibrium beliefs.

From an advertiser’s perspective, the expected quality of a publisher who has a quality signal $\hat{y}$ and chooses $m$, $\hat{b}^m_{\hat{y}}$, is given by

$$\hat{b}^m_{\hat{y}} \equiv E[b|\hat{y}, m] = \sum_y \mu(y|\hat{y}, m) b_y$$

An advertiser’s estimate of a publisher’s quality changes with the realization of the publisher’s quality signal $\hat{y}$. A publisher chooses his payment scheme anticipating all the possible realization of his quality signal and advertiser’s estimates. For a $y$-type publisher who chooses a payment scheme $m$, the expected advertiser’s estimates of his quality, $\hat{b}^m_y$, is given by

$$\hat{b}^m_y \equiv E[\hat{b}] = \sum_{\hat{y}} P(\hat{y}|y) \hat{b}^m_{\hat{y}}$$

It is straightforward to verify that for any given publisher strategy, off-equilibrium belief (see Online Appendix), and payment scheme $m$,

$$b_l \leq \hat{b}^m_y \leq \hat{b}^m_{\hat{y}} \leq b_h$$

It shall be noted that both $\hat{b}^m_{\hat{y}}$ and $\hat{b}^m_y$ are affected by the composition of low- and high-quality publishers at payment scheme $m$. So they are functions of publisher’s strategy $\{\rho^m_y\}$, advertisers’ off-equilibrium belief, and payment scheme $m$. For simplicity, we will omit the superscript and/or subscript when the context is clear. The same applies to $\hat{a}_{\hat{z}}$.

3.2 The PPI Auction
Under the PPI auction, an advertiser’s valuation per impression is $v \equiv va\hat{b}$, where $\hat{b}$ is the estimated publisher quality as defined in (5). Advertisers have identical estimation of the publisher’s quality but different valuation $v$ and quality $a$. The PPI auction can be viewed as a seal-bid second-price auction with $v$ distributed according to

$$F_v(v) = P\left(v'a\hat{b} < va\hat{b}\right) = E_{a'}[F(va/a')]$$

Throughout the paper, we use notation $E_X[\cdot]$ to represent an expectation with respect to random variable $X$. In this case, $E_{a'}[F(va/a')] = \alpha F(va/a_h) + (1 - \alpha) F(va/a_l)$. Denote $T(v)$ as the equilibrium bidding function under the PPI auction. As in the standard second-price auctions, we have

**Lemma 1.** $T(v)$ is strictly increasing in $v$.

Because $T(v)$ is strictly increasing, an advertiser $v$ outbids another advertiser $v'$ in equilibrium if and only if $v > v'$. Hence the equilibrium winning probability of an advertiser with valuation $v$ and quality $a$ is

$$\phi^I(v,a) = [F_v(v)]^{n-1} = \{E_{a'}[F(va/a')]\}^{n-1}$$

As with standard auctions, an advertiser’s equilibrium payoff and the publisher’s equilibrium revenue can be determined once the equilibrium winning probability is known. We can show that the publisher’s equilibrium expected revenue from all advertisers is (see Appendix)

$$\pi^I = n\hat{b}_yE_{a}\left[ a \int_{0}^{1} \phi^I(v,a) J(v) f(v) dv \right] \equiv \hat{b}_y\pi^I_{base}$$

where $J(v) \equiv v - \frac{1-F(v)}{f(v)}$ and $\hat{b}_y$ (the expected advertisers’ estimates of a $y$-type publisher’s quality) is given by (6). We term $\pi^I_{base} \equiv nE_{a}\left[ a \int_{0}^{1} \phi^I(v,a) J(v) f(v) dv \right]$ as the base revenue of the PPI auction.

**3.3 The PPS Auction**
Denote $T(v, w)$ as the equilibrium bidding function under the PPS auction. The following Lemma helps us establish an equilibrium ordering of all advertisers.

**Lemma 2.** The equilibrium bid function $T(v, w)$ is strictly increasing in $v$. Furthermore, the advertiser $(v, w)$ and $(v', w')$ tie in equilibrium (i.e., $T(v, w) = T(v', w')$) if

$$vw = v'w'$$

Observing (33), we find that two advertisers tie when they have the same “truthful” score, i.e., the score an advertiser will get when she bids her true willingness-to-pay. This finding was first established for weighted unit-price auctions by Liu and Chen (2006). This Lemma implies that the equilibrium bids are proportional to the underlying true unit-valuation, a relationship holds even though advertisers may not bid truthfully. The intuition behind this result is that when (33) holds and two advertisers bid the same score, their utility functions differ only by a scaling factor, which implies that they must tie in their optimal bidding.

By Lemma (33), we can write the equilibrium winning probability $\phi^S(v, \hat{a})$ for an advertiser with valuation-per-sale $v$ and weighting factor $\hat{a}$ ($\hat{a} \in \{\hat{a}_l, \hat{a}_h\}$) as:

$$\phi^S(v, \hat{a}) = \{E_{\hat{a}'} [F(v\hat{a}/\hat{a}')]\}^{n-1}$$

This winning probability is similar to the winning probability under PPI (9) except that under PPS advertisers are compared in terms of estimated quality $\hat{a}$ rather than true quality $a$.

Once we have the advertiser’s equilibrium winning probability, we derive the publisher’s total expected revenue as (see Appendix)

$$\pi^S = nb_yE_{\hat{a}} \left[ \hat{a} \int_0^1 \phi^S(v, \hat{a}) J(v) f(v) dv \right] \equiv b_y\pi^S_{base}$$

---

5 Please note that an advertiser’s type $x$ does not affect her equilibrium bidding because her quality enters her utility as a scaling factor which does not change the solution to the advertiser’s optimal bidding problem.  
6 Advertisers will bid truthfully when there is a single slot but they will not bid truthfully when there are more than one slot.
where $J(v)$ is as defined before and $\pi^S_{\text{base}} = nE_{\hat{a}} \left[ \hat{a} \int_0^1 \phi^S(v, \hat{a}) J(v) f(v) dv \right]$ denotes the base revenue under the PPS auction.

It is clear from (10) and (13) that the publisher revenue across payment schemes consists of a publisher-quality component ($b_y$ and $\hat{b}_y$) and a base-revenue component ($\pi^I_{\text{base}}$ and $\pi^S_{\text{base}}$). The former captures the impact of the publisher’s quality on the expected revenue. The later captures the difference between PPI and PPS auctions in terms of their ability to capture revenues. The main difference between PPI and PPS auctions lies in their different allocation rules: the former ranks advertisers by valuation-per-sale times true quality (Lemma 1) whereas the PPS auction ranks advertisers by valuation-per-sale times estimated quality (Lemma 2).

Because $b_l \leq \hat{b}_l$ (7), a low-quality publisher may receive a quality markup under PPI. In contrast, because $\hat{b}_h \leq b_h$, a low-quality publisher may receive a quality markdown under PPI. The sizes of the quality markup and markdown are affected by the market composition and publisher strategy.

3.4 Comparing PPI and PPS Auctions

To gain more insights into the PPI and PPS auctions, we compare them in terms of their allocation efficiency and profitability. The (ex ante) allocation efficiency $W$ is measured by the total expected valuation realized by an auction.

**Proposition 1.** The PPI auction is more efficient than the PPS auction. However, $\pi^I_{\text{base}}$ and $\pi^S_{\text{base}}$ cannot be unanimously ranked.

Similar results have been obtained by Liu and Chen (2006). Here we only note the intuition behind these results. Because all advertisers have the same estimation of the publisher’s quality, the maximum allocation efficiency is reached when the advertiser with highest $va$ is selected. This is the case in the PPI auction but not in the PPS auction – ranking under PPS is based on $\hat{va}$ instead of $va$. The distortion in efficiency comes in two ways: first, a high-quality (low-quality) advertiser may receive a low-quality (high-quality) signal and thus
a low weighting factor \( \hat{a}_l \) (\( \hat{a}_h \)). Second, the weighting factors \( \hat{a} \) may not be aligned with true quality \( a \), judging from the fact that \( a_l \leq \hat{a}_l \leq \hat{a}_h \leq a_h \).

Misclassifying advertisers has two kinds of impacts on the publisher’s revenue. First, the misclassification has a misallocation effect (e.g., when the winner turns out to be low quality) that may lead to a reduction in revenue. Second, misclassification equalizes advertiser’s weighting factors (note that \( a_l \leq \hat{a}_l \leq \hat{a}_h \leq a_h \)), which produces an incentive effect for high-quality advertisers to compete more aggressively. The incentive effect is more pronounced when high-quality advertisers lack competition under efficient allocation. The following table summarizes the effect of several factors on the relative efficiency (\( W_S/W_I \)) and profitability (\( \pi_S^{\text{base}}/\pi_I^{\text{base}} \)) of the PPS auction.

As shown in table 2, the relative efficiency and profitability of the PPS auction decrease with \( n \) (the number of advertisers) and \( a_h/a_l \) (the quality spread among advertisers). As the number of advertisers increases, both PPI and PPS auctions become more efficient due to the law of large numbers, but the relative efficiency of the PPS auction decreases because the misclassification error does not decrease with \( n \). On the other hand, the incentive effect of the PPS auction diminishes as the competition among advertisers intensifies (as a result of a large \( n \)). So the relative profitability of PPS decreases. As the quality spread among advertisers widens, the misclassification error under PPS becomes more detrimental, leading to lower relative efficiency and profitability. Even though the relative efficiency of the PPS auction decreases with the probability of misclassification \( \gamma_A \), its relative profitability may decrease or increase. This is because as \( \gamma_A \) increases, the incentive effect also increases, which may dominate the negative effect of efficiency loss. Finally, as the proportion of high-quality advertisers \( \alpha \) increases, the relative efficiency first decreases then increases because the misclassification error is less detrimental when advertisers are more homogeneous in

<table>
<thead>
<tr>
<th>PPS auction’s</th>
<th>( n )</th>
<th>( a_h/a_l )</th>
<th>( \gamma_A )</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative efficiency</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>Relative profitability</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Table 2. comparative statics (\(-/-\): decreases with; +/-: is not monotonic in)
quality. Correspondingly, the relative revenue of PPS may increase or decrease depending on relative strength of the incentive effect and the misallocation effect.

3.5 Publisher’s Equilibrium Strategy

By choosing PPS, the publisher reveals his quality but may incur a lower base revenue. A high-quality publisher must trade off two opposite effects. Next we discuss the equilibrium in the publisher’s game. For the publisher’s game, we adopt the Trembling Hand Perfect (THP) equilibrium concept (Kohlberg and Mertens 1986). THP equilibrium is a refinement of the Perfect Bayesian equilibrium to allow some degree of irrationality in the sense that players may play off-equilibrium strategies with negligible probability (resembles a “trembling hand”). THP equilibrium effectively eliminates all weakly-dominated strategies.

Clearly, when \( \pi^I_{base} < \pi^S_{base} \), both publisher types prefer PPS so the only equilibrium is a pooling equilibrium \((S, S)\) (see Online Appendix). We will not focus on the case of \( \pi^I_{base} < \pi^S_{base} \) for two reasons: first, the case \( \pi^I_{base} < \pi^S_{base} \) occurs only in a restrictive condition where the incentive effect caused by misclassification is strong enough to overcome the misallocation effect. Second, with \( \pi^I_{base} < \pi^S_{base} \), it is a dominate strategy for both publisher types to choose PPS. This does not seem to be the case in real world settings. Therefore, we make the following assumption:

**Assumption 2.** \( \pi^I_{base} \geq \pi^S_{base} \)

**Proposition 2.** Under the assumption 2, the publisher’s game has a separating THP equilibrium \((I, S)\) when

\[
(14) \quad b_h\pi^S_{base} \geq b_I\pi^I_{base}
\]

and a pooling THP equilibrium \((I, I)\) when

\[
(15) \quad b_h\pi^S_{base} \leq \hat{b}_h\pi^I_{base}
\]

where \( \hat{b}_h \), defined in (6), is a high-quality publisher’s expected quality estimates when the publisher plays \((I, I)\).
Proof. Under assumption 2, a low-quality publisher gets an expected revenue of $\hat{b}_l \pi_{base}^I$ under PPI and an expected revenue of $b_l \pi_{base}^S$ under PPS. Because $\hat{b}_l \geq b_l$, PPS is a weakly dominated strategy for the low-quality publisher. Because a weakly dominated strategy is never THP, a low-quality publisher always play PPI. So we only consider the remain two strategy profiles $(I, S)$ and $(I, I)$.

For $(I, S)$ to be THP, we only need to show that a high-quality publisher optimally choose PPS. A high-quality publisher’s expected revenue under PPS is $b_h \pi_{base}^S$. By deviating to PPI, he gets $b_l \pi_{base}^I$. The condition (14) ensures that a high-quality publisher is worse off deviating to PPI.

For $(I, I)$ to be THP, we only need to show that a high-quality publisher optimally chooses PPI. A high-quality publisher’s expected revenue under PPI is $\hat{b}_h \pi_{base}^I$. If a high-quality publisher deviates to PPS, her expected revenue is $b_h \pi_{base}^S$. A high-quality publisher would not deviate from PPI if the condition (15) holds.

When both (14) and (15) hold, both $(I, I)$ and $(I, S)$ are THP. But we find that $(I, I)$ Pareto dominates $(I, S)$ in the sense that both publisher types are better off under $(I, I)$ than under $(I, S)$.

**Corollary 1.** When both (14) and (15) hold, the equilibrium $(I, S)$ is Pareto-dominated by equilibrium $(I, I)$.

Proof. Under the pooling equilibrium, the expected revenues for the high-quality and the low-quality publisher are, respectively, $\hat{b}_h \pi_{base}^I$ and $\hat{b}_l \pi_{base}^I$. Under the separating equilibrium, the expected revenues are, $b_h \pi_{base}^S$ and $b_l \pi_{base}^I$. Because of the condition (15) and (7), we have $\hat{b}_h \pi_{base}^I > b_h \pi_{base}^S$ and $\hat{b}_l \pi_{base}^I > b_l \pi_{base}^I$. So the separating equilibrium is Pareto-dominated.

We further assume that a Pareto-dominated equilibrium strategy is not played:

**Assumption 3.** A Pareto-dominated equilibrium is not played.
**Corollary 2.** Under assumptions 2 and 3, the strategy profile \((I, S)\) is an equilibrium when

\[
\hat{b}_h \pi^S_{\text{base}} > \hat{b}_h \pi^I_{\text{base}}
\]

otherwise the strategy profile \((I, I)\) is an equilibrium.

The following corollary summarizes the impact of several factors on a high-quality publisher’s equilibrium choice.

**Corollary 3.** A high-quality publisher tends to separate from a low-quality publisher by choosing PPS when (a) the probability of high-quality publisher \((\beta)\) decreases, (b) \(b_h/b_l\) increases, (c) the probability that a publisher is misclassified \((\gamma_P)\) increases, and (d) \(\pi^I_{\text{base}}/\pi^S_{\text{base}}\) decreases.

*Proof.* the high-quality publisher will choose PPS if and only if the condition (16) holds. The first three results are evident from the fact that \(\hat{b}_h\) decreases in \(\beta\) and \(\gamma_P\) and increases in \(b_l\). The fourth result is also evident from the condition (16). \(\square\)

The result (a) in Corollary 3 implies that in a market where high-quality publishers are rare, they more likely separate themselves by adopting PPS. The second result (b) implies that when there is large quality spread among publishers, a high-quality publisher is more likely to adopt PPS. According to (c), we are more likely to see PPS in a market where the publishers are new and their qualities are not known to the advertisers. Lastly, when the relative profitability of the PPS auction improves, PPS is more likely used. Table 2 illustrates the factors that impact the relative profitability of the PPS auction. For example, when there is strong competition among the advertisers (e.g., when there are many advertisers or the equality spread among advertisers is narrow), PPS is less likely.

Figure 2 illustrates the high-quality publisher’s equilibrium payment scheme as a function of misclassifying rates \(\gamma_A\) and \(\gamma_P\). This figure shows that as the probability of misclassifying publishers \(\gamma_P\) increases, a high-quality publisher is more likely to adopt PPS. As the probability of misclassifying advertisers increases, a high-quality publisher is *generally* more likely...
Figure 2. The high-quality publisher's equilibrium choice ($n = 6, \alpha = 0.7, a_l = 0.1, a_h = 0.2, \beta = 0.7, b_l = 0.1, b_h = 0.12$) to adopt PPI. The boundary between PPI and PPS is not monotonic in $\gamma_A$ because a high $\gamma_A$ also creates a stronger incentive effect which may dominate the misallocation effect.

4 Offering Two Payment Schemes Simultaneously

In this section, we turn to the case where a publisher may offer PPI and PPS simultaneously. We call such payment scheme a portfolio (or PPI/S) scheme and index it with $IS$. We investigate the rationale for the portfolio scheme and its impact on publishers. If the publisher offers PPI/S, the advertisers must choose the bid as well as the payment scheme. The publisher must specify how PPS bids and PPI bids will be compared. We assume that the publisher uses the following scoring rule:

Assumption 4. Under a PPI/S scheme, an advertiser who chooses payment scheme $m$ and bids $t$ is assigned a score of

$$score_z = \begin{cases} 
t, & \text{if } m = I \\
t \times w, & \text{if } m = S \\
\end{cases}$$

The weighting factor $w$ is for PPS advertisers only and may depend on the signal on advertiser's quality. As before, the winner pays a lowest price to stay above the second-highest bidder and the timeline of the game remains the same.
4.1 The Advertiser’s Problem under the PPI/S

We can first examine how advertisers will choose payment schemes and bids.

Lemma 3. Under a portfolio scheme, an advertiser with quality \( a \) and weighting factor \( w \) will choose PPI when \( \hat{a}b > w \), PPS when \( \hat{a}b < w \), and either when \( \hat{a}b = w \).

Proof. Denote \( F_z(\cdot) \) as the CDF of the equilibrium score and \( z^{(2)} \) as the second highest score. For any advertiser with a score \( z \), her winning probability is \( \phi(z) = [F_z(z)]^{n-1} \) and the second highest score is \( p(z) = E[z^{(2)}|z^{(2)} < z] \). Let \( t^I \) and \( t^S \) denote an advertiser’s PPI bid and PPS bid respectively. Let \( w \) denote her PPS weighting factor, her expected utility under PPI and PPS are, respectively,

\[
U^I(t^I, v) = \phi(t^I) [va\hat{a} - p(t^I)] \\
U^S(t^S, v) = \phi(t^S) b[a - p(t^S) / w] = \phi(t^S) [va - p(t^S) \hat{a}b / w]
\]

when \( \hat{a}b = w \), \( U^I(t^I, v) = U^S(t^S, v) \), which means that the advertiser can get the same expected utility under two payment schemes. So the advertiser is indifferent between the two. Now suppose \( \hat{a}b < w \). Let \( t^*_I \) be the advertiser’s optimal bid under PPI. By bidding \( t^*_I / w \) under PPS, her expected utility is higher than \( U^I(t^I, v) \) (noticing that \( \hat{a}b / w < 1 \)). Her optimal payoff under PPS must be no less than \( U^I(t^*_I, v) \) because bidding \( t^*_I / w \) may be sub-optimal for her. So she prefers PPS when \( \hat{a}b < w \). Similarly, we can show that when \( \hat{a}b > w \), she prefers PPI. □

Lemma 3 suggests that advertisers with the same quality \( a \) and the PPS weighting factor \( w \) must choose the same payment scheme, regardless of their valuation. Because high-quality advertisers have higher \( a \), they tend to choose PPI whereas low-quality advertisers tend to choose PPS. In the fact when \( w \in [a\hat{b}, a_h\hat{b}] \), all high-quality advertisers choose PPI and all low-quality advertisers choose PPI.

Next result extends Lemma 7 to the portfolio scheme. Like before, the equilibrium bid must strictly increase in valuation \( v \) and advertisers with the same “truthful score” tie in equilibrium.
Lemma 4. Under PPI/S, (a) the equilibrium bid strictly increases in $v$. (b) A PPI advertiser $(v,a)$ and a PPS advertiser $(v',w)$ tie in equilibrium if

$$v a \hat{b} = v'w.$$  

Proof. The proof for (a) is analogous to that for Lemma 1 and Lemma 2 (notice that advertisers with the same $(a,w)$ choose the same payment scheme). The proof for (b) is also similar. Notice that (18) can be rewritten as $U_S(t^S,v') = \phi(t^S w) [v'w - p(t^S w)] / w$. When $v a \hat{b} = v'w$ the two utility functions $U_I(t^I,v)$ and $U_S(t^S,w,v')$ differ only by a constant scaling factor. The implies that the two functions must reach optimality at the same time, i.e., $w T^S(v',w) = T^I(v,a)$. In other words, the two advertisers tie in equilibrium. \qed

4.2 The Publisher’s Problem

Because the purpose of the portfolio scheme is to separate high- and low-quality advertisers, we are more interested in the non-degenerate case where the two types of advertisers are separated. A range of weighting factors can achieve this goal but only one is efficient.

Lemma 5. The efficient weighting factor that perfectly separates the high- and low-quality advertisers is $w = a_l \hat{b}$.

Proof. According to Lemma 3, if $w$ is higher than $a_h \hat{b}$, both advertiser types will choose PPS. If $w$ is lower than $a_l \hat{b}$, both types will choose PPI. For any given $w \in [a_l \hat{b}, a_h \hat{b}]$, high-quality advertisers choose PPI and low-quality ones choose PPS. By Lemma 4, a low-quality advertiser with valuation $v_l$ and a high-quality advertiser with valuation $v_h$ will tie if $v_h a_h \hat{b} = w v_l$. For the auction to be efficient, the two advertisers must have an equal valuation per impression, i.e., $v_h a_h \hat{b} = v_l a_l \hat{b}$. This implies that $w = a_l \hat{b}$. \qed

Based on Lemma 5, we assume the following weighting factor being used.

Assumption 5. $w = a_l \hat{b}$. 
With the efficient weighting factor, a publisher ignores the quality signals $\hat{x}$ completely. This makes sense because with the efficient weighting factor, all PPS advertisers are expected to be low-quality.

We next consider the publisher’s game. Now each type of publisher has three choices, $m \in \{I, S, IS\}$. We are interested in a few questions. How does the addition of the portfolio scheme affect publisher’s equilibrium? What’s the impact on the publisher’s revenue?

We first consider an advertiser’s equilibrium winning probability. According to Lemma 4, to win the PPI/S auction, an advertiser must have the highest “truthful score” among all advertisers. The “truthful score” of a high-quality advertiser is $a_h \hat{b}v$ and that of a low-quality advertiser $a_l \hat{b}v$. So the equilibrium winning probability of an advertiser with quality $a$ is,

$\phi^{IS}(v, a) = \{E_{a'}[F(va'/a')]\}^{n-1}$

This winning probability is the same as $\phi^I(v, a)$, the winning probability under the PPI case. This is because both schemes achieve the most efficient allocation of the advertising slot.

Given the advertisers’ equilibrium winning probability, we can follow similar steps as before to derive the overall expected revenue for the publisher:

$$\pi^{IS} = n\alpha\tilde{b}_y a_h \int_0^1 \phi^{IS}(v, a_h) \ J(v) \ f(v) \ dv + n(1 - \alpha) \ b_y a_l \int_0^1 \phi^{IS}(v, a_l) \ J(v) \ f(v) \ dv$$

$$= \hat{b}^{IS}_y \pi^h + b_y \pi^l$$

Where $\pi^h = n\alpha\int_0^1 \phi^{IS}(v, a_h) \ J(v) \ f(v) \ dv$ and $\pi^l = n(1 - \alpha) \ b_a \int_0^1 \phi^{IS}(v, a_l) \ J(v) \ f(v) \ dv$ are the base revenues from high- and low-quality advertisers respectively. Because $\phi^{IS}(v, a) = \phi^l(v, a)$, the base revenues from high- and low-quality advertisers under PPI/S are the same as those under PPI. Hence, we have

$$\pi^{base} = \pi^h + \pi^l$$
**Proposition 3.** Under assumptions 1, 2, 4, and 5, (a) \((I, I)\) is a THP equilibrium when

\[
\hat{b}_I^I \pi_{base}^I \geq \hat{b}_h^I \pi_{h}^I + b_h \pi_{I}^I, \quad \hat{b}_l^I \pi_{base}^I \geq \hat{b}_I^I \pi_{I}^I + b_l \pi_{I}^I, \quad \text{and} \quad \hat{b}_h^I \pi_{base}^I \geq b_h \pi_{base}^I.
\]

(b) \((IS, IS)\) is a THP equilibrium when

\[
\hat{b}_h^I \pi_{h}^IS + b_h \pi_{I}^IS \geq \hat{b}_l^I \pi_{base}^I, \quad \hat{b}_l^I \pi_{base}^I \leq \hat{b}_h^I \pi_{I}^IS + b_l \pi_{I}^IS, \quad \text{and} \quad \hat{b}_h^I \pi_{h}^IS + b_h \pi_{I}^IS \geq b_h \pi_{base}^I.
\]

(c) \((I, S)\) is a THP equilibrium when

\[
b_h \pi_{base}^S \geq b_l \pi_{I}^IS + b_h \pi_{I}^IS \quad \text{and} \quad \mu (l|IS) = 1.
\]

(d) \((IS, S)\) is a THP equilibrium when

\[
b_h \pi_{base}^S > b_l \pi_{I}^IS + b_h \pi_{I}^IS \quad \text{and} \quad \mu (l|I) = 1.
\]

**Proof.** Note that PPS is weakly dominated for a low-quality publisher by PPI and PPI/S. As a result, \((S, I)\), \((S, IS)\), and \((S, S)\) are not THP. \((I, IS)\) and \((IS, I)\) are not equilibrium either because the low-quality publisher is better off by mimicking the high-quality one in either case. We show that the remaining four strategy profiles are THP under appropriate conditions.

(a) Under \((I, I)\), a low-quality publisher gets \(\hat{b}_I^I \pi_{base}^I \) with PPI and \(\hat{b}_h^I \pi_{h}^I + b_l \pi_{I}^I \) by deviating to PPI/S. A high-quality publisher gets \(\hat{b}_h^I \pi_{base}^I \) under PPI, \(\hat{b}_h^I \pi_{I}^IS + b_l \pi_{I}^IS \) by deviating to PPI/S, and \(b_h \pi_{base}^S \) by deviating to PPI. Conditions in (23) ensure all deviations are unprofitable. The proof for (b) is analogous to that for (a) thus omitted.

(c) Under \((I, S)\), a low-quality publisher gets \(b_l \pi_{base}^I \) with PPI and \(\hat{b}_I^I \pi_{I}^IS + b_l \pi_{I}^IS \) by deviating to PPI/S. The deviation is unprofitable only when \(\mu (l|IS) = 1\) (implying \(\hat{b}_I^I = b_l\)). A high-quality publisher gets \(b_h \pi_{base}^S \) under PPS, \(b_l \pi_{base}^I \) by deviating to PPI, and \(\hat{b}_h^I \pi_{I}^IS + b_l \pi_{I}^IS \) by deviating to PPI/S. Because \(\pi_{base}^I = \pi_{h}^I + \pi_{I}^IS\), deviating to PPI is clearly dominated by deviating to PPI/S. Conditions in (24) ensure that other deviations are unprofitable. (d) The proof this case is analogous to that for (c) thus omitted.
We can verify that (a) through (d) survive the trembling hand perfection. The verification is omitted due to space limitation.

\((IS, IS)\) and \((I, I)\) result in the same allocation because both PPI and PPI/S are efficient auctions. But they generate different revenues for high- and low-quality publishers. The reason is that in the former case, only high-quality advertisers use noisy signals to infer publisher’s quality whereas in the latter case, all advertisers do. Under \((I, I)\), the revenues for low- and high-quality publishers are \(\hat{b}_l^I \pi_{base}^l\) and \(\hat{b}_h^I \pi_{base}^h\) respectively. In contrast, under \((IS, IS)\), \(\hat{b}_l^{IS} \pi_h^{IS} + b_l \pi_l^{IS}\) and \(\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS}\). Note that \(\hat{b}_y^I = \hat{b}_y^{IS}\) for \(y \in \{l, h\}\) and \(\pi_{base}^I = \pi_{base}^{IS}\) \((22)\), so the high-quality publisher is better off under \((IS, IS)\) whereas the low-quality publisher is better off under \((I, I)\).

Not all the THP equilibria in Proposition 3 are plausible. The equilibrium \((I, I)\) is “unstable” in the following sense. Its first condition in \((23)\) requires \(\mu(h|IS) < \beta.\) That is, this equilibrium requires advertisers to believe that a low-quality publisher is more likely to deviate to the off-equilibrium strategy \(IS\) than a high-quality publisher. However, this contradicts the observation that a high-quality publisher prefers \((IS, IS)\) and a low-quality publisher prefers \((I, I)\). We formalize this notion of stability by extending Mailath, Okuno-Fujiwara, and Postlewaite (1993)’s (MOP for short) notion of defeated equilibrium.

MOP view off-equilibrium messages as attempts to overturn the current equilibrium by player types who seek a better alternative equilibrium for themselves. MOP require players to assign probabilities to off-equilibrium actions according to an alternative equilibrium, if the off-equilibrium action is taken in the alternative equilibrium exactly by the player types who strictly prefer the alternative equilibrium. This way of refining the off-equilibrium beliefs allows an equilibrium to be “defeated” by an alternative equilibrium. We extend the MOP’s idea to the case where the off-equilibrium action in the alternative equilibrium may also be taken by player types who are worse off in the alternative equilibrium.

\(^{7}\)To see this, suppose \(\mu(h|IS) \geq \beta.\) We have \(\hat{b}_h^{IS} \geq \hat{b}_h^I\). So \(\hat{b}_l^I \pi_{base}^l = \hat{b}_l^I (\pi_h^{IS} + \pi_l^{IS}) \leq \hat{b}_h^{IS} (\pi_h^{IS} + \pi_l^{IS}) \leq \hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS}\), which contradicts the first condition in \((23)\).
**Definition 1.** (Compromised Equilibrium) A compromised equilibrium refinement requires players to believe that an off-equilibrium action in the original equilibrium is taken by player types who take such an action in an alternative equilibrium and are better off, if such an alternative equilibrium exists. The original equilibrium is *compromised* by the alternative equilibrium if it fails the compromised equilibrium refinement.

**Lemma 6.** (a) \((I, I)\) is compromised. (b) \((I, S)\) is compromised when

\[
\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS} \geq b_h \pi_{base}.
\]

(c) \((IS, IS)\) and \((IS, S)\) are not compromised.

**Proof.** (a) It is evident from earlier discussion that \((I, I)\) is compromised by \((IS, IS)\). (b) A high-quality publisher gets \(b_h \pi_{base}^{S}\) under \((I, S)\) and \(\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS}\) under \((IS, IS)\). When condition (27) holds, a high-quality publisher is better off with \((IS, IS)\). By the notion of compromised equilibrium, \(\mu (h|IS) = 1\), which contradicts the condition for \((I, S)\). So \((I, S)\) is compromised under the condition (27). (c). It is easy to verify that \((IS, IS)\) and \((IS, S)\) are not compromised by any other equilibrium.

According to the compromised equilibrium, the publishers under \((IS, IS)\) should hold an off-equilibrium belief \(\mu (l|I) = 1\). Under this belief, the first two conditions of (24) hold naturally. So (24) simplifies to (27).

\((IS, IS)\) and \((IS, S)\) are both equilibrium when \(\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS} \geq b_h \pi_{base}^{S} > b_l \pi_l^{IS} + b_h \pi_l^{IS}\). \((IS, IS)\) generates an expected revenue of \(\hat{b}_h^{IS} \pi_h^{IS} + b_l \pi_l^{IS}\) and \(\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS}\) for low- and high-quality publishers respectively. \((IS, S)\) generates an expected revenue of \(b_l (\pi_l^{IS} + \pi_l^{IS})\) and \(b_h \pi_{base}^{S}\) respectively. Clearly \((I, I)\) Pareto-dominates \((I, S)\).

**Corollary 4.** Under assumptions 1-5, \((IS, IS)\) is an uncompromised THP equilibrium with supporting belief \(\mu (l|I) = 1\) when

\[
\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_l^{IS} \geq b_h \pi_{base}^{S}.
\]
When otherwise, \((I, S)\) is an uncompromised THP equilibrium with supporting belief \(\mu(l|IS) = 1\) and \((IS, S)\) with \(\mu(l|I) = 1\).

\((I, S)\) and \((IS, S)\) are equivalent for both advertisers and the publisher because in the latter case, advertisers fully infer the publisher’s quality and a low-quality publisher with PPI/S achieves the same allocation and revenue as in the first case with PPI. So for practical purposes, when the publishers can choose the portfolio strategy, the publisher choose between two equilibria, \((IS, IS)\) and \((IS, S)\).

Comparing the case with and without the portfolio scheme, we observe the following:

- Both support a separating equilibrium and a pooling equilibrium. Without portfolio, the two equilibrium choices are \((I, I)\) and \((I, S)\). With portfolio scheme, the two equilibrium choices are \((IS, IS)\) and \((IS, S)\) (or \((I, S)\)).
- With the portfolio scheme, the pooling equilibrium \((I, I)\) is replaced by \((IS, IS)\). In such a case, the high-quality publisher is better off and the low-quality publisher is worse off.
- \((I, S)\) is also replaced by \((IS, IS)\) when \(\hat{b}_h^{IS} \pi_h^{IS} + b_h \pi_I^{IS} \geq b_h \pi_{base}^{IS} \geq b_l \pi_{base}^{IS} = b_l \left( \pi_h^{IS} + \pi_I^{IS} \right)\). In such a case, both the high- and low-quality publishers are better off. Moreover, the overall allocation efficiency improves (because PPI/S is efficient whereas PPS is not).

In sum, when the portfolio scheme is available, the high-quality publisher has the incentive to induce the \((IS, IS)\) equilibrium instead of \((I, I)\) and to displace \((I, S)\) with \((IS, IS)\) under certain parameter range, sometimes at the cost of low-quality publishers.

5 Extensions

5.1 Extension to Pay-per-click

To incorporate the pay-per-click (PPC) payment scheme, we consider a two-stage model of advertising. In stage 1, impressions turn into clicks with probability \(a_1b_1\), where \(a_1 \in \{a_{1l}, a_{1h}\}\) is the advertiser’s stage-1 quality and \(b_1 \in \{b_{1l}, b_{1h}\}\) is the publisher’s stage-1 quality. In stage 2, clicks turn to sales with probability \(a_2b_2\), where \(a_2 \in \{a_{2l}, a_{2h}\}\) is the
advertiser’s stage-2 quality and \( b_2 \in \{b_{2l}, b_{2h}\} \) is the publisher’s stage-2 quality. The stage-1 quality corresponds to the click through rate. The stage-2 quality corresponds to the conversion rate. Once again, we assume that \( a_1, a_2, b_1, b_2 \) are independently distributed. Correspondingly, the weighting factor for the PPI auction is \( w = \hat{a}_1 \) and the weighting factor for the PPS auction is \( w = \hat{a}_1 \hat{a}_2 \).

\[
\text{Probability of sale} = a_1 b_1 \times a_2 b_2
\]

\[
\text{click through rate} \quad \text{conversion rate}
\]

Our previous results are applicable to the case of PPI vs. PPC with a simple redefinition of variables (i.e., redefining \( v \) as valuation per click and replacing \( a \) with \( a_1 \) and \( b \) with \( b_1 \)). For example, under the assumption that \( \pi^I_{base} \geq \pi^C_{base} \), a high-quality (i.e., high click-through-rate) publisher may prefer PPC when there are relatively few high-quality publishers, when there is a large spread in publisher quality, and when advertisers have little information about publisher quality (Corollary 3). It is also profitable for a high-quality publisher to induce the equilibrium \((IC, IC)\), in which both publisher types let advertisers choose between PPI and PPC (Proposition 4). Similar conclusions can also be drawn for the case of PPC vs. PPS if publisher’s stage-1 quality is identical.

With the two-stage framework, it is also feasible to examine the case where the publisher can choose from three payment schemes \( \{I, C, S\} \). A strategy profile in this case is a mapping from four publisher types \( \{ll, lh, hl, hh\} \) to three payment schemes \( \{I, C, S\} \). With 81 (=3^{4}) possible strategy profiles (compared with just 4 in the base model), the equilibrium analysis in this case is more complex. Nevertheless, our analytical and numerical analysis show that the intuition for the publisher’s equilibrium payment choice is similar to our base model (See Online Appendix for a derivation of publisher revenues under the two stage model). For example, under the assumption that \( \pi^I_{base} \geq \pi^C_{base} \geq \pi^S_{base} \), we find that the \( ll \)-type publisher always choose PPI, \( lh \)-type chooses between PPI and PPS, \( hl \)-type chooses between PPI and PPC, and \( hh \)-type chooses between PPI, PPC, and PPS. Numerical results that demonstrate the equilibrium payment scheme choice for each type of publisher are available upon request.
5.2 Advertisers’ Uncertainty About Own Quality

In Section 3, we have assumed that the advertiser knows precisely her own quality $a$ and the publisher only observes a signal of it. We may relax this assumption by assuming that the advertiser only knows an imperfect signal about her quality $a$. We use $\hat{x}_P \in \{l, h\}$ and $\hat{x}_A \in \{l, h\}$ to denote the publisher’s signal and the advertiser’s signal respectively. The advertiser’s signal $\hat{x}_A$ may be correlated with the publisher’s $\hat{x}_P$. Denote $\tilde{a}_P = E[a|\hat{x}_P]$ and $\tilde{a}_A = E[a|\hat{x}_A]$ as the publisher’s and the advertiser’s estimate of an advertiser’s quality. Let $\lambda_A \equiv P(\hat{x}_A = h|x = l) = P(\hat{x}_A = l|x = h) \geq 0$ denote the probability for an advertiser to misclassify own quality type. $\lambda_A$ is common knowledge.

Under the PPI auction, an advertiser’s valuation per impression becomes $v \equiv v\hat{a}_A\hat{b}$. Let $t$ be the advertiser’s bid. The advertiser’s expected utility is given by $U(t, v) = \phi^I(t) \left[ v\hat{a}_A\hat{b} - p(t) \right]$. Because $\hat{b}$ is the same across all advertisers, an advertiser with $(v, \tilde{a}_A)$ must bid the same as an advertiser with $(v', \tilde{a}'_A)$ in equilibrium whenever $v\hat{a}_A = v'\hat{a}'_A$. Following Lemma 1, we can show that the equilibrium bid must increase with $v$ if we hold $\tilde{a}_A$ constant. So the equilibrium winning probability of an advertiser is completely determined by her $v$ and $\tilde{a}_A$

$$\phi^I(v, \tilde{a}_A) = \left\{ E_{\tilde{a}_A'} \left[ F \left( v\hat{a}_A/\tilde{a}_A' \right) \right] \right\}^{n-1}.$$  

With the equilibrium winning probability, we can similarly derive the equilibrium revenue for the publisher as

$$\pi^I = nb\tilde{y}E_{\tilde{a}_A} \left[ \tilde{a}_A \int_0^1 \phi^I(v, \tilde{a}_A) J(v) f(v) \, dv \right] \equiv \hat{b}_y\tilde{\pi}^I_{base}$$

Under the PPS auction, advertiser’s signal $\hat{x}_A$ does not affect the optimal bidding because it only scales an advertiser’s expected payoff. So the equilibrium bidding and allocation under the PPS auction remain the same as before. The publisher’s expected revenue is also the same because $\hat{x}_A$ is advertiser’s private information and the publisher relies on his signal $\hat{x}_P$ in estimating the expected revenue.
It becomes clear that the impact of the advertiser’s imperfect information about her quality is confined to $\tilde{\pi}_{\text{base}}^l$. In this case, the PPI auction is no longer efficient. But as long as $\tilde{\pi}_{\text{base}}^l > \pi_{\text{base}}^S$, the subsequent results on the publisher’s equilibrium in Section 3 still apply. Clearly, if $\tilde{\pi}_{\text{base}}^l < \pi_{\text{base}}^l$, then a high-quality publisher is more likely to adopt PPS when advertisers are uncertainty about own quality (Corollary 3). It is worth noting that, it does not matter the advertiser’s quality signal is correlated with the publisher’s because both PPI and PPS use only one signal. In contrast, the portfolio scheme may take advantage of the correlation between the two signals. The reason is that under PPI/S, advertisers can get a second opinion about their quality by observing the weighting factor assigned to them by the publisher. So they can update their belief about own quality and make a better-informed bid and payment scheme choice. A complete analysis of the publisher’s equilibrium with the portfolio scheme is not attempted in this paper due to space limitation. But we conjecture that the PPI/S payment schemes are more likely to be used in equilibrium when advertisers are unsure about their quality, due to their benefits from pooling information.

5.3 Extension to Multiple Slots

We can also extend our model to allow multiple slots. Suppose there are $k \leq n$ slots and an advertiser’s valuation for slot $j$ is $v a \hat{b} \delta_j$ with $1 \geq \delta_1 \geq \delta_2 \geq \delta_3, \ldots \geq \delta_k \geq 0$. $\delta_j$ can be interpreted as the number of (effective) impressions received on slot $j$. The auction allocates the slots in a way that the highest bidder gets slot 1, the second highest gets slot 2, and so on. Under the PPI auction, an advertiser $(v, a)$’s expected utility by bidding $t$ is,

$$U(t, v, a) = \sum_{l=1}^{k} \delta_l C_{n-k}^l [F_T(t)^{n-k} [1 - F_T(t)]^{k-1} [v a \hat{b} - p(t)]]$$

where $C_{n-k}^l \equiv \frac{n!}{(n-k)k!}$, $F_T(t)$ is the bid distribution, and $p(t)$ is the expected payment per impression by bidding $t$. Lemma 1 still holds in the multiple slot case. So we may redefine $\phi^l(v, a) = \sum_{l=1}^{k} \delta_l C_{n-k}^l [E_{a'} [F(v a / a')]]^{n-k} [1 - E_{a'} [F(v a / a')]]^{k-1}$ as the new equilibrium “winning probability”. Similarly, under the PPS auction, we may derive an advertiser’s new equilibrium “winning probability” as $\phi^S(v, \hat{a}) = \sum_{l=1}^{k} \delta_l C_{n-k}^l [E_{a'} [F(v \hat{a} / a')]]^{n-k} [1 - E_{a'} [F(v \hat{a} / a')]]^{k-1}$
The rest of the analysis in Section 3 will follow through except that the old $\phi^I(v, a)$ and $\phi^S(v, a)$ are replaced with new ones.

6 Discussion and Conclusions

Our study is among the first to examine the optimality of different payment schemes in the presence of information asymmetry between publishers and advertisers. Despite the rapid growth of pay-for-performance advertising models, traditional pay-per-impression mechanisms continue to exist and thrive alongside these newer payment schemes. Our study examines the conditions under which publishers find it optimal to offer these different payment schemes. In particular, given the presence of different qualities of advertisers and publishers, the need for not only separating advertisers of different qualities, but also signaling one’s own quality arises for publishers. Our findings illustrate how the different payments schemes can be used by publishers to not only differentiate among advertisers but also credibly signal their own quality. PPI leverages advertisers’ private information about their quality, while pay-for-performance schemes (PPS and PPC, to some extent) reveal the quality of the publishers to advertisers; hence it follows that ceteris paribus, advertisers would prefer pay-for-performance payment schemes and publishers prefer PPI. However, this does not take into account the dispersion in quality among publishers or among advertisers. Given the presence of low-quality publishers, while low quality publishers benefit from pooling using PPI, high quality publishers benefit from separating themselves using the more transparent pay-for-performance schemes. On the other hand, given the dispersion in advertiser quality, the use of pay-for-performance schemes creates allocational inefficiencies - the higher the dispersion in advertiser quality, the greater the loss for a publisher from misclassification errors. This tradeoff between the cost and benefits of the different payment schemes lies at the heart of our analyses.

A high quality publisher finds it optimal to choose PPS/PPC when the benefits of separating from other (low quality) publishers outweigh the loss from misclassification errors. Thus, in a market with little information about publisher quality (for instance, a market with a large
number of nascent publishers), or in a market with a large number of low-quality publishers, a high-quality publisher benefits from the transparency afforded by pay-for-performance schemes. However, this might not be necessary for well-established high-quality publishers. Many web portals, including many large websites (such as nytimes.com) still uses PPI as the payment scheme of choice for their banner ads, despite the dramatic advances in PPC technology. Our model offers a theoretical explanation for this phenomenon. These large and reputable sites have very low probability of being mistaken as a low-quality publisher (i.e., small $\gamma_P$) but face considerable risk of misjudging advertisers’ quality if they switch to PPC. In addition, the choice of PPI eliminates the need for the publisher to estimate advertiser quality. Consequently, PPI becomes a dominant choice of these high-quality publishers.

While several high-quality content publishers such as nytimes.com, washingtonpost.com, wsj.com use PPI (in the form of banner ads), other high-quality publishers - most notably, search engines such as Google, and Yahoo! - resort to the use of pay-for-performance schemes - PPC, in particular. As is well known, search engines accumulate large amounts of click-through data making them very adept at estimating advertisers’ performance. Information about advertisers’ performance help reduce any potential allocational inefficiencies that stem from the use of pay-for-performance schemes, particularly PPC. However, while search engines tend to have detailed information on click-through rates, information on conversion rates is less reliable. More importantly, advertisers typically have a larger role than publishers (i.e., search engines) when it comes to conversion. Adopting PPS is likely to increase misclassification errors for publishers; hence, search engines have little incentive to adopt PPS.

Traditionally, advertising has been dominated by pay-per-impression, largely due to the lack to adequate targeting and measurement technologies. With the advent of the technologies that enable granular targeting of advertisements and measurement of their quality, publishers now have a choice of offering multiple payment schemes. When publishers have a choice of multiple payment schemes, our findings suggest that publishers are more likely to offer a portfolio of payment schemes. For instance, firms such as Yahoo! and Google
offer both PPI and PPC for content-based advertisements, while firms such as Amazon offer PPI in the form of banner advertisements, PPC in the form of sponsored search advertisements, as well as PPS in the form of third-party sellers. Our results show that while offering multiple payment schemes reduces a high-quality publisher’s ability to separate, offering a portfolio of payment schemes benefits the publisher. Firstly, any loss from pooling with low-quality publishers is reduced as only part of transactions - those that use PPI - suffer from the inability of advertisers to distinguish amongst publishers. Secondly, any loss from misclassification errors are reduced as well, as high-quality advertisers self-select into PPI.

Our study also highlights two opposing effects of advances in advertising technologies. Technologies that better measure and predict advertiser quality facilitate the adoption of pay-for-performance mechanisms as they reduce the cost of separating. On the other hand, technologies that track publishers’ quality and dynamically adjust bids eliminate the need for a high-quality publisher to separate from low-quality publishers, and thus tend to favor PPI. As publishers being to take greater advantage of the measurement technologies, pay-for-performance mechanisms are more likely, especially among the new and distinct high-quality publishers.

We model payment schemes in advertising as a means of leveraging private information. PPI fully exploits advertisers’ private information but none of the publisher’s; PPS fully leverages the publisher’s private information but none of advertisers’ private information. Thus the payment schemes differ in how the responsibility of monitoring (or estimating) is divided among the publisher and the advertiser. Who should assume the responsibility of monitoring would depend not only on the monitoring technology, but also on the relative contribution of two parties to performance. If performance is influenced more by publishers than advertisers ($b_h/b_l$ increases), PPS, which leverages publisher’s information and imposes monitoring responsibility on the publisher, is more likely to be used. If performance is affected more by advertisers than the publisher ($a_h/a_l$ increases), PPI, which leverages advertisers’ information and imposes monitoring responsibility on advertisers, is
more likely to be used. Our argument thus, provides an alternative theoretical perspective to the conventional "risk-sharing" model of payment schemes.

The weighted ranking rule are presented as a way for publishers to leverage their (incomplete) information about advertisers’ quality in pay-for-performance mechanisms. Notice that under such a ranking rule, when the publisher has perfect information about advertisers, PPS achieves the same allocation efficiency as PPI. In fact, when the two parties have perfect information about each other, different payment schemes are equivalent. The adoption of weighted ranking rule allows us to conclude that the inefficiency associated with PPS is not an inherent deficiency of the pay-for-performance scheme (which is often assumed by the literature) but rather a result of imperfect information (about advertisers).

Our study is a first step in understanding the role of information asymmetry in the evolving landscape of advertising payment schemes. Our study and the findings pave the way for a number of possible extensions. Our model focuses on a representative publisher drawn from a heterogeneous pool. It would be interesting to examine the role of competition in a duopolistic setting. Our study examines a static setting. It would be important to understand the dynamics of payment scheme choices. It would also be valuable to understand the optimal allocation of resources among different payment schemes for a publisher. Our study models issues of adverse selection (hidden information) rather than moral hazard (hidden action) issues such as in Zhu and Wilbur (2010). It would be interesting to consider the impacts of hidden action in addition to those of hidden information. Our model assumes risk-neutral advertisers and publishers. It would be important to examine the impact of risk aversion: risk averse advertisers are likely to prefer PPS, and risk averse publishers PPI. While the inclusion of risk aversion is likely to shift the optimality conditions, it is unlikely to reverse the broad findings of our analyses. Finally, given the increased availability of data on advertising payment schemes, empirical tests of these theoretical models would add to our understanding of this complex landscape.
References


**Appendix A. Appendix**

**A.1 Proof of Lemma 1**

Denote $t^{(2)}$ as the second highest bid under PPI and $F_T(\cdot)$ as the distribution of the equilibrium bids. Denote $\phi(t) = [F_T(t)]^{n-1}$ and $p(t) = E[t^{(2)}|t^{(2)} < t]$ as the winning probability and the expected payment per impression of an advertiser with bid $t$. The expected payoff of an advertiser with valuation per impression $v$ and bid $t$ is

$$U(t, v) = \phi(t) [v - p(t)]$$

Let $t = T(v)$ and $t' = T(v')$ for some $v' \neq v$. By the equilibrium definition, we have $U(t', v) \leq U(t, v)$ and $U(t, v') \leq U(t', v')$. Substituting (28) and reorganizing, we get

$$\phi(t') v - \phi(t) v \leq \phi(t') p(t') - \phi(t) p(t)$$

and

$$\phi(t') v' - \phi(t) v' \geq \phi(t') p(t') - \phi(t) p(t).$$
Combining two inequalities, we have \([\phi(t') - \phi(t)](v' - v) \geq 0\), which implies for any \(v' > v\), \(\phi(t') \geq \phi(t)\) and therefore \(T(v) \geq T(v')\).

Next we show by contradiction that \(T(v) = T(v')\) can not hold in equilibrium. Suppose for some \(v < v'\) such that \(T(v) = T(v')\). Because \(T(v)\) is nondecreasing, we must have \(T(v)\) is a constant on \([v, v']\). We argue that the advertiser \(v\) is strictly better off by bidding \(T(v) + \epsilon\) where \(\epsilon\) is an infinitely small number. Doing so increases her expected payment (conditional on winning) by an infinitely small amount \((p(T(v) + \epsilon) - p(T(v)))\) but increases her winning probability by a significant amount because she can beat everyone within \([v, v']\) rather than tie them. This contradicts that \(T(v)\) is equilibrium.

A.2 Derivation of The Equilibrium Revenue under PPI (10)

Consider an advertiser with valuation-per-impression \(v\) and bid \(T(v')\). Denote \(\phi^I(v')\) and \(p^I(v')\) as the advertiser’s winning probability and the expected pay per impression conditional on winning. The advertiser’s expected payoff is \(U(v, v') = \phi^I(v') [v - p(v')]\), where \(\phi^I(v)\) is given by (9). Denote \(V(v) = U(v, v)\) as the advertiser’s equilibrium expected payoff. Notice that

\[
\frac{dV(v)}{dv} = \left[ \frac{\partial}{\partial v} U(v, v') + \frac{\partial}{\partial v'} U(v, v') \frac{\partial v'}{\partial v} \right] |_{v' = v}
\]

\[
= \frac{\partial}{\partial v} U(v'|v) |_{v' = v} = \phi^I(v)
\]

where the second step is due to the envelope theorem. Suppose the advertiser with the lowest valuation \((v = 0)\) has zero expected payoff. Then, \(V(v) = \int_0^v \phi^I(u) \, du\). Noticing that the total expected payment from advertiser is \(M(v) = \phi^I(v) v - V(v)\), we can write the publisher’s expected revenue as

\[
\pi' = nE_{\theta} [E_{v} [M(v)]] = nE_{\theta} \left[ E_{v} \left[ \phi^I(v) v - \int_0^v \phi^I(u) \, du \right] \right]
\]
Substituting $v = v^a \hat{b}_I$ and (9), we have

$$\pi^I = nE_y \left[ E_{v,a} \left[ \phi^I (v, a) v^a b_I - a \hat{b}_I \int_0^v \phi^I (u, a) du \right] \right]$$

$$= nE_y \left[ \hat{b}_I \right] E_a \left[ a \int_0^1 \left( \phi^I (v, a) v - \int_0^v \phi^I (u, a) du \right) f (v) dv \right]$$

$$= n \hat{b}_y E_a \left[ a \int_0^1 \phi^I (v, a) \left( v - \frac{1 - F (v)}{f (v)} \right) f (v) dv \right]$$

where the third equality is due to integration by parts.

A.3 Proof of Lemma 2

(1). Denote $U (t, v, w)$ as the expected payoff per unit sale of an advertiser with unit valuation $v$, weighting factor $w$, and bid $t$. Denote $\phi^S (t, w)$ and $p^S (t, w)$ as the advertiser’s winning probability and expected payment per sale conditional on winning. The expected payoff of an advertiser with valuation $v$, weighting factor $w$, and bid $t$ is

$$U (t, v, w) = \phi^S (t, w) \left[ v - p^S (t, w) / w \right]$$

(29)

Following the same steps in the Proof of Lemma 1, we can establish that for many $v' > v$, we must have $T (v, w) \geq T (v', w)$ and that $T (v, w) = T (v', w)$ cannot be equilibrium.

(2) For an advertiser with valuation $\lambda v$, weighting factor $w / \lambda$, and bid $\lambda t$,

$$U (\lambda t, \lambda v, w / \lambda) = \phi (\lambda t, w / \lambda) \left( \lambda v - p (\lambda t, w / \lambda) / (w / \lambda) \right)$$

(30)

$$= \phi (t, w) \left( \lambda v - \lambda p (t, w) / w \right) = \lambda U (t, v, w)$$

where the second equality is because $\phi (\lambda t, w / \lambda) = \phi (t, w)$ and $p (\lambda t, w / \lambda) = p (t, w)$.

(30) implies that if $t^*$ maximizes $U (t, v, w)$ then $\lambda t^*$ must maximize $U (t, \lambda v, w / \lambda)$, i.e., $\lambda T (v, w) = T (\lambda v, w / \lambda)$. Substituting $v' = \lambda v$ and $w' = w / \lambda$, we get $wT (v, w) = w' T (v', w')$.

A.4 Derivation of the Publisher Revenue under PPS (13)

Consider an advertiser with valuation $v$, weighting factor $w$, and bid $T (v', w)$. Denote $\phi^S (v', w)$ and $p^S (v', w)$ as the advertiser’s winning probability and expected pay per sale
conditional on winning. The advertiser’s expected payoff per sale is

\[ U(v, v', w) = \phi(v', w) [v - p(v', w)] \]  

where \( \phi^S(v', w) \) is given by (12). Denote \( V(v, w) = U(v, v, w) \) as the advertiser’s equilibrium expected payoff per sale. Following the steps in A.2, we can similarly obtain \( V(v, w) = \int_0^v \phi(u, w) du \). The total expected payment per sale from the advertiser is \( M(v, w) = \phi(v', w) v - V(v) \). So the publisher’s expected revenue is

\[ \pi^S = E_{v,w} [bwM(v, w)] \]

\[ = nbE_w \left[ wE_v \left[ \phi(v, w) v - \int_0^v \phi(u, w) du \right] \right] \]

\[ = nbE_w \left[ w \int_0^1 \left[ \phi(v, w) v - \int_0^v \phi(u, w) du \right] f(v) dv \right] \]

\[ = nbE_w \left[ w \int_0^1 \phi(v, w) \left[ v - \frac{1 - F(v)}{f(v)} \right] f(v) dv \right] \]

A.5 When \( \pi^S_{I_{base}} < \pi^S_{I_{base}} \), \( (S, S) \) is the only pure-strategy equilibrium.

By adopting PPS, the high-quality publisher gets \( b_h \pi^S_{I_{base}} \), the highest she can get. So it is a dominant strategy for the high-quality publisher to choose PPS. Given the high-quality’s strategy, a low-quality strictly prefers PPS, which generates \( b_l \pi^S_{I_{base}} \), to PPI, which generates \( b_l \pi^S_{I_{base}} \) (a publisher whose chooses is deemed as a low-quality because PPI is a dominated strategy for high-quality publisher).

A.6 The Publisher’s Total Expected Revenue under PPI/S

First consider a low-quality advertiser who chooses PPS (17). Her expected payoff of bidding \( t \) per sale is \( U^S_l(t, v) = \phi^S(tw) [v - p(tw)/w] \). Denote \( V_l(v) = U^S_l(v, v) \). Following the similar steps as before, we get \( V_l(v) = \int_0^v \phi^S(u, a_l) du \). Her expected payment per-sale is \( E[M(v, a_l)] = \int_0^1 [\phi^S(v, a_l) v - V_l(v)] f(v) dv = \int_0^1 \phi^S(v, a_l) J(v) f(v) dv \). The publisher’s expected revenue from a low-quality publisher is \( b_l a_l \int_0^1 \phi^S(v, a_l) J(v) f(v) dv \).

Next we consider a high-quality advertiser who chooses PPI with a expected payoff of \( U^S_h(t, v) = \phi^S(t, a_h) \left[ va_h \hat{b} - p(t) \right] \) per impression. Following similar steps as before, we
have that the expected payment from a high-quality advertiser, conditional on publisher’s type \( y \), is

\[
\hat{b}_y a_h \int_0^1 \hat{\phi}^{IS} (v, a_h) J (v) f (v) dv.
\]

So the publisher’s total expected revenue is

\[
\pi^{IS} = n \alpha a_h \hat{b}_y \int_0^1 \hat{\phi}^{IS} (v, a_h) J (v) f (v) dv + n (1 - \alpha) b_y a_i \int_0^1 \hat{\phi}^{IS} (v, a_i) J (v) f (v) dv
\]
B.1. Proof of \( a_l \leq \hat{a}_l \leq \hat{a}_h \) (4). By (3), \( \hat{a}_l = \frac{(1-\alpha)(1-\gamma_A)}{\alpha\gamma_A+(1-\alpha)(1-\gamma_A)}a_l + \frac{\alpha\gamma_A}{\alpha\gamma_A+(1-\alpha)(1-\gamma_A)}a_h \). In other words, \( \hat{a}_l \) is a weighted average of \( a_l \) and \( a_h \). Clearly, \( \hat{a}_l \geq a_l \). Similarly, \( \hat{a}_h = \frac{(1-\alpha)\gamma_A}{\alpha(1-\gamma_A)+(1-\alpha)\gamma_A}a_l + \frac{\alpha(1-\gamma_A)}{\alpha(1-\gamma_A)+(1-\alpha)\gamma_A}a_h \) and we have \( \hat{a}_h \leq a_h \). \( \hat{a}_l \leq \hat{a}_h \) is because \( \frac{\alpha\gamma_A}{\alpha\gamma_A+(1-\alpha)(1-\gamma_A)} \leq \frac{\alpha(1-\gamma_A)}{\alpha(1-\gamma_A)+(1-\alpha)\gamma_A} \), which is implied by \( \gamma_A > 0.5 \).

B.2. Proof of \( b_l \leq \hat{b}_l^{m} \leq \hat{b}_h^{m} \leq b_h \) (7). By (5) and (6), \( \hat{b}_l = (1-\gamma_P)E[b|\hat{y}=l,m] + \gamma_P E[b|\hat{y}=h,m] \) where \( E[b|\hat{y}=l,m] = \frac{\beta\rho_h^{m}(1-\gamma_P)b_h+(1-\beta)\rho_l^{m}(1-\gamma_P)b_l}{\beta\rho_h^{m}P_l^{m}+\(1-\beta)\rho_l^{m}P_h} \) and \( E[b|\hat{y}=h,m] = \frac{\beta\rho_h^{m}(1-\gamma_P)b_h+(1-\beta)\rho_l^{m}\gamma_P b_l}{\beta\rho_h^{m}(1-\gamma_P)+\(1-\beta)\rho_l^{m}\gamma_P} \). Clearly \( E[b|\hat{y},m] \) is a weighted average of \( b_l \) and \( b_h \) and so is \( \hat{b}_l \). Thus we have \( b_h \geq \hat{b}_l \). Similarly, we also have \( b_h \geq \hat{b}_h \). To prove that \( \hat{b}_l \leq \hat{b}_h \), noting that \( \hat{b}_h = \gamma_P E[b|\hat{y}=l,m]+(1-\gamma_P)E[b|\hat{y}=l,m] \) and \( \gamma_P > 0.5 \), we only need to show \( E[b|\hat{y}=l,m] \leq E[b|\hat{y}=h,m] \). The proof of \( E[b|\hat{y}=l,m] \leq E[b|\hat{y}=h,m] \) is analogous to that of \( \hat{a}_l \leq \hat{a}_h \) and therefore omitted.

B.3. The Semi-separating Equilibrium. We claim that when both conditions (14) and (15) hold, there also exists a mixed-strategy semi-separating equilibrium \((\rho_l^*,\rho_h^*) = (1,\rho_0)\) where \( \rho_0 \) is the solution to

\[
(32) \quad b_h\pi_{base}^S = \hat{b}_h^l\pi_{base}^l.
\]

Because in the semi-separating equilibrium the high-quality publisher chooses PPI and PPS with positive probability, the publisher must have equal expected revenue under PPI and PPS auctions, which implies that (32) must hold. Note that \( \hat{b}_l^l \) is an increasing function of \( \rho_l^* \).

For (32) to have a solution, we must have \( b_h\pi_{base}^{S} \geq \hat{b}_l^l|\rho_l^* = 0\pi_{base}^l \) and \( b_h\pi_{base}^{S} \leq \hat{b}_l^l|\rho_l^* = \pi_{base}^l \). The former leads to (14) and the latter leads to (15). Because \( \hat{b}_l^l \) increases in \( \rho_l^* \), it is straightforward that when (14) and (15) both hold, (32) must have a solution, thus permitting a semi-separating equilibrium.

B.4. The Two-Stage Model with Pay-Per-Click. We consider a general setting where advertisers pay for each unit of outcome, which can be either impression, click, or sale. The advertiser’s valuation per unit outcome is \( v \) and valuation per sale is \( v \). The distribution
of $v$ is as defined before. We denote $F(\cdot)$ and $f(\cdot)$ as the cumulative distribution function and the probability density function of $v$ respectively. We denote $a_1 b_1$ as the probability of outcome per impression. $a_1$ is the advertiser’s impression-to-outcome quality and $b_1$ is the publisher’s impression-to-outcome quality. We denote $a_2 b_2$ as the probability of sale per unit outcome. $a_2$ is the advertiser’s outcome-to-sale quality and $b_2$ is the publisher’s outcome-to-sale quality. Let $x_1, x_2, y_1,$ and $y_2$ denote the advertiser and the publisher’s types that correspond to $a_1, a_2, b_1,$ and $b_2$ respectively. Let $\hat{x}_1$ and $\hat{y}_2$ be signals for $x_1$ and $y_2$ respectively. Let $w_{\hat{x}_1} = E[a_1 | \hat{x}_1]$ (we will write $w$ for short) be the weighting factor for an advertiser with signal $\hat{x}_1$. Such a framework can be used to represent PPI, PPC, or PPS auctions. In particular, PPI auctions can be represented by setting $a_1 = b_1 = 1$ and $a_2 = a_2, b_2 = b_2$. PPS auctions can be represented by setting $a_1 = a_1, b_1 = b_1, a_2 = a_2,$ and $b_2 = b_2$. The PPC auctions can be represented by letting $a_1 = a_1, b_1 = b_1, a_2 = a_2,$ and $b_2 = b_2$.

**Lemma 7.** (1) The equilibrium bid $t^*(v, w)$ is strictly increasing in $v$. (2) An advertiser $(v, w)$ ties with advertiser $(v', w')$ in equilibrium if

$$vw = v' w'$$

**Proof.** Available upon request. $\square$

Given Lemma 7, we can immediately derive the equilibrium winning probability. Consider an advertiser whose valuation per sale is $v$, outcome-to-sales performance is $a_2$, and weighting factor is $w$ (base on the signal $\hat{x}_1$ about her type $x_1$). Her equilibrium winning probability is :

$$\phi(v a_2, w) = \left\{ E_{a_2' w'} \left[ F\left( \frac{v a_2 w}{a_2' w'} \right) \right] \right\}^{n-1}$$

**Proposition 4.** The publisher’s expected revenue in the above general setting is

$$\pi = n b_1 E[b_2 | \hat{y}_2] E_w \left[ w E_{a_2} \left[ a_2 \int_0^1 \phi(v a_2, w) J(v) f(v) dv \right] \right]$$
where $J(v) = v - [1 - F(v)]/f(v)$ and $\phi(va_2, w)$ is the equilibrium winning probability defined in (34).

Proof. Available upon request. □

B.4.1. The PPI Revenue. Because the impressions are independent of the characteristics of advertisers and the publisher, $b_1$ and $a_1$ degenerate into constant 1. The auction is not weighted and signal $x_1$ does not exist. On the other hand, $a_1 = a_1 a_2$ and $b_2 = b_1 b_2$. An advertiser’s valuation per impression is $v = va_2 E[b_1 b_2 | y_1, y_2]$, where $E[b_1 b_2 | y_1, y_2] \equiv \hat{b}_1 \hat{b}_2$ is the same for every advertiser and takes a similar form as (5). Correspondingly,

$$F(v) = E_{a_1' a_2'} \left[ F \left( \frac{va_1 a_2 \hat{b}_1 \hat{b}_2}{a_1' a_2' b_1 b_2} \right) \right] = E_{a_1' a_2'} \left[ F \left( \frac{a_1 a_2}{a_1' a_2'} \right) \right]$$

$$\phi(va_2, w) = \{F(v)\}^{n-1}.$$

According to the revenue formula in Proposition 4, the total expected revenue under PPI is (denote $\hat{b}_1 \hat{b}_2 \equiv E[E[b_1 b_2 | y_1, y_2] | y_1, y_2]$)

$$\pi = n E[E[b_2 | y_2] E_{x_2} \left[ \int_0^1 \phi(va_2, 1) J(v) f(v) dv \right]$$

$$= nb_1 b_2 E_{a_1, a_2} \left[ a_1 a_2 \int_0^1 \left\{ E_{a_1' a_2'} \left[ F \left( \frac{v a_1 a_2}{a_1' a_2'} \right) \right] \right\}^{n-1} J(v) f(v) dv \right]$$

B.4.2. The PPS Revenue. Because the outcomes are measured by the number of sales, $b_2$ and $a_2$ degenerate into constant 1. On the other hand, $a_1 = a_1 a_2$ and $b_1 = b_1 b_2$. An advertiser’s valuation per unit outcome is $v = v$. The quality signal of the advertiser is given by $x_1 (\hat{x}_1, \hat{x}_2)$ and the weighting factor for an advertiser with observed types $(\hat{x}_1, \hat{x}_2)$ is $w = E[a_1 a_2 | \hat{x}_1, \hat{x}_2] \equiv \hat{a}_1 \hat{a}_2$. No quality signal for the publisher is used. Correspondingly,

$$\phi(va_2, w) = \left[ E_{w'} \left[ F \left( \frac{v w}{w'} \right) \right] \right]^{n-1} = \left\{ E_{a_1' a_2'} \left[ F \left( \frac{v \hat{a}_1 \hat{a}_2}{a_1' a_2'} \right) \right] \right\}^{n-1}$$. 

According to the revenue formula in Proposition 4, the total expected revenue under PPS is
\[
\pi^S = nb_1 E_w \left[ w \int_{\hat{v}}^v \phi(v a_2, w) J(v) f(v) dv \right]
\]
\[
= nb_1 b_2 E_{\hat{a}_1, a_2} \left[ \hat{a}_1 \hat{a}_2 \int_0^1 \left\{ E_{\hat{a}_1', a_2'} \left[ F \left( \frac{\hat{a}_1 \hat{a}_2}{\hat{a}_1' \hat{a}_2'} \right) \right] \right\} \right] \left\{ J(v) f(v) dv \right\}^{n-1}
\]

B.4.3. The PPC Revenue. Because the outcomes are measured by the number of clicks, \( a_1 = a_1, b_1 = b_1, a_2 = a_2 \) and \( b_2 = b_2 \). The quality signal for the advertiser is \( \hat{x}_1 = \hat{x}_1 \) and the weighting factor for an advertiser with observed type \( \hat{x}_1 \) is \( w = E[a_1|\hat{x}_1] \equiv \hat{a}_1 \). The quality signal for the publisher is \( \hat{y}_2 = \hat{y}_2 \). An advertiser’s valuation per unit outcome is \( v = va_2 E[b_2|\hat{y}_2] \equiv v a_2 \hat{b}_2 \). Correspondingly,
\[
\phi(v a_2, w) = \left[ E_{w', a_2'} \left[ F \left( \frac{a_2 w}{a_2' w'} \right) \right] \right]^{n-1} = \left[ E_{\hat{a}_1', a_2'} \left[ F \left( \frac{\hat{a}_1 a_2}{\hat{a}_1' a_2'} \right) \right] \right]^{n-1}
\]

According to the revenue formula in lemma (4), the total expected revenue under PPC is (denote \( E[E[b_2|\hat{y}_2]|\hat{y}_2] = \hat{b}_2 \))
\[
\pi = nb_1 E[E[b_2|\hat{y}_2]|y_2] E_w \left\{ w E_{a_2} \left[ a_2 \int_0^1 \phi(v a_2, w) J(v) f(v) dv \right] \right\}
\]
\[
= nb_1 \hat{b}_2 E_{\hat{a}_1} \left[ \hat{a}_1 E_{a_2} \left[ a_2 \int_0^1 \left\{ E_{\hat{a}_1', a_2'} \left[ F \left( \frac{\hat{a}_1 a_2}{\hat{a}_1' a_2'} \right) \right] \right\} \right] \right] \left\{ J(v) f(v) dv \right\}^{n-1}
\]
\[
= nb_1 \hat{b}_2 E_{\hat{a}_1, a_2} \left[ \hat{a}_1 a_2 \int_0^1 \left\{ E_{\hat{a}_1', a_2'} \left[ F \left( \frac{\hat{a}_1 a_2}{\hat{a}_1' a_2'} \right) \right] \right\} \right] \left\{ J(v) f(v) dv \right\}^{n-1}
\]