Online Advertising, Data Sharing, and Consumer Control*

Justin P. Johnson† Thomas Jungbauer‡ Marcel Preuss§

October 21, 2023

Abstract. Invoking a property rights approach, we examine how competition between advertising exchanges influences the targeting options that these exchanges make available to advertisers. When advertisers have strong property rights over data regarding consumers' active purchase interests, competition between ad exchanges leads to too little sharing of data. This may harm consumers, who receive too few pertinent ads, and advertisers themselves can also be harmed due to a situation resembling a prisoner's dilemma. We find that reallocating property rights to consumers, i.e., giving them the right to opt out of tracking may also benefit consumers who allow tracking, by altering the incentives of ad exchanges to offer improved targeting options. In addition, we show that initiatives by Apple and Google to limit third-party tracking and to introduce alternative tracking systems such as Google's Topics, might benefit consumers by weakening the data property rights of advertisers. Because more data is shared by default under such systems, this can be true even if these systems are less accurate than third-party tracking systems.

Invoking a property rights approach we examine how competition between advertising exchanges influences the targeting options that these exchanges make available to advertisers. We are motivated by recent and ongoing market changes such as the introduction of the consumer privacy-rights legislation General Data Protection Regulation (GDPR) in the European Union, the deprecation of certain third-party tracking technologies by companies such as Apple and Google, and the introduction of new category-based tracking technologies such Google's Topics system.

Our analysis considers a particular type of online advertising, intent advertising. This refers to a situation in which a consumer has recently taken an action that indicates current intent to purchase in some category. In this context it is plausible that consumers prefer to see more rather than fewer relevant ads. For example, a consumer interested in buying athletic gear who has visited the Nike website may also benefit by receiving an advertisement from Adidas, even if the consumer did not visit the Adidas website. Yet, if advertisers have strong property rights over data reflecting consumer intent, ad exchanges may limit the targeting tools available to advertisers in a way that consumers receive fewer such ads.

*We thank John Asker, Alessandro Bonatti, Heski Bar-Isaac, Joyee Deb, Andrei Hagiu, Paul Heidhues, Maarten Janssen, Bruno Jullien, Heiko Karle, Jan Kraemer, Sherif Nasser, Kris Nimark, Martin Peitz, Markus Reisinger, Wilfried Sand-Zantman, Amin Sayedi, Christian Schmid, and Curtis Taylor, and participants of various conferences and seminars for valuable discussions and comments on earlier drafts of the paper, and Si Zuo for valuable research assistance. We also thank the Editor Joshua Gans, an anonymous associate editor and three anonymous referees.

†Cornell University, justin.johnson@cornell.edu.

‡Cornell University, jungbauer@cornell.edu.

§Cornell University, preuss@cornell.edu.

However, in this paper we show that category-based advertising systems (such as Google's Topics) that weaken advertisers' property rights may rectify the situation, ensuring consumers see additional relevant ads, while also potentially enhancing consumer privacy. Similarly, we show that endowing consumers with property rights over data reflecting their intent by giving them the right to opt out of tracking can cause ad exchanges to offer further targeting options. This benefits not only consumers with strong privacy preferences who opt out of tracking but even those who continue to allow tracking, as they receive additional relevant ads.

Our results can be easily understood by taking a property rights perspective on data ownership, and by understanding the structural features inherent in current third-party tracking systems. Consider the example of a consumer interested in purchasing athletic gear who has visited the Nike website. Although this consumer may benefit from also seeing an Adidas ad, Nike itself may not wish Adidas to advertise to this consumer; Nike would like to limit the flow of information regarding this consumer's current purchase interests. If an ad exchange insists on facilitating Adidas' targeting of consumers who visited Nike's website, then Nike may refuse to share data with the ad exchange, and instead share data only with an exchange that does not allow such "cross-targeting." That is, to the extent that Nike has property rights over the data indicating a consumer's current purchase interests, it may refuse to share that data with an ad exchange. Because neither Nike nor Adidas considers how their individual data-sharing decisions affect the profits of their rival, a prisoners' dilemma can result: neither shares data even though each would be better off if they both shared data. It is important to note that in reality ad exchanges do currently not enable cross-targeting, even though it is technically feasible.

In fact, under current third-party tracking systems advertisers do possess strong data property rights. For example, under the cookie system, ad exchanges that enable advertising on websites are able to track consumers only when websites consent to place tracking cookies into a consumer's browser on behalf of the ad exchange. Similarly, cross-app tracking on mobile devices requires that app developers use specific software development kits developed by ad exchanges.

Weakening the data property rights of advertisers can therefore benefit consumers if it induces ad exchanges to allow cross-targeting (that is, more widely disseminate data about a consumer's intent to purchase). As an application of this idea, consider Google's current proposal to replace the third-party cookie system (on Chrome) with the category-based system Topics, which uses in-browser machine learning tools to assign consumers to different categories based on their browsing history. Google argues that this will increase consumer privacy because the different categories are large enough that it will not be possible for advertisers to know exactly which websites any individual consumer visited. Another important aspect of Topics is that it collects browsing data by default, rather than relying on individual opt-in decisions of websites. In other words, Google becomes the default owner of data indicating a consumer's current intent to purchase, and Google will benefit from letting both Adidas and Nike target consumers based on that data.

Note that this beneficial outcome can emerge even if the Topics technology is less accurate than the current third-party cookie system, as advertisers often allege it will be. The reason is that, as already explained, under the current cookie system advertisers may withhold important data, whereas this cannot happen under Topics. This system may even benefit advertisers. The reason is that when advertisers would jointly benefit if they all agreed to share information with ad exchanges that enabled cross-targeting, then weakening the property rights of advertisers would eliminate any prisoner's dilemma type situation.

That said, there may be other costs associated with new advertising technologies that weaken advertisers' data property rights. We show that ad exchanges controlling these technologies may manipulate them to their own advantage. This is in line with the popular arguments that Google is eliminating third-party cookies to make Google's other advertising channels more attractive. Similarly, Apple's App Tracking Transparency policy makes cross-app tracking more difficult, which might benefit Apple's App Store ad business, or other ad businesses that Apple might develop in the future.

We now explain why granting consumers property rights over data reflecting their intent to purchase by allowing them to opt out of tracking can alter the data-sharing incentives of advertisers in a way that benefits both consumers who have strong privacy preferences and those who do not. When more consumers with strong privacy preferences opt out of tracking, advertisers worry less about the downsides of cross-targeting. Instead, advertisers become more concerned with reaching "look-alike audiences." Look-alike audience targeting is extremely common online, and refers to a situation in which, for example, a consumer is targeted with ads for athletic gear by Nike or Adidas not because they have visited either company's website but rather because they exhibit online browsing habits more generally associated with consumers interested in athletic gear. When advertisers share more data, they are able to target a larger look-alike audience. Therefore, when consumers have veto rights over how their data is used for tracking, advertisers are more willing to share data, and consumers who do not opt out of tracking see additional relevant ads.

We emphasize that the main problem identified in our baseline model is not dependent on the use of third-party cookies as tracking tools as such. Rather, it is the default advertiser property rights, the de facto ownership of intent-to-purchase data, which give advertisers some veto power over cross-targeting. Alternative tracking tools that are becoming pervasive, such as digital fingerprinting, pixel tracking, or the sign-in based technology Unified ID 2.0, maintain this ownership structure.

Our article is related to work on privacy, online advertising, and data intermediaries. Early work on privacy includes Taylor (2004) and Villas-Boas (2004), who study dynamic pricing when firms can track consumers over time. Montes, Sand-Zantman, and Valletti (2019) study price discrimination with endogenous privacy choices.[1] As in our study, privacy is often related to questions about data property and control rights, with early work from Hermalin and Katz (2006) and more recent contributions by Ichihashi (2020), Dosis and Sand-Zantman (2022) and Markovich and Yehezkel (2021). Choi, Jeon, and Kim (2019) examine privacy with data externalities.

Iyer, Soberman, and Villas-Boas (2005) argue that targeted advertising may lead to increased prices, but that it reduces wasted advertisements. Johnson (2013) and Bergemann and Bonatti (2011) both explore targeted advertising, examining the effect of changes in targeting accuracy. Athey, Calvano, and Gans (2018) identify how advertisers' desire to avoid reaching a given consumer too many times affects market outcomes; the possibility for duplicative or wasted ad impressions is an important part of our model. Goldfarb and Tucker (2011) empirically examine the effect of privacy regulations on the efficacy of online advertising. Villas-Boas and Yao (2021) consider dynamic retargeting.

Work on data intermediaries includes De Corniere (2016) and De Corniere and De Nijs (2016) who study search intermediaries. Hagiu and Jullien (2011) study search diversion. Bergemann and Bonatti (2015) study the pricing of information. Choi, Mela, Balseiro, and Leary (2020), Bergemann, Bonatti, and Gan (2022) and Ichihashi (2021) also investigate markets for data. The distinguishing feature of our work is our focus on practices specific to advertising markets. Other work that investigates the details of the modern advertising market includes Sayedi (2018), Kraemer, Schnurr, and Wohlfarth (2020), and D'Annunzio and Russo (2020).

1. Model

Here we present a model of information sharing and advertising. There are two advertisers, A_1 and A_2, and two advertising exchanges, ADX_1 and ADX_2. There is also a unit mass of consumers.

[1] Earlier work on targeting consumers and price discrimination that is not specific to an online environment includes Villas-Boas (1999), Fudenberg and Tirole (2000), and Chen and Iyer (2002).

We suppose that a mass $\alpha < 1/2$ consumers have recently and exclusively visited A_1's website, and α consumers have recently and exclusively visited A_2's website. The remaining $1 - 2\alpha > 0$ consumers have recently visited both websites.[2] Thus, a total of $1 - \alpha$ consumers have visited a given advertiser's website, where α of these consumers have exclusively visited this website.

The act of visiting one or more websites indicates that a consumer is actively interested in the product category served by both A_1 and A_2. However, even though consumers visit an advertiser because they are interested in the product, we assume that they do not make a purchase immediately.[3] Completing a sale is only possible if the consumer is also displayed an advertisement on a third-party content publisher website.[4]

The ability of advertisers to reach consumers with advertisements depends on the information sharing decisions of ADX_1 and the advertisers. The timeline of these decisions is given in Figure 1. In the first stage, ADX_1 publicly commits to either allowing cross-targeting or not (explained just below), whereas ADX_2 by assumption does not allow cross-targeting.[5] In the second stage, the two advertisers simultaneously and publicly decide which advertising exchange to (costlessly) share information with.[6] We assume each advertiser shares data with only one exchange but allowing it to share data with both has no effect on our results.

Advertising options in stage 3 are as follows. First, A_i can *retarget* those $1 - \alpha$ consumers who have visited its website. It can do this on ADX_j if and only if it agreed to share data with ADX_j in stage 2. For example, perhaps A_i placed a tracking cookie on behalf of ADX_j when consumers visited A_i's website.[7] Second, A_i can *cross-target* those $1 - \alpha$ consumers who have visited A_{-i}'s website. This option is only available if ADX_1 allowed cross-targeting in stage 1, and if A_i's rival A_{-i} shared data with ADX_1 in stage 2.

---

[2]Although we use the term "website", our model is more general. The important assumption here is that some consumers have revealed their interest (intent to purchase) in a category to some advertiser(s) by their actions.

[3]The analysis is identical if consumers who purchase immediately can be identified and excluded from the set of consumers who receive ads. If this cannot be done, the analysis nonetheless remains similar: because these consumers will not be responsive to ads, the value of reaching a randomly selected consumer who visited a website (defined as C or M below) will be smaller. But as long as this value still exceeds the cost of an advertisement, the analysis goes through unchanged. In fact, data shows that visit-to-purchase conversion rates of online shoppers in the US are in fact below 3% (Statista, 2022). The consumers' reasons for not buying immediately, apart from not wanting the product, may involve time restrictions, distractions, forgetfulness, or fatigue (see Ursu, Zhang, and Honka, 2022, who find evidence of search fatigue).

[4]In Section 4.3, we consider an extension in which retargeting is not necessary to convert a website visit into a sale.

[5]In an earlier version of this paper, we considered the possibility that both ad exchanges can allow cross-targeting. The analysis of that model was more intricate but led to the same qualitative results presented here.

[6]Assuming public observability is for simplicity. We could dispense with this assumption so long as advertisers can indicate which consumers they want to target (if available) and are only charged if these consumers are actually targeted.

[7]As noted earlier, our analysis in much of this article applies to situations in which advertisers have some control over whether information they collect about consumers is shared with other parties. This is also true of some newer advertising initiatives such as tracking pixels or digital fingerprinting, not just existing cookie technology.

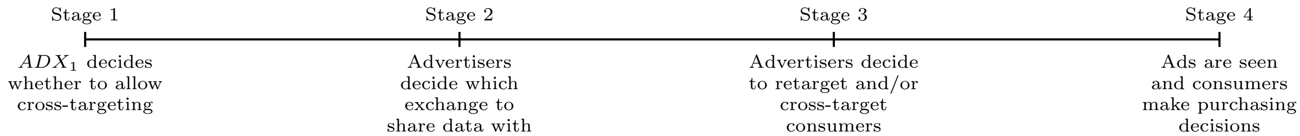

FIGURE 1. The Timeline

There is an additional advertising option available only through ADX_1. In particular, ADX_1 has the ability to reach an entirely separate group of consumers (who have visited neither the website of A_1 nor A_2), namely "look-alike" consumers.[8] Look-alike consumers have similar online habits to those consumers that have visited one of the two advertisers' websites, and ADX_1 is more able to reach those consumers with advertisements whenever it has more data to work with.[9] Precisely, if one advertiser has chosen to share data with ADX_1, then an additional $\eta_1 > 0$ consumers can be targeted, but if both advertisers have chosen to share data with ADX_1 then an additional $\eta_2 > \eta_1$ consumers can be targeted. Both advertisers can target these consumers, irrespective of which exchange they share data with.[10]

Placing an ad costs an exogenously determined value $w > 0$, where this cost is the same for each consumer.[11] For advertisers that choose to both retarget and cross-target, there is an advantage to concentrating these advertisements on ADX_1. In particular, an advertiser that retargets on ADX_2 and cross-targets on ADX_1 reaches $1-\alpha$ consumers on each exchange, thus wasting $2(1-\alpha)-1 = 1-2\alpha$ impressions (we assume showing a consumer more than one ad from the same firm has no effect on the consumer's decision). These wasted impressions do not occur when an advertiser concentrates its advertisements on one ad exchange (ADX_1).[12]

In the fourth and final stage, each consumer makes purchasing decisions. We assume that prices are exogenously fixed, and let M denote the "monopoly profit" that accrues to advertiser A_i from each consumer that exclusively sees an advertisement from A_i (gross of advertising fees). Let $C < M$ denote the "competition profit" that accrues to A_i from a consumer that sees an advertisement

[8] Targeting users who are "like" one's own buyers is extremely common. To this end, many firms share lots of customer data with Google or Facebook.

[9] For example, an intermediary such as Google controls a content platform (Youtube), email services (Gmail), and even a smartphone operating system (Android), all of which provide extensive unique data. To identify look-alike audiences, Google can compare its unique data with the data of website visitors it collects from advertisers who share data with Google.

[10] We could instead suppose that an advertiser can reach the look-alike audience only if it shares data with ADX_1. Then, the deviations we consider below entail a loss of look-alike consumers of size η_2 instead of $(\eta_2 - \eta_1)$, with little qualitative effect.

[11] In an extension presented in Section 4.2, we consider the possibility that the advertising cost for a given consumer is higher if both advertisers are trying to reach that consumer with an ad.

[12] In the real world ad exchanges in fact offer "frequency capping" services to limit how many times any one consumer receives the same ad from that particular exchange (see Buchbinder, Feldman, Ghosh, and Naor (2014)), thus avoiding wasted impressions.

from both advertisers.[13] We suppose that consumers who see both ads generate higher joint profits for A_1 and A_2, so that $2C > M$. This would be true if consumers are more likely to make a purchase if they have two choices rather than one, and prices are taken as given.

To close the model, we specify payments to the advertising exchanges. Each exchange earns an exogenously given fraction $\phi \in (0, 1]$ of the market price of an ad, which is exogenously set to w, for each advertisement that they facilitate. The remaining $(1 - \phi)w$ is assumed to go to the publisher.

We are now ready to analyze the model described above. Our primary interest is identifying the circumstances under which cross-targeting emerges, which can only happen if ADX_1 allows it and if advertisers share information with ADX_1. There are thus two relevant subgames to consider, corresponding to whether ADX_1 allows cross-targeting or not.

1.1. **ADX_1 does not allow cross-targeting.** Consider the subgame in which ADX_1 (in stage 1) does not allow cross-targeting. This means that the only option for each advertiser is to retarget those consumers who have visited its website (and also to target any available look-alike consumers on ADX_1). By definition, if A_i wishes to retarget consumers on ADX_j, then it must agree to share data with ADX_j (in stage 2). Because both advertisers will always choose to target all look-alike consumers, each of these consumers is worth $C - w$ to a given advertiser.[14]

If both advertisers share data with ADX_1, then each advertiser has profits of

$$\alpha M + (1 - 2\alpha)C - (1 - \alpha)w + \eta_2(C - w). \tag{1}$$

The intuition for this expression is as follows. A_i is retargeting all α consumers who have exclusively visited its website, earning $M - w$ from each of them (factoring in the advertising cost w and the fact that A_i's rival cannot target these exclusive consumers when cross-targeting is not allowed). It will also retarget all consumers who have visited both websites, of which there are $1 - 2\alpha$, earning $C - w$ from each (given that these consumers are also being retargeted by A_i's rival). Finally, it will target all available look-alike consumers, of which there are η_2 given that both advertisers are sharing data with ADX_1.

The only available deviation is for A_i instead to share data with ADX_2. However, the only effect of this deviation on A_i's profits is a reduction in the targetable look-alike audience from η_2 to η_1, which is unprofitable. In fact, when neither ad exchange allows cross-targeting, it is a dominant

[13] We are assuming it is either not profitable or not feasible for a single advertiser to purchase all ad slots for a given consumer so as to prevent that consumer from seeing a rival ad.

[14] We relax this assumption in Corollary 1 below.

strategy to share data with ADX_1 because doing so strictly expands the size of the look-alike audience. Thus, the unique equilibrium exhibits both advertisers sharing data with ADX_1.

1.2. ADX_1 **allows cross-targeting.** Now consider the subgame in which cross-targeting is allowed by ADX_1. We first investigate the conditions required for the existence of an equilibrium in which each advertiser shares data with ADX_1 and cross-targets the other advertiser, that is an equilibrium in which there is "full cross-targeting." In such an equilibrium, A_i earns

$$(1+\eta_2)(C-w). \tag{2}$$

The reason is that each advertiser reaches the mass one of consumers that have visited at least one advertiser's website, and additionally reaches the η_2 look-alike consumers that are available to target through ADX_1. The full cross-targeting outcome occurs whenever both advertisers share data with ADX_1 because A_i's only available deviations in the advertising stage (stage 3) are to forgo targeting its own visitors, the exclusive visitors of A_{-i}, or the look-alike audiences. But since neither of these three activities leads to wasted impressions when A_i has shared data with ADX_1, it would never be profitable to quit any of them.

There are two possible deviations for A_i from such an equilibrium, each involving A_i sharing data only with ADX_2; this ensures that A_i's rival A_{-i} cannot cross-target A_i. The deviating advertiser A_i must decide whether or not it will cross-target A_{-i}, which is possible because A_{-i} shares information with ADX_1 in the proposed equilibrium.

If A_i deviates but continues to cross-target A_{-i} ("leave and snipe"), it earns

$$\alpha M+(1-\alpha)C-2(1-\alpha)w+\eta_1(C-w). \tag{3}$$

Comparing (3) to (2), we see that the deviation is not profitable if and only if

$$\alpha(M-C)\leq(1-2\alpha)w+(\eta_2-\eta_1)(C-w). \tag{No Leave \& Snipe}$$

The left-hand side of this inequality represents the benefit of preserving monopoly power on the α consumers that exclusively visit A_i's website. The right-hand side represents the two associated costs of leaving and sniping. First, there are $(1-2\alpha)w$ costs wasted on duplicative impressions, due to the fact that this advertiser retargets $1-\alpha$ on ADX_2 and cross-targets $1-\alpha$ consumers on ADX_1, but there are only a total of (mass) 1 unique website consumers. Second, by not sharing data with ADX_1, there are only η_1 rather than $\eta_2>\eta_1$ look-alike consumers available.

Alternatively, A_i can share data with ADX_2 and not cross-target, giving it profits of

$$\alpha M + (1-2\alpha)C - (1-\alpha)w + \eta_1(C-w). \tag{4}$$

Comparing (4) to (2), this deviation to simply "leave" is not profitable if and only if

$$\alpha(M-C) \leq \alpha(C-w) + (\eta_2 - \eta_1)(C-w). \tag{No Leave}$$

The left-hand side of this inequality represents the benefit of preserving monopoly power on the α consumers that exclusively visit A_i's website. The right-hand side represents the two associated costs. First, this advertiser can no longer cross-target its rival's exclusive customers, which are worth $\alpha(C-w)$. Second, by not sharing data with ADX_1, there are only η_1 rather than $\eta_2 > \eta_1$ look-alike consumers available.

We conclude that if ADX_1 allows cross-targeting, a subgame with full cross-targeting is an equilibrium if and only if both the "No Leave & Snipe" and "No Leave" conditions hold. In the Appendix we further show that if an equilibrium with full cross-targeting exists, then it is Pareto dominant in terms of the advertisers' payoffs.

1.3. **Overall Equilibrium.** We are now ready to describe what happens in the overall game. To simplify the analysis, we restrict attention to Pareto undominated equilibria:

Assumption 1. *If a subgame has multiple equilibria, we remove any equilibrium which is strictly Pareto dominated by another equilibrium, in terms of the payoffs of the advertisers (who are the only players taking actions at this stage).*

The profits of ADX_1 are determined by how many advertisements are placed with it, recalling that it receives ϕw for each such advertisement. Hence, the ideal outcome for ADX_1 is that both advertisers share data and do both their retargeting and cross-targeting with it; this also maximizes the size of the look-alike audience, ensuring A_1 and A_2 each place an additional η_2 ads with ADX_1.

However, from our earlier analysis we know that it may not be an equilibrium for both advertisers to indeed share data with ADX_1, if ADX_1 allows cross-targeting. Instead, one or both advertisers may choose to work with ADX_2 in such a situation, which not only directly reduces how many re- or cross-targeting ads are placed with ADX_1 but also limits the size of the look-alike audience. This competitive pressure sometimes leads ADX_1 to not allow cross-targeting.

Proposition 1. *In equilibrium, both advertisers share their data with* ADX_1*, and retarget their exclusive website visitors and the look-alike audience of size* η_2 *using* ADX_1*. In addition,* ADX_1 *allows cross-targeting if and only if* (No Leave & Snipe) *and* (No Leave) *both hold, in which case both advertisers also cross-target on* ADX_1*.*

Proposition 1 indicates that, in equilibrium, ADX_1 ensures that both advertisers share data with it and purchase all of their ads from it. It does this by allowing cross-targeting if and only if it is certain that both advertisers will subsequently share their data with it (that is, if (No Leave & Snipe) and (No Leave) both hold). If, instead, allowing cross-targeting would cause one or both advertisers to not share data with ADX_1, then, to prevent this outcome, ADX_1 optimally does not allow cross-targeting. The reason it is so important to ADX_1 to have both advertisers share data with it is that, when an advertiser does not share data with it, it cannot sell retargeting ads to that advertiser (instead, ADX_2 does) and also achieves a smaller look-alike audience, which reduces profits received from both advertisers.

The cross-targeting outcome becomes less likely as α or M increase, or as C decreases. Intuitively, the reason why an advertiser might not share its data with ADX_1 when cross-targeting is allowed is that the advertiser wishes to protect the incremental monopoly profits $\alpha(M-C)$ of its exclusive website visitors. This explains why an increase in M makes cross-targeting less likely. These incremental monopoly profits are also larger when C decreases, but there is an additional effect, which is that the reduced size of the look-alike audience (as a result of not sharing data with ADX_1) matters less to advertisers (because those consumers are worth C to each advertiser).[15]

An implication of Proposition 1 is that cross-targeting may fail to emerge in equilibrium even when both A_1 and A_2 would prefer that it did.[16] To see this, note that advertiser profits if they cross-target each other are given in (2) while the profits if neither cross-targets (but both share data with ADX_1) are given by (1). From these expressions, advertisers prefer cross-targeting if and only if

$$M - C < C - w.$$

This condition says that the loss of monopoly power over an advertiser's own exclusive customers is made up for by the ability to reach the exclusive customers of its rival (where the mass of

[15] Mathematically, the comparative statics follow from the (No Leave & Snipe) and (No Leave) conditions and, for the most part, are apparent from inspection. To verify the effect of α on the (No Leave) condition, observe that if $C - w \geq M - C$, then (No Leave) holds for any α—changes in α do not affect whether (No Leave) holds. If, by contrast, $C - w < M - C$, then the left-hand side increases faster than the right-hand side, thus making it strictly harder for the (No Leave) condition to hold.

[16] In those cases, too little data is shared, which is reminiscent of the potentially insufficient connectivity between asymmetric internet backbone providers (Crémer, Rey, and Tirole, 2000). However, we consider symmetric advertisers, and find that inefficiency arises due to the unique incentives to leave and snipe that we identify in online advertising markets.

each group of consumers is α). If this condition holds, then (No Leave) also holds. However, (No Leave & Snipe) can still fail, and if it does then each advertiser is individually motivated to share data with ADX_2, so that it can cross-target its rival while its rival cannot cross-target. In particular, (No Leave & Snipe) tends to fail when w is smaller, whereas the condition for both advertisers to prefer cross-targeting, $M - C < C - w$, always holds for w sufficiently small, because we assumed earlier that $2C - M > 0$.[17] In this case, advertisers face a prisoner's dilemma.

More generally, an issue is that online advertising markets lack more sophisticated contracting options which make the eligibility to cross-target dependent on sharing data as well or on paying a fee. We conjecture that this is because even if cross-targeting required that an advertiser share its own data, it would be easy to circumvent this restriction. For example, an advertiser could simply share data from any barely visited website it owns (or even create a new website solely for this purpose). In return, the advertiser would obtain the right to cross-target all other advertisers that share data with the same ad exchange without giving away access to the exclusive visitors of its main website. Alternatively, the monetary transfers that compensate the advertiser who shares data may be difficult to verify and monitor both for the advertiser that shares the data and the advertiser that pays the fee. It may even be be difficult to ascertain the correct price for the data given how complex the environment could be.[18]

We now turn to the impact of ADX_1's decision on consumers. We begin by noting that there may be several effects on consumers that our analysis does not capture. For example, we do not consider the possibility that the cost of advertising is passed through to consumers. Also, we assume that the privacy costs borne by the consumer do not depend too much on whether they receive one targeted advertisement or more than one, but rather simply on whether or not they are being tracked. However, if instead some consumers suffer additional privacy-related costs when cross-targeting occurs, then this negative welfare effect must be factored in.

Under these maintained assumptions, there is a very clear consumer welfare prediction. In particular, consumers benefit from increased cross-targeting, because this provides them with more purchasing options. Therefore, consumers may be harmed by the equilibrium data sharing practices of ad exchanges in those cases where Proposition 1 predicts that cross-targeting does not

[17] As noted earlier, $2C - M > 0$ would hold in any reasonable model of symmetrically differentiated products in which prices are given and there is an outside option, as it says in effect that total sales would be higher when consumers see both products.

[18] For example, some other advertisers interested in the data may sell substitute products as we have assumed but these substitution patterns might be unknown in detail, and depend on how effective the creative content of advertisers is. And other advertisers may be simply interested in selling unrelated products to the same consumers.

emerge in equilibrium. In other words, there can be too little sharing of data.[19] The reason that too little data may be shared is that advertisers have property rights over data regarding individual consumers' current intent to purchase, and that ad exchanges cater to the desires of the advertisers. In the following sections, we investigate both technological and policy measures that weaken advertisers' property rights over consumer intent data and might encourage the sharing of more data so as to benefit consumers, while also giving consumers with strong privacy preferences the ability to avoid tracking.

1.4. **Comparative statics in the data advantage.** Here we investigate how ADX_1's ability to target look-alike audiences influences the equilibrium outcome. It is easy to see that both (No Leave & Snipe) and (No Leave) hold more easily as $\eta_2 - \eta_1$ increases. Intuitively, this is because a larger $\eta_2 - \eta_1$ makes deviating from the full cross-targeting equilibrium more costly as it implies a larger decrease in look-alike audiences that can be targeted. In addition, we have so far assumed that advertisers place the same value $C - w$ on a look-alike consumer as on other consumers. But whether this is true or not may well depend on how effective ADX_1's technology for identifying these look-alike consumers is. To study this effect, suppose instead look-alike consumers are valued at a possibly different amount π_L.

Corollary 1. *Improvements in look-alike audience targeting that increase $\eta_2 - \eta_1$ or the value π_L of the look-alike audience make it more likely that full cross-targeting emerges in equilibrium.*

In practice, improvements in targeting could either increase or decrease $\eta_2 - \eta_1$. If improvements primarily correspond to an expansion of the overall population that ADX_1 bases its look-alike targeting on, then it seems reasonable that both η_1 and η_2 would increase in a proportional manner, satisfying the assumption. On the other hand, improvements in such techniques might decrease $\eta_2 - \eta_1$ and, thus, make cross-targeting less likely. To see why, recall that the reason for assuming $\eta_2 > \eta_1$ was that ADX_1 has more data about website visitors if both advertisers share data with ADX_1, which it relies on to identify look-alike audiences. Thus, if improving its technology means that ADX_1 obtains better predictions with less data, $\eta_2 - \eta_1$ might shrink.

[19] A potential caveat is that when consumers are exposed to more cross-targeting they may in reality also receive fewer other advertisements; we have for simplicity ignored these other ads, or equivalently assumed them to have zero value for consumers. However, if outside ads have substantial value for consumers, then a complete accounting of consumer welfare would need to recognize the prospect that consumers might receive fewer such ads when they receive more targeted ads from A_1 and A_2. We consider this possibility in Section 3.

2. Consumer Data Property Rights

In this section, we study the effect of granting website visitors property rights over data reflecting their intent to purchase (here their browsing history), enabling them to opt out of tracking. Each website visitor has the option of completely blocking all tracking, which has the effect of ensuring that they receive no advertisements from A_1 or A_2 (by preventing advertisers to take ownership of data reflecting the consumer's intent to purchase).

We assume that a mass $q \in (0,1)$ of website visitors value receiving relevant ads more than they mind being tracked, and so allow tracking. However, $1-q$ website visitors value privacy more than they value receiving relevant ads, and so block tracking. From the viewpoint of advertisers, the consumers that block advertisements become irrelevant. Nonetheless, as we will show, the absence of these consumers may alter the equilibrium decision of ADX_1 to allow cross-targeting or not.

To derive the conditions under which a full cross-targeting equilibrium exists, we must again analyze the subgames that follow if ADX_1 allows cross-targeting or not. Due to the similarity with the previous section, we keep the analysis here short. If ADX_1 does not allow cross-targeting, then advertisers earn $\alpha q(M-w) + (1-2\alpha)q(C-w) + \eta_2(C-w)$ if both share data with ADX_1. This outcome remains the unique equilibrium in this case because whenever an advertiser does not share data with ADX_1, it could increase the targetable look-alike audience by deviating and sharing data with ADX_1 since $\eta_2 > \eta_1 > 0$. (Note that η_1 and η_2 might well depend on how many consumers are blocking tracking, that is on q; we accommodate this possibility in our formal results below.)

On the other hand, if ADX_1 allows cross-targeting, then each advertiser earns $(q+\eta_2)(C-w)$ if both advertisers share data with ADX_1, which induces full cross-targeting. As before, there are two deviations available for A_i: either share data with ADX_2 and cease cross-targeting its rival, or share data with ADX_2 and continue cross-targeting its rival while bearing some wasted ad impressions. Deriving the corresponding conditions here for the case in which consumers control their intent data is straightforward and yields the following two conditions as being individually necessary and jointly sufficient for an equilibrium with full cross-targeting to emerge (given that ADX_1 is allowing cross-targeting).

$$\alpha(M-C) \leq \alpha(C-w) + \frac{\eta_2 - \eta_1}{q}(C-w) \qquad \text{(C: No Leave)}$$

as well as

$$\alpha(M-C) \leq (1-2\alpha)w + \frac{\eta_2 - \eta_1}{q}(C-w). \qquad \text{(C: No Leave \& Snipe)}$$

Proposition 2. *In equilibrium, both advertisers share their data with ADX_1, and retarget their exclusive website visitors and the look-alike audience of size η_2 using ADX_1. In addition, ADX_1 allows cross-targeting if and only if* (C: No Leave & Snipe) *and* (C: No Leave) *both hold, in which case both advertisers also cross-target on ADX_1.*

We now consider the effect of granting consumers data property rights on the overall equilibrium. Observe that the case without consumer data property rights is equivalent to a situation in which $q = 1$, at least from the perspective of the advertisers and ad exchanges. Hence, to understand whether consumers holding property rights over their intent-reflecting data results in more or less cross-targeting in equilibrium, we wish to understand the impact of a reduction in q. Inspecting (C: No Leave) and (C: No Leave & Snipe), we see that each condition is more likely to hold for smaller q if

$$\frac{\eta_2 - \eta_1}{q}(C - w) \tag{5}$$

is decreasing in q. If η_1 and η_2 do not depend on q, then it is immediate that a reduction in q (that is, an increase in the number of consumers who block tracking) makes it easier to sustain an equilibrium with full cross-targeting. The reason is that as q becomes smaller, the size of the targetable group of website visitors relative to the size of the look-alike audience becomes smaller. Thus, advertisers care more about the look-alike audience as more consumers block, and in particular care more about increasing its size from η_1 to η_2, which is accomplished by sharing data with ADX_1, which also facilitates cross-targeting.

However, η_1 and η_2 are to likely depend on q. To explore the effect of consumer data property rights in this case, we write $\eta_2 = \eta(q)$ and $\eta_1 = \eta(q(1 - \alpha))$, where η is an increasing function. To understand these expressions, first consider η_2, the size of the look-alike audience when both advertisers share data with ADX_1. In this case, the data associated with all q consumers who do not block tracking is available to generate look-alike audiences, yielding $\eta_2 = \eta(q)$. But if only A_i shares data with ADX_1, then only the data of those consumers who do not block and who went to A_i's website is available to generate look-alike audiences. There are $q(1 - \alpha)$ such consumers, yielding $\eta_1 = \eta(q(1 - \alpha))$.

Proposition 3. *Consumer data property rights make it more likely that full cross-targeting emerges in equilibrium, if additional data has a decreasing marginal effect on the number of look-alike consumers that can be targeted (that is, if η is concave).*

As consumers receive property rights (or as q decreases), both η_2 and η_1, and possibly the difference between them, change. Concavity of η, however, guarantees that the difference between η_2 and η_1 does not shrink too much, so that the overall effect is that advertisers place relatively more weight on the look-alike audience compared to website visitors.

It is plausible that η is concave. Recall that η maps data from a pool of website visitors to a pool of look-alike users. Thus, the shape of η is determined by the returns to data, which are often considered to be decreasing (Agrawal, Gans, and Goldfarb, 2018).

We can now study the effect of consumer data property rights on consumer welfare. For the q website consumers that do not block tracking, let $u_1 > 0$ be the value of receiving just one targeted ad and $u_2 > u_1$ the value of receiving two targeted ads. Note that we are assuming that the look-alike audience either cannot or does not block ads. This is reasonable so long as the technology used to generate look-alike audiences is not as invasive as the tracking used on website visitors, or if these consumers are both tracked and targeted on a single platform that is also owned by ADX_1, such as YouTube, Facebook, or an app store, where it is more difficult to block tracking.

The welfare effects on consumers can be understood as follows. First, for those $1-q$ website visitors who strongly dislike tracking, they are now able to avoid being tracked, which clearly benefits them. Second, those q website visitors who do not block ads may receive more ads than in the case without consumer data property rights. This happens if the introduction of consumer data property rights shifts the equilibrium to one with full cross-targeting, causing these consumers to receive two ads (and utility u_2) rather than possibly just one ad (and utility $u_1 < u_2$). Third, the total size of the look-alike audience is smaller (at least if η_1 and η_2 depend on how many consumers block as assumed above). In particular, in the absence of consumer data property rights, data of all (mass one) website visitors is used by ADX_1 to generate a mass $\eta(1)$ of look-alike consumers. But, with consumer data property rights there is only a mass q of consumers whose data can be used, giving a look-alike audience base of size $\eta(q)$.

Therefore we conclude that endowing consumers with property rights over data reflecting their intent to purchase weakly benefits all website visitors, but harms some look-alike audience members who no longer receive pertinent advertisements.

Proposition 4. *Giving consumers data property rights (weakly) raises the surplus of all website visitors if additional data has a decreasing marginal effect on the number of look-alike consumers* *that can be targeted (that is, if η is concave). But, consumer data property rights lower the look-alike audience's welfare.*

As noted above, although the $1 - q$ consumers who block ads always benefit from consumer data property rights, those q consumers who do not block ads benefit only if the equilibrium switches to a full cross-targeting one. We can thus imagine even stronger data-control options for consumers that would ensure these q consumers benefit as well. In particular, this would be so if these consumers could costlessly share their data, causing both advertisers to target them with ads.

3. Category-Based Advertising

In recent years, market-leading companies that facilitate internet browsing have announced moves away from third-party tracking. Apple's Safari browser, for example, now requires that users actively opt in, and otherwise blocks third-party cookies by default. Google announced in 2020 that it would eventually discontinue the use of third-party tracking cookies on its industry-leading web browser, Chrome.

An alternative available to these companies is category-based advertising. In category-based advertising, the company that facilitates browsing offers advertisers the option to target consumers via categories that represent consumer interests or characteristics derived from their online behavior such as their browsing history (techcrunch.com, 2022). This is possible for these companies precisely because consumers are using their browsers to access online content, allowing the companies to directly track consumers' online activities and use this data, along with machine learning techniques, to assign consumers to categories.

Google has indeed decided to implement category-based advertising. A common argument in favor of category-based advertising is that it will better protect user privacy. One reason is that by definition these categories represent more aggregated data. Advertisers only know which categories a consumer has been assigned to, compared to current third-party tracking systems which allows advertisers to infer personal information based on the entire browsing history of each targeted consumer. Another reason is that, in Google's implementation, the data about an individual's browsing will only reside in the browser which itself assigns the consumer to the appropriate categories. Additionally, Google will refresh the categories (or 'Topics') a consumer is assigned to every three weeks, and may allow consumers to change their Topics themselves.[20]

[20]Google initially planned to replace third-party tracking using a technology called FLoC (Federated Learning of Cohorts). FLoC would have assigned consumers to a subset of as many as 30,000 different categories, using machine learning to analyze

The anticipated move away from third-party tracking cookies has created much consternation amongst advertisers (who argue that category-based advertising will be inferior to the current third-party tracking system and therefore harm consumers and advertisers) and some regulators (who welcome a higher level of consumer privacy but worry that the real goal of big tech companies may be to further advantage their own advertising businesses). The advertising industry is developing and improving third-party tracking alternatives, such as digital fingerprinting techniques and consumer sign-in systems (such as Unified 2.0), that seek to replicate the individual-level tracking afforded by third-party cookies. These technologies may help advertisers maintain the targeting accuracy they are accustomed to, and may also help them retain property rights over data reflecting consumer intent.

In this section, we build on our base model to allow for a look at some potential advantages and disadvantages of category-based advertising. We argue that the popular discussion has neglected a fundamental issue, which is that the strong market position of the companies that facilitate browsing give them the ability, in principle, to prevent advertisers from opting out of category-based advertising systems. That is, because such companies control much of the browser market, they can gain direct access to data about which websites consumers visit, and so advertisers are no longer technologically guaranteed to have property rights over data corresponding to an individual's interest in a particular product. More precisely, recall that under the third-party cookie tracking system any ad exchange that wishes to track consumers requires that advertisers allow the ad exchange to place a cookie in the browser of consumers who visit an advertiser's website. In contrast, under category-based advertising, data about a customer's browsing history is collected by the browser by default and used to assign the consumer to categories.

As we will show, removing the property rights of advertisers leads readily to an equilibrium in which more information is shared about consumer's current purchase interests. We note that this consumer benefit is completely separate from any benefit that consumers may accrue from the greater privacy afforded by such a system.

Consistent with our earlier approach, we will continue to assume that $1 + \eta_2$ is the total number of consumers interested in a certain category. However, conceptually, all consumers are now identified using a category-based system. Due to the nature of such a system, advertisers cannot differentiate between consumers who previously visited their website and those who did not. An implication is

consumers' browsing histories (wired.com, 2022). The sheer amount of potential categories, including sensitive ones such as gender and race, however, raised privacy concerns. Google has since announced that it will replace FLoC with Topics, and that the number of Topics will (initially) be limited to 300; Topics will enhance privacy relative to FLoC because Topics is based on more-aggregated consumer data than FLoC. Both FLoC and Topics have been born out of Google's Sandbox initiative.

that an advertiser who works with both exchanges cannot limit the consumers targeted on ADX_1 to exclude those targeted on ADX_2 (unlike in the base model where an advertiser could retarget on ADX_2 and then separately target η_2 distinct look-alike consumers on ADX_1).

Additionally, we assume that ADX_1 puts category-based advertising into effect. This system sometimes results in wasted impressions due to its imperfect nature. Precisely, we suppose that it takes $\tau > 1$ ad impressions to reach a mass one of consumers on ADX_1. Turning to ADX_2, we suppose it uses one of the newer third-party tracking alternatives to category-based advertising to offer retargeting.[21] Because such systems may require consumer opt-in (as Unified ID 2.0 does), we assume that an advertiser using this tracking technology on ADX_2 can reach a fraction $\sigma \leq 1$ of the $1-\alpha$ consumers who recently visited its website. The timing is as follows. First, advertisers purchase advertisements. Then, ads are seen and consumers make purchasing decisions.

We now describe the possible outcomes of the category-based advertising model analyzed in this section. One possibility is that both advertisers work with ADX_1, so that each advertiser is able to reach the entire market of $1+\eta_2$ consumers, each of which is valued at C. However, due to the inaccuracy of category-based advertising, this requires a total of $\tau(1+\eta_2)$ ads, implying total profits per advertiser of

$$(1+\eta_2)(C-\tau w).$$

Alternatively, if both advertisers retarget their website visitors on ADX_2, each has a profit of

$$\sigma\left[\alpha M + (1-2\alpha)C - (1-\alpha)w\right].$$

Except for the scaling parameter σ, these are the same profits that advertisers would get in the base model if they both used only ADX_2.[22]

The third possibility is that A_1, say, uses ADX_1 while A_2 uses ADX_2. In this situation, A_1 earns

$$(1+\eta_2)M - \sigma(1-\alpha)(M-C) - (1+\eta_2)\tau w.$$

The intuition is that A_1 reaches a total of $1+\eta_2$ consumers, but faces competition only on those $\sigma(1-\alpha)$ consumers that its rival A_2 also targets. The profit of A_2 in this case is

$$\sigma(1-\alpha)(C-w),$$

[21] Note that buying ads on ADX_1 using the category-based system does not involve a decision about sharing data. In addition, advertisers do not mind sharing data with ADX_2, as it does not support cross-targeting. We therefore omit a formal analysis of the data sharing stage in this section.

[22] Note that it cannot be optimal for an advertiser to use both exchanges because the set of consumers reached on ADX_2 is a strict subset of those reached on ADX_1.

which says that A_2 faces competition on all $\sigma(1-\alpha)$ consumers it reaches. Note that, unlike in the base model, A_2 cannot avoid being cross-targeted by A_1 simply by doing its retargeting on ADX_2. Our first result identifies the condition under which both advertisers bid only on ADX_1.

Proposition 5. *In the equilibrium of the category-based advertising model, both advertisers cross-target if category-based advertising is not too inaccurate. In particular, both advertisers cross-target on ADX_1 if and only if*

$$\tau < \tilde{\tau} \equiv \frac{1+\eta_2-\sigma\,(1-\alpha)}{1+\eta_2}\frac{C}{w} + \frac{\sigma\,(1-\alpha)}{1+\eta_2}, \tag{6}$$

where $\tilde{\tau} > 1$.

Observe that cross-targeting is less efficient than in the base model because $\tau > 1$ (so that $\tau - 1$ impressions are wasted). Nonetheless, full cross-targeting is certain to emerge under category-based advertising so long as this inefficiency is not too great, whereas without category-based advertising cross-targeting might not occur. Intuitively, under category-based advertising, the leave and snipe option is no longer available to advertisers. In terms of property rights, advertisers no longer have any exclusive rights to the data which consumers are interested in making a purchase; instead ADX_1 by default owns this data (for example, due to its control over the browser). An advertiser therefore cannot prevent its exclusive website visitors from being cross-targeted by leaving ADX_1 as in the base model. Hence, the only benefit of leaving ADX_1 is to reduce wasted impressions. If the number of wasted impressions is sufficiently small (τ close to 1), however, no advertiser leaves ADX_1 and full cross-targeting emerges. Note that this logic applies even if $\eta_2 = 0$.

We now turn to an assessment of consumer welfare under category-based advertising. As just discussed, cross-targeting is more likely to arise as long as category-based advertising is not too inaccurate. In turn, this suggests that consumers interested in the product category are better off under category-based advertising. Another advantage of category-based advertising is that some consumers may prefer the additional privacy associated with it. The reason for improved privacy is that the reduced tracking accuracy of category-based advertising, as compared to third-party cookies, makes it more difficult for advertisers to track consumers across the web and identify them individually from their browsing behavior.

Nonetheless, category-based advertising does carry an additional cost. In particular, because it is imperfect ($\tau > 1$) at identifying consumers who are interested in the product category, some consumers may receive ill-targeted advertisements. Suppose that these consumers incur a nuisance cost $\kappa > 0$ when receiving an ad from an advertiser in the product category.[23]

Recall that $u_1 > 0$ is the value of receiving one targeted ad, while $u_2 > u_1$ describes the value of receiving two targeted ads for a consumer interested in the product category. The following result characterizes the net welfare effect of adopting category-based advertising if this leads to a full cross-targeting equilibrium ($\tau < \tilde{\tau}$) instead of one in which only retargeting occurs. (This result does not take into account the additional possible benefit of greater consumer privacy under Topics.)

Proposition 6. *Suppose that full cross-targeting arises if category-based advertising is adopted by ADX_1, but that otherwise it does not. The net welfare effect of introducing category-based advertising is positive if the system is sufficiently accurate, that is if*

$$\tau < 1 + \frac{\alpha}{1+\eta_2}\frac{u_2 - u_1}{\kappa}.$$

The net welfare benefit of adopting category-based advertising is positive if the value of additional ads received by consumers interested in the product category exceeds the nuisance cost incurred by consumers not interested in the category. More precisely, the 2α consumers who exclusively visit a single website receive an additional ad under category-based advertising, which they each value at $u_2 - u_1$. On the other hand, $(\tau - 1)(1 + \eta_2)$ consumers who are not interested in the category each now receive two ads for it, thus suffering a nuisance cost 2κ. Equating these costs and benefits and solving for the critical τ value gives the result above.

Although consumers incur the nuisance cost κ when being targeted with an irrelevant ad, we now show that ADX_1 may benefit from reducing the accuracy of category-based advertising. Doing so leads to excess impressions and, hence, more ad sales for ADX_1. The only constraint is that category-based advertising must not be too inaccurate because advertisers will not use it otherwise (Proposition 5).

Proposition 7. *In the region of category-based advertising accuracy where both advertisers bid on ADX_1, i.e., $\tau < \tilde{\tau}$, (i) ADX_1's profits strictly increase as category-based advertising becomes less accurate, that is as τ increases. Moreover, (ii) the threshold $\hat{\tau}$ is larger if its rival ADX_2 is less efficient (σ is smaller).*

[23]One might imagine that this cost might also arise (only) with the look-alike audience in the base model. We believe the issue is much more salient here because all consumers are targeted using a category-based system; it is as if all consumers are now look-alike consumers.

In summary, as long as both advertisers bid on its ad exchange, ADX_1 benefits from category-based advertising being less accurate. At the same time, we see that ADX_1 may benefit from reducing the accuracy of its rival's system allowing it to reduce tracking accuracy further to sell even more ads. We have already identified that making third-party cookies less effective may be one means of accomplishing this. But consider an alternative context, that of in-app mobile advertising. Recently, Apple adopted its App Tracking Transparency program, which makes cross-app tracking of consumers more difficult and reduces the efficiency of existing in-app ad systems. Some commentators have suggested Apple's motivation for this initiative is less about privacy and more about giving an advantage to its own App Store ad system (vox.com, 2022).

Our final result concerns a potential regulatory constraint on τ. To the extent that higher τ values suggest consumer data being more disaggregated, regulators interested in privacy might impose a minimum level of τ. We now show that imposing too high a level of τ can have detrimental effects.

Corollary 2. *Suppose that there is an exogenously given requirement that category-based advertising not be too accurate, that is, if ADX_1 uses category-based advertising it must set $\tau > \bar{\tau}$. Then, if $\bar{\tau}$ is too large, ADX_1 is better off not introducing category-based advertising.*

If τ is large, both advertisers opt to retarget their website visitors on ADX_2. As a consequence, ADX_1 is better off not adopting category-based advertising. Note that this insight does not follow directly from Proposition 5 because ADX_1 may still prefer to adopt category-based advertising if $\tau > \tilde{\tau}$. In fact (as detailed in the proof of Proposition 5), if τ just exceeds $\tilde{\tau}$, one advertiser bids on ADX_1, while the other advertiser retargets on ADX_2 in equilibrium. If the number of excess impressions is sufficiently large, ADX_1 prefers this outcome over advertisers retargeting only in the base model even though full cross-targeting does not emerge in under category-based advertising.

4. Extensions

In this section we consider several extensions of the base model considered earlier.

4.1. **Higher profits from consumers who visit both websites.** In this extension, we assume that consumers who visit both websites are more likely to make a purchase than consumers who visit one website only. We model this difference in behavior by assuming that expected profits from targeting such a consumer are $\hat{C} > C$ with competition and $\hat{M} > M$ without competition.

It turns out that this change has no effect on the decisions of advertisers or ad exchanges, and hence does not alter any equilibrium predictions (except for the level of advertiser profits). The reason is that the consumers who visit both websites are always targeted, regardless of the data sharing choices. Hence, the gross profits associated with consumers who visited both websites, $(1-2\alpha)\hat{C}$, are always realized in any subgame, and hence do not factor into the attractiveness of any deviations.

To see this in a concrete example, consider an advertiser contemplating a deviation from a potential equilibrium with full cross-targeting. In the base model, we know that for this to be an equilibrium requires that (No Leave) and (No Leave & Snipe) both hold. For convenience we repeat the (No Leave & Snipe) condition here:

$$\alpha(M-C) \leq (1-2\alpha)w + (\eta_2 - \eta_1)(C-w).$$

The term $\alpha(M-C)$ represents the gain of avoiding competition for the consumers who have exclusively visited A_i, and hence does not change in this extension. The term $(1-2\alpha)w$ represents the costs of duplicative ads from operating on two ad exchanges, and does not change. Finally, $(\eta_2 - \eta_1)(C-w)$ represents profits from the look-alike audience, who have not visited either advertiser's website, and so this term does not change either.[24] Repeating this process for all potential deviations of advertisers in various subgames always leads to the conclusion that nothing changes.

4.2. **Higher cost of cross-targeting.** In this extension, we relax the assumption that the cost of targeting is independent of the number of advertisers targeting the same consumer. We instead assume that the cost of targeting a website visitor that is also targeted by one's rival is $\hat{w} > w$. We maintain the assumption that it is profitable to target all customers interested in the product category, so that $C > \hat{w}$. For consistency, we also assume that targeting look-alike audiences also costs $\hat{w}$, given that both advertisers target them.

This change does not affect the outcome when ADX_1 chooses not to allow cross-targeting. The reason is that the consumers who receive two ads are always the same whether advertisers work with ADX_1 or ADX_2. Only those $1-2\alpha$ consumers who visited both websites receive both ads. Hence, the fact that ADX_1 offers a look-alike audience that increases in size with more data ensures both advertisers share data with ADX_1.

[24]One might ask whether changing the profits associated with look-alike consumers might affect the equilibrium. The answer is yes, as shown in Corollary 1.

But when ADX_1 allows cross-targeting, the decisions of advertisers may well change. Consider the potential equilibrium in which full cross-targeting occurs. As explained earlier, one potential deviation for A_i is to instead share data with ADX_2 and not cross-target its rival. In the base model, one effect of this is a loss of $\alpha(C - w)$ from not cross-targeting, and this loss now reduces to $\alpha(C - \hat{w})$. A second effect of this deviation in the base model is the gain of $\alpha(M - C)$ from protecting this advertiser's exclusive customers from its rival. In this extension, this gain still exists but there is an additional benefit which is that the costs of reaching these customers fall from $\hat{w}$ to w, exactly because the rival is now incapable of targeting these consumers. Overall, also factoring in changes in profits due to the look-alike audience, this deviation is not profitable if and only if

$$\alpha(M - C) + \alpha(\hat{w} - w) \leq \alpha(C - \hat{w}) + (\eta_2 - \eta_1)(C - \hat{w}). \tag{7}$$

Because the left-hand side is larger than in the base model and the right hand side is smaller, overall this deviation is attractive on a larger set of parameters than in the base model.

The leave and snipe deviation also changes. The additional effect is that the duplicative costs associated with this deviation are now higher, precisely because the cost of reaching a consumer with advertising is higher when both advertisers are targeting that consumer. Consequently, this deviation is not profitable if and only if

$$\alpha(M - C) + \alpha(\hat{w} - w) \leq (1 - 2\alpha)\hat{w} + (\eta_2 - \eta_1)(C - \hat{w}). \tag{8}$$

This condition may be either easier or harder to satisfy, depending on the parameters. In particular, the left-hand side becomes larger, but whether the right-hand side becomes smaller or larger depends on the parameters. Nonetheless, we obtain a result that resembles our earlier result: If ADX_1 is open, there is full cross-targeting if and only if both (7) and (8) hold. The remainder of the equilibrium analysis from the base model goes through, except for one thing. In certain circumstances, allowing cross-targeting can be more profitable than not allowing it, even if the resulting subgame only exhibits one advertiser engaging in cross-targeting. In particular, there may be an equilibrium in which A_1 shares data with ADX_1 and is cross-targeted by A_2, but A_2 only shares data with ADX_2 and so cannot be cross-targeted by A_1.

To see why, suppose that ADX_1 allows cross-targeting and it leads to this "one-sided cross-targeting". Compared to not allowing any cross-targeting, the loss is that each advertiser only spends $\hat{w}$ on η_1 rather than η_2 look-alike consumers, for a total loss (for ADX_1) of $2\hat{w}\phi(\eta_2 - \eta_1)$. The gains accrue from each advertiser: A_1 now pays an additional $\alpha(\hat{w} - w)$ because its exclusive customers are now cross-targeted, while A_2 also now pays an additional $\hat{w}-w$ on an α consumers.[25] Overall, this implies that allowing cross-targeting is profitable in this case if $\alpha(\hat{w}-w) \geq (\eta_2-\eta_1)\hat{w}$.

However, if $\alpha(\hat{w}-w) < (\eta_2-\eta_1)\hat{w}$, then ADX_1 does not favor the one-sided cross-targeting outcome. In this situation, as in the base model, it can be shown that cross-targeting prevails in equilibrium if and only if there is full cross-targeting when ADX_1 allows cross-targeting.

Overall, this extension shows that, when it is more costly to target consumers that are also targeted by one's rival, advertisers have incentives to exploit their data property rights that are similar to those in the base model. As in the base model, this limits equilibrium cross-targeting, harming consumers.

4.3. **Consumers buy without retargeting.** In this extension, we relax the assumption that advertisers must retarget their website visitors to sell to them. To examine the effect as starkly as possible, we make the extreme opposite assumption and suppose that consumers have perfect recall so that there is no value in an advertiser retargeting.[26]

As before, we assume that consumers do not buy immediately so that cross-targeting serves the same purpose as in the base model. That is, by cross-targeting, an advertiser may change the final purchase decision of a consumer. Precisely, if a consumer visits only one advertiser, this advertiser earns M if the rival does not cross-target and C otherwise. In addition, each advertiser can expect profits of C from a consumer who visits both websites (without retargeting). Lastly, cross-targeting on ADX_1 costs $(1-\alpha)w$ if an advertiser does not share data with ADX_1 and αw otherwise. This is consistent with the base model, where we assumed that sharing data with ADX_1 allows advertisers to avoid wasted impressions since the exchange has access to the advertisers' data.

Overall, the advertisers' strong data property rights lead to the same qualitative predictions we obtained from the base model. In fact, this model is very similar to the base model. Intuitively, in the base model advertisers used retargeting both in equilibrium and when deviating. Hence, in effect removing retargeting as in this extension has no effect on the advertisers' behavior. To see this formally, consider the potential equilibrium with full cross-targeting. This exhibits profits for

[25]Precisely, A_2 now pays $\hat{w}$ on all $1-\alpha$ customers it cross-targets on ADX_1 rather than only paying $\hat{w}$ on the $1-2\alpha$ consumers who visited both websites. As $1-\alpha-(1-2\alpha)=\alpha$, its total increase in outlay is $\alpha\hat{w}$.

[26]It would be straightforward to model the case that consumers buy without being retargeted with a probability r less than one. If $rC > w$, sellers would continue to retarget all visitors and the results are similar to the base model. If $rC < w$, it is more profitable not to retarget, implying that the analysis is similar to the case we present here.

an advertiser of $(1+\eta_2)C - (\alpha+\eta_2)w$. In comparison, the deviation to leave and snipe yields

$$\alpha M + (1-\alpha+\eta_1)C - (1-\alpha+\eta_1)w$$

whereas the deviation to leave without cross-targeting yields

$$\alpha M + (1-2\alpha+\eta_1)C - \eta_1 w.$$

Comparing these terms yields necessary and sufficient conditions for cross-targeting (if ADX_1 allows cross-targeting) that are equivalent to (No Leave & Snipe) and (No Leave) from the base model.

Although the decisions of advertisers do not change, it turns out that ADX_1 may wish to act differently. In particular, ADX_1 might choose to allow cross-targeting even it does not lead to full cross-targeting but instead leads to only one advertiser cross-targeting the other. This is because not allowing cross-targeting only leads to a total number of ads on ADX_1 equal to $2\eta_2$ due to the lack of both retargeting and cross-targeting. By contrast, if one advertiser does not share data with ADX_1 but cross-targets on ADX_1, then ADX_1 sells $1-\alpha+2\eta_1$ ads. Thus, ADX_1 prefers the outcome with "one-sided cross-targeting" over not allowing cross-targeting if

$$1-\alpha > 2(\eta_2-\eta_1). \tag{9}$$

Intuitively, if allowing cross-targeting leads to one-sided cross-targeting, ADX_1 must weigh the benefit of guaranteeing a look-alike audience of size η_2 instead of η_1 against the profits from having at least one advertiser cross-target. If (9) holds, then the profits from enabling cross-targeting dominate and ADX_1 will allow cross-targeting even if the resulting subgame exhibits only one-sided cross-targeting. Of course, ADX_1 also chooses to allow cross-targeting if full-cross-targeting is an equilibrium since this guarantees a look-alike audience of size η_2 and leads to cross-targeting, thus strictly dominating not allowing cross-targeting, just like in the base model.

Overall, there is weakly more cross-targeting in equilibrium if advertisers do not need retargeting. This change is solely driven by the possibility of one-sided cross-targeting being more profitable for ADX_1 than not allowing cross-targeting, which does not happen in the base model because of the presence of retargeting.

4.4. **Competition in ad prices.** In this extension, to understand how competition in ad prices might affect our results, we suppose that ad exchanges choose the price of ads sold on their exchanges (in addition to choosing a data sharing regime). The main insight from this exercise is that the same economic forces are at work as in the base model, so that in equilibrium advertisers' data property rights limit cross-targeting even if advertisers would jointly prefer it. The key comparative statics also remain the same.

Let w_1 denote the price an advertiser pays for ads on ADX_1, and w_2 the price on ADX_2, respectively. To study the maximum impact that competition can have on the targeting options made available to advertisers in equilibrium, we consider the extreme case of $w_2 = 0$.[27]

The ad price w_1 is set after ADX_1 commits to allowing cross-targeting or not and before advertisers decide which advertising exchange to share data with. To simplify the analysis, we assume throughout here that $\eta_1 = \eta_2/2 = \eta$, so that there are constant returns to data.[28]

To derive the equilibrium outcome when advertising costs are endogenous, we compare ADX_1's profits if it allows cross-targeting to its profits if it does not allow cross-targeting. If ADX_1 does not allow cross-targeting, then both advertisers will share data with it if and only if w_1 is such that

$$(1-\alpha)w_1 \leq \eta(C - w_1). \tag{10}$$

The left-hand side depicts the cost savings of w_1 per ad if an advertiser instead uses ADX_2 to reach the $1-\alpha$ consumers who have visited its website. As in earlier analysis, this results in a smaller look-alike audience, which reduces this advertiser's profit by the amount on the right-hand side (recall that $\eta_2 - \eta_1 = \eta$ here). Thus, when this inequality holds, advertisers individually prefer to share data with ADX_1, so that this is the equilibrium.

Similarly, if (10) does not hold, neither advertiser uses ADX_1 and so ADX_1 gets zero profits. This cannot be optimal, and so when ADX_1 does not allow cross-targeting it will always choose w_1 such that (10) holds exactly. ADX_1's ad price thus satisfies

$$w_1 = w^{nx} := \frac{\eta}{1-\alpha+\eta}C > 0. \tag{11}$$

Next we consider the case that ADX_1 allows cross-targeting. As a preliminary matter we note that for any given value of w_1, Lemma 5 (in the Appendix) ensures that there exists a pure strategy equilibrium in the resulting subgame. There are three potential pure strategy equilibria here: (i)

[27] This assumption is consistent with our previous analysis in which ADX_2 is essentially passive, and holds, for example, if ADX_2 represents a competitive fringe. Moreover, as is readily verified, in equilibrium it would not be profitable for ADX_2 to instead charge any $w_2 > 0$ (if we allowed w_2 to be chosen), as so doing will not allow it to capture any demand from ADX_1.

[28] An analysis with arbitrary η_1 and η_2 leads to qualitatively similar conclusions but requires a more complicated analysis.

both advertisers share data with ADX_1, (ii) advertisers share data with different exchanges, or (iii) both advertisers share data with ADX_2.

For both advertisers to choose to share data with ADX_1, it must be that both (No Leave & Snipe) and (No Leave) hold, after modifying these conditions to account for the different values of w_1 and $w_2 = 0$. Such an equilibrium exists if and only if w_1 satisfies

$$w_1 \leq w^{cx} := \min\left\{\frac{\alpha(2C - M) + \eta C}{1+\eta}, C - \frac{\alpha}{\alpha+\eta}M\right\}, \tag{12}$$

where the first term is obtained from (No Leave) and the second one from (No Leave & Snipe). Note that for some parameters it may be that $w^{cx} \leq 0$, in which case it is clearly not optimal for ADX_1 to select such a price. Supposing, however, that $w^{cx} > 0$, then by setting $w_1 = w^{cx}$ ADX_1 can ensure that both advertisers share data with it.[29]

If instead ADX_1 chooses $w_1 > w^{cx}$, then the (subgame) equilibrium exhibits either zero advertisers or one advertiser sharing data with ADX_1.[30] For one advertiser to share data with ADX_1 it must also be that $w_1 \leq w^{nx}$, which is the same price that ADX_1 would optimally set if it did not allow cross-targeting.[31] But this means that any (subgame) equilibrium in which ADX_1 allows cross-targeting and in which $w_1 > w^{cx}$ earns ADX_1 strictly lower profits than what it can assure itself by not allowing cross-targeting and setting $w_1 = w^{nx}$.[32]

It follows that, if it is optimal for ADX_1 to allow cross-targeting, then it is also optimal to set $w_1 = w^{cx}$ and ensure that both advertisers share data with it. Hence, to determine when it is optimal to allow cross-targeting, we simply compare the profits from not allowing cross-targeting (in which case each advertiser buys $1 - 2\alpha + 2\eta$ ads from ADX_1 at price w^{nx}) to the profits from allowing cross-targeting (in which case each advertiser buys $1 + 2\eta$ ads from ADX_1 at price w^{cx}).

[29] In Lemma 7 in the Appendix we confirm that, for such $w_1 \leq w^{cx}$, the equilibrium in which both advertisers share data with ADX_1 is in fact Pareto-preferred by advertisers to any other equilibrium that might exist, and hence this equilibrium is selected for such values of w_1. It follows that, amongst all prices w_1 that lead to both advertisers sharing data with ADX_1, $w_1 = w^{cx}$ is optimal.

[30] This follows from Lemma 5 and the fact that, as just shown, for prices $w_1 > w^{cx}$, both advertisers sharing data with ADX_1 is not an equilibrium. Also, Lemma 6 shows that if one advertiser shares data with ADX_1 in the subgame equilibrium, then there is no cross-targeting. Thus, ADX_1's profits in this case come entirely from re-targeting by one advertiser and from selling ads for the look-alike audience.

[31] To see this, first note that, as shown in Lemma 6, if one advertiser shares data with ADX_1 in the subgame equilibrium, then there is no cross-targeting. To show that no price higher than w^{nx} can sustain such an equilibrium, consider the deviation incentive of the advertiser A_i sharing data with ADX_1. By sharing data with ADX_1, A_i receives a payoff of $\alpha M + (1-2\alpha)C + \eta C - (1-\alpha+\eta)w_1$. Instead deviating and sharing data with ADX_2 generates a payoff of $\alpha M + (1-2\alpha)C$. Algebra then shows that such a deviation is not profitable if and only if (10) holds, which w^{nx} is derived from. Thus, the definition of w^{nx} applies to the current circumstance, allowing us to conclude that $w_1 \leq w^{nx}$ if one advertiser is sharing data with ADX_1.

[32] To be entirely clear, this is because setting $w_1 = w^{nx}$ when cross-targeting is not allowed ensures both advertisers share data with ADX_1, whereas at most one advertiser shares data with ADX_1 when cross-targeting is allowed and $w^{nx} \geq w_1 > w^{cx}$.

That is, it is optimal to allow cross-targeting if and only if $(1-2\alpha+2\eta)w^{nx} \leq (1+2\eta)w^{cx}$, which is the case whenever the following holds:

$$\frac{(1-\alpha+2\eta)\eta}{1-\alpha+\eta}C \leq (1+2\eta)\min\left\{\frac{\alpha(2C-M)+\eta C}{1+\eta}, C-\frac{\alpha}{\alpha+\eta}M\right\}. \tag{13}$$

Notably, if ADX_1 does not allow cross-targeting in equilibrium, then $w^{nx} > w^{cx}$ because if this inequality failed, then ADX_1 would earn more from allowing cross-targeting (as it would earn a higher price per ad and also sell more ads). Examining inequality (13), we also see that the key comparative statics from the base model continue to hold. Specifically, cross-targeting becomes more likely if C increases, and less likely if M or α increase.[33] These results are intuitive. If α is larger, then each advertiser controls more exclusive data, forcing ADX_1 to cater to the advertisers' interests (and lower w^{cx}) to preserve their exclusive visitors. The effect of an increase in M is similar because an increase of monopoly profits also reduces each advertiser's willingness to share data with ADX_1 if ADX_1 allows cross-targeting. Lastly, an increase in C means that advertisers benefit more from cross-targeting and lose less if rivals cross-target their exclusive visitors (because $M-C$ is smaller), which allows ADX_1 to set a higher w_1 when allowing cross-targeting.

The analysis above indicates that the same economic forces are at play when ad exchanges set prices as was the case when prices were exogenous (in the base model). In the base model, when advertisers' incentives to secure monopoly profits from its exclusive visitors are strong, they do not share their visitors' data with ADX_1 if cross-targeting is allowed, lowering ADX_1's profits and therefore discouraging ADX_1 from allowing cross-targeting. With endogenous prices, ADX_1 can always convince advertisers to share their data, but only by sufficiently lowering w_1. When the required w_1 is too low, ADX_1 prefers not to allow cross-targeting because doing so results in lower profits—just as is the case when prices are exogenous.

To complete our discussion of the case with endogenous prices, we next investigate whether the advertisers' behavior leads to an inefficiency and a prisoner's dilemma for advertisers. To do so, we look for a condition such that advertisers jointly prefer cross-targeting but it does not occur in equilibrium. Note that, holding $w_1 = w^{nx}$ fixed, advertisers are jointly better off under full cross-targeting on ADX_1 if and only if $2C - M > w^{nx}$. Recall that $w^{nx} > w^{cx}$ if ADX_1 does

[33]The comparative statics for M follow immediately from (13). Concerning C, divide everything by C; then the right-hand side increases in C while the left-hand side is constant. To understand the comparative statics regarding α, suppose first that the second term on the right-hand side is the smaller one. Then, taking the derivative shows that the left-hand side increases in α whereas the right-hand side decreases, making it less likely that (13) continues to hold if α increases. Suppose next that the first term within the minimum function is the smaller one. In this case, bring everything on the left-hand side, which must now be less than zero if (13) holds. After multiplying by both denominators and collecting terms, notice that increasing α changes the left-hand side by $(2C-M)(1+2\eta) > 0$. Thus, raising α makes it less likely that (13) continues to hold.

not allow cross-targeting in equilibrium. Thus, if advertisers jointly prefer full cross-targeting at $w_1 = w^{nx}$, then they prefer it at a price of w^{cx} as well. Combining $2C - M > w^{nx}$ with (13), we thus obtain a sufficient condition for when cross-targeting does not emerge in equilibrium even though advertisers would jointly prefer it:

$$2C - M \geq \frac{\eta}{1-\alpha+\eta} C > \frac{1+2\eta}{1-\alpha+2\eta} \min\left\{ \frac{\alpha(2C-M)+\eta C}{1+\eta}, C - \frac{\alpha}{\alpha+\eta} M \right\}. \tag{14}$$

This sufficient condition holds in a number of cases. For example, consider the specification with $M = 1$ and $\eta = 0.2$. Algebra reveals that, for any $\alpha > 0.18$ there is always a range of C such that this condition holds, and the size of this region grows as α does. For example, at $\alpha = 0.25$, this condition holds whenever $C \in [0.56, 0.67]$, whereas at $\alpha = 0.4$, it holds whenever $C \in [0.58, 0.81]$. Additionally, outside of this specification, this condition always holds so long as η is sufficiently small. This is because if the look-alike audience is small, ADX_1 needs to charge a negative price for its ads to convince advertisers to share their data if cross-targeting is allowed.[34] Thus, there is no cross-targeting if η is small even though advertisers may well prefer it. This stands in contrast to the base model, where cross-targeting can emerge despite a small look-alike audience. The reason for this difference is the competitive ad price of ADX_2 in this extension, which prevents ADX_1 from making positive profits if it allows cross-targeting for small η.

5. Conclusion

In this paper, we build a model to analyze the economics of a particular type of targeted online advertising. Consumers who have shown intent to purchase (for instance by visiting a website) in some product category are targeted by advertisers. In equilibrium, ad exchanges may adopt data-sharing policies that give rise to a fundamental inefficiency: because of advertisers' strong property rights too little data about the consumers' purchase intentions is shared and, as a consequence, mutually beneficial market transactions do not take place. This prediction is consistent with real-world observations that ad exchanges do not facilitate cross-targeting, that is do not allow advertisers to directly target visitors of a rival advertiser's website.

We argue that the current industry model grants excessive data property rights to advertisers since the ability of ad exchanges to track consumers generally requires consent by website owners. An ad exchange may respond to this by designing restrictive data-sharing policies, so as to entice

[34] In terms of (14), when η is small, the left-hand term is positive, the middle term is positive but close to zero, and the right-hand term is negative because $C < M$.

advertisers to work with it. The advertisers' strong property rights are not confined to the present third-party cookie system but are also inherent to new third-party tracking systems that continue to require advertiser consent such as digital fingerprinting, pixel tracking, or Unified 2.0. We show that granting consumers property rights over data reflecting their intent to purchase such as the ability to opt out of third-party tracking can mitigate the issue. Likewise, moves by companies such as Apple and Google that circumvent the third-party tracking system—widely criticized as a power grab by advertisers—may actually benefit both consumers and advertisers by weakening advertisers' property rights. On the other hand, category-based tracking (such as Google's Topics system) may also increase the number of excess impressions that do not benefit consumers or advertisers, but nonetheless generate revenue for the ad exchange.

Appendix: Omitted Proofs

To simplify exposition in the proofs below, we sometimes use the term "open" to mean that ADX_1 has allowed cross-targeting, and "closed" to mean that it has not. Moreover, because the fee ϕ charged by ADX_1 is a scalar that multiplies advertising revenue generated through targeted ads, we henceforth omit ϕ or, equivalently, set $\phi = 1$ without loss of generality.

The proofs of the next two lemmas follow from arguments in the main text.

Lemma 1. *If ADX_1 is closed, both advertisers share data with ADX_1 in the unique equilibrium.*

Lemma 2. *Suppose ADX_1 is open. Then, full cross-targeting is an equilibrium if and only if both* (No Leave) *and* (No Leave & Snipe) *holds.*

Lemma 3. *Suppose ADX_1 is open. If full cross-targeting is an equilibrium, there is no other equilibrium in which at least one advertiser shares data with ADX_1.*

Proof of Lemma 3: As shown in the main text, if both advertisers share data with ADX_1 in stage 2, then it always leads to full cross-targeting in stage 3. Thus, to prove the result, we can restrict attention to ruling out equilibria in which exactly one advertiser, say A_i, shares data with ADX_1. Suppose that A_{-i} shares data with ADX_2 and cross-targets A_i's visitors. Then, A_{-i} earns profits of $\alpha M + (1-\alpha)C - 2(1-\alpha)w + \eta_1(C-w)$. By deviating and sharing data with ADX_1, A_{-i} would induce full cross-targeting and earn profits of $(1+\eta_2)(C-w)$. Thus, deviating is profitable if $\alpha M + (1-\alpha)C - 2(1-\alpha)w + \eta_1(C-w) \leq (1+\eta_2)(C-w)$ or, equivalently, if (No Leave & Snipe) holds. Since full cross-targeting is an equilibrium by assumption, Lemma 2 implies that (No Leave & Snipe) does hold, ruling out that A_{-i} cross-targets in equilibrium.

Suppose next that A_{-i} shares data with ADX_2 but does not cross-target A_i's visitors. That is, A_{-i} only retargets on ADX_2 and targets look-alike audiences on ADX_1, implying profits of $\alpha M + (1-2\alpha)C - (1-\alpha)w + \eta_1(C-w)$. By deviating and sharing data with ADX_1, A_{-i} would induce full cross-targeting, leading to profits of $(1+\eta_2)(C-w)$. Thus, the deviation is profitable if $\alpha M + (1-2\alpha)C - (1-\alpha)w + \eta_1(C-w) \leq (1+\eta_2)(C-w)$ or, equivalently, if (No Leave) holds. Since full cross-targeting is an equilibrium by assumption, Lemma 2 implies that (No Leave) does hold, ruling out that this can be an equilibrium. ■

Lemma 4. *Suppose ADX_1 is open. If full cross-targeting is an equilibrium, then it strictly Pareto-dominates any other equilibrium for advertisers.*

Proof of Lemma 4: If full cross-targeting is an equilibrium, (No Leave) implies that

$$(1+\eta_2)(C-w) \geq \alpha M + (1-2\alpha)C - (1-\alpha)w + \eta_1(C-w), \tag{15}$$

where the left-hand side represents each A_i's profits in the full cross-targeting equilibrium. By Lemma 3, both advertisers must share data with ADX_2 in the only alternative equilibrium that may exist. In this equilibrium, each A_i earns $\alpha M + (1-2\alpha)C - (1-\alpha)w$, which involves no look-alike audience targeting since look-alike audiences cannot be identified if no A_i shares data with ADX_1. The profits are thus strictly less than the expression on the right-hand side of (15) and, thus, less than profits with full cross-targeting. ■

Lemma 5. *If ADX_1 is open, there is always an equilibrium in pure strategies.*

Proof of Lemma 5: The claim follows if the data sharing game played by advertisers is a symmetric 2 strategy game, which must have a pure strategy Nash equilibrium always by Cheng, Reeves, Vorobeychik, and Wellman (2004). As argued in the main text, if both advertisers share data with

ADX_1, then full cross-targeting ensues, implying identical payoffs for advertisers. Likewise, their payoffs are identical if both share data with ADX_2 and only re-target their own visitors.

Finally, we establish that the off-diagonal payoffs of the normal form game are well-defined and symmetric. The payoffs resulting from asymmetric data sharing decisions depend on whether the advertiser that shares data with ADX_2, say A_i, cross-targets A_{-i}'s visitors via ADX_1 or not. This decision, in turn, is uniquely pinned down. Specifically, A_i prefers cross-targeting A_{-i}'s visitors over not doing so if and only if

$$\begin{aligned} \alpha M + (1-\alpha)C - 2(1-\alpha)w + \eta_1(C-w) \geq& \alpha M + (1-2\alpha)C - (1-\alpha)w + \eta_1(C-w) \\ \iff \quad \alpha C \geq& (1-\alpha)w \end{aligned} \tag{Sniping}$$

If (Sniping) holds and A_i shares data with ADX_2 whereas A_{-i} shares data with ADX_1, then A_i earns unique profits of $\alpha M + (1-\alpha)C - 2(1-\alpha)w + \eta_1(C-w)$ while A_{-i} earns unique profits of $(1-\alpha+\eta_1)(C-w)$. Analogously, if (Sniping) holds and A_i shares data with ADX_1 whereas A_{-i} shares data with ADX_2, the payoffs are reversed. Lastly, if (Sniping) fails, then each advertiser retargets its website visitors via the ad exchange it shares data with and also targets the look-alike audience of size η_1. Thus, the payoffs are uniquely pinned down and symmetric in this case as well.

Consequently, whether (Sniping) holds or not, the data sharing game played by advertisers is a symmetric 2 strategy game and thus must have a Nash equilibrium in pure strategies. Note that the same logic applies if we allow w_1 and w_2 to differ, as in an extension of ours in Section 4, and hence we can also conclude that a pure strategy subgame equilibrium exists in that case as well. ■

Proof of Proposition 1: By Lemma 1, both advertisers share data with ADX_1 if ADX_1 is closed. Total profits of ADX_1 are thus

$$2w(1-\alpha+\eta_2). \tag{16}$$

Next, consider profits if ADX_1 is open and assume that both (No Leave & Snipe) and (No Leave) hold. By Lemma 2, both advertisers sharing data with ADX_1 is an equilibrium in this case. By Lemma 4, this equilibrium strictly Pareto-dominates (for advertisers) any other equilibrium that may exist. Thus, Assumption 1 allows us to to restrict attention to the full cross-targeting equilibrium, in which total profits for ADX_1 are equal to

$$2w(1+\eta_2). \tag{17}$$

Since this is strictly higher than (16), ADX_1 allows cross-targeting if both (No Leave & Snipe) and (No Leave) hold, proving the "if" part.

For the "only if" part, assume that ADX_1 is open but that (No Leave & Snipe) or (No Leave) fail. Lemma 5 guarantees that a pure strategy equilibrium still exists, though at most one advertiser may share data with ADX_1 or else there would be full cross-targeting, contradicting Lemma 2. Thus, profits of ADX_1 are bounded from above by

$$2w(1-\alpha+\eta_1), \tag{18}$$

which equals ADX_1's profits if one advertiser shares data with ADX_1, the other cross-targets on ADX_1, and both advertisers target η_1 look-alike consumers. This is less than the profits from being closed as given in (16) because $\eta_2 > \eta_1$. Thus, ADX_1 chooses closed if (No Leave & Snipe) or (No Leave) fail, proving the "only if" part of the Proposition.

The remainder of the proposition follows from the discussion in the main text and the proof above. If ADX_1 allows cross-targeting as (No Leave & Snipe) and (No Leave) hold, both advertisers share data with ADX_1 by Lemma 2, implying a look-alike audience size of η_2. Moreover, if ADX_1 does not allow cross-targeting, both advertisers share data with ADX_1 as well according to Lemma 1, again implying a look-alike audience size of η_2. ■

Proof of Proposition 2: Given the similarity to Proposition 1, we keep this proof short. Note that Lemma 4 and Lemma 5 readily extend to the case in which consumers can block tracking.

Consequently, Assumption 1 implies that being open always leads to full cross-targeting if and only if both (C: No Leave) and (C: No Leave & Snipe) hold. Thus, choosing open if both (C: No Leave) and (C: No Leave & Snipe) hold leads to profits for ADX_1 of

$$2w(q+\eta_2). \tag{19}$$

Next, we must consider the outcome if ADX_1 chooses open and full cross-targeting is not an equilibrium. This subgame has a pure strategy equilibrium since the logic of Lemma 5 still holds. Furthermore, the best possible outcome for ADX_1 if it chooses open but full cross-targeting is not an equilibrium involves A_i sharing data with and retargeting on ADX_1, A_{-i} cross-targeting, and both targeting η_1 look-alike consumers. In this case, allowing cross-targeting leads to profits of

$$2w((1-\alpha)q+\eta_1). \tag{20}$$

Lastly, as argued in the main text, both advertisers share data with ADX_1 if it chooses closed, implying profits of

$$2w((1-\alpha)q+\eta_2). \tag{21}$$

As in the base model, it is easy to verify that the profits of ADX_1 if it allows cross-targeting and if full cross-targeting is an equilibrium exceed its profits if it is closed, which, in turn, exceed its profits if ADX_1 is open but full cross-targeting is not an equilibrium. Analogously to the proof of Proposition 1, this is is sufficient to establish that ADX_1 allows cross-targeting if and only if (C: No Leave) and (C: No Leave & Snipe) hold, in which case it leads to full cross-targeting. ■

Proof of Proposition 3: The result follows if the term in (5) decreases in q upon substituting $\eta_2 = \eta(q)$ and $\eta_1 = \eta(q(1-\alpha))$. Let $\eta'(x)$ denote the derivative of η with respect to x. Then, the derivative of $\dfrac{\eta(q)-\eta(q(1-\alpha))}{q}(C-w)$ with respect to q equals

$$\frac{\left[\eta'(q)-(1-\alpha)\eta'((1-\alpha)q)\right]q-\left[\eta(q)-\eta((1-\alpha)q)\right]}{q^2}(C-w), \tag{22}$$

which is negative if and only if

$$\eta'(q)q - \eta(q) < (1-\alpha)q\eta'((1-\alpha)q) - \eta((1-\alpha)q). \tag{23}$$

Since $q > (1-\alpha)q$, the inequality above is implied if $\eta'(x)x - \eta(x)$ decreases in x. Taking the derivative shows that $\eta'(x)x - \eta(x)$ decreases in x if and only if $\eta''(x) < 0$ or, equivalently, if η is concave. ■

Proof of Proposition 4 : Proposition 4 directly follows from the discussion on page 18. ■

Proof of Proposition 5: Only if: If both advertisers bid on ADX_1, their respective payoffs equal $(1+\eta_2)(C - \tau w)$, while a unilateral deviation to the ADX_2 accrues $\sigma(1-\alpha)(C-w)$. It follows that both advertisers bidding on ADX_1 constitutes an equilibrium if $w < w' \equiv \frac{(1+\eta_2)-\sigma(1-\alpha)}{\tau(1+\eta_2)-\sigma(1-\alpha)} C \Leftrightarrow \tau < \frac{(1+\eta_2)-\sigma(1-\alpha)}{(1+\eta_2)w} C + \frac{\sigma(1-\alpha)}{1+\eta_2}$. If: Furthermore, note that if both advertisers bid on the ADX_2, their respective payoffs equal $\sigma[\alpha M + (1-2\alpha)C - (1-\alpha)w]$. When playing asymmetric strategies, the advertiser bidding on ADX_1 earns $(1+\eta_2)M - \sigma(1-\alpha)(M-C) - (1+\eta_2)\tau w$, while the advertiser bidding on the ADX_2 earns $\sigma(1-\alpha)(C-w)$. It follows that both advertisers bidding on the ADX_2 constitutes an equilibrium only if $w > w'' \equiv \frac{[(1+\eta_2)-\sigma]M+\alpha\sigma C}{\tau(1+\eta_2)-\sigma(1-\alpha)}$, while asymmetric strategies do not exhibit a profitable unilateral deviation if $w' \leq w \leq w''$. Since $w'' = \frac{[(1+\eta_2)-\sigma]M+\alpha\sigma C}{\tau(1+\eta_2)-\sigma(1-\alpha)} = \frac{(1+\eta_2)-\sigma(1-\alpha)}{\tau(1+\eta_2)-\sigma(1-\alpha)} + \frac{1+\eta_2-\sigma}{\tau(1+\eta_2)-\sigma(1-\alpha)}(M-C) > w'$, it follows that if $\tau < \frac{(1+\eta_2)-\sigma(1-\alpha)}{(1+\eta_2)w} C + \frac{\sigma(1-\alpha)}{1+\eta_2}$ the unique equilibrium has both advertisers bid on ADX_1. Finally, using the fact that $C > w$, simple algebra shows that $\tilde{\tau} > 1$. ■

Proof of Proposition 6: Recall from Proposition 1 that if full-cross targeting does not obtain in the base model, both advertisers share data with ADX_1, retarget their own website visitors and target the look-alike audience of size η_2. As such, $1 - 2\alpha + \eta_2$ consumers receive two targeted ads while the 2α exclusive website visitors receive a single targeted ad. Under full cross-targeting, on the other hand, everyone receives two targeted ads. Thus, the total benefit of consumers interested in the product category when moving from an overall retargeting to an overall full cross-targeting equilibrium is $2\alpha(u_2 - u_1)$. Under full cross-targeting in the Topics model $\tau(1+\eta_2)$ consumers receive both ads, a fraction $\tau - 1$ of which is not interested in the product category. It follows that the total cost of introducing the Topics system is $(\tau-1)(1+\eta_2)2\kappa$. The claim follows from

$$2\alpha(u_2 - u_1) > (\tau - 1)(1+\eta_2)2\kappa \Leftrightarrow \tau < 1 + \frac{\alpha}{1+\eta_2}\frac{u_2 - u_1}{\kappa}.$$

■

Proof of Proposition 7: (i) There are $\tau(1+\eta_2)$ consumers under consideration. In the region where both advertisers bid on ADX_1, ADX_1 is selling two targeted ads for the $1+\eta_2$ consumers interested in the product category, and for the $(\tau-1)(1+\eta_2)$ who are not. It follows that ADX_1 earns $\tau(1+\eta_2)2w$ from selling targeted ads, which is clearly increasing in τ. (ii) Note that

$$\frac{\partial\tilde{\tau}}{\partial\sigma} = \frac{1-\alpha}{1+\eta_2}\left(1 - \frac{C}{w}\right) < 0,$$

which proves the claim. ■

Proof of Corollary 2: Let

$$w > w'' = \frac{[(1+\eta_2) - \sigma]M + \alpha\sigma C}{\tau(1+\eta_2) - \sigma(1-\alpha)} \Leftrightarrow \tau > \frac{[(1+\eta_2) - \sigma]M + \alpha\sigma C}{w(1+\eta_2)} + \frac{\sigma(1-\alpha)}{1+\eta_2},$$

and recall from the proof of Proposition 5 that this condition implies the equilibrium under the Topics system to have advertisers retargeting on ADX_2, which means ADX_1 earns zero profits. By Proposition 1, the worst possible outcome under the system without Topics is for ADX_1 to not allow cross-targeting, which nonetheless ensures that both advertisers retarget their own website visitors, as well as the η_2 look-alike consumers, on ADX_1. As such, if τ is large enough such that $w > w''$, ADX_1 prefers not to adopt the Topics system. ■

Lemma 6. *In the model extension in which ad prices are endogenous, suppose ADX_1 allows cross-targeting. Then, for a given value of $w_1 > 0$, if it is an equilibrium for A_i to share data with ADX_1 and for A_{-i} to share data with ADX_2, then there is no cross-targeting.*

Proof of Lemma 6: Suppose there is cross-targeting in equilibrium when the advertisers share data with different ad exchanges. Then, the advertiser sharing data with ADX_1 must not want to deviate by sharing data with ADX_2 instead, which would prevent being cross-targeted. Formally, $(1 - \alpha + \eta)(C - w_1) \geq \alpha M + (1 - 2\alpha)C$ or, equivalently,

$$(1 + \eta)(C - w_1) \geq \alpha M + (1 - 2\alpha)C + \alpha(C - w1). \tag{24}$$

In addition, the advertiser sharing data with ADX_2 must not want to deviate by sharing data with ADX_1, which would lead to full cross-targeting. Formally, $\alpha M + (1 - \alpha)C - (1 - \alpha)w_1 + \eta(C - w_1) \geq (1 + 2\eta)(C - w_1)$ or, equivalently,

$$\alpha M + (1 - \alpha)C - (1 - \alpha)w_1 \geq (1 + \eta)(C - w_1). \tag{25}$$

However, taken together, inequalities (24) and (25) imply

$$\alpha M + (1 - \alpha)C - (1 - \alpha)w_1 \geq \alpha M + (1 - 2\alpha)C + \alpha(C - w1) \Leftrightarrow (1 - 2\alpha)(C - w_1) \geq (1 - 2\alpha)C,$$

a contradiction for any $w_1 > 0$. ■

Lemma 7. *In the model extension in which ad prices are endogenous, suppose ADX_1 allows cross-targeting. If $w^{cx} > 0$ and $w_1 \in (0, w^{cx}]$, then both advertisers sharing data with ADX_1 strictly Pareto-dominates any other equilibrium that exists.*

Proof of Lemma 7: By construction of w^{cx}, $w_1 \in (0, w^{cx}]$ ensures that both advertisers sharing data with ADX_1 and engaging in full cross-targeting is an equilibrium. We first show that any potential equilibrium in which both advertisers share data with ADX_2 is strictly Pareto-dominated by the full cross-targeting equilibrium. This follows from the modified (No Leave) condition, which holds by construction because $w_1 \leq w^{cx}$ and implies that

$$(1 + 2\eta)(C - w_1) \geq \alpha M + (1 - 2\alpha)C + \eta(C - w_1), \tag{26}$$

where the term on the left-hand side depicts each advertiser's full cross-targeting equilibrium profits. If both advertisers share data with ADX_2, by contrast, they each earn $\alpha M + (1 - 2\alpha)C$, which is strictly less than the term on the right-hand side in the inequality above.

We next show that any alternative equilibrium in which one advertiser, say A_i, shares data with ADX_2 is strictly Pareto dominated. Note that by Lemma 6, we can rule out any cross-targeting in any such equilibrium. Again using the fact that the modified (No Leave) holds, the advertiser sharing data with ADX_2 would be weakly better off if it instead shared data with ADX_1. Note that the effect of this switch to ADX_1 on its profits comes from (i) full cross-targeting occurring rather than only re-targeting, (ii) an increased look-alike audience size, and (iii) an increase in its advertising costs from zero to w_1. Because the advertiser weakly prefers to make this switch, the other advertiser (on ADX_1) is strictly better off as a result. The reason is that its change in profits includes items (i) and (ii) just above, but not the higher costs in (iii). Hence any equilibrium with advertisers using different exchanges is stirctly Pareto-dominated by both using ADX_1. ■

References

Agrawal, A., J. Gans, and A. Goldfarb (2018): *Prediction machines: the simple economics of artificial intelligence.*

Athey, S., E. Calvano, and J. S. Gans (2018): "The impact of consumer multi-homing on advertising markets and media competition," *Management Science*, 64(4), 1574–1590.

Bergemann, D., and A. Bonatti (2011): "Targeting in Advertising Markets: Implications for Offline versus Online Media," *RAND Journal of Economics*, 42, 417–443.

——— (2015): "Selling Cookies," *American Economic Journal: Microeconomics*, 7, 259–294.

Bergemann, D., A. Bonatti, and T. Gan (2022): "The economics of social data," *The RAND Journal of Economics.*

Buchbinder, N., M. Feldman, A. Ghosh, and J. Naor (2014): "Frequency capping in online advertising," *Journal of Scheduling*, 17, 385–398.

Chen, Y., and G. Iyer (2002): "Consumer Addressability and Customized Pricing," *Marketing Science*, 21, 197–208.

Cheng, S.-F., D. M. Reeves, Y. Vorobeychik, and M. P. Wellman (2004): "Notes on equilibria in symmetric games," .

Choi, H., C. F. Mela, S. R. Balseiro, and A. Leary (2020): "Online Display Advertising Markets: A Literature Review and Future Directions," *Information Systems Research*, 31, 556–.

Choi, J. P., D.-S. Jeon, and B.-C. Kim (2019): "Privacy and Personal Data Collection with Information Externalities," *Journal of Public Economics*, 173, 113–124.

Crémer, J., P. Rey, and J. Tirole (2000): "Connectivity in the commercial Internet," *The Journal of Industrial Economics*, 48(4), 433–472.

De Corniere, A. (2016): "Search advertising," *American Economic Journal: Microeconomics*, 8(3), 156–88.

De Corniere, A., and R. De Nijs (2016): "Online Advertising and Privacy," *RAND Journal of Economics*, 47, 48–72.

Dosis, A., and W. Sand-Zantman (2022): "The Ownership of Data," *The Journal of Law, Economics, and Organization*.

D'Annunzio, A., and A. Russo (2020): "Ad networks and consumer tracking," *Management Science*, 66(11), 5040–5058.

Fudenberg, D., and J. Tirole (2000): "Customer Poaching and Brand Switching," *RAND Journal of Economics*, 31, 634–657.

Goldfarb, A., and C. E. Tucker (2011): "Privacy Regulation and Online Advertising," *Management Science*, 57, 57–71.

Hagiu, A., and B. Jullien (2011): "Why do intermediaries divert search?," *The RAND Journal of Economics*, 42(2), 337–362.

Hermalin, B. E., and M. L. Katz (2006): "Privacy, property rights and efficiency: The economics of privacy as secrecy," *Quantitative marketing and economics*, 4(3), 209–239.

Ichihashi, S. (2020): "Online Privacy and Information Disclosure," *The American Economic Review*, 110, 569–595.

——— (2021): "Competing Data Intermediaries," *RAND Journal of Economics*, 52.

Iyer, G., D. Soberman, and M. Villas-Boas (2005): "The Targeting of Advertising," *Marketing Science*, 24, 461–476.

Johnson, J. P. (2013): "Targeted Advertising and Advertising Avoidance," *RAND Journal of Economics*, 44, 128–144.

Kraemer, J., D. Schnurr, and M. Wohlfarth (2020): "Winners, Losers, and Facebook: The Role of Social Logins in the Online Advertising Ecosystem," *Management Science*, 65, 1678–1699.

Markovich, S., and Y. Yehezkel (2021): ""For the public benefit": who should control our data?," *Available at SSRN 3945108*.

Montes, R., W. Sand-Zantman, and T. Valletti (2019): "The Value of Personal Information in Online Markets with Endogenous Privacy," *Management Science*, 65, 1342–1362.

Sayedi, A. (2018): "Real-Time Bidding in Online Display Advertising," *Marketing Science*, 37.

Statista (2022): "Conversion rate of online shoppers in the United States from 2nd quarter 2021 to 2nd quarter 2022," www.statista.com/statistics/439558/us-online-shopper-conversion-rate/.

Taylor, C. R. (2004): "Privacy and the Market for Customer Information," *RAND Journal of Economics*, 35, 631650.

techcrunch.com (2022): "Google kills off FLoC, replaces it with Topics," www.techcrunch.com/2022/01/25/google-kills-off-floc-replaces-it-with-topics.

Ursu, R. M., Q. Zhang, and E. Honka (2022): "Search Gaps and Consumer Fatigue," *Marketing Science*.

Villas-Boas, J. M., and Y. J. Yao (2021): "A Dynamic Model of Optimal Retargeting," *Marketing Science*, forthcoming.

Villas-Boas, M. (1999): "Dynamic Competition with Customer Recognition," *RAND Journal of Economics*, 30, 604–631.

——— (2004): "Price Cycles in Markets with Customer Recognition," *RAND Journal of Economics*, 35, 486–501.

vox.com (2022): "The winners and losers of Apple's anti-tracking feature," www.vox.com/recode/23045136/apple-app-tracking-transparency-privacy-ads.

wired.com (2022): "Google Has a New Plan to Kill Cookies. People Are Still Mad," https://www.wired.com/story/google-floc-cookies-chrome-topics/.