Buy-it-now or Take-a-chance: Price Discrimination through Randomized Auctions *

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

October 17, 2012

Abstract

Online tracking technology allows platforms to offer advertisers targeted consumer demographics, improving match quality, but thinning markets. Bidding data from Microsoft Advertising Exchange exhibits a large gap in the top two bids, consistent with this intuition. This motivates our new mechanism. Bidders can "buy-it-now", 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. Running counterfactual simulations on our data, we find it improves revenue by 4.4% and consumer surplus by 14.5% compared to an optimal second-price auction.

*Department of Computer Science, University of Washington (ecelis@cs.washington.edu), Department of Economics, Harvard University (glewis@fas.harvard.edu), Microsoft Research New England (mobius@microsoft.com), and USC Marshall School of Business (nazerzad@marshall.usc.edu) respectively. All authors are grateful to Microsoft Research New England for their hospitality, and Greg would also like to thank NET Institute (www.NETinst.org) for financial support. Josh Feng and Danyang Su provided excellent research assistance. We thank Dan Ackerberg, Susan Athey, Alessandro Bonatti, Michael Grubb, Jakub Kastl, Jon Levin, Mallesh Pai, Jimmy Roberts, Hal Varian, Ricky Vohra and Lee Zen for fruitful discussions that improved this work.

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 increasing amounts of 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 2010, Levin and Milgrom 2010).

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

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

[1] Section 4 discusses the data in detail.

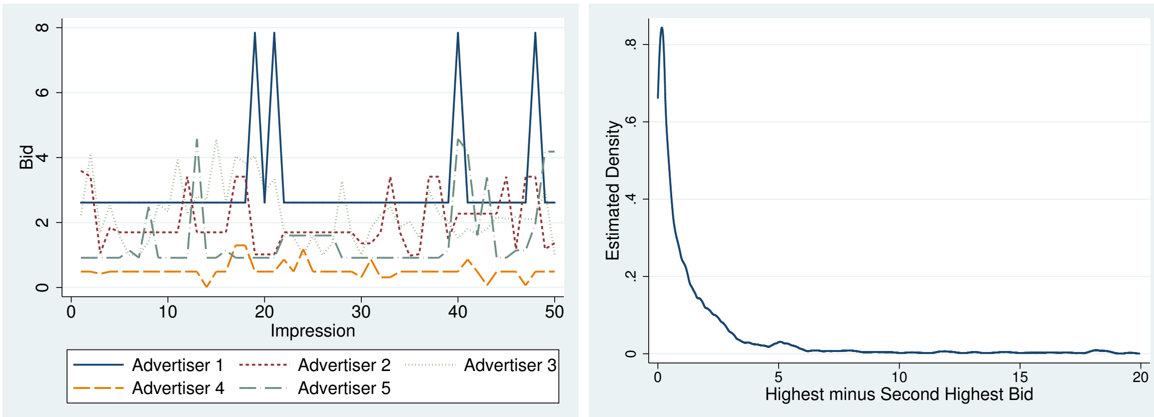

Figure 1: **Bids over Time and Bidding Gap.** The left pabel shows the (rescaled) bids of five advertisers in our data, selected at random from the top 50 advertisers (ranked by purchases) on 50 randomly chosen successive impressions of the most popular product. Note that the set of impressions differs across bidders (there are no impressions on which all 5 participate). The right panel shows a kernel density estimate of the probability density function of the (normalized) gap between the highest and second highest (rescaled) bids in auctions for the product with the highest sales volume in our dataset.

probabilities are independent across bidders, and sufficiently low that the probability that *any* bidder matches is relatively small. Then a second-price auction will typically get low revenue, since the probability of two "matches" occurring in the same auction is small. On the other hand, setting a high fixed price is not effective since the probability of zero "matches" occurring is relatively large and many impressions would go unallocated. Hence, allowing targeting creates asymmetries in valuations that can increase efficiency, but decrease revenue. Because of this, some have suggested it is better to create thicker markets by not disclosing information, thus "bundling" many different impressions together (Ghosh, Nazerzadeh and Sundararajan 2007, Even-Dar, Kearns and Wortman 2007, McAfee, Papineni and Vassilvitskii 2010). The question of how to optimally bundle is a subject of ongoing research (Bergemann, Bonatti and Said 2011).

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 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 auction mechanism ensures the impression is still allocated, thereby earning revenue.

BIN-TAC is simple, both in that it is easy to explain to advertisers 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. As we show both analytically and through monte carlo simulation, BIN-TAC generally outperforms the two leading alternatives: a second price auction with reserve, or the "bundling" solution in which the platform withholds targeting information. One could do better still by using the revenue-optimal mechanism suggested in Myerson (1981), which is considerably more complicated. We demonstrate via numerical simulations that BIN-TAC in simple environments approximates the allocations and payments of the optimal mechanism, achieving "almost optimal" performance. This makes BIN-TAC a practical alternative.

To analyze its performance in a real-world setting, we turn to historical data from the Microsoft Advertising Exchange. By estimating the distribution of advertiser valuations, we can simulate the effect of introducing the BIN-TAC mechanism. We also consider a bundling strategy in which all impressions on a given webpage browsed by a user located in a particular geographic region are sold as identical products. 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 consumer 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. The Myerson mechanism does better still, but requires the use of seven ironing regions: a complex price discrimination scheme to be sure. All of them outperform the bundling strategy, although we cannot rule out better performance from an optimal bundling strategy.

We view the main contribution of our paper as introducing and analyzing a new and simple price discrimination mechanism that makes use of randomized auctions, and then testing its performance in a realistic environment. While our focus is on the display advertising market, we note that there are other markets in which randomized allocations are used as a screening tool. For example, Priceline offers users the choice between a hotel of their choice at a high fixed price, or the opportunity to bid for a random hotel room of certain guaranteed characteristics (e.g. location, star rating).

A secondary contribution of the paper is to document participation and bidding behavior in the display advertising market. 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. We document that there is a large gap between the highest and second highest valuations in these auctions, consistent with targeting creating thin markets. We also show that advertisers vary their bids based on the location of their users, using the user demographics provided by the platform to achieve better matches. Finally, we believe we are the first to calculate and simulate revenues from the Myerson (1981) mechanism with ironing using a real data set, which may be of interest to the many people working on problems in optimal mechanism design.

Related Work: Our work is related to the literature on price discrimination and screening. Here we consider a mechanism that treats all bidders symmetrically, and proceeds sequentially. Other papers have suggested sequential screening approaches. In one setting, the buyers themselves learn their type dynamically, in two stages (Courty and Li 2000). In this case, offering contracts after the first type revelation but before the second may be optimal; see Bergemann and Said (2010) for a survey on dynamic mechanisms. In the static setting, sequential screening and posted-price mechanisms can be used to design optimal (or near-optimal) mechanisms when the bidders have multi-dimensional private information (see for example Rochet and Chone (1998) and Chawla, Hartline, Malec and Sivan (2010)).

More generally, the question of whether sellers should provide information that allows buyers to target their bids arises in the analysis of optimal seller disclosure (see for example Lewis and Sappington (1994) and Bergemann and Pesendorfer (2007)). The idea of bundling goods together to take advantage of negative correlation in valuations — in this case the negative correlation in the valuations from "match" or "no match" — dates back to Adams and Yellen (1976); see also McAfee, McMillan and Whinston (1989). Our paper is similar in style to

Chu, Leslie and Sorensen (2011), who combine theory, simulations and empirics to argue that bundle-size pricing is a good approximation to the more complicated (but theoretically superior) mixed bundling pricing scheme for a monopolist selling multiple goods.

We focus on a private values setting, while Abraham, Athey, Babioff 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. Our paper is also related to the literature on buy-now auctions (Budish and Takeyama 2001, Reynolds and Wooders 2009).

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 proceeds 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 monte carlo simulation. Finally, in the third part we provide an empirical analysis of a display advertising marketplace, including counterfactual simulations of our mechanism's performance. All proofs are contained in the appendix.

2 The Display Advertising Market

This paper proposes a new second degree price discrimination strategy for advertising platforms such as Microsoft, Google and Facebook. In these markets, advertisers care about the characteristics of the users they advertise to, but it is up to the platform to choose whether or not to disclose what they know about their users. The online display advertising market is an example of such a market. Its organization 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. 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 (e.g. 1.5% of total worldwide Internet pageviews are on Yahoo![2]), 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.[3] Advertising exchanges are a minor technological wonder. They work 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 pre-

[2]Source: alexa.com

[3]"In Sept 2009, RightMedia averaged 9 billion transactions a day with hundreds of thousands of buyers and sellers." Muthukrishnan (2010)

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

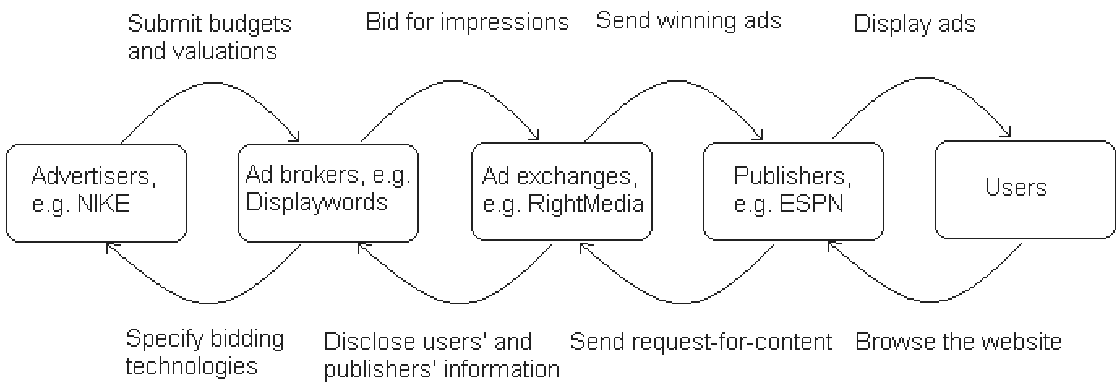

Figure 2: **The Display Advertising Market.**

viously 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 placed in the auction 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 have two main advantages over direct negotiation. First, they economize on transaction costs, by creating a centralized market for selling ad space. Second, they allow for very detailed products to be sold, such as the attention of a male 15-24 year old living in

[5]To 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, which is what we analyze 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 through a single expected revenue number.

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. We will focus on developing a real-time pricing mechanism for display advertising exchanges.

3 Model and Analysis

3.1 The Environment

A seller (publisher) has an 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. 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 $[\underline{\omega}_L, \overline{\omega}_L]$ and F_H has support $[\underline{\omega}_H, \overline{\omega}_H]$, and that these supports are disjoint (so $\overline{\omega}_L < \underline{\omega}_H$). 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 $\alpha \in (0, 1)$.

[6]Rougly speaking, a cookie refers to the information sent from the browser of the user to the website visited by user and can be used to store the state of the communication between them and other information about the user (RFC6265 2011).

The bidder type is a triple (X, V_L, V_H), drawn identically and independently across bidders. 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 $E[V] = (1-\alpha)V_L + \alpha V_H$.

For simplicity of the presentation, we also make some technical assumptions on the *virtual valuations* $\psi(v) = v - \frac{1-F(v)}{f(v)}$.[7] We assume that $\psi(v)$ is continuous and increasing over the regions $[\underline{\omega}_L, \overline{\omega}_L]$ and $[\underline{\omega}_H, \overline{\omega}_H]$. We additionally assume that $\psi(v)$ single-crosses zero, that this intersection occurs in the low valuation region $[\underline{\omega}_L, \overline{\omega}_L]$, and that $\psi(\overline{\omega}_L) \leq \psi(\underline{\omega}_H)$. Overall, our environment is fully characterized by the tuple (n, α, F_L, F_H).

Discussion: We assume that the match random variables X and the valuations V_L and V_H are independent across bidders. We focus on independence for two reasons. First, it is an assumption that is often made in the screening and mechanism design literatures, and so is a natural starting point. Second, in the log data examined in this paper we observe little correlation in bids.[8]

We also will focus on environments where α is small, since this implies 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 then there are essentially 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 for what follows. First though we need to describe our BIN-TAC mechanism.

[7] Without these assumptions we would have to analyze multiple cases, which is straightforward but tedious.

[8] That bids are not positively correlated certainly does not prove that the underlying valuations are not positively correlated — it could just be that informational and technological constraints prevent advertisers from fully expressing their preferences — but it does provide some support for our assumption.

3.2 Pricing Mechanisms

Our BIN-TAC mechanism 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).

A third benchmark is the revenue-optimal mechanism within the class of those that allow targeting (i.e. those that commit to reveal the match information to all bidders for free).[9] Usually this mechanism is the second-price auction with an optimally chosen reserve price. However in this case the virtual valuations $\psi(v)$ are not increasing over the whole support of F — indeed they are (infinitely) negative over the region $(\overline{\omega}_L, \underline{\omega}_H)$. The optimal mechanism may require ironing (Myerson 1981).

In plain terms, ironing implies that sometimes the allocation will be randomized among bidders with different valuations. Just as in our TAC auction, the winner of the auction 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. By contrast, in BIN-TAC this randomization occurs whenever no-one takes the BIN option. The differences will be clearer later when we compare the performance of the mechanisms. Right now, we start the analysis with the two type case.

[9] A seller may potentially do better by withholding match information from everyone (bundling), or by giving different information to different bidders — see Bergemann and Pesendorfer (2007).

3.3 The Two-Type Case

In this simplified environment, 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:

$$\max_{d,p} \;=\; \min\left\{\frac{n}{d},1\right\}(1-\alpha)^n v_L + n\alpha(1-\alpha)^{n-1}p + \left(1-(1-\alpha)^n - n\alpha(1-\alpha)^{n-1}\right)v_H$$

$$\text{subject to} \qquad v_H - p \geq \frac{1}{d}(v_H - v_L) \qquad (IC) \tag{1}$$

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 the case where no-one else takes 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 $\left(q_0 + \frac{q_1}{n}\right) v_L + \left(1 - q_0 - \frac{q_1}{n}\right) v_H$, 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 $\left(q_0 + \frac{q_1}{n}\right)$ in BIN-TAC.[10]

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 than under SPA-B (i.e. $\left(q_0 + \frac{q_1}{n}\right) < (1-\alpha)$), 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.

This last revenue result is perhaps surprising: bundling removes all private information, eliminating all information rents and extracting all the surplus from the buyers. But of course revealing the match information creates additional surplus, and since running the TAC auction causes no distortion in the allocations at the bottom, BIN-TAC is able to extract most of the rents it creates by revealing the match information.

Finally, it is not hard to show that BIN-TAC is in fact revenue-optimal within the class of all mechanisms that commit to targeting.[11] 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). So it may be a good auction mechanism, which motivates an analysis of the general case with a continuum of types.

[10] This 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.

[11] As 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.

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.[12] Truth-telling is also weakly dominant in the TAC auction.[13] Thus, overall, BIN-TAC is Bayes-Nash incentive compatible; both BIN and TAC are dominant-strategy incentive compatible, and the choice of which of the two to participate can be maximized based on other agent's expected behavior.

Taking these strategies as given, we turn to the decision of which auction, buy-it-now or take-a-chance, an agent will choose. 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: $\exists\, \overline{v}$ such that types with $v \geq \overline{v}$ 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^* be the maximum of Y^d and the TAC reserve r.

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

$$\overline{v} = p + \frac{1}{d} E\left[\overline{v} - Y^* | Y^1 < \overline{v}\right] \tag{2}$$

where types with $v \geq \overline{v}$ 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?

[12] Since 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.

[13] The 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 of the object 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.

If 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 $\frac{1}{d}E\left[v - Y^*|Y^1 < v\right]$, since they only win with probability $\frac{1}{d}$, although their payment of Y^* is on average much lower. Equating these two to find the indifferent type $\overline{v}$ yields Equation (2).[14]

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:

Proposition 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*

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

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

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

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 the usual incentive compatibility trade-offs apply, as the TAC reserve is relevant for the BIN choice. 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 $\overline{v}$ 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. This is why we get the usual solution.

[14] The above theorem does not require our mixture distribution environment: it remains true for an IPV environment with arbitrary absolutely continuous F.

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 $\overline{v}$. 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 $\overline{v} - p(\overline{v}, r^*)$. Finally, if the highest bidder is above the reserve and the second highest is below, an increase gains the seller $\frac{\partial p(\overline{v}, r^*)}{\partial \overline{v}}$. Working out the probabilities of these various events, expanding $\frac{\partial p(\overline{v}, r^*)}{\partial \overline{v}}$ 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. $\underline{\omega}_H \gg \overline{\omega}_L$). 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 any efficiency or revenue performance 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 the general model: BIN-TAC trades-off increased revenues from high types (by providing match information) against lower revenues and inefficient allocations to low types. This can generate less revenue than limiting information rents through bundling. But we will see that in Monte Carlo simulations below, BIN-TAC generally outperforms bundling.

Clearly BIN-TAC will have (weakly) worse revenue performance than the revenue-optimal mechanism with targeting. The question is how close BIN-TAC gets. Let $\widetilde{v}$ solve $-F(\widetilde{v})^2 + (2-\alpha)F(\widetilde{v}) + \alpha(\underline{\omega}_H - \widetilde{v})f(\widetilde{v}) = 1-\alpha$. and define the ironed virtual valuations as follows[15]:

$$\phi(v) = \begin{cases} 0 & v \in [\underline{\omega}_L, r^*) \\ \psi(v) & v \in [r^*, \widetilde{v}] \\ \psi(\widetilde{v}) & v \in (\widetilde{v}, \underline{\omega}_H) \\ \psi(v) & v \in [\underline{\omega}_H, \overline{\omega}_H], \end{cases} \tag{5}$$

The allocation procedure works as follows: award the good to the bidder with the highest ironed virtual valuation, breaking ties at random, provided the virtual valuation is positive. Payments are uniquely determined by ex-post incentive compatibility and the requirement that losers pay zero. Notice that all types between $\widetilde{v}$ and $\underline{\omega}_H$ get the same ironed virtual valuations, and therefore if they tie, the winner is selected at random. Like BIN-TAC, this is inefficient, but allows additional revenue extraction from higher types.

Having obtained this characterization, we can compare BIN-TAC with the optimal mechanism. For now, let us focus on a simple environment, where F_L is uniform over $[0,1]$ and F_H is uniform over $[\Delta, \Delta+1]$. Figure 3 shows the interim allocation probabilities (top panel) and expected payments by type (bottom panel) as a function of bidder type, in the case where $\Delta = 3$, $\alpha = 0.05$ and $n = 5$ (with optimal parameter choices). The optimal mechanism has a discontinuous jump in the allocation probability at $\widetilde{v} = 0.676$, 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 $\widetilde{v}$ with a smooth curve, by randomizing the allocation in that region using the TAC auction. By contrast, the slope of the SPA-T allocation schedule is steep on this region and so the SPA-T cannot extract revenue from the high types (who could easily pretend to be a lower type while barely changing their probability of winning).

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

[15] We derive the optimal mechanism in the online appendix, showing that the interesting case occurs when $\alpha\underline{\omega}_H < r^*(1-F(r^*))$, for r^* the optimal reserve of equation (3) (when this fails, the optimal mechanism is just a second price auction with $r^* \in [\underline{\omega}_H, \overline{\omega}_H]$). We assume this condition in what follows.

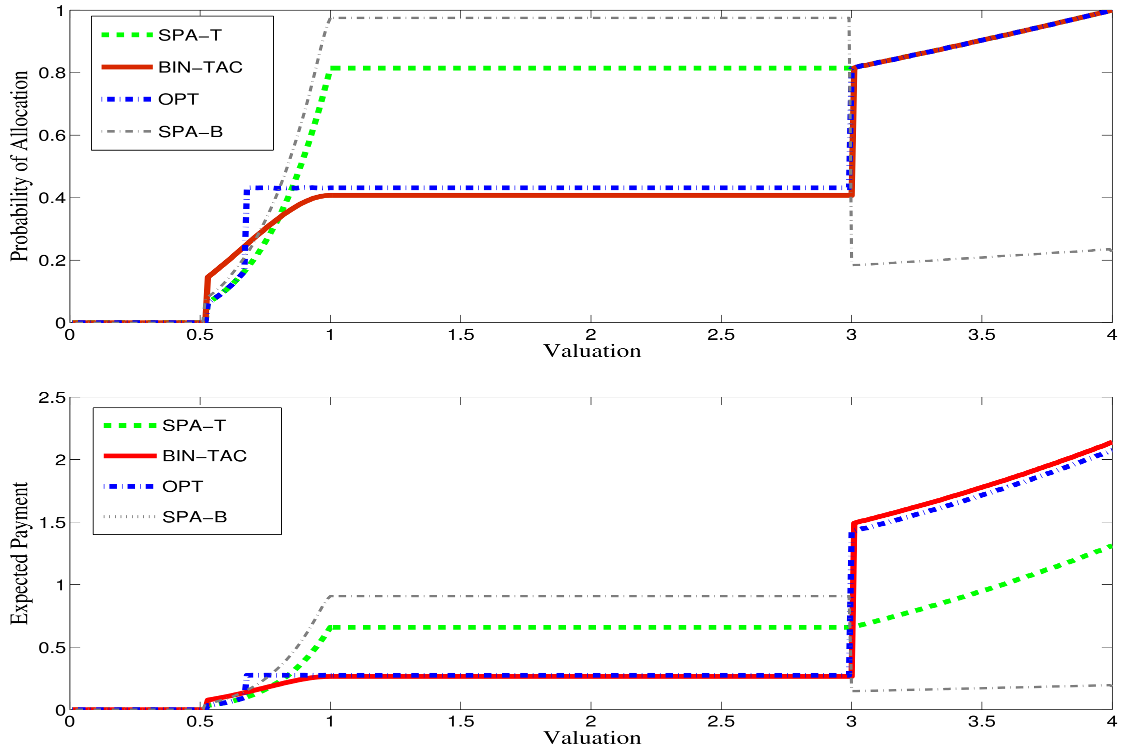

Figure 3: **Comparison of Allocations and Payments.** Allocation probabilities (top panel) and expected payments (bottom panel) for the OPT, SPA-T, SPA-B and BIN-TAC mechanisms when the distributions F_L and F_H are uniform. The x-axis corresponds to the bidder's valuation (unknown to them under bundling).

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 don't know they haven't matched), and substantially lowers them in the case of a match.

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

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 $n = 5$.

3.6 Monte Carlo Simulations

We would like to test our mechanism against the benchmarks in a variety of other settings. We drop the assumption that F_L and F_H have disjoint support. The optimal BIN-TAC mechanism remains easy to calculate. Nothing in the proof of Theorem 3 required the disjoint supports for determining r^* and p^*, and so these can be solved for numerically for each d. Thus the optimization problem reduces to a one dimensional discrete optimization problem in the randomization parameter d, which can be quickly solved. Finding the optimal mechanism is more challenging, but can be done using standard optimization techniques.

For our simulations, we restrict ourselves to location families where the distribution $F_H(\cdot) = F_L(\cdot - \Delta)$ for some shift-parameter Δ, as in the uniform case above. 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. Each experiment is repeated for 100000 impressions, and we calculate the average revenues.

The results are presented in Figures 5, 6 and 7. In all cases, on the y-axis we plot the revenue as a fraction of the revenue from the optimal mechanism. Recall that BIN-TAC generalizes SPA-T, so its performance is always at least as good, and often significantly better. In all cases, the BIN-TAC extracts at least 90% of the optimal revenue, compared to a worst-case performance of around 82% for the SPA-T.

The SPA-B in some cases does even better than OPT, but its performance sharply degrades as the probability or value of a match gets large. We see this in Figures 5 and 6. The expected number of matches is αn, and so as either α or n increases, the performance of the mechanisms that allow targeting improves relative to the SPA-B.

Figure 7 shows the dependence on the gap Δ. As expected, the performance of BIN-TAC increases while that of SPA-T falls as Δ gets larger, over some range. Since there is more revenue to be gained from high-valued bidders, BIN-TAC can only perform better with a large Δ. For sufficiently high Δ though, both BIN-TAC and the SPA set high reserves, "throwing away" low-valued impressions and extracting all their revenue from matches, with equal revenue performance.

Overall, the performance of BIN-TAC is very good, at least for the distributions and parameters chosen. The main caveats are that it doesn't perform well with very few bidders (when bundling is preferable), and has little to recommend it when matches are highly probable or very valuable (a second-price auction would do as well). Its niche is in markets with relatively large numbers of bidders but low match probabilities, so that markets are "thin" in the sense of having relatively low matches in expectation.

3.7 Extensions and Related Models

Risk Aversion: We know that in many contexts bidders exhibit risk aversion. Risk aversion is generally helpful for BIN-TAC, as it makes the BIN option more attractive. Specifically, consider an equilibrium of the model with risk neutral bidders, characterized by a threshold bidder $\overline{v}$, and consider making the bidders marginally risk averse. Bidding one's valuation remains a weakly dominant strategy in the BIN or TAC auctions, should they occur. But now the threshold type $\overline{v}$ strictly prefers the BIN option, as in the pivotal case when there are no higher valuation types the BIN option delivers the object with certainty, whereas the TAC auction involves risk. This implies that the new equilibrium threshold is $\overline{v}' < \overline{v}$, which in turn implies increased revenue for the same choices of parameters. Thus our mechanism has good performance under risk aversion.

Selling the information: Up to now, we have considered only two binary extremes: the platform doesn't reveal the targeting information, or it does. Eso and Szentes (2007) have argued that a dynamic mechanism may do better still by "selling the information": the platform commits to reveal the targeting information, but charges a participation fee that essentially extracts the expected information rents that this disclosure creates in advance. There are two problems with applying this result in our environment. The first is practical:

in a competitive environment, charging participation fees may quickly cause advertisers to defect to rival platforms. The second is technical: here, bidders start with two pieces of private information V_L and V_H. These become a single one-dimensional private valuation V under either bundling (as $(1-\alpha)V_L + \alpha V_H$) or targeting (as $(1-X)V_L + XV_H$), so that the static mechanism design problem is entirely standard. But in the dynamic mechanism, the future promises and fees for participation must be functions of both V_L and V_H, which takes us into the realm of multidimensional mechanism design and outside of the original Eso and Szentes (2007) model. It is an open question as to whether there is a simpler implementation of this idea that, like BIN-TAC, approximates the optimal sequential mechanism.

Inefficient Rationing and Sales: Our mechanism screens types using the threat of inefficient rationing, which is not an element of classic screening models (e.g. Maskin and Riley (1984)). Sales have the same logic: by committing to sell an object at a discounted price in the future to the first person willing to pay the sale price, the seller may be able to induce purchase in the pre-sale period by high valuation buyers who know that their chance of getting it on sale is small. Just as in our model, the sale is useful to guarantee that the object is sold in the case where there are no high valuation buyers. In the online appendix, we formalize this logic in a two-period two-type model, where in each period buyers with valuation either v_H or v_L arrive in a random order and have the opportunity to buy at a price set by the monopolist. We show that the monopolist can maximize revenue by committing to a declining price sequence in the first period (lower for each successive buyer), followed by a flat price of v_L in the "sale period". This suggests a rationale for sales as a price discrimination tool.[16]

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

[16]Nocke and Peitz (2007) make the same point in a model with aggregate demand uncertainty, where the number of high type buyers depends on the state of the world.

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, and call that a region. We use controls for product and region throughout the descriptive regressions. Unfortunately, we don't 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.[17] We observe who won each auction and the final price. We drop auctions in which the eventual allocation was determined by biased bids and modifiers.[18] We also restrict attention to impressions that originate in the US, and where the publisher content is in English. Finally, we restrict only to reasonably 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

[17]In 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.

[18]When the advertiser has a technologically complex kind of 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 purposes, and downward for payment). It is hard to know how to treat these auctions without taking some kind of stand on whether valuations vary with the kind of ad being displayed, what contract each bidder has, what other bidders know about these contracts etc.

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.

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 out 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.[19] 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

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

region should participate more frequently and bid higher on users from those regions.[20] 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 top 10% of advertisers.[21] 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 find 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 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 virtual valuations $\psi(v)$ as a function of the estimated valuations.

[20] Since the pageview data dates from a period before our exchange data we are not worried about reverse causality (i.e. advertisers who win more impressions from region X later get more views from region X).

[21] Fortunately 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.

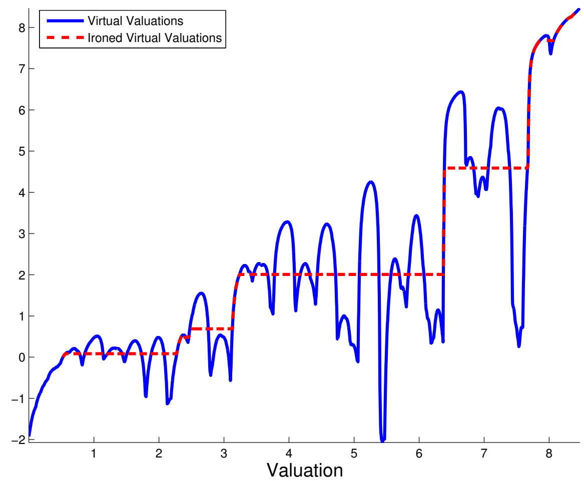

Figure 4: **Virtual Valuations.** This shows the virtual valuations (solid) as a function of the estimated valuations. It also shows the ironed virtual valuations (dashed), using ironing regions that would be optimal if the environment was symmetric independent private values.

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. We test this in the next section.

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. To do this, we proceed in a number of steps. First, we make a sequence of assumptions that allow us to infer individual bidder valuations from the observed bids in the data. Second, we show how to use these valuations to calculate counterfactual revenues for particular sales mechanisms. Third, we optimize over the various parameter choices to get the revenue-optimal version of each sales mechanism. In this optimization step we only consider a single choice of parameters for all auctions, ruling out different reserves or randomization parameters by product or user-region. 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 is explained more carefully below.

Step 1: Inference: Our theory is for a single auction rather than a sequence of auctions over time, and so in order to perform inference we need to enrich the model. As is standard in the literature, we make a number of assumptions that allow us to analyze each auction in isolation.[22] A sequence of impressions is auctioned over a time period. We assume there is a fixed set of N advertisers who are present throughout that time period, and who have zero costs of participating in an auction. Let $i = 1 \dots N$ index the advertisers, and let $t = 1 \dots T$ index the impressions. Valuations $v_{i,t}$ for each impression are private, and the payoff to winning an impression is $v_{i,t} - p_t$, where p_t is the price paid (losers get nothing). We assume that advertisers seek to maximize the sum of their payoffs across all impressions.

Under these assumptions, and focusing on equilibria in weakly undominated strategies, we can make quite strong inferences from the observed second-price auction data. Whenever an advertiser has a strictly positive valuation for an impression, it is weakly dominant to participate in the auction and bid their valuation. Consequently, 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 $v_{i,t} = b_{i,t}$.

How reasonable are our assumptions? One important assumption is that valuations are private. Following our discussion in section 3.7, it may be that bidders are asymmetrically informed about the users, due to re-targeting technology, implying common values. We stick with 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; 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 $\pi(v_{i,t}) \le c_i$, 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.

[22] But see Backus and Lewis (2012) for a model of a dynamic auction market with directed search.

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.[23]

Step 2a: Simulation for Robust Mechanisms: Two of the benchmark mechanisms — the second-price auction with targeting, and the Myerson mechanism with ironing — have the attractive property that bidding one's valuation is weakly dominant. So for these any particular version of these mechanisms (i.e. varying reserve prices and ironing regions), we can calculate revenue directly by substituting the inferred valuations in as bids.

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 affiliated private values. Advertiser behavior is characterized by a threshold value $\overline{v}_j = \overline{v}_j(p, d, r)$ for each product, above which they will take the BIN option, and below which they will TAC. This threshold solves the implicit equation $\overline{v}_j - p = \frac{1}{d}E[\overline{v}_j - Y^*|Y^1 < \overline{v}_j, \overline{v}_j]$. Notice that we need to condition on $\overline{v}_j$ itself on the RHS now, since under affiliation a bidder's own valuation is informative as to rival valuations and thus bids.

To solve this equation for fixed (p, d, r), we estimate the expected TAC payment as a function of the highest bid in the auction. Under symmetry, this payment can be non-parametrically estimated product-by-product (Pagan and Ullah 1999):

$$\text{Expected TAC Payment}(s, r, j; h) = \frac{\sum_{t \in J(j)} K\left(\frac{b_t^{(1)} - s}{h}\right) \max\{b_t^{(d+1)}, r\}}{\sum_{t \in J(j)} K\left(\frac{b_t^{(1)} - s}{h}\right)}$$

where $K(\cdot)$ is a kernel function, h is a bandwidth parameter and $J(j)$ is the set of all auctions for the product of interest. We can then solve for the equilibrium $\overline{v}_j(p, d, r)$ for each set of

[23] Budget 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.

BIN-randomization parameters (p, d, r), and get an average revenue estimate as follows:

$$\begin{aligned}\text{Revenue}^{\text{BIN-TAC}}(p,d,r) \;=\;& \frac{1}{T}\sum_j \Bigg(\sum_{t\in J(j)} 1(b_t^{(2)} \geq \overline{v}_j(p,d,r))b_t^{(2)} \\ &+ \sum_{t\in J(j)} 1(b_t^{(1)} \geq \overline{v}_j(p,d,r) > b_t^{(2)})p \\ &+ \sum_{t\in J(j)} 1(b_t^{(1)} < \overline{v}_j(p,d,r)) \sum_{j=1}^{d} 1(b_t^{(j)} \geq r) \max\{b^{(d+1)}, r\} \Bigg)\end{aligned} \tag{6}$$

This approach follows the empirical auctions literature in treating bidder's valuations as private information.[24] A different modeling approach was suggested in an influential paper by Edelman, Ostrovsky and Schwartz (2007). They proposed a complete information model of sponsored search auctions. Their logic was that since these players compete with high frequency and can potentially learn each others' valuations, a complete information model may be a better approximation to reality than an incomplete information model. Following this intuition, we also consider counterfactual simulations under complete information. This can create multiple equilibria, but employing a trembling hand perfection refinement eliminates this multiplicity.[25] In this case the threshold value becomes auction-specific: knowing the d-th highest valuation, the highest type declines the BIN option only when their expected payoff from taking a chance exceeds their surplus from buying-it-now.

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, when they believe 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 impression 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

[24] Athey and Nekipelov (2010) suggest a model of sponsored search auctions in this tradition.

[25] See 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 first two of these, we find these parameters by maximizing simulated revenue, using standard optimization methods.[27] In the case of the Myerson mechanism, since the estimated valuations are neither symmetric nor independent, the standard ironing approach is not revenue optimal. Nonetheless, we apply the standard approach, finding the ironing regions that would be optimal under symmetric independent private values. In some sense, this is the relevant benchmark, since asymmetric reserves or Cremer and McLean's (1988) side bets are not really practical. To get standard errors on our revenue and consumer surplus estimates, we bootstrap the estimation sample and re-run the simulation procedure, holding the parameter choices fixed.[28]

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 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

[26]The 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.

[27]This 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.

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

six bidders in an average auction. The Myerson mechanism is quite complex, as we earlier conjectured: as Figure 4 shows, there are 7 ironing regions (some are barely visible).

The welfare performance of these mechanisms is detailed in Table 5. The SPA without reserve earns revenue of 0.98 per auction, and leaves substantial consumer surplus — on average 1.97 per auction.[29] 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%.[30] 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 Myerson mechanism, though complex, does even better. It outperforms BIN-TAC on both revenues and consumer surplus, which suggests that it may be a good option provided it can be explained to advertisers. By contrast, 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

We have introduced the BIN-TAC mechanism, designed to allow sellers to capture the surplus created by providing match information. This mechanism outperforms the second-price auction mechanism in this setting, and is preferable to bundling goods together by withholding information, at least when there is a reasonable size population of potential bidders. Moreover, we demonstrated through an example that the mechanism can closely approximate Myerson's optimal mechanism with ironing, despite its relative simplicity.

[29]The 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.

[30]As the table shows, the results are quite robust to the information structure assumed.

Our analysis of the exchange marketplace revealed that it has many features that make it a good place to apply our mechanism: large differences between the highest and second highest bid, and evidence of matching on user characteristics that the platform has chosen to make available to advertisers. Although the market does not fit our stylized model, we found that the BIN-TAC mechanism would nonetheless improve revenues and consumer surplus relative to the existing mechanism, a second price auction with reserve; although not as much as implementing the Myerson mechanism.

Due to data limitations we were not able to compare our mechanism to an optimal bundling strategy. Instead, we looked at what would happen if the platform only provided advertisers with product and user location information, rather than more detailed demographics. This bundling strategy performed poorly, but it is an interesting and open research question as to whether switching mechanisms to BIN-TAC is in fact better than retaining the SPA with a more thoughtful bundling strategy.

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.

Adams, William J. and Janet Yellen, "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 1976, *90* (3), 475–498.

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," 2010. Cowles Foundation Discussion Paper 1758R.

——— **and Maher Said**, "Dynamic Auctions: A Survey," *Wiley Encyclopedia of Operations Research and Management Science*, 2010.

——— **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.

Chawla, Shuchi, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan, "Multi-parameter mechanism design and sequential posted pricing," in "Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC)" 2010, pp. 311–320.

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.

Cremer, 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.

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.

Maskin, Eric and John Riley, "Monopoly with Incomplete Information," *The RAND Journal of Economics*, 1984, *15* (2), pp. 171–196.

McAfee, Preston, "The Design of Advertising Exchanges," *Review of Industrial Organization*, 2011, *39* (3), 169—185.

____ **, John McMillan, and Michael D. Whinston**, "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *Quarterly Journal of Economics*, 1989, *104* (2), 371—384.

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.

Nocke, Volker and Martin Peitz, "A Theory of Clearance Sales," *The Economic Journal*, 2007, *117* (522), pp. 964–990.

Ostrovsky, Michael and Michael Schwarz, "Reserve Prices in Internet Advertising Auctions: A Field Experiment," 2009. Working Paper.

Pagan, Adrian and Aman Ullah, *Nonparametric Econometrics*, Cambridge, UK: Cambridge University Press, 1999.

Reynolds, Stanley S. and John Wooders, "Auctions with a buy price," *Economic Theory*, 2009, *38*, 9–39.

RFC6265, "Http state management mechanism," *http://tools.ietf.org/html/rfc6265*, April 2011.

Rochet, Jean-Charles and Philippe Chone, "Ironing, Sweeping, and Multidimensional Screening," *Econometrica*, July 1998, *66* (6), 788–826.

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 $\prod_{j\neq i}\left(\int 1(a_j(v)=0)dF(v)\right)$. 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:

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

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

$$\begin{aligned}\pi_B(\overline{v}) &= \mathbb{E}\left[1(Y^1 < \overline{v})(\overline{v} - p)\right] \\ &= \mathbb{E}\left[1(Y^1 < \overline{v})\frac{1}{d}(\overline{v} - Y^*)\right] = \pi_T(\overline{v}).\end{aligned} \tag{7}$$

Next, we show a $\overline{v}$ satisfying Eq. (2) exists and is unique. The right hand side of Eq. (2) is a function of $\overline{v}$ with first derivative $\frac{1}{d}(1 - \frac{\partial}{\partial \overline{v}}\mathbb{E}[Y^*|Y^1 < \overline{v}]) < 1$. Since at $\overline{v} = 0$ 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 $\overline{v}$. For any reserve price r, let $p(\overline{v}, r)$ denote the BIN-TAC price for threshold $\overline{v}$. By Equation (2) we have

$$p(\overline{v}, r) = \frac{d-1}{d}\overline{v} + \frac{1}{d}E\left[Y^*|Y^1 < \overline{v}\right] \tag{8}$$

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

$$\begin{aligned}
& nF(\overline{v})^{n-1}(1-F(\overline{v})) \times \frac{\partial p(\overline{v},r)}{\partial r} \\
= \;& nF(\overline{v})^{n-1}(1-F(\overline{v}))\frac{1}{d}\Pr\left(Y^d \leq r | Y^1 < \overline{v}\right) \\
= \;& nF(\overline{v})^{n-1}(1-F(\overline{v}))\frac{1}{d}\left(\frac{1}{F(\overline{v})^{n-1}}\left(\sum_{k=0}^{d-1}\binom{n-1}{k}(F(\overline{v})-F(r))^k F(r)^{n-k}\right)\right) \\
= \;& (1-F(\overline{v}))\left(\sum_{k=1}^{d}\frac{n}{d}\binom{n-1}{k-1}(F(\overline{v})-F(r))^{k-1}F(r)^{n-k}\right) \qquad (9)
\end{aligned}$$

- *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 $\overline{v}$ (and no bids above $\overline{v}$), 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

$$\begin{aligned}
& \left(\sum_{k=1}^{d}\frac{k}{d}\binom{n}{k}(F(\overline{v})-F(r))^k F(r)^{n-k}\right) \\
= \;& \left(\sum_{k=1}^{d}\frac{n}{d}\binom{n-1}{k-1}(F(\overline{v})-F(r))^k F(r)^{n-k}\right) \\
= \;& (F(\overline{v})-F(r))\left(\sum_{k=1}^{d}\frac{n}{d}\binom{n-1}{k-1}(F(\overline{v})-F(r))^{k-1}F(r)^{n-k}\right) \qquad (10)
\end{aligned}$$

- *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

$$\begin{aligned}
& r\left(\sum_{k=0}^{d-1}\frac{n}{d}f(r)\binom{n-1}{k}(F(\overline{v})-F(r))^k F(r)^{n-k-1}\right) \\
= \;& rf(r)\left(\sum_{k=1}^{d}\frac{n}{d}\binom{n-1}{k-1}(F(\overline{v})-F(r))^{k-1}F(r)^{n-k}\right) \qquad (11)
\end{aligned}$$

Summing up expressions (9), (10), and (11) we have

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

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 $[\underline{\omega}_L, \overline{\omega}_H]$), and so r^* is unique. Hence, for any $\overline{v}$, the optimal reserve is either r^*, or is one of the boundaries $\underline{\omega}_L$, $\overline{\omega}_L$, $\underline{\omega}_H$, and $\overline{\omega}_H$. Note that the derivative is positive at $\underline{\omega}_L$. Also, for reserve equal to $\overline{\omega}_H$, the item never sells. Moreover, observe that the reserve equal to $\overline{\omega}_L$ 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 $\overline{v}$. There are three effects on the revenue of the mechanisms by marginally increasing $\overline{v}$.

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

Therefore, we have:

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

$$
\begin{aligned}
= \quad & \frac{n}{d}(1 - F(\overline{v}))F(\overline{v})^{n-1} \times \\
& \left(- \left(\frac{(d-1)f(\overline{v})}{1 - F(\overline{v})} + \frac{(n-1)f(\overline{v})}{F(\overline{v})} \right) E\left[\overline{v} - Y^* | Y^1 < \overline{v}\right] \right. \\
& \left. + \left(d - 1 + \frac{\partial E[Y^* | Y^1 < \overline{v}]}{\partial \overline{v}} \right) \right)
\end{aligned}
$$

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

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

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

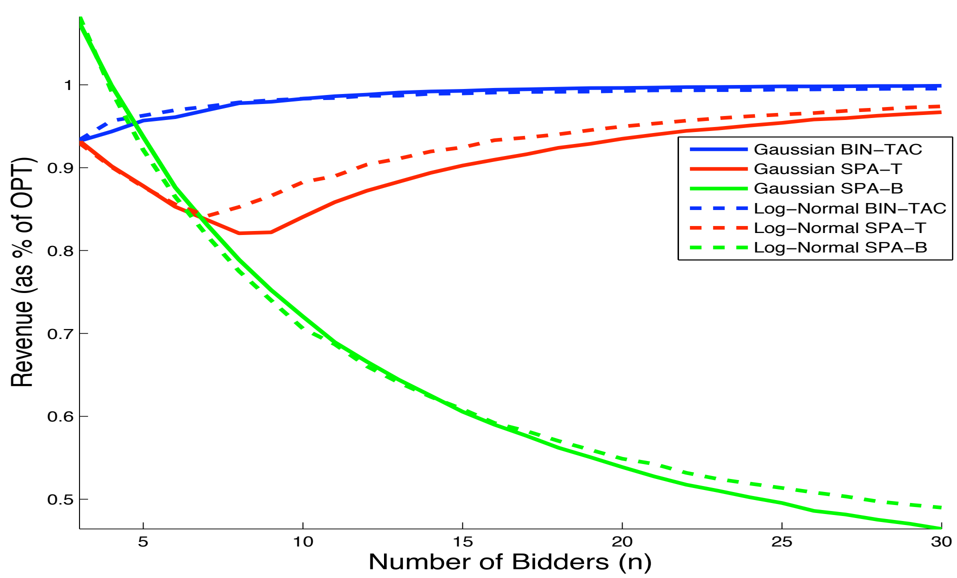

Figure 5: **Revenue Performance vs Number of bidders.** Simulated expected revenues for different mechanisms as the number of bidders n varies, in an environment where F_L has mean 1 and standard deviation 0.5, the match probability is 0.05 and the match increment is 5.

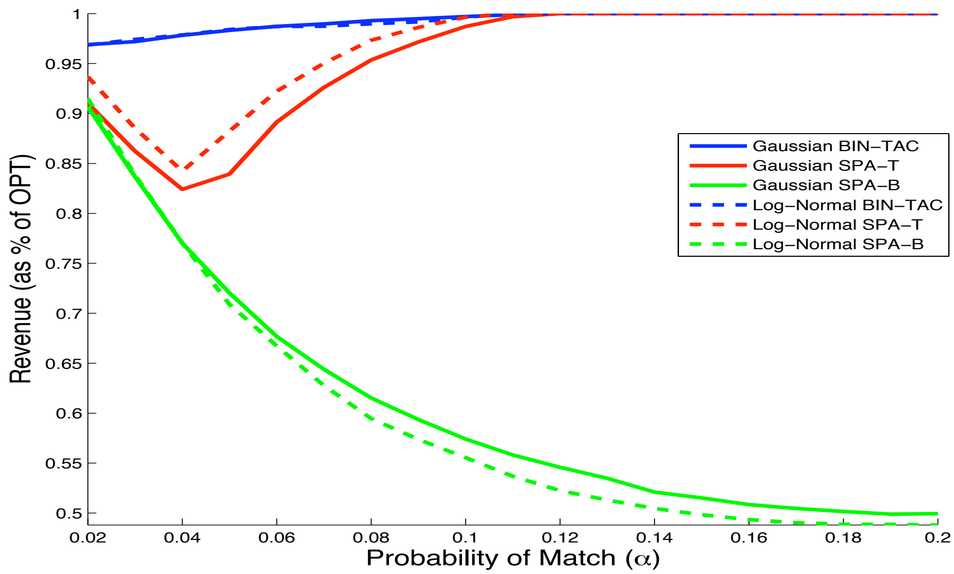

Figure 6: **Revenue Performance vs Match Probability.** Simulated expected revenues for different mechanisms as the probability of a match α varies, where F_L has mean 1 and standard deviation 0.5, the number of bidders is 10 and the match increment is 5.

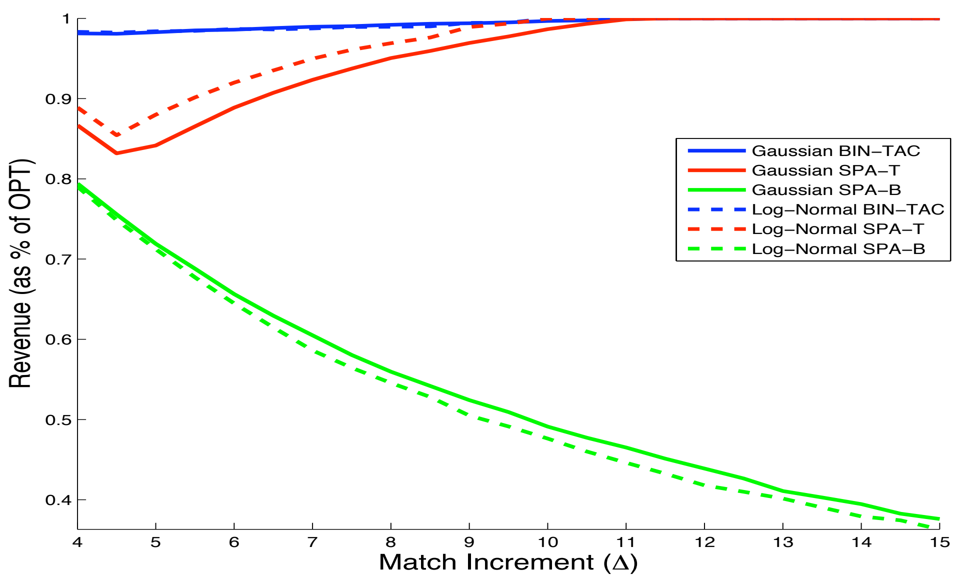

Figure 7: **Revenue Performance vs Match Increment.** Simulated expected revenues for different mechanisms as the match increment Δ varies, where F_L has mean 1 and standard deviation 0.5, the match probability is 0.05 and the number of bidders is 10.

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	0.329***	0.264***	0.286***
	(0.022)	(0.015)	(0.052)	(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
SPA	-	1.96	-
BIN-TAC (incomplete information)	3	0.26	1.83
BIN-TAC (complete information)	3	0.44	1.88
BIN-TAC (rationalizable worst case)	4	0.65	2.11

Revenue-maximizing parameter choices. For each of the above mechanisms, we find these by maximizing the revenue functions defined in the main text over the available parameters numerically using a grid search.

Table 5: Counterfactual Revenues and Welfare

Policy	Revenue	Consumer Surplus	Total Welfare
SPA (no reserve)	0.983	1.974	2.957
	(0.004)	(0.019)	(0.020)
SPA (optimal reserve)	1.028	1.471	2.499
	(0.005)	(0.019)	(0.020)
BIN-TAC (incomplete information)	1.073	1.685	2.757
	(0.005)	(0.019)	(0.020)
BIN-TAC (complete information)	1.074	1.641	2.715
	(0.004)	(0.019)	(0.020)
BIN-TAC (rationalizable worst case)	1.069	1.526	2.594
	(0.005)	(0.019)	(0.020)
Myerson (7 ironing regions)	1.147	1.759	2.906
	(0.005)	(0.015)	(0.017)
Bundling by Product-User Region	0.644	0.730	1.374
	(0.006)	(0.016)	(0.015)

Counterfactual simulations of average advertiser revenues, consumer surplus and total welfare (sum of producer and consumer surplus). All statistics reported outside parentheses are averages across impressions; those in parentheses are standard errors computed by bootstrapping the full dataset (i.e. they reflect uncertainty over the true DGP). Seven different simulations are run. The first is of a second price auction without reserve, while the second is of a second price auction with optimal (revenue-maximizing) reserve. The third is of the BIN-TAC mechanism, under the incomplete information structure outlined in the text. The fourth and fifth are robustness checks, varying the informational assumptions made for BIN-TAC. In the complete information case, bidders know the valuations of the other participants, and made BIN decisions accordingly. In the rationalizable 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 implements the Myerson (1981) mechanism with ironing, which is optimal under certain conditions. Here 7 ironing regions are used. Finally the seventh 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.