Taking the biscuit: how Safari privacy policies affect online advertising*

Simeon Duckworth[†] Mateusz Myśliwski[‡] Lars Nesheim[§]

April 2023

Abstract

Many controversies that beset the digital economy turn on the role of advertising and its use of personal data. We document several new stylised facts about the global digital advertising marketplace and examine the trade-off between privacy and ad targeting accuracy from the advertisers' perspective. We exploit a novel dataset with billions of observations of online ads spanning multiple countries, advertisers, and websites. Our focus is to estimate the impact of Apple's gradual restriction and ultimate abolition of ad tracking in its Safari browser called Intelligent Tracking Prevention (ITP). We analyse how much advertisers are willing to pay for third-party cookies and how tightening privacy policies affects market outcomes. Our empirical strategy treats Apple's policy changes as exogenous shocks to the supply of tracking opportunities and uses a series of event study models to estimate their causal impact. We find that the estimated treatment effects around the ITP introduction dates are small in magnitude on average but differ markedly across countries, advertising campaigns, and type of marketplace. These finding are consistent with a theoretical literature showing that changes in ad targeting have ambiguous general equilibrium effects. Moreover, our results suggest that markets failed to adjust immediately to new, more privacy-sensitive equilibria.

*We thank Garrett Johnson and seminar participants at NHH and the University of Bergen. We are grateful to an anonymous advertising agency for sharing the data. The views expressed in this paper belong to the authors and do not necessarily reflect those of any of the parties supplying data. All errors are ours. Myśliwski acknowledges financial support from the Research Council of Norway's KULMEDIA grant.

[†]UCL, simeonduckworth@gmail.com
[‡]NHH, mateusz.mysliwski@nhh.no
[§]UCL and IFS, l.nesheim@ucl.ac.uk

1 Introduction

In little over two decades, digital advertising has grown to $350bn in revenue - just over half of all advertising globally (GroupM, 2022). While a significant proportion of this growth is driven by the shift in consumer attention to web and mobile devices, it is the capability for user-level ad tracking that has become the defining characteristic of digital advertising (Goldfarb and Tucker (2019); Goldfarb (2014)). Tracking technology, such as cookies and mobile advertising identifiers, have enabled advertisers of all sizes to target, measure and buy advertising at relatively low cost. For proponents, this ability to link and share user data is more efficient for advertisers than traditional site-based targeting and has both driven demand for publishers and made content more accessible for consumers (Deighton and Kornfeld, 2020). But, as the industry has grown, ad tracking has also become increasingly controversial, triggering widespread concerns about privacy and misaligned content incentives (Information Commissioners Office, 2019; Aral, 2020).

In response to the rising controversy, both regulators and key industry actors have moved to protect consumer privacy and reduce data sharing without consumer consent. Through landmark laws, such as the General Data Protection Regulation (GDPR) in Europe and the California Consumer Protection Act (CCPA), regulators have sought to establish legal frameworks that permit consumers to control and limit the use of their data. More directly, several digital platforms have either reduced, or publicly committed to reduce, ad tracking.[1] For example, cross-site tracking with cookies is severely limited or prohibited in most web browsers[2] and Google, the market leader, has committed to eliminate third-party ad tracking in Chrome browsers from 2024.

As the industry moves toward a more privacy-sensitive equilibrium, regulators and industry participants are grappling with a new set of questions related to the trade-off between consumer privacy protection, publishers' revenue and advertisers' surplus. In this paper, we address the issue that underpins all of these questions: how sensitive is advertiser demand to user ad tracking and what is its impact on market outcomes?

In theory, a reduction in ad tracking has an ambiguous impact on ad prices. It reduces advertiser demand and willingness-to-pay for generic impressions, but at the same time thickens the ad auction market, increases competition between advertisers for less differentiated ad slots and, consequently, can drive prices up (Levin and Milgrom, 2010). Moreover, given advertiser and publisher heterogeneity, the relationship between the supply of targeting opportunities and ad prices is likely to be non-monotonic (Bergemann and Bonatti, 2011). In this paper we test these theoretical predictions and provide large-scale empirical evidence on how limiting tracking opportunities affects market equilibrium.

Our paper exploits the gradual abolition of ad tracking in Safari browsers to identify the impact

[1] See e.g.: https://blog.google/products/ads-commerce/a-more-privacy-first-web/, accessed December 2022.

[2] As of the end of 2022, third-party cookies are fully blocked in two main competitors of Chrome: Safari (https://webkit.org/blog/10218/full-third-party-cookie-blocking-and-more/) and Firefox (https://blog.mozilla.org/en/products/firefox/firefox-news/todays-firefox-blocks-third-party-tracking-cookies-and-cryptomining-by-default/), accessed December 2022.

of information restrictions on ad prices and advertiser demand. Apple first introduced Intelligent Tracking Prevention (ITP) in September 2017 with the stated aim of preventing cross site ad tracking in Safari browsers. ITP works by identifying which domains are cross site trackers and limiting the amount of information they can access. Over seven browser updates from 2017 to March 2020, ITP became progressively more restrictive as Apple combated industry workarounds and unintended consequences until ad tracking was completely eliminated in Safari browsers in March 2020. We provide more background in Section 4.

We measure the impact of ITP on ad prices and quantities using a large advertiser panel. Our dataset covers programmatic advertising spend for over a thousand advertisers from 2017 to 2020. Programmatic advertising is broadly defined as advertising on the open-web, excluding search and advertising on social platforms. Refer to Section 3 for further details. In contrast with most existing studies, which focus on individual publishers or ad exchanges, our data covers all advertiser spending on open-web advertising across multiple intermediaries, ad exchanges and publishers over several years and markets.

We treat the introduction of ITP as an exogenous and comparatively sudden change to advertisers' media choice set. Its impact on Safari ad prices and quantities is measured relative to Chrome using a difference-in-difference design, comparing outcomes before and after for the same campaigns to control for unobserved marketing goals and targeting strategies. We estimate this design separately for each country and ITP launch event - that is nearly 150 ITP launch events. The richness of our big dataset allows us to evaluate the introduction of Apple's privacy change on a previously unprecedented scale and study various margins of heterogeneity.

Our main finding is that there is a relatively small overall impact from ITP on average ad prices in Safari relative to Chrome. There is, however, considerable heterogeneity. For example, the causal impact of ITP on prices in open auctions, the most competitive segment, is around 5% - a larger impact than for other deal and market types. We subject our findings to several robustness checks using more granular, event-level advertising data. The headline finding remains the same.

While - as far as we are aware - there has been no academic research on the causal impact of ITP, several industry studies have looked at the baseline difference in average prices between Safari and Chrome. These estimates range from nearly zero[3] to 50-60% (Bidswitch, 2020). In addition to the causal impact of ITP, we also estimate the baseline Safari premium and find that Safari can be up to 20% cheaper. While our estimates fall in the range reported in the industry, the lack of a clear time trend suggests that gradual tightening of privacy rules does not have an obvious cumulative effect on ad prices.

Our research design mitigates concerns about the impact of unobserved campaign-level variables which can affect the result and the results pass a wide range of robustness checks. We cannot, however, rule out the possibility that unobserved time-varying impression quality biases the results, since we do not directly observe cookies. Intermediaries are incentivised to spend budgets and so may respond to price cuts by increasing the quality of the ad buy, for example by bidding on publishers with greater reach. The main results are robust to the inclusion of publisher and

[3] See https://fouanalytics.substack.com/p/media-cpms-cpms-paid-to-publisher.

creative fixed effects. A second possible explanation for our finding of small effects on average is that the equilibrium impact of reduced targeting is dampened by a gradual switch to private marketplaces (PMPs) where improved contextual targeting can offset reduced user-level targeting. We are investigating this hypothesis in ongoing work exploring the growth of PMPs.

The paper makes three contributions to the literature. Firstly, we explore a novel dataset with billions of observations which, apart from evaluating Apple's policies, allows us to document several new stylised facts about the advertising industry related to price dynamics, dispersion across and within campaigns and market structure. Secondly, existing studies estimating the value of user data have relied on marginal decisions in ad auctions or small-scale experimentally-induced variation. For policy makers and industry participants, our evaluation of a sizeable policy rather than a small experiment provides greater external validity. Finally, our empirical exercise can be seen as a test for the theoretical predictions of Levin and Milgrom (2010). Our results are broadly in line with their *conflation* hypothesis – privacy restrictions create thicker auction markets and while advertisers might still be willing to reduce their bids for untargeted impressions, increased competitive pressure attenuates the effect on ad prices.

The rest of the paper is organised as follows. Section 2 outlines the relevant literature. Section 3 describes the market context: programmatic advertising, the key intermediaries and the role of data. ITP, Apple's policy to reduce third-party ad tracking in Safari browsers, is discussed in Section 4. We introduce the datasets in Section 5. Section 6 documents a handful of programmatic advertising stylised facts. Section 7 outlines our econometric methodology. Our results are presented in Sections 8 and 9, where the latter explores various dimensions of heterogeneity. Section 10 discusses the limitations of our main findings and proposes several robustness tests using more granular data. We conclude in Section 11, drawing out the implication of our results for further work.

2 Related Literature

This paper focuses on how ad tracking restrictions impact advertiser demand, substitution and prices. We draw on three distinct fields of research: a theoretical literature on the impact of targeting on ad prices; a reduced-form literature estimating the value of a cookie in ad auctions; and an emerging structural literature estimating the trade-off between ad tracking and market outcomes.

In theory, less accurate targeting has an ambiguous impact on equilibrium ad prices. More accuracy increases differentiation between advertisers, softens product competition and ultimately, by increasing willingness-to-pay, raises media demand (Iyer et al. (2005); Johnson (2013); Chen and Stallaert (2014); Esteves and Resende (2016); Marotta et al. (2019)). But user-tracking also impacts the degree of competition. It increases competition between publishers by widening the potential ad inventory for key consumer segments (*data leakage*) (Ghosh et al., 2015) and it thins the number of bidders in auction-markets increasing advertiser information rents (*conflation*) (Levin and Milgrom (2010); Hummel and McAfee (2015); Bergemann et al. (2022)).

Targeting increases advertising efficiency. Several papers show that the impact of competing targeting technologies on market outcomes also depends on the composition of demand. For example, Bergemann and Bonatti (2015) investigate an advertiser's choice of targeting strategy and its impact on platforms' data pricing policies. In their model, advertisers choose between three targeting regimes: positive targeting where advertisers act positively to offer personalised content or media exposure to an audience, negative targeting that excludes an audience, and no targeting in which there is a residual untargeted audience. Media and data can be complements (positive targeting) or substitutes (negative targeting) and the impact of data on ad prices will depend on which strategy predominates. The value of targeting depends on the price of media and the quality of the match residual set. Crudely, cookieless Safari impressions could increase the expected value of non-targeted solutions (e.g. premium IOS users) or increase media demand (e.g. to compensate for reduced frequency-capping).

More subtly, Bergemann and Bonatti (2011) emphasise the importance of advertiser and publisher heterogeneity for market outcomes. They study the competition between online and offline media, showing that ad prices are non-monotone in targeting capability. At first, prices increase with more accurate targeting and then decrease as the advertising demand shifts from the extensive to the intensive margin.

Heterogeneity among advertisers and publishers is a key characteristic of the market and a reduction in ad tracking will have distributional effects. Targeting accuracy also has welfare effects. The trade-off between data accuracy and consumer (or firm) privacy is a key feature of differential privacy (Dwork and Roth, 2013) and central to Google's proposal to replace current cookie-based data sharing protocols (Ravichandran and Vassilvitskii, 2020). Essentially, differentially private algorithms add noise to the data to prevent identification of users when it is shared. Increasing the coarseness of disclosed information in ad auctions affects the distribution of surplus between the buyer and the seller (Bergemann et al., 2022) and between advertisers and consumers (Elliott et al., 2021). Large advertisers buy less targeted media, valuing broad reach more highly than niche targeting (i.e. avoiding *false negatives* more than *false positives*) (Gal-Or et al., 2006; Bergemann and Bonatti, 2011).

Athey et al. (2016) show that imperfect tracking and fragmenting audiences combine to frustrate the measurement of audience reach and consequently reduce the value of an incremental ad. In the context of our paper, coarser ad tracking not only reduces ad prices but also decreases demand for smaller publishers.

There is a broad empirical literature on the impact of ad tracking on market outcomes, summarised in Table 1. The literature covers experiments, counterfactual simulations and policy event studies. A key strand of the literature uses display ad auction data to estimate prices of impressions with and without cookies. Selection bias is an obvious concern in these studies. Advertising is targeted toward users based on a wide set of consumer characteristics that are available to bidders but not to researchers. Despite different approaches to selection bias, there is a degree of concurrence that cookies earn publishers a 2-3x premium over a cookieless impressions. Marotta et al. (2019) is an exception, finding only a modest 4% premium.

In theory, properly designed experiments should be free of selection bias. Johnson (2022) provides a survey of display advertising field experiments. Google (Ravichandran and Korula, 2019) randomly

Table 1: Empirical literature: Estimates of the value of user-tracking

Study	Data	Method	Metric	Estimate
Beales and Eisenach (2014)	2 ad exchanges + large, multi-side publisher	Regression	Exchange/publisher price	66%[1]
Marotta et al. (2019)	Large multi-site publisher	IPW	Publisher revenue	4%[1]
Johnson et al. (2020)	Large ad exchange	Regression	Exchange price	52%[1]
Ravichandran and Korula (2019)	Google top 500	Experiment	Publisher revenue	52%[1]
CMA (2020)	Google top 500	Experiment + ML	Publisher revenue	72%[1]
Laub et al. (2022)	Large ad exchange	AIPW	Publisher revenue	24%[1]
Rafieian and Yoganarasimhan (2021)	Mobile ad exchange (APAC)	ML algorithm + structural auction model	Bids, revenue & surplus	-1%[2]
Alcobendas et al. (2022)	Yahoo (multiple DSPs)	Structural auction model	Bids & revenue	35%[2]
Aziz and Telang (2016)	Large e-commerce retailer	Regression	Sales from re-targeting	85%[3]
Goldfarb and Tucker (2011)	Consumer surveys	Natural experiment (e-Privacy directive)	User purchase intent	65%[3,4]
Cecere and Lemaire (2021)	Facebook Marketing API	Natural experiment (ATT)	Ad prices	10%[4]
Aridor et al. (forthcoming)	Travel intermediary	Natural experiment (GDPR)	Average bids (search advertising)	-12%[4]

Note: (1) Reduced form estimates of cookie value, (2) Structural model estimates of cookie value, (3) Impact of user-tracking on advertising effectiveness, (4) Impact of policy intervention.

deleted cookies from its bidding platform in auctions for the top 500 (non Google) publishers. It found that, without cookies, short term revenue fell by 52% on average (60% for the median publisher), but with marked heterogeneity. The UK Competition and Markets Authority (CMA) re-analysed Google's experiment and estimated the cookie-premium to be 70% (CMA, 2020, Appendix F). It also noted that this provides an estimate only of the short run effect because the equilibrium value of user information would be lower as the market adjusts. An advantage of our multi-year panel dataset is that we can capture market dynamics in response to Safari's ad tracking restrictions.

As access to the technology necessary to randomise assignment in ad auctions is limited, most researchers have relied on large scale observational studies. Selection is controlled by modelling the impact of cookies on winning bids with covariates collected and shared by either the ad exchange (Beales and Eisenach, 2014; Johnson et al., 2020) or publishers (Marotta et al., 2019). Other approaches have looked at the demand side. Aziz and Telang (2016) estimate the value of a cookie on purchase propensities in re-targeting and Goldfarb and Tucker (2011) estimate the impact of the e-Privacy Directive on survey-based measures of advertising response.

Rafieian and Yoganarasimhan (2021) use data from a mobile ad server to estimate the welfare impacts from greater user privacy. They exploit an unusual auction-type, where ads are served randomly to users conditional on observable characteristics, to estimate underlying match rates between advertisers and impressions. In support of Levin and Milgrom (2010), they find that the platform gains from a reduction in user targeting as it induces more competition between advertisers for each impression and, consequently, higher ad prices. This is a surprising result as it would imply, for example, that Facebook would gain from the elimination of tracking on iPhones and iPads (McGee, 2021). In the same vein, Aridor et al. (forthcoming) find that introduction of GDPR increased advertising prices on travel websites through a different mechanism. They conclude that opting into GDPR made the average consumer more trackable as she substituted away from ad blocking.

Using data from one-day's trading on Yahoo, Alcobendas et al. (2022) estimate a structural model of Demand Side Platforms (DSPs) bidding for users in different browsers. Although their data does not capture a change in browser privacy policy, they use their structural model to estimate that, if Google bans cookies in Chrome browsers, ad prices would fall by 30%. Google's DSP, DV360, would also be informationally advantaged, but the impact on its probability of winning the auction would be only about 1% point. A limitation of the analysis is that it treats each DSP as a proxy for an advertiser and therefore cannot fully control for advertiser and campaign-level heterogeneity.

This paper takes a different approach to all those discussed above. We investigate the impact on advertiser demand from a policy event, the introduction of Intelligent Tracking Prevention (ITP) in Safari browsers. This browser policy change had an exogenous and comparatively sudden impact on about 15% of the market's inventory. The nearest comparison to our paper is Cecere and Lemaire (2021), who estimate the impact of Apple's App Tracking Transparency (ATT) on Facebook and find that (predicted) ad prices on iOS declined by 10% relative to Android.

3 Programmatic advertising

This section provides a brief overview of the most important features of the industry, its key terminology and explains acronyms which are used in the later parts of the paper. We only focus on aspects which we deem relevant for understanding the main findings of our empirical analysis, so the whole picture is necessarily simplified.

Our paper focuses on programmatic advertising, a segment of the digital display advertising market. Programmatic advertising is predominantly advertising on the *open web*, for example on news and general interest sites. In 2022, advertisers spent around $125bn worldwide on programmatic advertising, about a quarter of all their spending on digital display advertising.[4] The broader digital display market, in addition to programmatic, also includes social media, online video and open display but excludes search.

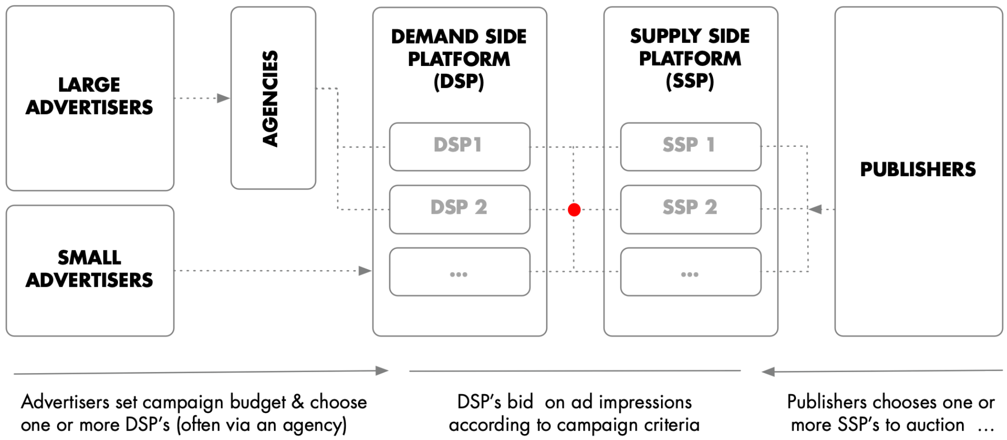
Programmatic advertising is highly automated. To process millions of user views (called impressions), match buyers and sellers, complete the auctions, and serve the ads in real time, requires enormous quantities of user and site-level data (IAB, 2014). Each auction is completed in about 200 millisecond enabling the winning ad to be displayed before the user's web page loads. Over the last few years, open display has largely transitioned from second to first-price auctions (Goke et al., 2022). Google started using first price auctions in September 2019 (Google, 2019).

The key participants in programmatic advertising are advertisers (ad buyers), demand side platforms (DSPs), supply side platforms (SSPs) and publishers (ad sellers).

Figure 1 provides a simplified schematic of the programmatic marketplace. Most advertisers work with an agency to select and manage programmatic advertising. Advertising agencies help advertisers manage campaigns across multiple media channels - for example, programmatic, social, search and TV - and set and monitor campaign parameters. Advertisers divide activity into campaign flights. A campaign flight typically has a single marketing goal, budget, audience targeting and

[4] See https://www.insiderintelligence.com/insights/programmatic-digital-display-ad-spending/.

Figure 1: Programmatic Advertising Industry in a Nutshell



measurement strategy. Target ad prices are set as media cost per thousand impressions, or *cost per mille* (CPM). Campaign flights are briefed to a DSP along with campaign parameters, such as site 'white' and 'black' lists. Most campaign-flights are only briefed to one DSP. DSPs bid on behalf of advertisers. They determine which auctions to enter and how much to bid for each impression within the campaign parameters provided by advertisers and agencies (Wang et al., 2017). Impression auctions are managed by SSPs on behalf of the publishers.

Advertisers are diverse. Most obviously they differ in the nature of the product they sell (automotive, finance, consumer good etc) and the size of their budget. But campaign goals are also heterogenous. Specifically, campaigns are often characterised as being either *brand* or *performance* orientated. *Brand* campaigns aim to increase salience or positive perceptions. Typically, they monitor reach and frequency of exposure among a broad target audience and measure effectiveness through consumer surveys or modelling sales. These campaigns tend to rely less on ad tracking. *Performance* campaigns are optimised toward media attributed outcomes such as *cost per action* (CPA). In these campaigns, target audiences are defined granularly based on predicted responses and ad tracking tends to be higher in performance campaigns.

Advertisers target advertising to achieve the campaign goal at lowest cost. Broadly, there are two types of targeting strategies: contextual and behavioural targeting. Contextual signals include the website and page content (for example content on electric cars). Behavioural signals use the history of web-surfing behaviour to infer interests, demographics, as well as specific browsing behaviour (for example, clicked on a car ad last week). For marketers the line between the two can be blurry, but, for our purposes, the important distinction is that behavioural targeting relies on ad tracking.

Our paper focuses on the role that data, and ad tracking in particular, has on market outcomes. Broadly defined, ad tracking is activity that collects, links, stores and shares user data across multiple websites, devices and contexts for targeting and measuring advertising.

Cookies are small text files or pixels with a unique identity. They are the most commonly used ad tracking technology. Other forms of tracking include Mobile ID's on smartphones and *fingerprinting*, a type of probabilistic matching.[5]

Cookie tracking works as follows. If a user visits "What Car" magazine online, the publisher saves a cookie to the user's browser. The cookie is used to collect and store data about the user's visit, such as the content they read, the cars they clicked on or ads they saw. This information can be used by "What Car" in a first-party context to improve the user's experience on its website. But it can also be used for advertising. By linking with third-party databases, cookies can be used to track users on other websites, to append demographic and interest data and serve them car advertising. In the context of our paper, it is important to note that cookies can be used for purposes other than advertising, that there are alternative methods of tracking (e.g. *fingerprinting*) and that the browser plays a gatekeeper role by processing, storing and sharing user information.

There are principally two types of programmatic advertising marketplaces: Real Time Bidding (RTB) and Private Marketplaces (PMP). In RTB, any advertiser can bid for each impression in an open-auction (Wang et al., 2017). While RTB is the archetypal programmatic advertising marketplace, programmatic also includes more direct trading relationships between publishers and advertisers where some element of the deal is fulfilled programmatically. PMPs are used to sell premium inventory to a selected group of advertisers through invitation-only auctions for deals with prespecified parameters (e.g. audience demographics). In general, prices are higher in PMP than RTB (Adshead et al., 2019).

4 Intelligent Tracking Prevention

Around the world, legislators have responded to widespread privacy concerns in the digital economy with increased consumer protection regulation.[6] The most notable examples are the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA). Under both policies, consumers are given more control over their data and options to opt-out of targeted advertising. GDPR, in particular, challenges many established practices in programmatic advertising around transparency, consent and data sharing (Information Commissioners Office, 2019).

Whether in response to the legislation or to the same consumer pressures, mobile devices and internet browsers have increased their default privacy protection. Apple's Safari and Mozilla's Firefox now block all third-party ad tracking by default. Google also announced its intention to ban all third-party cookies in Chrome, although the implementation of the ban has been significantly delayed while it consults on less privacy-invasive alternatives.[7] Besides Safari, Apple has more

[5] See e.g. https://www.wired.co.uk/article/browser-fingerprinting-tracking-explained

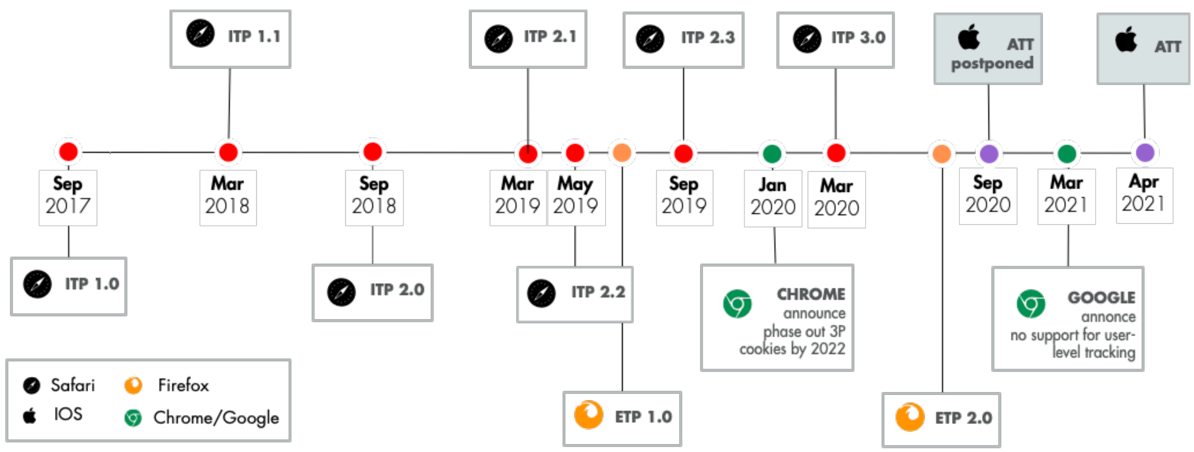
[6] For a summary, see https://piwik.pro/privacy-laws-around-globe/, accessed December 2022.

[7] Google's original plan, announced in January 2020, was to deprecate ad tracking with cookies by early 2022. This plan has been delayed by at least three years while Google consults on an alternative set of protocols called the Privacy Sandbox (Google, 2022). The Privacy Sandbox is a set of proposed open-standards that allows some level of targeting and measurement for advertisers while affording greater privacy protection for consumers. The proposals envisioned by the Privacy Sandbox coarsen information flows between advertisers, publishers and intermediaries. The implications for competition and welfare are sensitive to information design (Elliott et al., 2021). Google is working with the Competition and Markets Authority to win its approval for the Sandbox changes (CMA, 2021).

generally restricted ad tracking on all of its IOS devices. App Tracking Transparency (ATT) was introduced in April 2021 and requires all third-party applications to ask users for permission to track them for advertising. Peer-reviewed analysis of ATT's impact are still emerging (Cecere and Lemaire, 2021; Kesler, 2022), but industry sources believe the impact on advertisers and social media platforms is large. For example, ATT's impact on social media platforms is estimated by industry sources to be ~$18bn.[8]

Our paper focuses on Apple's introduction of Intelligent Tracking Prevention (ITP) which preceded ATT, affected both desktop browsers and mobile devices and was gradually rolled out between September 2017 and March 2020. Figure 2 summarises the timeline for ITP releases in the context of other browser and device-lead ad tracking restrictions.

Figure 2: ITP Timeline



ITP affects third party tracking in Safari browsers. Safari holds about 10-15% of the browser market globally, but with marked variation across countries particularly on smartphones and tables (see Figure 3). Chrome is the dominant browser with over 45% of the market.

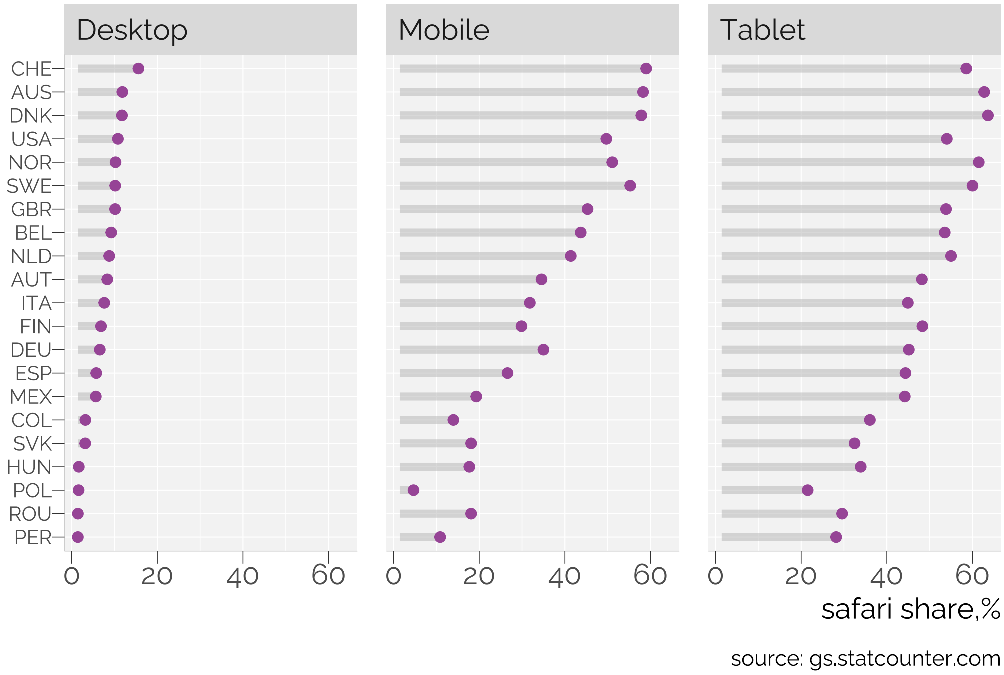
Apple launched ITP in September 2017. The aim was to severely limit and eventually prevent cross-site tracking in Safari browsers. ITP works within the Safari browser by classifying whether cookies can be used for tracking and, if so, blocking them. See for example Zawadziński (2017).

With every browser update, ITP became progressively more restrictive as Apple combated industry workarounds, unintended consequences and moved toward automatically blocking all third-party cookies. Table 2 shows the updates to the policy and the key changes.

ITP limits the user information available to advertisers for targeting and measuring advertising. Because each update had a different focus, we can separately identify the impact of changes in available information. For example, ITP 2.1 principally impacts measurement. It reduces client-side storage for first-party cookies from 30 to 7 days. So, if a Safari user visits a website more than 7 days after their previous visit, they will be counted as a new customer, which will inflate CPA metrics.

[8]See The Financial Times: https://www.ft.com/content/1959b06d-0a6e-4c37-9528-476f83626a86, accessed December 2022.

Figure 3: Safari Browser Share, 2020



In contrast, we might expect ITP 2.0 to have a smaller or even positive impact on programmatic advertising, as it was designed to reduce the advantages for Google and Facebook have over the publishers in the open web.

Table 2: Intelligent Tracking Prevention Versions

ITP Version	Released	Safari Version	Key changes
1.0	19-Sep-17	11.0	Limit retargeting. Create 24 hour grace window. Purge cookies after 30 days
1.1	29-Mar-18	11.3	Ease restrictions for 3rd Party content providers (eg video and Social Sign On)
2.0	17-Sep-18	12.0	Removal of 24 hour grace window. Limit workarounds. More significant impact on Facebook, Google and other 'walled gardens'
2.1	25-Mar-19	12.1	Limit measurement. Client-side cookies can only be stored for 7 days. Further limit workarounds
2.2	13-May-19	12.2	Further limitations on measurement. Reduce time for client-side storage to 1 day. Limit link-decoration
2.3	19-Sep-19	13.0	Further restrictions to link-decoration. Most cross-site tracking workarounds blocked
3.0	24-Mar-20	13.4	All cookies for cross-site tracking are blocked.

Finally, ITP 3.0 was released in March 2020 and so coincided with COVID-19 lockdowns in many markets. Demand for online advertising grew strongly and campaign periods shortened (to control costs). We might expect it to be more difficult to untangle a causal effect for ITP 3.0.

While we discuss our empirical strategy in detail in section 7, it is worth signaling here that we will treat each ITP version here as a separate event and compare ad prices within the same campaign before and after ITP introduction. We do not strive to estimate a cumulative causal effect of the entire rollout, as any such attempt would be notoriously plagued by selection, attrition and existence of highly nonlinear and non-parallel trends. We use our data to document some novel findings regarding long-term market trends in section 6 without making causal claims.

5 Data: Diode advertiser panel

Our dataset is a sample of programmatic ad auctions for 1100 anonymised advertisers. The data is randomly selected from Diode (DSP Insight Organisation DatabasE),[9] a commercial database owned and operated by a large advertising agency.

The Diode dataset is an advertiser panel. Data is available by advertiser on a large number of features including costs, volumes, ad clicks, ad context (e.g. publisher site), creative (e.g. format and size) and user device, browser and operating system. Diode is available from 2016 onwards and currently collects about 2bn auctions per day.

To minimise storage costs, the data is available at two levels of aggregation - event-level data and aggregated daily data. An event is typically an impression, but can also be a click or a post-view conversion. Event-level data is more detailed, but aggregated data is available for longer.

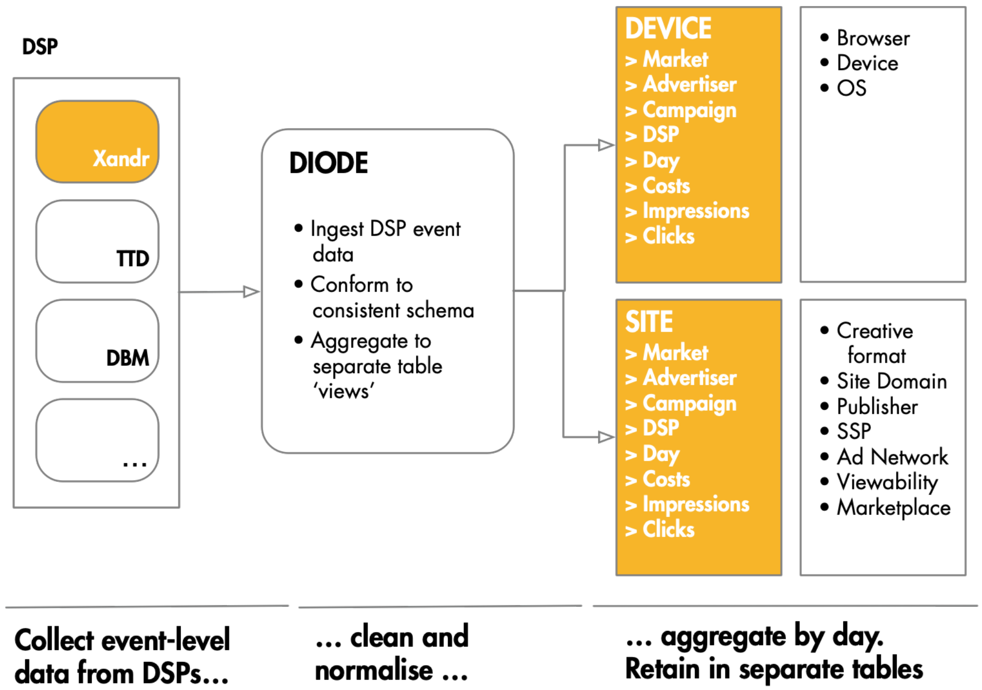
Figure 4 explains the principal data flows and the aggregation process. For each advertiser, data is collected from the DSP for each impression, click and conversion. This event-level data is then transformed to conform to a consistent schema and aggregated for efficient storage in separate tables of related metrics. The two main tables are the device table (e.g. data are aggregated at the level of device, browser and OS) and the site table (e.g. data are aggregated by site, publisher, creative format, SSP). Data on costs, impressions, clicks and conversions are common to both tables.

Our baseline analysis uses aggregate daily data from the device table with campaign-level information appended from the site table (for example, creative format and marketplace). We also use a sample of event-level data from one DSP (Xandr) and one ITP window (2.1) to test the robustness of our empirical approach. The Diode and the event-level Xandr dataset are described in more detail in Appendix A.

While extensive, the data is limited in three important ways. Firstly, in our sample, we only have a limited set of user-level data. This is partly because targeting variables, such as user interests, are not routinely collected by Diode, but also because our sample is fully anonymised to be compliant with GDPR. We cannot identify users nor do we have any information on cookies. We do have data on devices and browsers, which is sufficient for our study. Secondly, our data is limited to programmatic advertising. We do not observe advertiser spend on other media - such as search, Facebook or addressable TV. This limits the patterns of substitution we can investigate. Thirdly, Diode is for advertisers with agency representation and, hence, is skewed toward larger advertisers.

[9] Diode is not the real name of the database, which we cannot reveal in order to protect the identity of our data provider.

Figure 4: Diode Programmatic Dataset



5.1 Sample summary

Our total sample covers just over 1100 advertisers in 35 countries between 2017-2020, the period that ITP was introduced. The analysis breaks this period into seven 90-day windows, one for each ITP launch giving us 147 separate country-level ITP events with complete data. Table 3 shows that each window has typically more than 1000 campaign flights, the unit of our analysis, from advertisers in a diverse range of categories.

At a total sample level, average advertising prices are around $3.45 per 1000 impressions (Cost per Mille (CPM)) and have trended slightly up between ITP 1.1 and 2.3. As expected, clicks rates are low at about 7 clicks per 1000 impressions and do not change too much over time. Advertisers spend 7-12% of media on data for targeting.

6 Stylised Facts

Before proceeding with the main empirical part of our paper, we exploit the novelty of our data to highlight some key features of the dynamics and main features of the programmatic advertising market which have not been documented in the academic literature up to this point. Those stylised facts have informed our empirical approach to estimating the causal impact of ITP on ad prices and impressions.

Table 3: Diode Data Summary

	ITP Version						
	1.0	1.1	2.0	2.1	2.2	2.3	3.0
Counts							
Markets	15	27	37	38	39	36	36
Advertiser	76	147	168	201	246	246	179
Campaigns	153	403	570	664	825	812	517
Campaign flight	244	888	1 241	1 281	1 505	1 631	1 205
Advertiser category	14	17	17	17	18	19	18
DSP	3	4	5	3	3	3	3
Average Campaign Flight							
Impressions per flight	11.24M	3.87M	5.06M	7.81M	13.24M	9.44M	7.76M
CPM	\$5.15	\$3.89	\$2.79	\$3.47	\$3.22	\$3.51	\$3.38
Click rate	0.13%	0.17%	0.07%	0.08%	0.06%	0.07%	0.06%
Data Share (% CPM)	-	7.20%	6.32%	11.32%	11.73%	8.52%	8.78%

6.1 Safari-Chrome price premium

Programmatic CPM's have generally increased over time for our sample of advertisers. However, as shown in Figure 5, the Safari-Chrome premium, calculated as a ratio of average daily ad prices on Safari and Chrome browsers, has remained relatively flat. While these are simple averages and take no account of advertiser heterogeneity or impression characteristics, it is striking that the Safari-Chrome CPM did not decline and may even have risen slightly on desktops, suggesting that advertisers' willingness to pay to show their ads to Safari users does not exhibit an obvious decline.

6.2 CPM distribution

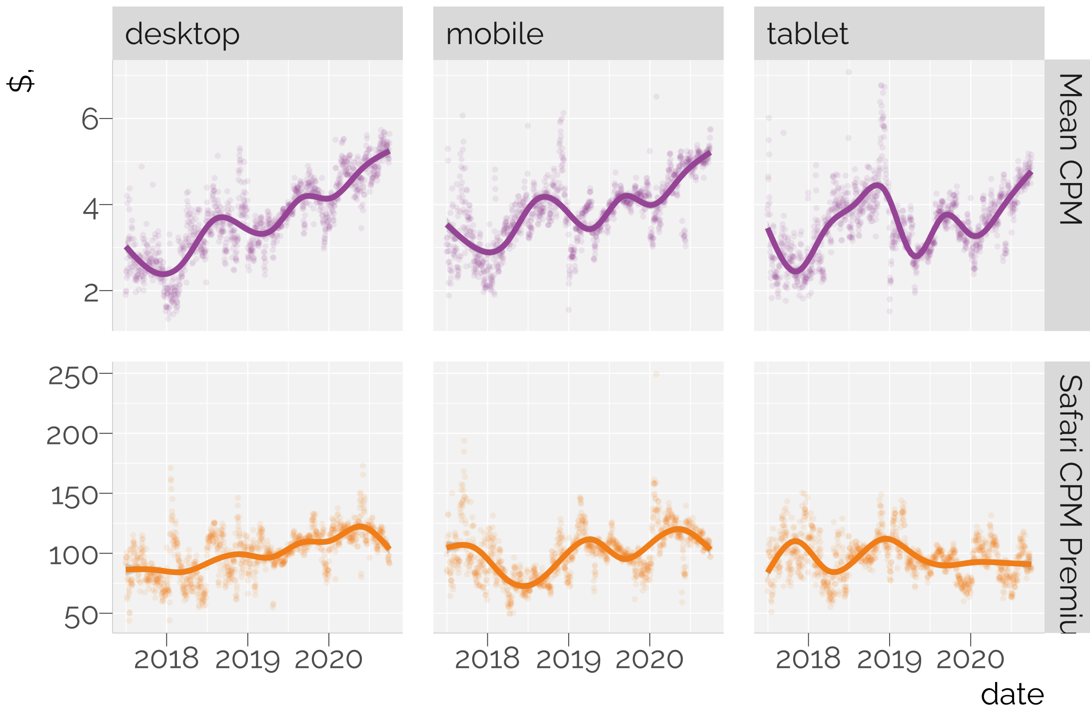
Heterogeneity matters a lot in digital advertising. There is a very large variation in CPMs across advertisers, across campaigns and within campaigns. The left panel in Figure 6 shows that CPMs vary by 3 orders of magnitude across a random selection of advertisers. It clearly shows the diversity of campaigns, even for the same advertiser, and justifies our approach (described below) to use campaign-flight fixed effects instead of comparing CPMs across different campaigns.

Even within a single campaign there remains significant variation in CPMs because of, *inter alia*, targeting, re-targeting and bid density. These are all elements that might be impacted by ITP and other privacy-enhancing policies. This heterogeneity is shown in the right panel in Figure 6.

6.3 Sites and publishers

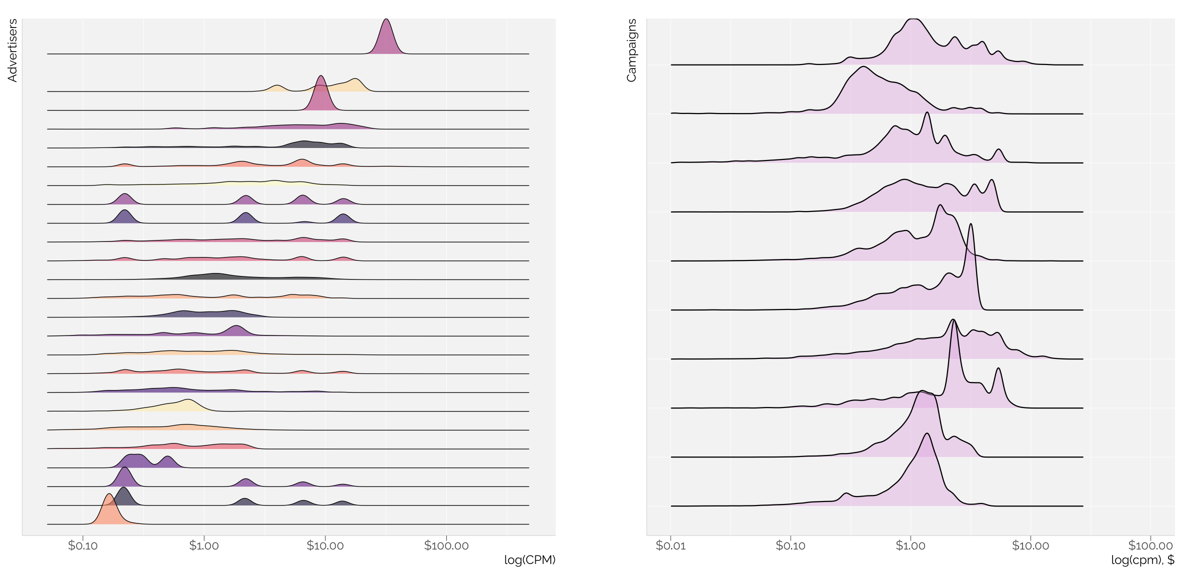
Despite the large number of sites and publishers, each campaign uses a relatively small number of publishers and sites. Figure 7 shows that top 1000 sites account for 80% of all impressions in the dataset, even in large markets like the US and GB, and the concentration is even more pronounced in smaller countries. We exploit this observation in our robustness checks to capture variation in

Figure 5: Advertising Prices Trend, Cost Per Mille (CPM), US



Note: the top panel shows the evolution of average daily CPMs for US advertisers and campaigns in our sample, split by device (desktop/mobile/tablet). Bottom panel displays the changes in the ratio between average price on Safari and Chrome browsers, which equals 100 when they are the same. Source: own calculations with Diode data.

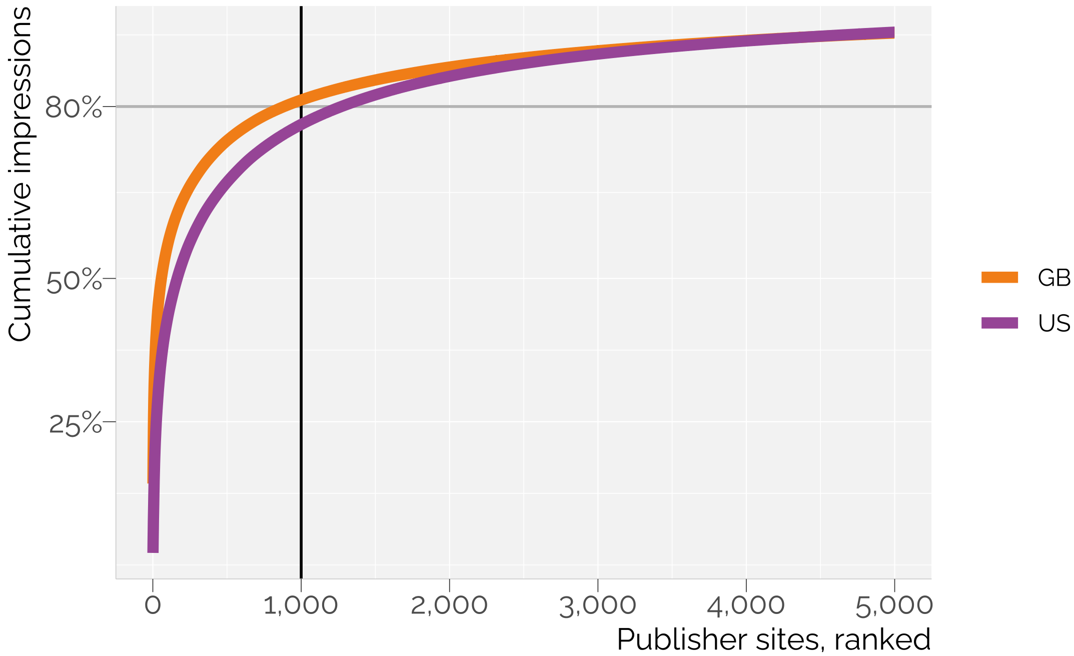
Figure 6: Ad price dispersion, across and within advertisers.



Note: left graph plots the density of log CPMs for a random sample of GB advertisers. The figure on the right-hand side shows the distribution of log CPMs for a random sample of 10 campaigns for one advertiser. Source: own calculations with Diode data.

placement quality which is not controlled for in models estimated with aggregate data.

Figure 7: Impressions by site, GB and US



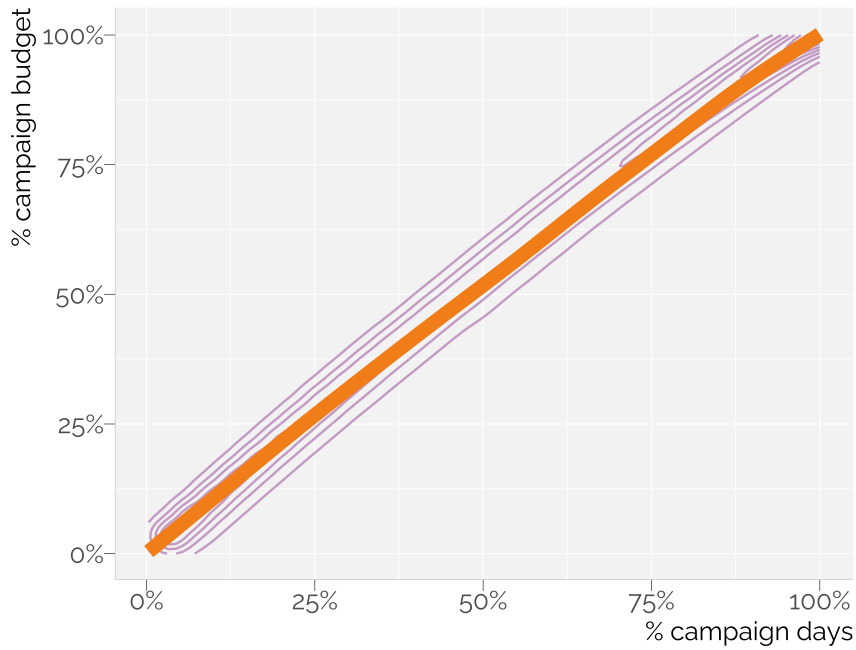
6.4 Campaign pacing

Advertiser budgets are typically evenly paced over time, with low variation in the daily budget (see Figure 8 for US campaigns). Intermediaries are often paid as a percentage of CPMs, so incentives to underspend campaign budgets are very low. This underspend asymmetry, in combination with the complexity of predicting the bid landscape, may induce excessive smoothing relative to consumer demand and media cost seasonality. One implication is that advertisers and DSPs might respond to (particularly unexpected) ITP-induced price reductions by bidding on higher quality sites (e.g. popular websites (e.g. national news outlets) or those attracting more desirable demographic profiles). This is one of the issues we test in our robustness section.

7 Empirical strategy

We chose our empirical approach to suit the heterogeneous nature of advertising campaigns, markets and ITP releases. We estimate the causal effect of ITP on programmatic ad prices and quantities using a difference-in-differences framework, separately for each ITP release and country. We select campaigns which began at most 30 days before the implementation date and ran for at most 60 more days and compare market outcomes before and after the change. In our setup, ITP-enabled Safari browsers are the treatment and other browsers are the control group. As described in Section 3, goals, audience strategy and budgets are set at the campaign flight level. As a result, campaign fixed effects can account for a large degree of heterogeneity. By looking at outcomes for the same

Figure 8: Budget pacing: campaign day vs. budget, US



Note: the figure shows that campaign budgets are spend equally throughout the duration of the campaign. The thick orange line shows the mean % of the total budget spent after a given % of campaign days. Source: own calculations with 23,539 US campaigns in the Diode data.

campaign, we mitigate any confounding impact of campaign- and advertiser-level unobservables and minimize the possibility that our estimates capture any other events happening in a dynamically changing market.

For valid causal inference, we require that there is no anticipation, that consumer adoption of new Safari versions is unrelated to ad outcomes, and that Apple's ITP policy is an exogenous shock to the supply of (potentially) targeted advertising. If anticipation were an issue, we would expect increased demand for Safari and a spike in prices prior to the ITP event date. We find no evidence that this is the case. In terms of adoption of new Safari versions, it seems likely that the speed of updating is driven by a variety of random factors unrelated to ad outcomes.

Average treatment effects on the treated (ATET)[10] are estimated separately for each market and ITP event using aggregated daily data. There are 147 such launch events in our dataset. We also put special emphasis on ITP 2.1 and estimate additional models using event-level data from a single DSP to test whether our baseline results are sensitive to the inclusion of additional characteristics of the publishers and the fact that adoption of new versions of Safari might be staggered.

7.1 Econometric specification

While our aim is to estimate the overall impact of ITP on programmatic advertising, we expect the effect to vary markedly by marketplace, advertiser and consumer characteristics. Accounting for

[10]We abbreviate Average Treatment Effect on the Treated to ATET so as not to create confusion with Apple's successor to ITP on iOS, App Tracking Transparency (ATT).

and understanding this heterogeneity is a central challenge for our analysis. Wooldridge (2021) recommends a consistent and extensible modelling framework to estimate heterogeneous treatment effects in difference-in-differences models. A key feature of the Wooldridge approach is that the two-way fixed effects (TWFE) estimator is shown to be numerically equivalent to pooled OLS - a significant computational advantage in our big data setting, where we have a large number of samples and models.

We analyse two outcome variables, log advertising prices ($\log(CPM)$) and log quantities ($\log(impressions)$. We could have also analysed clicks and conversions. However, conversions are not accurately recorded by the DSPs and, hence, are likely mismeasured in our dataset. We explored click models, but found that, given the low empirical probability of clicking an impression (typically around 0.2%, as seen in Table 3), the results were not sufficiently robust.

The most general specification we estimate for campaign flight i, browser b and day t is given by the following equation:

$$\begin{aligned} y_{ibt} &= \tau w_{ibt} + \omega(d_{i,Safari} \cdot t) + (d_{i,Safari} \cdot t)(\mathbf{x}_i - \boldsymbol{\mu}_{Safari})\boldsymbol{\theta} \\ &\quad + w_{ibt}(\mathbf{x}_i - \boldsymbol{\mu}_{Safari})\boldsymbol{\gamma} + \mathbf{z}_{it}\delta + \xi_i + \psi_b + \lambda_t + e_{ibt} \end{aligned} \tag{1}$$

where:

- y_{ibt} is either log advertising price or log quantity
- $w_{ibt} = d_{i,Safari} \cdot p_t$ is the treatment indicator where $d_{i,Safari}$ is an indicator for Safari and p_t is the date of ITP introduction
- $(\mathbf{x}_i - \boldsymbol{\mu}_{Safari})$ are time-invariant covariates, centered against Safari subsample (treated population)
- $\mathbf{z}_{it}$ are time-varying controls
- ξ_i, ψ_b and λ_t are campaign flight, browser and time FEs
- t is a linear time-trend

τ, the average treatment effect on the treated, is the main parameter of interest. Equation (1) shows the most general specification which allows for heterogeneous treatment effects through $\boldsymbol{\gamma}$ and allows for potentially heterogeneous trends through $\boldsymbol{\theta}$, as described in section 8 of Wooldridge (2021). We estimate four variations of this core model, from the basic with only treatment indicator, campaign, browser and time fixed effects to the richest one which includes all terms in equation (1). Adding time invariant covariates centred around Safari should not impact the estimate of the ATET, but will allow us to investigate key aspects of heterogeneity, e.g. across advertiser industry, budget size and type of marketplace.

To investigate whether the industry takes time to adjust to the new privacy policies, we also estimate a standard event study model. This is a variation of the regression in (1), where we replace the term $w_{ibt}(\mathbf{x}_i - \boldsymbol{\mu}_{Safari})\boldsymbol{\gamma}$ with w_{ibt} interacted with a post-treatment time dummy for each period to allow for dynamic treatment effects. We do not make more general use of the event model because our unbalanced panel complicates estimation and interpretation. Many campaigns in our data are scheduled to run for a calendar month, so if the ITP release date is e.g. in the last week of the month, we do not observe many campaign-days more than a week after the event.

Chrome is taken as the reference browser in all models. This makes it easier to interpret coefficients from the models directly as estimates of the Safari-Chrome relative price or impression volume.

7.2 Presentation of results

We present our findings in three sections. Firstly, we describe the impact of each ITP update on overall CPMs and impressions. For simplicity, we show results for Great Britain and the United States, the two largest markets. Secondly, we investigate heterogeneity in ITP's effect, notably between Real-Time-Bidding (RTB) and Private Marketplaces (PMP). And thirdly, we discuss some of the limitations of our approach and show how the findings are robust to different data sources, model specifications and estimation approaches.

8 Impact on price and quantities

8.1 Model specification

We estimate the effects of the introduction of ITP on 147 separate occasions spread across 35 markets. For each of these ITP events, we estimate the ATET on prices (CPMs) and demand (impressions) using different versions of (1). We focus our discussion on the ITP 2.1 launch event. For this event, we also have event level data, allowing to compare results from the aggregate data to results from the event level data.

Table 4 shows a summary of the CPM models for ITP 2.1 in the US. It shows that both the ATET estimates and the baseline estimate of the Safari-Chrome CPM premium are stable and small across all model specifications. In this example, and as we shall see more generally, the overall impact of ITP is negligible and not significantly different from zero apart from the first specification, where the introduction of ITP reduced the price of Safari ads by 2%. The effect goes away after we allow for a violation of the common trends assumption by including additional trend terms, however it does not really change between models (2)-(4). Given the importance of heterogeneity, our preferred model specification is (4), which includes heterogeneity in both the effect and in pre-existing trends. We therefore decided to present the remaining results only for (4) and compare them with the most parsimonious and easiest to interpret specification (1).

Staying with ITP 2.1, Figure 9 shows coefficients from a CPM event model in the US and GB. The charts illustrate three points. Firstly, the parallel trends assumption is not obviously rejected. Secondly, the size of the ITP effect is negative on most post-ITP days but small and imprecisely estimated.[11] Thirdly, coefficient standard errors increase markedly, but not as ITP is introduced but

[11]The event model is estimated without heterogeneity, so is broadly comparable to model (1), the simple TWFE. Estimating the model with heterogeneity requires interaction between each time dummy and the complete set of covariates. For 60 post-intervention days and 15 covariates, this is an additional 900 coefficients. We have run these

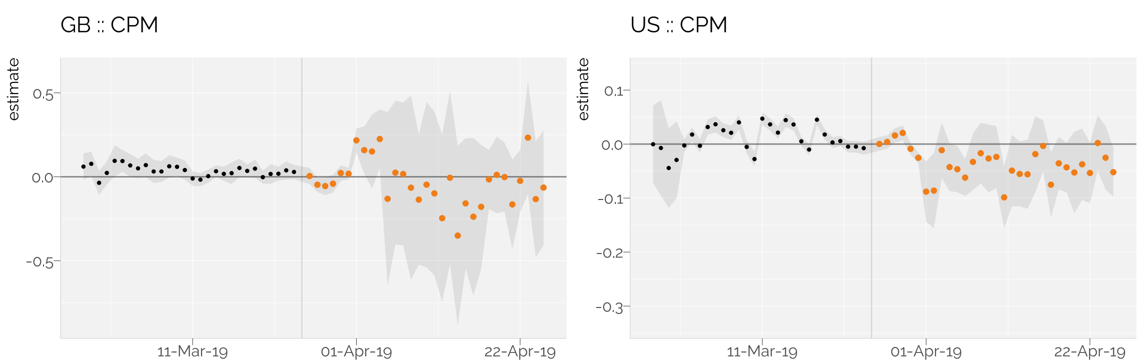
Table 4: Model specifications, US ITP 2.1

	(1)	(2)	(3)	(4)
ITP (τ)	-0.02*(0.01)	0.00(0.01)	0.00(0.01)	0.00(0.01)
Safari (ψ_{Safari})	-0.02*(0.01)	-0.03**(0.01)	-0.03**(0.01)	-0.03**(0.01)
N	445 386	445 386	444 164	444 164
Adj. R^2	0.84	0.84	0.84	0.84
Safari $\cdot\, t$		✓	✓	✓
Safari $\cdot\, t \cdot x$			✓	✓
ITP $\cdot x$				✓

Note: dependent variable: $\log(cpm)$. The 'ITP' row displays the estimate of τ, the average treatment effect on the treated. The row below shows the baseline difference between Safari and Chrome prices before the introduction of ITP 2.1. All models contain campaign, date, device, DSP and weekday fixed effects. Specification (1) only includes the treatment indicator and fixed effects, (2) adds a linear trend interacted with Safari dummy, (3) allows for conditional trends (interactions with covariates), (4) is the full model allowing for heterogeneous treatment effects. Standard errors clustered by campaign and day. * p < 0.1, ** p < 0.05, *** p < 0.01

rather as the new calendar month starts. This is because campaign flights mostly run on calendar months, so our sample of campaigns experiences significant amount of attrition at the end of the month. The issues of estimation with unbalanced panels are discussed in Section 10.3 below.

Figure 9: Safari-Chrome CPM Premium, GB and US ITP 2.1



Note: dependent variable: $\log(cpm)$. The figure shows estimates of dynamic causal effects of the introduction of ITP 2.1 on Safari ad prices with 95% confidence intervals, for Great Britain (left) and the US (right).

8.2 CPM: GB & US, all ITP versions

We now move on to the main set of results for all ITP versions introduced in the two biggest markets in our data. Table 5 summarises the impact of ITP releases on ad prices in Safari browsers. There is scant evidence that ITP had a large and significant impact on CPMs, at least for our sample of advertisers. In the US, the estimates are negative but small for ITP 2.1, 2.2 and 2.3. Even just taking the point estimates at their face value and ignoring statistical significance, one does not find an effect on price higher than -15%. As discussed in Section 4 above, the positive impact from ITP 2.0 might be expected because it impacted Facebook and Google more significantly, and hence is theoretically

more complex models for a subset of events to verify that the results are similar. They are.

Table 5: ITP impact (ATET) on Safari-Chrome CPM premium

	ITP Version						
	1.0	1.1	2.0	2.1	2.2	2.3	3.0
GB							
(1) Base TWFE	-0.042	-0.037	0.04	-0.034	-0.064	0.032	0.011
(4) TWFE w/ hetero	0.002	0.046	0.022	-0.149**	0.129	-0.025	0.033
US							
(1) Base TWFE	0.018	0.018	0.05***	-0.02**	-0.056***	-0.025**	0.106***
(4) TWFE w/ hetero	0.018	-0.073	0.099***	-0.001	-0.032	-0.018	0.019

Note: dependent variable: $\log(cpm)$. The table presents estimates of τ, the average treatment effect on the treated for two markets: Great Britain and the US and all versions of ITP. All models contain campaign, date, device, DSP and weekday fixed effects. Specification (1) only includes the treatment indicator and fixed effects, (4) is the full model allowing for heterogeneous trends and treatment effects. Standard errors clustered by campaign and day. * p < 0.1, ** p < 0.05, *** p < 0.01

positive for programmatic advertising. ITP 3.0 was released as COVID-19 lockdowns started in most countries. Besides ITP 2.1, there is no evidence of an impact in Great Britain. Overall, we conclude that, at least on average, ITP had modest or no impact on prices paid by advertisers for Safari ads. This result is surprising and so we investigate further in the remaining sections of the paper.

8.3 Impressions: GB, & US, all ITP versions

We now turn to the other outcome of interest - quantities of Safari impressions sold. Given that we find virtually no effect on prices, no visible change in demand would suggest that the industry remained in the old equilibrium and ITP did not shift advertisers' preferences. If we do, however, find a reduction in demand but no change in prices, it can mean that advertisers substituted towards higher quality impressions on Safari. Keeping the overall market size of Safari fixed around the ITP release dates, a negative effect on demand suggests that those impressions were now sold to smaller advertisers who are not in our sample.

Indeed, the impact of ITP on impression volumes appears more sizeable and significant. Table 6 shows that Safari impression volumes fell significantly in the US when ITP 1.1 was released and again for ITP 2.0. Safari volumes also decreased in Great Britain, particularly around ITP 2.2. Surprisingly, we do not observe any synchronisation between the two major markets, which suggests the importance of unobservables and their interaction with campaign-level characteristics, be they advertiser or marketplace mix or compliance with ITP (e.g. fingerprinting).

Any ATET estimates on impression volumes will be potentially biased if ITP encourages advertisers not only to buy fewer Safari impressions but to stop buying Safari altogether, i.e. if the effect goes through the extensive margin. However, as Table 7 shows, the proportion of campaigns which never use Safari is low and has, with the notable exception of ITP 2.3, been declining over time.

Table 6: ITP impact on Safari-Chrome relative impressions

	ITP Version						
	1.0	1.1	2.0	2.1	2.2	2.3	3.0
GB							
(1) Base TWFE	0.45**	0.497	-0.287**	0.204	0.063	0.141	0.276
(2) TWFE w/ hetero	-0.214	0.118	-0.391**	-0.194	-0.773***	-0.357	-0.034
US							
(1) Base TWFE	0.254	-0.91***	-0.157**	-0.046	0.032	-0.092	0.146**
(2) TWFE w/ hetero	-0.277	-0.528***	-0.042	0.114	0.057	0.08	-0.2**

Note: dependent variable: $\log(impressions)$. The table presents estimates of τ, the average treatment effect on the treated for two markets: Great Britain and the US and all versions of ITP. All models contain campaign, date, device, DSP and weekday fixed effects. Specification (1) only includes the treatment indicator and fixed effects, (2) is the full model allowing for heterogeneous trends and treatment effects. Standard errors clustered by campaign and day. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 7: Campaign flights with Safari impressions

ITP Version	Always	Some Days	Never
Before 1.0	60.2%	6.1%	33.7%
1.0	48.3%	4.7%	47.0%
1.1	67.5%	12.7%	19.8%
2.0	91.7%	6.3%	2.0%
2.1	90.9%	6.6%	2.5%
2.2	89.8%	8.2%	2.0%
2.3	56.2%	14.3%	29.5%
3.0	91.4%	6.0%	2.6%

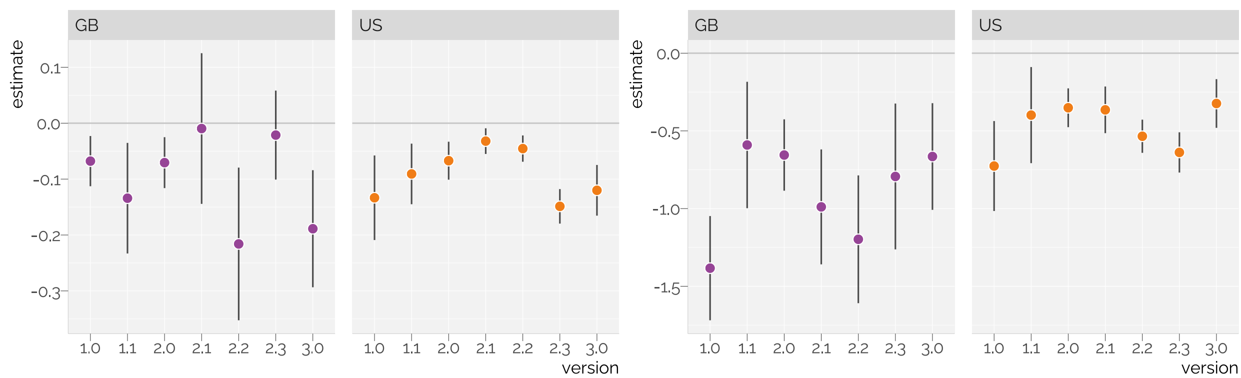
Note: for every campaign in our dataset, table shows the proportions of campaign-days with purchases of Safari impressions. The second column is the fraction of campaigns which advertise in Safari browsers always (i.e. on all days), the third column is the % of campaigns which have some days with no Safari impressions, while the last column is the % of campaigns never buying Safari.

8.4 Safari baseline

Our empirical approach is designed to estimate the causal impact of ITP in the 60 days following its introduction. It is not suited to estimate the cumulative impact. However, for each window, we also consistently estimate the baseline Safari-Chrome premium (see second row in Table 4). So, while we cannot ascribe it a causal interpretation, we can observe the evolution of the premium as an indication of the impact of ITP for our set of advertisers.

Figure 10 shows that both Safari CPMs and impressions are at a significant discount to Chrome over the period; baseline CPMs, for example, are up to 20% lower in Great Britain. Note that this result is not the same as Figure 5, because the latter is only a set of unconditional means over time, while here we estimate the mean difference in Safari and Chrome prices before the event controlling for an array of campaign and advertiser characteristics. However, there is little evidence that the premium is trending down over time. This is consistent with our main conclusion that ITP has had a small impact on our sample of advertisers.

Figure 10: Safari vs. Chrome baseline premium



Note: figures show the estimates of the Safari-Chrome difference before each of the ITP events in prices (left panel for GB and US) and quantities (right panel). Horizontal bars are 95% confidence intervals constructed using standard errors clustered by campaign-day.

9 Heterogeneity

Our base results suggest that ITP had a small impact on overall programmatic CPMs and a larger, but inconsistent impact on impression volumes.

While we believe that this result is interesting on its own, in the following section we explore various dimensions of heterogeneity to see how it varies across different types of advertisers and marketplaces, intermediaries and countries. Specifically, we address four key questions about the ITP effect:

- Is the effect larger in more competitive marketplaces which allow for instantaneous adjustments of what impressions to bid (RTB) on than in Private Marketplaces (PMP, see Section 3) where there are fewer bidders and some parameters of the deal are fixed?
- Does the effect vary across creative formats, i.e. display (banners) vs. video due to differences in scalability and cost?
- Is the ITP effect larger in countries where Safari has a greater market share?
- Do DSPs respond in a similar way to the ITP information shock?

Answers to the questions posed above will deepen our understanding of the effect of privacy policies, especially when it comes to the role of different initial conditions (Safari penetration), institutional setting (type of marketplace) and intermediaries' response to the change. We also explored how the effect varies by advertiser industry but found that results were mostly noisy and thus relegated this set of results to Appendix B.

9.1 Type of marketplace and creative format

Table 8 breaks down the average treatment effect into subpopulation-specific estimates along two dimensions: creative format and type of marketplace. We find that prices are more likely to be responsive to changing market environment in Real Time Bidding (RTB) than in Private Marketplaces (PMP). This is intuitive, since in RTB prices are determined in an auction with more competition between advertisers. Moreover, advertisers participating in open auction are less likely to be concerned about quality criteria, such as brand safety or integration with site content, and thus are typically more price-sensitive. While there is variation from window to window, the median impact on RTB CPMs is -21% in GB and -13% lower compared to PMPs.

Display advertising is traditionally cheaper to produce and hence more flexible than video. In addition, incorporating information contained in cookies (e.g. past browsing history) is easier in static images rather than videos. In Table 8 we can see that it responds negatively to most ITP introductions (notably, ITP 2.1). At the same time, the strong positive effect of ITP 3.0 relative to video likely captures the effect of COVID-19, when most advertisers froze their marketing budgets and expensive forms of advertising were strongly negatively affected. Evidence of this from an April 2020 survey of marketing professionals is documented on *The Drum* which notes that "the key brand-building category (which includes online video, TV, cinema and radio) had recorded its strongest downward revision since 2009".[12]

9.2 Geographic variation in Safari share

So far we have only focused on the two biggest markets in our data, but through our estimation approach, we have separate ITP estimates for up to 35 countries. Appendix B shows estimates for top 12 markets. In this section we investigate whether there is a relationship between the strength of ITP response and Safari's browser share.

A priori we might expect markets with a higher Safari share to be less responsive to ITP because of a trade-off with reach. Figure 11 shows that there is no strong relationship between the size of the ITP effect and browser share. At best there is a weak positive correlation between Safari share and the size of the median ITP effect for RTB.

Generally speaking, Figure 11 reveals that while the two biggest markets we focused on in the previous section experienced a near-zero (US) or small negative (GB) impact of ITP, in some other countries, notably in Scandinavia, Germany and Austria, the estimated effect is positive and significant. Figures 16 and 17 show the dispersion across ITP versions, with ITP 2.1 being most likely to have a negative effect on prices.

9.3 DSPs

DSPs are critical to price discovery in programmatic advertising. However, competitive pressure between DSPs and incentives to respond to market events, like ITP, may be muted because of asymmetric information and switching costs. For example, each DSP has its own idiosyncratic data feed which creates switching costs attributable to advertiser investment in systems and training. In

[12] See https://www.thedrum.com/news/2020/04/22/uk-ad-spend-hits-lowest-ebb-2009-crash-recovery-set-2021.

Table 8: Heterogeneity: RTB vs. PMP and Display vs. Video

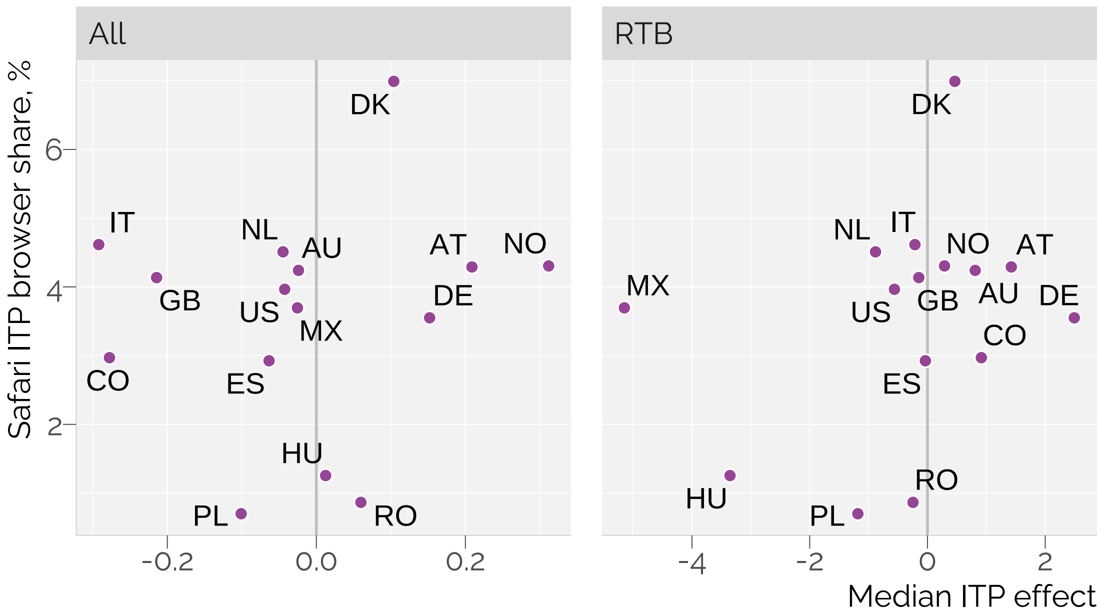
	CPM		Impressions	
	GB	US	GB	US
ITP 1.0				
Display	-0.002	-0.045	-0.86	0.386
RTB	-0.11	-0.352	1.845***	-1.624***
ITP 1.1				
Display	-0.063	0.182	-1.191	0.635
RTB	-0.598**	-0.188**	3.374***	-0.323
ITP 2.0				
Display	-0.183***	0.1**	0.997***	0.717***
RTB	-0.083	-0.451***	-0.147	-0.561
ITP 2.1				
Display	-0.385***	-0.184***	1.075**	1.428***
RTB	-0.247	-0.131	0.451	-0.909**
ITP 2.2				
Display	-0.193	0.05	-0.002	1.073***
RTB	-0.213	0.08**	-0.89	0.317
ITP 2.3				
Display	0.005	-0.09***	1.155***	1.387***
RTB	-0.214	-0.134**	-1.055	0.32
ITP 3.0				
Display	0.938***	0.453***	2.256**	0.169
RTB	-0.573**	0.042	-0.235	-0.955***

Note: the table presents estimates of γ as explained below (1, i.e. the treatment effect on the treated for different subpopulations relative to the reference category (video and PMP, respectively). We estimate these for two markets: Great Britain and the US and all versions of ITP. Standard errors clustered by campaign and day. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

our sample, over 90% of advertisers use only one DSP and even those that multihome, typically work with one DSP per campaign flight.

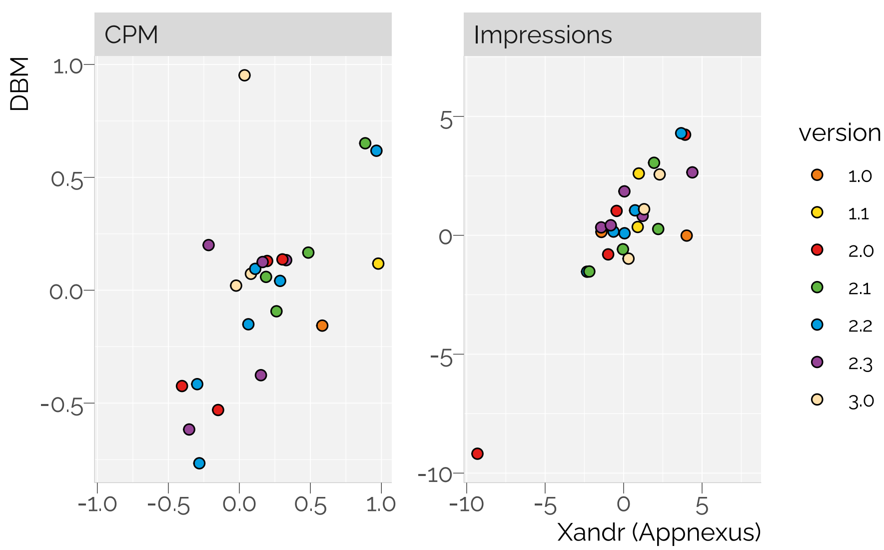
High correlation in ITP response is a tentative indicator of DSP competition. Figure 12 plots ATET estimates from country-ITP window models and shows there is a high positive correlation between ITP effects for Xandr and DBM. This correlation is indicative rather than conclusive. Firstly, it doesn't explicitly model The Trade Desk (which by constraint is negatively related). Secondly, it could be due to an unobserved campaign or advertiser characteristic that relates to how the DSP was selected. We therefore leave this result as preliminary evidence for strategic interactions between DSPs and between advertisers and DSPs. Investigating the role of DSP choice and studying the effects of DSPs market power on equilibrium outcomes is a promising topic for future research.

Figure 11: ITP impact by Safari penetration



Note: Safari browser share calculated for ITP versions over estimation period using data from Statcounter's global browser version share data (see https://gs.statcounter.com/). Browser share for ITP-enabled versions is summed for each country and averaged across the estimation period of our models. ITP impact is the median of the estimated coefficient. Two outliers (Sweden and Finland) have been removed. Note also that several other countries have been dropped from the analysis because of a lack of Statcounter browser data.

Figure 12: ITP impact by DSP across ITP windows and markets



Note: scatter plots of DSP-specific treatment effects for DBM (Google) and Xandr (Appnexus). Reference DSP: The Trade Desk. Results displayed only for ITP events where all three DSPs are present.

10 Robustness and limitations

Our main finding is that each ITP event had only a very small overall impact on programmatic ad prices in Safari browsers relative to Chrome. This section investigates how robust this finding is when we relax some of the assumptions in our empirical approach, particularly with more granular data.

We investigate three robustness issues that could bias our results - omitted dimensions of ad quality (for example, placement and creative); the accuracy of treatment exposure measurement (i.e. the adoption of browser versions supporting ITP); and the potential impact of attrition in an unbalanced panel.

We test robustness using an event-level dataset for one DSP (Xandr) and one ITP window (ITP 2.1). Besides being more disaggregated, the event-level dataset has both more impression and device characteristics than the Diode data and captures bids as well as the actual price paid (instead of CPM averaged over many websites). The data is described in more detail in Appendix A.

10.1 Controlling for publisher and creative formats

Advertising intermediaries have a strong incentive to minimise the risk of underspending campaign budgets (see Section 6). So one potential reason why average Safari CPMs do not fall is that advertisers or their intermediaries bid on higher quality impressions - for example on more premium sites.

Our aggregate Diode dataset has limited information on impression quality and site placement. The Xandr dataset records the site placement and creative format for each impression. We re-estimate our core model adding fixed effects for the top 100 sites, 8 creative categories and the interactions between the two.

Table 9 shows that estimates of the ITP effect are not sensitive to the addition of even several hundred fixed effects. The numbers can be compared with those from the ITP 2.1 column in Table 5. If anything, controlling for publisher and creative characteristics allows to explain more variation in CPMs but attenuates the effect. In the richest specification with interactions, the main effect remains indistinguishable from zero. We therefore conclude that while our baseline results are possibly biased due to omitted quality component, the direction of the bias is downward. Therefore, our main results overstate the effect of ITP.

10.2 Delayed adoption of new Safari versions

Our baseline model averages the ITP impact over 60 days post ITP release. It has no data on, and makes no assumption about, the actual rollout of ITP. That is, in the aggregate data we do not observe the version of Safari and e.g. do not know what proportion of ads were displayed to users of Safari 12.1 after the ITP 2.1 event (cf. Table 2). Our approach, therefore, effectively assumes that the treatment and ever-treated group are the same. This implicit assumption exposes our estimates both to aggregation and selection bias.

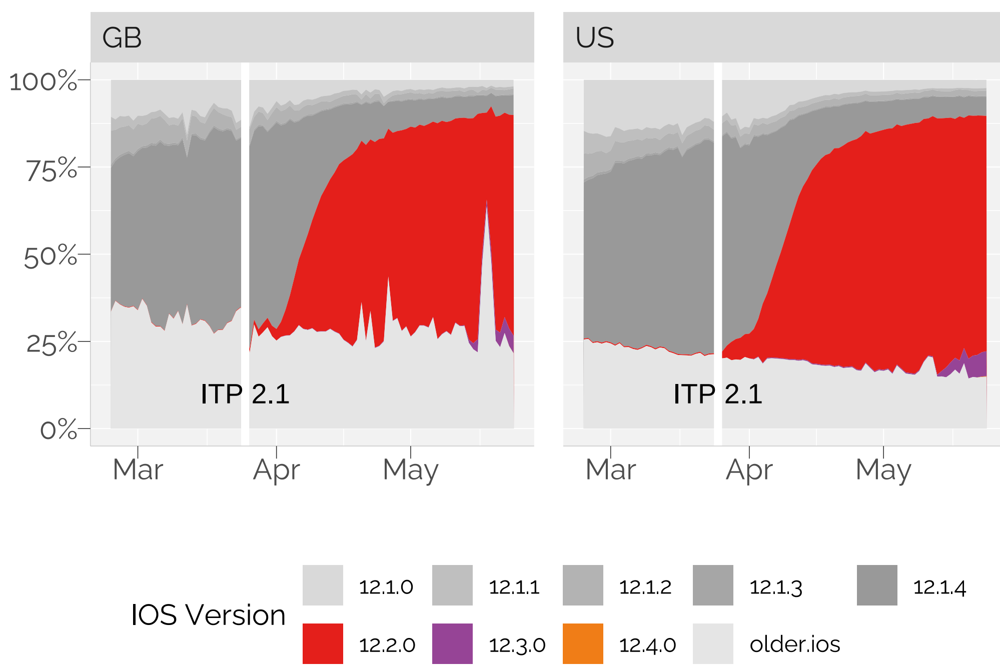
While we have no consistent browser version information in the Diode dataset, we do observe Safari version on iOS, mobile and tablet, in the Xandr event-level dataset. Figure 13 shows version share among Safari browsers. Safari 12.1, compatible with ITP 2.1, came with the introduction of iOS 12.2 which became the the dominant version within a month.

Table 9: ITP impact with site and creative controls, Xandr data, ITP 2.1

	GB			US		
	(1)	(2)	(3)	(1)	(2)	(3)
ITP 2.1	-0.026	-0.001	-0.01	-0.084**	0.002	0.005
N	602,447	594,600	594,600	3,698,256	3,638,518	3,638,518
R^2	0.655	0.705	0.714	0.561	0.651	0.673
Fixed Effects (#)						
Campaign	408	408	408	1447	1447	1447
Date	90	90	90	89	89	89
Week Day	7	7	7	7	7	7
Hour	24	24	24	24	24	24
Top Site	-	101	-	-	101	-
Creative Format	-	8	-	-	8	-
Site $\times$ Creative Fmt	-	-	543	-	-	630

Note: the table presents estimates of τ using event-level data from a single DSP (Xandr). The diff-in-diff model mimics the empirical approach employed in Section 7.1. Standard errors clustered by campaign and day. * p < 0.1, ** p < 0.05, *** p < 0.01

Figure 13: iOS version adoption among Safari users following ITP 2.1



Note: red area is the proportion of ads shown to iOS 12.2 users which came with ITP 2.1-enabled version of Safari. Source: Xandr data.

Table 10 compares ITP estimates from the full sample (cf. specification (3) in Table 9), where treatment group is defined as any Safari version, with estimates obtained from estimating the same model on a subsample of iOS-only data, where we precisely define the treated ads as those displayed in updated Safari browsers. It shows that improving precision in defining the treatment group does not impact the overall finding that ITP has had a negligible effect. In the bottom row, we have also shown that this is as true for bids as it is for actual price paid.

Table 10: ITP 2.1 impact full sample vs. iOS only, Xandr data

	GB		US	
Metric	All	iOS only	All	iOS only
CPM	-0.01	-0.035	0.005	0.008
Bid	0.028	-0.079	0.031	-0.003

Note: the table presents estimates of τ using event-level data from a single DSP (Xandr). The diff-in-diff model mimics the empirical approach employed in Section 7.1 with heterogeneous effects and trends. Standard errors clustered by campaign and day. * p < 0.1, ** p < 0.05, *** p < 0.01

10.3 Unbalanced panel and attrition

Our baseline model is estimated over a 90 day window, 30 days before the ITP release and 60 days after. Campaign flights are selected into the sample if they span the release date. However, the average qualifying campaign flight is only 35 days long, typically running over a calendar month. This is summarized in Table 11. The consequence is that the flight sample is fairly balanced for a month, and then sharply declines. Besides the impact on coefficient standard errors (evident in the event charts in Section 8), the sample attrition may trigger bias. For example, brand advertisers are more likely to run longer campaigns and to to be less sensitive to ITP. Estimates may be biased if this is not fully captured by the flight fixed effects.

Table 11: Campaign flight duration, days

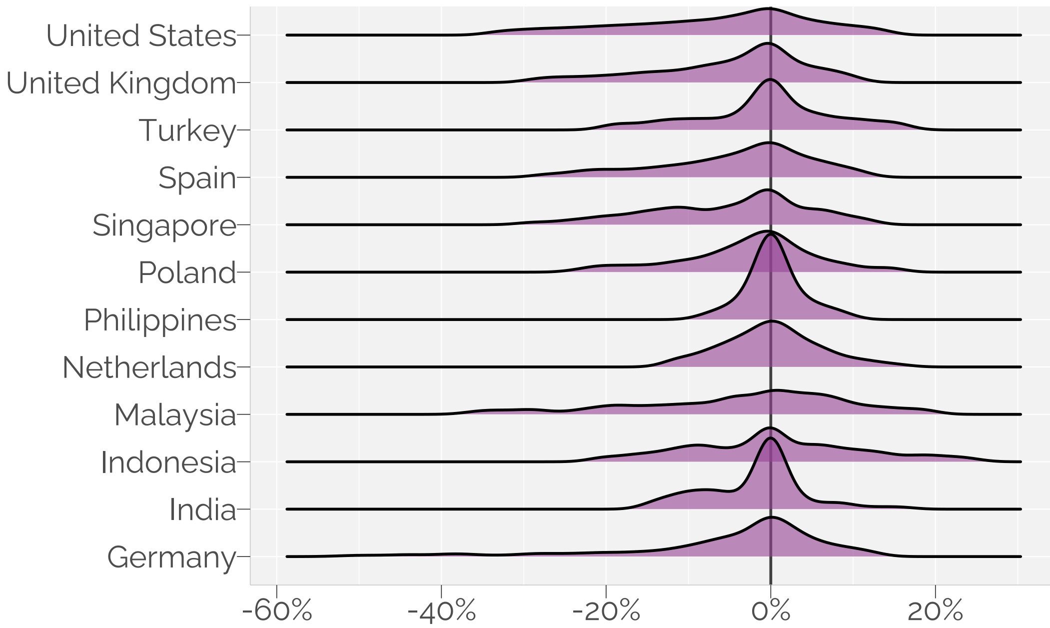
	N	Mean	Median	P0	P25	P50	P75	P100	SD
All Campaigns	42474	23.95	28	0	13	28	30	550	21.02
ITP Window	6681	34.84	29	0	23	29	30	550	41.20
Days pre ITP	6681	20.73	16	0	12	16	23	406	24.29
Days post ITP	6681	14.10	11	0	6	11	13	507	27.12

Note: summary statistics on the duration of campaign flights for the universe of all campaigns in our data.

In order to test whether our results are sensitive to attrition, we follow a different approach. We estimate a separate model for each of the nearly 6,700 campaign flights in our dataset. Figure 14 shows the distribution of ITP coefficients from these models. Almost as many campaigns witness Safari price increases following an ITP introduction as declines; the mode is zero with a slight left-hand skew. The impression-weighted mean CPM declines by 4.5% across all markets and ITP windows. This effect is significantly different from zero, but still relatively small in magnitude.

We can also use the estimated campaign-specific effects to check the robustness of our main heterogeneity findings. Figure 15 shows that the different distribution of ITP's effect between RTB

Figure 14: ITP CPM impact, campaign-level models

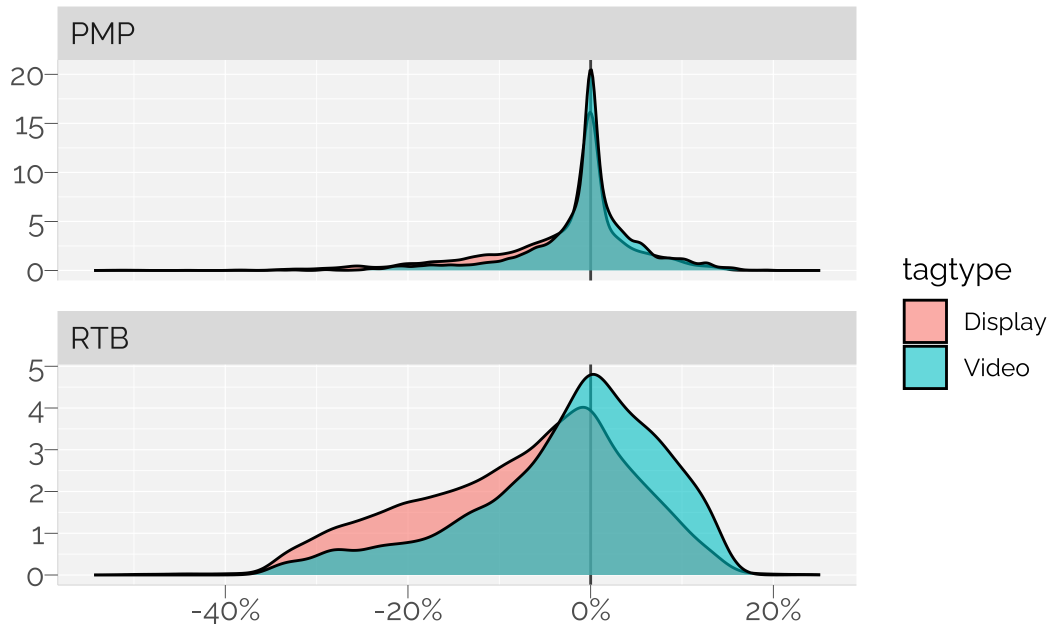


Note: Figure shows distributions of ITP effects estimated using 6681 campaign-level models and grouped by country. Bottom and top 10% of coefficients were trimmed.

and PMP is marked. As expected, the effect on Private Marketplaces is much more concentrated around 0 than it is the case for RTB. In open auctions, on average, Safari CPMs fell by 5.8% relative to Chrome, with the effect stronger for display ads than video (6.8%).

Overall, whether we use the approach of pooling all campaigns for a given country and ITP window and adding campaign FEs or estimate the effect of ITP separately for each campaign does not impact our main finding that on average, the impact of Apple's privacy policies on ad prices has been very small and in most markets nearly indistinguishable from zero. When looking at the distribution of campaign-specific coefficients, we observe large dispersion (see e.g. the distribution for US campaigns in Figure 14), with some campaigns experiencing price cuts around 20-40%, with others having to pay up to 10-15% more for Safari impressions. Uncovering the the importance of various dimensions of heterogeneity in online advertising would not be possible without our novel, comprehensive data and is thus one of the key contributions of this paper. If, as in the most of the existing literature, we only had data on a small snippet of the market (e.g. single publisher, DSP, or a very limited time horizon), our conclusions and, consequently, policy implications could well be completely altered.

Figure 15: ITP CPM impact: deal and creative type, campaign-level models



Note: Figure shows distributions of ITP effects estimated using 6681 campaign-level models and grouped by type of marketplace: PMP (top panel) and RTB (bottom panel) and creative format: video (turquoise) and display (light red). Bottom and top 10% of coefficients were trimmed.

11 Conclusions

This paper provided a large-scale evaluation of the impact of Apple's rollout of Intelligent Tracking Prevention – a new feature in Safari browsers, which allowed Internet users to protect their privacy and led to abolishment of third-party cookies – on prices and demand for online advertising. Using novel data spanning over a thousand large advertisers from different countries and industries we estimated causal effects of introducing seven, progressively more restrictive versions of ITP in a difference-in-differences framework, separately for each time window and country. Our main conclusion is that ITP itself had a negligible impact on prices paid by advertisers for most ITP updates and countries. In the US, for example, on average, the average effect was centered around zero. We did find a stronger effect on demand for Safari impressions, where updates to privacy policies were, in some cases, met with a short-lived, but sharp decline in advertisers' willingness to bid on ads shown to Safari users. Overall, our main findings corroborate two previously documented features of the advertising industry: great importance of campaign heterogeneity even if the average effect is negligible (Shapiro et al., 2021) and slower than expected adjustment to a new equilibrium (Goke et al., 2022).

The richness of our data allowed us to unpack the seemingly null result along various dimensions of heterogeneity. First, we find a predominantly negative effect of ITP on prices in open auctions (Real Time Bidding). We also observe that prices of display (static) ads are more likely to respond than videos. Further decomposition of the estimates revealed that there is at most weak correlation between the market share of Safari and size of the ITP effects across different geographic markets.

While, to the best of our knowledge, this is the first industry-wide academic evaluation of the effects of Intelligent Tracking Prevention on market outcomes, several industry sources, particularly media publishers, reported previously sizeable revenue losses which they attributed to tightening of privacy policies, including ITP. Our results are only seemingly at odds with such claims. Firstly, when estimating campaign-specific treatment effects, we found that for some advertisers, ITP caused a significant price decrease, even though the mode of the distribution was close to zero. A second possible explanation is that the equilibrium impact of reduced targeting is dampened by the increased popularity of PMPs where improved contextual targeting substitutes for third-party user data.

Despite the breadth of our data, our sample may not be fully representative of the market. Advertisers have been randomly selected from a large agency client list. By virtue of having an agency, they are likely to be larger and more complex brands. While programmatic advertising is less skewed toward SMEs than social media or search, it is possible that our sample is less focused on targeting and performance optimisation than the market as a whole. However, we see no reason from our advertiser mix why this would be the case. More importantly, if Safari prices were lower market-wide, we would expect this to filter through to Safari prices regardless of advertiser size. Consequently, we are not convinced that sample selection is an important reason for the difference.

Furthermore, our methodology might undervalue the cumulative impact of the updates. To identify causal effects, we have limited our analysis to campaigns that span an ITP launch and evaluated the impact in the 60 days after the launch. But if learning occurs primarily as campaigns are launched, with ITP effects embedded into campaign goals and targeting strategy, the impact might take longer than 60 days to fully materialise. However, our non-causal baseline estimate of the Safari-Chrome price discount is only 10-15%. If this was a significant reason for the difference, the baseline would be higher.

With both policymakers and other market participants, such as Google, announcing further changes in how user data can be used in advertising, there is still great need to understand market consequences of privacy protection. The evidence presented in this paper suggests that, as advertisers were not willing to pay much less for non-targeted ads, digital advertising and privacy protection can coexist. However, a radical, industry-wide shift to a privacy-sensitive regime would likely trigger responses from all sides of the market. Developing a structural equilibrium model of the industry which captures those different forces and allows for counterfactual simulations of market outcomes under different levels of user targeting is a fruitful direction for future research.

References

ADSHEAD, S., G. FORSYTH, S. WOOD, AND L. WILKINSON (2019): “Online Advertising in the UK,” Tech. rep., Plum.

ALCOBENDAS, M., S. J. KOBAYASHI, AND M. SHUM (2022): “The Impact of Privacy Protection Measures on Online Advertising Markets,” Working Paper.

ARAL, S. (2020): *The Hype Machine*, New York: Currency, first edition ed.

ARIDOR, G., Y.-K. CHE, AND T. SALZ (forthcoming): “The Effect of Privacy Regulation on the Data Industry: Empirical Evidence from GDPR,” *RAND Journal of Economics*.

ATHEY, S., E. CALVANO, AND J. S. GANS (2016): “The Impact of Consumer Multi-homing on Advertising Markets and Media Competition,” *Management Science*, 64, 1574–1590.

AZIZ, A. AND R. TELANG (2016): “What Is a Digital Cookie Worth?” Tech. rep.

BEALES, J. H. AND J. A. EISENACH (2014): “An Empirical Analysis Of The Value Of Information Sharing In The Market For Online Content,” Tech. rep.

BERGEMANN, D. AND A. BONATTI (2011): “Targeting in Advertising Markets: Implications for Offline versus Online Media,” *The RAND Journal of Economics*, 42, 417–443.

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

BERGEMANN, D., T. HEUMANN, S. MORRIS, C. SOROKIN, AND E. WINTER (2022): “Optimal Information Disclosure in Auctions,” Working Paper DP 16858.

BIDSWITCH (2020): “Programmatic Insights 2019,” Tech. rep., Bidswitch.

CECERE, G. AND S. LEMAIRE (2021): “Have I Seen You Before? Measuring the Value of Tracking for Digital Advertising,” 33.

CHEN, J. AND J. STALLAERT (2014): “An Economic Analysis of Online Advertising Using Behavioral Targeting,” *MIS Quarterly*, 38, 429–A7.

CMA (2020): “Online Platforms And Digital Advertising - Market Study Final Report,” Final Report, Competition & Markets Authority, London.

CMA (2021): “Investigation into Google’s ‘Privacy Sandbox’ Browser Changes,” .

DEIGHTON, J. AND L. KORNFELD (2020): “The Socio-Economic Impact of Internet Tracking,” Tech. rep., IAB.

DWORK, C. AND A. ROTH (2013): “The Algorithmic Foundations of Differential Privacy,” *Foundations and Trends® in Theoretical Computer Science*, 9, 211–407.

ELLIOTT, M., A. GALEOTTI., A. KOH, AND W. LI (2021): “Market Segmentation through Information,” Working Paper.

Esteves, R.-B. and J. Resende (2016): “Competitive Targeted Advertising with Price Discrimination,” *Marketing Science*, 35, 576–587.

Gal-Or, E., M. Gal-Or, J. H. May, and W. E. Spangler (2006): “Targeted Advertising Strategies on Television,” *Management Science*, 52, 713–725.

Ghosh, A., M. Mahdian, R. P. McAfee, and S. Vassilvitskii (2015): “To Match or Not to Match: Economics of Cookie Matching in Online Advertising,” *ACM Transactions on Economics and Computation*, 3, 1–18.

Goke, S., G. Y. Weintraub, R. Mastromonaco, and S. Seljan (2022): “Bidders’ Responses to Auction Format Change in Internet Display Advertising Auctions,” SSRN Working Paper.

Goldfarb, A. (2014): “What Is Different About Online Advertising?” *Review of Industrial Organization*, 44, 115–129.

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

——— (2019): “Digital Economics,” *Journal of Economic Literature*.

Google (2019): “Rolling out First Price Auctions to Google Ad Manager Partners,” .

——— (2022): “The Privacy Sandbox: Technology for a More Private Web.” https://privacysandbox.com/.

GroupM (2022): “This Year Next Year 2022,” Tech. rep., GroupM.

Hummel, P. and R. P. McAfee (2015): “When Does Improved Targeting Increase Revenue?” in *Proceedings of the 24th International Conference on World Wide Web - WWW ’15*, Florence, Italy: ACM Press, 462–472.

IAB (2014): “The Programmatic Handbook,” Tech. rep., Internet Advertising Bureau, London.

Information Commissioners Office (2019): “Update Report into Adtech and Real Time Bidding,” Tech. rep., Information Commissioner’s Office.

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

Johnson, G. (2022): “Inferno: A Guide to Field Experiments in Online Display Advertising,” *SSRN Electronic Journal*.

Johnson, G. A., S. K. Shriver, and S. Du (2020): “Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry?” *Marketing Science*, 39, 33–51.

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

Kesler, R. (2022): “The Impact of Apple’s App Tracking Transparency on App Monetization,” .

LAUB, R., K. MILLER, AND B. SKIERA (2022): "The Economic Value of User Tracking for Publishers," *SSRN Electronic Journal*.

LEVIN, J. AND P. MILGROM (2010): "Online Advertising: Heterogeneity and Conflation in Market Design," *American Economic Review*, 100, 603–607.

MAROTTA, V., K. ZHANG, AND A. ACQUISTI (2019): "Who Benefits from Targeted Advertising?" Tech. Rep. 1012763.

MCGEE, P. (2021): "Apple Takes on the Internet: The Big Tech Battle over Privacy," *Financial TImes*.

RAFIEIAN, O. AND H. YOGANARASIMHAN (2021): "Targeting and Privacy in Mobile Advertising," *Marketing Science*, 40, 193–218.

RAVICHANDRAN, D. AND N. KORULA (2019): "Effect of Disabling Third-Party Cookies on Publisher Revenue," Tech. rep.

RAVICHANDRAN, D. AND S. VASSILVITSKII (2020): "Evaluation of Cohort Algorithms for the FloC API," Google White Paper.

SHAPIRO, B. T., G. J. HITSCH, AND A. E. TUCHMAN (2021): "TV Advertising Effectiveness and Profitability: Generalizable Results From 288 Brands," *Econometrica*, 89, 1855–1879.

WANG, J., W. ZHANG, AND S. YUAN (2017): "Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting," *arXiv:1610.03013 [cs]*.

WOOLDRIDGE, J. M. (2021): "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators," Working Paper, Michigan State University.

ZAWADZIŃSKI, M. (2017): "What Is Intelligent Tracking Prevention (ITP)? Versions 1.0 - 2.3 Explained," https://clearcode.cc/blog/intelligent-tracking-prevention/.

Appendix

A Data

This Appendix provides additional information about our data sources.

A.1 Diode panel

Diode is a large commercial database of programmatic advertising spend for over 2000 advertisers in more than 60 countries. It is owned and operated by a large advertising agency group.

Data is collected from Demand Side Platforms (DSPs), processed to a conformed schema and stored by the agency group. The data covers key aspects of programmatic advertising, including advertisers and campaign parameters; advertisement and creative characteristics; user device and browser; publisher, site and ad view ability; intermediaries and auction types; media prices and associated costs.

Data is held in four aggregate tables - by device, by publisher, by location and by time of day and is aggregated to the lowest level of the expanded table. While these tables have many matching variables, such as campaign identifiers and auction costs, some features cannot be linked across tables. For example, it is not possible to look at advertiser creative and context choices by browser.

A.2 Xandr panel

We have also collected a more granular, event-level dataset to test the robustness of the aggregate models. The data is collected from one DSP (Xandr) and for one ITP window (2.1). As even this limited dataset is several billions of rows, we have extract a sample of 10,000 random impressions per campaign per day. Appropriate weights have been calulated to reflect campaign size in the full dataset. Table 12 summarises the Xandr dataset used to estimate models presented in Section 10.

Table 12: Xandr data summary

N	1,713,353,943
Start	2019-02-23
End	2019-05-24
Days	91
Advertisers	477
Campaigns	20,174
Countries	188
Sites	155,279
Media Spend	\$4,785,929
Impressions	1,705,742,699
Click Rate	0.221%

A.3 Variable description

Table 13 shows the variables used in the models.

Table 13: Model Variable Summary

Variable Description	Aggregate Data	Event Data
Media CPM	x	x
Impressions	x	
Clicks	x	x
Bid		x
Predicted winning bid		x
Anonymised advertiser & campaign IDs	x	x
Advertiser category	x	x
Date & time stamp	x	x
Country	x	x
Browser	x	x
Device type	x	x
Operating System	x	x
DSP	x	
Creative format & size	x	x
Marketplace & dealtype	x	x
Campaign media & data budgets	x	x
Site domain		x
Site category		x
Remarketing & conversion tags		x

Note: Aggregate models estimated with Diode daily data. Event-level models ('robustness') estimated with Xandr data.

B Supplementary results

This Appendix contains additional tables and figures referenced in the main body of the paper.

B.1 DSPs

Table 14 below is the decomposition of the average treatment effect across three DSPs, relative to The Trade Desk. Coefficients correspond to the dots in Figure 12 presented in Section 9.3.

Table 14: Heterogeneity: DSP

	CPM		Impressions	
	GB	US	GB	US
ITP 1.0				
DBM	0.584**	1.211***	-1.414	4.018**
Xandr	-0.157	0.288	0.141	-0.012
ITP 1.1				
DBM	-0.026	0.978	0.966	0.899
Xandr	-1.088***	0.118	2.6***	0.35
ITP 2.0				
DBM	0.303***	0.195	-0.829**	-0.984**
Xandr	0.137	0.129	0.402	-0.803**
ITP 2.1				
DBM	0.485	0.261***	-2.185	-0.043
Xandr	0.167	-0.093	-1.522	-0.588
ITP 2.2				
DBM	-0.282	0.063	0.729	-0.645
Xandr	-0.766	-0.151**	1.05	0.156
ITP 2.3				
DBM	-0.354	0.164	-1.428**	-0.819**
Xandr	-0.617**	0.125	0.332	0.422
ITP 3.0				
DBM	0.036	-0.023	2.297***	1.306***
Xandr	0.952***	0.02	2.56**	1.097***

Note: Estimates from the model with a full set of interactions and fixed effects relative to The Trade Desk. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Device

Table 15 below shows the decomposition of the average treatment effect of different ITP versions by device–mobile contrasted with tablet and desktop. We can see that in GB relative CPMs of ads on mobile phones went down as a result of ITP introduction for every version of ITP. The pattern is not as clear in the US, where the introduction of ITP increased demand for mobile relative to tablet and desktop.

Table 15: Heterogeneity: Device type

	CPM		Impressions	
	GB	US	GB	US
ITP 1.0				
Mobile	-0.469**	0.805***	3.621**	7.228***
ITP 1.1				
Mobile	-2.315***	0.332	2.951	5.26**
ITP 2.0				
Mobile	-0.611	-0.892**	1.979	4.725**
ITP 2.1				
Mobile	-0.24	-0.126	0.566	3.799***
ITP 2.2				
Mobile	-1.389	0.113	-2.168	0.959
ITP 2.3				
Mobile	-0.678	-0.24	-6.402***	3.581***
ITP 3.0				
Mobile	-1.1	0.118	0.585	1.776**

Note: Estimates from the model with a full set of interactions and fixed effects relative to tablet and desktop. * p < 0.1, ** p < 0.05, *** p < 0.01.

B.3 Advertiser industry

The results in Table 16 show how the ITP effects varies across advertiser industry. Numbers are missing from the table when a given combination of ITP window-country-advertiser industry was not observed in the estimation subsample.

B.4 Countries

Figures 16 and 17 present ITP effects for the top 12 countries in our dataset, as discussed in Section 9.2.

Table 16: Heterogeneity: Advertiser industry

	CPM		Impressions	
	GB	US	GB	US
ITP 1.0				
Automotive	-0.012	0.333***	0.057	0.718
Household Durable	0.03	-0.054	-0.772	-0.547
Household Product	0.001	0.061	-0.583	0.075
Household Service	0.129	-0.198	-0.326	-0.291
Retail	0.081	0.045	-0.469	-0.17
ITP 1.1				
Automotive	0.233	0.474	-3.212***	-1.042
Household Durable	-0.41	-	-1.354	-
Household Product	0.435**	-0.176	-0.4	-0.924
Household Service	-0.707***	-0.121	-0.175	-0.755**
Retail	0.219	0.015	0.102	-0.885
ITP 2.0				
Automotive	0.144	0.047	0.85**	-0.057
Household Durable	0.238	0.016	-1.034**	-0.476
Household Product	0.366***	-0.297***	1.411***	0.073
Household Service	0.041	0.11	0.022	-1.173***
Retail	-	0.167	-	-0.923
ITP 2.1				
Automotive	0.172	-0.198**	-0.577	0.473
Household Durable	-0.632***	-0.255***	-0.096	-2.212***
Household Product	0.139	0.028	-0.003	0.31
Household Service	-	-0.046	-	0.231
Retail	-	-0.033	-	-0.315
ITP 2.2				
Automotive	-0.2	-0.054	1.793**	1.329***
Household Durable	0.184	-0.24***	-0.476	-0.042
Household Product	0.099	-0.058	1.473	0.417
Household Service	-	-0.114	-	0.564
Retail	-	-0.094	-	-0.01
ITP 2.3				
Automotive	-0.403	0.085	-0.781	1.233***
Household Durable	0.073	0.153	1.565**	-0.764
Household Product	0.081	0.126	0.804	-0.468
Household Service	-	-0.056	-	0.002
Retail	-	0.088	-	-0.019
ITP 3.0				
Automotive	-0.735	0.068	1.589	-1.089
Household Durable	1.135***	-0.465***	4.14***	-0.492
Household Product	-	-0.161	-	-1.008
Household Service	-	0.158	-	-0.79
Retail	-	-0.28**	-	-0.558

Note: Estimates from the model with a full set of interactions and fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 16: ITP impact by market

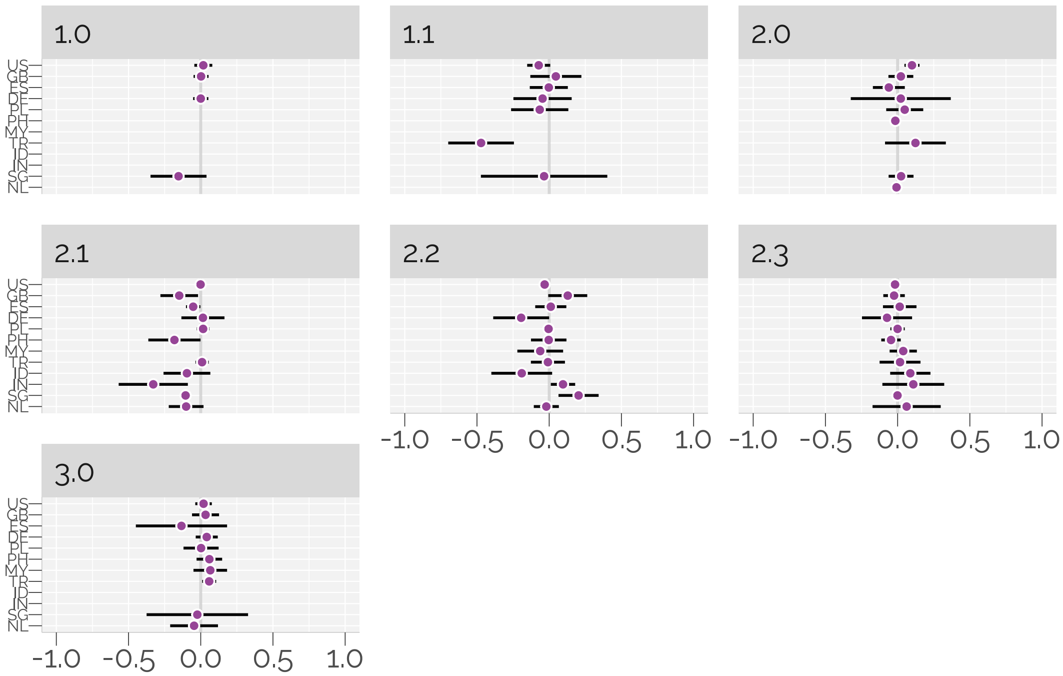


Figure 17: ITP impact by market/country: RTB only