

NBER WORKING PAPER SERIES

BUY-IT-NOW OR TAKE-A-CHANCE:
PRICE DISCRIMINATION THROUGH RANDOMIZED AUCTIONS

L. Elisa Celis
Gregory Lewis
Markus M. Mobius
Hamid Nazerzadeh

Working Paper 18590
<http://www.nber.org/papers/w18590>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2012

Markus Mobius is currently employed by Microsoft Research New England as a senior researcher. The remaining authors have been employed there in the past (Elisa Celis as a summer intern; Gregory Lewis as a visiting and consulting researcher; Hamid Nazerzadeh as a postdoctoral researcher). The authors have also benefitted from access to proprietary data from Microsoft Advertising Exchange through Microsoft, and have published a patent (with Lee-Ming Zen) held by Microsoft, based on this research. Gregory Lewis received a summer grant for this research through NET Institute. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by L. Elisa Celis, Gregory Lewis, Markus M. Mobius, and Hamid Nazerzadeh. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Buy-it-now or Take-a-chance: Price Discrimination through Randomized Auctions
L. Elisa Celis, Gregory Lewis, Markus M. Mobius, and Hamid Nazerzadeh
NBER Working Paper No. 18590
December 2012
JEL No. D4,D44,D47,D82

ABSTRACT

Increasingly detailed consumer information makes sophisticated price discrimination possible. At "fine levels of aggregation, demand may not obey standard regularity conditions. We propose a new randomized sales mechanism for such environments. Bidders can "buy-it-now" at a posted price, or "take-a-chance" in an auction where the top $d > 1$ bidders are equally likely to win. The randomized allocation incentivizes high valuation bidders to buy-it-now. We analyze equilibrium behavior, and apply our analysis to advertiser bidding data from Microsoft Advertising Exchange. In counterfactual simulations, our mechanism increases revenue by 4.4% and consumer surplus by 14.5% compared to an optimal second-price auction.

L. Elisa Celis
Xerox Research
ecelis@cs.washington.edu

Gregory Lewis
Department of Economics
Harvard University
125 Littauer Center
1805 Cambridge Street
Cambridge, MA 02138
and NBER
glewis@fas.harvard.edu

Markus M. Mobius
Department of Economics
Iowa State University
460A Heady Hall
Ames, IA 50011
and NBER
mobius@fas.harvard.edu

Hamid Nazerzadeh
University of Southern California
Marshall School of Business
3760 Trousdale Pkwy
Los Angeles, CA 90089
hamidnz@marshall.usc.edu

An online appendix is available at:
<http://www.nber.org/data-appendix/w18590>

1 Introduction

Advertising technology is changing fast. Consumers can now be reached while browsing the Internet, playing games on their phone or watching videos on YouTube. The large companies that control these new media — household names like Google, Facebook and Yahoo! — generate a substantial part of their revenue by selling advertisements. They also know more and more information about their users. This allows them to match advertisers to potential buyers with ever greater efficiency. While this matching technology generates surplus for advertisers, it also tends to create thin markets where perhaps only a single advertiser has a high willingness to pay. These environments pose special challenges for the predominant auction mechanisms that are used to sell online ads because they reduce competition among bidders, making it difficult for the platform to extract the surplus generated by targeting.

For example, a sportswear firm advertising on the New York Times website may be willing to pay much more for an advertisement placed next to a sports article than one next to a movie review. It might pay an additional premium for a local consumer who lives in New York City and an even higher premium if the consumer is known to browse websites selling sportswear. Each layer of targeting increases the sportswear firm’s valuation for the consumer but also dramatically narrows the set of participating bidders to fellow sportswear firms in New York City. Without competition, revenue performance may be poor (Bergemann and Bonatti 2011, Levin and Milgrom 2010).

Suggestive empirical evidence for this is provided in Figure 1. The left panel shows the re-scaled bids in 50 auctions by five large advertisers for the most popular webpage slot sold by a large publisher through a second-price auction on Microsoft Advertising Exchange.¹ The bids exhibit considerable variation, even though all these impressions were auctioned within a 3-hour period. We attribute this to matching on user demographics (evidence is provided later in the paper). Moreover, there is often a large gap between the highest and second highest bid in the second-price auction: on average the gap is bigger than the sales price.

Consider a simple model that formalizes this narrative: When advertisers “match” with users, they have high valuation; otherwise they have low valuation. Assume that match probabilities are independent across bidders, and sufficiently low that the probability that *any* bidder matches is small. Then a second-price auction will typically get low revenue, since the probability of two “matches” occurring in the same auction is very small. On the

¹Section 4 discusses the data in detail.

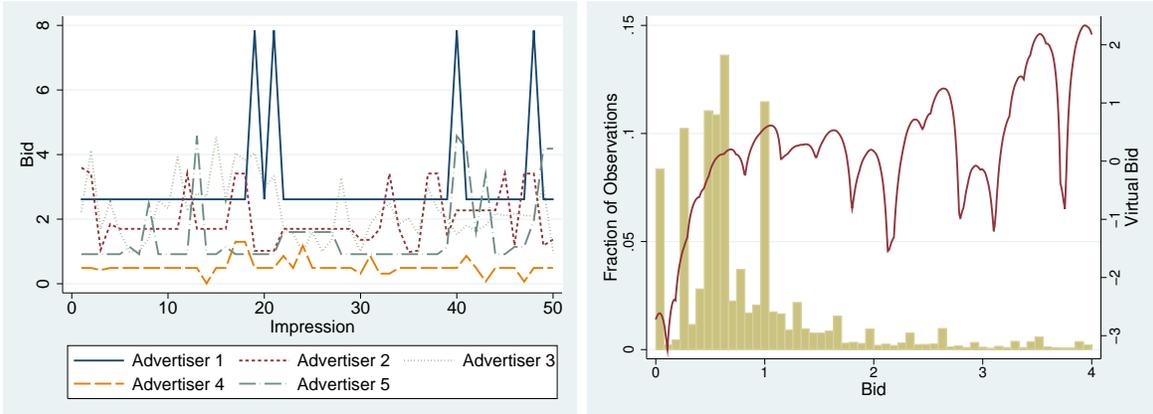


Figure 1: **Bids over Time, Bid Distribution and Virtual Bids.** The left panel shows the (rescaled) bids of five advertisers in our data, selected at random from the top 50 advertisers (ranked by purchases) on 50 randomly chosen successive bids for the most popular product. The set of impressions differs across advertisers; there are no impressions on which all 5 participate. The right panel overlays a histogram of the bid data from the same product with a line plot of the “virtual bid” function, which is defined analogously to the “virtual valuation” function as $\hat{\phi}(b) = b - \frac{1 - \hat{F}(b)}{\hat{f}(b)}$. The pdf f and cdf F are estimated by kernel smoothing with bandwidth 0.075.

other hand, setting a high fixed price is not effective since the probability of zero “matches” occurring is large and many impressions would go unallocated. Allowing user targeting creates asymmetries in valuations that increase efficiency, but decrease revenue.

But since targeting increases total surplus, platforms would like to allow targeting while still extracting the surplus this creates. This paper outlines a new and simple mechanism for doing so. We call it *buy-it-now or take-a-chance* (BIN-TAC), and it works as follows. Goods are auctioned with a buy-it-now price p , set relatively high. If a single bidder chooses buy-it-now, they get the good for price p . If more than one bidder takes the buy-it-now option, a second price auction is held between those bidders with reserve p . Finally, if no one participates in buy-it-now, an auction is held in which the top d bidders are eligible to receive the good, and it is randomly awarded to one of them at the $(d + 1)$ -st highest price.

In this manner, we combine the advantages of an auction and a fixed price mechanism. When matches occur, advertisers self-select into the fixed-price buy-it-now option, allowing for revenue extraction. Advertisers are incentivized to take the “buy-it-now” option because in the event that they “take-a-chance” on winning via auction, there is a significant probability they will not win the impression, even if their bid is the highest. On the other hand, when no matches occur, the impression is still allocated, thereby earning revenue.

Online advertising is just one example of a market in which randomized price discrimination mechanisms may be useful. Priceline offers consumers the chance to “name your own price” when booking a hotel room. But in return they are guaranteed only a certain star rating and rough location of the hotel — from their perspective, the actual allocation is random. This allows hotels to dispose of excess inventory, while still providing an incentive to customers with specific preferences to book at higher fixed prices. More generally, sales have a similar logic to our mechanism. During a sale, goods are inefficiently rationed: the first person willing to pay the sales price gets the object. Knowing this, high valuation buyers may buy at the higher pre-sale prices, knowing that their chance of getting the good on sale is small. Yet when no-one has high valuation, the sale ensures that the inventory is nonetheless sold.

Despite these examples, randomization is not part of the standard price discrimination toolkit. This is especially true of the canonical single-unit auction environment under symmetric independent private values (SIPV), where Myerson (1981) and Riley and Samuelson (1981) famously established that the seller’s optimal sales mechanism is just a (deterministic) second price auction with a reserve price. But in deriving this result it is assumed that the distribution of valuations is *regular*: the virtual valuations $\psi(v) = v - \frac{1-F(v)}{f(v)}$ are increasing. And although this holds for many standard distributions such as the normal and log-normal, it is not hard to violate this assumption in richer environments.

The right panel of Figure 1 shows a histogram of the bids for the most popular webpage slot, as well as the estimated virtual bids $\hat{\phi}(b) = b - \frac{1-\hat{F}(b)}{\hat{f}(b)}$. Notice that the virtual bids are certainly not increasing, even in regions where we have large numbers of observations. Interpreting bids as valuations — reasonable since it is weakly dominant for advertisers to bid their valuations — we can conclude that the distribution is irregular. We suspect that irregular values are actually quite common in practice. Any kind of “matching” story of the kind we outlined earlier could lead to bimodal distributions and possible irregularity. Moreover, it is hard to write down a sensible demand system in which valuations are not correlated, violating the IPV assumption. For example, if advertisers target users from a specific geographic region, this will induce positive correlation between advertisers with the same target region, and negative correlation across advertisers with different target regions.

Randomization will generally be necessary for optimal price discrimination with irregular values. Moreover, this is exactly the case when price discrimination is most valuable. Bulow and Klemperer (1996) proved that when valuations are regular, revenues from a simple second price auction with $n + 1$ bidders exceeds the revenue from the revenue optimal mechanism

with n bidders. This bounds the gain to design. But in the irregular environment this need not be true: good design can make up for fewer bidders.

With this motivation in mind, we spend the rest of the paper investigating our BIN-TAC mechanism. We have chosen to focus on it for a number of reasons. First, we believe it is *simple*, both in that it is easy to explain to participants and in that it requires relatively little input from the mechanism designer: a choice of buy-it-now price, randomization parameter d and a reserve in the take-a-chance auction. It is also reasonably easy to play, as only the buy-it-now decision requires some strategic thinking (it is weakly dominant to bid one’s valuation in the auctions).

Second, we will argue that it is *approximately optimal* for the symmetric independent private values environment with irregular valuations and *robust* to different informational environments. We show that in the simple two-type matching environment outlined earlier, BIN-TAC is revenue optimal, outperforming the leading alternatives: a second price auction with reserve, or the “bundling” solution in which the platform withholds targeting information, as suggested in a number of recent papers.² In a more general SIPV environment with irregular valuations, BIN-TAC will not generally be revenue optimal. The optimal mechanism, as suggested by Myerson (1981), is a direct mechanism that involves pooling groups of types through “ironing”.³ Numerical simulations show that in mixture distribution environments, BIN-TAC approximates the allocations and payments of this mechanism, achieving “almost optimal” performance. In fact, when valuations are negatively correlated across bidders, inducing bid gaps, BIN-TAC outperforms the ironing mechanism, despite its simplicity.

It is hard to motivate a mechanism based on simulations alone. So to analyze its performance in a real-world setting, we turn to historical data from Microsoft Advertising Exchange. Assuming that bids are equal to valuations, we can simulate the effect of introducing the BIN-TAC mechanism. We find that the optimal BIN-TAC mechanism generates 4.4% more revenue than the optimal second-price auction, while at the same time improving advertiser surplus by 14.5%. This is possible because the optimal second-price auction uses a high reserve to extract surplus from the long tail of valuations, whereas the BIN-TAC mechanism does this through a high buy-it-now price, which avoids excluding low valuation bidders.

²See Ghosh, Nazerzadeh and Sundararajan (2007), Even-Dar, Kearns and Wortman (2007), McAfee, Papineni and Vassilvitskii (2010) and Bergemann, Bonatti and Said (2011).

³The Myerson (1981) mechanism is optimal when the buyers know their valuations. If the seller can control the flow of information, it can theoretically do better by asymmetrically revealing information (Bergemann and Pesendorfer 2007) or selling the information dynamically (Eso and Szentes 2007).

Surprisingly, BIN-TAC also generates more revenue than the Myerson mechanism in our simulations, regardless of whether we follow the standard algorithm for ironing the virtual valuations, or find an optimal ironing region by numerical optimization. This is theoretically possible because the independent values assumption fails to hold in our data. BIN-TAC handles the large gap between the highest and second highest valuations better, since the randomization threat induces buy-it-now even when the gap is large. By contrast, in the Myerson mechanism the payment charged depends on the bids of the other bidders, and when the gap is large — which is often the case — the payment is necessarily smaller.

We view the main contribution of our paper as introducing and analyzing a new, simple and robust price discrimination mechanism that makes use of randomized auctions, and then evaluating its performance in a realistic environment. Although we have motivated the mechanism through a particular story about targeting in advertising exchanges, our insights apply more generally to optimal design in the single unit auction environment with irregular valuations. A secondary contribution of the paper is to document participation and bidding behavior in the display advertising market specifically. While there has been theoretical work on this market (Muthukrishnan 2010, McAfee 2011), and empirical work on the search advertising market (Ostrovsky and Schwarz 2009, Athey and Nekipelov 2010), there has been little empirical work of this sort on display advertising. Finally, we believe we are the first to calculate and simulate revenues from the Myerson (1981) mechanism with ironing using data from a real market, which may be of interest to the many people working on problems in optimal mechanism design.

Related Work: Our work is fundamentally related to the work on optimal auction design (Myerson 1981, Riley and Samuelson 1981, Maskin and Riley 1984, Crémer and McLean 1988). There has also been some recent work on simple designs, showing that they may be approximately optimal (Hartline and Roughgarden 2009); and on screening problems where the principal approximates the type space (Madarász and Prat 2010). In the online advertising context, the seller has an additional choice variable: the decision of whether or not to reveal information that allows advertisers to target their bids. This is related to work on optimal information disclosure (Lewis and Sappington 1994, Bergemann and Pesendorfer 2007, Eso and Szentes 2007). The BIN-TAC mechanism itself has the feel of a sequential screening mechanism as in Courty and Li (2000), although our environment is static. The buy-now option has been studied in other contexts, although typically its usage

elsewhere is to allow buyers to bypass a lengthy English auction and/or avoid risk (Budish and Takeyama 2001, Reynolds and Wooders 2009).

We focus on a private values setting, while Abraham, Athey, Babiouff and Grubb (2010) consider an adverse selection problem that arises in a pure common value setting when some bidders are privately informed. This is motivated by the case when some advertisers are better able to utilize the user information provided by the platform. They show that asymmetry of information can sometimes lead to low revenue in this market.

In style our paper is close to Chu, Leslie and Sorensen (2011), who combine theory, simulations and empirics to argue that bundle-size pricing is a good approximation to the optimal mixed bundling pricing scheme for a monopolist selling multiple goods. Finally from an empirical perspective, our paper contributes to the growing literature on online advertising and optimal pricing. Much of the work here is experimental in nature — for example, Lewis and Reiley (2011) ran a randomized experiment to test advertising effectiveness, while Ostrovsky and Schwarz (2009) used an experimental design to test the impact of reserve prices on revenues. There has also been recent empirical work on privacy and targeting in online advertising (Goldfarb and Tucker 2011a, Goldfarb and Tucker 2011b).

Organization: The paper is in three parts. First, we give an overview of the market for display advertising. In the second part we introduce a stylized environment, and prove existence and characterization results for the BIN-TAC mechanism. We also provide analytic results concerning the revenue maximizing parameter choices, and compare our mechanism to others using both theory and simulations. Finally, in the third part we provide an empirical analysis using data from Microsoft Advertising Exchange, including counterfactual simulations of our mechanism’s performance. All proofs are contained in the appendix.

2 The Display Advertising Market

The organization of the display advertising market is depicted in Figure 2. On one side of the market are the “publishers”: these are websites who have desirable content and therefore attract Internet users to browse their sites. These publishers earn revenue by selling advertising slots on these sites. The other side of the market consists of advertisers. They would like to display their advertisements to users browsing the publisher’s websites.

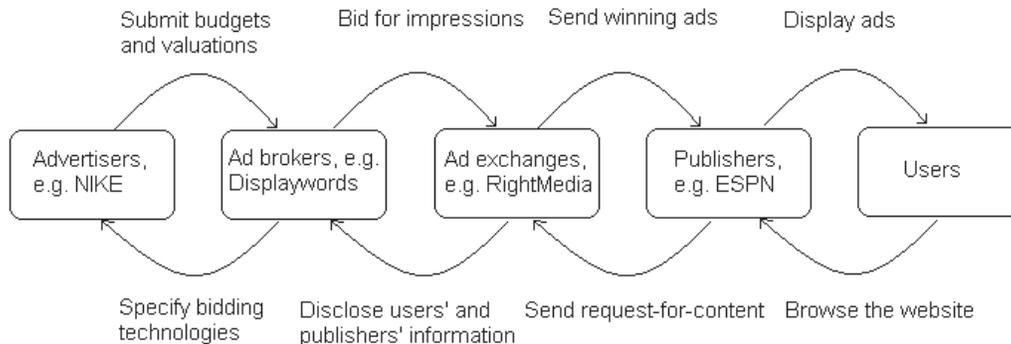


Figure 2: **The Display Advertising Market.**

They are buying user attention. Each instance of showing an advertisement to a user is called an “impression”. Advertiser demand for each impression is determined by which user they are reaching, and what the user’s current desires or intent are. For example, a Ferrari dealer might value high income users located close to the dealership. A mortgage company might value people that are reading an article on “how to refinance your mortgage” more than those who are reading an article on “ways to survive your midlife crisis”, while the dealership might prefer the reverse.

Some large publishers, primarily AOL, Microsoft and Yahoo!, sell directly to advertisers. Since the number of users browsing such publishers is extremely large, they can predict with high accuracy their user demographics. Consequently, they think of themselves of having a known inventory, consisting of a number of products in well-defined buckets: for example, male 15-24 year olds living in New York City viewing the Yahoo! homepage. They can thus contract to sell 1 million impressions delivered to a target demographic to a particular advertiser. Provided they have the inventory, they should be able to fulfill the contract. Transactions of this kind are generally negotiated between the publisher and the advertiser.

Alternatively, content is sold by auction through a centralized platform called an advertising exchange. Examples of leading advertising exchanges include the Microsoft Advertising Exchange (a subset of which we examine in this paper), Google’s DoubleClick, and Yahoo’s RightMedia. These are large marketplaces: for example, in September 2009, RightMedia averaged 9 billion transactions a day with hundreds of thousands of buyers and sellers (Muthukrishnan 2010). Advertising exchanges are a minor technological wonder, as they run all these auctions in real time. When a user loads a participating publisher’s webpage,

a “request-for-content” is sent to the advertising exchange. This request will specify the type and size of advertisement to be displayed on the page, as well as information about the webpage itself (potentially including information about its content), and information about the user browsing the page.⁴

The advertising exchange will then either allocate the impression to an advertiser at a previously negotiated price, or hold a second-price auction between participating advertisers. If an auction is held, all or some of the information about the webpage and user is passed along to ad brokers who bid on behalf of the advertisers. These ad brokers can be thought of as proprietary algorithms that take as input an advertiser’s budget and preferences, and output decisions on whether to participate in an auction and how much to bid. The winning bidder’s ad is then served by the ad exchange, and shown on the publisher’s webpage.⁵

The bids are jointly determined by the preferences advertisers have, the ad broker interface and the disclosure policies of the ad exchanges or the publishers they represent. The ad brokers can only condition the bids they place on the information provided to them: if the user’s past browsing history is not made available to them, they can’t use it in determining their bid, even if their valuation would be influenced by this information. Similarly, the advertisers are constrained in expressing their preferences by the technology of the ad broker: if the algorithm doesn’t allow the advertiser to specify a different willingness to pay based on some particular user characteristic, then this won’t show up in their bids.

Ad exchanges economize on transaction costs, by creating a centralized market for selling ad space. They also allow for very detailed products to be sold, such as the attention of a male 15-24 year old living in New York City viewing an article about hockey that has previously browsed articles about sports and theater. There is no technological reason why the products need to be sold in “buckets”, as publishers tend to do when guaranteeing sales in advance. This “real-time” sales technology is often touted as the future of this industry, as it potentially improves the match between the advertiser and their target audience. Currently ads are sold both through negotiation and advertising exchanges, with the most valuable impressions sold in advance, and the remainder sold on the exchange — a sequential sales mechanism that is similar to the BIN-TAC mechanism we will now analyze.

⁴For example, it may include their IP address and cookies that indicate their past browsing behavior.

⁵To make things yet more complicated, in some ad exchanges — though not Microsoft Advertising Exchange — two different pricing models coexist. The first is pay-per-impression, analyzed in the current paper; the second is pay-per-click, where the payment depends on whether or not the user clicks on the advertisement. Ad exchanges use expected click through rates to compare these different bids.

3 Model and Analysis

3.1 The Environment

A seller (publisher) has a single impression to sell in real time, and they have information about the user viewing the webpage, summarized in a cookie.⁶ The seller is considering one of two policies: either disclosing the cookie content to the advertiser (the “targeting” policy), or withholding it (the “bundling” policy). When they allow targeting, bidders know whether the user is a “match” for them or not. When a match occurs, the bidder has a high valuation. But the probability of a match is low and matches are assumed independent, so it is likely that everyone in the auction has a low valuation. Allowing targeting may make the market “thin” in the sense of bids being relatively low.

Instead the seller may choose to withhold the cookie, so that bidders are uncertain about whether the user is a match for them or not. The seller thus bundles good impressions with bad ones, so that bidders have intermediate valuations. This reduces match surplus, but also reduces the bidder’s information rents and so may be good for revenue.

The formal model is as follows. There are $n \geq 2$ symmetric bidders who participate in an auction for a single good which is valued at zero by the seller. Bidders are risk neutral.⁷ They have value V_H for the good when a match occurs, and value V_L for the good if no match occurs, where $V_L \sim F_L$ and $V_H \sim F_H$. We assume that F_L has support $[\underline{\omega}_L, \bar{\omega}_L]$ and F_H has support $[\underline{\omega}_H, \bar{\omega}_H]$, and that these supports are disjoint (so $\bar{\omega}_L < \underline{\omega}_H$). We assume both F_L and F_H have continuous densities f_L and f_H . The Bernoulli random variable X indicates whether a match has occurred, and the event $X = 1$ occurs with probability $\alpha \in (0, 1)$.

The bidder type is a triple (X, V_L, V_H) , drawn identically and independently across bidders, and independently of each other. This implies that a user who is a match for one advertiser need not be a match for the others. In the case with targeting, each advertiser’s realized valuation $V = (1 - X)V_L + XV_H$ is private information, known only to the advertiser. Instead if the seller bundles all impressions, the advertiser knows V_L and V_H but does not know the realization of X , implying their expected valuation is $E[V] = (1 - \alpha)V_L + \alpha V_H$.

⁶Roughly speaking, cookies allow advertisers and data brokers to track the browsing behavior of Internet users; they can also be linked to basic demographic information.

⁷Risk aversion will in general improve the performance of our mechanism relative to a second-price auction, since risk-averse bidders will be more willing to pay a high buy-it-now price to avoid the risk of the take-a-chance auction.

Most of our analysis concerns the case where the seller discloses the targeting information. In this case our model is equivalent to a standard SIPV model in which each bidder’s single-dimensional valuation is drawn from the mixture distribution F , with $F(v) = (1 - \alpha)F_L(v)$ when $v \in [\underline{\omega}_L, \bar{\omega}_L]$, and $F(v) = (1 - \alpha) + \alpha F_H(v)$ when $v \in [\underline{\omega}_H, \bar{\omega}_H]$. To avoid analyzing multiple cases, we make some technical assumptions on the virtual valuations $\psi(v) = v - \frac{1-F(v)}{f(v)}$. We assume that $\psi(v)$ is continuous and increasing over the regions $[\underline{\omega}_L, \bar{\omega}_L]$ and $[\underline{\omega}_H, \bar{\omega}_H]$, crossing zero once in the low valuation region; and that $\psi(\bar{\omega}_L) \leq \psi(\underline{\omega}_H)$. The virtual valuations are (infinitely) negative over the region $(\bar{\omega}_L, \underline{\omega}_H)$ since F is unsupported on this region. So the virtual valuations are not increasing, and the distribution F is *irregular*.

Discussion: We assume that the match random variables X and the valuations V_L and V_H are independent across bidders. We focus on the independent case because it is canonical in the price discrimination and mechanism design literatures, and so is a natural starting point (we will allow for correlated values in later simulations). We also restrict attention to environments where α is small, so that the probability of zero or a single match is high. This is the interesting case, reflecting the industry concern that providing “too much” targeting information reduces competition and hurts revenues. As shown in Figure 1 we observe a large gap between the highest and second highest bid in our data, which motivates this choice.

An important special case occurs when the distributions F_L and F_H are degenerate, with all their mass in atoms at v_L and v_H respectively. We call this the *two-type case*, since there are two types of bidders: those who matched, and therefore have valuation v_H , and those who didn’t, with valuation v_L . We will analyze this case before diving into analysis of the full model, since it provides useful intuition. But first we describe our mechanism.

3.2 Pricing Mechanisms

BIN-TAC works as follows. A *buy-it-now price* p is posted. Buyers simultaneously indicate whether they wish to *buy-it-now* (*BIN*). In the event that exactly one bidder elects to buy-it-now, that bidder wins the auction and pays p . If two or more bidders elect to BIN, a second-price sealed bid auction with reserve p is held between those bidders. Bidders who chose to BIN are obliged to participate in this auction. Finally, if no-one elects to BIN, a sealed bid *take-a-chance* (*TAC*) auction is held between all bidders, with a reserve r . In that auction, one of the top d bidders is chosen uniformly at random, and if that bidder’s

bid exceeds the reserve, they win the auction and pay the maximum of the reserve and the $(d + 1)$ -th bid (if it exists). Ties among d -th highest bidders are broken randomly prior to the random allocation. We call r the TAC-reserve, and d the randomization parameter.

To analyze the performance of BIN-TAC, it will be useful to have some benchmarks for comparison. A natural benchmark is the pricing mechanism that is most commonly used in practice, the second price auction (SPA). We distinguish between when an SPA is used and targeting is allowed (SPA-T), and when it is used with bundling (SPA-B).

Another benchmark is Myerson’s (1981) revenue optimal mechanism for the SIPV environment, which may require ironing.⁸ Ironing requires that sometimes the allocation is randomized among bidders with different valuations, and so — just as in our TAC auction — the winning bidder need not have the highest valuation. The difference is that in the optimal mechanism, the randomization only takes place when two or more bidders — including the highest valuation bidder — have valuations in a given “ironing” region.

3.3 The Two-Type Case

To gain intuition, we start by analyzing the two-type environment. Here the goal is to set the BIN-TAC parameters in such a way that bidders who match take the buy-it-now option, and the rest take-a-chance. Then whenever at least two bidders match, they will bid up the price to their common valuation v_H ; whenever a single bidder matches, the revenue will be the buy-it-now price p ; and whenever no-one matches, the revenue will be v_L (provided $d \leq n - 1$). Since bidders match independently according to the Bernoulli random variable X , the number of matches is Binomial(n, α). The buy-it-now price is constrained by incentive compatibility, as high types must prefer to buy-it-now. Optimally setting the TAC reserve to v_L to economize on math, the remaining BIN-TAC parameters are chosen to solve the following expected revenue maximization problem:

$$\begin{aligned} \max_{d,p} &= \min \left\{ \frac{n}{d}, 1 \right\} (1 - \alpha)^n v_L + n\alpha(1 - \alpha)^{n-1} p + (1 - (1 - \alpha)^n - n\alpha(1 - \alpha)^{n-1}) v_H \\ \text{subject to} & \quad v_H - p \geq \frac{1}{d}(v_H - v_L) \quad (IC) \end{aligned} \quad (1)$$

⁸As earlier remarked, the seller potentially achieve even higher revenue by selling the information. But the results of Eso and Szentes (2007) do not immediately apply here, since the buyers start with two pieces of information V_L and V_H ; whereas in their model buyers have only one piece of information initially.

The revenue expression follows the logic outlined above, allowing for the possibility that $d > n$, so that the good may not be allocated to any of the TAC bidders. The LHS of the IC constraint is the payoff to a high type from buying-it-now assuming no other types buy-it-now; and the RHS is the expected payoff from taking-a-chance, again assuming no other types buy-it-now. Conditioning on no-one else taking the BIN option does not affect the IC constraint, since if someone does, expected surplus is zero regardless of whether the buyer elects to BIN or TAC.

Clearly the seller wants to set the BIN price so the IC constraint holds with equality; which implies a BIN price equal to $\frac{d-1}{d}v_H + \frac{1}{d}v_L$. Notice this is strictly increasing in d , which makes sense since it is the threat of randomization in the TAC auction that makes the BIN option attractive. Substituting out for the price in the objective function yields an integer optimization problem in d . The objective function is strictly increasing in d until $d = n$, since randomization in the TAC auction is costless (all bidders are willing to pay the reserve v_L) and the BIN price is increasing. It has a bang-bang solution: either $d = n$ or $d = \infty$. The latter is equivalent to setting a reserve at v_H and never selling to the low-types, optimal when $\alpha v_H \geq v_L$ (i.e. when a seller selling to a single agent would prefer a price of v_H to v_L).

The interior solution is more interesting: the seller randomly allocates the object among the n bidders if no-one takes the BIN option, charging the low valuation v_L . The mechanism achieves the efficient allocation, since if there is a high type they take the BIN option and get the object, and otherwise the object is allocated to a low type. Letting q_i be the probability of exactly i matches, expected revenue simplifies to $(q_0 + \frac{q_1}{n})v_L + (1 - q_0 - \frac{q_1}{n})v_H$, a weighted average of the low and high valuations.

How does this compare to other mechanisms? The SPA-T achieves identical efficiency and revenue in the case where the gap in valuations $v_H - v_L$ is sufficiently high that $d^* = \infty$. But in the other case where the low types are not excluded, the SPA-T is equally efficient but achieves lower revenue. This is because with probability $q_0 + q_1$ there will be zero or one high types, and revenue will be v_L , as compared to a weight on v_L of $(q_0 + \frac{q_1}{n})$ in BIN-TAC.⁹

Under bundling the analysis is trivial since all bidders have expected valuations of $\alpha v_H + (1 - \alpha)v_L$, implying random allocation and revenue from the SPA-B of $\alpha v_H + (1 - \alpha)v_L$. One can show that in the interior case the weight on v_L under BIN-TAC is lower than

⁹This is a neat illustration of the well-known fact that the revenue equivalence theorem fails when the distribution of valuations is discrete: BIN-TAC and the SPA-T have identical allocations, extract identical expected payments from the lowest type, and yet generate different revenues.

under SPA-B (i.e. $(q_0 + \frac{q_1}{n}) < (1 - \alpha)$), strictly for $n > 2$. This implies BIN-TAC has higher revenues than bundling (and clearly this remains true in the case where exclusion is optimal). In addition, bundling is less efficient, since the allocation is random.

Finally, BIN-TAC is in fact revenue-optimal within the class of all mechanisms that commit to targeting (see our online appendix for a proof).¹⁰ To summarize, in the two-type case, BIN-TAC is both efficient and revenue optimal, and dominates both second-price auction mechanisms on both revenue and efficiency (strictly in some cases). These good properties motivate an analysis of the general case with a continuum of types.

3.4 General Case: Equilibrium Analysis

Returning to the general environment, we proceed by backward induction to characterize equilibrium strategies under BIN-TAC. If multiple players choose to BIN, the allocation mechanism reduces to a second-price auction with reserve p . Thus, it is weakly dominant for players to bid their valuations.¹¹ Truth-telling is also weakly dominant in the TAC auction.¹²

Taking these auction strategies as given, we turn to the decision whether to buy-it-now or take-a-chance. Intuitively, the BIN option should be more attractive to higher types: they have the most to lose from either random allocation (they may not get the good even if they are willing to pay the most) or from rivals taking the BIN option (they certainly do not get the good). This suggests that in a symmetric equilibrium, the BIN decision takes a threshold form: $\exists \bar{v}$ and types with $v \geq \bar{v}$ elect to BIN, and the rest do not. This is in fact the case.

Prior to stating a formal theorem, we introduce the following notation. Let the random variable Y^j be the j -th highest draw in an i.i.d sample of size $n - 1$ from F (i.e., the j -th highest rival valuation) and let Y^d be the maximum of Y^d and the TAC reserve r .

¹⁰As an example of the Bergemann and Pesendorfer (2007) argument that it may be even better to give different information to different bidders, consider giving the match information to $(n - 1)$ of the bidders and making them sequential take-it-or-leave-it offers at price v_H , selling to the last (uninformed) bidder at $\alpha v_H + (1 - \alpha)v_L$ if none of the others buy. This has higher expected revenue than BIN-TAC.

¹¹Since participation is obligatory at this stage, the minimum allowable bid is p , but no bidder would take the BIN option unless they had a valuation of at least p . Note, however, that even if a bidder's value is above p , they may not choose to take the BIN auction.

¹²The logic is standard: if a bidder with valuation v bids $b' > v$, it can only change the allocation when the maximum of the d -th highest rival bid and the reserve price is in $[v, b']$. But whenever this occurs, the resulting price is above the bidder's valuation and if she wins she will regret her decision. Alternatively, if she bids $b' < v$, when she wins the price is not affected, and her probability of winning will decrease.

Theorem 1 (Equilibrium Characterization) *Assume $d > 1$. Then there exists a unique symmetric pure strategy Bayes-Nash equilibrium of the game, characterized by a threshold \bar{v} satisfying:*

$$\bar{v} = p + \frac{1}{d}E[\bar{v} - Y^* | Y^1 < \bar{v}] \quad (2)$$

where types with $v \geq \bar{v}$ take the BIN option; and all types bid their valuation in any auction that may occur.

Equation (2) is intuitive: Which type is indifferent between the BIN and TAC options? Since strategies are increasing, the choice is relevant only when there are no higher valuation bidders (since otherwise those bidders would BIN and win the resulting auction). So if a bidder has the highest value and chooses to BIN, they get a surplus of $v - p$. Choosing to TAC gives $\frac{1}{d}E[v - Y^* | Y^1 < v]$, since they only win with probability $\frac{1}{d}$, although their payment of Y^* is on average much lower. Equating these two gives the indifferent type \bar{v} .

Now we consider the revenue-maximizing choices of the design parameters: the BIN price p , the TAC reserve r and the randomization parameter d . It is hard to characterize the optimal d , as it is an integer programming problem which does not admit standard optimization approaches. However for a given d , the optimal BIN price and TAC reserve can be characterized using first order conditions:

Theorem 2 (Optimal Buy Price and Reserve) *For any randomization parameter d , the revenue-maximizing TAC reserve r^* is either equal to $\underline{\omega}_H$ or is the unique solution of*

$$r = \frac{1 - F(r)}{f(r)} \quad (3)$$

The optimal BIN price is given by $p(\bar{v}^, r)$ where $p(\bar{v}, r) = \bar{v} - \frac{1}{d}E[\bar{v} - Y^* | Y^1 < \bar{v}]$ and \bar{v}^* is either equal to $\underline{\omega}_H$ or is a solution of the equation below:*

$$d - \frac{\partial E[\bar{v} - Y^* | Y^1 < \bar{v}]}{\partial \bar{v}} = \left(\frac{(d-1)f(\bar{v})}{1-F(\bar{v})} + \frac{(n-1)f(\bar{v})}{F(\bar{v})} \right) E[\bar{v} - Y^* | Y^1 < \bar{v}] \quad (4)$$

Equation (3) is familiar: the optimal TAC reserve is exactly the standard reserve in Myerson (1981), ensuring that no types with negative virtual valuation are ever awarded the object. This is a little surprising, since BIN-TAC is not the optimal mechanism. But raising the TAC reserve lowers the surplus from participating in the TAC auction, and so the seller can

also raise the BIN price while keeping the indifferent type \bar{v} constant. So the trade-off is the usual one: raising the TAC reserve increases expected payments from types above r^* — even those who take the BIN option — at the cost of losing revenue from the marginal type.

On the other hand, the implicit equation for the optimal BIN price is new. Notice that the BIN price in some sense sets a reserve at \bar{v} . If two bidders meet the reserve, the seller gets the second highest bid; if only one, the BIN price; and if none, he gets the TAC revenue. So a marginal increase in the threshold has three effects. First, if the highest bidder has valuation exactly equal to the threshold, following an increase she will shift from BIN to TAC. This costs the seller the difference between the BIN price and the expected revenue from the TAC auction (which is lower). Second, if the second highest bidder has valuation equal to the threshold, an increase will knock her out of the BIN auction, and the seller’s revenue falls by $\bar{v} - p(\bar{v}, r^*)$. Finally, if the highest bidder is above the reserve and the second highest is below, an increase gains the seller $\frac{\partial p(\bar{v}, r^*)}{\partial \bar{v}}$. Working out the probabilities of these various events, expanding $\frac{\partial p(\bar{v}, r^*)}{\partial \bar{v}}$ and equating expected costs and benefits, we get the result.

We cannot rule out a corner solution for the buy-price, where it is set equal to $\underline{\omega}_H$. This can easily occur if the value of a match is high (i.e. $\underline{\omega}_H \gg \bar{\omega}_L$). In this case it is not profitable to randomize the allocation for any of the high types: the BIN price is set at $p(\underline{\omega}_H, r^*)$ so that the lowest high type at $\underline{\omega}_H$ elects to BIN.

3.5 General Case: Performance Comparisons

We now compare the BIN-TAC mechanism to each of the other benchmark mechanisms. Our first observation is that SPA-T is just a special case of BIN-TAC, and so any outcome achievable by the SPA-T is also achievable by BIN-TAC. Notice that when $d = 1$, the TAC auction is just a second price auction. So any SPA-T with optimal reserve r can be mimicked by a BIN-TAC mechanism with the same reserve, $d = 1$ and $p = \infty$. Moreover, as we saw in the two-type case, when the objective is revenue maximization, BIN-TAC can sometimes do strictly better — in fact, we will see in the empirical application a case where the revenue-optimal BIN-TAC beats the revenue-optimal SPA-T on *both* revenue and consumer surplus.

We saw earlier that BIN-TAC also dominated the SPA-B when there were only two types. This is unfortunately not true in general: BIN-TAC trades-off increased revenues from high types (by providing match information) against lower revenues and inefficient allocations to low types. This may be less profitable than limiting information rents through bundling. But

in the Monte Carlo simulations presented in Section 3.6, BIN-TAC generally outperforms bundling.

Clearly BIN-TAC will have (weakly) worse revenue performance than the Myerson mechanism, which is revenue optimal. The question is how close BIN-TAC gets. As we show in the online appendix, whenever the low types are not excluded, the Myerson mechanism has an ironing region that pools types on $(\tilde{v}, \underline{\omega}_H)$, for \tilde{v} the solution of $-F(\tilde{v})^2 + (2 - \alpha)F(\tilde{v}) + \alpha(\underline{\omega}_H - \tilde{v})f(\tilde{v}) = 1 - \alpha$; and has a reserve $r^* < \tilde{v}$ that solves equation (3). The allocation and payments work as follows: if the top two bidders are not in the ironing region, the highest bidder wins and pays the second highest bid. When the highest bidder is not in the ironing region, but the second highest bidder is, the high bidder pays $\frac{k}{k+1}\underline{\omega}_H + \frac{1}{k+1}\tilde{v}$ for k the number of other bids in $(\tilde{v}, \underline{\omega}_H)$. Finally, if the highest bid is in the ironing region, the good is allocated uniformly at random to a bidder in the ironing region at price \tilde{v} . Like BIN-TAC, this is inefficient, but allows additional revenue extraction from higher types. In equilibrium bidders bid their valuations. The mechanism is ex-post incentive compatible in the following sense: no bidder would have an incentive to change their bid if they were told the other bids prior to the outcome of any randomization.

Having obtained this characterization, we can compare BIN-TAC with the Myerson mechanism (OPT). For now, let us focus on a simple environment, where $n = 5$, $\alpha = 0.05$, F_L is uniform over $[0, 1]$ and F_H is uniform over $[3, 4]$. Figure 3 shows the interim allocation probabilities (top panel) and expected payments by type (bottom panel) as a function of bidder type. The optimal mechanism has a discontinuous jump in the allocation probability at $\tilde{v} = 0.676$, and then irons until the high valuation region on $[3, 4]$. As you can see, BIN-TAC is able to approximate the discontinuous increase in allocation probability at \tilde{v} with a smooth curve, by randomizing the allocation in that region using the TAC auction. The important thing is that neither BIN-TAC nor OPT assign the good with high probability to *any* of the types in $[0, 1]$; thereby incentivizing the high types to take the BIN option / report their high type truthfully. By contrast, the SPA-T offers high allocation chances to the highest types on $[0, 1]$, and so can't extract high revenues from the types on $[3, 4]$.

For SPA-B, the figures depict the allocation probabilities and expected payments as a function of the *realized* valuations of the agent (which are unknown to the agent under bundling). For instance, for a given realized high valuation v_H (i.e. on $[3, 4]$), the allocation probability plotted on the y-axis is equal to the average allocation probability across all types $(1 - \alpha)V_L + v_H$, who will bid their expected valuations in the second-price auction under

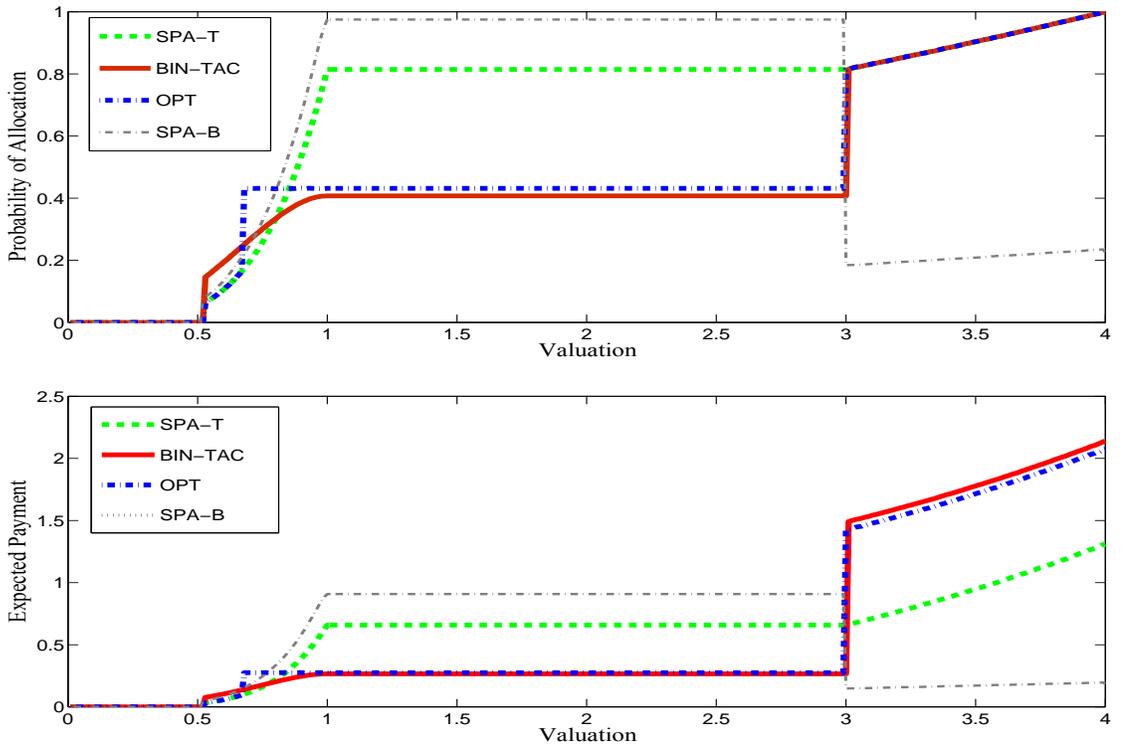


Figure 3: **Comparison of Allocations and Payments.** Allocation probabilities (top panel) and expected payments (bottom panel) for the OPT, SPA-T, SPA-B and BIN-TAC mechanisms when the distributions F_L and F_H are uniform. The x -axis corresponds to the bidder’s valuation (unknown to them under bundling).

bundling. As you can see, this implies that the allocation probability is no longer monotone: since matches are ex-ante unlikely ($\alpha = 0.05$), the most aggressive bidders in the bundling auction are those who have high valuations even without matching (high v_L), and therefore they are most likely to get the object. This makes the trade-off clear: the bundling mechanism raises expected payments when there is no match (because bidders do not know they haven’t matched), and substantially lowers them in the case of a match.

Table 1 compares the outcomes of the mechanisms. The expected revenue performance of BIN-TAC is close to the optimal mechanism (about 96% of OPT), much better than the optimal SPA-T (85%). SPA-B performs less well than both BIN-TAC and OPT, especially in terms of consumer surplus. This is because it fails to match advertisers and users correctly.

3.6 Monte Carlo Simulations

We now test our mechanism in a wider range of environments. We drop the assumption that F_L and F_H have disjoint support, and consider location families where $F_H(\cdot) = F_L(\cdot - \Delta)$ for some shift-parameter Δ . Δ is the difference in mean valuation between the high and low groups, which we call the “match increment”. We consider two location families: one where V_L is normal, and another where V_L is log normally distributed. In both cases V_L has mean 1 and standard deviation 0.5. We allow Δ , n and α to vary across experiments, and compute r^* , p^* and d^* as discussed. The default parameters we consider are $n = 10$, $\Delta = 5$, and $\alpha = .05$, and we vary one parameter at a time. Expected revenues are calculated by simulation: for each case, we sample T auctions (i.e. nT iid valuations); optimize the choices of parameters such as reserve or buy-price according to the analytic results we or others have derived, and report the average revenue across the T auctions at the optimal parameters for each mechanism.¹³

The results are presented in Figures 5, 6 and 7. On the y-axis we plot the “gain to design”:

$$\text{GD}\%(M, \theta) = \frac{E[\text{Rev}(M, \theta)] - E[\text{Rev}(SPA, 0)]}{E[V^1] - E[\text{Rev}(SPA, 0)]}$$

where M indexes mechanisms and θ parameterizes M . $E[\text{Rev}(SPA, 0)]$ is the revenue from a second-price mechanism without reserve and $E[V^1]$ is the expected value of the highest bidder. This measures the fraction of the available surplus a more sophisticated mechanism is able to capture, relative to a second price auction without reserve. We use it because it is invariant under affine transformation of the valuation distribution, which would not be true of raw revenue numbers or ratios.

This irregular environment has high potential gains to design, often over 20%. For comparison, adding an extra bidder to the default environment and then running an SPA without reserve (as in Bulow and Klemperer (1996)) is worth between 6-7%, depending on the distribution. BIN-TAC often performs much better than SPA-T, which is not surprising given our earlier analysis. More strikingly, BIN-TAC performs very comparably to the optimal Myerson mechanism, which obtains the maximum gains to design. The bundling mechanism generally performs very poorly, except in the case where the number of bidders is small.

¹³We use the optimal reserve in all cases; iron the virtual valuations using the techniques of Myerson (1981); and get the BIN-TAC parameters (p, r) from Theorem 2 for each d , computing revenues at $d \in \{2, 3, 4, 5\}$ to find the best one.

We also run a more complex simulation in which we allow for correlated values. To model correlation, we sample from a Gaussian copula with variance-covariance matrix having on-diagonal elements equal to 1, and off diagonal elements equal to ρ . When ρ is negative, elements are pairwise negatively correlated, so that valuations are “spaced out” and in expectation there is a bigger “gap” between the highest and second highest valuation. Conversely, when ρ is positive, valuations are clustered. We then inverse transform the sampled quantile vectors using the mixture distribution $F = \alpha F_H + (1 - \alpha)F_L$, where F_L and F_H are as above. Holding the baseline $n = 10$, $\Delta = 5$ and $\alpha = 0.05$, we vary ρ in Figure 8.¹⁴ We find that BIN-TAC outperforms the Myerson mechanism with negative correlation, and performs worse when valuations are positively correlated.

Overall, the performance of BIN-TAC is very good, at least for the distributions and parameters chosen, and particularly when valuations are “spaced out”. The main caveats are that it doesn’t perform well with very few bidders (when bundling is preferable), or when matches are highly probable or very valuable (a second-price auction would do as well). Its niche is in markets with many bidders but low match probabilities, so that markets are “thin” in the sense of having relatively low matches and bid gaps in expectation.

4 Empirical Application

Our theoretical analysis has shown that there are cases in which BIN-TAC performs well. We now test our mechanism’s performance in a real-world setting. We have historical data from Microsoft Advertising Exchange, one of the world’s leading ad exchanges. Our data comes from a single large publisher’s auctions on this exchange and consists of a 0.1% random sample of a week’s worth of auction data from this publisher, sampled within the last two years. This publisher sells multiple “products”, where a product is a URL-ad size combination (e.g. a large banner ad on the sports landing page of the New York Times).

The data includes information from both the publisher and the advertiser. On the publisher side, we see the url of the webpage the ad will be posted on, the size of the advertising space and the IP address of the user browsing the website. We form a unique identifier for the url-size pair, and call that a product. We determine which US state the user IP originates from,

¹⁴Because we now sample directly from F , comparison to SPA-B is not possible. There is also no guarantee that the ironing algorithm will work well under correlated values, and so we add in an additional version of the Myerson mechanism where the ironing region is chosen by numerical optimization, rather than ironing.

and call that a region. We use controls for product and region throughout the descriptive regressions. Unfortunately, we do not have more detailed information on the product or the user, as the tags and cookies passed by the publisher to the ad exchange were not stored.

On the advertiser side, we see the company name, the ad broker they employed, a variable indicating the ad they intend to show, and their bid.¹⁵ We see who won each auction and the final price. We drop auctions in which the eventual allocation was determined by biased bids and modifiers.¹⁶ We also restrict attention to impressions that originate in the US, and where the publisher content is in English. Finally, we only consider frequently sold products, those with at least 100 sales in the dataset. This leaves us with a sample of 83515 impressions.

The dataset is summarized in Table 2. For confidentiality reasons, bids have been rescaled so that the average bid across all observations is equal to 1 unit. Bids are very skew, with the median bid being only 0.57 units. Perhaps as a consequence of this skewness, the winning bid — which is more heavily sampled from the right tail of the bid distribution — is much higher at 2.96 units. There are on average 6 bidders per auction, but there is considerable variation in participation, with a standard deviation of nearly 3. Bids are not strongly correlated: as the table shows, the correlation between a randomly selected pair of bids from each auction is only 0.01. This is not statistically significant at 5% (p-value 0.116, $N = 15827$).

The advertisers are themselves quite active in the market. On average they bid on 0.7% of all impressions, and win nearly 40% of those they bid on. But participation is quite skew, and the median advertiser is far less active, bidding on only 0.02% of impressions; while the most active advertiser participates in nearly 90% of auctions. Some advertisers choose to participate in relatively few auctions, but tend to bid quite highly and therefore win with relatively high probability. Others bid lower amounts in many auctions, and win with lower probability. The first strategy is followed by companies who want to place their advertisements only on webpages with specific content or to target specific demographics, while the latter strategy is followed by companies whose main aim is brand visibility.

¹⁵In the overwhelming majority of cases there is a single ad for each company, but some larger firms have multiple ad campaigns simultaneously. We treat these as being a single ad campaign in what follows because each firm should have the same per impression valuation across campaigns.

¹⁶When the advertiser has a technologically complex ad to display, their bid is modified down (for allocation purposes) and up (for payment purposes). When the advertiser has a previously negotiated contract with the platform, their bid may be biased (usually upward for allocation, and downward for payment). It is hard to know how to treat these auctions since valuations may vary with the kind of ad being displayed; and bids may depend on what contract each bidder has, what other bidders know about these contracts etc.

4.1 Descriptive Evidence

Before proceeding to the main estimation and simulations, we provide some evidence that advertisers bid differently on different users (i.e. there is matching on user demographics). We also show that the platform is doing poorly in extracting this match surplus as revenue. We introduced our first piece of evidence in the introduction in Figure 1. As we noted, there is significant variation in the bids of large advertisers over a short time horizon. While this could be driven by decreases in the advertisers’ available budget, since the bids go both up and down it seems more likely that this variation arises from matching on user demographics.

A more direct test of advertiser-user matching is to look for the significance of advertiser-user fixed effects in explaining bids. We estimate an unrestricted model where the dependent variable is bids and the controls are advertiser-user dummies, versus a restricted model with just advertiser and user fixed effects, but not their interaction. The restricted model is overwhelmingly rejected by the data (p-value ≈ 0). This suggests matching on demographics.

Proving that this matching is motivated by economic considerations is a little more difficult. The only user demographic we observe is the user region, and it is hard to know a priori what the advertisers’ preferences over regions are. To get a handle on this, we turn to another proprietary dataset that indicates how often an advertiser’s webpage was viewed by Internet users in different regions of the country during the calendar month prior to the auction.¹⁷ Our intention is to proxy for the advertisers’ geographic preferences (insofar as these exist) using this pageview data: firms who operate in only a few regions probably attract most of their pageviews from those regions, and also mainly want to advertise in those regions. If this is right, advertisers who attract a large fraction of their pageviews from a particular region should participate more frequently and bid higher on users from those regions.¹⁸ We normalize the pageviews from a particular state by the state population to get a per capita pageview measure, and construct the fraction of normalized pageviews each region receives, calling this the “pageview ratio”.

In Table 3, we present results from regressions of auction participation (a dummy equal to one if the advertiser participated), and bid (conditional on participation) on the pageview ratio, as well as a number of fixed effects. Because the sheer size of our dataset makes it difficult to run the fixed effect regressions, we run this on a subsample consisting of the

¹⁷For example, if these auctions were in May, the pageview data would be taken from April.

¹⁸Since the pageview data dates from a period before our exchange data we are less worried about reverse causality (i.e. advertisers who win more impressions from region X later get more views from region X).

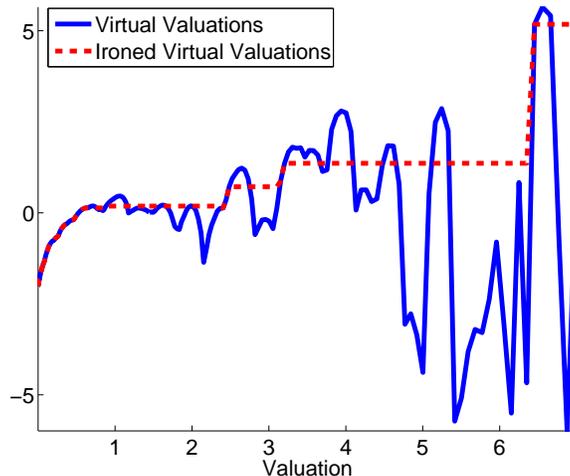


Figure 4: **Virtual Valuations.** This shows the estimated virtual valuations (solid), calculated from a non-parametric estimate of the marginal distribution of valuations. It also shows the ironed virtual valuations (dashed), derived by applying the Myerson (1981) ironing procedure.

top 10% of advertisers.¹⁹ The first column shows participation as a function of the pageview ratio, as well as product-region fixed effects, and time-of-day fixed effects (since participation and bids may vary with the user’s local time). We find a positive but insignificant effect. But when we include advertiser fixed effects to control for different participation frequencies across advertisers, we find a much bigger and now highly significant effect. All else equal, an advertiser is 3.3% more likely to bid on a user from a state that contributes 10% of the population-weighted pageviews for their site than one that contributes none. This is a large increase, as the average probability of participation is only around 1%.

Turning to the bids, we get similar estimates and significance levels from the specifications with and without advertiser fixed effects. We find that firms bid higher on users from more relevant regions, although this effect is relatively modest in economic terms. Given that our proxy for advertiser preferences is relatively crude, it is notable that we find these effects. This provides some evidence that the matching is surplus increasing, in that advertisers are able to target regions where their most valuable customers are.

Next we ask whether the platform is able to extract most of the consumer surplus. As discussed in the introduction, the average bid in an auction is 0.88, while the mean gap between the highest and second highest bid is much larger at 1.89, indicating that there is a

¹⁹Fortunately since participation is highly skewed, these advertisers account for 90% of the bids. With only bidder fixed effects we could use a within transformation to reduce the computational burden; but unfortunately this is not possible with multiple non-interacting fixed effects.

lot of money left on the table by a second-price mechanism (see Table 2 for other summary statistics). That gap itself is extremely skewed.

Assuming bids are equal to valuations — an assumption we will motivate in the next section — figure 4 shows the estimated virtual valuations $\widehat{\psi}(v)$, calculated from a non-parametric estimate of the marginal distribution of bids. Although the virtual valuations are never infinitely negative, as in our stylized model, they are certainly non-monotone. This implies that BIN-TAC may be able to extract more revenue than a second price auction.

4.2 Estimation and Counterfactual Simulations

We want to calculate how platform revenues would change if the platform instead used a different sales mechanism with optimally chosen parameters. As in Haile and Tamer (2003), we use an incomplete model of the environment for inference. Because the existing mechanism is a second price auction, we get pointwise identification of valuations under relatively weak assumptions that we outline below. But for some — but not all — of the counterfactual simulations we will need to complete the model by specifying the information structure in detail. Since we are not sure what the most realistic model is, we consider several possibilities, including a complete information model, and a symmetric private values model with correlation. For each of these models, we have to optimize over the various parameter choices to get the revenue-optimal version of each sales mechanism. In this optimization step we enforce a uniform parameter choice for all auctions, ruling out different reserves or randomization parameters by product or user-region.²⁰ The BIN-TAC mechanism performs well even when the parameter choices are coarse in this way, which alleviates the concern that our mechanism requires the seller to have detailed knowledge in order to design the mechanism. Each step in this process is explained more carefully below.

Step 1: Inference: Our original model was of a single auction rather an entire marketplace. We now offer a richer model that allows us to make sense of the dataset.²¹ A sequence of impressions is sequentially auctioned over a time period by second-price sealed bid auctions. Each impression corresponds to a user-product tuple. We assume there is a

²⁰Although the revenue of all mechanisms increases when we allow the parameters to be chosen “finely” (product-by-product), BIN-TAC continues to outperform the second price alternatives. We have not tested it against the Myerson mechanism on a product-by-product basis.

²¹We rule out search and dynamic behavior — see Backus and Lewis (2012) for a model of that form.

fixed set of N advertisers who are present throughout that time period, and who have zero costs of participating in an auction. Letting $i = 1 \dots N$ index the advertisers, $j = 1 \dots J$ index the products, and $t = 1 \dots T$ index the impressions, we assume that advertisers have valuations of the form $v_{i,t} = f_i(x_t, j_t)$ where x_t is some vector of user characteristics (largely unobserved by us), and j_t is the product being sold (known to us). Conditional on product, user characteristics x_t are sampled identically and independently across impressions (i.e. if s and t are both auctions of product j , x_s is independent of x_t). This implies that the valuation vectors $v_t = (v_{1,t}, v_{2,t} \dots v_{N,t})$ are also conditionally independent, with conditional distributions $\{F_j\}_{j=1}^J$. Advertisers know their own valuations. They get a payoff of $v_{i,t} - p_t$ when they win an impression, where p_t is the price paid. Losing bidders get nothing. They seek to maximize the sum of their payoffs across all T impressions.

We make a single assumption on advertiser behavior: whenever they have a strictly positive valuation for an impression, they participate and bid their valuation. Given the payoff structure, this strategy is weakly dominant. This one assumption allows us to make strong inferences from the data. If advertiser i did not participate in auction t , we deduce their valuation is $v_{i,t} = 0$; and if they did, we infer a valuation $v_{i,t} = b_{i,t}$. Since valuations vectors are conditionally independent given product identity, we can consistently estimate F_j from the empirical distribution of bids on that product. This is a high dimensional object, and so in what follows we will generally use the empirical distribution $\hat{F}_{j,T}$ directly in simulation, rather than attempting to smooth it using a nonparametric estimation approach.

We get a lot of mileage out of our assumptions, so we briefly note some important caveats. In assuming that advertisers know their payoffs, we implicitly rule out common values. As Abraham et al. (2010) note, advertisers can be asymmetrically informed about the user characteristics (e.g. because this user has bought a product from one of the advertisers before), which can lead to common values. We make the private values assumption for three reasons. First, we know from our earlier regressions that there is certainly an important private value component driven by geographic matching. Second, re-targeting was less prevalent in the advertising industry during the time period of our auctions; and as a result advertisers may have been less worried about adverse selection, implying minimal bid shading. Finally, from a practical perspective, it immensely simplifies the inference problem.

Another important assumption is that participation costs are zero. In fact advertisers may find it costly to fully express their preferences, and therefore issue instructions not to bid on certain less desirable user groups, rather than precisely delineating how much less those

groups are worth. Instead of inferring a valuation of zero from non-participation, we should infer $\pi(v_{i,t}) \leq c_i$, where $\pi(v_{i,t})$ is the expected payoff from participating in an auction with valuation $v_{i,t}$ and c_i is the cost of participation for bidder i . But the zero cost assumption may not be too bad. The 5th percentile of bids in our data is equal to 0.013, which is a tiny bid in the sense that it has almost zero chance of winning and even lower surplus. So certainly some bidders have participation costs closely approximated by zero, although we cannot rule out asymmetries in participation costs across large and small bidders.

The last big assumption is that bidders maximize the sum of their payoffs across all impressions, *without any budget constraint*. We ignore budget constraints partly to avoid the dynamic complications they create (but see Balseiro, Besbes and Weintraub (2012) for a simplifying approach), partly because they are unobserved to us in the data, and partly we are uncomfortable with treating them as fixed in counterfactuals.²²

Step 2a: Simulation for Robust Mechanisms: In two of the benchmark mechanisms — the second-price auction with targeting, and the Myerson mechanism with ironing — bidding one’s valuation is weakly dominant, regardless of the reserve price or ironing regions. So for these mechanisms, we can calculate revenue directly by substituting the inferred valuations in as bids, without taking a stance on the information structure.

Step 2b: Simulation of BIN-TAC: BIN-TAC is harder, as a bidder’s decision to take the BIN option depends on their beliefs about the distribution of rival valuations, which in turn depends on the information structure. We consider three possible approaches. The first is to generalize the incomplete information model from earlier in the paper, moving from symmetric independent private values to symmetric correlated private values.²³ We continue to focus on equilibria in threshold strategies: there is some $\bar{v}_j = \bar{v}_j(p, d, r)$ for each product, above which advertisers will take the BIN option, and below which they will TAC. As before, the threshold bidder is indifferent between BIN and TAC, implying $\bar{v}_j - p = \frac{1}{d}E[\bar{v}_j - Y^*|Y^1 < \bar{v}_j, \bar{v}_j]$. Because of the correlated values, we have to condition on \bar{v}_j itself on the RHS now. To solve this equation for fixed (p, d, r) , we estimate the expected

²²Budget constraints are not fixed in the long-run: an advertiser who hits their desired ad metrics with room to spare may decrease their ad budget, while one who underperforms may increase it.

²³The correlation structure in the data is extremely complex and we cannot assume affiliation. Symmetry is almost certainly violated, but without it we would have to calculate a fixed point in advertiser-specific thresholds, which is computationally infeasible.

TAC payment as a function of the highest bid in the auction. We do this product-by-product, using a kernel estimate of the conditional expectation. We then solve for the equilibrium $\bar{v}_j(p, d, r)$ for each set of BIN-randomization parameters (p, d, r) , getting an average revenue estimate as follows:²⁴

$$\begin{aligned} \text{Revenue}^{\text{BIN-TAC}}(p, d, r) = & \frac{1}{T} \sum_j \left(\sum_{t \in J(j)} 1(b_t^{(2)} \geq \bar{v}_j(p, d, r)) b_t^{(2)} + \sum_{t \in J(j)} 1(b_t^{(1)} \geq \bar{v}_j(p, d, r) > b_t^{(2)}) p \right. \\ & \left. + \sum_{t \in J(j)} 1(b_t^{(1)} < \bar{v}_j(p, d, r)) \sum_{j=1}^d 1(b_t^{(j)} \geq r) \max\{b^{(d+1)}, r\} \right) \end{aligned} \quad (5)$$

We also consider a complete information environment. Here we follow the logic of Edelman, Ostrovsky and Schwartz (2007), who suggested that since these players compete with high frequency and can potentially learn each others’ valuations, a complete information model may better approximate reality than an incomplete information model. Under complete information, there can be multiple equilibria: a bidder who knows they will lose is indifferent between BIN and TAC. We refine away this multiplicity by assuming a positive probability that a bidder will “tremble” and play TAC instead of BIN; then a bidder will take BIN whenever this yields higher expected surplus than TAC, assuming all other bidders TAC.²⁵

Finally, following the computer science literature, we perform a worst-case analysis. This is agnostic as to the particular information structure, and assumes that bidders hold the “worst-case” beliefs for BIN-TAC revenue. Specifically, each bidder believes that all other bidders will choose to TAC and then bid zero, making TAC relatively attractive. This implies that incentives to take the BIN option must be provided directly by the design, through the randomization parameter d and the reserve price r in the TAC auction.

Step 2c: Bundling Simulation: As we do not observe all the user characteristics that are provided to advertisers, we cannot consider the optimal bundling strategy. But we can examine bundling by product and user region, where the platform strips away all user characteristics except for the region, so that advertisers are buying a random impression of

²⁴Under correlation, finding threshold types $\bar{v}_j(p, d, r)$ no longer suffices for equilibrium, and so we additionally check that types above and below the threshold do not want to deviate (they don’t). In theory there may be multiple solutions $\bar{v}_j(p, d, r)$ for each product-parameter pair, but in practice we found only one.

²⁵See the online appendix for a discussion. Multiplicity arises also in the generalized second price auction — see Edelman et al. (2007) and Varian (2007).

a given size on a given website viewed by a user from a particular US state. This is unlikely to be optimal, but provides a lower bound on the revenues from the bundling strategy.

Our estimate of a bidder’s willingness to pay for this “bundled” impression is their average bid across all auctions of this product-region combination, taking their implicit bids when they didn’t participate as equal to zero. In the counterfactual, we assume that the bundled impressions are sold by second-price auction without reserve, under full participation.²⁶

Step 3: Optimization: The above steps allow us to get an estimate of the revenue from BIN-TAC and the benchmark mechanisms for various parameter choices. We want to optimize those parameter choices at a platform-wide level: finding the optimal reserve for the SPA-T, the optimal BIN-TAC parameters, and the optimal ironing regions (as explained earlier, we cannot optimize the bundle). For the SPA-T and BIN-TAC, we find these parameters by maximizing simulated revenue.²⁷ For the Myerson mechanism, we try two approaches. In one, we assume that only a single ironing region is used, and find it by numerical optimization.²⁸ The second is to estimate the marginal distribution of valuations non-parametrically (not conditioning on product), form the virtual valuations and iron. This may not be optimal given that the estimated valuations are neither symmetric nor independent. But in some sense this is the relevant benchmark, since asymmetric reserves or Crémer and McLean’s (1988) side bets are not really practical. Recently, Papadimitriou and Plerakos’s (2011) showed that computing the optimal ex-post individually rational mechanism can be computationally intractable. To get standard errors on our revenue and consumer surplus estimates, we bootstrap the sample of auctions and re-run the simulation procedure, holding the parameter choices fixed.²⁹

4.3 Results

The results are in Tables 4 and 5. We find that the optimal reserve when running a second price auction is high: nearly twice as high as the second highest bid. By contrast, BIN-TAC

²⁶The platform could in principle set a reserve to extract all the consumer surplus (since the willingness to pay of each advertiser for a bundled impression can be calculated here), but this doesn’t seem realistic.

²⁷This raises an over-fitting concern, in that the parameters are optimized for this specific realization of the data generating process. However given our sample size, the bias this introduces is likely to be small.

²⁸In contrast to our earlier simulations, this 3D optimization problem turns out to be quite hard; different starting points converge to different local optima.

²⁹We use 100 bootstrap samples (i.e. samples of T impressions drawn randomly with replacement).

always uses relatively low reserves (all well below the average bid), and instead offers a high buy price (which is close in magnitude to the optimal SPA reserve). To make this buy price attractive, the platform threatens to randomize between the top three bidders in each auction (four in the worst-case scenario), which is significant given that there are only six bidders in an average auction. The numerically optimized version of the Myerson mechanism has an ironing region that starts at the reserve, which is a little unusual. The ironed mechanism shown in Figure 4 is quite complex, with many ironing regions (some are barely visible).

Table 5 shows the mechanisms’ performance. The SPA without reserve earns revenue of 0.98 per auction, and leaves substantial consumer surplus — on average 1.97 per auction.³⁰ Adding the large optimal reserve improves revenue slightly (to 1.03 per auction), but hurts consumer surplus substantially (it falls to 1.47). Using the incomplete information results, BIN-TAC does better than both of these mechanisms in terms of revenue: it outperforms the SPA-T with optimal reserve by 4.4%. As the table shows, the results are quite robust to the information structure assumed. Interestingly, it also does better on consumer surplus than the SPA-T (an increase of 14.5%). This happens because the optimal SPA-T reserve price is very high — to extract revenue from the long right tail — and so many impressions are not sold, resulting in inefficiency and lower total welfare. By contrast, BIN-TAC has the BIN price to extract this revenue, and so the reserve is much lower, and more impressions are sold. Even accounting for distortions owing to the TAC auction, this is a welfare improvement.

The results on the Myerson mechanism are intriguing. The ironing approach performs terribly; it achieves lower revenue than the SPA-T with optimal reserve. This may be because the noisy data makes it hard to calculate good ironing regions (even though we used sophisticated adaptive bandwidth estimates of the marginal valuation distribution); or because the correlation structure makes ironing ineffective. Though the numerically optimized Myerson mechanism does much better, BIN-TAC still beats it.

To investigate why this occurs, we formed a comparison dataset by sampling bids independently from the data, holding the distribution of the number of bidders in each auction constant. This dataset satisfies the SIPV assumptions by construction. When we simulated revenue on this data set, using parameters optimized for this data, we found that Myerson achieved higher revenues, consistent with the theory. The main difference between the real and simulated data lay in the conditional distribution of the second highest bid given the

³⁰The per auction revenue of 0.98 is lower than the average second highest bid of 1.07 in Table 2 because some auctions in the data have a single bidder, which will realize zero revenue in an SPA without reserve.

highest bid. Whenever the highest bid is above 2.5, the second-highest bid in the real data is lower than that in the (independently sampled) simulated data, in the sense of first order stochastic dominance. In simple terms, the bid “gap” is higher than it would be under independence. As we saw in the simulations with negative correlation, this makes BIN-TAC more attractive, as BIN-TAC doesn’t rely on competition from other bidders to set prices — it relies solely on the threat of randomization.

Finally, the bundling strategy underperforms. Revenues are much lower than the SPA-T, and consumer surplus falls even more dramatically. This is because there is considerable variation in match surplus across impressions even after conditioning on product and region, and so bundling along only these two dimensions destroys a lot of surplus.

5 Conclusion and Future Work

Increasingly detailed consumer information makes sophisticated price discrimination possible. But at fine levels of aggregation, demand may not obey standard regularity conditions. We have seen evidence of this on Microsoft Advertising Exchange, where there are large differences between the highest and second highest bid, and irregular valuations.

Motivated by this, we have introduced a new price discrimination mechanism, designed for revenue maximization in environments with irregular valuations. As we have shown through both theory and simulations, it has a number of advantages: it is simple to explain, has good revenue performance in a variety of settings and can closely approximate Myerson’s optimal mechanism with ironing. When we tested our approach out on the advertising data, we found that the BIN-TAC mechanism would improve revenues and consumer surplus relative to the existing mechanism, a second price auction with reserve; and would even beat the Myerson mechanism on revenues. We conclude that randomization is an important price discrimination tool in irregular environments.

References

Abraham, Ittai, Susan Athey, Moshe Babioff, and Michael Grubb, “Peaches, Lemons and Cookies: The Impact of Information Structure on Revenue in Common Value Auctions,” 2010. working paper, Microsoft Research.

- Athey, Susan and Denis Nekipelov**, “A Structural Model of Sponsored Search Advertising Auctions,” 2010. Working Paper, Harvard University.
- Backus, Matthew and Gregory Lewis**, “A Demand System for a Dynamic Auction Market with Directed Search,” 2012. Working paper, Harvard University.
- Balseiro, Santiago R., Omar Besbes, and Gabriel Y. Weintraub**, “Auctions for Online Display Advertising Exchanges: Approximations and Design,” 2012. Working paper, Columbia GSB.
- Bergemann, Dirk, Alessandro Bonatti, and Maher Said**, “Product Design in Ad Exchanges,” 2011. Working paper, MIT Sloan.
- and —, “Targeting in Advertising Markets: Implications for Offline vs. Online Media,” *Rand Journal of Economics*, 2011, *42*, 417–443.
- and **Martin Pesendorfer**, “Information Structures in Optimal Auctions,” *Journal of Economic Theory*, 2007, *137*, 580–609.
- Budish, Eric and Lisa Takeyama**, “Buy prices in online auctions: irrationality on the internet,” *Economics Letters*, 2001, *72* (3), 325—333.
- Bulow, Jeremy and Paul Klemperer**, “Auctions Versus Negotiations,” *The American Economic Review*, 1996, *86* (1), pp. 180–194.
- Chu, Sean, Phillip Leslie, and Alan Sorensen**, “Bundle-size Pricing as an Approximation to Mixed Bundling,” *American Economic Review*, 2011, *101* (1).
- Courty, Pascal and Hao Li**, “Sequential Screening,” *The Review of Economic Studies*, 2000, *67* (4), pp. 697–717.
- Crémer, Jacques and Richard McLean**, “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions,” *Econometrica*, 1988, *56* (6), 1247–1257.
- Edelman, Benjamin, Michael Ostrovsky, and Michael Schwartz**, “Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars of Keywords,” *American Economic Review*, 2007, *97* (1), 242—259.
- Eso, Peter and Balazs Szentes**, “Optimal Information Disclosure in Auctions and the Handicap Auction,” *Review of Economic Studies*, 2007, *74* (3), 705—731.

- Even-Dar, Eyal, Michael J. Kearns, and Jennifer Wortman**, “Sponsored Search with Contexts,” in “Internet and Network Economics, Third International Workshop (WINE)” 2007, pp. 312–317.
- Ghosh, Arpita, Hamid Nazerzadeh, and Mukund Sundararajan**, “Computing Optimal Bundles for Sponsored Search,” in “Internet and Network Economics, Third International Workshop (WINE)” 2007, pp. 576–583.
- Goldfarb, Avi and Catherine Tucker**, “Online Display Advertising: Targeting and Intrusiveness,” *Marketing Science*, 2011, *30* (3), 389–404.
- and — , “Privacy Regulation and Online Advertising,” *Management Science*, 2011, *57* (1), 57–71.
- Haile, Philip A. and Elie Tamer**, “Inference with an Incomplete Model of English Auctions,” *The Journal of Political Economy*, feb 2003, *111* (1), 1–51.
- Hartline, Jason D. and Tim Roughgarden**, “Simple versus Optimal Mechanisms,” in “Proceedings of the 10th ACM Conference on Electronic Commerce” 2009.
- Levin, Johathan and Paul Milgrom**, “Online Advertising: Heterogeneity and Conflation in Market Design,” *American Economic Review, Papers and Proceedings*, 2010, *100* (2).
- Lewis, Randall A and David H Reiley**, “Does Retail Advertising Work? Measuring the Effects of Advertising on Sales via a Controlled Experiment on Yahoo!,” 2011. Yahoo! Research Working Paper.
- Lewis, Tracy and David Sappington**, “Supplying Information to Facilitate Price Discrimination,” *International Economic Review*, 1994, *35* (2), 309–327.
- Madarász, Kristóf and Andrea Prat**, “Screening with an Approximate Type Space,” Dec 2010. LSE Working Paper.
- Maskin, Eric and John Riley**, “Optimal Auctions with Risk Averse Buyers,” *Econometrica*, nov 1984, *52* (6), 1473–1518.
- McAfee, Preston**, “The Design of Advertising Exchanges,” *Review of Industrial Organization*, 2011, *39* (3), 169–185.

- McAfee, R. Preston, Kishore Papineni, and Sergei Vassilvitskii**, “Maximally Representative Allocations for Guaranteed Delivery Advertising Campaigns,” 2010. Yahoo! Research Working Paper.
- Milgrom, Paul and Chris Shannon**, “Monotone Comparative Statics,” *Econometrica*, 1994, *62* (1), 157–180.
- Muthukrishnan, S.**, “Ad Exchanges: Research Issues,” 2010. manuscript.
- Myerson, Roger B.**, “Optimal Auction Design,” *Mathematics of Operations Research*, 1981, *6* (1), 58–73.
- Ostrovsky, Michael and Michael Schwarz**, “Reserve Prices in Internet Advertising Auctions: A Field Experiment,” 2009. Working Paper.
- Papadimitriou, Christos H. and George Pierrakos**, “On optimal single-item auctions,” in Lance Fortnow and Salil P. Vadhan, eds., *STOC*, ACM 2011, pp. 119–128.
- Reynolds, Stanley S. and John Wooders**, “Auctions with a buy price,” *Economic Theory*, 2009, *38*, 9–39.
- Riley, John G. and William F. Samuelson**, “Optimal Auctions,” *American Economic Review*, 1981, *71* (3), 381–392.
- Varian, Hal**, “Position Auctions,” *International Journal of Industrial Organization*, 2007, *25* (7), 1163–1178.

Appendix

Proof of Theorem 1

Let a be a binary choice variable equal to 1 if the bidder takes BIN and zero if TAC. Fix a player i , and fix arbitrary measurable BIN strategies $a_j(v)$ for the other players. Let q be the probability that no other bidder takes the BIN option, equal to $\prod_{j \neq i} (\int 1(a_j(v) = 0) dF(v))$. Let $\pi(a, v)$ be the expected payoff of action a for type v given that the bidder bids their valuation in any auction that follows. Then we have that $\frac{\partial}{\partial v} \pi(1, v) \geq q$, as a marginal

increase in type increases the payoff by the probability of winning, which is lower bounded by q when taking the BIN option. Similarly we have that $\frac{\partial}{\partial v}\pi(0, v) \leq \frac{q}{d}$, as the probability of winning when taking-a-chance is bounded above by q/d . Then $\pi(a, v)$ satisfies the strict single crossing property in (a, v) ; it follows by Theorem 4 of Milgrom and Shannon (1994), the best response function must be strictly increasing in v , which in this case implies a threshold rule. It follows that any symmetric equilibrium must be in symmetric threshold strategies. So fix an equilibrium of the form in the theorem, and let the payoffs to taking taking BIN be $\pi_B(v)$ and to TAC be $\pi_T(v)$. They are given by:

$$\begin{aligned}\pi_B(v) &= E [1(v > Y^1 > \bar{v})(v - Y^1)] + E [1(Y^1 < \bar{v})(v - p)] \\ \pi_T(v) &= \mathbb{E} \left[1(Y^1 < \bar{v})1(Y^* < v)\frac{1}{d}(v - Y^*) \right]\end{aligned}$$

The threshold type \bar{v} must be indifferent, so

$$\begin{aligned}\pi_B(\bar{v}) &= \mathbb{E} [1(Y^1 < \bar{v})(\bar{v} - p)] \\ &= \mathbb{E} \left[1(Y^1 < \bar{v})\frac{1}{d}(\bar{v} - Y^*) \right] = \pi_T(\bar{v}).\end{aligned}\tag{6}$$

Next, we show a \bar{v} satisfying Eq. (2) exists and is unique. The right hand side of Eq. (2) is a function of \bar{v} with first derivative $\frac{1}{d}(1 - \frac{\partial}{\partial \bar{v}}\mathbb{E}[Y^*|Y^1 < \bar{v}]) < 1$. Since at $\bar{v} = 0$ it has value $p > 0$ and globally has slope less than 1, it must cross the 45° line exactly once. Thus there is exactly one solution to the implicit Eq. (2).

Proof of Theorem 2

To prove the first claim, let us fix d and \bar{v} . For any reserve price r , let $p(\bar{v}, r)$ denote the BIN-TAC price for threshold \bar{v} . By Equation (2) we have

$$p(\bar{v}, r) = \frac{d-1}{d}\bar{v} + \frac{1}{d}E [Y^*|Y^1 < \bar{v}]\tag{7}$$

Using first order conditions, we consider the effects of the marginal increase in reserve r on the revenue of BIN-TAC mechanism denoted by $\text{Rev}_{\text{BIN-TAC}}$. There are three cases:

- *The item is allocated via BIN:* If there are two bidders above \bar{v} , then increase of \bar{v} does not change the revenue. But the revenue increases by $\frac{\partial p(\bar{v}, r)}{\partial r}$ if a bidder wins the item at the buy-it-now price. This happens with probability $nF(\bar{v})^{n-1}(1 - F(\bar{v}))$. Hence the marginal increase in revenue from BIN auctions is equal to

$$\begin{aligned}
& nF(\bar{v})^{n-1}(1 - F(\bar{v})) \times \frac{\partial p(\bar{v}, r)}{\partial r} \\
&= nF(\bar{v})^{n-1}(1 - F(\bar{v})) \frac{1}{d} \Pr(Y^d \leq r | Y^1 < \bar{v}) \\
&= nF(\bar{v})^{n-1}(1 - F(\bar{v})) \frac{1}{d} \left(\frac{1}{F(\bar{v})^{n-1}} \left(\sum_{k=0}^{d-1} \binom{n-1}{k} (F(\bar{v}) - F(r))^k F(r)^{n-k} \right) \right) \\
&= (1 - F(\bar{v})) \left(\sum_{k=1}^d \frac{n}{d} \binom{n-1}{k-1} (F(\bar{v}) - F(r))^{k-1} F(r)^{n-k} \right) \tag{8}
\end{aligned}$$

- *The item is allocated via TAC:* The revenue of TAC changes only if the price is equal to r . In the event of a bidder winning an item at TAC and then paying the reserve price r , the revenue increases by the marginal increase of r . Observe that if there are k ($1 \leq k \leq d$) bidders with valuation between r and \bar{v} (and no bids above \bar{v}), then the revenue of the auction is equal to r with probability $\frac{k}{d}$. By this observation, the probability that the revenue is equal to r (and hence the marginal increase in the revenue) is given by

$$\begin{aligned}
& \left(\sum_{k=1}^d \frac{k}{d} \binom{n}{k} (F(\bar{v}) - F(r))^k F(r)^{n-k} \right) \\
&= \left(\sum_{k=1}^d \frac{n}{d} \binom{n-1}{k-1} (F(\bar{v}) - F(r))^k F(r)^{n-k} \right) \\
&= (F(\bar{v}) - F(r)) \left(\sum_{k=1}^d \frac{n}{d} \binom{n-1}{k-1} (F(\bar{v}) - F(r))^{k-1} F(r)^{n-k} \right) \tag{9}
\end{aligned}$$

- *The item is not allocated:* In the event that the bidder chosen by TAC cannot receive the item because his valuation was equal to r (before marginal increase) the revenue

decreases by r . In this case, the marginal decrease in the revenue is equal to

$$\begin{aligned} & r \left(\sum_{k=0}^{d-1} \frac{n}{d} f(r) \binom{n-1}{k} (F(\bar{v}) - F(r))^k F(r)^{n-k-1} \right) \\ &= r f(r) \left(\sum_{k=1}^d \frac{n}{d} \binom{n-1}{k-1} (F(\bar{v}) - F(r))^{k-1} F(r)^{n-k} \right) \end{aligned} \quad (10)$$

Summing up expressions (8), (9), and (10) we have

$$\frac{\partial \text{Rev}_{\text{BIN-TAC}(d,p(\bar{v},r))}}{\partial r} = (1 - F(r) - r f(r)) \left(\sum_{k=1}^d \frac{n}{d} \binom{n-1}{k-1} (F(\bar{v}) - F(r))^{k-1} F(r)^{n-k} \right)$$

Therefore, the derivative is equal to zero for the solution of $r = \frac{1-F(r)}{f(r)}$ denoted by r^* . By assumption, $\psi(v)$ single-crosses zero exactly once from below (in the region $[\underline{\omega}_L, \bar{\omega}_H]$), and so r^* is unique. Hence, for any \bar{v} , the optimal reserve is either r^* , or is one of the boundaries $\underline{\omega}_L$, $\bar{\omega}_L$, $\underline{\omega}_H$, and $\bar{\omega}_H$. Note that the derivative is positive at $\underline{\omega}_L$. Also, for reserve equal to $\bar{\omega}_H$, the item never sells. Moreover, observe that the reserve equal to $\bar{\omega}_L$ is dominated by reserve equal to $\underline{\omega}_H$ since no low-type bidder would receive the item. Therefore, the optimal reserve is either equal to r^* or $\underline{\omega}_H$.

Again, we use first order conditions again to find the optimal choice of \bar{v} . There are three effects on the revenue of the mechanisms by marginally increasing \bar{v} .

- *The highest valuation bidder now declines to take BIN:* This reduces revenue by $p(\bar{v}, r) - E[Y^* | Y^1 < \bar{v}] = \frac{d-1}{d} E[\bar{v} - Y^* | Y^1 < \bar{v}]$, and happens with probability $n f(\bar{v}) F(\bar{v})^{n-1}$.
- *The second highest valuation bidder now declines to take BIN:* This decreases revenue by $\bar{v} - p(\bar{v}, r) = \frac{1}{d} E[\bar{v} - Y^* | Y^1 < \bar{v}]$, and happens with probability $n(n-1) f(\bar{v}) (1 - F(\bar{v})) F(\bar{v})^{n-2}$.
- *Only the highest bidder takes BIN, and pays slightly more:* With probability $n(1 - F(\bar{v})) F(\bar{v})^{n-1}$, the highest bidder may have valuation above \bar{v} and the second highest below it. In this case revenue increases by $\frac{\partial p(\bar{v}, r)}{\partial \bar{v}} = \frac{d-1}{d} + \frac{1}{d} \frac{\partial E[Y^* | Y^1 < \bar{v}]}{\partial \bar{v}}$

Therefore, we have:

$$\begin{aligned}
\frac{\partial \text{Rev}_{\text{BIN-TAC}(d,p(\bar{v},r))}}{\partial \bar{v}} &= n(1 - F(\bar{v}))F(\bar{v})^{n-1} \times \\
&\quad \left(-\frac{f(\bar{v})}{1 - F(\bar{v})} \frac{d-1}{d} E[\bar{v} - Y^* | Y^1 < \bar{v}] - \frac{f(\bar{v})}{F(\bar{v})} \frac{n-1}{d} E[\bar{v} - Y^* | Y^1 < \bar{v}] \right. \\
&\quad \left. + \left(\frac{d-1}{d} + \frac{1}{d} \frac{\partial E[Y^* | Y^1 < \bar{v}]}{\partial \bar{v}} \right) \right) \\
&= \frac{n}{d} (1 - F(\bar{v}))F(\bar{v})^{n-1} \times \\
&\quad \left(-\left(\frac{(d-1)f(\bar{v})}{1 - F(\bar{v})} + \frac{(n-1)f(\bar{v})}{F(\bar{v})} \right) E[\bar{v} - Y^* | Y^1 < \bar{v}] \right. \\
&\quad \left. + \left(d-1 + \frac{\partial E[Y^* | Y^1 < \bar{v}]}{\partial \bar{v}} \right) \right)
\end{aligned}$$

Re-arranging terms, the optimal choice for \bar{v} is either at the boundaries or is a solution of the equation below

$$d - 1 + \frac{\partial E[Y^* | Y^1 < \bar{v}]}{\partial \bar{v}} = \left(\frac{(d-1)f(\bar{v})}{1 - F(\bar{v})} + \frac{(n-1)f(\bar{v})}{F(\bar{v})} \right) E[\bar{v} - Y^* | Y^1 < \bar{v}]$$

Similar to the previous argument, it is easy to see that the only boundary that would be optimal is $\underline{\omega}_H$.

Table 1: Revenue Comparison: Uniform Environment

Mechanism	OPT	SPA-T	BIN-TAC (d=2)	BIN-TAC (d=3)	SPA-B
Expected Revenue	0.89	0.76	0.85	0.83	0.81
Expected Consumer Surplus	0.51	0.67	0.48	0.40	0.16

Expected revenue and welfare under different mechanisms, for the uniform environment with $\Delta = 3$, $\alpha = 0.05$ and the number of bidders $n = 5$.

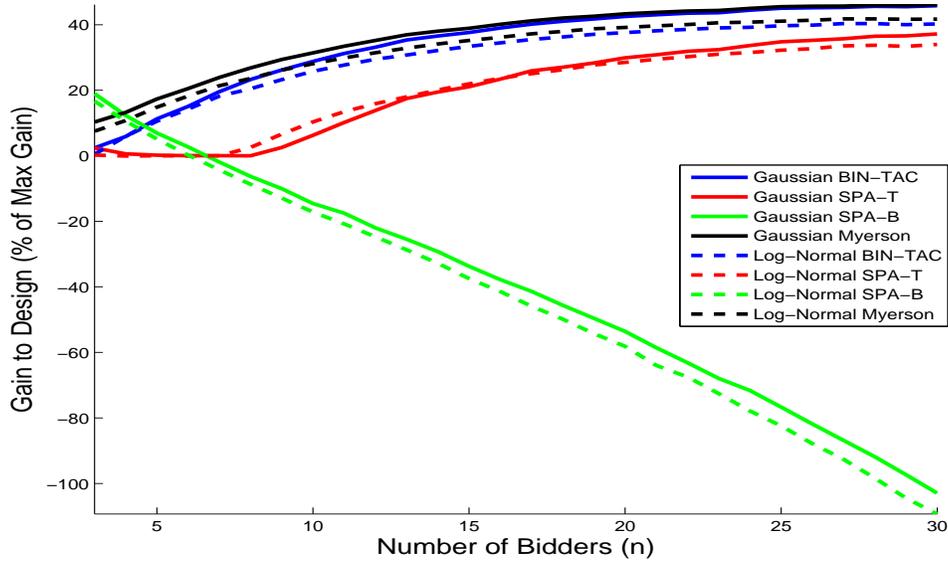


Figure 5: **Revenue Performance vs Number of bidders.** Simulated expected revenues for different mechanisms as the number of bidders n varies, in an environment where valuations are independent, F_L has mean 1 and standard deviation 0.5, the match probability is 0.05 and the match increment is 5.

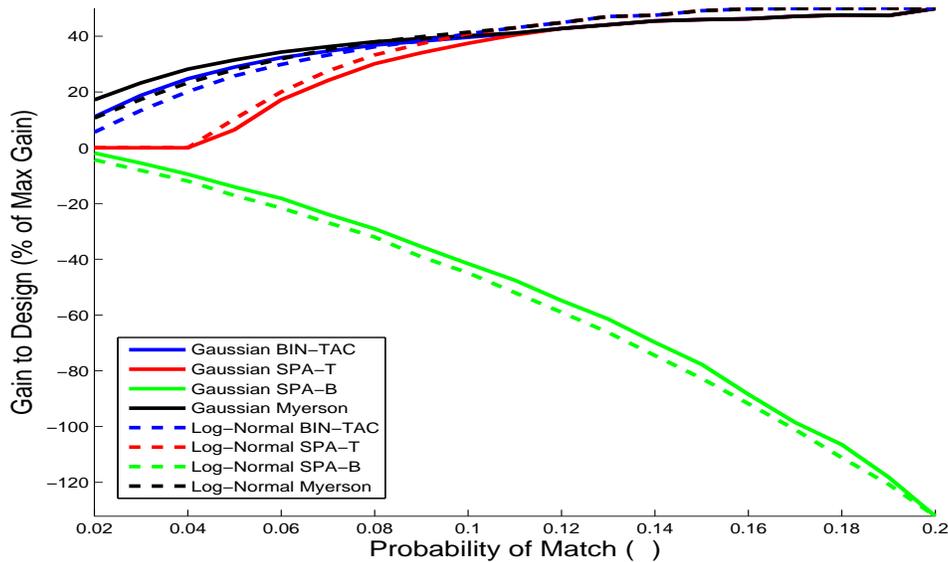


Figure 6: **Revenue Performance vs Match Probability.** Simulated expected revenues for different mechanisms as the probability of a match α varies, in an environment where valuations are independent, F_L has mean 1 and standard deviation 0.5, the number of bidders is 10 and the match increment is 5.

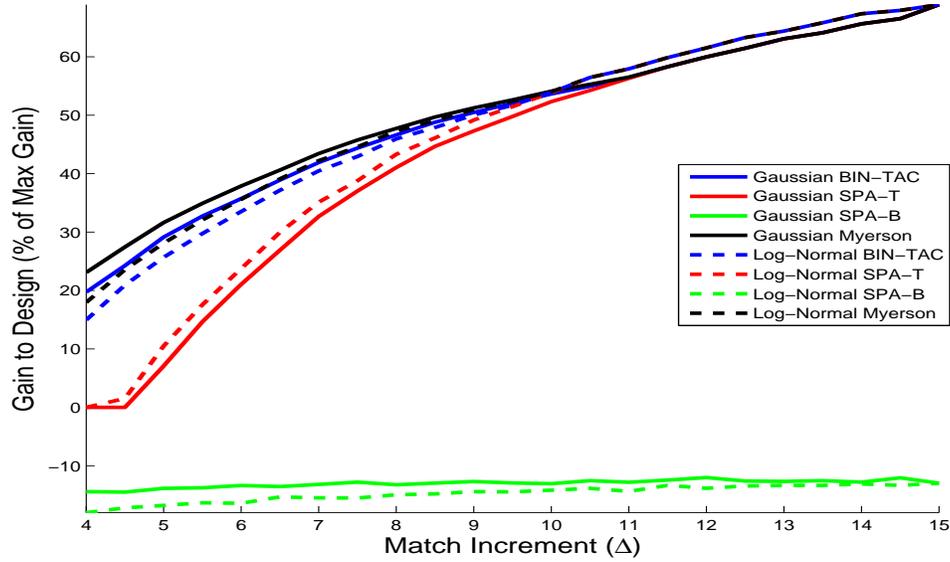


Figure 7: **Revenue Performance vs Match Increment.** Simulated expected revenues for different mechanisms as the match increment Δ varies, in an environment where valuations are independent, F_L has mean 1 and standard deviation 0.5, the match probability is 0.05 and the number of bidders is 10.

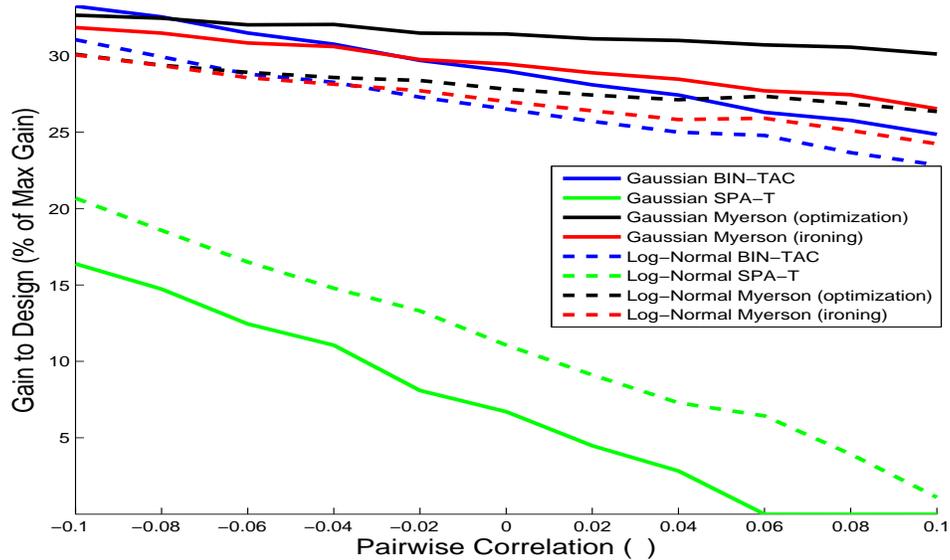


Figure 8: **Revenue Performance vs Pairwise Correlation.** Simulated expected revenues for different mechanisms as the pairwise correlation ρ varies, where F_L has mean 1 and standard deviation 0.5, the match probability is 0.05, the match increment is 5 and the number of bidders is 10.

Table 2: Summary Statistics: Microsoft Advertising Exchange Display Ad Auctions

	Mean	Median	Std. Dev.	Min	Max
Bid-Level Data					
Average bid	1.000	0.565	2.507	0.0000157	130.7
Number of bids	508036				
Auction-Level Data					
Winning bid	2.957	1.614	5.543	0.00144	130.7
Second highest bid	1.066	0.784	1.285	0.00132	39.22
Number of bidders	6.083	6	2.970	1	15
Bid correlation	0.01				
Number of auctions	83515				
Advertiser-Level Data					
% of auctions participated in (p1)	0.697	0.0251	4.641	0.00120	88.28
% of auctions won if participated (p2)	38.90	29.59	35.50	0	100
Correlation of (p1,p2)	-0.09				

Summary statistics for the full dataset, which is a 0.1 percent sample of a week's worth of auction data sampled within the last two years. An observation is a bid in the top panel; an auction in the middle panel; and an advertiser in the last panel. Bids have been normalized so that their average is 1, for confidentiality reasons. The bid correlation is measured by selecting a pair of bids at random in every auction with at least two bidders, and computing the correlation coefficient.

Table 3: Matching on Region

	Participation		Bids	
Advertiser Website Pageview Ratio	0.029 (0.022)	0.329*** (0.015)	0.264*** (0.052)	0.286*** (0.053)
Time-of-Day Fixed Effects	yes	yes	yes	yes
Product-Region Fixed Effects	yes	yes	yes	yes
Advertiser Fixed Effects	no	yes	no	yes
N	5581749	5581749	417557	417557
R^2	0.02	0.34	0.04	0.26

Results from OLS Regressions. In the first two columns, the dependent variable is a dummy for participation. The sample used in the regressions consists of all auction-bidder pairs, limited to the 10% of bidders who participate most often. In the last two columns, the dependent variable is the bid. The sample used in the regressions only includes bids from the 10% of bidders who bid most often. The independent variable is the population-weighted fraction of pageviews of the advertiser's website that come from the region the user is in. Time-of-day fixed effects refer to a dummy for each quarter of the day, starting at midnight. Product-region fixed effects are dummies for the page-group advertised on, and the state the user is located in. Standard errors are robust. Significance levels are denoted by asterisks (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 4: Optimal Parameter Choices

Policy	d	r	p	\underline{v}	\bar{v}
SPA	-	1.96	-	-	-
BIN-TAC (incomplete information)	3	0.26	1.83	-	-
BIN-TAC (complete information)	3	0.44	1.88	-	-
BIN-TAC (worst case)	4	0.65	2.11	-	-
Myerson (single numerically optimized ironing region)	-	0.65	-	0.65	2.59

Revenue-maximizing parameter choices. We find these by maximizing the revenue functions defined in the main text over the available parameters using a global optimization algorithm based on differential evolution. For the Myerson mechanism \underline{v} and \bar{v} are the left and right endpoints of the ironing region.

Table 5: Counterfactual Revenues and Welfare

Policy	Revenue	Advertiser Surplus	Total Surplus
SPA (no reserve)	0.983 (0.004)	1.974 (0.019)	2.957 (0.020)
SPA (optimal reserve)	1.028 (0.005)	1.471 (0.019)	2.499 (0.020)
BIN-TAC (incomplete information)	1.072 (0.006)	1.685 (0.019)	2.757 (0.020)
BIN-TAC (complete information)	1.074 (0.004)	1.641 (0.018)	2.715 (0.020)
BIN-TAC (worst case)	1.069 (0.005)	1.526 (0.018)	2.594 (0.020)
Myerson (single optimized region)	1.056 (0.004)	1.713 (0.019)	2.769 (0.020)
Myerson (ironing algorithm)	1.023 (0.005)	1.822 (0.020)	2.846 (0.021)
Bundling by Product-User Region	0.644 (0.006)	0.730 (0.016)	1.374 (0.015)

Counterfactual simulations of average platform revenues, advertiser surplus and total surplus. Statistics outside parentheses are averages across impressions; those in parentheses are standard errors computed by bootstrapping. Eight different simulations are run. The first is a second price auction without reserve, while the second is a second price auction with optimal (revenue-maximizing) reserve. The third is the BIN-TAC mechanism, in a symmetric correlated private values model. In the fourth case we re-simulate the mechanism, assuming complete information. In the worst-case model, bidders assume they will only have to pay the reserve price in TAC auction, and therefore take the BIN option more rarely. The sixth and seventh implement the Myerson (1981) mechanism with ironing. The first time we allow only one region of pooled virtual valuations, found by numerical optimization; the second time we iron the virtual valuations in the standard way. The eighth is a no-targeting counterfactual where the impressions are bundled according to the product (i.e. URL and ad size) and user region, and sold by second-price auction. Where applicable, the parameters used are the optimal parameters from Table 4.