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

L. Elisa Celis
Xerox Research
Bangalore, India

ecelis@cs.washington.edu

Gregory Lewis
Department of Economics
Harvard University
Cambridge, MA

glewis@fas.harvard.edu

Markus Mobious
Microsoft Research New England
Cambridge, MA

mobius@microsoft.com

Hamid Nazerzadeh
Marshall School of Business
University of Southern California
Los Angeles, CA

nazerzad@usc.edu

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

1. Introduction

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

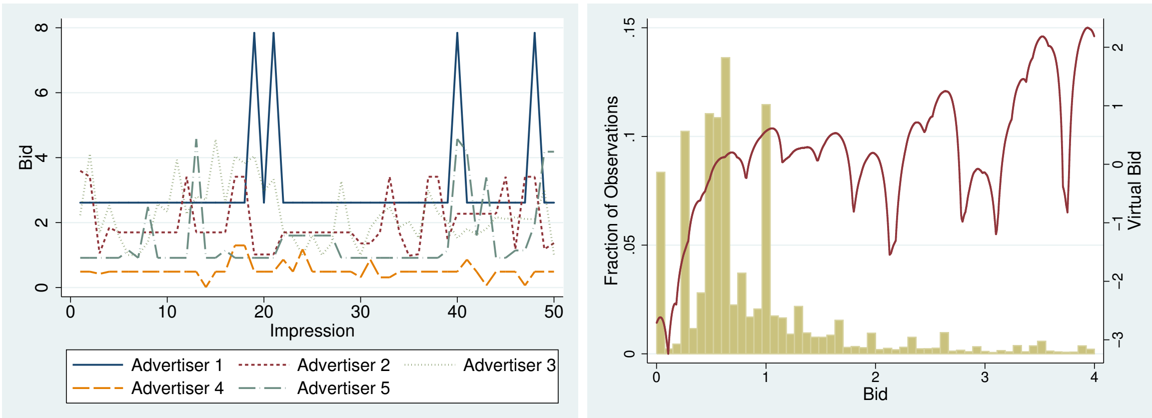

Figure 1 Bids over Time — Bid Distribution and Virtual Bids. The left panel shows the (rescaled) bids of five advertisers in our data, selected at random from the top 50 advertisers (ranked by purchases) on 50 randomly chosen successive bids for the most popular product. The right panel overlays a histogram of the bid data from the same product with a line plot of the "virtual bid", $\widehat{\phi}(b) = b - \frac{1-\widehat{F}(b)}{\widehat{f}(b)}$. The pdf f and cdf F are estimated by kernel smoothing with bandwidth 0.075.

pose special challenges for the predominant auction mechanisms that are used to sell online ads because they reduce competition among bidders, making it difficult for the platform to extract the surplus generated by targeting.

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

Suggestive empirical evidence for this is provided in Figure 1. The left panel shows the re-scaled bids in 50 auctions by five large advertisers for the most popular webpage slot sold by a large publisher through a second-price auction on the Microsoft Advertising Exchange; see Section 4 for details. The bids exhibit considerable variation, even though all these impressions were auctioned within a three hour period. We attribute this to matching on user demographics (evidence is provided later in the paper). Moreover, there is often a large gap between the highest and second highest bid in the second-price auction: on average, the gap is bigger than the sales price.

Consider a simple model that formalizes this narrative: when advertisers "match" with users, they have high valuation; otherwise, they have low valuation. Assume that match probabilities are independent across bidders and sufficiently low that the probability that *any* bidder will match is small. Then, a second-price auction will typically yield low revenue since the probability of two "matches" occurring in the same auction is very small. On the other hand, setting a high fixed price is not effective since the probability of zero "matches" occurring is large and many impressions would go unallocated. Allowing user targeting creates asymmetries in valuations that increase efficiency, but decrease revenue.

However, 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. [1] If more than one bidder takes the buy-it-now option, a second-price auction is held between those bidders with reserve p. Finally, if no one participates in buy-it-now, an auction is held in which the top d bidders are eligible to receive the good, and it is randomly awarded to one of them at the $(d+1)$-st highest price.

In this manner, we combine the advantages of an auction and a fixed price mechanism. When matches occur, advertisers self-select into the fixed-price buy-it-now option, allowing for revenue extraction. Advertisers are incentivized to take the "buy-it-now" option because in the event that they "take-a-chance" on winning via auction, there is a significant probability that they will not win the impression, even if their bid is the highest. [2] On the other hand, when no matches occur, the impression is still allocated, thereby earning revenue.

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

The right panel of Figure 1 shows a histogram of the bids for the most popular webpage slot, as well as the estimated virtual bids, $\widehat{\phi}(b) = b - \frac{1-\widehat{F}(b)}{\widehat{f}(b)}$. Notice that the virtual bids are certainly not increasing, even in regions where there are large numbers of observations. Interpreting bids as valuations — reasonable since it is weakly dominant for advertisers to bid their valuations — we can conclude that the distribution is irregular. We suspect that irregular values are actually quite common in practice. Any kind of "matching" story of the kind we outlined earlier could lead to bimodal distributions and possible irregularity. Moreover, it is hard to formulate a sensible

[1] This is similar to the buy-now option used commonly by eBay: `http://pages.ebay.com/help/buy/how-buy-bin.html`.

[2] In the eBay auctions, the "buy-now" incentive for a bidder who arrives to the auction is that, in the future, buyers with higher valuations may show up that would purchase the item (cf. Board and Skrzypacz (2013)). There is no temporal component in our model; see **??** for a discussion on the observability of the arrivals.

demand system in which valuations are not correlated, violating the IPV assumption. For example, if advertisers target users from a specific geographic region, this will induce a positive correlation between advertisers with the same target region and a negative correlation across advertisers with different target regions.

Randomization will generally be necessary for optimal price discrimination with irregular values. Moreover, this is exactly the case when price discrimination is most valuable. Bulow and Klemperer (1996) proved that when valuations are regular, revenues from a simple second-price auction with $n+1$ bidders exceeds the revenue from the revenue-optimal mechanism with n bidders. This bounds the gain to design. But in the irregular environment, this need not be true: good design can make up for fewer bidders.

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

Second, we will argue that it is *approximately optimal* for the symmetric independent private values environment with irregular valuations and *robust* to different informational environments. We show that in the simple two-type matching environment outlined earlier, BIN-TAC is revenue-optimal, outperforming the leading alternatives: a second-price auction with reserve, or the "bundling" solution in which the platform withholds targeting information, as suggested in a number of recent papers; See Ghosh, Nazerzadeh and Sundararajan (2007), Even-Dar, Kearns and Wortman (2007), McAfee, Papineni and Vassilvitskii (2010) and Bergemann, Bonatti and Said (2011). In a more general SIPV environment with irregular valuations, BIN-TAC will not generally be revenue-optimal. The optimal mechanism, as suggested by Myerson (1981), is a direct mechanism that involves pooling groups of types through "ironing".[3] Numerical simulations show that in mixture distribution environments, BIN-TAC approximates the allocations and payments of this mechanism, achieving "almost optimal" performance. In fact, when valuations are negatively correlated across bidders, inducing bid gaps, BIN-TAC outperforms the ironing mechanism, despite its simplicity.

It is hard to motivate a mechanism based on simulations alone. Therefore, to analyze its performance in a real-world setting, we turn to historical data from the Microsoft Advertising Exchange.

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

Assuming that bids are equal to valuations, we can simulate the effect of introducing the BIN-TAC mechanism. We find that the optimal BIN-TAC mechanism generates 4.4% more revenue than the optimal second-price auction, while at the same time improving advertiser surplus by 14.5%. This is possible because the optimal second-price auction uses a high reserve to extract surplus from the long tail of valuations, whereas the BIN-TAC mechanism does this through a high buy-it-now price, which avoids excluding low valuation bidders.

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

We view the main contribution of our paper as introducing and analyzing a new, simple, and robust price discrimination mechanism that makes use of randomized auctions and then evaluating its performance in a realistic environment. Although we have motivated the mechanism through a particular story about targeting in advertising exchanges, our insights apply more generally to optimal design in the single-unit auction environment with irregular valuations. A secondary contribution of the paper is to document participation and bidding behavior in the display advertising market specifically. While there has been theoretical work on this market (Muthukrishnan 2010, McAfee 2011, McAfee and Vassilvitskii 2012, Balseiro, Besbes and Weintraub 2012) 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 the game theoretic aspects of display advertising and exchange markets; see Balseiro, Feldman, Mirrokni and Muthukrishnan (2011), Feldman, Korula, Mirrokni, Muthukrishnan and Pál (2009), Radovanovic and Zeevi (2010), and Bhalgat, Feldman and Mirrokni (2012) on the optimization aspects of these markets. Finally, we believe, we are the first to calculate and simulate revenues from the Myerson (1981) mechanism with ironing using data from a real market, which may be of interest to the many people working on problems in optimal mechanism design.

Related Work: Our work is fundamentally related to the work on optimal auction design (Myerson 1981, Riley and Samuelson 1981, Maskin and Riley 1984, Crémer and McLean 1988). There has also been some recent work on simple designs showing that they may be approximately optimal (Hartline and Roughgarden 2009) and on screening problems where the principal approximates the type space (Madarász and Prat 2010). In the online advertising context, the seller has an additional choice variable: the decision of whether or not to reveal information that allows advertisers to target their bids. This is related to work on optimal information disclosure (Lewis and Sappington 1994, Bergemann and Pesendorfer 2007, Eso and Szentes 2007, Kakade, Lobel and Nazerzadeh 2013). The BIN-TAC mechanism itself has the feel of a sequential screening mechanism, as in Courty and Li (2000), although our environment is static. The buy-now option has been studied in other contexts, although typically its usage elsewhere aims to allow buyers to bypass a lengthy English auction and/or avoid risk (Budish and Takeyama 2001, Reynolds and Wooders 2009). For more on the general literature on price discrimination, see McAfee, McMillan and Whinston (1989)

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. Mahdian, Ghosh, McAfee and Vassilvitskii (2012) show that the revenue may also suffer from sharing cookies with advertisers due to information leakage.

In style, our paper is close to Chu, Leslie and Sorensen (2011), who combine theory, simulations, and empirics to argue that bundle-size pricing is a good approximation of the optimal mixed bundling pricing scheme for a monopolist selling multiple goods. Finally, from an empirical perspective, our paper contributes to the growing literature on online advertising and optimal pricing. Much of the work here is experimental in nature — for example, Lewis and Reiley (2011) ran a randomized experiment to test advertising effectiveness, while Ostrovsky and Schwarz (2009) used an experimental design to test the impact of reserve prices on revenues. There has also been recent empirical work on privacy and targeting in online advertising (Goldfarb and Tucker 2011a, Goldfarb and Tucker 2011b, Johnson 2013).

Organization: The paper is organized 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 the existence and characterization results for the BIN-TAC mechanism. We also provide analytic results concerning the revenue maximizing parameter choices and compare our mechanism to others using both theory and simulations (additional simulations are offered in the online appendix available at `http://www-bcf.usc.edu/~nazerzad/pdf/bintac-online-appendix.pdf`). Finally, in the third part, we provide an empirical analysis using data from the Microsoft Advertising Exchange, including counterfactual simulations of our mechanism's performance. All proofs are contained in the appendix.

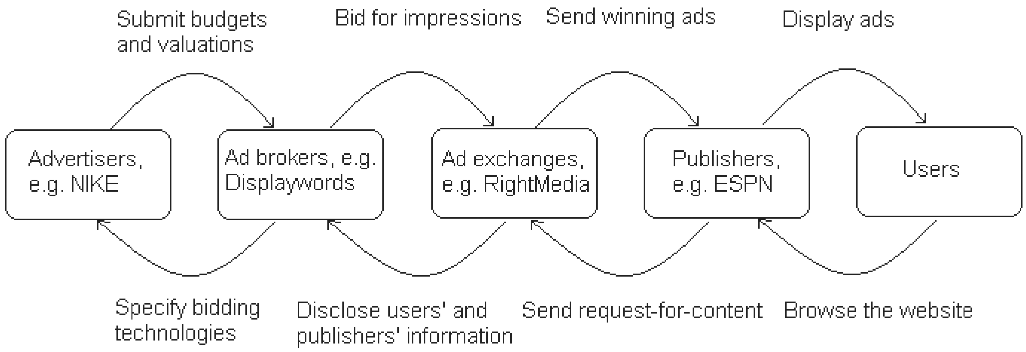

Figure 2 The Display Advertising Market.

2. The Display Advertising Market

The organization of the display advertising market is depicted in Figure 2. On one side of the market are the "publishers": these are websites that 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 intents are. For example, a Ferrari dealer might value high income users located close to the dealership. A mortgage company might value people reading an article on "how to refinance your mortgage" more than those reading an article on "how to survive your midlife crisis", while the dealership might prefer the reverse.

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

Alternatively, content is sold by auction through a centralized platform called an advertising exchange. Examples of leading advertising exchanges include the Microsoft Advertising Exchange (a subset of which we examine in this paper), Google's DoubleClick, and Yahoo's RightMedia.

These are large marketplaces: for example, in September 2009, RightMedia averaged 9 billion transactions a day with hundreds of thousands of buyers and sellers (Muthukrishnan 2010). Advertising exchanges are a minor technological wonder, as they run all these auctions in real time. When a user loads a participating publisher's webpage, a "request-for-content" is sent to the advertising exchange. This request will specify the type and size of advertisement to be displayed on the page, as well as information about the webpage itself (potentially including information about its content), and information about the user browsing the page.For example, it may include their IP address and cookies that indicate their past browsing behavior.

The advertising exchange will then either allocate the impression to an advertiser at a previously negotiated price, or hold a second-price auction between participating advertisers. If an auction is held, all or some of the information about the webpage and the 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.[4]

The bids are jointly determined by the preferences of the advertisers, 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 cannot 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 does not allow the advertiser to specify a different willingness to pay based on some particular user characteristic, then this will not show up in their bids.

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

[4] 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, analyzed in the current paper; the second is pay-per-click, where the payment depends on whether or not the user clicks on the advertisement. Ad exchanges use expected click through rates to compare these different bids.

3. Model and Analysis

3.1. The Environment

A seller (publisher) has a single impression to sell in real time, and they have information about the user viewing the webpage, summarized in a cookie.[5] 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. However, 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 that bids may be 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 thus 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 that is valued at zero by the seller. Bidders are risk-neutral.[6] 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 that 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 the probability $\alpha \in (0,1)$.

The bidder type is a triple (X, V_L, V_H), drawn identically and independently across bidders, and independently of each other. This implies that a user who is a match for one advertiser need not be a match for the others. In the case of 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 that their expected valuation is $E[V] = (1-\alpha)V_L + \alpha V_H$.

Most of our analysis concerns the case where the seller discloses the targeting information. In this case, our model is equivalent to a standard SIPV model in which each bidder's single-dimensional valuation is drawn from the mixture distribution F, with $F(v) = (1-\alpha)F_L(v)$ when $v \in [\underline{\omega}_L, \overline{\omega}_L]$, and $F(v) = (1-\alpha) + \alpha F_H(v)$ when $v \in [\underline{\omega}_H, \overline{\omega}_H]$. To avoid analyzing multiple cases, we make

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

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

some technical assumptions on the virtual valuations, $\psi(v) = v - \frac{1-F(v)}{f(v)}$. We assume that $\psi(v)$ is continuous and increasing over the regions $[\underline{\omega}_L, \overline{\omega}_L]$ and $[\underline{\omega}_H, \overline{\omega}_H]$, crossing zero once in the low valuation region, and that $\psi(\overline{\omega}_L) \leq \psi(\underline{\omega}_H)$. The virtual valuations are (infinitely) negative over the region $(\overline{\omega}_L, \underline{\omega}_H)$ since F is unsupported on this region. Therefore, the virtual valuations are not increasing, and the distribution F is *irregular*.

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

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

3.2. Pricing Mechanisms

BIN-TAC works as follows. A *buy-it-now price* $p \in (\underline{\omega}_L, \overline{\omega}_H)$ 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, but the number of bidders in the BIN auction is not revealed. Finally, if no-one elects to BIN, a sealed bid *take-a-chance* (*TAC*) auction is held between all bidders, with a reserve r ($\underline{\omega}_L \leq r \leq p$). In that auction, one of the top d bidders is chosen uniformly at random, and if their bid exceeds the reserve, they win the auction and pay the maximum of the reserve and the $(d+1)$-th bid (if it exists). Ties among d-th highest bidders are broken randomly prior to the random allocation. We call r the TAC-reserve, and d the randomization parameter.

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

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

3.3. The Two-Type Case

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

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

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

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 that no other types buy-it-now; and the RHS is the expected payoff from taking-a-chance, again assuming that no other types buy-it-now. Conditioning on no-one else taking the BIN option does not affect the IC constraint, since if someone does, expected surplus is zero regardless of whether the buyer elects to BIN or TAC.

Clearly, the seller wants to set the BIN price so the IC constraint holds with equality, which implies a BIN price equal to $\frac{d-1}{d}v_H + \frac{1}{d}v_L$. Notice that 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, the 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 $d^* = \infty$. However, 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.[7]

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 that BIN-TAC has higher revenues than bundling (and clearly this remains true in the case where exclusion is optimal). In addition, bundling is less efficient, since the allocation is random.

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

3.4. General Case: Equilibrium Analysis

Returning to the general environment, we proceed by backward induction to characterize equilibrium strategies under BIN-TAC. If multiple agents choose to BIN, the allocation mechanism reduces to a second-price auction with reserve p. Thus, it is weakly dominant for agents to bid their valuations. Truth-telling is also weakly dominant in the TAC auction.[8]

Taking these auction strategies as given, we turn to the decision whether to buy-it-now or take-a-chance. Intuitively, the BIN option should be more attractive to higher types: they have the most to lose from either random allocation (they may not get the good even if they are willing to pay the most) or from rivals taking the BIN option (they certainly will not get the good). This suggests

[7] 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.

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

that in a symmetric equilibrium, the BIN decision takes a threshold form: $\exists\, \overline{v}$ and 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.

Theorem 1 (Equilibrium Characterization) *Assume $d > 1$. Then, there exists a unique symmetric pure strategy Bayes-Nash equilibrium, characterized by a threshold $\overline{v} \in (\underline{\omega}_L, \overline{\omega}_H]$. If $\overline{\omega}_H - p < \frac{1}{d} E\left[\overline{\omega}_H - Y^* | Y^1 < \overline{\omega}_H\right]$, then $\overline{v} = \overline{\omega}_H$ and all types take-a-chance. Otherwise, $\overline{v}$ is the solution of*

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

and types with $v \geq \overline{v}$ take the BIN option and types $v < \overline{v}$ TAC. All types bid their valuation in any auction that may occur.

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

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

Theorem 2 (Optimal Buy-It-Now and Reserve Price) *For any randomization parameter d, the revenue-maximizing TAC reserve r^* is either equal to $\underline{\omega}_H$ or 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 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 raising the TAC reserve lowers the surplus from participating in the TAC auction, and the seller can also raise the BIN price while keeping the indifferent type $\overline{v}$ constant. Thus, the trade-off is the usual one: raising the TAC reserve increases the expected payments from types above r^* — even those who take the BIN option — at the cost of losing revenue from the marginal type.

On the other hand, the implicit equation for the optimal BIN price is new. Notice that the BIN price in some sense sets a reserve at $\overline{v}$. 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. Therefore, a marginal increase in the threshold has three effects. First, if the highest bidder has a 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 a 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 the expected costs and benefits, we obtain the result.

We cannot rule out a corner solution for the buy-price. 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 SPA-T is just a special case of BIN-TAC; therefore, any outcome achievable by the SPA-T is also achievable by BIN-TAC. Note that when $d = 1$, the TAC auction is just a second-price auction. So any SPA-T with optimal reserve r can be mimicked by a BIN-TAC mechanism with the same reserve, $d = 1$ and $p = \infty$. Moreover, as we saw in the two-type case, when the objective is revenue maximization, BIN-TAC can sometimes do strictly better — in fact, we will see in the empirical application a case where the revenue-optimal BIN-TAC beats the revenue-optimal SPA-T on *both* revenue and consumer surplus.

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

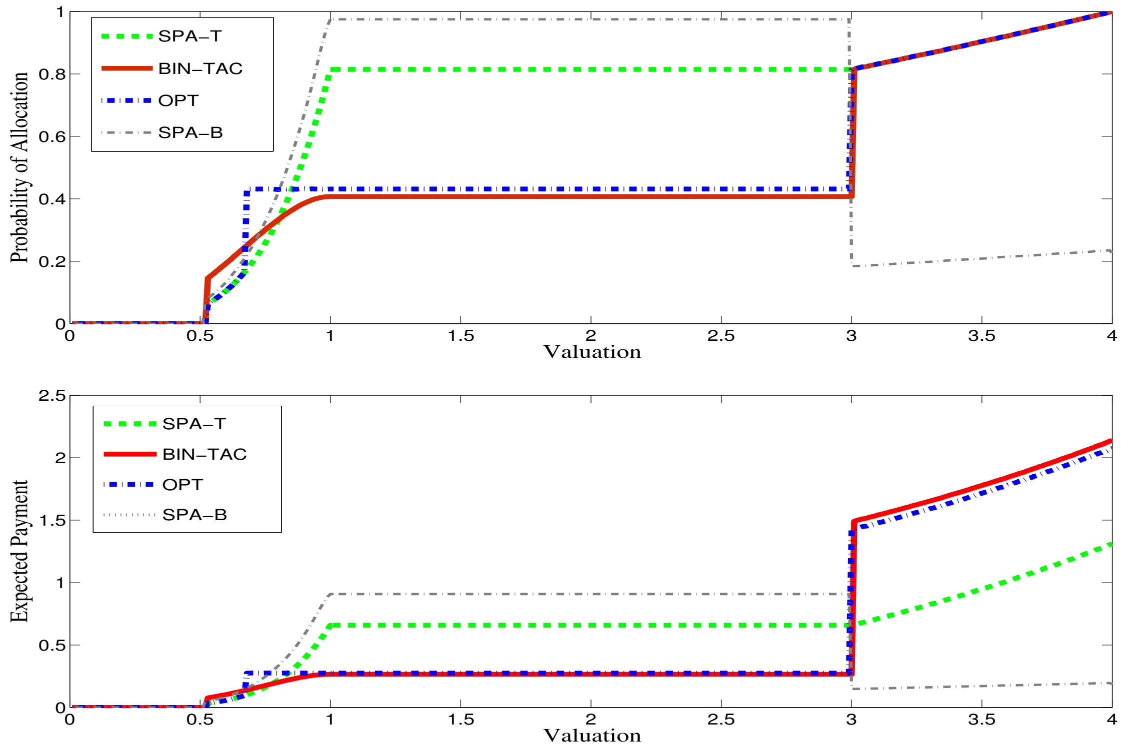

Figure 3 Allocation probabilities (top panel) and expected payments (bottom panel) when the distributions F_L and F_H are uniform. The x-axis corresponds to the bidder's valuation (unknown to them under bundling).

may be less profitable than limiting information rents through bundling. But in the simulations presented below and in the online appendix, BIN-TAC generally outperforms bundling.

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

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

For SPA-B, the figures depict the allocation probabilities and expected payments as a function of the *realized* valuations of the agent (which are unknown to the agent under bundling). For instance, for a given realized high valuation v_H (i.e. on $[3,4]$), the allocation probability plotted on the y-axis is equal to the average allocation probability across all types $(1-\alpha)V_L + v_H$, who will bid their expected valuations in the second-price auction under bundling. This implies that the allocation probability is no longer monotone: since matches are ex-ante unlikely ($\alpha = 0.05$), the most aggressive bidders in the bundling auction are those who have high valuations even without matching (high v_L), and therefore they are most likely to get the object. This makes the trade-off clear: the bundling mechanism raises expected payments when there is no match (because bidders do not know that they have not matched), and substantially lowers them in the case of a match.

The table below, compares the outcomes of the mechanisms in the uniform environment with $\Delta = 3$, $\alpha = 0.05$ and 5 bidders. The expected revenue performance of BIN-TAC is close to the optimal mechanism (about 96% of OPT), much better than the optimal SPA-T (85%). SPT-B performs less well than both BIN-TAC and OPT, especially in terms of consumer surplus. This is because it fails to match advertisers and users correctly.

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

3.6. Extensions

In the empirical application in the next section, we maintain our assumption of private values, but allow for correlated valuations. To facilitate a later discussion of how these assumptions drive our simulation methodology and results, we analyze the correlation and common values using the two-type model introduced in Section 3.3.

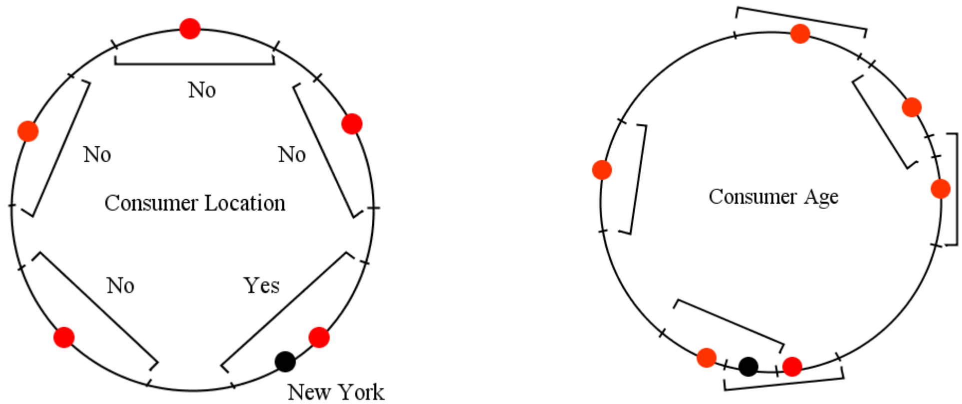

Figure 4 **The left figure depicts an example of perfectly differentiated targeting where $\alpha n < 1$. The red dots represent the advertisers' locations and the black dot corresponds to the location of the user; the intervals around the advertisers represent their target match range. The right figure is an example of independent target demographics.**

Correlation: Let the advertisers jointly draw random locations on a unit circle according to some copula (i.e., a multivariate distribution with uniform marginals) and let the user independently draw a random location uniformly. We will say that a user and advertiser match if the user is within a distance of $\alpha/2$ of the advertiser so that the probability of any one advertiser matching with a user is α. As before, when a match occurs, the value to the advertiser is v_H; otherwise, it is v_L. What we have in mind is that advertisers have tastes concerning a particular user characteristic that may be revealed by the platform, and they may agree in their tastes (positive correlation of locations) or disagree (negative correlation of locations). Thus, for example, if the platform reveals user income to the advertisers and they all agree that high-income consumers are more likely to purchase their products, this would be represented by a clustering of the advertiser locations. If the platform reveals the geographic location of the user, and the advertisers are spatially differentiated and want to target local consumers, this could be represented by a spacing of the advertiser locations; see Figure 4.

We focus on two cases: perfectly positively correlated preferences (all advertisers have an identical random location) and perfect differentiated preferences (the advertisers are evenly spaced around the unit circle, a form of negative correlation, see Figure 4). We assume $\alpha < \frac{1}{n}$ so that even with perfect differentiation there may be no matches. In the perfectly correlated case, the revenue from all of the mechanisms considered above — bundling, second-price with targeting and BIN-TAC — is the same. Either everyone matches or no one does, and so revealing information always results in revenue of $\alpha v_H + (1-\alpha)v_L$, just as with bundling. BIN-TAC has no advantage here because it is designed to extract revenue when only a single match occurs, which is never the case.

On the other hand, in the case of perfect preference differentiation BIN-TAC continues to outperform the other mechanisms. The reason is that, because of the spacing, there can only be one or zero matches. Hence, the SPA-T returns revenue of v_L, and (with simple calculations) BIN-TAC returns revenue of $v_L + \alpha(n-1)(v_H - v_L)$. [9] By inspection, this is higher than the bundling revenue of $\alpha v_H + (1-\alpha)v_L$. In addition, the gain from switching from another mechanism to BIN-TAC is bigger than in the case of independent preferences, suggesting that BIN-TAC is more relatively attractive under negative correlation in advertiser valuations, and less so under positive correlation.

Common Values: We now return to the case of independent type draws, but introduce the possibility of common values. Let us reinterpret each agent's realized valuation $V = (1-X)V_L + XV_H$ as a signal. The true valuation U_i of a user to advertiser i depends on the signals of the other agents, as $U_i = \beta V_i + (1-\beta)\overline{V}_{-i}$, where $\overline{V}_{-i}$ is the average of the other players' signals, for $\beta \in [\frac{1}{n}, 1]$. If $\beta = 1$, we restore private values; if $\beta = \frac{1}{n}$, we have the case of pure common values. Importantly for the analysis, the model remains symmetric (see Abraham et al. (2010) for the complications that result with common values and asymmetry).

Revenue from bundling remains $\alpha v_H + (1-\alpha)v_L$ since, in the absence of any signal, the expected value of the user has not changed. The unique symmetric pure-strategy equilibrium of the SPA-T is that bidders bid v_L if they draw v_L, and bid $u^\star < v_H$ if they draw v_H, where $u^\star = E[U_i | V_i = v_H, \max_{j \in -i} V_j = v_H]$. [10] This is the equilibrium identified by Milgrom and Weber (1982), and the shading reflects the winner's curse. Revenue from SPA-T is thus $(q_0 + q_1)v_L + (1 - q_0 - q_1)u^*$. Recall that q_i is the probability of exactly i matches.

BIN-TAC is a bit more complicated. Suppose that the buy-it-now price is set so that an advertiser with a high signal takes the BIN. Then, if multiple advertisers take the BIN, at the auction stage it becomes common knowledge among whoever takes the BIN option that at least one rival bidder has a high signal (since the number of BIN participants is not revealed, further inference is impossible). All the bidders at the BIN stage will bid their expected valuation, u^*. Similarly, if no one takes the BIN, it becomes common knowledge that everyone has low signals, and everyone bids v_L.

We now find the BIN price. A bidder with the high signal who wins outright by BIN gets a user with the expected valuation $\widetilde{u} = E[U_i | V_i = v_H, \max_{j \in -i} V_j = v_L]$, since they learn that no one else had a high signal. So a high type will take BIN iff $\widetilde{u} - p \geq \frac{1}{n}(\widetilde{u} - v_L)$, implying that the optimal BIN price is $p = \frac{n-1}{n}\widetilde{u} + \frac{1}{n}v_L$. Overall revenue is thus $q_0 v_L + q_1(\frac{n-1}{n}\widetilde{u} + \frac{1}{n}v_L) + (1 - q_0 - q_1)u^*$, which remains better than SPA-T since $\widetilde{u} > v_L$. On the other hand, BIN-TAC is no longer guaranteed to weakly better than bundling. For example, when $\beta = \frac{1}{n}$ (pure common values), the winner's curse

[9] The buy-it-now price is equal to $v_H - \frac{1}{n}(v_H - v_L)$ which yields revenue equal to

$$\alpha n(v_H - \frac{1}{n}(v_H - v_L)) + (1-\alpha n)v_L = \alpha(n-1)v_H + \alpha v_L + (1-\alpha n)v_L = v_L + \alpha(n-1)(v_H - v_L).$$

[10] All symmetric equilibria must consist of a pair of bids (b_L, b_H) (one for each of the types (v_L, v_H)) such that local deviations are unprofitable. One can quickly verify that the only bid pair of this form is the one given in the text.

effects make $\widetilde{u} = v_L + \frac{1}{n} v_H$, so revenue from BIN-TAC is almost equal to the SPA-T, and SPA-T and SPA-B cannot in general be ranked (it depends on the magnitudes of α and n). BIN-TAC remains revenue-improving relative to the SPA-T under common values, but bundling may attractive if the efficiency gains from targeting are small relative to the common value effects.

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 the Microsoft Advertising Exchange, one of the world's leading ad exchanges. Our data comes from a single large publisher's auctions on this exchange and consists of a 0.1% random sample of a week's worth of auction data from this publisher, sampled within the last two years. This publisher sells multiple "products", where a product is a URL-ad size combination (e.g. a large banner ad on the sports landing page of the New York Times).

The data includes information from both the publisher and the advertiser. On the publisher side, we see the URL of the webpage the ad will be posted on, the size of the advertising space and the IP address of the user browsing the website. We form a unique identifier for the url-size pair, and call it a product. We determine which U.S. state the user IP originates from, and call it a region. We use controls for product and region throughout the descriptive regressions. Unfortunately, we do not have more detailed information on the product or the user, as the tags and cookies passed by the publisher to the ad exchange were not stored. The information contained in these tags and cookies will vary across users, but may include information on the user's age, gender, and location as well as "behavioral characteristics" such as the fact that the user recently visited an online shopping site.

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. We see who won each auction and the final price. We drop auctions in which the eventual allocation was determined by biased bids and modifiers.[11] We also restrict attention to impressions that originate in the US, and where the publisher content is in English. Finally, we only consider frequently sold products, those with at least 100 sales in the dataset. This leaves us with a sample of 83515 impressions.

The dataset is summarized in Table 1. For confidentiality reasons, bids have been rescaled so that the average bid across all observations is equal to 1 unit. As shown in the top left panel of

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

	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				

Table 1 Summary Statistics. The data is a 0.1 percent sample of a week's worth of auction data from a single publisher 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. For confidentiality reasons, the bids have been normalized so that their average across all observations is equal to 1. The bid correlation is measured by selecting a pair of bids at random in every auction with at least two bidders.

Figure 5, the bids are very skew. The median bid is only 0.57 units, while the average winning bid is much higher at 2.96 units. This may be a consequence of the bid skewness since, as the highest order statistic, the winning bid is more heavily sampled from the right tail of the bid distribution.

The advertisers are quite active in this market. On average, they bid on 0.7% of all impressions, so that the average auction has 6 bidders. But as the top right panel of Figure 5 shows, participation is also quite skew. The median advertiser bids on only 0.02% of impressions, while the most active advertiser participates in nearly 90% of auctions. Advertisers who participate in few auctions tend to bid quite highly and therefore win with high probability. Others bid lower amounts in many auctions, and win with lower probability. We speculate that the first strategy is followed by companies that 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.

Our earlier theoretical analysis was of the canonical case where values are independently and identically sampled from some distribution F. At first glance, this looks like a reasonable approximation for our data: bids should be equal to values in a second-price auction and the pairwise correlation in bids is only 0.01. Yet, for the purposes of optimal mechanism design, it is the joint distribution of the highest two values that is of most importance, since these determine the price (in a second-price auction without a reserve) and whether the object sells (with a reserve).

The data is not consistent with these two order statistics being generated by i.i.d. sampling from some marginal distribution. To see this, we constructed a simulated dataset with the same number of observations by sampling bids independently from the data, holding the distribution of

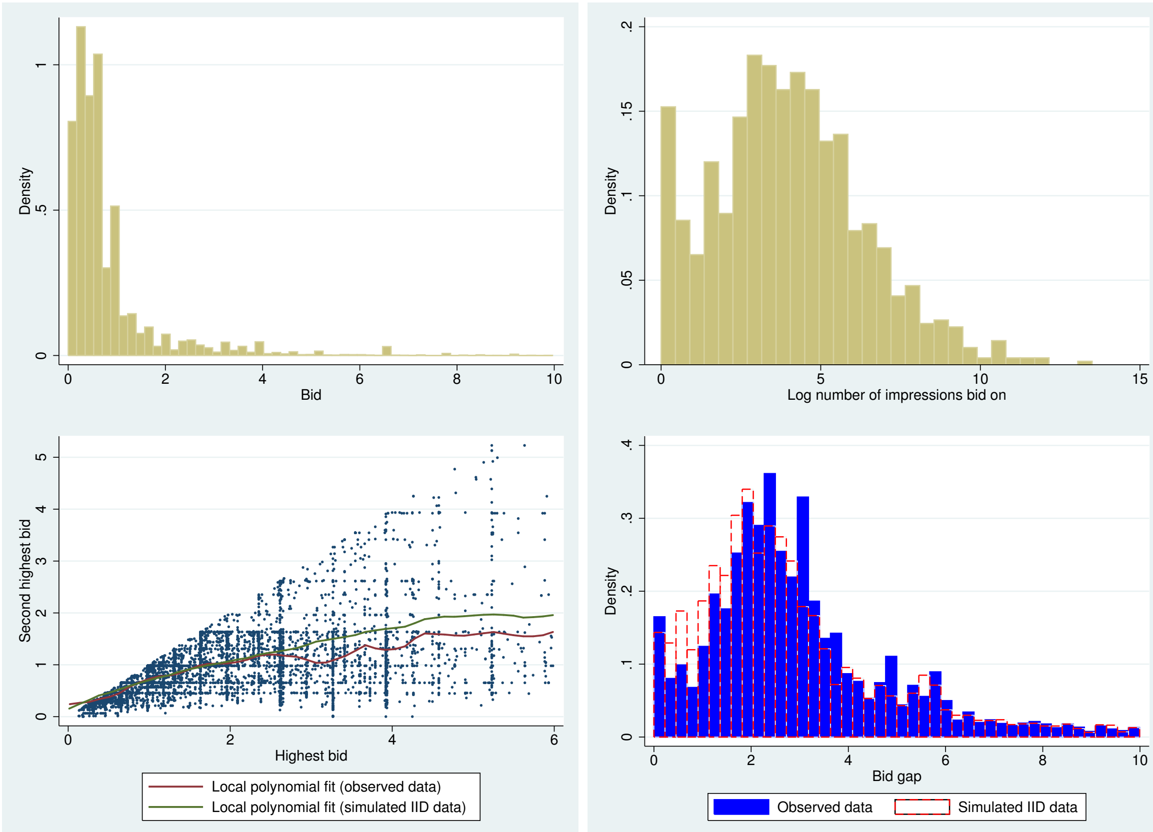

Figure 5 **Graphical data summary: The top left panel shows the distribution of bids in the data. The top right panel shows the distribution of the log number of impressions bid on. The bottom left panel shows a scatter plot of the highest and second highest bids in each auction with a local polynomial fit (in red); as well as the expected value of the second highest bid given the highest bid, if the bid were i.i.d. (in green). The bottom right panel compares the distribution of the bid gap in the observed and the simulated i.i.d. data, in both cases when the highest bid is at least 2.5.**

the number of bidders fixed (i.e. by independently sampling two bids with replacement for each two-bidder auction, three bids for each three-bidder auction, etc.). This simulated dataset satisfies the independence of bids by construction and has the same average bid (equal to one).

In the bottom left panel of Figure 5, we show a scatter plot of the highest bid against the second highest bid in the observed data, along with a local polynomial fit based on this data (i.e. $\widehat{E}[B^2|B^1]$) and another fit based on the simulated data ($\widehat{E}^{SIPV}[B^2|B^1]$) (using Stata's lpoly command with default options). The second highest bid is lower on average than one would have expected under independence over the range where the highest bid exceeds 2.5.

The bottom right panel shows this in another way. We plot the distribution of the bid gap in both the observed and simulated data for auctions in which the highest bid exceeds 2.5. It is pretty clear that the bid gap is stochastically higher in the data than in the simulated data. This is reminiscent of the two-type model with differentiated preferences offered as an extension earlier in the paper: when one bidder has a high value, it is relatively less likely (relative to the independent case) that the other bidders have matched and also have high values. Our analysis suggest that BIN-TAC will perform even better under differentiated preferences, and so one might expect BIN-TAC to perform particularly well in this environment; see Section 3.6.

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 does poorly in extracting this match surplus as revenue. We introduced our first piece of evidence in the introduction in Figure 1. As we noted, there is significant variation in the bids of large advertisers over a short time horizon. While this could be driven by decreases in the advertisers' available budget, since the bids go both up and down it seems more likely that this variation arises from matching on user demographics.

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

Proving that this matching is motivated by economic considerations is a little more difficult. The only user demographic we observe is the user region, and it is hard to know a priori what the advertisers' preferences over regions are. To get a handle on this, we turn to another proprietary dataset that indicates how often an advertiser's webpage was viewed by Internet users in different regions of the country during the calendar month prior to the auction.[12] 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 that attract a large fraction of their pageviews from a particular region should participate more frequently and bid higher on users from those regions.[13] 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".

To test for a significant correlation between the pageview ratio and participation, we run the following regression by OLS:

$$y_{i,t} = \beta \text{ pageview}_{i,r(t)} + d_{q(t)} + \gamma_{j(t),r(t)} + \mu_i + \varepsilon_{i,t} \tag{5}$$

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

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

	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

Table 2 Results from OLS Regressions. In the first two columns, the dependent variable is a dummy for participation. In the last two columns, the dependent variable is a bid. The estimation sample used in these regressions only includes bids from the 10% of bidders who bid most often. The pageview ratio 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 in which the user is located. In parentheses, we show the robust (Huber-White) standard errors. Significance levels are denoted by asterisks (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

where i indexes advertisers, j indexes products (i.e. webpage slots), q indexes quarters of the day (i.e. midnight to 6am, 6am to noon...), r indexes regions and t indexes impressions. $y_{i,t}$ is a binary variable indicating whether bidder i bid on impression t; pageview$_{i,r(t)}$ is the pageview ratio for bidder i and region $r(t)$, $d_{q(t)}$ is a quarter-of-day fixed effect, $\gamma_{j(t),r(t)}$ is a product-region fixed effect and μ_i is an advertiser fixed effect. An observation is an advertiser-impression pair. The sheer size of our dataset makes it difficult to run this regression on the full dataset. Therefore, we run this on a subsample consisting of the top 10% of advertisers. Fortunately, since participation is highly skewed, these advertisers account for 90% of the bids.

The first column of Table 2 shows the results. We find a positive but insignificant correlation between the pageview ratio and participation in the first specification (which does not consider the advertiser fixed effects). However, 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%.

To see whether the geographic preferences affect bidding as well as participation, we slightly amend the specification in Equation (5) using bid $b_{i,t}$ as the dependent variable. An observation is a bid and we continue to restrict the estimation sample to frequent bidders. 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 is evidence that advertisers are able to target regions where their most valuable customers are, thus increasing surplus.

4.2. Estimation and Counterfactual Simulations

There are a number of reasons to believe that BIN-TAC would perform well in this environment: there are large gaps between the highest and second highest bids and they are bigger than one would expect under independence. Moreover, the virtual valuations are non-monotone; see Figure 1. We now test BIN-TAC's revenue performance against three benchmarks: the second-price auction without reserve, the optimal second-price auction, and the Myerson mechanism.[14] The latter is a natural benchmark because of the non-monotone virtual valuations, even though the failure of the independence assumption means that it is no longer the revenue-optimal mechanism. The alternatives — asymmetric reserves, Crémer and McLean's (1988) side bets or the optimal ex-post individually rational mechanism — are either impractical, or in the latter case, computationally intractable (Papadimitriou and Pierrakos 2011).

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

Step 1: Inference: Our original model was of a single auction rather an entire marketplace. We now offer a richer model that allows us to make sense of the dataset.[15] A sequence of impressions is sequentially auctioned over a time period by second-price sealed bid auctions. Each impression corresponds to a user-product tuple. We assume there is a fixed set of N advertisers that are present throughout that period and have zero costs of participating in an auction. As before, let $i = 1 \dots N$ index the advertisers, $j = 1 \dots J$ index the products, and $t = 1 \dots T$ index the impressions. We assume that advertisers have valuations of the form of $v_{i,t} = f_i(x_t, j_t)$, where x_t is some vector of

[14] In a previous version, we also offered a benchmark where impressions were bundled by user region and product. This strategy generated low revenues, but this tells us little about how an optimal bundling strategy would perform.

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

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

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

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

Step 2b: Simulation of BIN-TAC: BIN-TAC is harder, as a bidder's decision to take the BIN option depends on their beliefs about the distribution of rival valuations, which in turn depends on the information structure. We consider three possible approaches. The first is to generalize the incomplete information model from earlier in the paper, moving from symmetric independent private values to symmetric correlated private values. We do this because the data is not consistent with symmetric independent private values (see Figure 5). Ideally, we would relax both the symmetry and independence assumptions, but finding an equilibrium under asymmetry would require calculating a fixed point in advertiser-specific thresholds, which is computationally challenging. We continue to focus on equilibria in threshold strategies: there is some $\overline{v}_j = \overline{v}_j(p,d,r)$ for each product, above which advertisers will take the BIN option and below which they will TAC. As before, the threshold bidder is indifferent between BIN and TAC, implying that $\overline{v}_j - p = \frac{1}{d}E[\overline{v}_j - Y^* | Y^1 < \overline{v}_j, \overline{v}_j]$. Because of the correlated values, we have to condition on $\overline{v}_j$ itself on the RHS now. 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. We do this product-by-product, using a kernel estimate of the conditional expectation (Pagan and Ullah 1999). We then solve for the equilibrium $\overline{v}_j(p,d,r)$ for each set of BIN randomization parameters (p,d,r), getting an average revenue estimate as follows: [16]

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

where $J(j)$ is the set of all auctions for product j and $b_t^{(k)}$ is the k-th highest bid in auction t.

We also consider a complete information environment. Here we follow the logic of Edelman, Ostrovsky and Schwartz (2007), who suggest that, since these players compete with high frequency and can potentially learn each others' valuations, a complete information model may better approximate reality than an incomplete information model. Under complete information, there can be multiple equilibria: a bidder who knows they will lose is indifferent between BIN and TAC. We refine away this multiplicity by assuming a positive probability that a bidder will "tremble" and play TAC instead of BIN; then, a bidder will take BIN whenever this yields higher expected surplus than TAC, assuming that all other bidders TAC.[17]

Finally, following the computer science literature, we perform a worst-case analysis. Here, we are agnostic as to the particular information structure, and assume that bidders hold the worst-case beliefs for BIN-TAC revenue. Specifically, each bidder believes that all other bidders will choose to TAC and then bid zero, making TAC relatively attractive. This implies that incentives to take the BIN option must be provided directly by the design, through the randomization parameter d and the reserve price r in the TAC auction. Again, simulated revenue is just the average across auctions — see (6).

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 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. For the SPA-T and BIN-TAC, we find these parameters by maximizing simulated revenue. We use a differential evolution routine for the optimization (Storn and Price 1997).

For the Myerson mechanism, we try two approaches. The first we dub "naïve" or Hamiltonian ironing. In this approach, we follow the algorithm suggested in Myerson (1981):

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

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

1. Estimate the marginal distribution of valuations F using kernel-based smoothing.
2. Construct an estimate of the virtual valuation function $\psi(v) = v - \frac{1-F(v)}{f(v)}$.
3. Integrate up the virtual valuation function to obtain the Hamiltonian $H(v) = \int_0^v \psi(s)ds$.
4. Obtain the convex hull of $H(v)$ and get the ironing regions as the linear segments of the hull.

This algorithm is computationally straightforward; the main difficulty is choosing the bandwidth when estimating the bid density. We employ an adaptive bandwidth approach, starting with a global bandwidth chosen by Silverman's rule-of-thumb, and locally increasing the bandwidth for any grid point for which there are fewer than 500 observations within the bandwidth. This sort of adaptive approach is necessary because the bids are so skew; using a uniform bandwidth would enforce different bias-variance tradeoffs over different parts of the support.

Estimating the marginal distribution F and ironing the resulting virtual valuations to get the ironing regions may not be revenue-maximizing when the data fails the symmetric independent private values assumptions. An alternative approach would be to choose the ironing regions using numerical optimization. Unfortunately, this turns out to be a genuinely difficult optimization problem. Each ironing region introduces two free parameters (each endpoint), and since the optimal number of ironing regions is not known in the general case, neither is the number of parameters. We found the general problem intractable, and instead solved for a single ironing region plus a reserve (three parameters).

To obtain standard errors on our revenue and consumer surplus estimates, we bootstrap at the auction level and re-run the simulation procedure, holding the parameter choices fixed. By bootstrapping at the auction level, we preserve the within-auction correlation structure of the data. We use 100 bootstrap samples.

4.3. Results

The optimal parameter choices are shown in Table 3. 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 (that is close in magnitude to the optimal SPA reserve). To make this buy price attractive, the platform threatens to randomize between the top three bidders in each auction (four in the worst-case scenario), which is significant given that there are only six bidders in an average auction. The numerically optimized version of the Myerson mechanism has an ironing region that starts at the reserve, which is a little unusual. We discuss why this may be below.

The Hamiltonian version of the Myerson mechanism has six ironing regions. This is because the estimated virtual valuation function is non-monotone in multiple places. The number and location of the ironing regions are sensitive to how we choose to estimate the bid density (e.g. bandwidth

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

Table 3 Revenue-Maximizing Parameter Choices. Derived by maximizing the revenue functions defined in the main text over the available parameters using a global optimization algorithm based on differential evolution. For the Myerson mechanism $\underline{v}$ and $\overline{v}$ are the left and right endpoints of the ironing region.

choice and adaptive versus non-adaptive estimator) — this is a practical econometric problem in the implementation of the Myerson mechanism with ironing. For this reason, we prefer the numerically optimized version because — like the estimated parameters for the SPA and BIN-TAC — it does not require an initial estimate of the bid density.

Table 4 shows the mechanisms' performance. The SPA without reserve earns revenue of 0.98 per auction, and leaves substantial consumer surplus — on average 1.97 per auction.[18] Adding the large optimal reserve improves revenue slightly (to 1.03 per auction) but hurts the consumer surplus substantially (it falls to 1.47). Based on the incomplete information results, BIN-TAC does better than both of these mechanisms in terms of revenue: it outperforms the SPA-T with optimal reserve by 4.4%. As the table shows, this finding is robust to the assumed information structure.

Interestingly, BIN-TAC 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 so the reserve is much lower and more impressions are sold. Even accounting for distortions owing to the TAC auction, this is a welfare improvement.

The results for the Myerson mechanisms are intriguing. The Hamiltonian approach performs terribly: it achieves lower revenue than the SPA-T with an optimal reserve. The numerically optimized Myerson mechanism with a single ironing region does much better, which is possible because the Hamiltonian approach may choose sub-optimal ironing regions (the algorithm guarantees optimal output if the environment is SIPV and the correct marginal value distribution is inputted, but both of these assumptions fail here).

However, BIN-TAC beats even the numerically optimized Myerson mechanism. Figure 5 gives us some insight into why this occurs. Bidders take the BIN option when their values exceed a

[18] The per auction revenue of 0.98 is lower than the average second highest bid of 1.07 in Table 1 because some auctions in the data have a single bidder, which will realize revenue of 0 in an SPA without reserve.

Mechanism	Revenue	Advertiser Surplus	Total Surplus
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)	1.072	1.685	2.757
	(0.006)	(0.019)	(0.020)
BIN-TAC (complete information)	1.074	1.641	2.715
	(0.004)	(0.018)	(0.020)
BIN-TAC (worst case)	1.069	1.526	2.594
	(0.005)	(0.018)	(0.020)
Myerson (single optimized region)	1.056	1.713	2.769
	(0.004)	(0.019)	(0.020)
Myerson (Hamiltonian ironing)	1.023	1.822	2.846
	(0.005)	(0.020)	(0.021)

Table 4 Results of counterfactual simulations. Statistics outside parentheses are averages across impressions; those in parentheses are standard errors computed by bootstrapping (100 samples) at the auction level. The average bid is normalized to be equal to 1.

threshold value. In the worst-case scenario, this threshold is constant at 2.6. Consider a scenario in which the highest bid is above the threshold and the second highest bid is below it. BIN-TAC will return the BIN price as revenue, where the BIN price is a weighted average of the reserve and the threshold, and the weights are given *exogenously* by the randomization parameter.

Now consider employing the Myerson mechanism. If the second highest bid falls outside of an ironing region, the high bidder pays the second highest bid, as in a second-price auction. But the stochastically larger bid gaps documented in Figure 5 make this something to avoid. So the optimal ironing region is big, ranging all the way from the reserve price of 0.65 to about 2.6. With this ironing region, revenues in the single-high-bid scenario are again a weighted average of the reserve and threshold, since the threshold and upper ironing endpoint are equal. The difference is that the weights are *endogenous*: they depend on the number of bidders in the ironing region.

In our data, whenever this scenario occurs, the average number of bidders in the ironing region is 2.29, whereas the worst-case randomization parameter is 4. Thus BIN-TAC extracts substantially more revenue in this scenario (2.11 on average versus 1.12). The downside of an exogenous randomization parameter is that sometimes the good is not allocated at all. In auctions in which the highest bid is below 2.6 but above the reserve, BIN-TAC fails to allocate the good in 40% of cases. The Myerson mechanism makes up some ground in these cases.

4.4. Discussion

In this section, we discuss the role of our estimation approach and assumptions in obtaining our results. Our data does not have many conditioning variables — e.g. user age and gender — and

Mechanism	Revenue
SPA (optimal reserve)	1.201
BIN-TAC (complete information)	1.291
BIN-TAC (worst case)	1.276
Myerson (single optimized region)	1.277

Table 5 Counterfactual simulations with product-specific parameter choices.

we have chosen to optimize our parameter choices at a very coarse level, for the platform as a whole. It is likely that both the bidders and the platform have more information at their disposal. This raises a couple of questions: are we modeling the bidder decision process correctly? And if the platform were to engage in finer design (e.g. by choosing separate parameters by product), would our revenue ordering across mechanisms still hold up?

We have tried to be agnostic as to how much bidders know, by offering on the one hand an incomplete information model in which they condition only on their own value and the product identity; and on the other hand, a complete information model in which they know all the other bidders' values. While the results from these two extremal information structures do not formally bound the results from arbitrary information structures, we find it reassuring that they are pretty similar, and in any case, the revenue ordering remains unchanged even in the worst case.

We investigate the question of product-by-product design in Table 5. In the table, we report the average revenues from four mechanisms: the second-price auction with reserve, the BIN-TAC mechanism under complete information and under the worst case, and the Myerson mechanism with a single numerically optimized ironing region. We omit incomplete information BIN-TAC and (Hamiltonian) Myerson because they require estimation of a conditional expectation and a bid density respectively, and it is time consuming to do this carefully for all 87 different products. In each case, the relevant parameters (e.g. BIN price and ironing endpoints) are optimized at the product level.

Revenues are improved under all mechanisms, as one would expect given the finer design. Moving from an SPA with a single reserve across all products to the product-specific reserves increases the average revenue from 1.028 units to 1.201 units, a 16.8% gain. Both BIN-TAC and the Myerson mechanism increase average revenue from 1.201 to at least 1.276 (in the worst case), a further increase of 6.2%. This gain is bigger than the 4.4% we obtained in the uniform parameter choice case, so the product-specific parameters and more complex mechanisms are complementary here. BIN-TAC outperforms Myerson slightly in all cases but the worst case. However, the performance ranking is not uniform; for some products, BIN-TAC substantially outperforms Myerson, while for others the reverse is true.

Of course, it is possible that with more conditioning variables we would be able to divide impressions within a product on the basic of user characteristics into smaller groups within which the

SIPV assumptions are approximately satisfied. Then second-price auctions with finely tailored reserves would be optimal. We do not know whether this is possible. Even if it is, one may be concerned about ratchet effects if the optimal reserves within each group are determined on the basis of the bids of a small number of bidders.

It is harder to assess the implications of our initial assumptions. By assuming that advertisers know their payoffs, we implicitly rule our common values. The most valuable users are those that frequently shop online, and advertisers may be able to assess past purchasing behavior by matching user cookies with their own databases (Abraham et al. 2010). In this case, values are interdependent: advertiser A would like to know what advertiser B knows about the user's past purchases. Our results on the common-value two-type case shed light on how this may bias our results: we expect that BIN-TAC remains more favorable than the SPA, but our simulations may upwardly bias the magnitude of the difference. The reason is that, under common values, taking the BIN option is less attractive due to a winner's curse, see Section 3.6.

We also assume that the 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. The bias introduced by this should be minimal as long as c_i is uniformly small across bidders since our estimated valuations will be close to the truth and bidders are unlikely to change their participation decisions under the counterfactual mechanisms. As evidence that c_i is small for many bidders, 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 an even lower expected surplus.

The last big assumption is that bidders maximize the sum of their payoffs across all impressions, *without any budget constraint*. We ignore budget constraints because they are unobserved by us in the data. Budget constraints, even in static settings, can create complications (cf. Dobzinski, Lavi and Nisan (2008)). Recently, Balseiro et al. (2012) showed that shading strategies, where advertisers shade their bid by a constant factor, can define fluid mean field equilibria. The intuition behind these strategies is that a budget-constrained advertiser could reduce her bid to purchase impressions at a lower price (only when the second highest bid is not high). The use of shading strategies would decrease revenues in all mechanisms; it is unclear to us how this biases the difference in revenues across mechanisms.

5. Conclusion and Future Work

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

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

Acknowledgment All authors are grateful to Microsoft Research New England for their hospitality. Greg Lewis would also like to thank NET Institute (www.NETinst.org) and the National Science Foundation for financial support. Josh Feng and Danyang Su provided excellent research assistance. Hamid Nazerzadeh's research was supported in part by a Google Faculty Research Award. We thank Dan Ackerberg, John Asker, Susan Athey, Alessandro Bonatti, Ying Fan, Michael Grubb, Tom Holmes, Ali Hortaçsu, Jakub Kastl, Jon Levin, Michael Ostrovsky, Mallesh Pai, Ariel Pakes, Jimmy Roberts, Hal Varian, Ricky Vohra, Joel Waldfogel, Glen Weyl and Lee Zen for fruitful discussions and useful comments. We also would like to thank the department editor, Assaf Zeevi, and the anonymous associate editor and referees for their valuable suggestions that improved this work.

References

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

Athey, Susan and Denis Nekipelov, "A Structural Model of Sponsored Search Advertising Auctions," 2010. Working Paper, Harvard University.

Backus, Matthew and Gregory Lewis, "A Demand System for a Dynamic Auction Market with Directed Search," 2012. Working paper, Harvard University.

Balseiro, Santiago, Jon Feldman, Vahab S. Mirrokni, and S. Muthukrishnan, "Yield optimization of display advertising with ad exchange," *ACM Conference on Electronic Commerce*, 2011, pp. 27–28.

Balseiro, Santiago R., Omar Besbes, and Gabriel Y. Weintraub, "Auctions for Online Display Advertising Exchanges: Approximations and Design," 2012. Working paper, Columbia GSB.

Bergemann, Dirk, Alessandro Bonatti, and Maher Said, "Product Design in Ad Exchanges," 2011. Working paper, MIT Sloan.

—— **and** —— , "Targeting in Advertising Markets: Implications for Offline vs. Online Media," *Rand Journal of Economics*, 2011, *42*, 417–443.

—— **and Martin Pesendorfer**, "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 2007, *137*, 580–609.

Bhalgat, Anand, Jon Feldman, and Vahab S. Mirrokni, "Online allocation of display ads with smooth delivery," in Qiang Yang, Deepak Agarwal, and Jian Pei, eds., *KDD*, ACM 2012, pp. 1213–1221.

Board, Simon and Andy Skrzypacz, "Revenue Management with Forward-Looking Buyers," 2013. Working Paper.

Budish, Eric and Lisa Takeyama, "Buy prices in online auctions: irrationality on the internet," *Economics Letters*, 2001, *72* (3), 325—333.

Bulow, Jeremy and Paul Klemperer, "Auctions Versus Negotiations," *The American Economic Review*, 1996, *86* (1), pp. 180–194.

Chu, Sean, Phillip Leslie, and Alan Sorensen, "Bundle-size Pricing as an Approximation to Mixed Bundling," *American Economic Review*, 2011, *101* (1).

Courty, Pascal and Hao Li, "Sequential Screening," *The Review of Economic Studies*, 2000, *67* (4), pp. 697–717.

Crémer, Jacques and Richard McLean, "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 1988, *56* (6), 1247–1257.

Dobzinski, Shahar, Ron Lavi, and Noam Nisan, "Multi-unit Auctions with Budget Limits," in "FOCS" IEEE Computer Society 2008, pp. 260–269.

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.

Feldman, Jon, Nitish Korula, Vahab S. Mirrokni, S. Muthukrishnan, and Martin Pál, "Online Ad Assignment with Free Disposal," in Stefano Leonardi, ed., *WINE*, Vol. 5929 of *Lecture Notes in Computer Science* Springer 2009, pp. 374–385.

Ghosh, Arpita, Hamid Nazerzadeh, and Mukund Sundararajan, "Computing Optimal Bundles for Sponsored Search," in "Internet and Network Economics, Third International Workshop (WINE)" 2007, pp. 576–583.

Goldfarb, Avi and Catherine Tucker, "Online Display Advertising: Targeting and Intrusiveness," *Marketing Science*, 2011, *30* (3), 389—404.

____ and ____ , "Privacy Regulation and Online Advertising," *Management Science*, 2011, *57* (1), 57—71.

Haile, Philip A. and Elie Tamer, "Inference with an Incomplete Model of English Auctions," *The Journal of Political Economy*, feb 2003, *111* (1), 1–51.

Hartline, Jason D. and Tim Roughgarden, "Simple versus Optimal Mechanisms," in "Proceedings of the 10th ACM Conference on Electronic Commerce" 2009.

Johnson, Garrett, "The Impact of Privacy Policy on the Auction Market for Online Display Advertising," 2013. Working Paper.

Kakade, Sham M., Ilan Lobel, and Hamid Nazerzadeh, "Optimal Dynamic Mechanism Design and the Virtual Pivot Mechanism," *Operations Research*, 2013, *61* (4), 837–854.

Levin, Johathan and Paul Milgrom, "Online Advertising: Heterogeneity and Conflation in Market Design," *American Economic Review, Papers and Proceedings*, 2010, *100* (2).

Lewis, Randall A and David H Reiley, "Does Retail Advertising Work? Measuring the Effects of Advertising on Sales via a Controlled Experiment on Yahoo!," 2011. Yahoo! Research Working Paper.

Lewis, Tracy and David Sappington, "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 1994, *35* (2), 309–327.

Madarász, Kristóf and Andrea Prat, "Screening with an Approximate Type Space," Dec 2010. LSE Working Paper.

Mahdian, Mohammad, Arpita Ghosh, Preston McAfee, and Sergei Vassilvitskii, "To match or not to match: Economics of cookie matching in online advertising," in "Proceedings of the 13th ACM Conference on Electronic Commerce" ACM 2012, pp. 741–753.

Maskin, Eric and John Riley, "Optimal Auctions with Risk Averse Buyers," *Econometrica*, nov 1984, *52* (6), 1473–1518.

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

____ and Sergei Vassilvitskii, "An Overview of Practical Exchange Design," *Current Science*, November 2012, *103* (9), 12.

____ , 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.

Milgrom, Paul R. and Robert J. Weber, "A Theory of Auctions and Competitive Bidding," *Econometrica*, 1982, *50* (5), 1089–1122.

Muthukrishnan, S., "Ad Exchanges: Research Issues," 2010. manuscript.

Myerson, Roger B., "Optimal Auction Design," *Mathematics of Operations Research*, 1981, *6* (1), 58—73.

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

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

Papadimitriou, Christos H. and George Pierrakos, "On optimal single-item auctions," in Lance Fortnow and Salil P. Vadhan, eds., *STOC*, ACM 2011, pp. 119–128.

Radovanovic, Ana and Assaf Zeevi, "Revenue maximization through "smart" inventory management in reservation-based online advertising," *SIGMETRICS Performance Evaluation Review*, 2010, *38* (2), 33–35.

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

Riley, John G. and William F. Samuelson, "Optimal Auctions," *American Economic Review*, 1981, *71* (3), 381—392.

Storn, Rainer and Kenneth Price, "Differential Evolution A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *Journal of Global Optimization*, 1997, *11* (4), 341–359.

Varian, Hal, "Position Auctions," *International Journal of Industrial Organization*, 2007, *25* (7), 1163–1178.

Appendix A: Proofs

Proof of Theorem 1 It is weakly dominant for bidders to bid their true valuation in any auction that occurs. So we focus on the BIN decision, taking the auction strategies as given. We first establish that any symmetric equilibrium must be in symmetric threshold strategies. Let a be a binary choice variable equal to 1 if the bidder takes BIN and 0 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 all bidders bid their valuations in the auctions. Then $\frac{\partial}{\partial v}\pi(1,v)\geq q$, as a marginal increase in type increases the BIN payoff at rate 1 whenever no one else takes the BIN option. Similarly $\frac{\partial}{\partial v}\pi(0,v)=\frac{q}{d}<q$. Then $\pi(a,v)$ satisfies the strict single crossing property in (a,v); it follows by Theorem 4 of Milgrom and Shannon (1994) that the best response function must be strictly increasing in v, which implies a threshold rule for the BIN decision.

Next we show that such a threshold $\overline{v}$ exists. The threshold type must be indifferent between taking the BIN option and declining it. The threshold type will only win the object if all other players decline BIN. In that event, taking BIN gives a payoff of $\overline{v}-p$. Declining it gives $\frac{1}{d}\mathbb{E}[\overline{v}-Y^*|Y^1<\overline{v}]$, where we condition on $Y^1<\overline{v}$ since all rivals declined BIN. Any threshold must thus solve the equation:

$$\overline{v}-p=\frac{1}{d}\mathbb{E}\left[(\overline{v}-Y^*)\,|\,Y^1<\overline{v}\right] \tag{7}$$

Now at $\overline{v}=\underline{\omega}_L$, we have that the $LHS<RHS$ since the LHS is negative ($p>\underline{\omega}_L$) and the RHS is equal to 0 since the reserve will bind. Moreover, the slope of the LHS in $\overline{v}$ is equal to 1, while the slope of the RHS is $\frac{1}{d}(1-\frac{\partial}{\partial\overline{v}}\mathbb{E}[Y^*|Y^1<\overline{v}])<1$. So either there is a unique solution to Equation (7) on the interior of $[\underline{\omega}_L,\overline{\omega}_H]$ (corresponding to Equation (2)), or there is no solution at all. In the former case, there is a unique equilibrium: by the single crossing property, no type above $\overline{v}$ wants to deviate to TAC, and no type below $\overline{v}$ wants to deviate to BIN. In the latter case, the unique equilibrium is for all types to TAC: since the LHS is everywhere lower, $\overline{\omega}_H$ prefers TAC to BIN and by single-crossing, so does everyone else.

□

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

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

- *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 $(1 \le k \le d)$ 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

$$r\left(\sum_{k=0}^{d-1}\frac{n}{d}f(r)\binom{n-1}{k}(F(\overline{v})-F(r))^k F(r)^{n-k-1}\right) = rf(r)\left(\sum_{k=1}^{d}\frac{n}{d}\binom{n-1}{k-1}(F(\overline{v})-F(r))^{k-1}F(r)^{n-k}\right) \tag{11}$$

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

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

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:

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

Re-arranging terms, the optimal choice for $\overline{v}$ is either at the boundaries or is a solution of the equation $d-1+\frac{\partial E[Y^*|Y^1<\overline{v}]}{\partial \overline{v}}=\left(\frac{(d-1)f(\overline{v})}{1-F(\overline{v})}+\frac{(n-1)f(\overline{v})}{F(\overline{v})}\right)E\left[\overline{v}-Y^*|Y^1<\overline{v}\right]$. Similar to the previous argument, it is easy to see that the only boundary that would be optimal is $\underline{\omega}_H$. □