Browser Tying and Data Privacy Innovation

Stephen Dnes[1]

Abstract

This paper explores the role of the browser in relation to competition for advertising. It explores the important role of the browser in collecting and transmitting data to servers via data storage and transmission functions, the mechanisms by which these data flows are enabled, and some of the competing uses for which these data flows are used on the server side. This reveals the need for server-side processing of some information.

In this context, profound concerns are raised that technological tying will imminently undermine competition and innovation in digital markets via browser-based restrictions. Indeed, there are proposals for de facto vertical integration of significant browser and server functionality under proposals from Apple and Google, which will strongly favour vertically integrated solutions (the so-called "walled gardens"). Both proposals are restrictive, but Google's appear to go even further than Apple in its tying proposals, without justification.

The paper concludes with some possible remedies to prevent anti-competitive technological tying of browser and server functionality. Indeed, there is an acute need to use established competition law tools to prevent anti-competitive foreclosure from locking down the browser.

Introduction

In the increasingly prominent "Big Tech" antitrust cases, there is increasing sensitivity to issues with pre-installation and exclusivity agreements denying scale to rival operations. Less noted, but perhaps even more important, is a set of proposals known as the Google Privacy Sandbox. These build on earlier proposals to retire third-party cookies, in the name of privacy. However, on a closer examination, they amount to anti-competitive technological tying of user information into browsers, with sigificant negative implications for competition and innovation in online markets.

There is early recognition of this issues on the part of some regulators. For instance, in the words of the UK CMA, there is a serious risk that the main Privacy Sandbox proposals will turn "Chrome (or Chromium browsers) into the key bottleneck for ad tech."[2] Most recently, the UK CMA published commitments proposals designed to address the concerns, which are currently out for consultation.[3] There is also sensitivity to the issue in a number of the U.S. Google cases; especially, *Texas v. Google*.[4] However, there is limited analysis of the tying issues involved; and limited analysis of the weighing of data privacy and competition concerns.[5] This paper seeks to further the analysis of privacy concerns under competition law by considering: (I) The role of the browser; (II) Limitations to the browser and the competitive implications of these; (III) Market power on the part of browser vendors; (IV) Proposed technological ties; (V) Analysis of these ties under tying law; and (VI) the absence of justification from a privacy point of view.

In summary, the paper finds justification for tying wanting, there being no robust case for restriction of *pseudonymous identifiers* unless *identity* is disclosed. In the context of high market shares, apparent barriers to entry, and anti-competitive pre-installation agreements for browsers and related search and

---

[1] Assistant Professor and Director of International Programmes, Northeastern University, London; Lecturer (Part-Time), Northeastern University School of Law, Boston. The author acted as an expert consultant to Marketers for an Open Web, a complainant before the UK CMA ("CMA") in its recent investigation into Google's Privacy Sandbox. The paper reflects experience of working with affected publishers and advertisers including several small businesses, but the positions taken are those of the author and the usual disclaimer applies.

[2] CMA Online Platforms and Digital Advertising Market Study, Final Report, hereafter *CMA Final Report*, para 5.327.

[3] CMA, Consultation on proposed commitments in respect of Google's 'Privacy Sandbox' browser changes, 11 June 2021.

[4] *Texas et al v. Google Inc.*, (E.D. Tex., 2020)

[5] For an early example of the analysis of the competitive effects of the GDPR, see Michal S Gal, Oshrit Aviv, The Competitive Effects of the GDPR, *Journal of Competition Law & Economics*, 16(3) September 2020, 349–391.

advertising functions, there is every reason to be careful and to apply a subtractive tying remedy by which users would be empowered to choose between the level of privacy protection and the ad monetisation on offer from publishers and advertisers. The paper concludes that this should not be the preserve of a dominant browser provider.

I. What do browsers do? The role of the browser

A web browser might simply be thought of as a window into the web. This is, however, a dangerous oversimplification. In addition to website rendering and display, the browser plays a central role in data transmission. Indeed, to present a website, much information must pass from the user to servers, including:

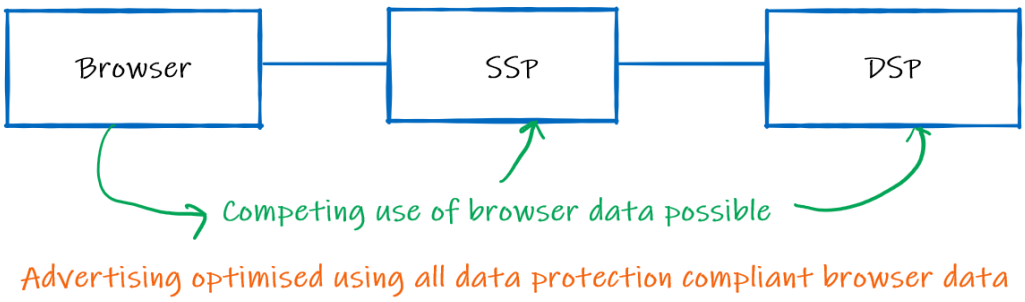
- Which site is to be accessed (www.example.com);
- Information about the browser to allow correct rendering (e.g. display size; mobile or desktop);
- Information about the user (language, location, etc.).

Far from being a dumb window, the browser is primarily an information conduit. This is always a two-way process as the browser communicates with the server.

The browser stores data as part of this process. The data in the browser is essential for offering targeted advertising. Stored text files of data known as cookies provide the best-known example; although there are others. For current purposes, one of the most important, besides cookies, is the User Agent String, which enables considerable competitive advertising technology.[6]

These data sources allow identification of a user and knowledge of user attributes to pass to the server. For targeted advertising, the process can be visualised in a slightly simplified format as follows:

Figure 1: Browser links to servers for targeted advertising



The browser passes information into the server stack (SSPs and DSPs; hereafter, "server-side"). The servers pass an advertisement back to the browser based on the information provided to them by it.[7]

This information allows important context and forecasting services to work. A simple example is a weather widget: the weather widget needs to know where the user is to display the correct weather forecast.

[6] Google, Privacy Sandbox announcement, available at https://www.chromium.org/Home/chromium-privacy/privacy-sandbox (outlining a range of changes to information handling by the browser, including WebID; User Agent string degradation; X Client Data discrimination; Cohort management proposals; Cookies handling changes and the abolition of third-party cookies; Trust tokens; and the Privacy Budget).
[7] This is a simplification for ease of exposition. In practice, publisher ad servers, data management systems, and ad exchanges interact in a more complex set of interactions. However, the core characteristic is information flow from the *viewer* (browser) to the *publisher* (supply side) to the *advertiser* (demand side).

For advertising deployments, the browser-based data takes on particular significance. The richer the data, the better the advertisement targeting. Instead of targeting to *Users who have visited a shoe website* a local shoe shop could target *People in a 20-mile radius who shop in person and who visited a shoe website three times in the past week*. Targeted advertising is useful to viewers, content publishers, and advertisers alike: for example, the shoe buyer who did not know about a shoe store in the locality. In principle, the richer the data the more value is provided to all system users.

This is especially important for smaller users (e.g. a specialist blog), who depend critically on occasional highly targeted, high value advertisement revenue from relatively small traffic flows, known as "spikes", rather than large volume branded sites (e.g. the New York Times). The latter can advertise to a large volume of users, but the smaller sites rely more on specialist advertising, although the large sites also benefit from targeted advertising for their non-homepage content. Even a large, brand-name website benefits from targeted access for its less prominent inventory (e.g. non home page articles).

For these highly targeted uses, there is a margin of targeting which advances and recedes based on the richness of the data set. Contextual data allows critical uses like forecasting, which helps to deploy advertising budget and increases revenue to publishers. Moreover, the data stream from the browser enables competing processing of information about the advertising campaigns to optimise them. These sales attribution and advertising retargeting services are currently competitive, with significant competition on algorithm design. However, they require browser-based data to work.

There is no suggestion that targeting of advertisements should be banned or undermined except to the extent specific, existing laws regulate (e.g. GDPR; CCPA); except, perhaps, by those with a conflict of interest in limiting access to the data sets by rivals. Nonetheless, severe restrictions on targeting are precisely the outcomes most likely from proposals to limit access to certain browser-based information.

II. What *can't* browsers do? The limits of the browser

Before moving on to proposed restrictions, it should be noted that much debate assumes that retargeting services can operate using local processing in the browser. This is not true. The best and richest data processing is resource intensive and requires server-side processing. It may not even be possible for the browser to undertake the processing required for optimal re-targeting services.

For example, the relevant Apple proposals outlined below include a 64-category limit to identification tags, in part because of browser computing restrictions. For the best retargeting, significant computing is required. A large advertising campaign might need hundreds or even thousands of characteristics to be compared and correlated against a detailed profile: that is, multi-variate comparisons between which campaigns have been seen, which websites visited, etc. Provided this is done pseudonymously, it should not raise privacy concerns. It allows significant value to be added by targeting (e.g., information that the *third* website in four visited is the most predictive of a purchase – implying a handsome reward for the publisher of that content). This type of information optimisation is well beyond the processing limits in user-facing devices.

It is unclear how the data tying proposals in the Sandbox address this issue. In practice, the need for server-side processing will not be eliminated. Thus any reference to browser-based processing may necessarily imply a less rich and less competitive targeting ecosystem, with all this implies for depressed payments to content producers.

III. Market power in browsers: the tying product

Under the proposals, the browser will increase control over the data flows in the browser. It is necessary to determine whether there is market power in this data source. Many factors raise significant concerns that the browser and the data collected within it confer significant market power.

a. Data: scope

The data that is captured by the browser is extensive. Much more data is captured that is not personal data. For example, the User Agent String and Ex-Client Data are coded strings of data that are not themselves personal data at all.

Scale is also derived from market share in terms of browser usage and collection via Google's systems. Chrome accounts for 63.6% of worldwide browser use, and the two-firm concentration ratio in browser use is 83.4%, suggesting serious risks of duopoly tipping towards an anti-competitive equilibrium.[8]

As the default search engine in Chrome and Safari, Google obtains access to search histories on both Apple and Android. In the case of Apple, the search bar provides data from pre-installation of Google search. Google's share of worldwide search is 91.8% and the nearest rival is Bing with 2.7%.[9]

There are further sources of insight into users. Android users must sign into Google, ensuring real time data access, regardless of the data policies of the browser.

This data provides an unmatched picture of the long tail for specialist advertising: data can be combined and provides unique advantages in advertising targeting. In principle, pro-competitive use can be made of the data gathered. However, serious risks arise from tying practices where the market shares are so high. There is every incentive to limit access to the search and browser-based information to prevent the emergence of competing products in vertically related markets.

Moreover, benefits from large data scale are possible without tying the browser data and the advertising servers. If there is a benefit from scale, it can simply compete with other competing server deployments.

Given the very high value of the long tail, any restriction of data is very likely to move the frontier of advertising targeting inwards.

b. Is the data replicable?

There is debate over the replicability of data and the scope to process data to compete. It should be noted that there are multiple touch points in a data set. The more sources of data are available, the more complete the picture.

There are many different dimensions of data. Technical, geographical, and speed dimensions are all important. This can be seen in the header bidding debate, visible in the pleadings in *Texas v Google*.[10]

Thus, an efficient rival might well need access to a range of data sources, to compete in a related market. This could be a simple as the piece of information that there is demand, prompting a bid, in a header bidding scenario. It might also be considerably more complex.

---

[8] https://gs.statcounter.com/browser-market-share/ last accessed 2/2/21 (83.4% Chrome and Safari combined share). Edge, Samsung and Opera might also be included as they run on Chromium; although they are open-source deployments, it is unknown how strong the firewalling between Chromium and Google is. They account for an additional 9.3% of the market.

[9] https://gs.statcounter.com/search-engine-market-share Search Engine Market Share Worldwide, last accessed 2/2/21.

[10] Header Bidding allows more advertising technology to compete for a bid. Decreasing access to the data required for it would favour those who compete in the bidding process, by limiting competition in it. *Texas et al v. Google Inc.*, (E.D. Tex., 2020), *3 and *59 et seq.

More data is one thing. More data against more dimensions from more touch points is really the issue. If the browser can control what these touch points are, and define them, it will tie in significant functionality that is currently competitive. In essence, the browser proposals risk handing controls the corner of the jigsaw puzzle through unmatched scale and scope.[11]

There is no guarantee that the browser will be the best or the quickest at completing the jigsaw puzzle, but if it is permitted to tie the data in the browser to the processing of the data, it effectively controls the corners of the puzzle. Indeed, there is evidence of this already taking place, as DoubleClick IDs are removed from the market and with them some of the richest sources of data needed for targeting.[12]

c. What are the privacy impacts of these data sources?

It is notable that some of the largest data sources may also be the most privacy invasive. Smaller, competing rivals generally access less data. Although access to data and problematic identity practices are distinct, as noted below, if there are concerns simply about the accumulation of data, these would seem to be loudest in the largest players: they are most likely to have access to rich data provided after log-in and have brands which are more trusted by consumers. Thus, any concerns based on scale alone are accentuated by the tie, because the data sets used by competing data users may be less privacy invasive.

IV. Limitations to browser-based information

In the name of an asserted privacy concern, the two major web browsers, Safari and Chrome, both propose significant reductions in the passage of data from browser to servers.

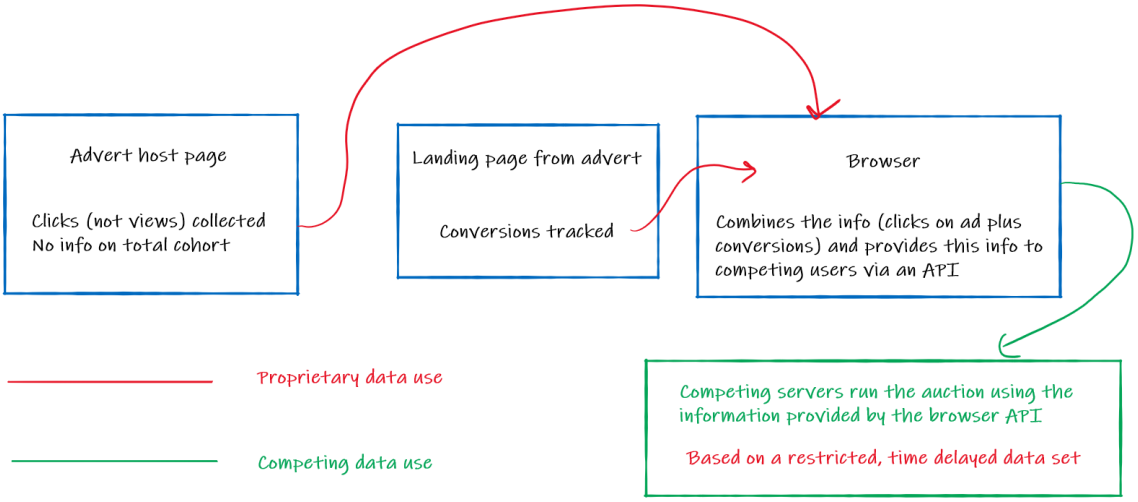
a. Safari: Webkit proposals and PPACA

Privacy Preserving Ad Click Attribution for the Web (PPACA) is an Apple proposal to change the way the Safari browser passes information to servers.[13] PPACA provides an Application Programming Interface (API) allowing some restricted server-side access to browser-based data.

It provides information based on: (1) the website hosting the advert; (2) the landing page to which the advert leads; and (3) matching information. In a nutshell, the browser reports out on the number of *clicks* (not views) on the advert, and the number of purchases made at the landing page, via an API:

Figure 2: Webkit (Safari – Apple)

[11] *CMA Final Report*, para 5.60 ("Google's extensive first-party data is likely to give it a substantial advantage over smaller rivals, creating a barrier to entry and expansion.")
[12] Google, *Important changes to Data Transfer*, Jan. 2021, available at: https://support.google.com/campaignmanager/answer/9006418?hl=en
[13] This outline follows the Webkit publication to be found at https://webkit.org/blog/8943/privacy-preserving-ad-click-attribution-for-the-web/ (22 May 2019)



There are significant limitations to the data set, notably a time-delay and a 64-item limit to the data categories provided (e.g. sight of *other* campaigns). These are major limitations, as they ban real-time retargeting and in practice provide very little relevant information on the sight of other adverts. Most users are routinely exposed to more than 64 campaigns, and the limitation here will strongly impede high quality retargeting.

b. Chromium: Privacy Sandbox

The data handling restrictions in Privacy Sandbox have been covered in other submissions.[14] The most relevant here is the proposal for the existing user identifiers to be retired: the third-party cookie and User Agent alterations.

Instead, data will be provided via interest-based targeting (referred to as FLoC) and bundled remarketing (TURTLE-DOV): a server-based deployment with an auction run within the browser itself. Server-to-server handling by third parties is restricted, because the browser runs the auction, based on parameters defined by Google. The current proposals exceed Apple's and tie significant portions of currently contested markets:

Figure 3: TURTLE-DUV (Chromium – Google)

[14] See above, fn. 2.

c. Comparison

Strikingly, the Chromium proposals go significantly further than the Webkit ones. Webkit is not a gold standard, since it also results in the tying of separate, currently competing functionality. However, the following functionality is still possible which is not possible under Privacy Sandbox:

- **At step 1: the host page,** third-party servers can still access the number of clicks on an advert via the Webkit API.
- **At step 2: the landing page,** third party servers can still access the number of conversions.

There are serious limitations at both stages: at step 1, the server can see only *clicks*, not *views*. Nor can the server see how many people were targeted by the advert, i.e. the data denominator, a crucial piece of retargeting information (known as the "ex cohort ratio").

At step 2, processing is again limited because step 1 acts as a denominator under the whole process: what good is conversion data if it cannot be matched against a meaningful cohort?

Nonetheless, it is very significant that an organisation like Apple which seeks to emphasise privacy concerns—at least in platform systems used by rivals–has not thought it necessary to go as far as the Privacy Sandbox proposals. The Privacy Sandbox proposals appear to exceed the restrictions on data processing undertaken by others seeking to address a similar (asserted) concern.

V. The proposals are a de facto tie between browser data and server-based services

As noted above, in the words of the UK CMA, the main Privacy Sandbox proposals will turn "Chrome (or Chromium browsers) into the key bottleneck for ad tech."[15]

The Privacy Sandbox proposals amount to a de facto tie between two separate products: (i) browser data storage and provision and (ii) processing services. These are distinct, but vertically related

[15] *CMA Final Report*, para 5.327.

services. A browser providing information to an advertising system is different from the service of targeting that advertisement to the user on behalf of a publisher.[16]

This can be seen in the existence of distinct markets for browsers, demand side platforms (DSPs), and supply side platforms (SSPs). The Sandbox proposal effectively ties at least the browser and the SSP, since the information on the audience is integrated into the browser. Depending on deployment, the DSP may be integrated as well.

There is currently a degree of competition in DSP and SSP services, but not in browsers where a duopoly prevails.[17] The motivation is obvious, and the leverage risk from the browser into server-side services is significant.

a. Consumer harm

It may be that a user has limited interest in the "plumbing" bringing them an advertisement, provided that the advertisement is relevant. However, this is not an argument for technological tying. It would justify any integration of technical functionality that is complicated.

For example, tying complex technologies into motor vehicles would not be justified solely on the basis that the user does not know, and perhaps does not often care, how the latest technological innovation works: a serious oversimplification of the competition, which does not necessarily all show up in the end price of the vehicle, given the transaction costs involved. Significant competition would be lost in the upstream, input markets. Effects from the tie should instead be considered in context.[18]

The relevant legal question from the leading cases is whether the integration is *necessary* to promote a pro-consumer outcome,[19] or at the very least that the integration is *relevant*.[20]

In fact, user ambivalence to the *processes* of advertising "plumbing" is a reason to be sceptical of the need for technological tying, since it would suggest that there is not a preference in favour of tying – and thus, that tying is neither *necessary* nor *relevant* to consumer outcomes.

Since the user does not see the processes involved, it is difficult to see how they benefit from tying them. Unlike the media player or web browser in the historic *Microsoft* cases, where there was at least some plausible benefit from consumer-facing integrations, the user only loses accurate targeting from these proposals: they see an advertisement, just not as well targeted.

If users have a privacy preference, that should be judged affirmatively based on evidence of privacy protection and the necessity of the tie for promoting it, and not based on the bare assertion of a benefit, or the bare assertion of a problem without a comparison with evidence demonstrating that the

---

[16] Case T-201/04 *Microsoft*, although no longer directly binding on UK competition law, suggests that the following factors should be taken into account: The situation at the time of the tying decision (para 942); The existence of consumer demand for the separate product (Para. 940); and the existence of competing suppliers (para 940). Significantly, commercial usage alone is not dispositive, since it may reflect anti-competitive practices (para 940). All of the factors point firmly towards different markets for browsers and server-side services.

[17] For Windows desktop users, this is a near-monopoly, since Apple no longer develops Safari for Windows.

[18] See e.g. Bishop and Walker, *The Economics of EC Competition Law: Concepts, Application and Measurement* (3 ed., 2010) at 6-072 (arguing that exclusionary practices driving effective price increases for consumers lie at the heart of concerns. Here, decreased competition from the tie harms advertisers, publishers, and content consumers by excluding competing advertising solutions currently in the market.)

[19] Case T-201/04 *Microsoft*

[20] *United States v. Microsoft Corp,* 147 F3d 935, 948 (D.C. Cir. 1998) "the question is not whether the integration is a net plus, but merely whether there is a plausible claim that it brings some advantage."

Less noted, but arguably more important, is the genesis of this rule, which applies a gating rule of technological relevance from the tie:

"Where there is a difference of opinion as to the advantages of two alternatives which can both be defended **from an engineering standpoint,** the court will not allow itself to be enmeshed `in a technical inquiry into the justifiability of product innovations.' " ILC Peripherals Leasing Corp. v. International Business Machines Corp., 458 F. Supp. 423, 439 (N.D. Cal. 1978), quoting Carolina, Inc. v. Leasco Response, Inc., 537 F.2d 1307, 1330 (5th Cir. 1976), aff'd per curiam sub nom. *Memorex Corp. v. IBM Corp.,* 636 F.2d 1188 (9th Cir. 1980) (emphasis added).

intervention will promote a better situation (a "greener grass" comparison).[21] As explored below, there is no obvious privacy enhancement from impeding *identification* systems provided that they do not disclose *identity*. Thus, the tie appears to be unnecessary from the consumer point of view.

This is particularly true since there is consumer harm from the tie: the unseen harm comes in the form of less content from lower payments to content providers (e.g. blogs), as retargeting restrictions undermine revenue flows to smaller publishers.

b. Anti-competitive discrimination

The technological tying between the browser and related data use is particularly troubling because it amounts to discrimination by the application of the data to a competing use, rather than for any innovative or pro-consumer reason:

- **Discrimination by data use.** Privacy Sandbox data restrictions address competing multi-site advertising. This discriminates between multi-site, non-multi-site and non-advertising use of the data. It also applies a different rule to *browser-based* and *app-based* data use, for no obvious reason (although this *is* very convenient for those who run app stores).

- **Discrimination by rivalry.** Privacy Sandbox restrictions introduce discrimination between those who use what are labelled *third party* data sets, and those using so-called *first party* data sets. This is an inherently discriminatory taxonomy, as it takes no account of the large walled gardens run by the tech platforms, which are treated as first-party systems in the eyes of data protection law. The walled gardens are only too happy to play along with this Nelsonian blindness, which simply ignores the fact that large competing advertising services can call themselves first-party and thus still see the necessary data.[22]

It is hoped is simply an inadvertent, if major, error on the part of the data protection regulators,[23] and not an exercise in regulatory capture or the troubling thought of unequal regulatory treatment of substantively similar cases. That would be the data protection equivalent of saying that first- and third-party systems are "separate but equal" and ignoring the substance of what the rule is doing.

Regardless of its genesis, this is a serious anti-competitive result from the consumer perspective. If consumers are taken to have interests in privacy and relevance, it is very doubtful that a consumer would favour a large first-party data system over a more privacy protective one that happens to be third-party. It should be clear that the quality of privacy protection does not depend on the categorisation of the system, but rather on the appropriate level of *identity protection*. By contrast, Privacy Sandbox ties *advertising identification* into the browser, without any substantive evidence that this helps with *identity protection*. This is explored further, below, in relation to justifications.

c. Incentives and anticompetitive effects

Google itself noted a 52% contribution to advertising revenue from targeting, and a 62% contribution in the case of news sites.[24] This accords with CMA estimates suggesting a slightly higher figure:

---

[21] Demsetz, Harold. "Information and Efficiency: Another Viewpoint." The Journal of Law & Economics 12, no. 1 (1969): 1, 3.

[22] The CMA is alive to discrimination between first- and third-party data sets: *CMA Final Report*, para 5.60: ("Google's extensive **first-party** data is likely to give it a substantial advantage over smaller rivals, creating a barrier to entry and expansion." (emphasis added)); 5.312 ("both Google and Facebook do not allow advertisers and independent third-party providers of measurement and attribution services to collect user level data from ads shown on their owned and operated inventory (ie in the walled garden). This hurts independent attribution providers and gives an advantage to Google and Facebook's own ad tech and analytics services"); 5.324-324 (noting discrimination between first- and third-party data sets from the retirement of third-party cookies). See also Appendix G on user identification.

[23] It is troubling that the ICO is "encouraged" by these developments. *Final Report*, para 5.328.

[24] Michael Kleber, introducing Privacy Sandbox, https://web.dev/digging-into-the-privacy-sandbox/ at 19' 40".

70%.[25] Google asserts that this creates an incentive not to foreclose, despite a clear ability to do so, on the basis that content helps drive traffic and thus the value of search advertising.

This ignores the obvious possibility that search advertising, a competing product in the eyes of many advertisers,[26] is likely to benefit from decreased competition from retargeting services. The same is likely to be true of competing DSP and SSP services. Even on the numbers presented, the "vertical arithmetic" is likely to favour Google: if demand is at all inelastic, decreased traffic at a higher price is likely to pay. There is no analysis of the volumetric impact of increased price, but it is doubtful that the advertising involved would be very elastic. This can be seen in the high prices paid for "spikes", where retargeted adverts command a substantial premium.[27]

As there is a high value to targeting, and the data needed for targeting is tied, downstream competing products are likely to benefit from decreased rivalry since the data is an input for both uses. Even if there is a margin of substitution, which in any event has not been shown, it is likely to be inframarginal substitution into no-consumption, i.e. a cellophane fallacy, because targeting exists at a margin.

Further, since the variable cost of a spike is moderate, the high price of a spike seems to imply less than perfect competition at least in the short term. The risks from technological tying in these circumstances are substantial, as the rents which appear to be present can be extracted via demand profiling tactics enabled by the tie.[28]

This explains why the tie is wanted despite the possibility of simply charging more for the data: restricted access allows leverage into a related market (targeted advertising) with differentiated demand profiles. The tie also harms rivalry from targeted advertising into a different, related market (search advertising).[29]

There is also a significant prospect of anti-competitive discrimination in favour of apps run by the browser operator. As above, data used in apps is treated differently.[30] This allows discrimination in favour of app-based provision, where the two largest browser providers sell competing products.[31]

VI. Is there a privacy justification?

It is for the tying product provider to justify a tie based on evidence.[32] However, an argument can be anticipated that technological tying promotes user privacy.

It should be noted that, at least under EU competition law, there is a duty for a dominant company to provide an unbundled version of a tied product wherever possible, to allow consumer choice between the tied and the non-tied version.[33] Here, that would mean allowing consumers to choose between a more, and less, locked-down version of the browser.[34]

---

[25] *CMA Final Report*, para 5.326 and Appendix F.
[26] *CMA Final Report*, para 2.8.
[27] *CMA Final Report*, para 2.5; para 5.162 ("Access to higher quality or more granular data allows for more precise targeting of more specific audiences. Granular data is particularly valuable when combined with high reach among different audience types using the platform, as this allows for relatively large numbers of very specific audiences to be targeted. These factors can allow platforms with better data to sell their advertising inventory at higher prices. This creates a substantial competitive advantage for Google and Facebook, both of which have access to much richer and higher quality datasets and benefit from much greater scale and reach than their rivals.")
[28] In a perfectly competitive market, demand profiling based on the data would not be possible; the tie enables it.
[29] Bishop and Walker, op. cit., 6-064 (noting strengthening of position in other related markets in some cases).
[30] E.g. Apple's IDFA restrictions.
[31] See e.g. *Epic Games, Inc. v. Apple Inc.* Case No. 20-cv-05640 (N.D. Cal., 2020); Case 1377/5/7/20 *Epic Games, Inc. and others v Apple Inc. and Another* (UK Competition Appeal Tribunal, 8 Dec 2020).
[32] Case T-201/04 *Microsoft*, paras 1144 to 1167: tying must be *essential* for the tie to be justified.
[33] Even some proven efficiency benefits would not themselves suffice to justify a tie if there are means to offer *both* a tied and a non-tied product (*Microsoft* at para 1152).
[34] Unless impossible, a non-bundled version should also be offered (*Microsoft* at para 1149).

This is a major problem for a justification, as the tying must meet the burden of being *essential*. It is also notable that even under a more permissive test, these justifications would still be very weak, as they make a fundamental error: they confuse *identifiers* with *identification*.

a. *The difference between identity and identification*

It may be that users have concerns about what is pejoratively called tracking, although as the CMA noted in its Online Markets report, there is currently a gap in the literature on the revealed preference for data protection, so it is somewhat speculative to make this claim.[35] It may be that consumers want more privacy, but this is not usually defined with any precision in surveys: see above on tying complex products. More importantly, this is not unlike asking if you would like a faster car, bigger house, more free time etc.: It does not identify a trade-off. The relevant question would be to compare properly defined privacy protection with its impact on advertising-funded services, and thus the cost of privacy protections to users.

Significantly, it is for the tying party to provide this evidence. Otherwise, the spontaneous order is to be protected.[36] Moreover, for such an argument to work, it must go no further than necessary to achieve its aims.[37] Here, there are serious questions over whether there is any user benefit to either the Apple or Google proposals.

There is a major difference between *revealing a user identity* and *using an identifier*. Identifying a user might raise concerns, e.g. John at house number 22 was bankrupted and therefore should have to pay more for car insurance. That would be especially true in the case of a protected characteristic like race or gender. However, specific laws address those *identification* problems, with reference to specifically prohibited *identity practices*.[38] Whether the practice is a problem does not turn on whether the party undertaking the activity is dominant, since the practice is always problematic, although it is true to say that dominant companies may undertake more of the practices.[39]

Identifiers, by contrast, do not pose such issues. They are simply a means to identify characteristics, and do not reveal individual *identity*. For example, *Insured Party X currently drives car model X and may be in the market for car model Y.* Nothing is said about John at number 22. The example does not reveal *identity* despite containing significant *identifiers*.

Server-side processing is emphatically an exercise in *identifiers*. The DSP and SSP process information about an individual, but this does not reveal who they are. It just refers to *Party X*. The servers are not playing Cluedo, but are simply trying to make sales. Even if they were playing Cluedo, they would never win the game. Just knowing that someone was in the study at some time with a revolver does not win, unless you can say it was Professor Plum as well. But data protection law prohibits that in all cases regardless of dominance—or it should do if there is an unaddressed *identity* concern.

This is clearest in the cookie string: user ABC123XYZ—an online avatar—might be identified many times, but there is no revelation that it is John at Number 22. That would be troubling whether done by a dominant company, or not. And if that were a concern, it would be a case for specific restrictions

[35] "Few surveys examine what UK consumers perceive the specific benefits or harms of data processing and targeted advertising to be. Instead, consumer surveys tend to focus on the high-level benefits and harms resulting from all forms of online targeting." (*CMA Final Report* Appendix L: para 285. *Summary of research on consumers' attitudes and behaviour*).
[36] Recital 5 to EU Regulation 1/2003 (burden is on the defence to prove its applicability); see also *Microsoft* para 1144 (dominant company to put forward evidence of the justification).
[37] *Microsoft*, para 1151 (obligation to provide a non-bundled version to allow consumer choice).
[38] E.g. Equality Act 2010; Car insurance gender discrimination ban under EU law.
[39] This issue will be familiar from the German Federal Cartel Office's litigation on the terms and conditions for Facebook users. This litigation ultimately was resolved in favour of the position that dominance can materially diminish the quality of data protection terms and conditions.

of the specific practice raising a concern, using generally applicable data protection law. It is not a justification for a technological tie of browser data by dominant browser providers.

The difference can be helpfully thought of in terms of outcome. Consider a camera pointed at traffic and noting the passage of one fast moving car. The same piece of information leads to very different results depending solely on whether *identity* is revealed:

- **Traffic information.** *Identifiers – **at 4pm, northbound** traffic is moving freely **at junction 15**.*
- **A speeding ticket.** *Identity – **John** speeding at junction 15 at 4pm.*

In both examples, the same thing has happened: John's car has passed the camera. Whether *identity* is revealed leads to a very different result! The example highlights that data use and abuse questions should be assessed by evidence-based outcomes, and not by using arbitrary definitions derived from the technology used, since this does not indicate the presence or absence of concerns.

b. Browser-based restrictions to competition in identification systems

In a hypothetical perfect market, there would be competition on the balance between perceived privacy and the funding of content via advertising. The real world may fall short of this ideal. It can be argued that good competition policy is not to think of ideals, but to compare practical, evidence-based equilibria and whether they are more, or less, competitive.[40]

In this case, the relevant comparison is between access to the browser-based data, and discrimination over it to prohibit third-party use. It is striking that third parties may have the strongest incentives to innovate on privacy. A large first-party system, with its fig leaf of consent from one interaction, may have little incentive to innovate on privacy. After all, it has been given a pass by EU data protection law: the enterprise-wide privacy policy is complied with.

Indeed, there is a dangerous assumption that this data protection compliance by a first-party system is adequate. In fact, there is no scope for user control over the first-party data sets once the consent is taken: unlike existing technologies like cookies, there is no guarantee that users can revoke consent given to a first-party system. The system is simply on trust to implement such a request.

By contrast, a third-party system provider has a strong incentive to innovate, precisely to avoid outcomes like the Privacy Sandbox. If third-party systems are at risk from reputational harm prompting regulation, they will be at pains to demonstrate that their *identification* systems do not allow problematic *identity* problems – since the alternative is regulation that will impede them compared with first-party systems.[41] Indeed, third-party systems currently compete on user control over the identifier.

In other words, the strongest competition for privacy is likely to come from the third-party systems, precisely because they are not first-party.

Other important metrics of consumer-facing competition include cookie consent repetition and consent fatigue, which may be addressed through innovation. A first-party system able to use first party cookies or other generally applicable consent mechanisms has no incentive to address this problem.

There is also significant advertising-facing competition in identification systems. If the browser succeeds in tying user data to server-side technology, it will prove very difficult to introduce some strongly desired advertiser functionality, notably:

[40] Again, avoiding a "greener grass" fallacy derived from the application of abstract comparisons as opposed to benchmarking evidence.
[41] Indeed, this is a particularly strong reason why regulation of any privacy concerns must not be in the hands of a dominant provider: there is a major conflict of interest.

- **Auditing.** The Sandbox proposals are strikingly light on any means by which advertisers might verify the views, clicks, and conversion mechanisms.
- **Retargeting optimisation.** As the data is an essential input into retargeting, preventing its flow to third parties prevents optimisation of retargeting methods through innovation in retargeting algorithms.

A universal ID of the sort mooted by the CMA, or even just competition on the introduction of competing identifiers, would also prove impossible if the data required for it is tied into the browser.

It is unclear why this level of control over the data-driven supply chain is required, especially in the absence of any robust data showing that *identification* mechanisms are harming consumers where *identity revelation* does not occur. Instead, there appears to be an assertion that control over the data is required without any serious assessment of the relevant trade-offs from a consumer point of view.

VII. Remedy

The safest path in the absence of affirmative evidence of consumer harm from privacy concerns would be not to tie the browser data and server-side processing. At the very least, an unbundled product should be available, i.e. a browser version without the restrictions. This would allow consumers to choose whether to adopt the tied functionality, and is a light touch, "subtractive" tying remedy.[42] This could be highly significant: ad-funded browsers might well gain access to more content, whereas those users favouring more protection, and thus less valuable advertising, might have to pay for content via a paywall. This increases utility for both groups.

It is not clear why an unbundled version cannot be offered. If however it were considered necessary to adopt the Privacy Sandbox proposals as a global deployment, it would be important to follow through with an anti-discrimination remedy designed to address the potential for anti-competitive discrimination from these controls. The rest of the paper considers these possibilities to address the clear risks of harm identified above.

a. Interconnection of servers

At a minimum, competing ad tech should be able to access data streams on equivalent terms to those of competing Google advertising servers. If there truly is a reason for data not to flow because of identity concerns, then that concern must be equally applicable to any *internal* servers. The browser and the server are undertaking different tasks, and the dominant browser provider must not prevent interconnection of rival server-side services, as this would be anti-competitive.

b. Non-discrimination in data collection

Considering the large data sets required for retargeting, it may be that simple interconnection of competing advertising technology on equivalent terms does not go far enough to enable server-side competition. It may prove necessary to prohibit restrictions on data collection in the browser, since they amount to an essential facility for an innovative product provided by a third party.

This is a broader remedy than interconnection and would be likely to open the market to more competition. If there were a privacy case requiring limitations to the scope of data access, these could be pleaded specifically by application to the regulatr for variation in the remedy, based on evidence of consumer benefit from the change.

[42] A subtractive remedy removes a function. It does not prevent innovation, but allows access to the non-tied product for those who wish to use it.

For example, the Apple proposals allow API-based access to browser-based information—although they still restrict significant functionality. Nonetheless, keeping Chromium at least this open to rivalry might be considered a bare minimum.

It should be noted that this amounts to conceding that browser-based processing is necessary for no obvious reason. However, if the 64-tag limit in the Apple proposal is necessary to enable local processing – assuming that there is evidence for the necessity of local processing – then the specification of those categories should be open to any provider. They should not be defined by the browser provider, as in TURTLE-DUV, as this would tie the advertising technology to the browser.

c. Ex-client data access

Google appears to have access to a unique identifier of its own, called Ex-Client Data. The status of Ex-Client Data is not currently known, but it appears to be a widespread identifier used by Google. Providing access to Ex-Client Data would open competition to equally or more efficient rivals.

Like any access remedy involving an existing resource whose creation is, in principle, a variable, there might be arguments about diminishing incentives to invest. These must be weighed against the foregone competition in the access market.

In this case, where ad tech market shares are so high, and where the Ex-Client Data is already deployed on an already widespread system, it is unclear why mandating access to it would have any serious impact on innovation incentives. On the contrary, providing this identifier to rivals would strongly enhance competition in the ad stack: whoever can make best use of the data in the eyes of system users would prevail.

d. Firewall

As above, it may not be technologically feasible to deploy browser-based solutions without foregoing significant innovation, because the scale of computer processing involved may well require server-side data handling.

If it is truly the case that evidence demonstrates a need to restrict the passage of data from the browser to servers in the user interest, a conflict of interest on data handling would remain: the browser and the ad tech layer being different products, but run in some instances by the same company, there is significant scope for anti-competitive discrimination that would be very challenging to police ex post. A firewall between browser and server functionality would robustly address this.

Functional separation would be a sensible and robust step to address the conflict of interest between the browser and server-side functions.

VIII. Conclusion

This paper has noted recent developments which appear to introduce a significant risk of a new technological tie between browsers and advertising data required for competition in advertising. In the context of potentially serious damage to competing advertising and the rich, free publishing it enables, urgent attention is needed to consider whether well-tested competition law remedies are required in order to prevent the application of a specious privacy concern by dominant companies in a discriminatory fashion, since their own first-party “walled gardens” are unencumbered by any such rules. The paper concluded with some suggestions on how a remedy might be structured to prevent the abuse of market power in the browser tying market from snuffing out the already somewhat

threatened competition in vertically related markets. In this way, competition in advertising can continue to fund free content.

Significantly, there is no reason why a more privacy-enhancing browser cannot be introduced; it is just that tying law requires a non-tied version to be offered so as to enable consumer choice between different levels of privacy, rather than applying a one-size-fits-none bundled option. Applying such a remedy has the promise of enabling more choice, differentiation, and competition between models that trade off privacy, content monetisation, and payment for content in different ways to suit different consumers. It remains to be seen whether the commitments offered by Google to the UK CMA, which are currently out for consultation, will meet the concerns.