

AN EMPIRICAL MODEL OF MOBILE ADVERTISING PLATFORMS

KHAI X. CHIONG AND RICHARD Y. CHEN

ABSTRACT. We study a new online advertising platform, created specifically for mobile app-to-app advertising. Both advertisers and publishers are mobile applications (apps). Advertisers seek to acquire new users for their mobile apps, while publishers seek to monetize their apps. Our data come from the intermediary who operates this two-sided platform, and who uses a centralized market-clearing mechanism to meet the demand with the supply of users' in-app impressions. Notably in this mechanism, advertisers bid for impressions, but only pay when impressions are won and when ads lead to user acquisitions (pay-per-install). We develop a model for the advertiser's optimal bidding problem, and use observed bids to recover the advertiser's valuation or willingness-to-pay for a new user.

We find certain segments of this market to be severely uncompetitive, resulting in large bid-shading, and lost profit for the intermediary and publishers. Using the estimated model, we discuss how the intermediary can make the platform more competitive. Interestingly, we also find that how much a user spends on in-app purchases contributed to only about 10% of the average advertiser's valuation for that user. We argue that this relatively large unobserved valuation partly stems from advertisers' resale motives: after acquiring users and selling in-app purchases to them, an advertiser then has the option to resell these users by becoming a publisher, selling these users' impressions to other advertisers. Thereby, the advertiser participates in both sides of the market as an advertiser and a publisher.

1. INTRODUCTION

The Internet and smart mobile devices are transforming how consumers receive information of new products. Digital advertising, which consists of online and mobile advertising among others, has become an increasingly important part of many businesses' marketing channels. In 2015, businesses' spending on digital

Date: October 2016.

Khai Chiong (corresponding author): USC Dornsife INET, Department of Economics, University of Southern California, kchiong@usc.edu.

Richard Chen: OpenAI, richardchen100@gmail.com. Acknowledgements are at the end of the document. All errors remain our own.

advertising¹ increased by 17.2% to \$160 billion. In fact, 2017 is widely projected to be the tipping point when businesses spend more on digital advertising than traditional media advertising such as TV and print.² Driving this stunning rise is mobile advertising. Mobile devices have now firmly replaced desktop and laptop computers as the primary means for individuals to access the Internet, and in many emerging economies, they are the only means. According to Google, more searches now take place on mobile devices than on computers in 10 countries including the U.S. and Japan. Moreover in the U.S., spending on mobile advertising amounts to \$31.59 billions in 2015, which is more than half of the total \$59.61 spending on all digital advertising.³

In this paper, we introduce a new online advertising platform, which has been specifically created for the purpose of mobile app-to-app advertising. In this two-sided platform, both advertisers and publishers are mobile apps (applications). For instance, we could have Uber advertises on Instagram, the Starbucks mobile app advertises on the Waze app (GPS-based navigational app), or a mobile gaming app advertises on another gaming app. Mobile apps have become the main interface between consumers and firms within mobile devices. On the consumer's side, users prefer to install a myriad of mobile apps for their different needs than to use a mobile web browser.⁴ Individual apps are able to provide better user experience and personalization.⁵ On the firm's side, firms including traditional retailers, are facing a growing need to develop and market their presences in mobile devices as mobile apps, which act as a gateway to branding and purchases.

In this new online advertising platform, advertisers and publishers participate in a new bidding mechanism whereby advertisers set bids which are promises-to-pay when (i) advertisers win the slot to show ads, and (ii) users take a certain action after ads. This is known as a CPA (cost-per-action) bid. Particularly in our app-to-app environment, this user's action is specifically defined as the installation of the advertiser's app. That is, advertisers bid for users' impressions,

¹Digital advertising includes all advertising (online and offline) that appears on digital media devices, which is defined by being interactive, personalized, or Internet-connected, such as desktop, laptop computers, mobile phones, tablets, gaming consoles.

²New York Times (December 7, 2015). Digital Ad Spending Expected to Soon Surpass TV. The Economist (March 26, 2016). Invisible ads, phantom readers.

³eMarketer (March 8, 2016). Digital Ad Spending to Surpass TV Next Year.

⁴ComScore: the average users of mobile devices in the U.S. spend 86% of their time on mobile apps, as opposed to 14% on mobile web browsers.

⁵App developers can maximize precious screen real estate by getting rid of navigation bars imposed by the web browsers. They can access phone features such as push-notification, camera, GPS, physical sensors. Once installed, apps also work faster than their web-counterparts.

but only pay this bid to the publisher in the event that their ads lead to a *user acquisition or install*. This platform mechanism is highly attractive to the advertisers, who only need to pay per user’s install. It aligns with the objective of the app developers who operate on the “freemium” business model (see [Lee et al. \(2015\)](#)), which seeks to acquire as many new users as possible. With this new bidding mechanism, publishers are now more accountable for ad effectiveness, as they would otherwise be selling impressions for free if ads do not lead to a desired user’s response. CPI (cost-per-install) advertising has seen stunning rise in the past few years. Businesses’ spending on CPI campaigns has increased by 80% from 2014 to 2015, and accounted for 10.3% of total mobile advertising spend in 2015.⁶

Another feature of this platform that we exploit is that mobile advertisers have the capability to track the behavior of users post viewing of ads. Advertisers use these user-tracking data to actively inform their valuations and to optimize bids. For instance, advertisers know how much different types of acquired users spend on in-app purchases, which are products offered by the mobile apps. These tracking capabilities contrast with traditional advertising channels such as TV, newspapers, and roadside billboards, which do not allow advertisers to understand users’ behavior after viewing an ad.⁷ Tracking user’s spending on in-app purchases is an important task due to the “freemium” business model adopted by a vast majority of mobile apps – installing the apps is free, monetization is achieved by a small fraction of users paying for additional premium content within the app (which includes monthly subscriptions and in-game currency).

The contribution of our paper is two-fold. Firstly, we take the managerial perspective that the intermediary is our client, and we seek to optimize the platform design in order to increase the intermediary’s revenue. Secondly, we advance our understanding of the behavior of advertisers in the context where they can also participate in the other side of the market as publishers. To achieve this, we develop a model for the advertiser’s optimal bidding problem. We then use observed bids to recover the advertiser’s valuation or willingness-to-pay for a new user. Our data come from the intermediary, which is the entity who operates this two-sided platform.

Knowing the advertiser’s valuation allows us to quantify the margin of strategic bid-shading – the difference between valuation and observed bids, or in another words, how much profit the intermediary and publishers are losing as a result

⁶eMarketer (December, 2015). Mobile Advertising and Marketing Trends Roundup

⁷Another important feature of digital advertising is its capability to target users with much higher precision ([Goldfarb \(2014\)](#))

of uncompetitive bidding among the advertisers. We find that certain segments of the market are severely uncompetitive. We discuss several policies the intermediary can pursue to increase competition among advertisers. We also use the estimated model to evaluate the counterfactual effectiveness of these policies. For instance, we show that in some cases, it is desirable to introduce some inaccuracy in the score. Specifically, it is sometimes desirable to boost and increase the scores of some uncompetitive and low-quality advertisers. On one hand, score boosting introduces inaccuracy and mismatch in the matching so that there are now fewer user acquisitions, but on the other hand, it stimulates more aggressive bidding by advertisers. Overall, score-boosting can increase the revenue of the intermediary.

Our numerical result: the estimated valuation for a new user is \$12.69 for the average advertiser. For the median advertiser, it is \$8.55. The interquartile range is [\$5.10, \$15.53]. By comparison, the advertisers ended up paying significantly less to acquire these users: the average winning bid is \$4.04, and the median at \$3.50. This leads to a large margin (difference between willingness-to-pay and the price paid) that ranges from \$1.87 to \$10.52 in the interquartile range. Our result strongly suggests that certain segments of the market are severely uncompetitive, where advertisers can shade their bids well below their valuations, and still win a large number of impressions. This leads to lost profits for the publishers and intermediary.

Further, we decompose the valuation for a new user into two parts: the first part is the sales revenue from in-app purchases made by the acquired user; the second part is the advertiser’s *unobserved valuation*, which represents the residual value that the acquired user brings to the advertiser. The first part is *observed* using data that track user’s spending on in-app purchases after acquisition.⁸ In-app purchases are products offered by the app developer within the app, which take the form of additional content, services or subscriptions within the app. In our sample, the 4,973,931 users that were acquired in the span of 47 days, have spent a total of \$679,470 on in-app purchases. *Methodologically, the advertiser’s valuation is partially observed, and we exploit the information in this observed component to better estimate the overall valuation.*

However we find that users’ spending account for about 9% of an average advertiser’s overall valuation. This is puzzling given the importance of users’ in-app purchases.⁹ We offer an explanation for what drives the advertiser’s valuation

⁸Which is routinely collected by advertising intermediaries, where they then inform advertisers of their Return on Investments (ROI) of advertising.

⁹One popular mobile gaming app, Clash of Clans, generates \$4 million a day from in-app purchases (the app developer was recently acquired for \$8.6 billion by Tencent).

for user acquisitions: due to the app-to-app advertising nature of our platform, an app developer can readily participate in both sides of the market, as both an advertiser and a publisher. This introduces **resale motive which partly drives up the advertisers' valuations for a new user**.

More concretely, the advertiser perceives a user as a **durable good**, whereby a user generates a stream of diminishing monetary benefits in terms of in-app purchases. Now the advertiser then has the option to resell a user by selling its impression and attention to other advertisers. By reselling the attention of the user, the user would potentially be drawn away to other apps (due to the scoring procedure used by the platform mechanism, the ads that are shown to this user are often competing apps that are closely relevant to the user). The crucial question this naturally leads to is, how then should firms optimally predict and measure the lifetime value of their customers, given this resale option? To our knowledge, this goes beyond the standard models of customer lifetime valuation.

The estimation procedure is based on the Generalized Method of Moments (GMM), where the moment conditions provide a link between (i) observed bids; (ii) observed valuation due to users' spending; (iii) unobserved valuations. Intuitively, we use a profit-maximization framework to derive advertisers' first-order optimality conditions trading off the expected benefit and cost of higher bids. Although the estimation procedure is frequentist in nature, our model also lends itself naturally to a full-information likelihood approach, which can be estimated using Bayesian methods.

1.1. Related literature

First, our paper is related to the vast literature of digital advertising. Mobile advertising includes keyword search advertising that takes place in mobile devices. Theoretical papers that analyse search advertising are [Amaldoss et al. \(2015\)](#); [Shin \(2015\)](#), while empirical papers in this area include [Athey and Nekipelov \(2010\)](#); [Hsieh et al. \(2014\)](#); [Jeziorski and Moorthy \(2016\)](#); [Rutz and Bucklin \(2011\)](#); [Yang and Ghose \(2010\)](#); [Yao and Mela \(2011\)](#)). Display advertising such as banner advertising ([Andrews et al. \(2015\)](#); [Bruce et al. \(2016\)](#); [Johnson \(2013\)](#); [Manchanda et al. \(2006\)](#)) has been studied extensively in the context of web browsers on desktop or laptop computers. Display advertising can also occur in mobile devices (see [Bart et al. \(2014\)](#) who study mobile display advertising (MDA)).

Our paper studies a growing form of mobile advertising called app-to-app advertising, where companies seek to promote their mobile apps within other mobile apps. The fast-rising prominence of the mobile app economy has been noted

and studied in a recent paper by [Ghose and Han \(2014\)](#), which uses a structural model to estimate the demand for mobile apps.

The intermediary we study in this paper is an example of an online advertising platform or network, which provides a common marketplace for advertisers and publishers to buy and sell impressions (see [Sriram et al. \(2015\)](#) for a survey of recent work related to advertising platforms). In a novel paper, [Wu \(2015\)](#) studies an online advertising platform which uses a decentralized matching mechanism, as opposed to the centralized mechanism here. An important role of an online advertising platform is the ability to facilitate the delivery of targeted ads to users ([Goldfarb and Tucker \(2011\)](#); [Iyer et al. \(2005\)](#); [Lambrecht and Tucker \(2013\)](#); [Sayedi et al. \(2014\)](#); [Zhang and Katona \(2012\)](#)). In mobile advertising, targeting can be achieved with even more degrees of freedom, due to built-in GPS sensors in mobile devices (see [Andrews et al. \(2014\)](#); [Fong et al. \(2015\)](#); [Grewal et al. \(2016\)](#); [Luo et al. \(2013\)](#); [Zubcsek et al. \(2015\)](#)). More generally, mobile advertising goes beyond mobile app-to-app advertising in this paper. It also includes firms using SMS to send promotional messages (see [Andrews et al. \(2016\)](#); [Shankar and Balasubramanian \(2009\)](#) for comprehensive surveys of mobile marketing).

The CPI (cost-per-install) advertising considered in this paper is a type of performance-based or CPA (cost-per-action) advertising which includes the popular CPC (cost-per-click). The CPC pricing scheme where advertisers pay per clicks have received much attention in the literature ([Agarwal et al. \(2009\)](#); [Asdemir et al. \(2012\)](#); [Ghose and Yang \(2009\)](#); [Hu et al. \(2015\)](#); [Liu and Viswanathan \(2014\)](#); [Zhu and Wilbur \(2011\)](#)). Another common pricing scheme is CPM where advertisers pay per impression. The CPI advertising we considered here is more recent and attributed to the recent rise of the mobile app economy and app-to-app advertising. The objective of a CPI advertising campaign is user acquisition.

Another related area is modeling customers' lifetime value. These papers are normative, i.e. prescribing how to measure CLV, while our paper here is descriptive in nature, i.e. asking what advertisers are doing to form their valuations. For primers on modeling CLV, we refer to [Fader and Hardie \(2005\)](#); [Fader et al. \(2005\)](#); [Gupta et al. \(2006\)](#); [Schmittlein et al. \(1987\)](#). Here, another important paper is [Chan et al. \(2011\)](#), which estimate the lifetime value for a firm's customers acquired through sponsored search advertising (cost-per-click campaigns) on Google based on the Pareto/NBD model.

Also noteworthy is the literature on measuring the returns on investment (ROI) of online advertising. The valuation of the advertiser for a new user is strongly

related to how much returns they expect to receive from advertising. [Lewis and Rao \(2015\)](#) highlighted the challenges for advertisers to evaluate the ROI from impression-based (CPM) advertising campaigns. A related paper on causal effectiveness and ROI of sponsored search advertising is [Blake et al. \(2015\)](#).

2. INDUSTRY BACKGROUND

A mobile app is a computer program designed to run on mobile devices such as smartphones and tablet computers. The mobile apps industry started with the introduction of the iPhone and Apple’s app store in 2008. App developers market their products through distribution platforms called app stores (Apple app store and the Google Play are the two largest), which takes a 30% cut out of the developer’s revenue.

By far the overwhelming fraction of mobile apps have adopted the ‘freemium’ business model,¹⁰ such that the users install the app for free and are given the option to make in-app purchases (see also [Lee et al. \(2015\)](#)). A well-known success stories of the freemium model is the mobile game Clash of Clans, which generates \$4 million a day in revenue, just from in-app purchases. Supercell, its app developer, posted \$2.4 billion in revenue in 2015, and was recently acquired by the Chinese internet giant, Tencent, for \$8.6 billion. This transaction is the seventh largest Chinese overseas acquisition on record.

The genres of mobile apps include photography (Instagram), social networking, health & fitness, shopping, travel & navigation (Uber), news, books, utilities, music. By far the most prominent is the gaming genre – global revenue from mobile games is on track to rise 21% to about \$37 billion this year.¹¹

A *primary concern* of mobile app developer is marketing. While mobile apps such as Instagram are worth \$1 billion, many mobile apps are not worth much. In 2015, Apple announced that there were over 1.5 million apps in the Apple app store, and over 100 billion apps had been downloaded.¹² Two stylized facts stand out: the large number of products available on the app store, and the large

¹⁰In 2014, freemium app revenue now accounts for 98% of worldwide revenue on Google Play, with Japan, U.S. and South Korea users contributing the most. <http://blog.appannie.com/google-io-special-report-launch-2014/>

¹¹Tencent President Martin Lau said “We are very bullish on the [mobile games] market” <http://www.wsj.com/articles/tencent-agrees-to-acquire-clash-of-clans-maker-supercell-1466493612>

¹²<http://www.statista.com/statistics/263794/number-of-downloads-from-the-apple-app-store/>

number of potential users dispersed worldwide. Marketing and advertising their products have become essential for mobile app developers.

Recognizing this challenge, several platforms have begun to fill this niche in the last several years. The San Francisco-based start-up from which our data are obtained exists in such a niche.

2.1. How the platform works

The San Francisco-based start-up, which we will call an *intermediary*, operates and designs the platform that brings together the buyers (advertisers) and sellers (publishers) of users' impressions. The platform runs a centralized market-clearing mechanism to meet the demand with the supply of users' in-app impressions. We now describe this mechanism.

For the publishers, the intermediary offers a SDK (software development kit) for app developers wanting to monetize their user-base by publishing ads. The developers then integrate the SDK into the infrastructure of their apps, which allow for video ads to be displayed within the app. When a user within the publisher's app reaches a pre-specified ad placement opportunity, the SDK pings the intermediary for an ad to be served to the user.

For the advertisers, the intermediary allows the advertisers to specify three parameters:¹³ (i) the publishers to advertise on, (ii) the CPI (cost-per-install) bid for each of the publisher, (iii) the active duration of these bids. The advertiser's objective is to acquire new users through advertising. *A CPI bid is the amount that the advertiser pays to the intermediary in the event that the advertiser (a) wins the slot to show an ad, and (b) the user installs the mobile app being advertised after watching the ad.*

The intermediary shares a fraction of this CPI payment to the publisher where this user's impression comes from. We summarize this in Figure 1 below.

2.1.1. Market-clearing mechanism

We now describe the most important function of the platform, which is to run a mechanism that decides which ads are matched or served to which users. Whenever a user's impression arises from the publisher's side, the mechanism first identifies all potential advertisers and the CPI bids they placed in the database.

¹³The advertisers can supply their own creatives, or the intermediary can provide a premium service for designing the creatives. The creatives are in video form, with a typical length up to 35 seconds.

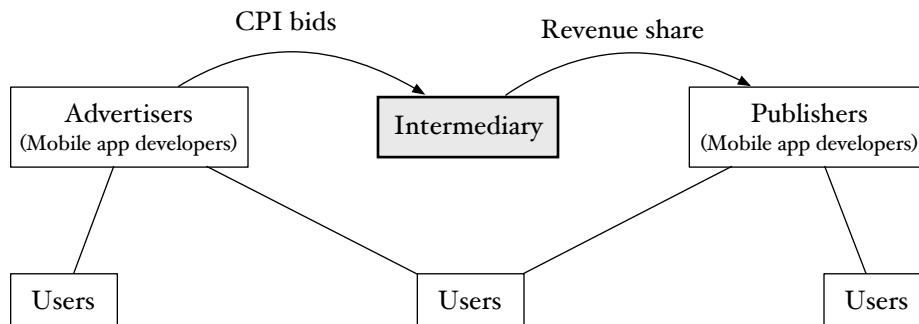


FIGURE 1. Overview of how the platform works

Next, the mechanism assigns a score to each ad based on the following formula: $CPI \times Score$, where $Score$ is the estimated probability that the user would install the advertised app conditional on viewing the ad. It is calculated by plugging in the user and advertiser's attributes into a prediction model.

The mechanism then selects the ad with the highest score-weighted CPI bid as the winner of that user's impression. This ad will be shown to the user. After the user watches the ad, he or she can click and proceed to the advertiser's page in the App Store. If the user installs the app, we say the user is acquired by the advertiser. If the user does not install, the advertiser is not charged, otherwise, the advertiser pays the CPI bid to the intermediary (who passes a fraction of this revenue to the publisher).

The score-weighted CPI bid associated with a pair of user and ad is just the expected amount that the advertiser pays for that user's impression. Intuitively then, for each user's impression that arises, the mechanism selects the advertiser with the highest expected payment to be the winner.

2.2. Advertiser's information set

In Figure 2, we show an actual example of what an advertiser would see during and after his bidding activities. The intermediary provides a graphical interface that reports the various outcomes of the advertiser's bidding activities.

In the example given in Figure 2, the advertiser had set a CPI bid of \$8.99 that was (i) held constant for two days, and (ii) specific to a given publisher. At that bid, the advertiser would know that he was able to win 192,000 impressions, and acquire 447 new users. At the end of those two days, the advertiser had paid a sum of \$4,020 to the intermediary, a fraction of which was passed on to the publisher.

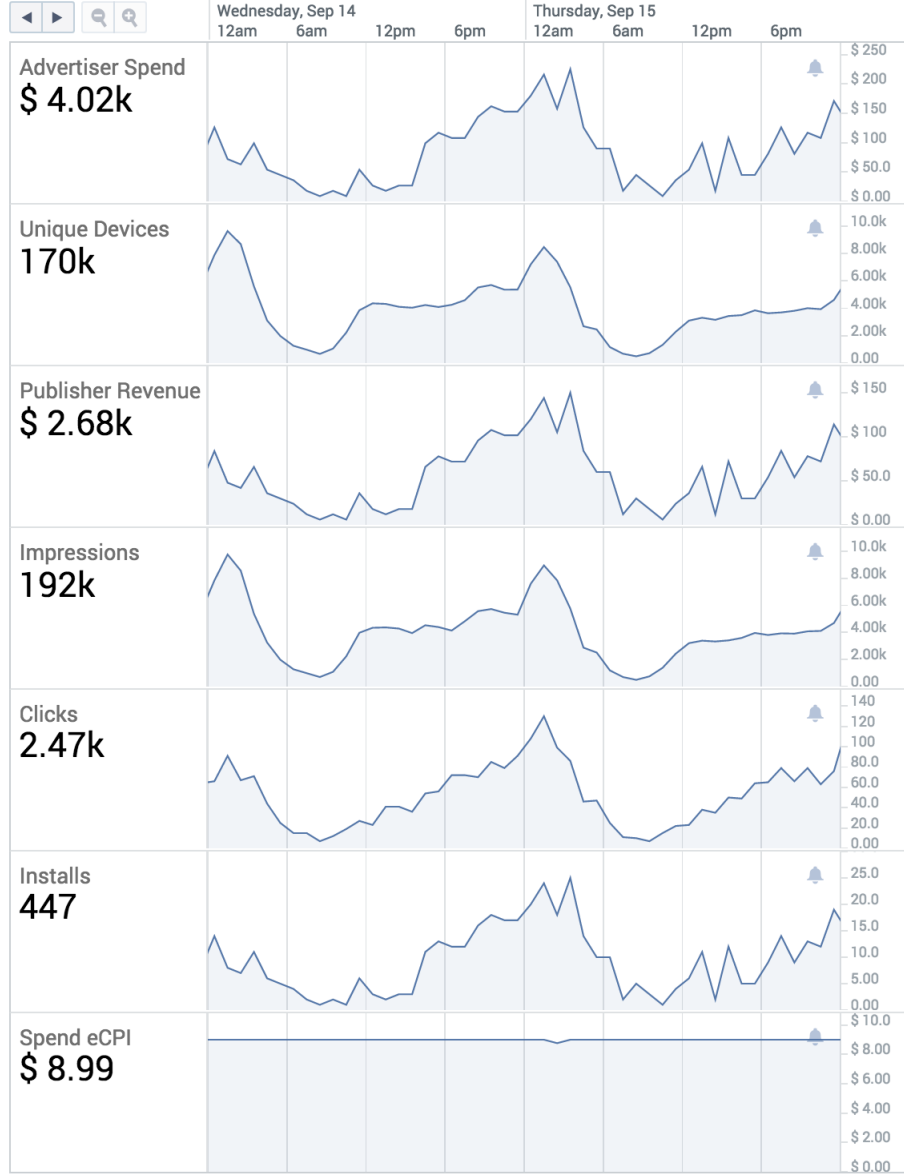


FIGURE 2. Advertiser's information set: what do advertisers see?

3. MODEL

Consider a setup between one advertiser and one publisher. In this section, we propose a model for how the advertiser forms his bid for that publisher, conditional on choosing to bid on that publisher.

The advertiser has a valuation v for the user. *This valuation is a random variable*, to reflect the fact that the advertiser is bidding to acquire a distribution of users who benefit the advertiser heterogeneously. Now the advertiser knows the probability distribution of v , but it is not known to us. Ultimately, we want to learn about v , which is also known as the advertiser’s willingness-to-pay for a new user.¹⁴

When the advertiser submits a bid of p , he expects to acquire $Q(p)$ number of users from the publisher. Now $Q(\cdot)$ is a random function, i.e. $Q(p)$ is a random variable whose probability distribution is parameterized by p . The expected profit of the advertiser when he bids p is:

$$(1) \quad U(p) = \mathbb{E}[(v - p)Q(p)]$$

The expectation in Equation (1) is taken over the joint probability density of $(v, Q(p))$. We will refer to $Q(p)$ as the supply curve. We refer to Section 3.2 for a detailed comparison with the standard auction setup.

3.1. Specifying the supply curve

We now specify $Q(p)$, which is the advertiser’s belief about the number of users he will successfully acquire at different levels of bid p . In formulating $Q(p)$, we model closely how the advertiser’s bids is translated into numbers of user acquisitions. See Section 2.1.

$$(2) \quad Q(p) = F(s \cdot p)^N \cdot \chi \cdot s$$

Now, the advertiser believes that the score-weighted bids of each of his competitor is distributed independently with the cumulative density function $F(\cdot)$. Therefore, the advertiser believes that he will win each impression with probability $F(s \cdot p)^N$ where N is the number of competing bidders, and s is the quality score associated with the advertiser. For a given user that arises from the publisher, the intermediary computes a quality score that equals to the estimated probability that the user would install after watching the advertiser’s ad. This score s is a random variable (since different users within the publisher would have different scores). For simplicity, we will first take s to be the average score of

¹⁴See Chan et al. (2007), where they define WTP as the maximum amount a bidder is willing to bid for an item such that she is indifferent between winning the item at this bid and not winning.

the advertiser specific to the publisher. In the empirical section, we show how to relax this assumption in the estimation.

χ is the random variable representing the advertiser's belief about the total number of impressions that would be supplied by the publisher. Therefore $F(s \cdot p)^N \cdot \chi$ is the number of impressions that the advertiser expects to win at bid p . Note that χ and s do not depend on p . Although the number of competing bidders N is fixed, in Section 4.7, we allow it to be a random variable to reflect entry uncertainty.

Finally, since the quality score s is the probability that a user would install after being shown the advertiser's app, $F(s \cdot p)^N \cdot \chi \cdot s$ is then the number of users that the advertiser expects to acquire at bid p .

For ease of notation, we will introduce the following:

$$(3) \quad \alpha(p) \equiv F(s \cdot p)^N$$

That is, $\alpha(p)$ is the probability that the advertiser expects to win an impression at bid p .

Proposition 1. *Suppose that the advertiser's expected profit is given by Equation 1. Then the optimal bid p^* satisfies:*

$$(4) \quad \frac{\alpha(p^*)/p^*}{\left. \frac{\partial \alpha(p)}{\partial p} \right|_{p^*}} = \frac{1}{p^*} \frac{\mathbb{E}[v\chi]}{\mathbb{E}[\chi]} - 1$$

Plugging in $\alpha(p) = F(s \cdot p)$, where $F(\cdot)$ is the advertiser's belief about the CDF of the adjusted CPI bids of other competing bidders, and $f(\cdot)$ is the corresponding PDF, we then have:

$$(5) \quad \frac{F(sp^*)}{sp^* N f(sp^*)} = \frac{1}{p^*} \frac{\mathbb{E}[v\chi]}{\mathbb{E}[\chi]} - 1$$

We relegate all proofs to the appendix. The LHS of Equation 4 is the inverse of the elasticity of $\alpha(p)$ with respect to p evaluated at p^* .

Proposition 2. *When $F(\cdot)$ is the CDF of the lognormal distribution, there is a **unique** optimal bid p^* that satisfies Equation 5. In particular, sufficient*

conditions for the optimal bid p^* satisfying Equation 5 to be **unique** are: (i) $\lim_{z \rightarrow 0_+} \frac{F(z)}{zf(z)} = 0$, and (ii) $\frac{F(z)}{zf(z)}$ is strictly monotonically increasing in z . Both of these conditions are satisfied when $F(\cdot)$ is the CDF of the lognormal distribution.

In particular, when $F(\cdot)$ is the CDF of the lognormal distribution with parameter (μ, σ) , that is, the random variable X drawn from the distribution F is such that $\log X$ is distributed $N(\mu, \sigma)$, then

$$(6) \quad \frac{\sqrt{2\pi}}{N} \sigma e^{-\frac{(\mu - \log(sp))^2}{2\sigma^2}} \Phi\left(\frac{\log(sp) - \mu}{\sigma}\right) = \frac{1}{p} \frac{\mathbb{E}[v\chi]}{\mathbb{E}[\chi]} - 1$$

where $\Phi(\cdot)$ is the CDF of the standard Gaussian. In the appendix, we show that the LHS of Equation 27 is strictly monotonically increasing in p , and converges to zero as $p \rightarrow 0_+$.

Proposition 3 below states the factors that determine the advertiser's optimal bid. Intuitively, bidding higher increases the advertiser's chance of winning, but the advertiser also ends up paying more if he does win. The optimal bid depends on this trade-off.

Proposition 3. *Assume that the conditions in Proposition 2 hold. The advertiser's optimal CPI bid p^* is larger when:*

- (1) $\frac{\mathbb{E}[v\chi]}{\mathbb{E}[\chi]}$ is larger;
- (2) N , the number of competing bidders is larger;
- (3) $\frac{F(x)}{xf(x)}$ is weakly smaller for all $x > 0$ and strictly smaller for some $x > 0$, where $f(\cdot)$ and $F(\cdot)$ are the probability and cumulative distribution functions associated with the advertiser's belief about the score-weighted CPI bids of other competing bidders.

In another words, the slope of the function $\frac{F(x)}{xf(x)}$ determines how competitive the market is. When the function $\frac{F(x)}{xf(x)}$ is steep, the market is less competitive. Moreover consider $\tilde{F}(\cdot)$ and $\tilde{f}(\cdot)$ such that $\tilde{F}(\cdot)$ is a *first-order stochastic dominant* shift of $F(\cdot)$. Then it follows that $F(x) \geq \tilde{F}(x)$ at all $x > 0$ and $F(x) > \tilde{F}(x)$ for some x . A first-order stochastic dominant shift can be achieved by taking a mass of bidders and increasing their score weighted CPI bids, holding all else constant. If this first-order stochastic dominant shift is large enough such that $\frac{F(x)}{xf(x)}$ is weakly larger than $\frac{\tilde{F}(x)}{\tilde{f}(x)}$ for all x and strictly larger for some x , then the advertiser would increase its' bid.

To give an example, we can verify that if F is the CDF of the triangular distribution with support $[0, 1]$ and mode of 0.25, and \tilde{F} is the CDF of the triangular distribution with support $[0, 1]$ and mode of 0.75, then it is true that $\frac{F(x)}{xf(x)} \geq \frac{\tilde{F}(x)}{x\tilde{f}(x)}$ for all $x \in [0, 1]$, and $\frac{F(x)}{xf(x)} > \frac{\tilde{F}(x)}{x\tilde{f}(x)}$ for all $x \in [0.25, 1]$. However, \tilde{F} first-order stochastically dominates F does not guarantee that $\frac{F(x)}{xf(x)} \geq \frac{\tilde{F}(x)}{x\tilde{f}(x)}$.

3.2. Remark: comparison with standard auction setups

Our setup differs from the standard auction setups due to the kind of information that is available to the advertisers in our platform (see Section 2.2). Notably, our advertisers can only set a single bid for the whole publisher – this bid applies to all users in that publisher. Moreover, this bid is held constant over a given time period.

Therefore, our advertisers do not bid for each user’s impression, but bid for a *distribution* of heterogeneous impressions. At the time the bid is placed, the bidder’s valuation is a random variable, *as opposed to a realized, fixed number*. To form the optimal bid, the bidder has to first integrate out his belief about the randomness in this valuation and obtains the expected profit equation in 1.

In comparison, in an online *ad exchange*, advertisers are able to bid for each impression in real-time, programmatically. The standard auction setup then applies to such online ad exchanges because we can treat each impression as having a known value to the advertiser at the time he forms his bid.

The final distinction here is that bidders do not know the identities of competing bidders. In contrast, in a keyword search advertising, the bidder can search for the keyword and find out *who* eventually won the top slots. The intermediary ensures *mutual anonymity* between participating advertisers (see Figure 2). Therefore, we will later take a *mean-field approximation* approach (Weintraub et al. (2005)) to modeling strategic interaction among bidders – by assuming that agents are best-responding to the correct distribution of opponents’ bids, instead of best-responding to opponents’ bidding strategies.

4. THE ECONOMETRIC MODEL

In this section, we develop an econometric model based on the previous theoretical model. The goal is to recover the advertiser’s valuation v using observed bids. In particular, v is a random variable, and we provide an estimator for $\mathbb{E}[v]$. In

the empirical application, we will estimate one such valuation parameter for 216 different advertiser-publisher pairs, obtaining 216 parameter estimates.

4.1. The econometrician’s uncertainty

In our previous theoretical model, there is no econometrician’s uncertainty. Recall that there is then a single optimal bid that is consistent with the model. As such, we are not able to rationalize the data where observed bids could be changing over time. Therefore we need to incorporate the econometrician’s uncertainty in order to fit the data to our model.

Before the advertiser sets the bid, he observes a shock to the supply curve ϵ . The advertiser then sets his bid to maximize the expected profit equation below:

$$(7) \quad U(p; \epsilon) = \mathbb{E}[(v - p)(Q(p) + \epsilon)]$$

We do not observe this shock. This represents the econometrician’s uncertainty. Moreover ϵ is unanticipated by the advertiser, as such, the expectation above is only taken with respect to the joint distribution of $(v, Q(p))$.

4.2. Partially observed valuation

Now for the purpose of estimation, we want to incorporate additional information that we (the econometrician and the intermediary) has about the advertiser’s valuation. We do this by formulating the valuation as being partially observed.

Specifically, let $v = x + \xi$, where x is the random variable representing the belief about how much a user acquired from the publisher would spend on *in-app purchases* in the first 14 days after acquisition; ξ is the residual unobserved valuation.¹⁵ The advertiser’s expected profit from bidding p can now be written as:

$$(8) \quad U(p; \epsilon) = \mathbb{E}[(x + \xi - p)(Q(p) + \epsilon)]$$

¹⁵For instance, ξ includes the positive network externalities of an acquired user (Ho et al. (2012)): – a new user either induces more new users to join, or induces existing users to spend more on in-app purchases. Mobile apps also have multiple channels to monetize their acquired user-base besides selling in-app purchases. The advantage of our estimation procedure is that we do not need to have data on all the other monetization channels of the advertiser.

Now the probability distribution of x is known to us, and ξ is a random variable with a probability distribution that is only known to the advertisers. The intermediary provides a service whereby the Return on Investment (ROI) of advertising is calculated with respect to how much acquired users spend on in-app purchases. In-app purchases are the additional content that users can buy within the apps.¹⁶ This a common service provided by most advertising platforms including Google’s Adwords.

Typically, users’ spending are tracked for 14 days post-acquisition. After 14 days, advertisers can then re-optimize their bids according to their 14-days ROI of advertising. As we would see in Section 6.2, user’s spending on in-app purchases typically decay quickly, so using the cutoff point of 14 days is appropriate. The advertisers can use these tracking data to inform their valuations.

4.3. Moment restriction

Since the advertiser maximizes his expected profit, the first-order condition of expected profit maximization $\frac{\partial U(p;\epsilon)}{\partial p}\Big|_{p^*} = 0$ gives rise to $\mathbb{E}[(x + \xi - p^*)\frac{\partial Q(p)}{\partial p}\Big|_{p^*} - Q(p^*) - \epsilon] = 0$. Now by assuming that ϵ has zero mean, the observed bid of the advertiser, p^* , satisfies Equation (9) below. Equation (9) is known as a population moment condition.

$$(9) \quad \begin{aligned} \mathbb{E}_\epsilon \left[\mathbb{E} \left[(x + \xi - p^*) \frac{\partial Q(p)}{\partial p} \Big|_{p^*} - Q(p^*) - \epsilon \mid \epsilon \right] \right] &= 0 \\ \implies \mathbb{E} \left[(x + \xi - p^*) \frac{\partial Q(p)}{\partial p} \Big|_{p^*} - Q(p^*) \right] &= 0 \end{aligned}$$

The expectation in Equation 9 is taken with respect to the advertiser’s belief about the joint distribution of $(x, \xi, Q(p))$. We assume that we can simulate this belief and hence, we can draw independent samples from this joint distribution. This allows us to compute the sample average corresponding to the left-hand side of Equation 9. Intuitively then, our estimation procedure involves finding ξ that best sets this sample moment condition to zero.

Now we show how to draw independent samples from the joint distribution of all the random variables in Equation 9. Recall that $Q(p) = F(s \cdot p)^N \chi s$. At the observed bid p^* , $Q(p^*)$ is a random variable, whose realization is just the observed

¹⁶They are subjected to 30% revenue sharing between the app developers and the app store. They include subscriptions and premium content (such as in-game currency or bonus content).

number of users acquired by the advertiser per period. Moreover, from Equation 10 below, $\frac{\partial Q(p)}{\partial p} \Big|_{p^*}$ is a random variable that is a constant multiple of $Q(p^*)$.

$$(10) \quad \frac{\partial Q(p)}{\partial p} \Big|_{p^*} = s \cdot N \cdot \frac{f(sp^*)}{F(sp^*)} \cdot Q(p^*)$$

Here we take s to be the average score of the advertiser over the sample period, which is estimated as the number of installs over the number of impressions won during the sample period. The case where s is a random variable is covered in Section 4.7. Moreover, we will take N to be the number such that $F(s \cdot p^*)^N = r$, where r is the observed probability that the advertiser wins an impression, which is estimated as the number of impressions won by the advertiser over all the impressions supplied by the publisher, during the sample period. The case where N is a random variable is covered in Section 4.7.

Now $F(\cdot)$ and $f(\cdot)$ are the CDF and PDF associated with the advertiser's belief about the score-weighted bids of his competitors. We will take $F(\cdot)$ to be the empirical distribution of score-weighted bids, observed over the sample period. Specifically, we fit a log-normal distribution to the observed score-weighted bids.¹⁷

This is the so-called *mean-field approximation*, where bidders best-respond to the (correct) empirical distribution of opponents' bids, instead of best-responding to opponents' bidding strategies. The justification for this is laid out in Section 3.2.

For convenience, we now denote $q \equiv Q(p^*)$ and $q' \equiv \frac{\partial Q(p)}{\partial p} \Big|_{p^*}$. By the law of iterated expectations, Equation (9) is equivalent to $\mathbb{E}[\mathbb{E}[(x + \xi - p^*)q' - q | x, q', q]] = 0$, which in turn is equivalent to the following.

$$(11) \quad \mathbb{E}[(x + \mathbb{E}[\xi | x, q', q] - p^*)q' - q] = 0$$

We see in Equation 10 that q' is a deterministic function of q . Hence it holds true that $\mathbb{E}[\xi | x, q', q] = \mathbb{E}[\xi | x, q]$. The population moment condition is then:

¹⁷First, we identify the competitors of the advertiser, which we denote by $\mathcal{K} = \{1, \dots, k, \dots, K\}$. The competitors of advertiser i are identified as the set of advertisers (excluding i itself) who has acquired at least one user from the publisher in the sample period. Secondly, we estimate the scores s_k for each $k \in \mathcal{K}$. Thirdly, we obtain $F(\cdot)$ as the CDF of the **log-normal distribution** fitted to the data $(s_k \cdot p_{kt}^*)_{k=1, \dots, K; t=1, \dots, T}$. Here, p_{kt}^* is the observed bid of advertiser k at time t .

$$(12) \quad \mathbb{E}[(x + \mathbb{E}[\xi|x, q] - p^*)q' - q] = 0$$

In the following sections, we will specify $\mathbb{E}[\xi|x, r]$, which then allows us to compute the sample analog of Equation (12).

4.4. Unobserved valuation

We now specify the model for ξ , the unobserved component of the valuation. Motivated by the fact that ξ takes only *positive* value, we assume that the conditional expectation of ξ follows Equation (13) below.

$$(13) \quad \mathbb{E}[\xi|x, q] = e^{\beta_0 + \beta_1 x + \beta_2 q}$$

The unknown parameters that we will later estimate in Section 4.6 are $\beta = (\beta_0, \beta_1, \beta_2)$. Substituting Equation (13) into the population moment condition in Equation (12):

$$(14) \quad \mathbb{E}[(x + e^{\beta_0 + \beta_1 x + \beta_2 q} - p^*)q' - q] = 0$$

4.5. Conditional moment restrictions

There are additional moment conditions that we have not utilized.¹⁸ When the advertiser forms its expected profit, it uses the information set \mathcal{F} to form such expectation. In our model, we assume that the advertiser's information set is $\mathcal{F} = \{\tilde{x}, \tilde{q}\}$, that is, the information set consists of \tilde{x} , the previous period's realization of user's spending; and \tilde{q} , the number of users acquired in the previous period.

The advertiser's expected profit as a function of bid is then:

$$(15) \quad U(p) = \mathbb{E}[(x + \xi - p)Q(p)|\mathcal{F}]$$

At the advertiser's optimal bid p^* , we have the following *additional* moment conditions:

¹⁸Conditional moment restriction is an important class of econometric models Chamberlain (1987); Newey (1985, 1993), and can be estimated using GMM (Hansen (1982); Hansen and Singleton (1982)).

$$\begin{aligned}\mathbb{E} \left[\left((x + \xi - p^*) \frac{\partial Q(p)}{\partial p} \Big|_{p^*} - Q(p^*) \right) \tilde{x} \right] &= 0 \\ \mathbb{E} \left[\left((x + \xi - p^*) \frac{\partial Q(p)}{\partial p} \Big|_{p^*} - Q(p^*) \right) \tilde{q} \right] &= 0\end{aligned}$$

Plugging in the specification for the unobserved valuation, $\mathbb{E}[\xi | x, q] = e^{\beta_0 + \beta_1 x + \beta_2 q}$, we can rewrite the above moment conditions as:

$$(16) \quad \mathbb{E} \left[\left((x + e^{\beta_0 + \beta_1 x + \beta_2 q} - p^*) q' - q \right) \tilde{x} \right] = 0$$

$$(17) \quad \mathbb{E} \left[\left((x + e^{\beta_0 + \beta_1 x + \beta_2 q} - p^*) q' - q \right) \tilde{q} \right] = 0$$

4.6. Generalized Method of Moments (GMM)

Estimating the valuation now boils down to estimating the vector of parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ using the moment conditions in (14), (16), (17). This is done using GMM. First, we turn the population moment conditions into sample moment conditions, as in (18) below. Then we find $\boldsymbol{\beta}$ that jointly minimizes these sample moment conditions.

$$\begin{aligned}\hat{m}_1(\boldsymbol{\beta}) &\equiv \frac{1}{T} \sum_{t=1}^T \left((x_t + e^{\beta_0 + \beta_1 x_t + \beta_2 q_t} - p_t^*) q'_t - q_t \right) \\ \hat{m}_2(\boldsymbol{\beta}) &\equiv \frac{1}{T} \sum_{t=1}^T \left((x_t + e^{\beta_0 + \beta_1 x_t + \beta_2 q_t} - p_t^*) q'_t - q_t \right) q_{t-1} \\ (18) \quad \hat{m}_3(\boldsymbol{\beta}) &\equiv \frac{1}{T} \sum_{t=1}^T \left((x_t + e^{\beta_0 + \beta_1 x_t + \beta_2 q_t} - p_t^*) q'_t - q_t \right) x_{t-1}\end{aligned}$$

Each t denotes a day. (x_t, q_t, q'_t) is an i.i.d. draw from the distribution of (x, q, q') . Now we have (i) x_t is the observed 14-days spending of an average user acquired at time t ; (ii) q_t is the observed daily number of users acquired at time t ; and (iii) $q'_t = s \cdot N \cdot \frac{f(sp^*)}{F(sp^*)} \cdot q_t$, where s, N and $F(\cdot)$ are defined as in Section 4.3. The observed advertiser's bid at time t is denoted by p_t^* .

Stacking the moment conditions as a vector, the parameters are estimated as in Equation 19 below, where \mathbf{W} is a positive-definite weighting matrix. We take \mathbf{W} to be the weighting matrix of a two-steps feasible GMM.

$$(19) \quad \hat{\boldsymbol{\beta}} = \operatorname{argmin}_{\boldsymbol{\beta}} \hat{\mathbf{m}}(\boldsymbol{\beta})^T \mathbf{W} \hat{\mathbf{m}}(\boldsymbol{\beta})$$

After obtaining an estimate $\hat{\boldsymbol{\beta}}$, we can compute the unobserved valuation as follows.

$$\hat{\mathbb{E}}[\xi] = \frac{1}{T} \sum_{t=1}^T e^{\hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{\beta}_2 q_t}$$

Finally, the estimator for the advertiser's valuation for a user is:

$$(20) \quad \hat{\mathbb{E}}[x + \xi] = \frac{1}{T} \sum_{t=1}^T (x_t + \hat{\mathbb{E}}[\xi])$$

4.7. Score and entry uncertainty

In this section, we show how to incorporate score and entry uncertainty. With score and entry uncertainty, the advertiser's belief about the supply curve he faces has the same form:

$$Q(p) = F(s \cdot p)^N \cdot \chi \cdot s$$

Except now that s and N are also random variables. The empirical distribution of score-weighted bids $F(\cdot)$ is now the lognormal fit of the data $(s_{it} \cdot p_{it}^*)$ for all competing bidder i , and across all time t , where s_{it} is the estimated conversion probability at time t of bidder i .

Here, s_{it} is calculated as the number of installs obtained by the advertiser i at time t divided by the number of impressions won by the advertiser i at time t . In the previous formulation, we use an average score instead of a time-varying score to reflect score uncertainty. Moreover,

The expected profit of the advertiser is now $\mathbb{E}[(x + \xi - p)(Q(p) + \epsilon)]$, where this expectation is taken with respect to the joint distribution of (s, N, x, ξ, χ) . The optimal bid maximizes this $\mathbb{E}[(x + \xi - p)(Q(p) + \epsilon)]$. As before, we have $\frac{\partial Q(p)}{\partial p} \Big|_{p^*} = s \cdot N \cdot \frac{f(sp^*)}{F(sp^*)} \cdot Q(p^*)$. However now $\frac{\partial Q(p)}{\partial p} \Big|_{p^*}$ is no longer a deterministic function of $Q(p^*)$.

Hence to sample from $\frac{\partial Q(p)}{\partial p} \Big|_{p^*}$, we use the following: $q'_t = s_t \cdot N_t \cdot \frac{f(s_t p_t^*)}{F(s_t p_t^*)} \cdot q_t$, where N_t , the number of competing bidders at time t , is given by $N_t = \log(r_t) / \log(s_t p_t^*)$, where r_t is the fraction of impressions won by the advertiser at time t out of the total impressions supplied by the publisher at time t . Recall that q_t is the number of users acquired by the advertiser at time t at the observed bid p_t^* .

5. DATA

As described in Section 2, our proprietary dataset comes from a San Francisco-based start-up, which is an intermediary who operates an online advertising platform that serves the purpose of mobile app-to-app advertising. The full dataset contains 7,613,445 observations or instances of user acquisitions from February 5, 2016 to March 22, 2016, spanning 47 days. A user acquisition is defined as the user installing the advertiser's mobile app.

During this period, 1,141 distinct advertisers have spent \$13,917,058 to acquire 7,613,445 (not necessarily unique) users located in 244 distinct countries or regions¹⁹ from 10,521 publishers. Both publishers and advertisers are mobile apps for either iOS or Android devices. The dataset are anonymized, but we know their app attributes such as genres, ratings, and languages, as scraped from Apple and Google's app stores.

As mentioned, the novelty of our dataset is that we also have information on how much users subsequently spend on in-app purchases after being acquired by the advertisers. Specifically, we tracked the users' daily spending on the advertisers' in-app purchases for 31 days after being acquired. However a subset of advertisers declined to disclose information on users' spending, and we omit these advertisers altogether, leaving us with 5,540,685 observations or instances of user acquisitions. The total amount of in-app purchases made by all acquired users in this sample is \$679,470. Within this sample, there are 1,126 distinct advertisers, who have acquired 4,973,931 unique users located from 10,037 publishers. The total number of impressions supplied by the publishers is close to 2.18 billion. The total ad spending is \$12,264,562.

Finally for the structural estimation, we select 222 advertiser-publisher pairs according to the criteria that each pair is sufficiently experienced with the bidding process.²⁰ In these 222 pairings, an advertiser may pair with more than one

¹⁹As denoted by their ISO 3166-1 alpha-2 codes.

²⁰There must be at least \$500 total payment from the advertiser to the publisher. The advertiser must acquire at least one user every day from the publisher. The observed valuation,

publisher and vice versa. For each pair, we separately estimate $\mathbb{E}[v]$ according to the GMM estimator in Section 4, thereby obtaining 222 parameter estimates of the mean valuation, $\mathbb{E}[v]$. Note that when we estimate each advertiser’s belief about how competitive the market is (the CDF $F(\cdot)$ in Equation 10), we have to use the full dataset.

In the Appendix 9.2, we further elaborate on the summary statistics concerning this selected sub-sample.

5.1. Variables Description and Construction

Each time period corresponds to one day, with a total of $T = 47$ days, spanning from February 5, 2016 to March 23, 2016. To implement the estimation procedure in the previous section (i.e. Equations 18), we need time-series i.i.d realization from (x, q) . We now define how we sample $(x_t, q_t)_{t=1}^T$ from (x, q) .

For each pair of advertiser i and publisher j , $(x_t)_{t=1}^T$ is the 14-days cumulative spending (on in-app purchases) of a user who was acquired at time t by advertiser i from publisher j . If there is more than one such users, we take x_t to be the average among users acquired at time t .

For each pair of advertiser i and publisher j , $(q_t)_{t=1}^T$ is the observed daily number of user acquisitions, that is, the number of users from publisher j who install advertiser i ’s app during time t .

6. ESTIMATION RESULT

We estimate $\mathbb{E}[v]$ separately for each of the 222 advertiser-publisher pair in our sample, obtaining 222 parameter estimates. Among these parameter estimates, 6 of them are poorly identified.²¹ As a result, the estimated $\mathbb{E}[\xi]$ for these pairs are much larger than 1.5 times the interquartile range. We dropped these pairs from the subsequent analysis, retaining 216 pairs. Among these pairs, there are 40 distinct advertisers.

We plot the histogram of these valuation estimates in Figure 3.

By comparison, we plot in Figure 4, the histogram of $\hat{\mathbb{E}}[x] = \frac{1}{T} \sum_{t=1}^T x_t$, which is the component of the overall valuation attributed to the user’s spending in the

$\hat{\mathbb{E}}[v] = \sum_{t=1}^T v_t$, of the advertiser for acquiring a user from the publisher is at least \$0.10. See the next section for construction of $\hat{\mathbb{E}}[v]$.

²¹For these 9 pairs, $\frac{\partial Q(p)}{\partial p}|_{p^*}$ in the moment condition (9) is almost zero.

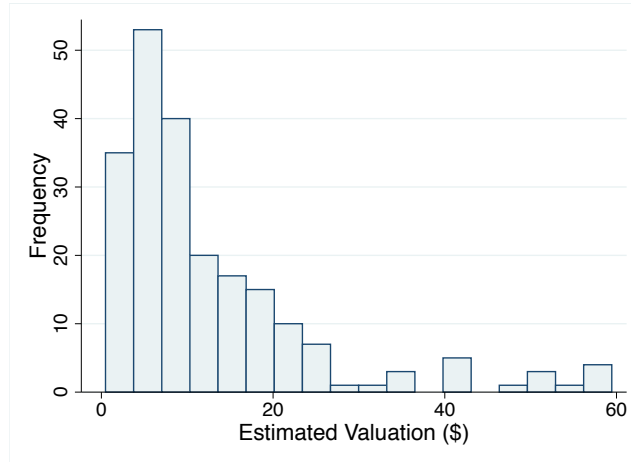


FIGURE 3. Histogram of estimated valuation for a new user across the 216 advertiser-publisher pair. For each pair, we recover the advertiser’s valuation $\mathbb{E}[v]$. The mean is \$12.69, the median is \$8.55, and the interquartile range is [\$5.10, \$15.53].

first 14 days of acquisition. We see that the recovered valuation differs drastically from the observed user’s spending.

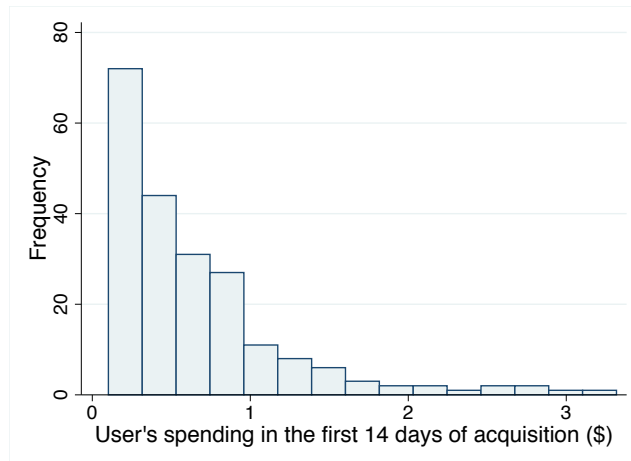


FIGURE 4. Histogram of $\hat{\mathbb{E}}[x] = \frac{1}{T} \sum_{t=1}^T x_t$, the average user’s spending in the first 14 days of acquisition, across the 216 advertiser-publisher pairs. The mean across all pairs is \$0.65, the median is \$0.48, the interquartile range is [\$0.26, \$0.84].

Now we see in Figure 5 that the estimated valuation for a new user is significantly greater than how much the advertisers end up paying. The observed successful

bids in the sample averaged to \$4.09 (the 25th, 50th and 75th percentiles are \$2.50, \$3.50, \$5.00). More formally, we estimate the margin: $\mathbb{E}[v - p^*]$, for each pair. The average of estimated margins is \$8.61, and the median margin is \$4.64. Our result strongly suggests that advertisers are paying well below their valuation, and that there is room for more transfer to the publisher and the intermediary.

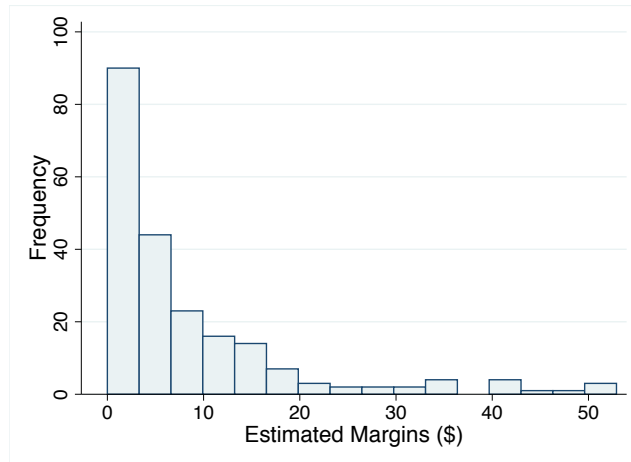


FIGURE 5. Advertisers are paying well below their valuation. Histogram of $\hat{\mathbb{E}}[v - p]$, the estimated margin across the 216 advertiser-publisher pairs. The mean is \$8.61, the median is \$4.64, while the interquartile range is [\$1.87, \$10.52].

6.1. Composition of valuation

In this section, we compute the fraction of the advertiser’s valuation not accounted for by in-app purchases, that is, $\hat{E}[\xi]/\hat{E}[v]$. We compute this fraction for each advertiser-publisher pair, and plot this as a histogram in Figure 6. We find that unobserved valuation dominates advertiser’s valuation. In particular, unobserved valuation accounts for 91.6% of the (average) advertiser’s valuation. In another words, user’s spending on in-app purchases account for roughly less than 10% of the (average) advertiser’s valuation for a new user.

6.2. User’s lifetime spending

In this section, we quantify and predict the lifetime spending on in-app purchases by the users in our sample. We want to determine whether the large unobserved valuations we found in the previous section can be attributed to users’ spending beyond the first 14 days of being acquired.

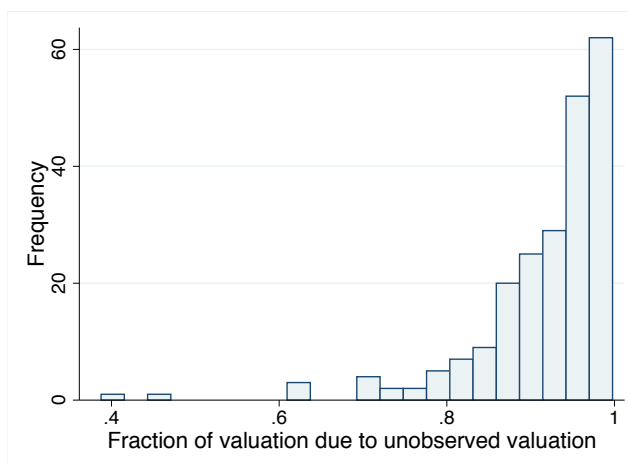


FIGURE 6. Histogram of the fraction of advertiser’s valuation attributed to unobserved valuation, among the 216 advertiser-publisher pairs. The mean across these pairs is 0.916, and the interquartile range is [0.88, 0.97].

Although we track users’ spending until time $t_0 = 14$, we observe the **rates of decay** of users’ spending, which can then tell us users’ lifetime spending beyond t_0 . *Specifically, for each advertiser-publisher pair, we will estimate $\mathbb{E}[w]$, which is the estimated or predicted lifetime spending (beyond the first 14 days) of an average user that the advertiser acquired from the publisher.* The details for this computation is relegated to Appendix 9.4.

In Figure 7, we plot the histogram of $\hat{\mathbb{E}}[w]$. We see that for many advertiser-publisher pairs, the spending of acquired users decay so fast that tracking users spending for the first 14 days of acquisition is sufficiently informative for the advertisers.

6.3. Advertisers and publishers’ attributes

In this section, we merge a dataset containing the attributes of the advertisers, which were scraped from the Google and Apple app stores at the beginning of the sample period. The list of attributes and their definitions are tabulated in Table 1.

Among the 216 advertiser-publisher pairs, only three of the publishers, and two of the advertisers, belong to non-gaming app genres. Even though in the full dataset, non-gaming apps have a significant presence, accounting for 1,280 unique

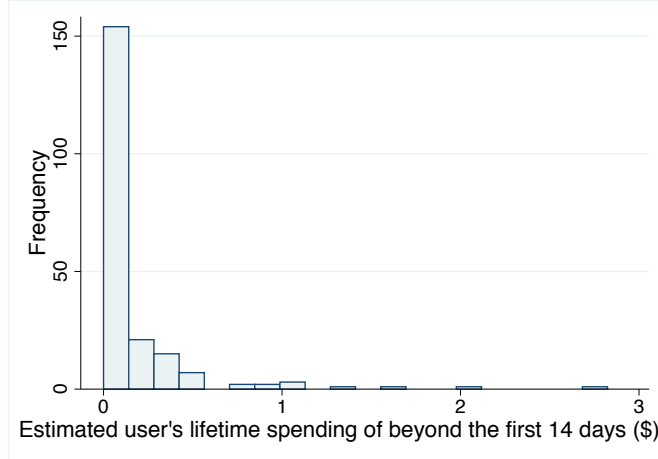


FIGURE 7. Histogram of $\hat{\mathbb{E}}[w]$, the predicted average user’s lifetime spending beyond the first 14 days. Across the 216 advertiser-publisher pairs, the mean is \$0.55, while the interquartile range is [\$0.00, \$0.19].

publishers’ apps out of the 10,521²², the major advertisers and publishers, as selected by advertising budget and other criteria in Section 5, all are gaming and entertainment apps.

The intermediary slices up a single app publisher into three publishers that an advertiser can bid separately on, by grouping users according to their geographical locations. For instance, if *Country Tier 2* = 1, it means that all the users from that publisher are from countries located in continental Europe. With the exception of *Country Tier* variables, all the other attributes in Table 1 are specific only to the advertisers.

We see in Table 2 that on average, casino-type app developers (e.g. poker, bingo, roulette, solitaire, blackjack apps) are willing to pay \$27.16 for a new user, the highest of any group. We compare this to role-playing-type app developers, who are willing to pay only \$6.29 for a new user. The difference is striking especially when we look at the pattern of users’ spending. From Table 3, these role-playing-type apps on average acquire a user that spends \$0.69, versus \$0.64 for casino-type apps – this difference is negligible.

On average, advertisers are willing to pay \$14.22 for a user from *Country Tier 1*, the highest of the three geographic regions. In comparison, advertisers are

²²Non-gaming genres in the full dataset include health & fitness, music, weather, photography, social networking, shopping, travel & navigation, news, utilities, productivity, etc

Variables	Definition
<i>Country Tier 1</i>	Target users originating from US, AU, CA, GB and NZ
<i>Country Tier 2</i>	Target users originating from DE, FR, IT, NL, BE, NO, SE, CH, ES, RU, FI, PT
<i>Country Tier 3</i>	Target users originating from the rest of the world
<i>English</i>	Whether the advertised app is available in the English language
<i>Chinese</i>	Whether the advertised app is available in the Chinese language (simplified or traditional)
<i>User Ratings</i>	Average user ratings of the advertised app in the app store
<i>Number of User Ratings</i>	Number of users who have rated the advertised app in the app store
<i>Age 12+</i>	Whether the content rating of the advertised app is age 12 and above
<i>Casino</i>	Whether the advertised app belongs to the “casino” or “card” genres
<i>Simulation, Role-playing, Strategy</i>	Whether the advertised app belongs to the “simulation”, “role-playing” or “strategy” genres

TABLE 1. Attributes characterizing the advertiser-publisher pairs. Note that genres are not mutually exclusive, so an app could fall into both role-playing and simulation genres.

	mean (\$)	median (\$)	sd	count
Country Tier 1	14.22	8.81	13.03	151
Country Tier 2	4.21	2.99	3.41	15
Country Tier 3	10.62	9.20	9.33	50
English	9.93	7.32	8.98	114
Chinese	8.24	6.02	6.44	119
User Ratings => 4.5 stars	11.24	8.44	9.90	109
Age 12+	15.61	11.85	14.22	101
Age < 12	10.13	6.83	9.22	115
Simulation	5.85	4.95	3.49	66
Casino	27.16	21.15	15.69	40
Role-playing	6.29	5.93	3.59	64
Strategy	11.22	8.91	7.72	56

TABLE 2. Summary statistics of the estimated valuation for a new user by attributes of the advertiser-publisher pairs as defined in Table 1.

willing to pay \$10.62 for a user coming from *Country Tier 3*, which consists predominantly of users from China, South Korea, Japan and India. However,

	mean (\$)	median (\$)	sd	count
Country Tier 1	0.72	0.59	0.57	151
Country Tier 2	0.43	0.31	0.37	15
Country Tier 3	0.61	0.37	0.81	50
English	0.58	0.45	0.49	114
Chinese	0.69	0.41	0.73	119
User Ratings => 4.5 stars	0.60	0.48	0.48	109
Age 12+	0.57	0.41	0.48	101
Age < 12	0.77	0.58	0.72	115
Simulation	0.48	0.32	0.48	66
Casino	0.64	0.55	0.44	40
Role-playing	0.69	0.52	0.57	64
Strategy	0.86	0.54	0.89	56

TABLE 3. Summary statistics of the average user’s spending in the first 14 days of acquisition, $\mathbb{E}[x]$, by attributes of the advertiser-publisher pairs as defined in Table 1.

advertisers are only willing to pay \$4.21 for a user from *Country Tier 2*, which consists of users from continental Europe.

7. COUNTERFACTUALS

7.1. Bid-shading

One managerial relevance of our model is the ability to estimate how much advertisers underbid relative to their valuations. We have showed that bid-shading is large and prevalent, but it turns out, bid-shading varies across segments of the market, and it is more severe in some. This has implication for the intermediary, who has a dedicated sales team who seeks out new advertisers to participate in the platform. The knowledge of which segment of the market is less competitive, and where the intermediary is forgoing the most profits, is useful for this task.

In Table 4, we examine which kind of advertisers shade their bids the most. We see in the columns of Table 4 that the coefficient on *Casino* is significantly positive. This implies that the market for impressions coming from these *Casino* apps is severely uncompetitive. The group of advertisers bidding on these publishers are able to shade their bids significantly. Moreover, publishers whose content are more mature also suffer from bid-shading and subsequently derive less revenue than they could have had advertisers bid competitively.

	(1) Dollar bid-shade (\$)	(2) Percentage bid-shade
Country Tiers=2	-4.772*** (1.386)	-4.417 (6.372)
Country Tiers=3	-2.090 (1.622)	-4.860 (4.906)
English	-1.421 (2.105)	-1.156 (3.835)
Chinese	0.201 (1.473)	2.680 (3.016)
Age 12+	1.245* (0.711)	2.647* (1.378)
User Ratings	-1.484 (2.722)	-6.034 (5.779)
Number of user ratings	-0.672* (0.377)	-1.093** (0.493)
Casino	9.008*** (1.869)	10.93*** (3.313)
Role-playing	-2.643 (2.475)	4.068 (5.037)
Simulation	0.795 (1.384)	7.316* (4.269)
Strategy	0.714 (2.594)	-3.411 (4.478)
R-squared	0.259	0.182
N	200	200

Cluster-robust standard errors in parenthesis.

Each cluster is a unique advertiser. There are 39 advertiser-cluster in total

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 4. The dependent variable in Column (1) is $\hat{\mathbb{E}}[v] - p^*$, while in Column (2), it is $(\hat{\mathbb{E}}[v] - p^*) / \hat{\mathbb{E}}[v] \times 100$. Constants not displayed.

7.2. Counterfactual bids: increasing competition

In this section, we use our estimated models to quantify the optimal responses (counterfactuals) of advertisers as a result of changes in the underlying model parameters.

More specifically, we want to determine which publisher the intermediary should target the most, in terms of bringing in new advertisers to bid on that publisher. In practice, the intermediary has a sales team who pitches the platform to potential advertisers. Moreover, once these advertisers joined the platform, the sales team can then frame which publishers the new advertisers can bid on. Therefore using our tools here, we can guide this decision-making process of where to insert competition. The technical procedure for calculating the counterfactual bids is deferred to Appendix 9.3.

Our counterfactual exercise is as follows: suppose each publisher were to gain an additional competing bidder, what are the advertisers' profit-maximizing bids given this new parameter? Using the procedure in Appendix 9.3, we compute the advertisers' counterfactual bids, and plot them as a histogram in Figure 8, which shows the distribution of counterfactual bids for the 216 advertiser-publisher pairs. Further, we show in Figure 9 that advertisers' bids would increase by about 14.4% on average. As evident in Figure 9, increasing an additional bidder has different effects across publishers – it is more effective for some publishers.

7.3. Score boosting: artificial competition

One tweak to the mechanism that the intermediary can perform is to *boost* the quality scores of those bidders who would otherwise have a lower chance of winning. This is inefficient in the sense that the impression does not always go to the bidder who promises the highest expected payment. On the surface then, the intermediary would suffer a loss. However, taking the endogeneity of bids into consideration, the other bidders with higher scores would now bid more aggressively due to this artificial increase in competition. Now if these advertisers increase their bids sufficiently that it offsets fewer number of user acquisitions, then overall, the intermediary would receive a higher revenue.

Our counterfactual experiment proceeds as follows. There are two group of advertisers, small and large. The large advertisers are the 40 advertisers we selected (among the 216 advertiser-publisher pairs) where their valuations are estimated. The second group are all the remaining advertisers.

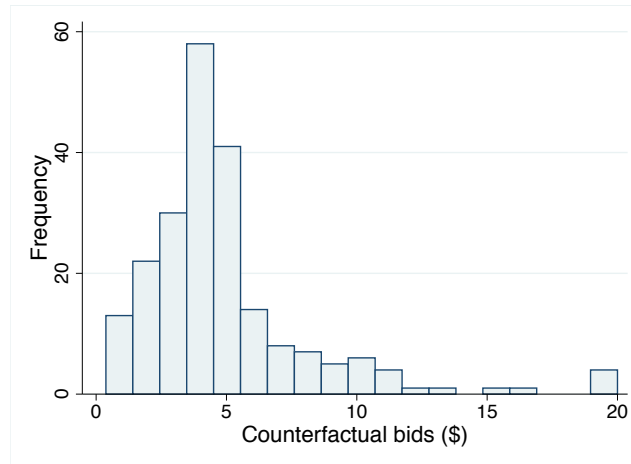


FIGURE 8. If each publisher were to gain an additional bidder, then the resulting counterfactual bids has the distribution above. The mean is \$4.91, and the interquartile range is [\$3.26, \$5.48]. In comparison, the observed bids have a mean of \$4.06, and an interquartile range of [\$3.00, \$4.87].

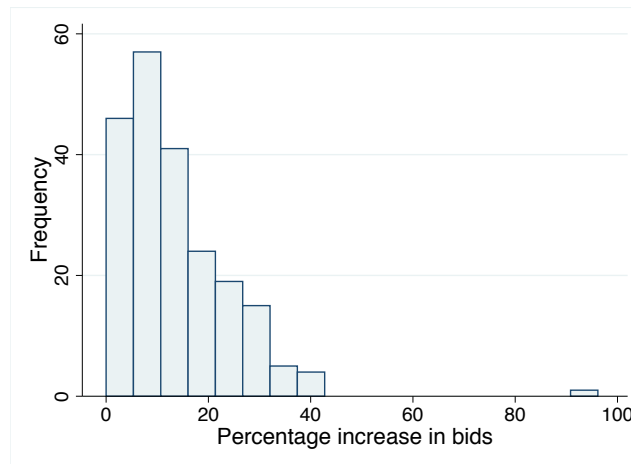


FIGURE 9. On average, an advertiser's bid would increase by about 14.47%, with a median of 10.92%, and interquartile range of [6.33%, 19.27%],

For each publisher among the 216 advertiser-publisher pairs, we look at all the small advertisers whose score-weighted bids lie in the bottom y -percentile of the

empirical score-weighted bids. We then boost the average score of these advertisers by a certain percentage. This results in a new empirical distribution of score-weighted bids $F_c(\cdot)$.

For each of the large advertiser, we calculate the counterfactual bid \tilde{p} in response to this new empirical distribution of bids $F_c(\cdot)$. This is done by implementing the *iterative procedure* described in Appendix 9.3. We then compute the expected payment to the intermediary: $\mathbb{E}[\tilde{p}\tilde{Q}(\tilde{p})]$, where $\tilde{Q}(\tilde{p}) = \tilde{F}(s \cdot \tilde{p})^N \chi s$ is the daily number of users that the advertiser would acquire in the counterfactuals. Here, $\tilde{F}(\cdot)$ is the counterfactual empirical distribution of score-weighted bids. Note that χ does not depend on p so it can be estimated using the original non-counterfactual dataset as $\chi = Q(p^*) / (sF(s \cdot p^*)^N)$.

In Table 5, we see that creating artificial competition by means of boosting scores (instead of letting score be the best prediction of conversion probability), allows the intermediary to extract more revenue from the advertisers. For instance, in the second row, if we boost the scores of the bidders at the bottom 20th percentile by a factor of 1.2, the intermediary extracts \$889 more revenue from the large advertisers (Column 1). In fact, there is a large heterogeneity in how effective this policy is among the advertisers. Column 2 shows that 70% advertisers generated higher revenues, which means the policy resulted in losses for the remaining 30%. In practice then, the intermediary can design specific policies for each publisher, including not intervening for those publishers that would result in losses.

In Column (3), we see that the artificial increase in competition ‘stole’ users away from these large advertisers. As a result of the score-boosting, these advertisers won fewer impressions (as they are awarded to bidders with the boosted scores), and subsequently acquire fewer users. If bids do not change, then the intermediary’s revenue from these advertisers would have decreased. After bids are re-optimized (Column 4), the increase in bids is sufficient to bring about overall revenue increase for the intermediary.

<i>Experiments</i>	(1)	(2)	(3)	(4)
<i>Boost bottom 40th by 1.2</i>	\$914.77	64.4%	-2.84	\$0.76
<i>Boost bottom 20th by 1.2</i>	\$889.38	65.3%	-2.53	\$0.69
<i>Boost top 20th by 1.2</i>	\$81.6	59.0%	-2.52	\$0.68
<i>Boost middle 20th by 1.2</i>	\$373.2	60.8%	-2.40	\$0.71

TABLE 5. Counterfactual experiments that are conducted for all publishers. Column (1) reports the sum of revenue changes across all advertisers. Column (2): the percentage of advertisers whose revenue changes are positive. Column (3): the median decrease in the daily number of users acquired by those advertisers in which revenue changes are positive. Column (4): the median increase in bid among those advertisers with positive revenue changes.

8. QUANTIFYING RESALE MOTIVES

8.1. Resale of users' attention

We have seen the importance of advertisers' unobserved valuations. In this section, we single out one possible explanation for what goes on in the advertiser's unobserved valuation.

Due to the app-to-app advertising nature of our platform, an app developer can readily participate in both sides of the market, as both an advertiser and a publisher. This introduces *resale motive* which partly drives up the advertisers' valuations for a new user. More concretely, the advertiser perceives a user as a *durable good*, whereby a user generates a stream of diminishing monetary benefits in terms of in-app purchases. Now the advertiser then has the option to resell a user by selling its impression and attention to other advertisers. By reselling the attention of the user, the user would potentially be drawn away to other apps (due to the scoring procedure used by the platform mechanism, the ads that are shown to this user are often competing apps that are closely relevant to the user).

Below we show more evidence that the resale motives of advertisers increase their valuations for new users.

8.2. Quantifying resale motives

For each pair of advertiser i and publisher j , we construct the variable π_{ij} capturing how much advertiser i expects to resell a user that is acquired from publisher j . Because this variable is not observed, we estimate π_{ij} using Equation (21) below.

$$(21) \quad \pi_{ij} = \eta_{ij} \times \delta_{ij}$$

Where η_{ij} is the number of impressions that advertiser i 's app can extract out from a user that was acquired from publisher j ; and δ_{ij} is the revenue per impression that advertiser i expects for a user that was acquired from publisher j .

Both η_{ij} and δ_{ij} are still not observed because only a subset of advertisers participate as publishers within our intermediary. Moreover an advertiser could resell users' attention and impression using other intermediaries, therefore we do not observe the reselling activities of these advertisers. However, we *do* observe the corresponding variables defined for *publishers* in our dataset. That is, we observe the number of impressions that a given publisher's app extracts out from a user in a day; and the revenue per impression obtains by a publisher's app in our sample.

The idea here is to assume that η_{ij} and δ_{ij} are functions of \mathbf{X}_i and \mathbf{Y}_j , as in Equations (22), (23) below, where \mathbf{X}_i is the vector of attributes of the app where the impression was generated (such as the app's genre, language, users' rating, maturity content rating, file size, cumulative number of installs); \mathbf{Y}_j is the attributes that characterized user j 's impression. We take \mathbf{Y}_j to be a scalar attribute such that $Y_j = k$ if and only if user j belongs to Country Tier k for $k = 1, 2, 3$.²³

²³See Table 1 for definitions of Country Tiers. The intermediary allows advertisers to "slice" up a single publisher into multiple ones – by targeting different users from the same publisher according to users' geographical locations.

$$(22) \quad \eta_{ij} = \eta(\mathbf{X}_i, \mathbf{Y}_j) + \epsilon_{ij} = \sum_{k=1}^3 \mathbb{1}\{Y_j = k\} \mathbf{X}_i \boldsymbol{\theta}_k + \epsilon_{ij}$$

$$(23) \quad \delta_{ij} = \delta(\mathbf{X}_i, \mathbf{Y}_j) + \nu_{ij} = \sum_{k=1}^3 \mathbb{1}\{Y_j = k\} \mathbf{X}_i \boldsymbol{\gamma}_k + \nu_{ij}$$

The parameters $\boldsymbol{\theta}_k$ and $\boldsymbol{\gamma}_k$ above can be estimated using the sample of *publishers* in our dataset. More precisely, we fit the data $(\tilde{\eta}_{ij}, \tilde{\delta}_{ij}, \mathbf{X}_i, Y_j)_{i \in \mathcal{P}, j \in \{1,2,3\}}$ to Equations (22) and (23), where \mathcal{P} is the sample of publishers' apps; $\tilde{\eta}_{ij}$ is the observed number of impressions per type- j user for app i ; $\tilde{\delta}_{ij}$ is the observed ad revenue per impression received by app i , when the impression is generated by a type- j user.

In this estimation (fitting publishers' impressions data to Equations 22 and 23), we use the full set of publishers available in our dataset, which totals to 15,562 observations (i.e. 15,562 pairs of publisher/country-tier). This full set of publishers includes thousands of publishers that we withheld from the structural estimation (see Section 5).

Note that the resale value constructed here is an *underestimate* of the actual resale value. There are two reasons. (1) We only observe impressions that are sold using our platform, which specialized in the niche of in-app video advertising. Publishers could enlist multiple platforms to sell impressions. (2) Our sample period only covers 47 days from February 5, 2016 to March 22, 2016, we did not account for the lifetime number of impressions generated per user.

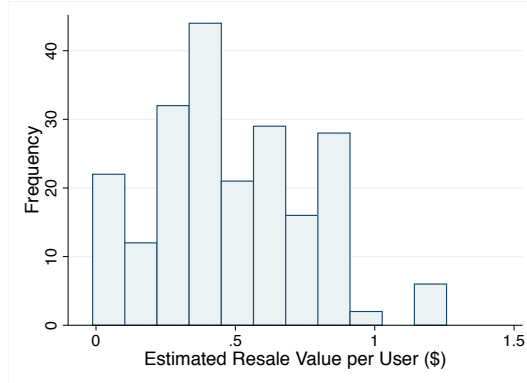


FIGURE 10. Estimated resale value of a user, $\hat{\pi}_{ij}$, across 216 advertiser-publisher pairs. The average resale value per user is \$0.49, compare to \$0.65 of in-app purchases generated per user.

We estimate the resale value per user for each of the 216 advertiser-publisher pairs in our sample. We plot the histogram in Figure 10, which has an average of \$0.49. We compare this to the average user’s spending on in-app purchases, which is \$0.65. Therefore, advertisers can gain a significant amount of monetary benefits by reselling acquired users, comparable to selling in-app purchases to these users.

Finally, we predict what would happen when we shut off the advertisers’ resale motives. In Column (4) of Table 6, if users have zero resale value, the average estimated unobserved valuation (across the 216 advertiser-publisher pairs) decreases to \$8.72 from \$12.26, while the median unobserved valuation (across the pairs) decreases to \$4.73 from \$7.89. This leads to a median *valuation* for a new user (across the pairs) of \$5.33, a decrease from \$8.56; and a median margin (across the pairs) of \$1.37, a decrease from \$4.66. In conclusion, the estimated resale value, $\hat{\pi}_{ij}$ seems to explain why advertisers have such high valuation for new users.

	(1)	(2)	(3)	(4)
	Unobserved Valuations, $\hat{\mathbb{E}}[\xi]$ (\$)			
Impressions per user, η_{ij}	22.05*** (7.712)			
Revenue per impression, δ_{ij}		64.34*** (14.82)		
Resale value per user, $\eta_{ij} \times \delta_{ij}$			24.24*** (4.817)	7.123*** (2.230)
Publishers’ Fixed Effects	No	No	No	Yes (134)
R-squared	0.175	0.279	0.311	0.934
N	216	216	216	216

Cluster-robust standard errors in parenthesis.

Each cluster is a unique advertiser. There are 40 advertiser-cluster in total

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 6. Estimated resale value per user (constructed using Equations 21, 22, 23) significantly and positively predicts higher unobserved valuation. Constants not shown.

	(1)	(2)	(3)	(4)
	Relative unobserved valuation, $\hat{\mathbb{E}}[\xi]/\hat{\mathbb{E}}[x + \xi]$			
Impressions per user, η_{ij}	0.0562 (0.0435)			
Revenue per impression, δ_{ij}		0.326*** (0.0572)		
Resale value per user, $\eta_{ij} \times \delta_{ij}$			0.113*** (0.0186)	0.0647* (0.0379)
Publishers' Fixed Effects	No	No	No	Yes (134)
R-squared	0.0237	0.149	0.142	0.712
N	216	216	216	216

Cluster-robust standard errors in parenthesis.

Each cluster is a unique advertiser. There are 40 advertiser-cluster in total

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 7. The dependent variable is the fraction of overall valuation due to unobserved valuation. The variable *Resale value per user* is significantly and positively related to relative unobserved valuation. Constants not shown.

9. APPENDIX

9.1. Proofs

We now present the proofs for Proposition 1 and 2. Recall that the first-order condition of profit maximization is:

$$(24) \quad \mathbb{E} \left[(v - p^*) \frac{\partial Q(p)}{\partial p} \Big|_{p^*} - Q(p^*) \right] = 0.$$

$$(25) \quad Q(p) = \alpha(p) \cdot \chi \cdot s,$$

Where $\alpha(p) = F(s \cdot p)^N$, and $F(\cdot)$ is the CDF representing the advertiser's belief about the score-weighted CPI bids of other competing bidders, and s is the conversion probability for the advertiser, which we assumed to be fixed and does not depend on p . The derivative of $Q(p)$ is

$$\frac{\partial Q(p)}{\partial p} \Big|_{p^*} = \frac{\partial \alpha(p)}{\partial p} \Big|_{p^*} \cdot \chi \cdot s$$

Rewrite (24) using conditional expectation:

$$\mathbb{E} \left[\mathbb{E}[(v - p^*)|\chi] \cdot \frac{\partial Q(p)}{\partial p} \Big|_{p^*} - Q(p^*) \right] = 0.$$

Substituting the derivatives into the equation, we have

$$\mathbb{E} \left[\mathbb{E}[(v - p^*)|\chi] \cdot \frac{\partial \alpha(p)}{\partial p} \Big|_{p^*} \cdot \chi - \alpha(p^*) \cdot \chi \right] = 0,$$

Rewriting gets

$$\mathbb{E}[\chi \cdot \mathbb{E}[(v - p^*)|\chi]] \cdot \frac{\partial \alpha(p)}{\partial p} \Big|_{p^*} = \alpha(p^*) \mathbb{E}[\chi]$$

and

$$\frac{\alpha(p^*)}{\frac{\partial \alpha(p)}{\partial p} \Big|_{p^*}} = \frac{\mathbb{E}[\chi \cdot \mathbb{E}[(v - p^*)|\chi]]}{\mathbb{E}[\chi]}$$

Finally we have,

$$\frac{\alpha(p^*)/p^*}{\frac{\partial \alpha(p)}{\partial p} \Big|_{p^*}} = \frac{1}{p^*} \frac{\mathbb{E}[v\chi]}{\mathbb{E}[\chi]} - 1$$

Where $\rho = \frac{\text{Cov}(\chi, v)}{\mathbb{E}[\chi]\mathbb{E}[v]}$. Plugging in $\alpha(p) = F(s \cdot p)^N$, we then have:

$$(26) \quad \frac{F(sp^*)}{sNp^*f(sp^*)} = \frac{1}{p^*} \frac{\mathbb{E}[v\chi]}{\mathbb{E}[\chi]} - 1$$

The RHS of Equation 26 is strictly decreasing in p^* , and we now show that the LHS of Equation 26 is strictly **increasing** in p^* when $F(\cdot)$ is the CDF of a lognormal distribution. Without loss of generality, we normalize $N = 1$.

In particular, when $F(\cdot)$ is the CDF of the lognormal distribution with parameter (μ, σ) , that is, the random variable X drawn from the distribution F is such that $\log X$ is distributed $N(\mu, \sigma)$, then from Equation 26, we have:

$$(27) \quad \sqrt{2\pi}\sigma e^{\frac{(\mu - \log(sp))^2}{2\sigma^2}} \Phi \left(\frac{\log(sp) - \mu}{\sigma} \right) = \frac{\mathbb{E}[v]}{p} (1 + \rho) - 1$$

where $\Phi(\cdot)$ is the CDF of the standard Gaussian. Denote the LHS of Equation 26 as $L(p)$. Taking the partial derivative of $L(p)$ with respect to p , we have that $L(p)$ is strictly increasing in p if

$$(28) \quad 1 > \sqrt{\pi} e^{y^2} y \operatorname{erfc}(y)$$

Where $y \equiv \frac{\mu - \log(ps)}{\sqrt{2}\sigma}$, and $\operatorname{erfc}(y)$ is the complementary error function, i.e. $\operatorname{erfc}(y) = 1 - \frac{2}{\sqrt{\pi}} \int_0^y e^{-t^2} dt$. Now denote $l(y) \equiv \sqrt{\pi} e^{y^2} y \operatorname{erfc}(y)$. We then show that $l(y)$ is strictly increasing in y , and that $\lim_{y \rightarrow \infty} l(y) = 1$. Therefore, the inequality in 28 above is true for all y .

We will invoke two useful properties of the complementary error function.

$$(29) \quad \operatorname{erfc}(y) \geq \frac{2}{\sqrt{\pi}} \left(\frac{e^{-y^2}}{y + \sqrt{y^2 + 2}} \right)$$

$$(30) \quad \operatorname{erfc}(y) \rightarrow \frac{e^{-y^2}}{y\sqrt{\pi}}, \text{ as } y \rightarrow \infty$$

Equation 29 is a known lower bound of the complementary error function (Durrett, Probability: Theory and Examples, 3rd edition)²⁴, while Equation 30 comes from the known asymptotic expansion of $\operatorname{erfc}(y)$. From Equation 28 and 30, we can immediately see that $\lim_{y \rightarrow \infty} \sqrt{\pi} e^{y^2} y \operatorname{erfc}(y) = 1$.

Now $l'(y) = \sqrt{\pi} e^{y^2} (2y^2 + 1) \operatorname{erfc}(y) - 2y$, and plugging in the lower bound in Equation 29, we have $l'(y) > 0$ if $\frac{2(2y^2+1)}{\sqrt{y^2+2+y}} - 2y > 0$. This is true for all y .

Finally, we want to show that the LHS of Equation 27, i.e. converges to zero as $p \rightarrow 0_+$. That is, $\lim_{p \rightarrow 0_+} L(p) = 0$. Again this can be shown by invoking Equation 30:

$$\begin{aligned} L(p) &= \sqrt{2\pi}\sigma e^{\frac{(\mu - \log(sp))^2}{2\sigma^2}} \operatorname{erfc}\left(\frac{\mu - \log(ps)}{\sqrt{2}\sigma}\right) \\ &\rightarrow e^{\frac{(\mu - \log(sp))^2}{2\sigma^2}} \frac{\exp\left(-\left(\frac{\mu - \log(ps)}{\sqrt{2}\sigma}\right)^2\right)}{\frac{\mu - \log(ps)}{\sqrt{2}\sigma} \sqrt{\pi}}, \text{ as } p \rightarrow 0_+ \\ &= \frac{\sqrt{\frac{2}{\pi}}\sigma}{\mu - \log(ps)} \end{aligned}$$

Therefore, $\lim_{p \rightarrow 0_+} L(p) = 0$.

²⁴Also, see <http://mathworld.wolfram.com/Erfc.html>

9.2. Summary statistics for the selected sample

We will present some summary statistics based on the sample of 222 advertiser-publisher pairs over the sample period of 47 days. There are 41 unique advertisers and 165 unique publishers among the 222 pairs. On average, (i) each advertiser spends a total of \$5,719 to acquire users from each publisher; (ii) each advertiser acquires a total of 1,399 users from each publisher; (iii) each advertiser receives \$796 in revenue generated by users acquired from each publisher (tracking in-app purchases for 14 days after user acquisition).

Overall in our sample, a single advertiser on average spends \$31,260 throughout the sample period to acquire 7,577 number of users. On average, the advertiser receives \$4,311 revenue from in-app purchases made by these new users (tracking revenue for the first 14 days). The winning CPI bids averaged over all incidences of user acquisitions is \$4.09, with a standard deviation of 2.72. The total number of users acquired in our sample is 310,657.

9.3. Procedure for computing counterfactual bids

Here, we show how the counterfactual bid is computed as we increase the number of competing bidders N by one. We implement an iterative procedure to compute the counterfactual bids as follows. For each publisher, there are two sets of advertisers bidding on it: large \mathcal{L} , and small \mathcal{S} . The large advertisers are the 40 advertisers we selected (among the 216 advertiser-publisher pairs) where their valuations are estimated. At the n -th iteration:

- (1) Solve for the optimal bids $p_i^{(n)}$ for $i \in \mathcal{L}$, given the empirical distribution of bids $F_i^{(n)}(\cdot)$ using Equation 31.
- (2) Re-compute the empirical distribution of bids $F_i^{(n+1)}(\cdot)$ with $p_i^{(n)}$ replacing $p_i^{(n-1)}$ for $i \in \mathcal{L}$.

Particularly for this counterfactual exercise, we always set $\tilde{N} = N + 1$. For a given advertiser-publisher pair and the corresponding (intermediate) counterfactual empirical distribution $F^{(n)}(\cdot)$, the counterfactual bid $p^{(n)}$ satisfies Equation 31 below.

$$\begin{aligned}
 & \mathbb{E} \left[\chi \left((x + \xi - p^{(n)}) \frac{\partial \alpha^{(n)}(p)}{\partial p} \Big|_{p^{(n)}} - \alpha^{(n)}(p^{(n)}) \right) \right] = 0 \\
 (31) \quad & \implies \frac{(\mathbb{E}[\chi x] + \mathbb{E}[\chi \xi])}{\mathbb{E}[\chi]} = p^{(n)} + \frac{\alpha^{(n)}(p^{(n)})}{\frac{\partial \alpha^{(n)}(p)}{\partial p} \Big|_{p^{(n)}}}
 \end{aligned}$$

Since we observe the time-series of χ and x , we can then directly estimate $\mathbb{E}[\chi x]$ and $\mathbb{E}[\chi]$. Now using Equation 13, we can compute $\mathbb{E}[\chi \xi]$ as follows.

$$\begin{aligned}
\mathbb{E}[\chi\xi] &= \mathbb{E}[\mathbb{E}[\chi\xi|x, \chi]] \\
&= \mathbb{E}[\chi\mathbb{E}[\xi|x, \chi]] \\
&= \mathbb{E}[\chi e^{\beta_0 + \beta_1 x + \beta_2 q}] \\
\hat{\mathbb{E}}[\chi\xi] &= \sum_{t=1}^T (\chi_t e^{\hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{\beta}_2 q_t})
\end{aligned}$$

The parameters $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ are the ones we estimated between advertiser i and publisher j from the two-steps GMM in Equation 19. The time-series q_t is the daily number of users acquired by the advertiser at the original non-counterfactual bid p^* . It remains to specify $\alpha^{(n)}(p)$ and $\left. \frac{\partial \alpha^{(n)}(p)}{\partial p} \right|_p$, which are given by:

$$\alpha^{(n)}(p) = F^{(n)}(s \cdot p)^{\tilde{N}}$$

$$\left. \frac{\partial \alpha^{(n)}(p)}{\partial p} \right|_p = \tilde{N} F^{(n)}(s \cdot p)^{\tilde{N}-1} f^{(n)}(s \cdot p) s$$

9.4. Quantifying user's lifetime in-app purchases

We begin with:

$$w = \int_{t=t_0}^{\infty} e^{-\rho(t-t_0)} \tilde{w}(t) dt$$

Where ρ is our discount rate and $\tilde{w}(t)$ is the spending of the acquired user at time t . Therefore, the conditional expectation of $\mathbb{E}[w|x]$, where x is the acquired user's cumulative spending until t_0 , is given by:

$$\mathbb{E}[w|x] = \int_{t=t_0}^{\infty} e^{-\rho(t-t_0)} \mathbb{E}[\tilde{w}(t)|x] dt$$

Further, we assume that:

$$(32) \quad \mathbb{E}[\tilde{w}(t)|x] = \exp(a_0 + a_1 t + a_2 t^2 + a_3 x + a_4 x^2 + a_5 t \cdot x)$$

The use of exponential form is appropriate here because spending takes positive values. We can estimate the unknown parameters (a_0, \dots, a_5) using Equation 32 since we recorded the spending $\tilde{w}(t)$ of each acquired user at different point in time, as well as

the corresponding x , the acquired user’s cumulative spendings at time t_0 .²⁵ Similar to Chan et al. (2011), we set the continuous discount rate, ρ , to be 0.000611. This corresponds to a 20% annual discount rate. We also examine whether the results are sensitive to using 15% discount rate.

Therefore, our predicted lifetime spending of an average user acquired from the publisher is given by:

$$(33) \quad \frac{1}{T} \sum_t^T (x_t + \hat{\mathbb{E}}[w|x_t])$$

where

$$(34) \quad \hat{\mathbb{E}}[w|x] = \int_{t=t_0}^{\infty} e^{-\rho(t-t_0)} \hat{\mathbb{E}}[\tilde{w}(t)|x] dt$$

and $\hat{\mathbb{E}}[\tilde{w}(t)|v]$ is Equation (32) fitted to the data.

ACKNOWLEDGEMENTS

We thank Matt Shum, Sha Yang, Botao Yang, Ben Gillen, Marcelo Fernandez, Kirill Pogorelskiy, as well as seminar participants at UCLA, University of Florida and UT Dallas.

REFERENCES

- Agarwal, N., Athey, S., and Yang, D. (2009). Skewed bidding in pay-per-action auctions for online advertising. *The American Economic Review*, 99(2):441–447.
- Amaldoss, W., Desai, P. S., and Shin, W. (2015). Keyword search advertising and first-page bid estimates: A strategic analysis. *Management Science*, 61(3):507–519.
- Andrews, M., Goehring, J., Hui, S., Pancras, J., and Thornswood, L. (2016). Mobile promotions: A framework and research priorities. *Journal of Interactive Marketing*, 34:15–24.
- Andrews, M., Luo, X., Fang, Z., and Ghose, A. (2014). Mobile crowdsensing. *Marketing Science forthcoming. Fox School of Business Research*, pages 15–040.
- Andrews, M., Luo, X., Fang, Z., and Ghose, A. (2015). Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science*.

²⁵In estimating Equation 32, we use tracking data that spanned the first 30 days of user acquisition.

- Asdemir, K., Kumar, N., and Jacob, V. S. (2012). Pricing models for online advertising: Cpm vs. cpc. *Information Systems Research*, 23(3-part-1):804–822.
- Athey, S. and Nekipelov, D. (2010). A structural model of sponsored search advertising auctions. In *Sixth ad auctions workshop*.
- Bart, Y., Stephen, A. T., and Sarvary, M. (2014). Which products are best suited to mobile advertising? a field study of mobile display advertising effects on consumer attitudes and intentions. *Journal of Marketing Research*, 51(3):270–285.
- Blake, T., Nosko, C., and Tadelis, S. (2015). Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica*, 83(1):155–174.
- Bruce, N., Murthi, B., and Rao, R. C. (2016). A dynamic model for digital advertising: The effects of creative formats, message content and targeting on engagement. *Message Content and Targeting on Engagement (April 10, 2016)*.
- Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics*, 34(3):305–334.
- Chan, T. Y., Kadiyali, V., and Park, Y.-H. (2007). Willingness to pay and competition in online auctions. *Journal of Marketing Research*, 44(2):324–333.
- Chan, T. Y., Wu, C., and Xie, Y. (2011). Measuring the lifetime value of customers acquired from google search advertising. *Marketing Science*, 30(5):837–850.
- Chen, Y., Wang, Q., and Xie, J. (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of marketing research*, 48(2):238–254.
- Chen, Y. and Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, 54(3):477–491.
- Fader, P. S. and Hardie, B. G. (2005). A note on deriving the pareto/nbd model and related expressions. <http://www.brucehardie.com/notes/009/>.
- Fader, P. S., Hardie, B. G., and Lee, K. L. (2005). ‘counting your customers’ the easy way: An alternative to the pareto/nbd model. *Marketing science*, 24(2):275–284.
- Fong, N. M., Fang, Z., and Luo, X. (2015). Geo-conquesting: Competitive locational targeting of mobile promotions. *Journal of Marketing Research*, 52(5):726–735.
- Ghose, A. and Han, S. P. (2014). Estimating demand for mobile applications in the new economy. *Management Science*, 60(6):1470–1488.
- Ghose, A. and Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, 55(10):1605–1622.
- Goldfarb, A. (2014). What is different about online advertising? *Review of Industrial Organization*, 44(2):115–129.

- Goldfarb, A. and Tucker, C. (2011). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, 30(3):389–404.
- Grewal, D., Bart, Y., Spann, M., and Zubcsek, P. P. (2016). Mobile advertising: a framework and research agenda. *Journal of Interactive Marketing*, 34:3–14.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., and Sriram, S. (2006). Modeling customer lifetime value. *Journal of service research*, 9(2):139–155.
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054.
- Hansen, L. P. and Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, pages 1269–1286.
- Ho, T.-H., Li, S., Park, S.-E., and Shen, Z.-J. M. (2012). Customer influence value and purchase acceleration in new product diffusion. *Marketing Science*, 31(2):236–256.
- Hsieh, Y., Shum, M., and Yang, S. (2014). To score or not to score? structural estimation of sponsored search auctions. Technical report, mimeo.
- Hu, Y., Shin, J., and Tang, Z. (2015). Incentive problems in performance-based online advertising pricing: Cost per click vs. cost per action. *Management Science*.
- Iyer, G., Soberman, D., and Villas-Boas, J. M. (2005). The targeting of advertising. *Marketing Science*, 24(3):461–476.
- Jeziorski, P. and Moorthy, S. (2016). Advertiser prominence effects in search advertising. Working Paper.
- Johnson, G. A. (2013). The impact of privacy policy on the auction market for online display advertising.
- Lambrech, A. and Tucker, C. (2013). When does retargeting work? information specificity in online advertising. *Journal of Marketing Research*, 50(5):561–576.
- Lee, C., Kumar, V., and Gupta, S. (2015). Designing freemium: Balancing growth and monetization strategies. Available at SSRN.
- Lewis, R. A. and Rao, J. M. (2015). The unfavorable economics of measuring the returns to advertising*. *The Quarterly Journal of Economics*, page qjv023.
- Liu, D. and Viswanathan, S. (2014). Information asymmetry and hybrid advertising. *Journal of Marketing Research*, 51(5):609–624.
- Luo, X., Andrews, M., Fang, Z., and Phang, C. W. (2013). Mobile targeting. *Management Science*, 60(7):1738–1756.

- Manchanda, P., Dubé, J.-P., Goh, K. Y., and Chintagunta, P. K. (2006). The effect of banner advertising on internet purchasing. *Journal of Marketing Research*, 43(1):98–108.
- Newey, W. K. (1985). Maximum likelihood specification testing and conditional moment tests. *Econometrica: Journal of the Econometric Society*, pages 1047–1070.
- Newey, W. K. (1993). 16 efficient estimation of models with conditional moment restrictions. *Handbook of statistics*, 11:419–454.
- Rutz, O. J. and Bucklin, R. E. (2011). From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research*, 48(1):87–102.
- Sayedi, A., Jerath, K., and Srinivasan, K. (2014). Competitive poaching in sponsored search advertising and its strategic impact on traditional advertising. *Marketing Science*, 33(4):586–608.
- Schmittlein, D. C., Morrison, D. G., and Colombo, R. (1987). Counting your customers: Who-are they and what will they do next? *Management science*, 33(1):1–24.
- Shankar, V. and Balasubramanian, S. (2009). Mobile marketing: a synthesis and prognosis. *Journal of Interactive Marketing*, 23(2):118–129.
- Shin, W. (2015). Keyword search advertising and limited budgets. *Marketing Science*, 34(6):882–896.
- Sriram, S., Manchanda, P., Bravo, M. E., Chu, J., Ma, L., Song, M., Shriver, S., and Subramanian, U. (2015). Platforms: a multiplicity of research opportunities. *Marketing Letters*, 26(2):141–152.
- Trusov, M., Bucklin, R. E., and Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. *Journal of marketing*, 73(5):90–102.
- Weintraub, G. Y., Benkard, C. L., and Van Roy, B. (2005). Oblivious equilibrium: A mean field approximation for large-scale dynamic games. In *NIPS*, pages 1489–1496.
- Wu, C. (2015). Matching value and market design in online advertising networks: An empirical analysis. *Marketing Science*, 34(6):906–921.
- Yang, S. and Ghose, A. (2010). Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? *Marketing Science*, 29(4):602–623.
- Yao, S. and Mela, C. F. (2011). A dynamic model of sponsored search advertising. *Marketing Science*, 30(3):447–468.
- Zhang, K. and Katona, Z. (2012). Contextual advertising. *Marketing Science*, 31(6):980–994.

Zhu, Y. and Wilbur, K. C. (2011). Hybrid advertising auctions. *Marketing Science*, 30(2):249–273.

Zubcsek, P. P., Katona, Z., and Sarvary, M. (2015). Social and location effects in mobile advertising. *NA-Advances in Consumer Research Volume 43*.