Competition and Targeting in Auctions for Online Ads *

Tulio Sousa†

November 1, 2024

Abstract

The digital display advertising industry is phasing out third party cookies due to growing concerns about data privacy, fundamentally altering how advertisers target and reach consumers. This paper investigates the potential impact of this shift on advertisers and platforms, focusing on the retargeting strategy, where advertisers reach consumers who have previously visited their website. Utilizing a dataset of 1.4 billion auctions paired with an experiment varying bids of live campaigns, I estimate a structural model of endogenous advertiser entry in the online advertising industry. The model disentangles two opposing forces influencing winning bids when third-party retargeting cookies are removed: the matching effect, where the inability to precisely identify and reach advertisers' most valued audiences decreases their bids, and the competition effect, where advertisers face tougher competition in auctions after adjusting their targeting strategies toward broader audiences, increasing the winning bids. Counterfactual simulations reveal that removing third-party cookies would decrease total ad spend by 60%, with an 84% reduction being due to lower match values, partially offset by a 24% increase arising from intensified competition.

Keywords: Online Advertising, Retargeting, Data Privacy, Third-Party Cookies, Auctions, Structural Estimation
JEL Codes: C15, C54, C55, D22, D44, L51, L81, L86, M31, M37

*I am grateful to Allan Collard-Wexler, Carl Mela, James Roberts and Daniel Xu for their guidance and support. I thank Pellumb Reshidi, Emily Cuddy, Michael Dinerstein, Srinivas Tunuguntla, John Shin, Seohee Kim, Siddharth Prusty, Levin Zhu, Tirza Angerhofer, several contacts at the advertising industry, and participants of the Duke IO Lunch, Duke Public/IO Seminar, UNC IO Workshop, and Triangle Micro Conference 2023 for helpful discussions and conversations.

†PhD Candidate at Duke University, Department of Economics: tds46@duke.edu

1 Introduction

The online ad industry matches advertisers to potential consumers through targeted ads allocated via auctions. In 2023, U.S. advertisers spent over $150 billion on digital display ads, with personalized auction-based ads accounting for $38 billion of that spend.[1] A key factor driving this market is how advertisers leverage the rich information available on consumers to target them with personalized ads. One targeting strategy is *retargeting*, where advertisers show ads to consumers who previously visited their website. For instance, if a consumer browses a mattress on a retailer's website, the retailer can later display ads about that specific mattress to the same consumer across other websites and platforms she visits. Retargeting relies on identifying individual consumers across different websites, which is largely enabled by third-party cookies. However, highly targeted ads that improve the match between advertisers and consumers come at the cost of reduced consumer online privacy. For this reason, third-party cookies have recently faced substantial scrutiny from regulators in Europe and the U.S., which has spurred a wave of incipient data privacy legislation. Anticipating this trend, major industry players have already begun phasing out third-party cookies. Both Safari and Firefox now block third-party cookies by default, and Google plans to deprecate them on Chrome by late 2025.[2]

This paper studies the equilibrium effects, for advertisers and platforms, of removing third-party cookies on the $38 billion personalized advertising market. The ability to identify specific individuals across different websites allows advertisers to target an audience that best represents the set of consumers more interested in their products. Because targeted ads are allocated via auctions, this implies that more granular information on consumers enables advertisers to enter fewer auctions, effectively reducing the number of competitors in each auction. Therefore, the removal of third-party cookies poses a tradeoff from a policy

[1] EMarketer Programmatic Ad Spending Forecast H1 2024, available at `https://image.insight.insiderintelligence.com/lib/fe2d11737164047b7c1576/m/1/Programmatic+Ad+Spending+Forecast+H1+2024_LiveRamp.pdf` (Archive.org, accessed on 06/21/2024)

[2] `https://developers.google.com/privacy-sandbox/3pcd` (Archive.org, accessed on 06/21/2024)

perspective: granular information on consumers improves the match that advertisers have with their audience, but it creates a thinner market that advertisers can explore by facing lower competition in auctions. Studying this tradeoff requires understanding the mechanisms behind the improved match and lower competition associated with the granular information advertisers have on consumers. In particular, I focus on the implications for retargeting, a highly valuable strategy used by advertisers. In the dataset, winning bids by advertisers employing retargeting are two orders of magnitude higher than those that do not. Moreover, retargeting accounts for 1.4% of the 1.4 billion auctions, but it represents 42.3% of the total ad spend. Despite its importance to advertisers, retargeting is one of the strategies most impacted by the removal of third-party cookies because it heavily relies on individual identification of users across websites.

More specifically, the equilibrium impact of eliminating third-party cookies on advertiser auction bidding behavior hinges on two countervailing forces affecting winning bids. On the one hand, the inability to precisely identify and reach advertisers' most valued audiences could decrease advertiser bids because of a lower match value: if advertisers are no longer able to recognize the most valuable users online, they will not bid as much in auctions as they did previously. On the other hand, advertisers will adjust their targeting strategies following the removal of cookies. This could intensify auction competition, pushing bids higher as advertisers target broader audiences. For example, if an advertiser is no longer able to target the precise set of users who visited its website in the past, it may attempt to reach a broader audience that it believes loosely matches the audience interested in its product. Therefore, the advertiser will compete in more auctions, intensifying competition in the industry and, as a consequence, increasing the value of winning bids.

This paper contributes to the literature in four key ways. First, it endogenizes advertiser participation in retargeting. There exists a body of literature analyzing the impact of data privacy policies on advertisers and platforms in the context of online targeted ads (Johnson 2013, Alcobendas et al. 2021). However, those papers focus on how the match

between advertisers and consumers will be impacted and neglect the fact that advertisers will reoptimize their entry in auctions following a data privacy policy that restricts information granularity on consumers. This paper highlights the separate contribution of the matching and competition effects on ad spend and advertiser welfare. Properly separating these effects requires endogenizing advertisers' decision to enter the targeted advertising market. This is important for several reasons. For one, regulators in the U.S. and Europe are concerned about potential unintended consequences on competition in the online advertising industry following the removal of third-party cookies. If the goal is to analyze policy effects on competition, one needs a model that endogenizes players' decision to enter the market. Second, given the massive scale of the online advertising industry — with billions spent annually on targeted ads — accurately distinguishing between matching and competition effects is crucial for understanding the potential economic impact of policy changes in this multi-billion dollar market. Finally, the tradeoff between thin and thick markets goes beyond the data privacy policy brought in this paper. It is common to observe conflation in commodity markets, where products with similar characteristics are grouped together to create a thicker market. For example, this happens on the wheat, cocoa and diamond markets (Levin and Milgrom 2010), as well as in the stock market (Budish et al. 2024).

To disentangle the matching and competition effects, I estimate a structural model of endogenous advertiser entry in retargeting audience segments. I build on the literature on auctions with an entry stage (Samuelson 1985, Levin and Smith 1994, Li and Zheng 2009, Athey et al. 2011, Roberts and Sweeting 2016). In the model, advertisers first decide whether to enter the segment of auctions for personalized retargeted ads, and then participate by submitting bids. The model includes fixed entry costs to capture managerial or attention costs associated with monitoring the performance of various campaigns.[3] Using a unique dataset that covers 1.4 billion auctions and more than 200 advertisers in the Brazilian market from November 2021 to March 2022, I recover the full equilibrium strategies governing adver-

[3]An ad campaign represents a configuration of targeting settings and bid strategies defined by the advertiser to achieve a specific goal. More on this in Section 3.1.

tiser entry and bidding decisions. The model is able to accurately replicate the fact that winning bids for retargeting campaigns are two orders of magnitude higher than those for non-retargeting campaigns - there is a large valuation premium for retargeting campaigns. In addition, the average of the entry cost distribution estimated by the model represents 12% of the typical retargeting winning bid, but as much as double the typical non-retargeting winning bid. The higher entry rate for retargeting predicted by the model is in line with the empirical regularity that advertisers simultaneously run various retargeting campaigns, whereas they usually run fewer broad, non-retargeting campaigns.

With the model estimates, I evaluate counterfactual scenarios where retargeting is either no longer possible or more costly, and disentangle the two opposing forces acting on equilibrium bids. I find that eliminating third-party cookies would decrease total ad spend by 60%, with an 84% reduction coming from a pure decrease in the match value between advertisers and consumers, counterbalanced by a 24% increase in bids from fiercer competition in auctions. Advertiser surplus from auctions would be reduced by 79%. This would be the result of a 72% decrease due to a lower match with consumers and a further 7% decrease from the fiercer auction competition. If the entry cost for retargeting is doubled, which simulates increased costs of adopting alternative targeting technologies such as Google's Privacy Sandbox, the model predicts that expected auction revenue would decrease by 12%.

Another contribution is that the model proposed in this paper can be used for targeting settings other than retargeting. This paper focuses on retargeting because it is the targeting strategy most impacted by the deprecation of third-party cookies. However, it is possible to utilize the same framework employed here if one has data on advertisers using other targeting strategies - for example, geographical, contextual or demographical targeting. This poses two benefits. First, it would increase the realism of the model by incorporating better alternative targeting strategies that advertisers can use if they are no longer able to employ retargeting. Second, it opens up possibilities for future research and more accurate policy analysis when evaluating the effects of data privacy policies. The

regulatory landscape of the online advertisement industry continues to evolve, and other policies that restrict different types of targeting strategies can be the focus in the future.

A third contribution of the paper is to provide identification of parameters governing advertisers' entry into retargeting. Typically, the literature states that identification of the entry cost distribution requires observing either some measure of entry costs or the number of actual and potential bidders (Roberts and Sweeting 2010). The dataset used in this paper has none of those. Instead, the entry cost distribution is identified by using variation in the winning probabilities of the ad campaigns ran by advertisers. Those winning probabilities come from an auction planner tool, which is a website that advertisers use to plan the performance of advertising campaigns before going live with them. Through simulation exercises, I claim that there is a direct relationship between the parameters of the entry cost distribution and the unobserved number of actual bidders in each auction. This relationship translates into a relationship between the distribution parameters and the observed winning probabilities. This is what identifies the parameters of the entry cost distribution.

Finally, this paper also contributes to the literature by pairing the structural model with exogenous variation from an experiment. Between February 22 and 25 2022, I varied the bids in 12 live campaigns from 6 different advertisers. The main goal of the experiment is to test whether these online ad auctions are more consistent with the assumption of independent private values (IPV) or common values (CV), as this poses potential endogeneity problems when estimating advertisers' valuations. I find that, conditional on the targeting strategy, auctions won after an increase in bids do not yield higher click-through rates, consistent with the IPV assumption, which I then use in the structural model.

Related Literature This paper relates to the empirical literature on targeted ads and data privacy (Dube et al. 2024). When banning third-party-cookies, advertisers' valuation and targeting strategies change simultaneously in equilibrium. Evaluating the effects of data privacy policies on advertisers and platforms requires understanding the equilibrium impacts

of advertisers' decision to enter the market following changes to cookie access. These effects can alter the competitive landscape of the industry, which is one of the major concerns regulators currently have with the deprecation of third-party cookies. This paper estimates a structural model of advertiser entry into targeted segments, endogenizing the equilibrium effects that follow in the counterfactual scenario where cookies are no longer available. By doing so, I can separately quantify the importance of the change in valuations from the change in competition when determining the final effect of the change in information available to advertisers on auction revenue and advertiser surplus.

The most relevant papers to my work are Johnson (2013) and Alcobendas et al. (2021). Both studies examine the impact of privacy regulations on the online advertising industry. Johnson (2013) estimates bidding strategies for advertisers that use consumer tracking technologies, with a focus on behavioral targeting. The author simulates counterfactuals where different fractions of the user population can no longer be tracked. Similar to my dataset, his includes bids from hundreds of different bidders placing ads on thousands of websites. However, the author does not observe the precise audience sizes and targeting settings of each advertiser. Consequently, his structural model includes audience sizes as parameters to be estimated, and the advertisers' valuation of impressions includes an unobserved heterogeneity term to account for the targeted user characteristics not observed by the econometrician. In contrast, I observe both audience sizes and advertisers' targeting specifications. This, paired with the data on winning probabilities, allows me to estimate an entry stage to the auction game, where advertisers first decide whether to enter the retargeting segment, and then submit bids. I estimate advertisers' entry strategy in equilibrium, which varies in the counterfactual scenario where targeting is no longer available.

More recently, Alcobendas et al. (2021) study the specific impact of Google's removal of third-party cookies. Similar to Johnson (2013), the authors also estimate bidding strategies of advertisers. I differ from Alcobendas et al. (2021) in two important ways. First, I focus on retargeting strategies, while those authors focus on demographic targeting such

as age and gender. My focus on retargeting is justified because retargeting heavily relies on third-party cookies and is not easily substitutable in their absence. Second, advertiser entry in Alcobendas et al. (2021) is interpreted as submitting a bid above the reserve price. That is, the authors take advertisers' decision of what targeting strategy to employ as given in the counterfactual scenario with no cookies. In contrast, I endogenize advertisers' decision of which targeting strategy to use following the removal of cookies and compute a new market equilibrium. This is important because advertisers are likely to change their entry rate in targeting strategies that remain viable in the absence of cookies. Such change happens because advertisers running broader, non-retargeting campaigns will no longer face high-valuation competitors engaging retargeting. Hence, their surplus in those non-retargeting auctions tends to increase, giving them greater ability to bear higher entry costs and more frequently enter non-retargeting campaigns.

This paper also relates to the general literature on targeted ads and the value of information leveraged by consumer tracking. Several papers either document or measure how valuable targeting is to advertisers (Wernerfelt et al. 2022, Goldfarb and Tucker 2011, Marotta et al. 2019, Lambrecht and Tucker 2013, Deisenroth et al. 2024). In particular, Wernerfelt et al. (2022) perform a large-scale experiment with thousands of advertisers on Meta. They document that advertisers' cost of acquiring a new customer would increase by 37% if they were not able to use off-site data in their targeting settings.[4] In the dataset used in this paper, retargeting provides a substantially higher premium on bids. If one compares campaigns that rely solely on retargeting to campaigns that do not target consumer characteristics, the bids in the dataset are as much as two orders of magnitude higher. Also related to my paper but in the context of Apple's ATT policy, Deisenroth et al. (2024) use difference-in-differences techniques to estimate the causal effect of ATT on several outcome variables, including ad spend and firm entry and exit. The authors find that firms tend to exit the market and overall ad spend decreases, but advertising levels of surviving firms

[4]In their setting, off-site data is defined as data originating off of Meta's sites and applications.

increase. This is in line with the conclusion I reach in this paper that the entry rate of advertisers in non-retargeting campaigns increases following the deprecation of third-party cookies.

I also relate to the recent empirical literature on the impact of data privacy regulations on advertisers and platforms (Johnson 2013, Alcobendas et al. 2021, Johnson, Shriver, et al. 2020, Goldberg et al. 2024, Aridor and Che 2024, Sun et al. 2024, Korganbekova and Zuber 2023, Peukert et al. 2022). I find that the advertising spend in ad auctions would decrease by 60% following the removal of third-party cookies. This number is similar to those found in the literature. For example, in 2019 Google ran an internal experiment by disabling cookies for a randomized set of publishers and found that the ad auction revenue received by those publishers decreased by 52% (Ravichandran and Korula 2019). Similarly, Johnson, Shriver, et al. (2020) find that auctions for users who opt out of cookies close at a price 52% lower than for opt-in users; Alcobendas et al. (2021) find a reduction in publisher revenue of 54% if cookies are removed; and Johnson (2013) finds that advertiser surplus would decline by 45% under a regulation that bans user tracking. I extend this analysis by estimating advertisers' full entry strategies for retargeting and non-retargeting settings, which allows a decomposition of the final effect into two opposing forces: I find that the 60% decrease in auction revenue is a result of an 84% decrease in match value between advertisers and consumers, offset by a 24% increase in bids submitted because of fiercer competition. Conversely, Peukert et al. (2022) find that Google's market share increased after the implementation of the General Data Protection Regulation (GDPR). This is the kind of effect that concerns regulators in Europe regarding the adoption of the Privacy Sandbox, Google's proposal to replace third-party cookies.

This paper also relates to the theoretical literature on thin markets in ad auctions, first introduced by Levin and Milgrom (2010) and followed by Chen and Stallaert (2014) and Hummel and Mcafee (2016). These papers present the idea that more finely targeted ads may not monotonically increase ad auction revenue. As one increases the level of infor-

mation available for targeting, fewer bidders participate in the auction for a specific user, which decreases auction competition and lowers advertisers' bids. Interestingly, Rafieian and Yoganarasimhan (2021), Lu and Yang (2020), Ahmadi et al. (2024) and Ada et al. (2022) provide results in this direction. Rafieian and Yoganarasimhan (2021) and Lu and Yang (2020) conclude that the optimal level of targeting to maximize platform revenue is not to provide as much information as possible to advertisers. Similarly, Ahmadi et al. (2024) show that, in some cases, ad campaigns that are not targeted to specific audiences can provide advertisers with more surplus than finely targeted campaigns. Ada et al. (2022) show that when bidders gain more contextual information about the impressions they are attempting to buy, they react by increasing their bids, and this effect is larger for thicker markets. The structural model I build and estimate in this paper allows the separate quantification of the matching and competition effects. I am able to provide the importance of each effect in a counterfactual scenario where third-party cookies are banned by endogenizing advertisers' entry choice into targeted segments. This is important in light of the growing regulatory scrutiny received by the industry and the changes advertisers will face if third-party cookies are phased out. I show that the competition effect would increase bids for non-retargeting campaigns by as much as 24%. This is a relevant effect, and competition regulators across Europe and the US should consider this effect when designing new data privacy policies.

Finally, the proposed empirical framework in this paper can be used in future research on the targeted ads industry. This paper exploits a unique dataset containing winning bids, targeting settings, audience sizes and winning probabilities for ad campaigns. The focus of this paper is on retargeting, as it heavily relies on the information provided by third-party cookies and is highly valuable for advertisers. However, the empirical framework provided here would allow future research to expand advertisers targeting options if the researcher has access to data containing different targeting strategies. Allowing advertisers to choose both retargeting and behavior, demographic or contextual targeting would improve the model's realism and fit to the data, providing more accurate counterfactual predictions.

The remainder of this paper is organized as follows. Section 2 provides some institutional background on the online ad auctions industry. Section 3 describes the datasets used in the paper and motivates the focus on retargeting strategies. Section 4 details the structural model used in the paper, and Section 5 details how it was estimated. Section 6 discusses what variation in the data identifies the parameters in the model. Section 7 presents the estimation results, while Section 8 shows the results of counterfactual scenarios where third-party cookies are banned or are more costly to utilize. Finally, I conclude the paper in Section 9.

2 Institutional Background

In this section, I provide a brief explanation of how the digital display ad industry works, the main players in it, and an overview of the regulatory landscape involving the deprecation of third-party cookies. A more detailed description of the industry can be found in Choi et al. (2020).

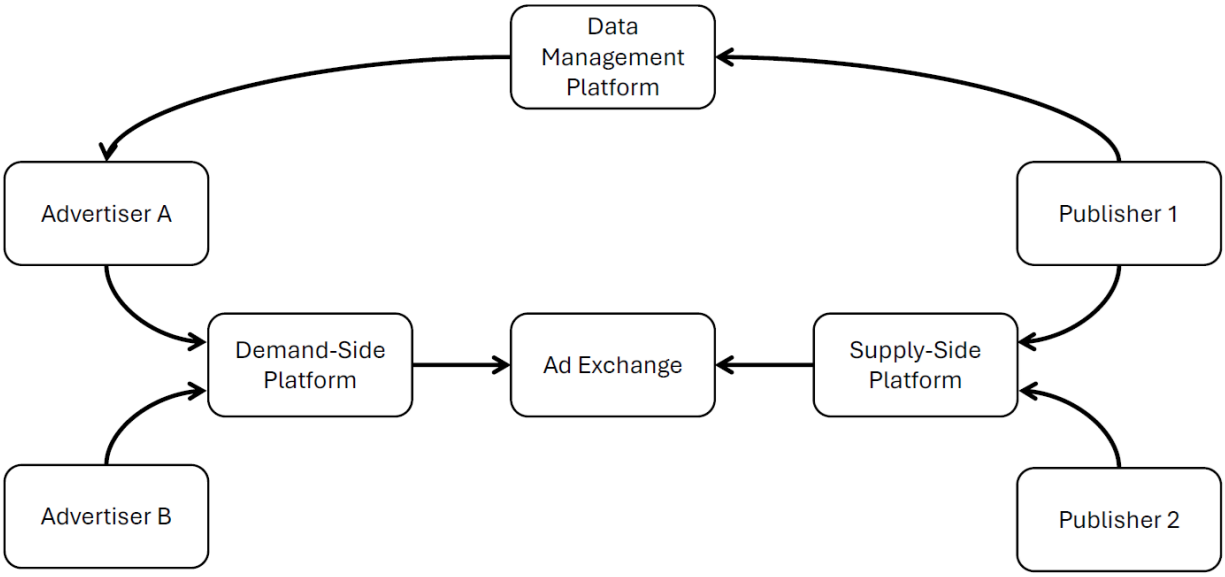
2.1 The Online Advertisement Ecosystem

The digital display advertising ecosystem is a complex network of various parties involved in the automated buying and selling of ad space on websites and apps. Figure 1 shows a simplified diagram of this ecosystem. On the demand side are advertisers, companies looking to display their ads to targeted audiences. Advertisers set up campaigns, which consist of specific targeting configurations (e.g., location, interests, demographics and retargeting) and a maximum bid amount they are willing to pay for reaching those targeted audiences. For instance, an advertiser may create a campaign targeting consumers in a particular region of the country with certain inferred interests, setting a maximum bid of, say, $1.50 per thousand auctions for that audience segment.

The other relevant players on the demand side of the market are the demand-side platforms. They work with advertisers, automating the bidding process across numerous auction opportunities based on the targeting settings and bidding strategies defined in the advertisers' campaigns. Typically, demand-side platforms offer advertisers data on consumer characteristics, but advertisers can also obtain these data from data management platforms, whose purpose in the ecosystem is to gather and sell information about consumers online.

On the supply side are publishers, the website and app owners seeking to monetize their ad inventory space by displaying ads. Publishers make their inventory available through supply-side platforms, which act as consolidated catalogs of potential ad placements. The ad exchanges connect the demand side (advertisers and demand-side platforms) with the supply

Figure 1: Digital Display Advertising Ecosystem.



This figure represents a simplified version of the digital display advertising ecosystem. Advertisers contact Demand-Side Platforms with their targeting settings and bidding strategies. Data Management Platforms collect user data from publishers and sell it to advertisers. Publishers make their inventory available through Supply-Side Platforms, which list it at the Ad Exchange. The Ad Exchange runs the auctions.

side (publishers and supply-side platforms) and run the auctions to allocate ad impressions to the highest bidders.

The process leading to an ad impression begins when a user visits a publisher's website. An auction opportunity arises when the page is loaded, and the supply-side platform sends the auction details, including available user data, to the ad exchange. The exchange queries the connected demand-side platforms, which respond with bids from advertisers targeting those specific user characteristics based on their campaign settings. The exchange then allocates the ad impression to the highest bidder. Traditionally, the standard auction format used in the allocation process has been the second-price auction. However, the industry is gradually shifting toward first-price auctions in an effort to provide greater transparency to advertisers and publishers.[5]

[5] `https://clearcode.cc/blog/first-price-second-price-auction/` (Archive.org, accessed on 06/21/2024)

When an ad auction opportunity arises, a critical piece of information sent by publishers to the ad exchange is the user's cookie ID. Cookies are small text files websites store on users' devices to track the latter's online activities and preferences. There are two main types: first- and third-party cookies. First-party cookies are set directly by the visited website, allowing that site to recognize users across visits and enhance their experience through features such as saved settings, shopping carts, and automatic logins. Third-party cookies, however, are set by third-party entities with code integrated into the website. When a user visits a site with such code, the third-party server places a cookie associated with the third-party's domain on the user's device in response to a request from the user's browser. This cookie can be accessed and updated when the user visits other websites with the same third-party code, allowing these companies to track users' browsing activities across multiple sites and build comprehensive user profiles. Advertisers leverage these third-party cookies to create targeted audiences based on users' online behavior and interests.

2.2 The Regulatory Landscape Surrounding Third-Party Cookies

While first-party cookies are generally regarded as beneficial for users, third-party cookies are not. Although they facilitate the delivery of more relevant ads to consumers, third-party cookies are often deemed invasive for several reasons. These cookies are frequently set without explicit user knowledge or consent, enabling companies to track user behavior across websites without the user's awareness or approval. By doing so, companies can create detailed user profiles by monitoring users across multiple websites, violating consumers' online privacy. Furthermore, these companies often share or sell these data to other entities, again without the user's consent. Such concerns have encouraged a wave of data privacy legislation across Europe and the United States in recent years.

The legislation that set the standards for this regulatory surge is the GDPR. It was implemented in Europe in 2018 and has become the main legislation for data privacy, regarded as the standard for subsequent rules. The GDPR sets standards for data collection,

processing, and storage for websites offering goods or services to European citizens, mandating that websites acquire users' consent to collect and use their personal data. Following the GDPR, several countries have enacted their own data privacy laws, such as the California Consumer Privacy Act (CCPA) in the United States, China's Personal Information Protection Law (PIPL), and Brazil's General Data Protection Law (LGPD).

In anticipation of stricter regulations concerning consumer privacy, major industry players have initiated movements to phase out the use of third-party cookies. Safari and Firefox have already blocked third-party cookies by default on their browsers. In 2020, Google Chrome, which accounts for approximately two-thirds of the global browser market share,[6] announced a plan to remove third-party cookies as well. This plan has garnered substantial attention from both advertisers and regulatory bodies in Europe. The primary concern among legislators is that the removal of third-party cookies from Chrome could solidify Google's already dominant market position. Advertisers would need to explore alternative methods to reach targeted audiences, potentially relying on Google's proprietary solutions,[7] further strengthening Google's market power.

This paper estimates a structural model of endogenous entry in targeted segments. The counterfactuals I provide shed light on the evolving regulatory landscape surrounding the use of highly targeted online advertising driven by third-party tracking and user data collection. Efforts to curtail advertisers' ability to personalize ads, whether by restricting third-party cookies or mandating alternative business models that offer users more privacy choices, could have profound impacts on the online advertisement industry. However, the potential ramifications, particularly the equilibrium implications for auction competition arising from constraining advertisers' access to user information for ad personalization, remain understudied in the economics and marketing literature. By examining these issues, the aim of this research is to contribute to the understanding of how data privacy regulations

[6] `https://gs.statcounter.com/browser-market-share` (Archive.org, accessed on 06/21/2024)

[7] Google's proposal to replace third-party cookies is the Privacy Sandbox. It is a set of solutions aimed at allowing advertisers to still run targeted ads while enhancing consumer privacy. For further details, see `https://privacysandbox.com/` (Archive.org, accessed on 06/21/2024)

may reshape the digital advertising ecosystem and competitive interactions therein.

3 Data

A necessary condition for the tension between matching and competition effects to exist is that markets are not too thick. In a thick market, changes in competition following the removal of third-party cookies would have little impact on final bids, and the total effect would be approximately equal to the matching effect. To demonstrate that online ad auctions involving retargeting are not exceedingly thick, I provide descriptive evidence showing that advertisers' winning probabilities are significantly influenced by whether they are running retargeting campaigns, even after controlling for the size of the audience reached. This finding suggests that advertisers engaging in retargeting face less competition than those running non-retargeting campaigns, as they are targeting a more specific and limited audience. Consequently, the reduced competition in retargeting campaigns can be attributed to their inherent characteristic, which focus on a narrower set of consumers who have previously interacted with the advertiser's website. This evidence supports the existence of a meaningful interplay between matching and competition effects in the context of online ad auctions.

In this section, I also present the results of a field experiment. Between February 22 and February 25, 2022, I varied bids of 12 live campaigns from 6 advertisers in the Brazilian market. The main purpose of the experiment is to test whether the online ad auctions are more consistent with the IPV or CV assumption, as this poses potential endogeneity problems when estimating advertisers' valuations. The main idea is that, under a CV setting, marginal auctions won after an exogenous increase in bids should yield higher payoffs compared to the infra marginal ones. I find that the click-through rates do not change following an exogenous increase in bids, consistent with the IPV assumption.

3.1 Data Sources

The data used in this paper come from a Brazilian company (henceforth, The Company) that acts as an intermediary between advertisers and demand-side platforms.[8] Their primary service is assisting advertisers in identifying the most suitable target audiences for their products, with a strong emphasis on retargeting strategies. Advertisers seeking to implement retargeting campaigns must integrate a code snippet into their websites, commonly referred to as a *pixel*. This pixel facilitates the exchange of cookies between the advertiser's domain and the ad servers. The Company possesses the technological infrastructure to support advertisers in this implementation process. Additionally, The Company helps advertisers in determining the most profitable segments within an advertiser's website for executing effective retargeting strategies.

I use two main datasets. First, I have access to auction-level data encompassing all ad auctions won by The Company's clients between November 2021 and March 2022. For each auction, I observe the winning bid, the advertiser's payment,[9] and identifiers for the advertiser, publisher, and user. Additionally, the dataset includes the date and time of the auction, indicators for whether the user viewed[10] or clicked the ad, and variables describing the user's device type, model, operating system, and geographical location. Notably, I also observe a variable called *campaign ID*. An ad campaign represents a configuration of targeting settings and bid strategies defined by the advertiser to achieve a specific goal. Advertisers can simultaneously run multiple campaigns to reach different target audiences. For instance, a shoe retailer may create one campaign targeting consumers who visited its "running shoes" section, with a maximum bid of $3.00 per thousand auctions. Concurrently, it could run another campaign targeting users who added items to their shopping cart, with

[8] For confidentiality reasons, I cannot disclose the name of the company from which I sourced the data.

[9] If the auction is first-price, those two quantities are the same.

[10] According to the Interactive Advertisement Bureau, an ad banner is considered to be viewed by a user if at least 50% of the ad was visible on the user's screen for at least one continuous second. See `https://www.iab.com/news/viewability-has-arrived-what-you-need-to-know-to-see-through-this-sea-change/` (Archive.org, accessed on 06/21/2024)

a maximum bid of $10 per thousand auctions, or a broad campaign aiming to reach all possible online consumers (i.e. a no-targeting campaign), with a maximum bid of $0.05 per thousand auctions. The presence of the campaign ID in the dataset allows me to observe all targeting configurations for each auction won by the advertisers.

I augment the auction-level data with three unique pieces of information for each campaign. I observe (i) all targeting settings chosen by advertisers in their campaigns; (ii) the resulting audience sizes each campaign reached; and (iii) a discretized CDF of each campaign's winning probability as a function of bids. These data come from the *auction planner*, a tool provided by the demand-side platform to advertisers. The auction planner is a platform that allows advertisers to input their campaign settings and simulate the campaign's performance for a given bidding strategy based on historical data gathered by the demand-side platform.[11] The data I gather from the auction planner complement the auction-level dataset, which does not contain the number of active bidders in each auction. Therefore, I use data on winning probabilities to infer the intensity of competition faced by each campaign.

3.2 Descriptive Analysis

Table 1 provides a summary of the frequency with which advertisers and publishers appear in the auction-level data. For a 1% random sample of the fourth week of November 2021, 196 different advertisers won at least one auction in the data. The average advertiser won 3,891 auctions, but this number varies significantly. Some advertisers won only a few dozen auctions, while the most frequent advertiser won more than 50,000 auctions in the sample. This heterogeneity demonstrates that my dataset contains advertisers running both small, finely targeted campaigns, and large, broad campaigns. Those 196 advertisers ran over 1,958 different campaigns. On average, each advertiser ran 10 different campaigns simultaneously.

[11] The auction planner is a common feature that demand-side platforms offer to advertisers. See, for example, the case of Yahoo (`https://www.advertising.yahooinc.com/our-dsp/plan`) (Archive.org, accessed on 06/21/2024)

Table 1: Frequency Distribution of Various IDs in the Dataset for a 1% Random Sample from Nov 20th to Nov 26th 2021

	Number of Unique IDs	Average Frequency	Std. Frequency	Min	25%	50%	75%	Max
Advertisers	196	3,891	8,409	22	560	1,226	3,084	52,942
Campaigns	1,958	390	2,323	1	2	5	41	41,577
Publishers	2,571	182	3,263	1	1	2	10	126,152
URLs	4,681	163	2,771	1	1	2	9	123,108

This table shows the distribution with which each unique player or object appears in the data. For example, the 1% random sample of the dataset contains 196 unique advertisers and, on average, each advertiser wins 3,891 auctions. The least frequent advertiser won 22 auctions, and the most frequent advertiser won 52,942 auctions.

On the supply side of the market, we observe ads being placed at 2,571 different publishers and 4,681 different URLs.[12] There is also substantial heterogeneity in these numbers. Some publishers are present infrequently in the data, while others appear as many as 120,000 times. This suggests that my dataset covers both niche websites and those with high traffic.

Table 2 shows summary statistics of the main numeric variables in the auction-level dataset for a random 1% of the auctions won by The Company's clients between November 20 and November 26. This accounts for more than 760,000 auctions in the period. There is substantial heterogeneity in the values of bids and payments. Throughout this paper, I multiply any monetary variables by a constant and denote them in an unknown numeraire currency per thousand impressions (N$) for confidentiality. The average bid is less than N$0.1, but the maximum bid found in the dataset is three orders of magnitude higher. The same pattern repeats for the payments. Virtually all auctions resulted in an impression.[13] Finally, 77% of impressions were viewed by the user, and 0.13% resulted in a click.

One of the unique features of the dataset used in this paper is the ability to observe targeting settings for each campaign. Advertisers in my dataset target their audi-

[12] To exemplify the distinction between publisher and URL, one can think of ESPN.com being a publisher, while ESPN.com/nfl and ESPN.com/nba being two different URLs within the same publisher.

[13] A closed auction may not result in an impression if, for some reason, the ad server could not deliver the ad creative to the website after the end of the auction. This mostly happens because of connectivity issues at the user's device.

Table 2: Descriptive Statistics of Numerical Variables in the Dataset for a 1% Random Sample from Nov 20th to Nov 26th 2021

	Mean	Std	Min	25%	50%	75%	Max
Buyer Bid (N$)	0.09	0.96	0.0003	0.01	0.02	0.03	254.14
Buyer Payment(N$)	0.06	0.50	0	0.01	0.02	0.03	133.68
Impression	1.00	0.03	0	1.00	1.00	1.00	1.00
View	0.76	0.42	0	1	1	1	1
Click ($\times$100)	0.13	4.00	0	0.00	0.00	0.00	100

This table shows summary statistics for all relevant numeric variables in the dataset. The total number of auctions in the 1% random subsample for one week of the data is 762,728. Monetary variables are multiplied by a constant and stated in terms of a numeraire currency per thousand auctions (N$).

ence using combinations of the following settings: (i) *retargeting*, tracking users who have visited the advertiser's website in the past; (ii) *geography*, targeting consumers based on their geographical location; (iii) *device type*, selecting to show ads depending on the device the consumer is using - for example, show ads only on smartphones and tablets or only on computers; (iv) *operating system*, choosing consumers that use specific operating systems, such as iOS or Android; and (v) *inventory type*, showing ads to consumers that are using either apps or browsers.

Table 3 shows how representative each targeting setting is in the data. Retargeting is present in 81.8% of the 2,592 campaigns in the full dataset, but it accounts for only 2.3% of the auctions won in the 1% random sample of the fourth week of November 2021. This shows how granular retargeting can be and demonstrates that advertisers run multiple retargeting campaigns to reach specific audiences. Despite representing only 2% of the auctions won, retargeting accounts for almost 40% of the value spent in the auctions. The average bid for retargeting campaigns in my data is two orders of magnitude higher than for campaigns that do not use any targeting settings. These figures show how valuable retargeting is to advertisers, and motivate my focus on retargeting audiences. Geographical targeting is present in almost one-quarter of the auctions and accounts for 10% of the ad spend. Device type and inventory type are highly correlated targeting settings: if an adver-

Table 3: Representativeness of Each Targeting Setting in the Dataset for a 1% Random Sample from Nov 20th to Nov 26th 2021

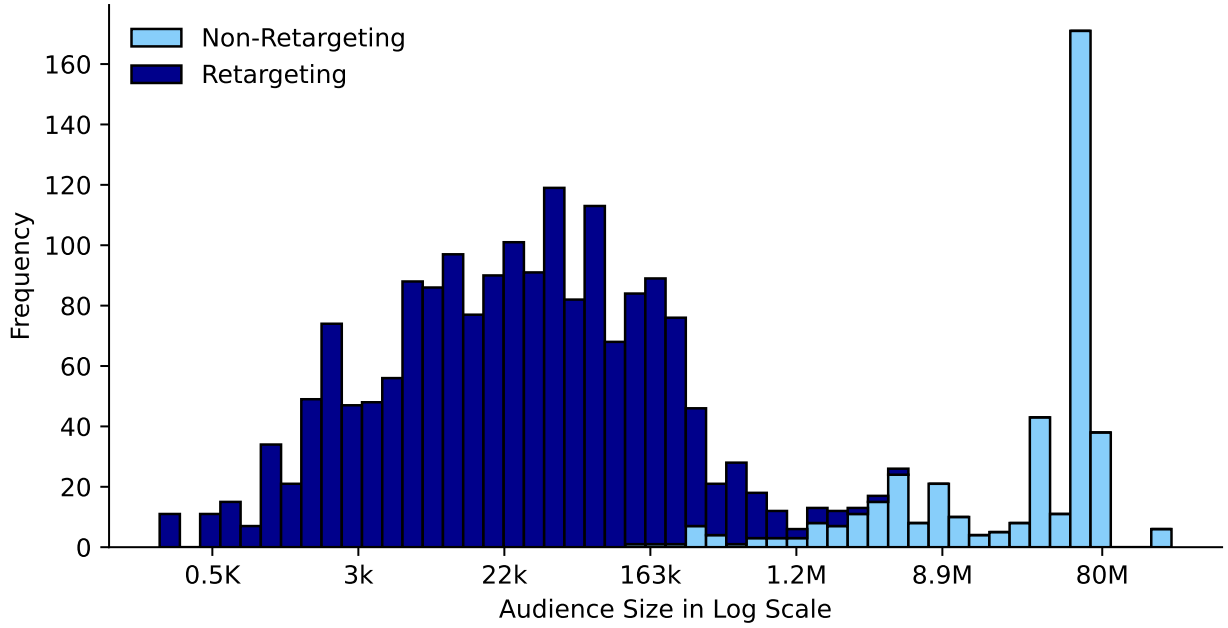
Targeting Setting	% Campaigns	% Auctions	% Value Spent
Retargeting	81.8	2.3	39.6
Geography	6.7	22.9	10.2
Device Type	68.1	99.2	73.1
Inventory Type	99.2	96.9	73.7
Operational System	1.4	1.8	0.6

This table shows how representative each targeting setting is in the dataset. Each row represents a dummy for whether campaigns contains the respective targeting setting. For example, 81.8% of all 2,592 campaigns involve retargeting. They represent 2.3% of the 762,728 auctions and account for 39.6% of the total value spent in the auctions.

tiser shows ads only for inventory in apps, it automatically restricts the device type to be smartphones. This causes the representativeness of those two settings to be similar. Both account for a large share of the value spent because they are also highly correlated with retargeting. In my dataset, all retargeting campaigns are run on browsers, either on desktops/laptops or mobile phones. Hence, retargeting is usually paired with either device type or inventory type targeting. Operational system targeting is niche in my dataset, accounting for only 1.8% of auctions won and only 0.6% of the value spent.

Another unique feature of my dataset is that I observe the audience size that advertisers reach with their campaigns. Figure 2 shows the distribution of audience size in log scale. There are two clear clusters of audience sizes in the data: one concentrated around 80M users, and another more heterogeneous around 100K users. The cluster around larger audience sizes consists of campaigns that do not use retargeting. The variation in audience size in this group is attributed mainly to advertisers selecting by user device type. Campaigns around 80M users target all smartphones in Brazil, whereas campaigns around 200M users (the last bin on the right) target all devices in Brazil. In contrast, the cluster of campaigns targeting approximately 100K users is composed mainly of advertisers using retargeting. These campaigns are more granular, as they track specific users who have previously visited the advertiser's website. Variation in audience size for these campaigns is mainly because of

Figure 2: Distribution of Audience Size in Log Scale



This figure shows the distribution of the audience each advertiser reaches with their campaigns. Audience size is displayed on the x-axis in log scale, and the frequency of each bin is displayed on the y-xis. Bins in dark blue represent retargeting campaigns, while bins in light blue represent non-retargeting campaigns.

two factors: the average traffic on the advertiser's website, and what part of the website the campaign is retargeting. For example, it is expected that campaigns that reach users who visited the homepage are larger than campaigns targeting only users who have reached the shopping cart. Non-retargeting campaigns with audience sizes below 80M but above 1M are mainly composed of combinations of targeting settings - for instance, all smartphones in a specific region of the country.

The major disadvantage of the auction-level dataset is the inability to observe losing auctions and the number of actual bidders. This information is typically used for identification and estimation of parameters that determine advertisers' entry into retargeting. To circumvent this, I have data on the winning probability of each ad campaign predicted by the auction planner. This is a tool given by the demand-side platform to advertisers so they can forecast the performance of their campaigns and determine which targeting strategies would work best. The auction planner allows advertisers to input their selected targeting

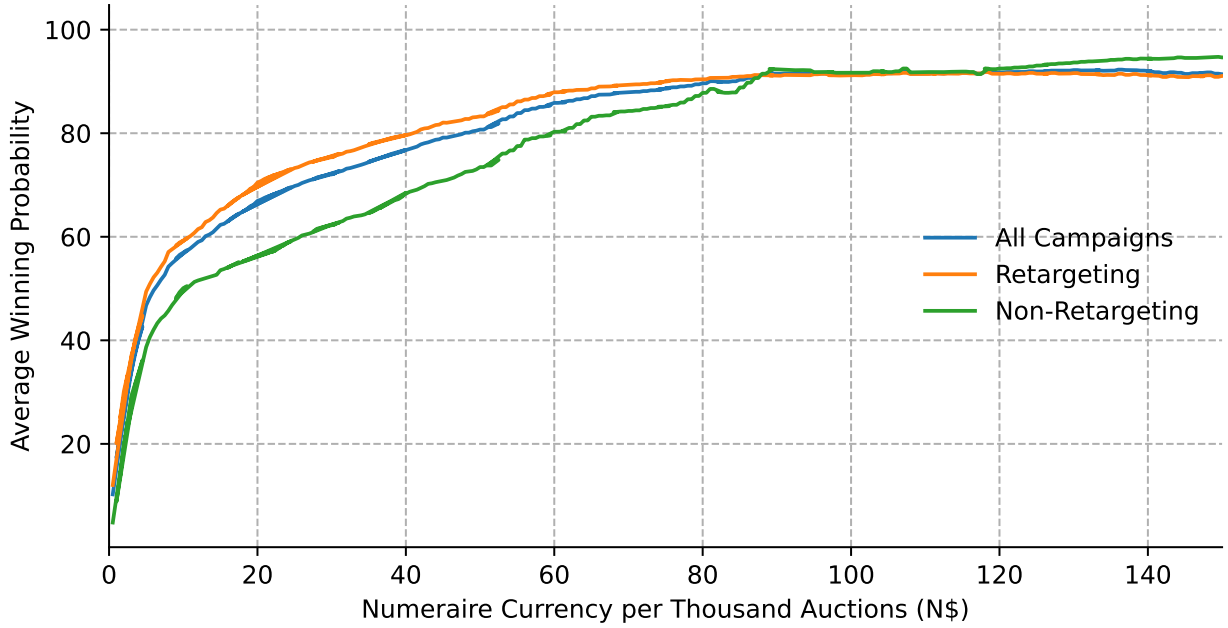
strategy along with their intended bid and, based on historical data collected by the demand-side platform, obtain an estimate of their winning probabilities in the auctions. I collected a discretized CDF of the winning probability distribution for each campaign. Figure 3 shows the average CDF across campaigns in my data as a function of the intended bid measured in a numeraire currency per thousand auctions.[14] I plot three lines: one with the average probability across all campaigns in the data, one for the retargeting campaigns alone, and another for the non-retargeting campaigns alone.

Two facts arise. First, recall that average bids are typically very low (Table 2), but bids from retargeting campaigns are two orders of magnitude higher than those from non-retargeting campaigns. This implies that non-retargeting campaigns reach a large audience by winning a small fraction of auctions, whereas retargeting campaigns reach a small audience by winning a large fraction of auctions. Second, for most of the left portion of the average CDFs, where bids are close to those observed in the data, winning probabilities for retargeting campaigns are above those for non-retargeting campaigns. This means that a given bid ensures, on average, a higher winning probability for retargeting than for non-retargeting campaigns. In my model, this is rationalized by fiercer competition faced in non-retargeting campaigns.

Retargeting is the strategy with the most substantial impact on advertisers' bids among the five targeting strategies examined in the data. Table 4 presents regression results on the logarithm of advertisers' bids using the auction-level dataset. Column 1 controls for dummy variables indicating the use of one of the five targeting settings, as well as for audience size. Note that those five categories are not mutually exclusive, i.e., campaigns can have more than one targeting strategy at the same time. For example, a campaign that does geographical and device-type targeting would have those two dummies turned on. Subsequent columns progressively introduce fixed effects for device type, inventory type and

[14] I multiply all monetary values in the dataset by a constant for confidentiality reasons, and state them in terms of a fictitious numeraire currency (N$).

Figure 3: Average Winning Probabilities Across Different Campaigns



This figure shows the average winning probability across different campaigns as a function of advertisers' bid. This data comes from the auction planner, a tool provided by the Demand-Side Platform to help advertisers plan their campaigns. On the x-axis I display the bids in a numeraire currency per thousand auctions (N$), and on the y-axis I display the corresponding average winning probability. The blue line is the average across all campaigns, the orange line is the average only for retargeting campaigns, and the green line is the average only for non-retargeting campaigns.

operating system, as well as for advertiser and publisher.[15] Across all specifications, retargeting is consistently associated with the largest increase in bids. In the most comprehensive specification (Column 4), retargeting is associated with bids that are 10.8 times higher than non-retargeting campaigns.[16] While geography and device type targeting are also associated with higher bids, their impact is relatively lower: geography targeting increases bids by 34%, while device type targeting increases bids by 90%. As anticipated, smaller audience sizes correlate with higher bids, with a 1% decrease in audience size corresponding to a 10.8%

[15]Fixed effects for device type differ from the dummy for device type targeting. The dummy merely captures that the campaign as a whole shows ads only for some specific device type. In contrast, the device type fixed effect is at the impression level, and controls for the specific type of device the consumer used. For example, a campaign that targets "mobile devices" would have 1 for the device type dummy, and the device type fixed effect would control for individuals using smartphones or tablets. The same logic applies to inventory type and operational system.

[16]Because the dependent variable is in logs, the effect of the independent variable is calculated by exponentiating the coefficient. In this case, $e^{2.38} = 10.8$.

increase in bids.

Considering the conclusions from Figure 2 and Table 4, the effect of retargeting on bids becomes even more evident. Retargeting campaigns are associated with significantly smaller audience sizes by design. The goal of these campaigns is to reach a specific group of consumers that visited the advertiser's website in the past. The coefficients from Table 4 indicate that the predicted bid for a retargeting campaign on mobile devices that reaches 100K consumers is N$5.92 whereas that for a non-retargeting campaign on mobile devices that reaches 80M consumers is N$0.26. These findings suggest that the heterogeneity in bids observed in the dataset cannot be solely attributed to a large dispersion of valuations across different bidders. When an advertiser chooses to employ retargeting, it is effectively attempting to buy a better match with consumers, which is embodied in the signal that the consumer visited the advertiser's website and is aware of its products. This provides the microfoundation I use to specify bidders' valuations in Section 4.

Retargeting is also the targeting setting with the most substantial impact on advertisers' winning probability in auctions. Table 5 presents regression results of winning probabilities retrieved from the auction planner on the five targeting settings, as well as on audience size and advertiser fixed effects.[17] In Column 1 audience size is not used as control; in Column 2 it is, but I remove the targeting dummy as control; and in Column 3, I control for both. Since retargeting is negatively correlated with audience size, not controlling for retargeting exacerbates the negative effect that a smaller audience size has on the winning probability. In the richer specification (Column 4), I use advertiser fixed effects. Retargeting is associated with a 23.6% increase in the winning probability of a campaign. This result remains stable across all specifications, showing that retargeting is an essential factor driving winning probabilities. Finally, a 1% decrease in audience size increases the winning probability by 5.4%. Those results taken together support the intuition from the discussion on thin markets in online advertising (Levin and Milgrom 2010, Chen

[17]The winning probability for each campaign is evaluated at the average bid observed for the campaign in the dataset.

Table 4: Regression on the Winning Bids in the Observed Dataset

Dependent Variable Log(Bid)	(1)	(2)	(3)
Intercept	-4.159***		
	(0.867)		
Retargeting	3.570***	4.335***	2.380***
	(0.265)	(0.421)	(0.332)
Geography	-0.062	-0.053	0.293**
	(0.174)	(0.161)	(0.096)
Device Type	0.577**	0.704***	0.646***
	(0.183)	(0.205)	(0.159)
Inventory Type	0.520	-0.230	-1.491***
	(0.294)	(0.412)	(0.398)
Operational System	-0.001	0.074	0.017
	(0.165)	(0.147)	(0.247)
Log(Audience Size)	-0.043	-0.050	-0.108**
	(0.046)	(0.031)	(0.036)
Fixed effects:			
Device Type		✓	✓
Inventory Type		✓	✓
Operational System		✓	✓
Advertiser			✓
Publisher			✓
R2	0.481	0.523	0.839
Cluster Level: Campaign			
Number of Auctions: 684,318			

This table shows results of three different regressions on the logarithm of the winning bids observed in the 1% random subsample of the dataset. The first five control variables represent dummies for whether the auction is part of a campaign that uses a specific targeting setting. The variable Log(Audience Size) controls for the number of potential consumers reached by the campaign. I gradually include auction-level fixed effects for each type of device, inventory and operational system, as well as fixed effects for advertisers and publishers. Standard errors are clustered at the campaign level. Significance levels: $* p < 0.05, ** p < 0.01, *** p < 0.001$.

Table 5: Regression on the Winning Probability in the Auction Planner Dataset

Dependent Variable Winning Probability	(1)	(2)	(3)	(4)
Intercept	0.102***	1.454***	0.796***	
	(0.026)	(0.048)	(0.121)	
Retargeting	0.582***		0.318***	0.236***
	(0.019)		(0.046)	(0.039)
Geography	0.028	-0.122**	-0.028	-0.132***
	(0.020)	(0.037)	(0.029)	(0.033)
Device Type	-0.091**	-0.113***	-0.089**	-0.050*
	(0.028)	(0.029)	(0.028)	(0.020)
Inventory Type	0.003	0.029	-0.009	-0.173**
	(0.012)	(0.022)	(0.012)	(0.064)
Operational System	0.004	-0.188***	-0.078***	-0.046
	(0.010)	(0.028)	(0.023)	(0.031)
Log(Audience Size)		-0.074***	-0.039***	-0.054***
		(0.003)	(0.007)	(0.005)
Fixed Effects:				
Advertiser				✓
R2	0.587	0.576	0.606	0.758
Number of Campaigns	1,807	1,739	1,739	1,739
Average Winning Probability (Retargeting): 62.8%				
Average Winning Probability (Non-Retargeting): 2.1%				
Cluster Level: Advertiser				

This table shows results of four different regressions on the campaign's winning probability, using campaign-level data collected from the auction planner. The first five control variables represent dummies for whether the campaign uses a specific targeting setting. The variable Log(Audience Size) controls for the number of potential consumers reached by the campaign. In the last column the regression includes fixed effects for advertisers. Standard errors are clustered at the advertiser level. Significance levels: $* \, p < 0.05, * * \; p < 0.01, * * * \, p < 0.001$.

and Stallaert 2014, Hummel and Mcafee 2016). When advertisers decrease the number of targeted consumers, they tend to face fewer competing bidders, thereby increasing their winning probability in the auction.

3.3 Experimental Variation on Bids

In partnership with The Company, I ran experiments by varying bids of 12 live campaigns from six active advertisers in the Brazilian market between February 22 and February 25 2022. For each day, I varied the bid advertisers submit to the demand-side platform, which represents the maximum amount the advertiser is willing to pay for winning 1,000 auctions.

The main goal of the experiment is to test whether these online ad auctions are more consistent with the assumption of independent private values (IPV) or common values (CV), as there are potential endogeneity issues when estimating advertisers' valuations under a wrong assumption. Under a CV setting, information about the competitors' valuation reveals information about its own valuation, whereas it does not under an IPV setting. Therefore, when a single bidder increases its bids, the marginal auctions it wins should give it a higher payoff than before, since those auctions were being closed at a higher price. This mechanism generates correlation between bids and the unobserved idiosyncratic shocks in advertisers' valuation, biasing estimates of its distribution if one mistakenly estimates them under an IPV setting.

For instance, consider advertisers running campaigns targeting both iPhone and Android users. iPhone users are generally more valuable to advertisers due to the positive correlation between iPhone adoption and higher income levels - some advertisers even create campaigns exclusively for iPhone users, paying a premium for these impressions. In a campaign that displays ads to both iPhone and Android users, an increase in bids could lead to winning impressions with higher value for the advertiser. This is because the more expensive impressions would likely be shown to iPhone users, as a result of the premium paid by advertisers targeting this specific audience. Consequently, the cheaper impressions would

be displayed to Android users, effectively segmenting the audience based on their perceived value for the impression.

However, under the IPV setting, the marginal auctions won after a unilateral increase in bids should not yield higher payoffs to the advertiser. I find that, for the case of retargeting, the auctions won after an increase in bids do not yield higher click-through rates, consistent with the IPV assumption. I use this conclusion in the model, where the IPV setting is assumed to hold.

The design of the experiment is as follows. For each of the six advertisers selected by The Company, two ad campaigns were selected: one that does retargeting, and one that does not - totalling 12 campaigns. I define the *base bid* as the bid submitted to the demand-side platform on Monday February 21, 2022. Then, I varied the submitted bid by a percent amount compared to the base bid. On Tuesday (Feb 22), I increase the submitted bid by 10%; on Wednesday, I do so by 50%; on Thursday, I decrease the submitted bid by 10% compared to the base bid; and on Friday I decrease it by 50%. Therefore, the treatment period is February 22 to February 25. For each treated day, I choose the control period to be the respective weekday of the previous week.[18] For example, I compare the treatment effects of increasing bids by 10% on Tuesday, February 22, to Tuesday, February 15. More details on the experiment design are in Appendix A.

Table 6 shows the effect of exogenous changes in bids on the number of auctions won, view rates and click-through rates. The coefficients reflect the aggregate effect of the experiment, which increased and decreased bids by 10% and 50% over four days. Units are standardized such that the coefficients show the effect on the dependent variable following an increase in bids by 10%. An exogenous change in bids does not show any significant impact for click-through rates. This conclusion supports the IPV assumption, as the marginal auctions won after a unilateral increase in bids do not seem to yield higher payoffs to the advertisers. There is a small but significant increase in view rates - but since advertisers can condition

[18]The base bid was stable over the previous week for all treated campaigns.

Table 6: Regressions Results from the Experimental Variation in Bids

Effect of a 10% Bid Increase on:	All Auctions (1)	Non-Retargeting (2)	Retargeting (3)
Log(Number of Auctions Won)	0.023	0.064**	-0.018
	(0.023)	(0.010)	(0.041)
View-Rate	-0.001	-0.005	0.004*
	(0.004)	(0.007)	(0.002)
Click-Through Rate	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Campaign Fixed Effects	✓	✓	✓
Weekday Fixed Effects	✓	✓	✓
Number of Campaigns	84	42	42
Cluster Level: Campaign			

This table shows the effect of exogenous changes in bids on the number of auctions won, view-rates and click-through-rates. Coefficients represent the change in the dependent variable after an increase in bids by 10%. All regressions include campaign and weekday fixed effects. Standard errors are clustered at the campaign level. Significance levels: $* \ p < 0.05, ** \ p < 0.01, *** \ p < 0.001$.

on the minimum historical view rate of publishers when creating their campaigns, this is not evidence of common values. In contrast, advertisers cannot condition on a guaranteed click-through rate. In Appendix A I perform a series of robustness checks, with alternative specifications and control groups to assess treatment effects. The main conclusion supporting the IPV assumption remains unchanged.

4 Model

In this section, I present the structural model employed in the paper. I build on the literature on auctions with an entry stage (Samuelson 1985, Levin and Smith 1994, Li and Zheng 2009, Athey et al. 2011, Roberts and Sweeting 2016). The model captures the two-stage decision process faced by advertisers in the context of personalized retargeted ad auctions. In the first stage, advertisers must decide whether to enter the segment of auctions for personalized ads, incurring a fixed managerial (or attention) cost to do so. This cost represents the resources required to monitor the performance of various advertising campaigns. In the second stage, those who choose to enter participate by submitting bids in second-price auctions with no reserve price.

The focus of the model is on advertisers' decision to either engage in retargeting or forgo targeting altogether. Although this simplification excludes other targeting strategies, such as geographic targeting, it preserves the core economics of the entry decision while reducing computational complexity. By concentrating on the fundamental choice between retargeting and no targeting, the model captures the essential tradeoffs faced by advertisers in the context of targeted online ads.

4.1 Advertiser Valuations

In my model there are N potential bidders indexed by i and J potential consumers indexed by j. I assume that the valuation of bidder i for showing an ad to potential consumer j is:

$$v_{ij} = \beta_r \mathbb{1}_{(j \in ret_i)} + \epsilon_{ij}, \quad \epsilon_{ij} \sim LogNormal(\mu_\epsilon, \sigma^2_{\epsilon,s}) \quad \text{iid} \quad \forall i, j, s \tag{1}$$

That is, valuations are composed of an idiosyncratic shock ϵ_{ij} and a retargeting premium β_r if potential consumer j is in the retargeting audience of advertiser i. I assume that the idiosyncratic shocks are private information to each bidder, known only at the time of

bidding, but their distribution is common knowledge. The specification in Equation 1 is motivated by the results of Section 3. Those reduced-form results demonstrate that auctions won by retargeting campaigns close at substantially higher prices than those won by non-retargeting campaigns. This price differential suggests that the observed gap cannot be solely attributed to valuation heterogeneity across bidders. Retargeting strategies provide advertisers with a highly informative signal, indicating that the targeted audience comprises the most valued consumers for the advertiser. By leveraging retargeting, advertisers can identify and focus their efforts on the subset of consumers who have previously expressed interest in their products or services, thereby justifying more competitive bidding behavior. Note that idiosyncratic shocks for advertisers in retargeting and non-retargeting are drawn from a lognormal with same mean but different variance parameters. I do not allow different mean parameters because retargeting valuations are already shifted by the retargeting premium, but their variances are flexibly different.

4.2 Auction Entry

Before bidding, advertisers must decide whether to enter the market for online ads. I allow them to enter auctions with a retargeting setting (indexed by r) and with a non-retargeting setting (indexed by nr).[19] Therefore, each advertiser can have 0, 1 or 2 campaigns running simultaneously. The decision to enter auctions under different targeting settings is made independently of each other, that is, advertisers' decisions are a binary outcome to enter each type of targeting setting. Participation in an auction with targeting setting $s \in \{r, nr\}$ requires paying a fixed cost κ_{ij}^s. This term represents an upfront managerial, or attention, resource required to monitor the performance of the auction. I assume that

$$\kappa_{ij}^s \sim LogNormal(\mu_\kappa, \sigma_\kappa^2) \quad \text{iid} \quad \forall i, j, s \tag{2}$$

[19]In the dataset, advertisers run on average 10 simultaneous retargeting campaigns. I abstract from this empirical regularity, as allowing for a higher number of simultaneous campaigns does not change the core economical tradeoffs captured in the model

This means that advertisers independently draw a value for the participation cost in each auction. I assume the participation costs are private information, known before the entry decision, but their distribution is common knowledge.

4.3 Retargeting Audiences

Each advertiser i is endowed with an audience size $J_i \leq J$. This represents the number of consumers who have visited advertiser i's website in the past. When advertisers have access to the third-party cookie technology, they can create a retargeting campaign to reach the exact group of J_i potential consumers who previously visited their website. Potential consumer j can belong to audiences of multiple advertisers at the same time.[20] I assume that J_i and J are common knowledge to all bidders. Both J_i and J are observed in the dataset.

4.4 Optimal Bidding

The optimal strategy in the bidding stage of the game is to bid one's expected valuation conditional on winning, as this is a second-price auction within the IPV setting (Krishna 2009):

$$b_{ij}(v_{ij}) = \mathbb{E}[v_{ij}|i \text{ wins}] = \beta_r \mathbb{E}[\mathbb{1}_{(j \in ret_i)}|i \text{ wins}] + \epsilon_{ij} \tag{3}$$

In particular, I define the nonrandom component of the bid as:

$$\bar{b}_{is} \equiv \beta_r \mathbb{E}[\mathbb{1}_{(j \in ret_i)}|i \text{ wins}] \tag{4}$$

In this way, I can rewrite the optimal bid in Equation 3 as

[20] I assume that this happens independently: I do not model any correlation between audiences of different advertisers. Therefore, when an advertiser i enters a non-retargeting campaign nr for a consumer j, the probability that consumer j belongs to advertiser i's audience J_i is $\frac{J_i}{J}$, whereas for a retargeting campaign this probability is 1.

$$b_{ij}(v_{ij}) = \bar{b}_{is} + \epsilon_{ij} \tag{5}$$

This notation will be important in Section 5, when I discuss estimation of the model. The nonrandom component of bids $\bar{b}_{is}$ is an equilibrium object, as it depends on the expectation that advertisers have about the chance of winning an auction for a consumer that belongs to their retargeting audience. This expectation depends on how other advertisers enter auctions, which is defined as an equilibrium in the first stage of the game.

4.5 Optimal Entry

In the entry stage, advertisers' optimal strategy is characterized by a cutoff threshold k^*_{is} for the entry costs. If the entry cost of advertiser i for participating an auction for potential consumer j with targeting setting s exceeds this threshold ($\kappa^s_{ij} > k^*_{is}$), than the advertiser does not enter the auction. This threshold is given by the value of the entry cost that makes advertiser i indifferent, in expectation, about entering an auction with targeting setting s. Two points about this cutoff strategy are noteworthy. First, k^*_{is} is advertiser and setting specific.[21] This means that, for a given advertiser with audience size J_i, there will be two thresholds determining whether i enters each targeting setting: one for retargeting ($k^*_{i,r}$) and one for non-retargeting ($k^*_{i,nr}$). Second, the entry decision is made without knowledge of the idiosyncratic shocks ϵ_{ij}. This implies that there is no selection on valuation for bidders that decide to participate in auctions. This timing closely matches how the online ad industry works. Advertisers must decide whether to create ad campaigns much before participating in auctions. At the time of this decision, they do not know their exact valuation for each of the millions of auctions in which they will participate.

In summary, the timing of this two-stage game is as follows. First, advertisers draw an entry cost κ^s_{ij} for both retargeting and non-retargeting campaigns. They compare

[21] This is a slight abuse of notation, as the entry strategy is not advertiser specific but rather audience-size specific. This means that if two advertisers have exactly the same audience size $J_a = J_b$, then their entry thresholds will be identical.

this value to the equilibrium thresholds k^*_{is} and decide whether to enter the auctions. Next, entrants draw their value for the idiosyncratic shock ϵ_{ij} and submit bids. Finally, the ad is allocated to the highest bidder, and payments are realized.[22][23]

4.6 Equilibrium Definition

Let $\pi_i(b_{ij}, \mathbf{k}^*_\mathbf{s})$ denote advertiser i's ex ante expected payoff of entering an auction using targeting setting s, given some bidding function b_{ij} and a vector of equilibrium entry cost thresholds $\mathbf{k}^*_\mathbf{s}$. An equilibrium in this two-stage game is composed of an entry rule $\sigma(\kappa, k^*_{is})$ and a bidding function $b(v, \mathbf{k}^*_\mathbf{s})$ such that:

1. The entry rule $\sigma(\kappa, k^*_{is}) : (\kappa, k^*_{is}) \rightarrow \{0, 1\}$ is given by an equilibrium threshold k^*_{is} such that $\sigma(\kappa, k^*_{is}) = 1$ if, and only if, $\kappa \leq k^*_{is},\ s \in \{r, nr\}$;

2. The equilibrium entry cost threshold k^*_{is} is given by the indifference condition

$$\pi_i(b_{ij}, \mathbf{k}^*_\mathbf{s}) = k^*_{is} \tag{6}$$

   for a given bidding function b_{ij}

3. The equilibrium bidding function $b(v, \mathbf{k}^*_\mathbf{s}) : v, \mathbf{k}^*_\mathbf{s} \rightarrow \mathbb{R}^+$ is given by Equation 3

4. The entry rule $\sigma(\kappa, k^*_{is})$ and the equilibrium entry threshold k^*_{is} are consistent with the ex ante expected payoffs generated by the bidding function $b(v, \mathbf{k}^*_\mathbf{s})$

5. The bidding function $b(v, \mathbf{k}^*_\mathbf{s})$ is consistent with the selection induced by the entry rule $\sigma(\kappa, k^*_{is})$

[22] The timing of this model closely resembles that in Levin and Smith (1994), where the entry decision is made without full knowledge of valuations. However, players in Levin and Smith (1994) use a mixed strategy, whereas in my model the strategy is deterministic conditional on the entry cost draw. This timing closely resembles model 2 in Li and Zheng (2009).

[23] One could also create another model where advertisers do not take two random draws, for ϵ and κ. Instead, entry would be contingent only on advertisers' idiosyncratic shocks, ϵ_{ij}, being higher than a threshold. I do not pursue this approach because this assumption requires advertisers to know beforehand the exact valuation of each impression. This is not a reasonable assumption, as individual valuations are determined on an impression basis, at the time of the auction, long after the entry decision has been taken.

5 Estimation

This section details the estimation procedure I use to recover the parameters of interest from the structural model. The parameters I estimate are $\Theta = \{\beta_r, \mu_\epsilon, \sigma^2_{\epsilon,r}, \sigma^2_{\epsilon,nr}, \mu_k, \sigma^2_k\}$, namely, the valuation premium for retargeting as well as the distribution of idiosyncratic shocks and entry costs. These parameters allow for the computation of the objects determining the equilibrium of the model: the optimal bidding policies from Equation 5 and the entry cost thresholds given by Equation 6.

Those objects are crucial in understanding the counterfactual effects of the removal of third-party cookies. When advertisers are no longer able to identify individual consumers online, two effects happen simultaneously to determine the final impact on their bids. The valuation advertisers have for each impression decreases, as the consumer-advertiser match is not as precise as before. This tends to decrease their bids. However, because of that, advertisers will likely change their targeting strategies, targeting more loosely targeted audiences. As a consequence, advertisers would participate in more auctions than before. This would increase competition in the auctions, which tends to increase advertisers' payments. Therefore, estimating advertisers' equilibrium entry policy in targeted segments is necessary to separate the matching effect from the equilibrium effect in a counterfactual world without third-party cookies.

5.1 Selected Data for Estimation

I use a subset of the data described in Section 3 to estimate the structural model. Since I focus on the decision to either employ retargeting or no targeting at all, I select a part of the dataset that allows me to make the cleanest possible comparison between retargeting and no-targeting campaigns.

In my dataset, all 2,121 retargeting campaigns are run on browsers, either on desktops/laptops or mobile devices. However, only 10.3% of these campaigns exclude mobile

devices, while 49.5% of them are run exclusively on mobile devices. For this reason, I only use campaigns that are run on mobile devices. I select mobile campaigns that either (i) only do retargeting; or (ii) do not target any other specific settings. This selection serves two purposes. First, I ensure that no other targeting settings confound the bids I use to estimate advertisers' valuations. Second, by requiring all campaigns to be run on the same device type, I standardize the maximum number of consumers reachable by the auctions. In the model, this number was J, the number of consumers. Since I observe the number of consumers in each advertiser's retargeted audience (J_i), I can compute the unconditional probability $\frac{J_i}{J}$ that a random auction for advertiser i in a non-retargeting setting is for a consumer in i's audience. The data show that, for mobile campaigns, $J = 80$ million consumers.

From the pool of mobile campaigns, I further use only those that appear for at least 7 days in the data. It is common practice in the online advertisement industry for advertisers to test various targeting strategies and keep those that they believe are profitable. By only using campaigns that are mature enough, I ensure that I do not consider experimental campaigns, which could still be in their initial learning stages. Those campaigns could have bidding behavior not compatible with an optimal bidding policy. Including these campaigns in the estimation could bias the results.

Finally, I only use second-price auctions. In my dataset, the auction format is a mix between first- and second-price auctions (Zeithammer 2019). This is decided at the moment of the impression, either by the publisher or by the exchange. I simplify the optimal bidding function in the structural model by only considering second price auctions.

5.2 Simulated Method of Moments

I use a Simulated Method of Moments (SMM) estimator to recover the parameter vector Θ. The basic intuition for the SMM is that I simulate a dataset using my model and match specific moments from the simulated data to the corresponding moments in the actual dataset. I find the set of parameters $\Theta = \{\beta_r, \mu_\epsilon, \sigma^2_{\epsilon,r}, \sigma^2_{\epsilon,nr}, \mu_k, \sigma^2_k\}$ that most closely approximate the

simulated moments to the true moments.

Formally, let M be a $k \times 1$ vector of moments from the true dataset, and $\widetilde{M}(\Theta)$ be the analogous vector of moments from the simulated dataset, which depends on the vector of model parameters. Let Ω be a $k \times k$ positive-definite weight matrix. Then, the SMM estimator is defined as:

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}} \left[M - \widetilde{M}(\Theta)\right]^\intercal \Omega \left[M - \widetilde{M}(\Theta)\right] \tag{7}$$

This paper uses an identity matrix for Ω. Nonetheless, the estimates are consistent regardless of the choice of weights. I use an Adaptive Nelder-Mead minimization algorithm in Python.

5.3 Moments Used in Estimation

There are two sets of moments I use to find the parameters of interest. The first set consists of the average and standard deviation of the first- and second-highest bids for retargeting and non-retargeting campaigns. This set comprises 8 moments. Table 7 shows the value of each moment in the selected data I use for estimation. Similar to the descriptive statistics of the full data, the average bids for retargeting campaigns are two orders of magnitude higher than those for non-retargeting campaigns. In particular, the average of the highest bid for retargeting is 100 times higher than that for non-retargeting, demonstrating how much advertisers value the informative signal embodied in retargeting audiences. Retargeting campaigns also show a substantially larger variance in bids, with a standard deviation of approximately N$1.98. In contrast, there is considerably less bid heterogeneity for non-retargeting campaigns, where the standard deviation of bids is approximately N$0.01.

The second set of moments I use in estimation is the winning probabilities conditional on bidding, taken from the auction planner. Those probabilities are computed at the average bid of each campaign in the dataset. For each simulated bidder in the algorithm, I

Table 7: First Set of Moments from the Data Used in Estimation

Moment	Value (N$)
Average highest non-retargeting bid	0.012
Average second-highest non-retargeting bid	0.011
Average highest retargeting bid	1.133
Average second-highest retargeting bid	0.699
Std. of highest non-retargeting bid	0.013
Std. of second-highest non-retargeting bid	0.009
Std. of highest retargeting bid	1.989
Std. of second-highest retargeting bid	1.137

This table shows the first 8 moments used in estimation. Those moments are computed using 12,672,902 auctions from 265 campaigns in the data, and then matched by the Simulated Method of Moments estimator.

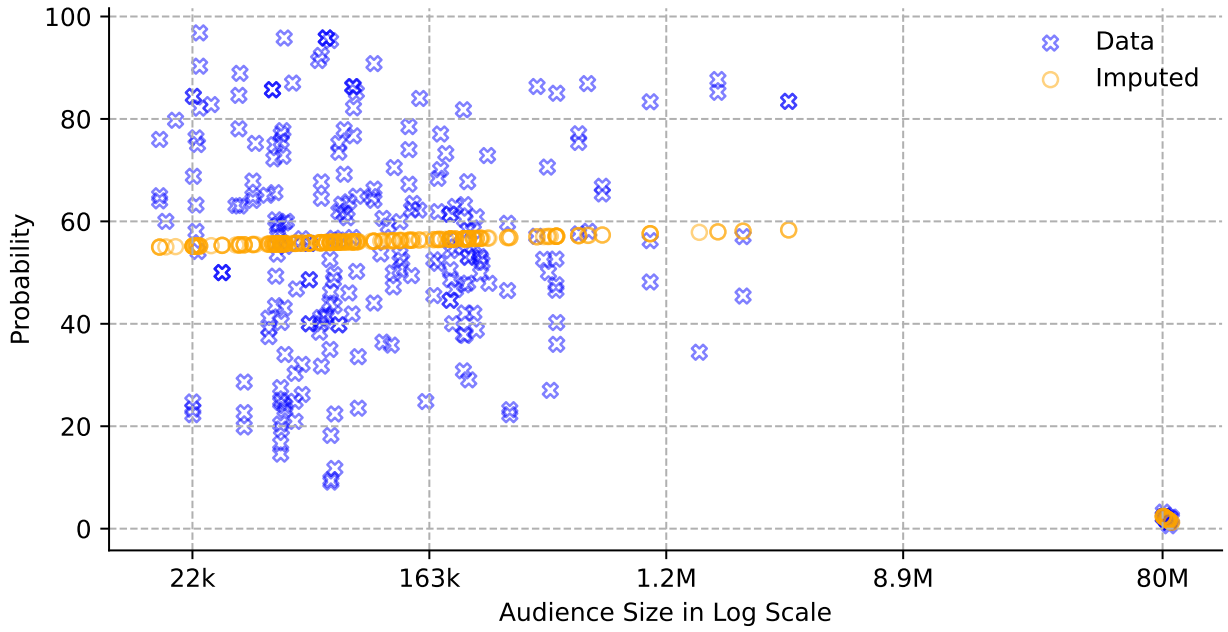
assign a predicted winning probability based on whether the bidder is adopting retargeting and on the bidder's audience size. Figure 4 shows how the imputation is done. The cluster of campaigns with an audience size of 80M are those non-retargeting campaigns that reach all mobile device users. The other cluster, with an audience size of approximately 100K, are the retargeting campaigns. For each simulated bidder in the dataset, I impute the predicted winning probability as a function of its audience size and whether it is adopting retargeting. Then, each simulated potential bidder is assigned the predicted winning probability from this projection.

5.4 Simulation Algorithm to Solve the Model

The equilibrium objects I solve for are the nonrandom component of bids $\bar{b}_{is}$ and the entry cost thresholds k^*_{is}. I use a nested fixed point algorithm to recover them.

I initialize the algorithm with a given set of parameters $\Theta = \{\beta_r, \mu_\epsilon, \sigma^2_\epsilon, \mu_k, \sigma^2_k\}$ and a number of potential bidders N. Then, for an initial guess of $\mathbf{k^*_{s,0}}$, I simulate a large number of auction opportunities. For each auction, bidders draw an entry cost κ_{is} and compare this value to their respective (guessed) entry threshold $k^*_{is,0}$ to make the entry decision. Then, each entrant draws its idiosyncratic shock ϵ_{ij} and realizes whether consumer

Figure 4: Imputed Winning Probabilities as a Function of Audience Size and Retargeting



This figure shows the campaigns' winning probability as a function of audience size. On the x-axis, it displays the size of the audience reached by each campaign in log scale. On the y-axis, it shows the winning probability associated with each campaign. Each blue cross marks a campaign's winning probability and audience size, collected from the auction planner. Each orange circle denotes a campaign's predicted winning probability using a simple OLS regression.

j belongs to its retargeting audience J_i. This happens with probability $\frac{J_i}{J}$ for each bidder. Next, each player bids and payoffs are realized.

I repeat the above procedure a sufficiently large number of times to compute the equilibrium nonrandom component of bids $\bar{b}^0_{is} \equiv \beta_r \mathbb{E}^0[\mathbb{1}_{(j \in ret_i)} | i \text{ wins}]$ (Equation 4). I repeat this computation until convergence of the nonrandom component of bids is achieved. This configures the inner loop of my algorithm. Note that this quantity depends on the guessed vector of entry thresholds $\mathbf{k}^*_{\mathbf{s},\mathbf{0}}$. With $\bar{b}^0_{is}$, I then update the equilibrium entry cost thresholds by computing each bidder's ex ante expected payoff $\pi^0_i(b^0_{ij}, \mathbf{k}^*_{\mathbf{s},\mathbf{0}}) = k^*_{is,0}$ (Equation 6). Finally, I repeat the updating of entry thresholds until it converges. This configures the outer loop of my algorithm, which recovers the equilibrium entry cost thresholds.

Ideally, the estimation of the structural model would be done with thousands of potential bidders. However, using a large number of potential entrants substantially increases

the computational complexity of the algorithm. Thus, the current version of this paper performs an approximation procedure based on the assumption that bidders are identical up to their audience size. The rationale for this approximation is the following.

Using campaigns that target mobile devices, the maximum number of consumers reached is $J = 80,000,000$. The average audience size for retargeting in the data is $J_i = 78,500$, which implies an average probability of 0.1% for a random consumer j of belonging to the audience of bidder i. If I were to simulate, for instance, 2,000 bidders each with a probability of 0.1% of facing a retargeted consumer, then the probability that a random consumer does not belong to the retargeted audience of any bidder is $0.999^{2000} \approx 13.5\%$. If instead I simulate $N = 10$ representative bidders, each with a probability of approximately 20% of bidding for a consumer in their retargeting audience, the probability that a random consumer does not belong to any bidder's audience is $0.8^{10} \approx 10.7\%$. By doing so, I substantially reduce the computational burden of solving the model. The only assumption needed for this is that bidders are identical up to their audience size. In Appendix B I perform robustness checks relaxing the approximation on the number of simulated bidders. The current version of this paper utilizes $N = 10$ for estimation.

6 Identifying Variation

This section discusses what variation in the data helps identify the key structural parameters of the model: retargeting valuation premium (β_r), mean and variance of idiosyncratic shocks ($\mu_\epsilon, \sigma^2_{\epsilon,r}$ and $\sigma^2_{\epsilon,nr}$), and mean and variance of entry costs (μ_k and σ^2_k). I argue that these parameters are identified from the variation in bids and winning probabilities across different types of advertisers and impressions, combined with the equilibrium restrictions implied by the model.

6.1 Valuation Parameters

First, the retargeting premium β_r is identified from the difference in average bids between retargeting and non-retargeting campaigns. The model assumes that valuations include an additive term β_r for retargeted impressions. Therefore, the average difference in bids between these two types of targeting settings identifies the premium that advertisers place on retargeting.

Second, the mean and variances of idiosyncratic shocks, $\mu_\epsilon, \sigma^2_{\epsilon,r}$ and $\sigma^2_{\epsilon,nr}$, are identified from the mean and variance of bids within retargeting and non-retargeting campaigns, after controlling for the retargeting premium β_r. Intuitively, μ_ϵ determines the average level of bids, while $\sigma^2_{\epsilon,r}$ and $\sigma^2_{\epsilon,nr}$ govern the dispersion of bids around this mean. By matching the observed bid moments within targeting setting, the model can separate the contribution of the idiosyncratic shocks from the retargeting premium.

6.2 Entry Cost Parameters

The identification of entry cost parameters, μ_ϵ and σ^2_ϵ, requires more careful consideration. In auctions with selected entry, identification of entry costs typically relies on observing either a measure of entry costs or the number of actual and potential bidders in each auction (Roberts and Sweeting 2010, Athey et al. 2011, Roberts and Sweeting 2016). However, nei-

ther is observed in the dataset used in this paper. Instead, the winning probabilities of each campaign are available. Consequently, the identification of entry costs in this case involves linking variation in the unobserved numbers of bidders, driven by differing parameters for the entry cost distribution, to variations in observed winning probabilities.

To illustrate the relationship between the entry cost parameters and key outcome variables of the model, I simulate datasets using the estimated parameters from Table 8. While the number of actual bidders is not observed in the data, they are observed in the simulated data. This allows me to trace the effect of varying the mean and standard deviations parameters from the (unobserved) number of actual bidders to the (observed) winning probabilities.

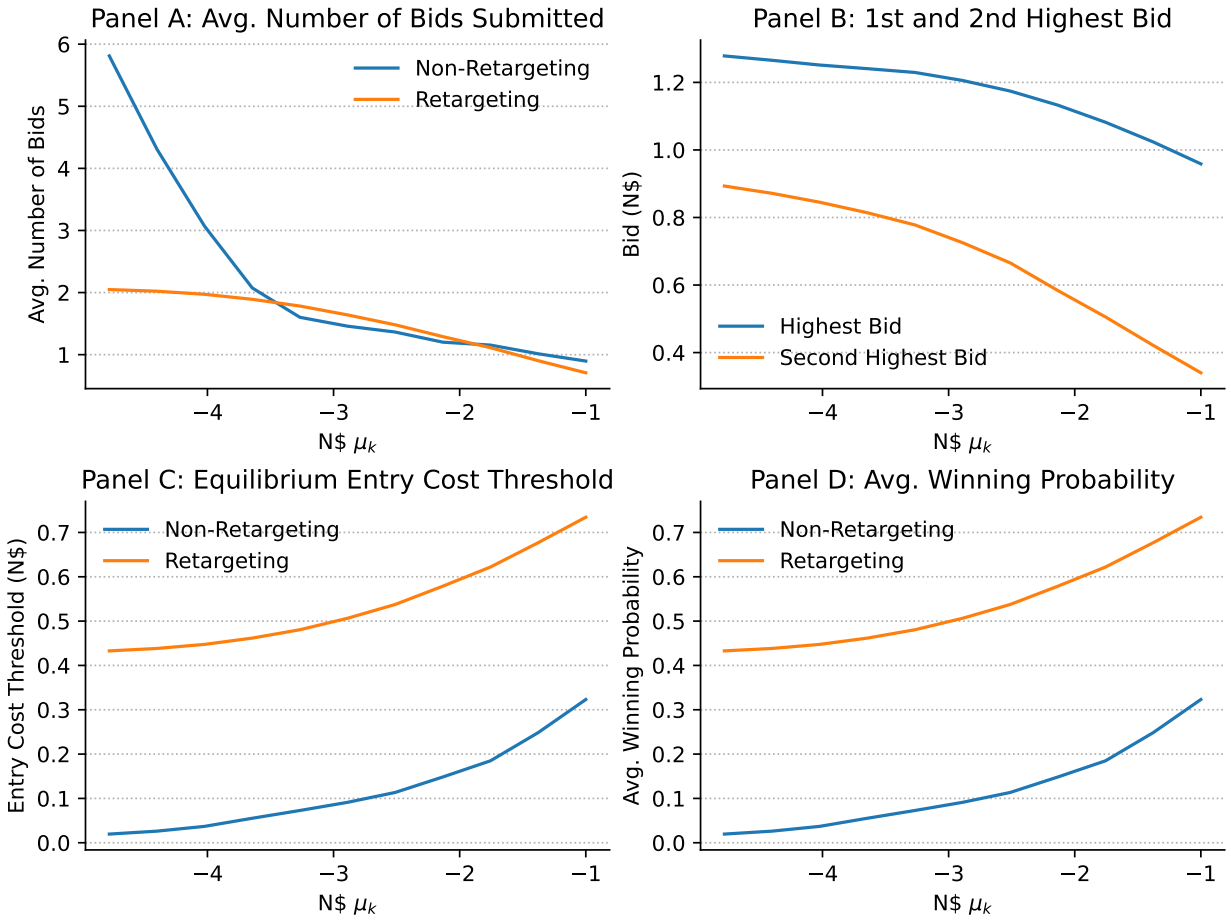
Figures 5 and 6 show simulation results for four quantities in the model: number of actual bidders (Panel A); average highest and second-highest bids (Panel B); equilibrium entry cost thresholds (Panel C); and winning probability conditional on entry (Panel D). A higher mean parameter of the entry cost distribution has an unambiguously negative effect on the number of bids submitted. Advertisers running both retargeting and non-retargeting campaigns tend to participate less frequently in auctions, as their entry costs are generally increasing. With fewer participants, the average highest and second-highest bids in the auctions tend to decrease. This leads to two consequences for advertisers: first, the surplus in the auctions they win tend to increase, which in turn increases their equilibrium entry cost thresholds; second, their winning winning probabilities tend to increase, as they face lower competing bids. Therefore, higher mean parameters for the entry cost distribution exhibit a positive correlation with the observed winning probabilities in the dataset used in this paper.

The effect of increasing the parameters for the standard deviation is shown in Figure 6. Contrary to what is observed when one varies the mean parameter of the entry cost distribution, a higher variance parameter has an ambiguous effect on the number of bids submitted. This happens because, as one increases the variance of the entry cost

distribution, some advertisers are more likely to receive very low entry cost draws, which would have previously been prohibitive of entry. Those tend to be the advertisers that would only enter under very low entry costs, which are typically advertisers running non-retargeting campaigns. Because of this ambiguous effect on participation, a higher variance parameter of the entry cost distribution also leads to an ambiguous effect on winning probabilities.

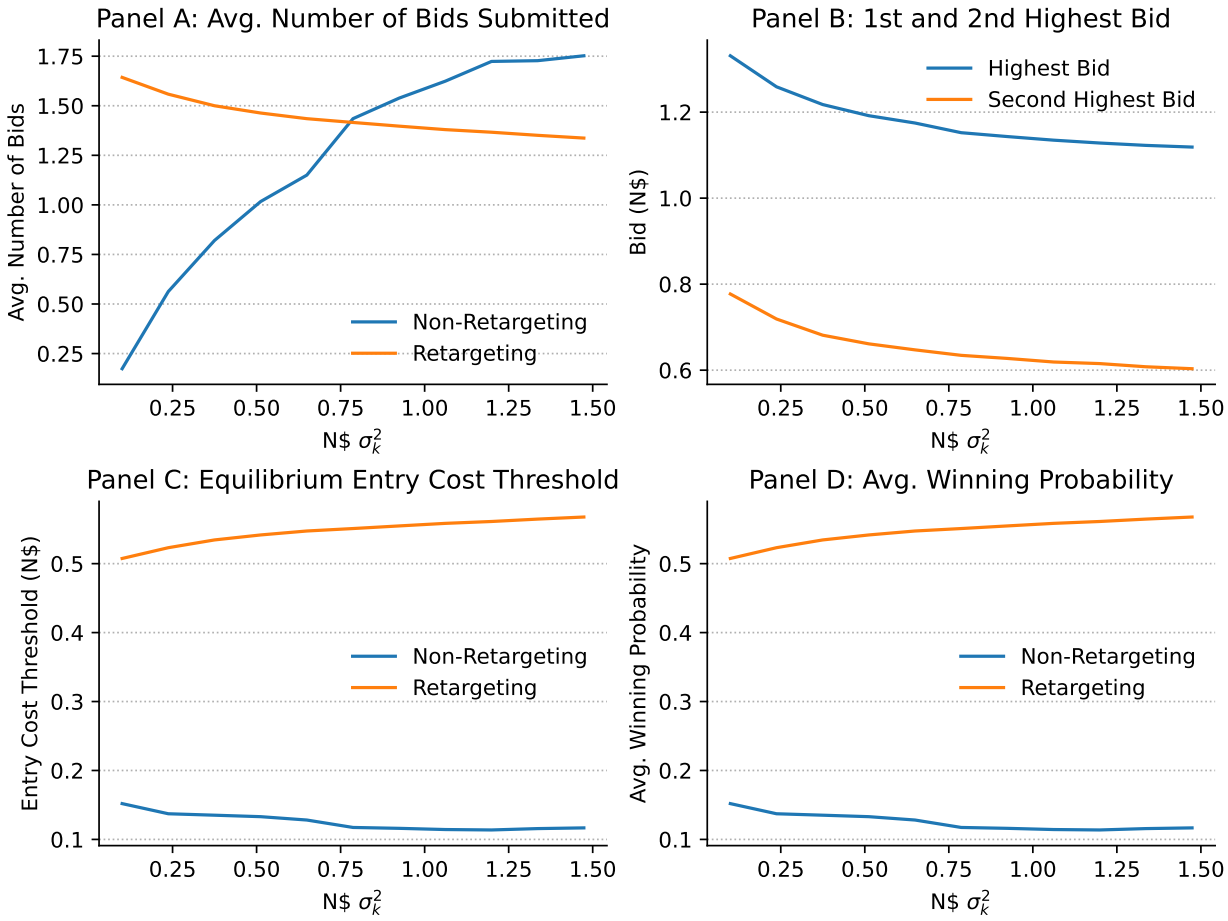
These simulation exercises establish a link between the effects of different parameters for the entry cost distribution on the unobserved number of bidders and the observed winning probabilities. A higher mean parameter has a negative effect on the number of bids submitted, leading to increased equilibrium entry cost thresholds and winning probabilities for advertisers. In contrast, a higher variance parameter has an ambiguous effect on participation and winning probabilities due to the increased likelihood of some advertisers receiving very low entry cost draws. Therefore, separate identification of the mean and variance of the entry cost distribution is possible by connecting the variations in unobserved bidder numbers to the variations in observed winning probabilities.

Figure 5: Simulations Results of Varying Mean Parameter of Entry Cost Distribution



This figure shows simulation results that indicate the effect of varying the mean parameter of the entry cost distribution on four key quantities of the model: average number of submitted bids (Panel A); highest and second highest bid (Panel B); equilibrium entry cost threshold (Panel C); and average winning probability conditional on entry (Panel D). The mean parameter of the entry cost distribution, on the x-axis of each panel, is multiplied by a constant for confidentiality and displayed in a numeraire currency per thousand auctions.

Figure 6: Simulations Results of Varying Variance Parameter of Entry Cost Distribution



This figure shows simulation results that indicate the effect of varying the variance parameter of the entry cost distribution on four key quantities of the model: average number of submitted bids (Panel A); highest and second highest bid (Panel B); equilibrium entry cost threshold (Panel C); and average winning probability conditional on entry (Panel D). The variance parameter of the entry cost distribution, on the x-axis of each panel, is multiplied by a constant for confidentiality and displayed in a numeraire currency per thousand auctions.

7 Results

In this section, I report the results of the SMM estimation. First, I present the estimated parameters that compose the model. Next, I show the equilibrium entry cost threshold policy used by advertisers in this market. This object is crucial in separating the matching from competition effects in counterfactuals.

7.1 Parameter Estimates and Moment Comparison

I depict the vector of estimated parameters in Table 8, multiplied by a constant for confidentiality reasons.[24] The parameter for the retargeting premium, β_r, is N$1.30. This represents the monetary difference in valuations between non-retargeting and retargeting. The estimated distribution for idiosyncratic shocks is Lognormal(-4.51, 0.76) for non-retargeting bidders and Lognormal(-4.51, 9.87) for retargeting bidders, while the entry costs distribution is Lognormal(-3.39, 1.25). This raises two observations. First, the model implies that the average bid for a non-retargeting advertiser is N$0.08 while that for a retargeting bidder is N$1.29. This is in line with the summary statistics presented in Section 3, where the winning bids for retargeting campaigns are two orders of magnitude higher than those for non-retargeting campaigns. This demonstrates that the model is able to capture the large premium advertisers derive from retargeting. Second, the average entry cost drawn by advertisers is N$0.16. This represents approximately 12% of the typical retargeting bid, but as much as double the typical non-retargeting bid. This is also in line with the empirical regularity that advertisers simultaneously run various retargeting campaigns. Because of the large retargeting premium in their valuations, advertisers generally derive a larger surplus in retargeting campaigns. Therefore, they can afford to pay higher entry costs to run those campaigns.

Figure 7 compares the first set of simulated moments against their counterpart in the true dataset. The model fits those moments well - both the level of bids as well

[24] Standard errors will be computed in a future version of this draft.

Table 8: Estimated Parameters

Parameter	Interpretation	Estimate (N$)
β_{ret}	Valuation premium for retargeting	1.29
μ_k	Mean parameter for entry cost distribution	-4.51
μ_ϵ	Mean parameter for idiosyncratic shock distribution	-8.18
σ_k^2	Variance parameter for entry cost distribution	5.15
$\sigma_{\epsilon,nr}^2$	Variance parameter for idiosyncratic non-retargeting	0.76
$\sigma_{\epsilon,r}^2$	Variance parameter for idiosyncratic retargeting	9.87

This table shows the parameters estimated in the model, their interpretation and estimated value in a numeraire currency per thousand auctions (N$).

Table 9: Winning Probability Conditional on Bidding

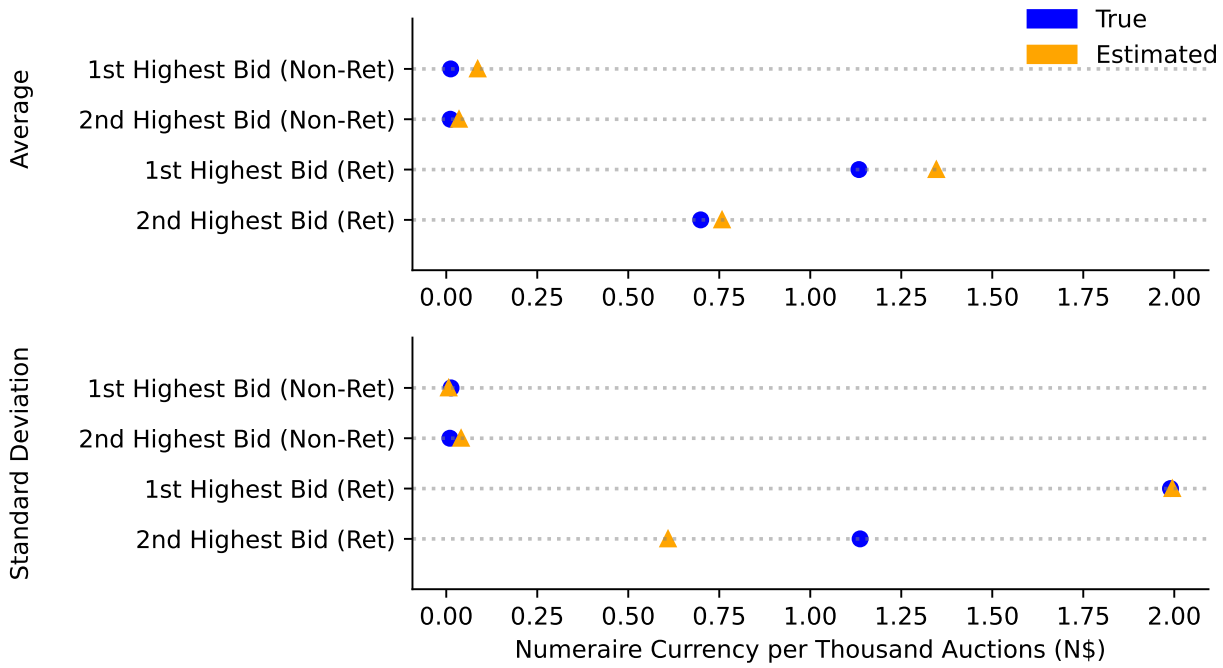
Equilibrium Object	Data	Model
Average winning probability for retargeting (%)	55.7	53.5
Average winning probability for non-retargeting (%)	1.9	11.0

This table compares the average winning probability conditional on bidding observed in the data to the one predicted by the model, in each targeting setting.

as their variation. For example, the average observed second highest bid for retargeting is N$0.70, and the model simulates an average of N$0.75. The observed average highest bid for non-retargeting is N$0.012 and the model simulates it at N$0.086.

Table 9 compares the average winning probabilities observed in the data to the averages predicted by the model for each targeting setting. The model is able to fit the winning probabilities for retargeting well, but it overestimates the winning probability for non-retargeting bidders. The average winning probability imputed from the data for the simulated potential bidders in retargeting is 55.7%, and the average predicted by the model is 53.5%. However, for non-retargeting campaigns the imputed mean is 1.9% while that simulated by the model is 11.0%.

Figure 7: True and Estimated Moments - Bids



This figure compares the true set of bid moments used in estimation to their simulated counterparts. The value of moments is displayed on the x-axis in a numeraire currency per thousand auctions (N$). The first subplot shows moments pertaining to the average for highest and second highest bids of retargeting and non-retargeting campaigns, while the last subplot shows the moments for standard deviations. Blue dots represent the true moments, and orange triangles represent the simulated moments.

7.2 Equilibrium Entry Cost Threshold

The equilibrium entry cost thresholds are computed as the value of the entry cost that would make bidders indifferent about entering an auction (Equation 6). Table 10 shows the average estimated thresholds for simulated advertisers in each targeting setting.

Advertisers that decide to employ retargeting exhibit higher entry cost thresholds than those not participating in retargeting. Specifically, the average entry threshold for the retargeting setting is N$0.064, whereas for non-retargeting it is N$0.002. This disparity arises because retargeting provides greater surplus for bidders. Advertisers running retargeting campaigns are assured that their ads will be shown to a relevant consumer if they win the auction. The high information value of this targeted audience justifies the higher entry cost thresholds for advertisers.

Table 10: Equilibrium Entry Cost Thresholds

Equilibrium Object	Value (N$)
Average equilibrium entry cost threshold for retargeting	0.064
Average equilibrium entry cost threshold for non-retargeting	0.002

This table shows the average equilibrium entry cost thresholds predicted by the model in each targeting setting. Values are multiplied by a constant and displayed in a numeraire currency per thousand auctions (N$).

Conversely, in non-retargeting campaigns, there are two factors influencing advertisers' low entry cost thresholds. First, their average valuation tends to be lower than in retargeting campaigns, as they only have a probabilistic chance of reaching a valuable consumer. As a result, they are inclined to submit lower bids and are unable to justify high entry costs for these auctions. Second, advertisers running non-retargeting campaigns often compete against other advertisers that are running retargeting campaigns. Facing these high valuation bidders tends to squeeze their surplus in the auctions: if they win against an advertiser that employs retargeting, this means that their payment in the auction tends to be higher than that if there were no competition from retargeting bidders. With lower margins, advertisers running non-retargeting campaigns cannot afford to pay high entry costs.

8 Counterfactual Exercises

This section presents the results of two counterfactual scenarios. First, I run an exercise where cookies are banned, following Google's plans to deprecate third-party cookies on Chrome. Without third-party cookies, it is not possible to track individual consumers online, which hinders advertisers' ability to target specific audiences. The targeting strategy that is most affected by the removal of cookies is retargeting, as it relies on identifying, across different sites, specific consumers who visited the advertiser's website in the past. Next, instead of banning cookies altogether, I run a counterfactual exercise where the entry cost of retargeting is doubled. This is motivated by two possible scenarios. For example, regulatory agencies could impose a "cookie tax" in an attempt to increase consumers' online privacy. In the model, only retargeting campaigns would be affected by a cookie tax. This would change advertisers' entry rate in equilibrium, altering both the value of the match they have with consumers and the competition they face in auctions.

Another motivation for the cookie tax is if advertisers have to pay high adjustment costs to adopt Google's proposal to replace third-party cookies, called Privacy Sandbox.[25] The primary goal of the Privacy Sandbox is to develop new technologies that can replace third-party cookies, particularly with respect to delivering personalized targeted ads. One of the proposals contained in the Sandbox is the Protected Audiences API, which could be a privacy-friendly alternative for retargeting.

The rollout of the Privacy Sandbox has prompted concerns from both advertisers and regulators (Johnson and Neumann 2024). According to a poll conducted by EMarketer with 204 U.S. advertising and marketing professionals in April 2024, approximately 50% of respondents affirm they do not intend to use the Privacy Sandbox at this time.[26] One of the possible reasons for such initial negative acceptance by the industry is the fact that using the Privacy Sandbox would require advertisers to substitute the tracking technology

[25] `https://privacysandbox.com/` (Archive.org, accessed on 06/21/2024)

[26] Report "The cookie-loss deadline: Three steps to take before the trackers go away", April 2024, available on request.

and digital infrastructure already in place, incurring high costs for the adoption of this alternative privacy-friendly solution[27].

Concomitantly, regulators have reservations about the rollout of the Privacy Sandbox. The Competition and Markets Authority (CMA) is the regulatory body in charge of overseeing the release of the Privacy Sandbox. Its main concern about Google's solution regards the possibility that Google would strengthen its already dominant position in the online ad industry. An illustration of this concern is stated on page 6 of their "Q1 2024 update report on the on implementation of the Privacy Sandbox commitment", released on April 2024.[28] It reports the following:

> *"In addition, we set out in our January 2024 report several key concerns that Google will need to resolve ahead of third-party cookie deprecation:*
>
> *(a) Ensuring that Google does not design, develop or use the Privacy Sandbox tools in ways that reinforce the existing market position of its advertising products and services (...)"*

8.1 Counterfactual A: Banning Third-Party Cookies

In the model, banning third-party cookies means that advertisers cannot choose to do retargeting, and are restricted to either run non-retargeting campaigns or to leave the market. I run the estimation simulation algorithm to solve the model described in Section 5.4 with the only change being that entry costs for retargeting are prohibitively high, that is, I set $\kappa_{ij}^{r} \approx \infty$ but do not change κ_{ij}^{nr}. The results for the new entry cost thresholds appear in Table 11.

There are two noteworthy facts about this result. First, the new entry cost thresholds for advertisers seeking retargeting is zero. This is a natural result in light of

[27] `https://www.aol.com/news/googles-privacy-sandbox-adoption-costs-100300733.html` (Archive.org, accessed on 09/09/2024)

[28] See their Q1 2024 Quarterly Report on Google's case at `https://assets.publishing.service.gov.uk/media/662baa3efee48e2ee6b81eb1/1._CMA_Q1_2024_update_report_on_Google_Privacy_Sandbox_commitments.pdf` (Archive.org, accessed on 06/21/2024)

Table 11: Equilibrium Entry Cost Thresholds - Cookies Banned

Equilibrium Object	Value (N$) Cookies Available	Cookies Banned
Avg. eqm. entry cost threshold for retargeting	0.064	0.000
Avg. eqm. entry cost threshold for non-retargeting	0.002	0.008

This table compares the average equilibrium entry cost thresholds predicted by the model in each targeting setting, in a counterfactual scenario where cookies are banned to the ones when cookies are available. Values are multiplied by a constant and displayed in a numeraire currency per thousand auctions (N$).

Equation 6. Because the cost of retargeting is prohibitively high under this counterfactual scenario, advertisers' ex ante expected payoff of doing so is zero, as there is no payoff large enough to justify participating these auctions.

Second, the new equilibrium entry cost thresholds for advertisers in non-retargeting campaigns is increased. The intuition for this result is as follows. Before the removal of cookies, advertisers participating in non-retargeting campaigns would frequently compete against other advertisers that used retargeting. However, those in retargeting campaigns hold a substantially higher valuation for the same impression - this premium is captured by β_r in Equation 1. Therefore, on average, the surplus enjoyed by advertisers in non-retargeting campaigns is typically small, evidenced by the average equilibrium entry cost threshold at N$ 0.002. When third-party cookies are removed, advertisers in non-retargeting campaigns no longer compete against those high-valuation advertisers adopting retargeting. Therefore, their surplus when winning auctions is not as small as before, as they only need to outbid advertisers that have a similar valuation as themselves. This allows advertisers in non-retargeting campaigns more room to pay higher entry costs, thereby increasing the equilibrium entry cost thresholds for non-retargeting in the counterfactual.

8.2 Disentangling Match from Competition

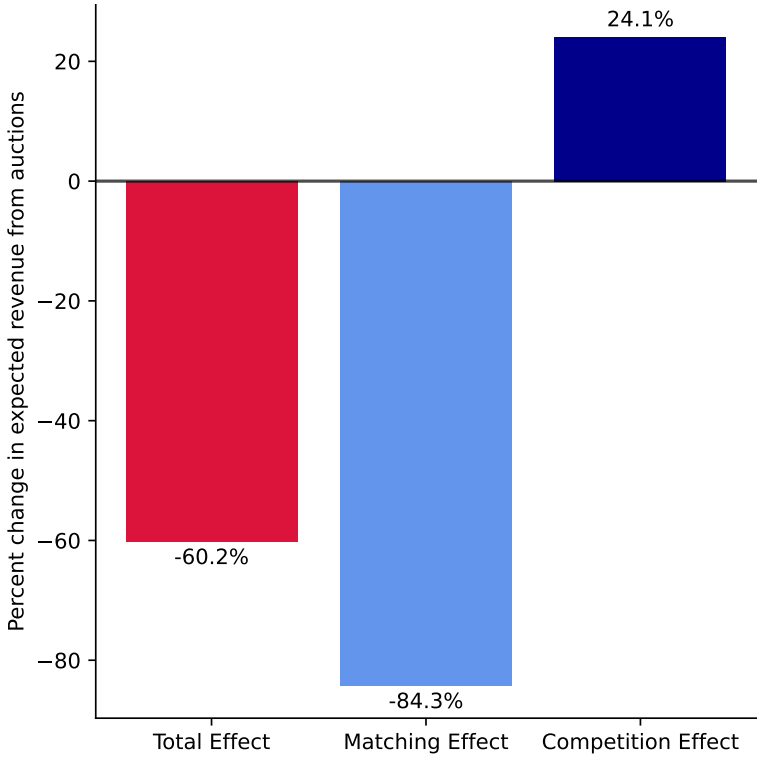
In light of the results of the previous subsection, it is clear that competition plays a role in the determination of advertisers' final bids following the removal of third-party cookies.

By estimating advertisers' valuations and equilibrium entry policy, I am able to separately quantify the effects of matching from competition on advertiser' final bids. This is particularly important for policymakers and regulators analyzing the impacts of the removal of third-party cookies on the industry.

Separately quantifying those two opposing forces is not trivial because they act simultaneously in equilibrium. Therefore, I disentangle them by computing the change in ad auction revenue in two steps. First, I compute the revenue under the *status quo*, where advertisers have access to third-party cookies and can employ retargeting. Next, I compute the ad auction revenue in an intermediate scenario where advertisers can no longer adopt retargeting, but they have not yet adjusted their equilibrium entry policies. This means that the only change between this intermediate scenario and the *status quo* is their lack of access to retargeting - their entry policies, and therefore the competition in auctions, remains the same. The difference between this intermediate scenario and the *status quo* is defined as the matching effect. Finally, I compute the ad auction revenue under the counterfactual scenario where advertisers do not have access to retargeting and have adjusted their entry policies. The difference in ad auction revenue between the counterfactual and the intermediate scenario is the competition effect: in those two scenarios, advertisers do not have access to retargeting, but their entry policies vary. The total effect is, by construction, the sum of the matching and competition effects.

Figure 8 shows the separate effects of matching and competition on ad auction revenue following the removal of third-party cookies. The model predicts that the revenue from ad auctions would decrease by 60.2%. This is in line with the existing literature on data privacy regulations studying the deprecation of cookies. For example, Johnson, Shriver, et al. (2020) find that auctions for users who opt out of cookies close at a price 52% lower than for opt-in users, and Alcobendas et al. (2021) find a reduction in publisher revenue of 54% when cookies are removed. Google also ran an internal experiment by disabling cookies for a randomized set of publishers and found that the ad auction revenue received by those

Figure 8: The Separate Effects of Match and Competition on Ad Auction Revenue Following a Ban on Third-Party Cookies



This figure shows the decomposition of the final effect of the removal of third-party cookies on bids. Bars represent the percent change in auction revenue, displayed on the y-axis. The red bar depicts the total effect of third-party cookies removal. The effect attributed to a lower match between advertisers and consumers is in light blue, and the effect resulting from tougher competition in auctions is in dark blue.

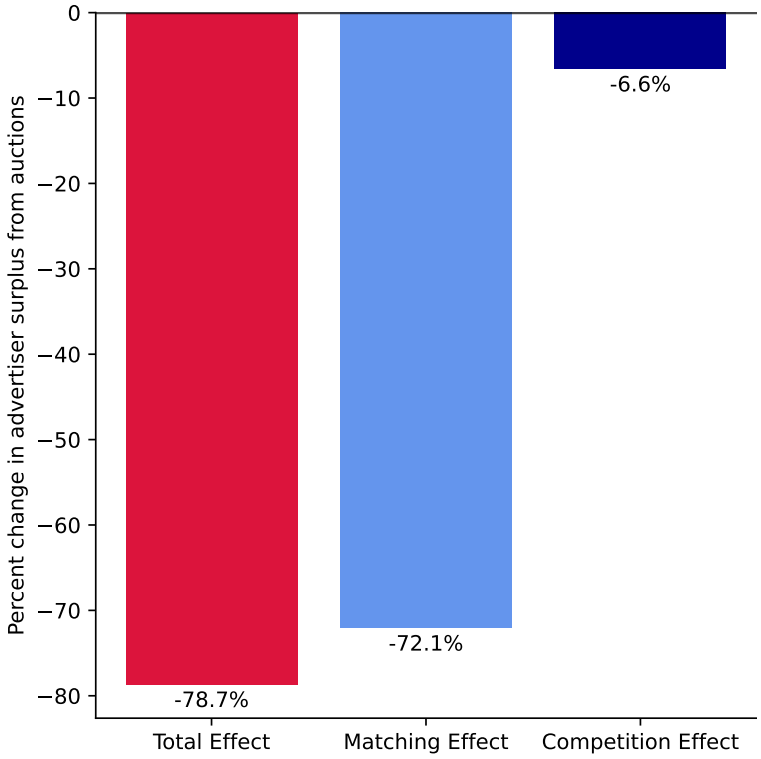
publishers decreased by 52% (Ravichandran and Korula 2019).

I extend this analysis by separately quantifying the importance of the matching and competition effects for the final change in revenue. In particular, the model predicts that if one only considers the lower match between advertisers and consumers resulting from the removal of cookies, ad auction revenue would decrease by 84.3%. However, this decrease in revenue is partially offset by a 24.1% increase in revenue because of fiercer competition. This raises two conclusions. First, publishers and platforms would be severely harmed by such a large decrease in revenue from ad auctions. Usually, the revenue from auctions is proportionally split between the publisher and the ad exchange. Therefore, the 60% decline in expected revenue predicted by the model would be translated into a 60% lower revenue

for publishers and platforms. Second, the decomposition of this number into a matching and competition effect shows that advertisers would be severely harmed if they no longer had access to retargeting. Advertisers would experience a large decrease in the match value they derive from this market and pay substantially more for non-retargeting impressions as a result of fiercer competition.

Figure 9 shows how the decomposition of the matching and competition effects applies to the change in advertiser welfare. I define advertiser welfare as the difference between their valuation derived from the auction and the sum of payments and entry costs. For instance, an advertiser that entered the auction but lost has a negative surplus represented by the sunk entry cost paid for participation. Contrary to what happens to ad auction revenue, in this case, both the matching and competition effects act to decrease advertiser welfare. The matching effect represents the lost valuation derived from auctions where advertisers could target the best possible audience for their products. Without third-party cookies, this is no longer possible. As a consequence, winners obtain a lower average surplus due to a lower match with the consumer to whom they show their ad. In addition, the surplus derived by advertisers from auctions is reduced by the fact that now they face a higher number of bidders: without the fierce competition from high-valuation retargeting bidders, advertisers conducting non-retargeting campaigns have more incentives to enter the market. This causes the second-highest bid in the auctions to increase, which further reduced advertisers surplus. In equilibrium, advertisers lose 78.7% of their surplus from the removal of third-party cookies. This effect consists of a 72.1% reduction due to a lower match with consumers and a further 6.6% surplus reduction due to fiercer competition in auctions. This means that the competition effect represents 8.3% of the total decrease in advertiser surplus from the removal of third-party cookies. This is a sizeable number that should not be disregarded by advertisers, platforms and policymakers that implement new data privacy regulations.

Figure 9: The Separate Effects of Match and Competition on Advertiser Welfare Following a Ban on Third-Party Cookies



This figure shows the decomposition of the final effect of the removal of third-party cookies on advertiser surplus, defined as the ex-post valuation derived from the auction (if win) subtracted by the payment (if win) and the entry costs (if participate). Bars represent the percent change in advertiser surplus, displayed on the y-axis. The red bar depicts the total effect of third-party cookies removal. The effect attributed to a lower match between advertisers and consumers is in light blue, and the effect resulting from tougher competition in auctions is in dark blue.

8.3 Counterfactual B: Taxing Third-Party Cookies

In this counterfactual exercise, rather than banning cookies I increase the entry cost for retargeting by 100%, that is, I set $\kappa_{ij}^{r,tax} = 2\kappa_{ij}^{r}$. This is motivated by two possible scenarios: a "cookie tax" imposed by the government intended to increase consumers privacy, or higher entry costs advertisers would have to pay to adopt Google's alternative solution to third-party cookies.

Table 12 displays the equilibrium entry cost thresholds under this counterfactual. The result that the entry cost thresholds for both retargeting and non-retargeting campaigns increases appears counterintuitive at first glance. However, recall that the entry cost paid

Table 12: Equilibrium Entry Cost Thresholds - Cookies Taxed

Equilibrium Object	Value (N$) Cookies Available	Cookies Taxed
Avg. eqm. entry cost threshold for retargeting	0.064	0.076
Avg. eqm. entry cost threshold for non-retargeting	0.002	0.004

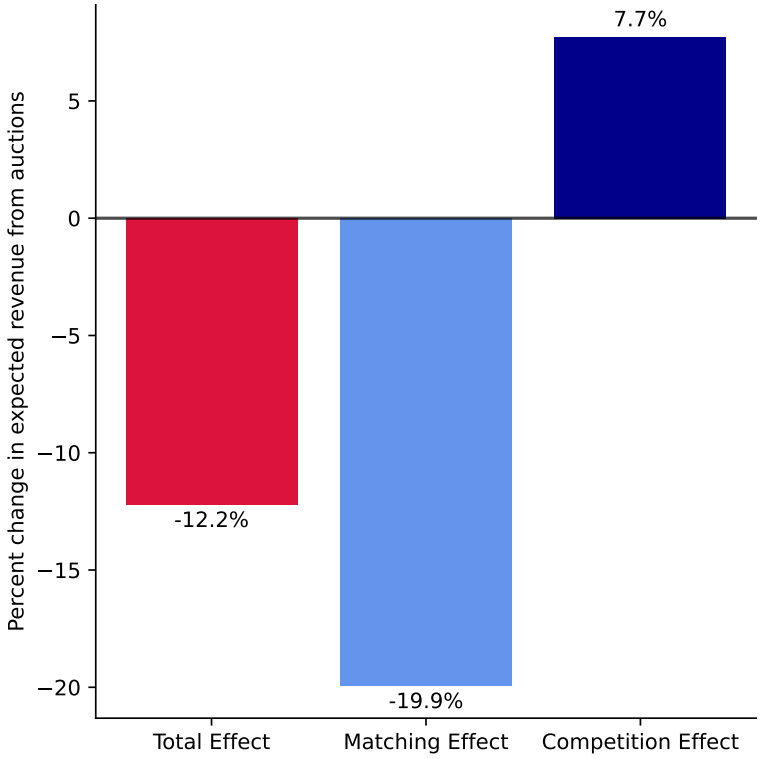
This table compares the average equilibrium entry cost thresholds predicted by the model in each targeting setting, in a counterfactual scenario where cookies are taxed to the ones when cookies are available as usual. Values are multiplied by a constant and displayed in a numeraire currency per thousand auctions (N$).

by retargeting campaigns is doubled. Therefore, participation in retargeting campaigns is reduced in this counterfactual exercise, as the increase in the threshold does not offset the effect of doubling entry costs paid for retargeting. As fewer retargeting campaigns compete, non-retargeting campaigns can experience higher surplus in the auctions, thereby increasing their entry cost thresholds and participation in auctions.

Figures 10 and 11 show the decomposition of the total effect of doubling entry costs for retargeting. Figure 10 shows that the Privacy Sandbox is effective in softening the negative effect of cookie deprecation on auction revenue. Under this new scenario, auction revenue decreases by 12.2%, a lower figure compared to the 60% decrease in case you ban cookies and do not replace them at all. Matching and competition effects are quantitatively relevant, with matching representing a 19.9% decrease in revenue, offset by a 7.7% increase due to tougher competition. These numbers are in line with recent tests conducted by Kobayashi et al. (2024) to assess the effectiveness of the Privacy Sandbox. In a series of experiments using thousands of advertisers in the US market, the authors find that Privacy Sandbox dollar conversions decrease by 14%.

Figure 11 shows that advertisers would have an average net zero effect on their surplus. The matching and competition effects offset each other, with a magnitude of 10.9% each. However, this happens because the survivors that still do retargeting are now facing an even thinner market than before. Whenever those bidders enter auctions, they now expect to have a higher surplus as they face lower competition from similar retargeting bidders, which

Figure 10: The Separate Effects of Match and Competition on Ad Auction Revenue Following Double Entry Costs for Retargeting

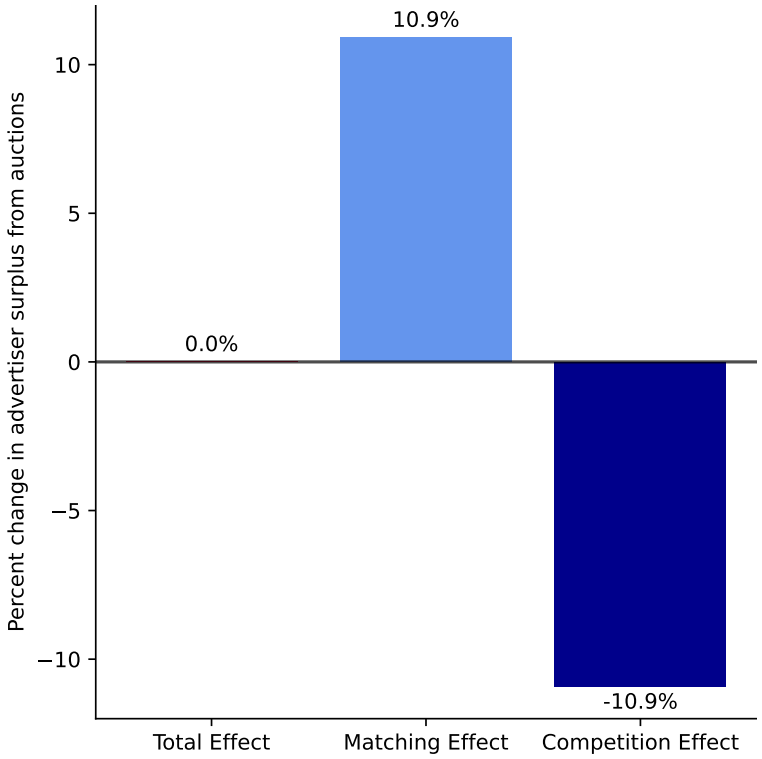


This figure shows the decomposition of the final effect on bids of doubling the entry cost to do retargeting. Bars represent the percent change in auction revenue, displayed on the y-axis. The red bar depicts the total effect. The effect attributed to a lower match between advertisers and consumers is in light blue, and the effect resulting from tougher competition in auctions is in dark blue.

effectively increases their expected surplus. Non-retargeting bidders, on the other hand, will still have their margins squeezed by tougher competition. Therefore, the net zero effect on advertiser surplus actually happens because remaining retargeting bidders that use the Privacy Sandbox are benefited at the expense of non-retargeting bidders. This shows that the CMA concerns are justified. Google's proposal for replacing cookies benefits their own customers. Since Google takes a share of revenue whenever advertisers win auctions using Google's products, the Privacy Sandbox has the potential to strengthen Google's position in the market even further.

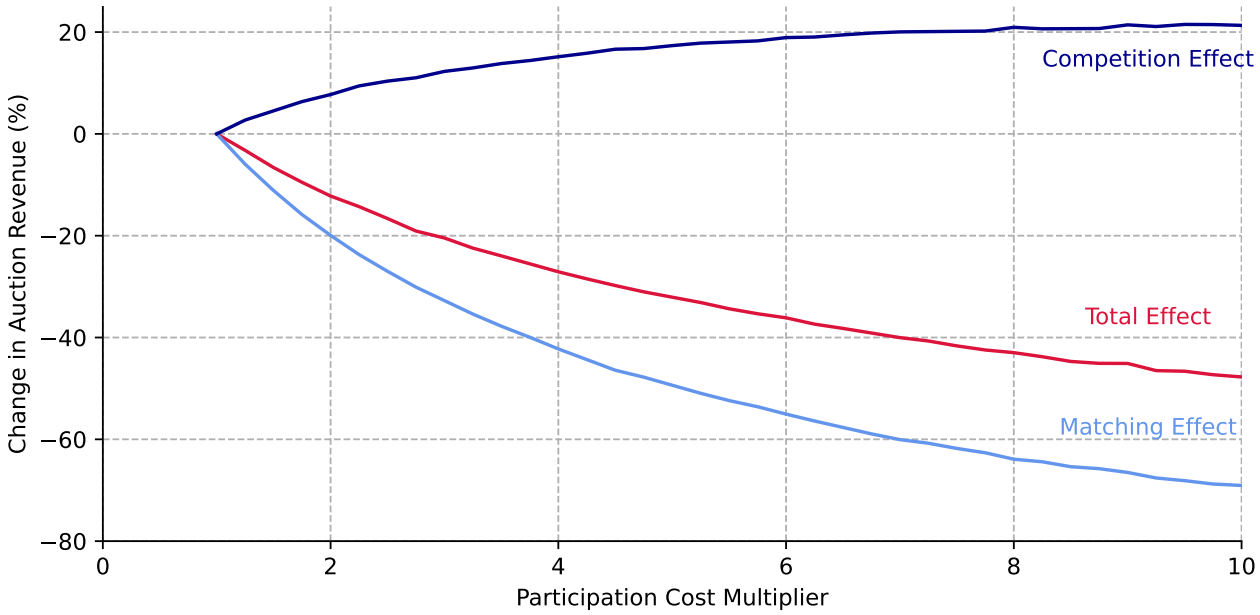
Multiplying entry costs by a factor of 2 is an approximation to the adoption costs faced by advertisers that seek to use the Privacy Sandbox. However, there is no variation in

Figure 11: The Separate Effects of Match and Competition on Advertiser Welfare Following Double Entry Costs for Retargeting



This figure shows the decomposition of the final effect on advertiser surplus of doubling the entry cost to do retargeting. Advertiser surplus is defined as the ex-post valuation derived from the auction (if win) subtracted by the payment (if win) and the entry costs (if participate). Bars represent the percent change in advertiser surplus, displayed on the y-axis. The red bar depicts the total. The effect attributed to a lower match between advertisers and consumers is in light blue, and the effect resulting from tougher competition in auctions is in dark blue.

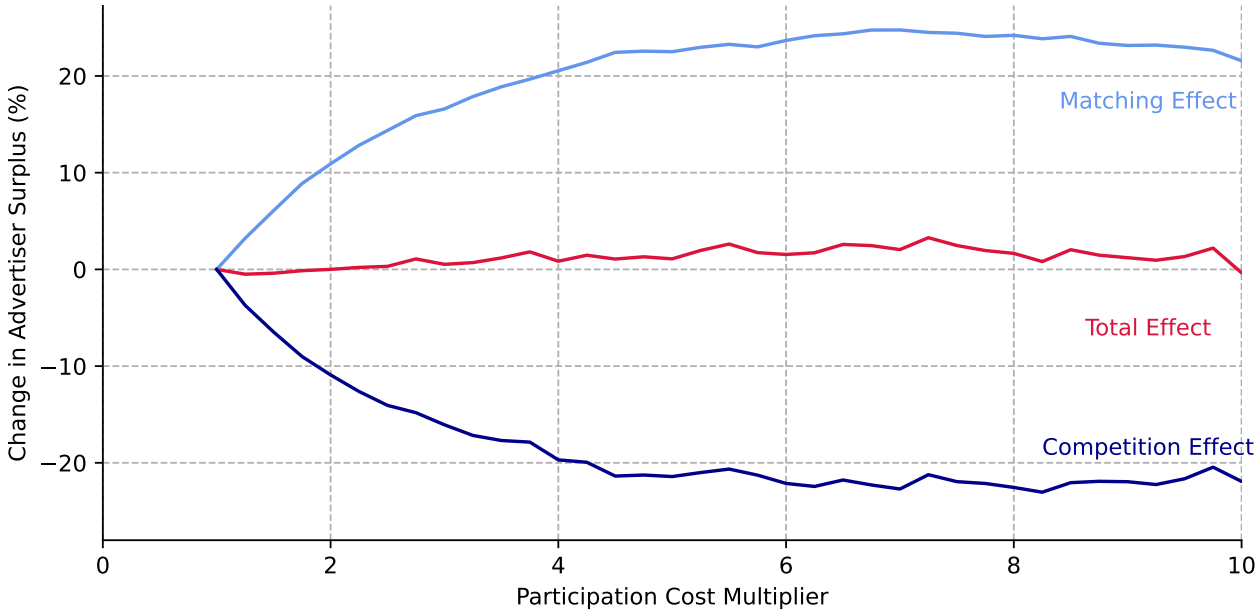
Figure 12: The Separate Effects of Match and Competition on Auction Revenue for Different Values of Entry Costs Multipliers



This figure shows the decomposition of the final effect on auction revenue (Y-axis) for different values of an entry cost multiplier (X-axis). The total effect is represented by the red line; the matching effect is represented by the red line; and the competition effect is represented by the dark blue line.

the dataset that allows the estimation of a precise cost multiplier that best represents the adoption costs of the Sandbox. For this reason, I decompose the matching and competition effects on auction revenue and advertiser surplus for different entry cost multipliers in Figures 12 and 13, respectively. Recall that banning cookies without any alternative technology, effectively removing retargeting, is estimated to decrease auction revenue by 60%. The industry could experience about half of that effect with an entry cost multiplier of as little as 4. This evidences that a substantial portion of the decrease in revenue happens early on, for not so large values of adoption costs. Therefore, if Google wants to provide a solution to cookieless targeting that does not harm the industry as much, it should focus on decreasing the adoption costs required to use it.

Figure 13: The Separate Effects of Match and Competition on Advertiser Surplus for Different Values of Entry Costs Multipliers



This figure shows the decomposition of the final effect on advertiser surplus (Y-axis) for different values of an entry cost multiplier (X-axis). The total effect is represented by the red line; the matching effect is represented by the red line; and the competition effect is represented by the dark blue line.

9 Conclusion

The digital display advertising industry has experienced remarkable growth, with advertisers in the United States spending over $150 billion on digital display ads in 2023 alone. A substantial portion of this spend, approximately $38 billion, is attributed to personalized auction-based ads that heavily rely on third-party cookies for retargeting consumers. However, the industry has come under increasing scrutiny due to data privacy concerns, leading to impending policy changes. One of the most significant changes on the horizon is the removal of third-party cookies by Google, the largest player in the industry.

The literature in economics and marketing has explored the impact of data privacy and the effect of providing bidders with more information on the match between advertisers and consumers prior to auctions. However, in light of the upcoming changes to how advertisers will have to adjust their targeting strategies following the removal of cookies, it is paramount to compute counterfactuals that endogenize advertisers' decision to enter specific targeting segments. This is crucial because changes to the access to and use of cookies can have two opposing impacts on final bids: on one hand, bids can decrease because of a lower match between advertisers and consumers. On the other hand, bids can increase as a result of fiercer competition in auctions, since advertisers may target broader audiences in the absence of the finer targeting enabled by cookies, making them compete in more auctions. This paper addresses this gap by proposing a structural model that accounts for the endogenous entry of advertisers in the online ad auctions market.

The structural model developed in this paper is a two-stage model of entry in auctions. In the first stage, advertisers decide whether to enter the market for online ads by paying a fixed cost. This cost represents the managerial and attention resources required to monitor the performance of their advertising campaigns. In the second stage, advertisers who have entered the market submit their bids. The model allows for the estimation of advertisers' valuations, which incorporate a retargeting premium and an idiosyncratic shock, as well as their entry cost distributions.

To estimate the model, I leverage a unique and comprehensive dataset spanning from November 2021 to March 2022. The data cover more than 1.4 billion auctions and over 200 advertisers in the Brazilian market. In addition to auction-level data, which include the two highest bids for each auction won by advertisers, I have access to data from the auction planner, a tool used by advertisers to predict the performance of their campaigns. This planner data provides valuable information on the winning probabilities associated with different targeting settings and bid levels, as well as the audience sizes reached by each advertiser's targeting configurations. I use this data to identify and estimates parameters pertaining to the entry costs paid by advertisers.

I estimate the model parameters using a simulated method of moments approach, matching key moments from the auction-level data and the planner data on winning probabilities. The results of the model estimation and counterfactual analysis provide valuable insights into the impact of policies intended to reduce the use of third-party cookies on the digital display advertising industry. In the absence of third-party cookies, the total ad spend would decrease by 60%. This overall effect can be decomposed into two opposing forces: an 84% decrease in ad spend due to a lower match value between advertisers and consumers and a 24% increase in ad spend resulting from heightened competition among advertisers. Advertisers would see their surplus reduced by 79%, with a 72% decrease from lower match amplified by a 7% decrease from fiercer competition. The increase in competition arises because those campaigns adopting retargeting tend to bid substantially higher than non-retargeting campaigns. Retargeting is no longer possible in the absence of cookies, so non-retargeting campaigns no longer compete against those high-valuation bidders anymore. Therefore, advertisers that do not employ retargeting have incentives to participate more often in auctions following the removal of cookies.

In addition, I run another counterfactual where the entry cost for retargeting is doubled. This is motivated by two possible scenarios. For example, the government could impose a "cookie tax" to discourage advertisers from violating consumers' online privacy.

Another motivation for imagining a scenario where entry costs for retargeting are increased is if advertisers need to incur high adjustment costs to adopt Google's Privacy Sandbox solution. In this counterfactual, auction revenue is expected to decrease by 12% and, interestingly, advertisers would have an average zero change in surplus. However, this happens because the remaining retargeting bidders that use the Sandbox would benefit from a thinner market at the expense of the non-retargeting bidders - an effect regulators are trying to avoid as the proposed alternatives to third-party cookies are primarily offered by Google, an already dominant player in the industry.

The findings in this paper have implications for both industry players and policymakers. Advertisers need to be prepared for a significant shift in the landscape of digital display advertising as third-party cookies are phased out. They will need to adapt their strategies and explore alternative methods for reaching their target audiences. Policymakers, on the other hand, should take into account the equilibrium effects that the removal of third-party cookies will have on the industry. While addressing data privacy concerns is important, policymakers should not disregard the potential impact of policies restricting the access of third-party cookies on competition.

Appendices

A Experiment Details

B Robustness Checks: Structural Model

References

Ada, Sıla, Nadia Abou Nabout, and Elea McDonnell Feit (2022). Context Information Can Increase Revenue in Online Display Advertising Auctions: Evidence from a Policy Change. *Journal of Marketing Research* 59.5, pp. 1040–1058.

Ahmadi, Iman, Nadia Abou Nabout, Bernd Skiera, Elham Maleki, and Johannes Fladenhofer (2024). Overwhelming Targeting Options: Selecting Audience Segments for Online Advertising. *International Journal of Research in Marketing* 41.1, pp. 24–40.

Alcobendas, Miguel, Shunto Kobayashi, and Matthew Shum (2021). The Impact of Privacy Measures on Online Advertising Markets. *SSRN Electronic Journal*.

Aridor, Guy and Yeon-Koo Che (2024). Privacy Regulation and Targeted Advertising: Evidence from Apple's App Tracking Transparency. *SSRN Electronic Journal*.

Athey, Susan, Jonathan Levin, and Enrique Seira (2011). Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions. *The Quarterly Journal of Economics* 126.1, pp. 207–257.

Budish, Eric, Robin S. Lee, and John J. Shim (2024). A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market? *Journal of Political Economy* 132.4, pp. 1209–1246.

Chen, Jianqing and Jan Stallaert (2014). An economic Analysis of Online Advertising Using Behavioral Targeting. *Mis Quarterly* 38.2, 429–A7.

Choi, Hana, Carl F. Mela, Santiago R. Balseiro, and Adam Leary (2020). Online Display Advertising Markets: A Literature Review and Future Directions. *Information Systems Research* 31.2, pp. 556–575.

Deisenroth, Daniel, Utsav Manjeer, Zarak Sohail, Steven Tadelis, and Nils Wernerfelt (2024). Digital Advertising and Market Structure: Implications for Privacy Regulation. *NBER Working Paper*.

Dube, Jean-Pierre H., Dirk Bergemann, Mert Demirer, Avi Goldfarb, Garrett Johnson, Anja Lambrecht, Tesary Lin, Anna Tuchman, Catherine E. Tucker, and John G. Lynch (2024).

The Intended and Unintended Consequences of Privacy Regulation for Consumer Marketing: A Marketing Science Institute Report. *SSRN Electronic Journal*.

Goldberg, Samuel G., Garrett A. Johnson, and Scott K. Shriver (2024). Regulating Privacy Online: An Economic Evaluation of the GDPR. *American Economic Journal: Economic Policy* 16.1, pp. 325–358.

Goldfarb, Avi and Catherine Tucker (2011). Online Display Advertising: Targeting and Obtrusiveness. *Marketing Science* 30.3, pp. 389–404.

Hummel, Patrick and R. Preston Mcafee (2016). When Does Improved Targeting Increase Revenue? *ACM Transactions on Economics and Computation* 5.1, pp. 1–29.

Johnson, Garrett A and Nico Neumann (2024). The Advent of Privacy-Centric Digital Advertising: Tracing Privacy-Enhancing Technology Adoption.

Johnson, Garrett A. (2013). The Impact of Privacy Policy on the Auction Market for Online Display Advertising. *SSRN Electronic Journal*.

Johnson, Garrett A., Scott K. Shriver, and Shaoyin Du (2020). Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry? *Marketing Science* 39.1, pp. 33–51.

Kobayashi, Shunto J, Garrett A Johnson, and Zhengrong Gu (2024). Privacy-Enhanced versus Traditional Retargeting: Ad Effectiveness in an Industry-Wide Field Experiment.

Korganbekova, Malika and Cole Zuber (2023). Balancing User Privacy and Personalization. *Working paper*.

Krishna, Vijay (2009). *Auction Theory*. Elsevier Inc.

Lambrecht, Anja and Catherine Tucker (2013). When Does Retargeting Work? Information Specificity in Online Advertising. *Journal of Marketing Research* 50.5, pp. 561–576.

Levin, Dan and James L. Smith (1994). Equilibrium in Auctions with Entry. *The American Economic Review* 84.3, pp. 585–599.

Levin, Jonathan and Paul Milgrom (2010). Online Advertising: Heterogeneity and Conflation in Market Design. *American Economic Review* 100.2, pp. 603–607.

Li, Tong and Xiaoyong Zheng (2009). Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions. *Review of Economic Studies* 76.4, pp. 1397–1429.

Lu, Shijie and Sha Yang (2020). Targeting Breadth in Behaviorally Targeted Display Advertising. *SSRN Electronic Journal*.

Marotta, Veronica, Vibhanshu Abhishek, and Alessandro Acquisti (2019). "Online Tracking and Publishers' Revenues: An Empirical Analysis". *Workshop on the Economics of Information Security*, pp. 1–35.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer (2022). Regulatory Spillovers and Data Governance: Evidence from the GDPR. *Marketing Science* 41.4, pp. 746–768.

Rafieian, Omid and Hema Yoganarasimhan (2021). Targeting and Privacy in Mobile Advertising. *Marketing Science* 40.2, pp. 193–218.

Ravichandran, Deepak and Nitish Korula (2019). Effect of Disabling Third-Party Cookies on Publisher Revenue. *Google Report*.

Roberts, James W and Andrew Sweeting (2010). Entry and Selection in Auctions. *Working Paper, Duke University*.

— (2016). Bailouts and the Preservation of Competition: The Case of the Federal Timber Contract Payment Modification Act. *American Economic Journal: Microeconomics* 8.3, pp. 257–288.

Samuelson, William F. (1985). Competitive Bidding with Entry Costs. *Economics Letters* 17.1–2, pp. 53–57.

Sun, Tianshu, Zhe Yuan, Chunxiao Li, Kaifu Zhang, and Jun Xu (2024). The Value of Personal Data in Internet Commerce: A High-Stakes Field Experiment on Data Regulation Policy. *Management Science* 70.4, pp. 2645–2660.

Wernerfelt, Nils, Anna Tuchman, Bradley Shapiro, and Robert Moakler (2022). Estimating the Value of Offsite Data to Advertisers on Meta. *SSRN Electronic Journal*.

Zeithammer, Robert (2019). Soft Floors in Auctions. *Management Science* 65.9, pp. 4204–4221.