Unearthing Privacy-Enhancing Ad Technologies (PEAT): The adoption of Google's Privacy Sandbox

Garrett A. Johnson*

September 15, 2025

Abstract

As privacy regulations reshape the digital advertising landscape, the industry is turning to privacy-enhancing ad technologies (PEATs) to reconcile relevant advertising with consumer privacy. Google's "Privacy Sandbox" is the leading initiative, offering tools designed to replace third-party cookies. These technologies, however, are already under regulatory scrutiny for their impact on market competition. We provide the first large-scale longitudinal evidence on Sandbox adoption, tracking almost 60,000 commercial websites from 2023 to mid-2025 with specialized detection methods. Adoption varied across technologies: the lightweight Topics API appeared on over 40% of sites, while the more complex Protected Audience API peaked at 23% on the advertiser side and 18% on the publisher side. Uptake accelerated through early 2024 but declined somewhat after Google paused—and later abandoned—cookie deprecation. Vendor adoption remains sparse, particularly for Protected Audience API, resulting in highly concentrated markets where Google retains a leading position. This persistent market power underscores the tension between improving privacy and ensuring competition in digital advertising. Our results, updated monthly on a public dashboard, provide timely, independent evidence for policymakers, practitioners, and researchers evaluating the role of PEATs in the future of marketing.

*Johnson: Questrom School of Business, Boston University <garjoh@bu.edu>. We gratefully acknowledge the financial support of the Digital Business Institute at Boston University Questrom School of Business, the Rafik B. Hariri Institute for Computing and Computational Science & Engineering at Boston University, and the Program on Economics & Privacy at George Mason University. We thank Zhengrong Gu, Klaus Miller, Nico Neumann, and Alisha Ukani as well as audiences at the 2024 ACM Symposium on Computer Science and Law, Consumer Financial Protection Bureau, George Mason University, and Topics W3C meeting for helpful comments. We thank Aman Ahmed, Yufan Biao, Henrik Gombos, Hurshvardhan Singh, Nakul Shahdadpuri, and Yun (Klaus) Zhu for their research assistance. We thank Jerrod Howlett, Matt Martin, Ian Meyers, and Michael O'Sullivan at Sincera for data assistance. Disclosures: Prof. Johnson served as an outside expert for the Competition and Markets Authority (CMA) in the United Kingdom for its Privacy Sandbox case. This research was conducted in his personal capacity and does not reflect the views of the CMA.

1 Introduction

Digitalization revolutionized advertising by enabling highly targeted, measurable, and optimized campaigns—capabilities underpinned by cross-site and cross-app identifiers such as cookies and mobile identifiers. Advertisers gained the ability to target users based on their interests inferred from past browsing data (Goldfarb, 2013) and to measure and optimize ads by linking user ad views to outcomes like purchases (Wernerfelt et al., 2025). These capabilities created value for advertisers (Aridor et al., 2024; Deisenroth et al., 2024) and for publishers through higher ad prices (Alcobendas et al., 2023; Johnson et al., 2020; Miller & Skiera, 2023). Yet privacy concerns have driven regulatory action, such as Europe's General Data Protection Regulation (GDPR), and prompted browsers like Safari and Firefox to block cross-site identifiers.

In response, the ad industry is turning to privacy-enhancing technologies (PETs) as new methods to build a more privacy-protective future for marketing. PETs like differential privacy and federated learning seek to better balance data utility and consumer privacy, though applications are often narrow in scope and restricted to large firms. In contrast, privacy-enhancing ad technologies (PEATs) must operate at large scale and at near real-time speed. PEATs are often integrated into browsers and serve as platforms enabling multiple types of ad tech firms to transact (Johnson et al., 2022). Google's "Privacy Sandbox" is the leading PEAT initiative, which was originally designed to replace cross-site and cross-app identifiers and address related use cases (Google, 2022). Given Google's dominance in both digital advertising and web browsing, the Sandbox initiative has significant implications for advertisers, consumers, and competition.

In this paper, we track the adoption of Privacy Sandbox by almost 60,000 websites and their ad tech vendors. As the online ad market features many interdependent firms, Sandbox needs broad adoption on all sides of the market to perform. Partnering with technology startup Sincera, we developed detection methods to measure Sandbox use from its full launch in 2023 through mid-2025. Furthermore, we identify the website's

choice of ad tech partners to illuminate the vendor-level market structure. We made our findings available on a public dashboard that we updated monthly to inform the contemporary discussion of these new technologies.

Google's Privacy Sandbox includes technologies for targeting, measurement, and fraud detection. We focus on two targeting tools that are most relevant to advertisers and publishers. The first technology, Protected Audience API, enables behavioral targeting (e.g., retargeting) but keeps sensitive browsing data on the user's device. In particular, the browser stores information about advertisers that wish to target the user and, when the user visits a participating publisher site, it runs an on-device auction to select an ad to display. The second, Topics API, provides ad tech vendors with interest segments (e.g., cats, golf) inferred from the user's recent browsing activity, allowing coarse interest-based targeting. Because Google operates both the dominant web browser and the largest digital advertising business, the UK's Competition and Markets Authority (CMA) is monitoring the Sandbox to ensure these changes do not unduly distort competition.

At its peak in 2024, 22.9% of websites used the Protected Audience API on the advertiser side, 18.0% on the publisher side, and 41.2% had at least one vendor request Topics API segments. Website opt-out signals from Topics were rare (~1%). Sandbox use grew after September 2023, when Google made Privacy Sandbox technologies generally available in Chrome, and peaked in early 2024 when Google deprecated cookies for 1% of Chrome users as part of the CMA (2023)'s Sandbox performance experiment. Adoption plateaued or declined after Google announced in July 2024 it would not fully deprecate cookies, but remained steady following its April 2025 decision to abandon deprecation entirely.

We also examine the market structure of Privacy Sandbox technology vendors. Google has the largest vendor share across both the Protected Audience API and Topics API, with a particularly dominant position in Protected Audience API auction selling. While the Topics API invites greater entry and remains less concentrated, the Protected Au-

dience API market is highly concentrated on the buy side (HHI over 2,500 points) and approaches monopolistic levels on the sell side (HHI over 7,000 points).

To better understand adoption patterns, we estimate descriptive regressions relating site-level adoption to site characteristics. Results show that adoption is more likely among high-traffic sites, those with higher advertising intensity, and those serving U.S. audiences, suggesting that scale, monetization focus, and audience location are important drivers of early adoption.

Our study contributes to multiple domains by empirically examining new applications of PETs in online advertising. Regulators note a potential role for PETs in privacy compliance (ICO 2022; OPCC 2017). PETs are now used in contexts such as government statistics (Kenny et al., 2024), medical research (Warnat-Herresthal et al., 2021), social science research (Messing et al., 2020), and industry (Hard et al., 2018). In marketing, prior work has focused on PETs that transform data (Huang & Ascarza, 2023; Li et al., 2022; Tian et al., 2024) or generate synthetic data (Anand & Lee, 2023; Ponte et al., 2024; Schneider et al., 2018) to mine consumer purchase data or target marketing efficiently. Johnson et al. (2022) introduce PEATs conceptually. We extend this work by studying PEATs empirically in a setting where multiple firms deploy PEATs and adoption can be measured directly. Despite case studies and isolated datasets, evidence on the broader industrial adoption of PETs remains scarce. More broadly, we contribute to the marketing literature on privacy (Dubé et al., 2025; Goldfarb & Que, 2023), online advertising (Choi et al., 2020; Gordon et al., 2021), and their intersection (Bleier et al., 2020; Choi & Jerath, 2022; Tucker, 2012).

We also contribute to the web measurement literature—a research area primarily in computer science,[1] but also in marketing—that examines websites' use of advertising vendors and the implications for user privacy. When a user visits a website, their browser often connects to multiple third-party vendors that provide services to the site. Prior

[1]The annual ACM Internet Measurement Conference is a major computer science conference whose proceedings are dedicated to this subject area: `https://www.sigcomm.org/events/imc-conference`.

work documents vendors' roles and technologies—such as cookies—that may threaten user privacy (e.g., Englehardt & Narayanan 2016; Karaj et al. 2018; Lerner et al. 2016; Libert 2015; Nikiforakis et al. 2013). More recently, marketing researchers have used these data to study data flows, market structure, and the impact of privacy regulations (Lefrere et al., 2022; Johnson et al., 2023; Miller et al., 2025; Peukert et al., 2022). However, growing regulatory and browser restrictions on cookies threaten the transparency these methods provide. We address this challenge with a novel measurement approach that detects both website and vendor use of PEATs. In this way, we can study important developments in the marketing industry and speak to related privacy and competition policy issues.

A growing multidisciplinary literature examines Google's Privacy Sandbox. Several researchers evaluate the performance of Privacy Sandbox, for instance, using a structural approach (Alcobendas et al., 2023; Sousa, 2024). Several firms report on Sandbox performance leveraging an open experiment on Chrome's users (CMA 2025) including academic work on advertiser performance (Kobayashi et al., 2024) and publisher revenue (Gu et al., 2025). Other research considers whether PEATs align with user privacy expectations (Martin et al., 2025; McGuigan et al., 2024) and uses surveys to gauge consumer perceptions (Cooper et al., 2023; Jerath et al., 2025; Lachner et al., 2023). Both industry and academic research offers guidance for marketers using noised and/or coarsened ad measurement signals that enhance user privacy (Dawson et al., 2023; Geng et al., 2023; Shchetkina, 2025). Law scholarship addresses the Sandbox's competitive implications (Geradin et al., 2021; Nottingham, 2021), its governance (Olejnik, 2023), as well as its fit with European privacy law (Olejnik, 2024).

Prior work investigating Sandbox adoption provides complementary evidence but is often limited to snapshots in time. Rumiński et al. (2022) and Calderonio et al. (2024) track Protected Audience API use in 2022 and 2023, respectively, whereas Philipse (2024) also examines Topics APIs in December 2023. Vekaria et al. (2024) study both targeting APIs as well as Attribution Reporting API in 2024 using data from HTTP Archive. Chrome's pub-

lic statistics report the share of page loads using Privacy Sandbox features over time, but do not identify individual sites or vendors. In contrast, we construct granular, site- and vendor-level data over a two-year period to document adoption trajectories and market structure.

The remainder of this paper is organized as follows. Section 2 describes Privacy Sandbox and the specific technologies that we measure. Section 3 describes our data and measurement approach. Section 4 provides descriptive results for Sandbox use and vendor market structure. Section 5 examines characteristics of sites that use Sandbox technologies. Section 6 discusses our findings and our results dashboard. Section 7 concludes.

2 Background: Google's Privacy Sandbox

Google Chrome is one of the last major browsers to allow third-party cookies. Safari began blocking them by default in 2018, followed by Firefox in 2019. Third-party cookies are small files containing pseudonymous identifiers that help recognize the same user across different websites. This capability underpins online behavioral advertising, where advertisers target people based on interests inferred from their past browsing activity. These cookies also facilitate ad measurement by linking a user's ad exposures with subsequent actions on the advertiser's site, such as purchases or store locator searches. Ad tech vendors place third-party cookies through networks of partner websites, combining real-time ad performance data with browsing histories to optimize campaigns. Beyond advertising, third-party cookies also support functions such as single sign-on authentication, personalized content across sites, and fraud prevention.

Google (2022) is developing replacement technologies for third-party cookies under its "Privacy Sandbox" initiative. The stated goal is to enhance user privacy while preserving the economic benefits of personalized advertising, such as funding websites and

apps.[2] Google also argues that simply blocking third-party cookies has pushed advertisers toward more covert tracking methods, such as browser fingerprinting.[3]

The Privacy Sandbox encompasses more than a dozen proposals for both the web and the Android mobile ecosystem. Advertising-related use cases include targeting (Protected Audience API, Topics API), measurement (Attribution Reporting API, Private Aggregation API), and fraud prevention (Private State Tokens API). It also proposes a cookieless approach to user authentication via the Federated Credential Management API. These technologies employ a range of privacy-enhancing technologies (PETs), including on-device computation, differential privacy, and zero-knowledge proofs (see, e.g., ICO 2022 for an introduction to PETs).

Beyond advertising, Privacy Sandbox includes technologies to limit cross-site data sharing (e.g., CHIPS, Fenced Frames, Related Websites) and to mitigate covert tracking (e.g., User-Agent reduction, IP protection). Together, these tools aim to re-architect the web's data flows to reduce privacy risks while maintaining key functionality for advertisers, publishers, and users.

Figure 1 summarizes the Privacy Sandbox timeline. Google first announced the initiative in 2019, alongside its intention to deprecate third-party cookies. The CMA opened a related competition investigation in January 2021. Over the following years, Privacy Sandbox evolved through industry feedback: some proposals were abandoned (e.g., Federated Learning of Cohorts), while others underwent multiple iterations and rebranding (e.g., TURTLEDOVE $\rightarrow$ FLEDGE $\rightarrow$ Protected Audience API). Chrome began limited trials of several Sandbox technologies in 2022, culminating in their general availability for Chrome users in September 2023.[4] In January 2024, Google deprecated third-party cookies for 1% of Chrome users, randomly selected as part of an industry-wide experiment to

[2] See `https://blog.google/products/chrome/building-a-more-private-web/`.

[3] See `https://privacysandbox.com/news/working-together-to-build-a-more-private-internet/`.

[4] See `https://privacysandbox.com/news/privacy-sandbox-for-the-web-reaches-general-availability`.

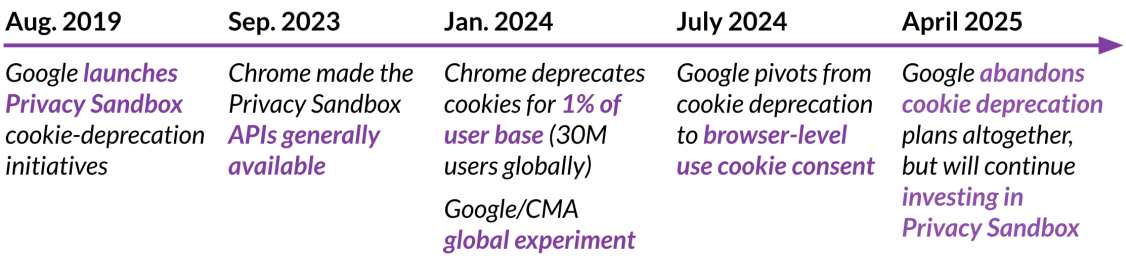


Figure 1: Privacy Sandbox timeline

assess the performance of Privacy Sandbox that was designed by Google and the CMA (2023).[5]

Google repeatedly postponed its planned deprecation of third-party cookies, issuing successive annual deadlines between 2022 and 2025. In July 2024, Google retreated from a full deprecation, instead proposing that Chrome users be required to choose whether to allow or block third-party cookies.[6] In April 2025, Google abandoned this forced choice plan as well, citing feedback from industry stakeholders and regulators.[7] Despite abandoning its cookie deprecation plans, Google continues to invest in developing and deploying Privacy Sandbox technologies.

Privacy Sandbox is part of a broader industry movement toward PEATs. Microsoft's Ad Selection API, based on Google's Protected Audience API, entered testing in the second half of 2024.[8] Apple has deployed PEATs for ad measurement on both the web (Private Click Measurement) and on mobile (AdAttributionKit), while Firefox offers Privacy-Preserving Attribution. The web standards body W3C is also considering PEATs for ad attribution measurement (Tholoniat et al., 2024, 2025), including an early joint proposal from Firefox and Meta.[9] We focus on Google's Privacy Sandbox because it remains the most comprehensive in scope and functionality, and is generally available on the web.

[5]See also `https://developers.google.com/privacy-sandbox/setup/web/chrome-facilitated-testing`.

[6]See `https://privacysandbox.com/news/privacy-sandbox-update/`.

[7]See `https://privacysandbox.com/news/privacy-sandbox-next-steps/`.

[8]See `https://blogs.windows.com/msedgedev/2024/03/05/new-privacy-preserving-ads-api/`.

[9]See `https://blog.mozilla.org/en/mozilla/privacy-preserving-attribution-for-advertising/`.

We study the adoption of two key ad-related Privacy Sandbox technologies that we discuss below: Protected Audience API and Topics API. We discuss other Privacy Sandbox technologies like Related Website Sets and their adoption in Appendix B.

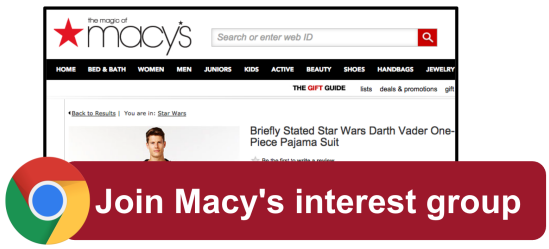
2.1 Protected Audience API

Protected Audience API (PA API) rearchitects online ad transactions so that auctions occur on a user's device.[10] Its canonical use case is retargeting. For example, as illustrated in Figure 2, `macys.com` might use PA API to retarget a user on `washingtonpost.com`. With cookies, both the advertiser and the publisher could learn that the same user visited both sites, creating a privacy risk. With PA API, the advertiser instructs the browser to add visitors to an "interest group" for future ad serving. This group, stored locally in the browser, contains the advertiser's bidding logic and a link to the ad creative, which never leaves the device. When the user later visits the publisher site, the site provides auction rules, and the browser runs the auction using the interest groups on the device to select a winning ad.

Figure 2 shows how ad tech vendors encode information to pass through PA API. On the buy-side, PA API interest groups contain up to 22 data fields (many with subfields) that define bidding functions, encode static inputs, and accommodate variations in ad size or creative components like product images. We identify buy-side ad tech vendors from the `owner` field in the interest group (e.g., Figure 2a).

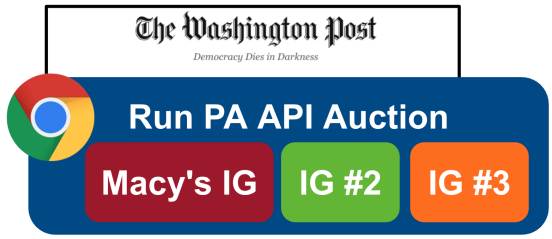
On the sell side, ad tech vendors set the `auctionConfig` for a PA API auction, which has up to 27 fields (again with subfields) specifying the auction logic and input data. The auction has a single top-level seller, identifiable via the `seller` field (e.g., Figure 2b), but may also include component auctions that feed into a top-level meta-auction. Sellers can also pre-select eligible buy-side vendors and send them signals to inform their bids.

[10] For a detailed description, see `https://privacysandbox.google.com/private-advertising/protected-audience`.




```
interestGroup = {
  owner: 'adtech-buyer.com',
  name: 'macys-pajamas',
  ads: [{renderURL=
            'adtech-buyer.com/pa?adId=123',
        sizeGroup: 'group1'}],
  // ...other interest group settings
}
```

(a) PA API interest group (buy side)




```
auctionConfig = {
  seller: 'adtech-seller.com',
  decisionLogicUrl: 'adtech-seller.com/pa_rules'
  interestGroupBuyers: ['adtech-buyerA.com',
                        'adtech-buyerB.com']
  componentAuctions: [
    {seller: 'component-sellerA.com',...}
    {seller: 'component-sellerB.com',...}
  // ...other auction settings
}
```

(b) PA API auctions (sell side)

Figure 2: Protected Audience API example with associated data fields

Notes: The right-hand side provides illustrative examples that represent the JSON elements (in part) passed to the PA API on the buy and sell sides respectively.

Early adoption shows both promise and challenges. PA API is technically complex and requires substantial engineering investment to implement. Yet by 2022, buy-side vendors such as Criteo and RTB House were already adding millions of users to PA interest groups and bidding on millions of impressions during Chrome's testing phase (Höring, 2022; Rumiński et al., 2022). These vendors reported slow adoption among supply-side platforms as a barrier, while others cited latency and debugging issues.[11] Because auctions occur locally, ads sometimes fail to load—reducing publishers' revenue (Gu et al., 2025). Buy-side vendors found that PA API reduced advertiser performance to varying degrees relative to cookie-based buying (Google, 2024b; Kobayashi et al., 2024; Parsons, 2024).

[11]See Index Exchange product team (2024), `https://github.com/grupawp/PAapi`, and `https://www.adexchanger.com/the-sell-sider/privacy-sandboxs-latency-issues-will-cost-publishers`.

2.2 Topics API

The Topics API assigns users to broad interest segments based on recent browsing activity.[12] When a user visits a participating website, ad tech vendors can call the API to learn the user's segment membership. Advertisers can use this segment information to prospect for potential customers or to refine their bidding models. Unlike PA API, ad tech vendors do not pass information into Topics API, but instead only obtain the interest segment outputs. This makes Topics API technically simpler and easier to integrate into existing programmatic ad systems.

Figure 3 visualizes how Topics API works. Each week, the browser reviews the list of domains the user visited. Topics API only classifies site domains—not the full URL address or the page content—using a system where Google manually labeled 50,000 top domains and applied a classifier model to others.[13] The interest segment taxonomy consists of 469 human-readable segments, such as "tennis" or "cats."[14] Google further splits this taxonomy by commercial relevance: e.g., "finance" and "shopping" versus "news" and "online communities."[15] From this, the browser selects the user's top five most visited categories, after first sorting for commercial relevance.

Topics API builds in several privacy protections. When queried, the Topics API returns either one of these top five categories (chosen with equal probability) or a random category (5% probability) to ensure each segment has a minimum population size ("hiding in the crowd"). Vendors only receive segments for users they have previously encountered on domains in that segment (see Figure 3), encouraging broad site coverage. Additionally, interest segments are randomized at both the vendor and domain level, so

[12] For a detailed description, see `https://privacysandbox.google.com/private-advertising/topics`.

[13] `https://developers.google.com/privacy-sandbox/relevance/topics/topic-classification`.

[14] Topics API originally included 349 segments, but its taxonomy was expanded to 469 segments in version two. The segments are also hierarchical: e.g., "/Sports/Tennis" and "/Pets & Animals/Pets/Cats." For the full taxonomy, see `https://github.com/patcg-individual-drafts/topics/blob/main/taxonomy_v2.md`.

[15] Google added this emphasis on commercial relevance in response to industry feedback (see e.g., Selman, 2022). See `https://developers.google.com/privacy-sandbox/blog/topics-enhancements`.

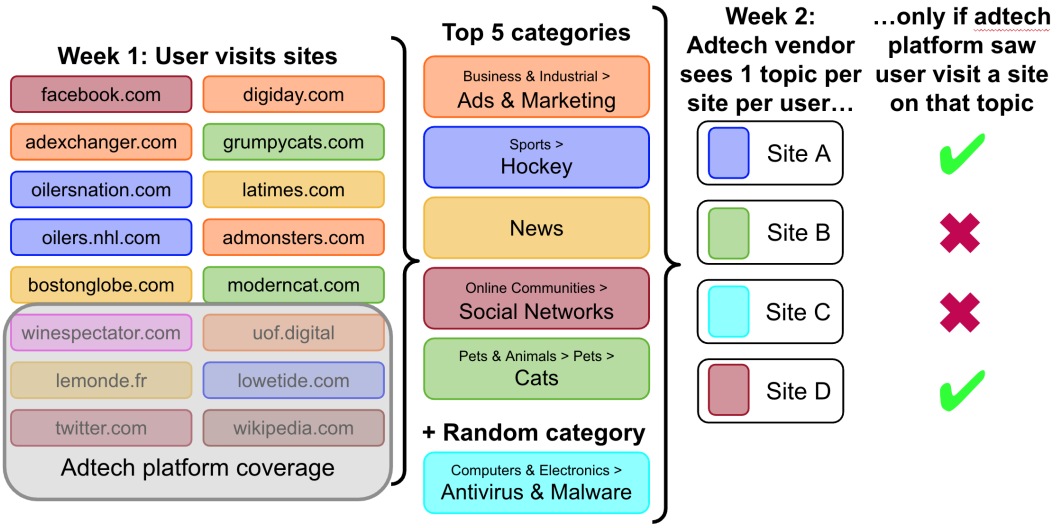


Figure 3: Topics API explanation

the same user may appear to belong to different segments across contexts, reducing fingerprinting risk (Alvim et al., 2023; Carey et al., 2023; Epasto et al., 2022). In practice, the API may return up to three segments per user, one from each of the past three weeks.

Feedback on Topics API has been mixed, with more criticism than for PA API. Thomson (2023) questioned both its privacy guarantees and advertiser utility. Beugin & McDaniel (2024) quantified its re-identification risks using browsing histories. Some vendors report alignment between Topics classifications and their own internal systems (Quéré, 2022; Selman, 2022). Nonetheless, usage in programmatic auctions has been low: RTB House found Topics segments in just 0.15% of impressions in late 2023 (Król, 2023). Google (2023a)'s internal testing suggested Topics API performs almost as well as cookie-based behavioral segments, but Criteo found its own cookie-based "in-market" segments outperformed Topics by a wide margin (Selman, 2022). Moreover, third-party testers ascribed minimal performance gains to Topics adoption in the CMA's (2025) Sandbox testing summary.

3 Data

We collect data on Privacy Sandbox technology use over time by a panel of almost 60,000 top websites. Below, we describe our data collection methodology and then describe our sample of websites.

3.1 Data collection

We partnered with Sincera (`sincera.io`) to track website adoption of Privacy Sandbox technologies over time. Sincera, acquired by ad tech vendor The Trade Desk in 2025, offers a paid service that scans the technologies and vendors used by websites, with a focus on advertising-related technologies. Similar to Builtwith (`builtwith.com`; see Stroube & Dushnitsky, 2023), which has been used in prior economics research (e.g., Koning et al., 2022), Sincera's data are better suited for in-depth evaluation of websites' advertising and privacy practices.[16] Its enterprise scanners stand out for their high scanning frequency and depth.

Sincera uses the Chrome DevTools Protocol to control a fresh instance of headless Chrome configured to mimic a real user. Scans cover hundreds of thousands of websites, typically from U.S.-based servers, with additional scans from European locations. Each site's homepage is scanned every few days, and five or more subpages are scanned every 1–2 weeks.[17] This approach goes beyond studies that only examine homepages (e.g., Libert, 2015), while offering more consistency than user-panel browsing data (Karaj et al., 2018). Like Libert (2015), Sincera examines HTML and JavaScript usage, but also inspects

[16] See `https://www.adexchanger.com/data-exchanges/meet-sincera-the-ad-tech-web-crawler-that-wants-to-disrupt-the-verification-market/` and `https://www.adexchanger.com/publishers/shadier-than-forbes-premium-publishers-are-partnering-with-content-farms-to-make-a-quick-programmatic-buck/`.

[17] Sincera randomly selects pages from links on the homepage, though the crawler's logic seeks out new content and avoids index pages (e.g., `wired.com/authors`). After June 2024, Sincera scans non-publisher websites less frequently as these sites make infrequent changes and are less relevant to Sincera's focus on advertising. Sincera uses the presence of the site's `ads.txt` file (Bashir et al., 2019) listing authorized ad sellers to classify a site as a publisher.

files such as `ads.txt` (Bashir et al., 2019).

For this study, we track adoption of the Protected Audience API (both buy and sell sides) and the Topics API (API calls and site-level opt-outs). Sincera intercepts JavaScript calls to Privacy Sandbox APIs, logging both attempted access and any content passed through the APIs (see Figure 2). We parse these raw logs to identify associated ad tech vendors by their domain names. While PA API's JSON input data is potentially rich, many fields are encrypted and thus not interpretable. Sincera also detects whether websites opt out of Topics API via their HTTP header. Note that adoption is often jointly determined by a site and the ad tech vendors it engages. Vendors may deploy PA API or Topics API on a site—sometimes without the site's prior knowledge—yet sites can opt out of Topics API and, via a similar mechanism, opt out of PA API. As such, detection of Sandbox use on a site is only a partial indicator of the site's own intent.

One notable strength of our data collection is that we capture Privacy Sandbox activity linked to the Google-CMA experiment (CMA 2023). Specifically, we randomize whether our scanner presents as a member of the experimental group that has Privacy Sandbox enabled but third-party cookies disabled.[18] This matters because some vendors—most notably Google Ad Manager, the leading publisher ad server—are far more likely to use the APIs in this condition.[19]

Our scanning approach has several limitations. First, we did not detect PA API use by Google's ad tech until February 2024, when Sincera adapted its scanners to respond to Google's pre-screening for API availability. Second, we miss API calls from vendors that block traffic from cloud data centers (e.g., non-residential IPs); in particular, this leads us to miss PA API interest group data from the vendor Criteo between November 2023 and

[18]From December 2023, Sincera represented the user as being part of this experimental group (i.e., treatment_1) half of the time and no controlled presentation for the other half. Chrome launched the relevant experiment on January 4, 2024. From May 2024, Sincera began presenting itself as part of the "label_only_1" group 25% of the time (and no controlled presentation the remaining 25% of the time). The "label_only_*" experimental group was created to allow ad tech vendors to voluntarily coordinate Privacy Sandbox testing on these users while leaving third-party cookies on. In practice, this group appears to have been used sparingly (CMA 2025).

[19]See `https://support.google.com/admanager/answer/13178817` for details.

March 2024. Third, we observe Topics API calls via JavaScript but not through other invocation methods.[20] Fourth, we may miss Sandbox activity that occurs only after explicit user interaction (e.g., consent prompts or sign in).[21]

3.2 Website panel data

We construct a website-month panel indicating whether a given site used a specific Privacy Sandbox API in that month. The panel begins in 2023 because Sincera does not retain historical logs of when a technology was first detected on a site, only the most recent detection. As a result, we build the panel ourselves by collecting repeated, contemporaneous cross-sections of Sincera's data over time. The panel ends in June 2025 when Sincera ceased operations following its acquisition by The Trade Desk.

Our starting point is Sincera's panel of top commercial websites, which grew to over 370,000 sites by the time the service shut down. Because Sincera's site selection—based on client requests and internal criteria—lacked transparency, we instead create a reproducible top site list using public rankings and documented filters. We begin with a Tranco top 100,000 site list, which combines Alexa and Majestic rankings to produce a more stable and reproducible sample for research (Le Pochat et al., 2019).

We then apply filters to align with Sincera's focus on high-quality, commercially relevant websites. First, we exclude sites from certain countries outside the Americas and Europe, based on their top-level domains (e.g., .ru, .cn). We also remove domains with non-commercial top-level domains such as .edu, .gov, and .org. In addition, we exclude

[20]Specifically, we lack data on HTTP header calls to the Topics API: see `https://developers.google.com/privacy-sandbox/relevance/topics/developer-guide` for details. Unfortunately, examining all HTTP headers for Topics API requests is impractical at Sincera's scale. Chrome's platforms statistics indicate that it detects Topics API calls via HTTP fetch requests on 6.95% of page loads and via iframe attributes on 2.82% of page loads as of March 2024. Monthly statistics for each are provided at `https://chromestatus.com/metrics/feature/timeline/popularity/4460` and `https://chromestatus.com/metrics/feature/timeline/popularity/4497`, respectively.

[21]However, our crawlers originate from the United States, which does not require consent. Moreover, research on jurisdictions that require prior affirmative user consent finds that most sites ignore this requirement for initiating third-party domain interactions (Sanchez-Rola et al., 2019).

domains containing terms associated with low-quality or unsafe content, including adult material, spam, or marketing firms. Finally, we omit sites that are unreachable, return HTTP errors, redirect to other domains, or belong to web technology vendors.

The resulting sample consists of 59,186 top sites. As of May 2023, Sincera had scanned 52,948 of these sites (89.5%), with coverage expanding to 99.8% by January 2024. Appendix A provides further detail on the sample construction process.

4 Descriptive results: Privacy Sandbox use

We document how websites and ad tech vendors adopt Privacy Sandbox technologies, focusing on the Protected Audience API (PA API) in Section 4.1 and Topics API in Section 4.2. We provide alternate measures of Sandbox adoption in Appendices B and C.

We also describe the market structure for vendors that operate Sandbox technologies. Following Johnson et al. (2023), we measure an ad tech vendor's market share for a given Sandbox technology using reach—the number of sites where the vendor is observed using the technology—divided by the total reach of all vendors using it. We quantify market concentration with the Herfindahl–Hirschman Index (HHI), which sums squared market shares on a 0 (perfect competition) to 10,000 (monopoly) scale.

4.1 Protected Audience API

PA API is a core Privacy Sandbox technology that enables interest-based ad targeting without cross-site identifiers (see Section 2.1). It operates on both the buy and sell sides of the market. On the buy side, advertisers (or their intermediaries) add users to interest groups (IGs) based on observed behavior, such as visiting a product page. On the sell side, publishers initiate on-device auctions to select ads from eligible IGs. These functions correspond to two API methods: `joinAdInterestGroup()`, which adds a browser to an IG, and `runAdAuction()`, which runs the auction locally to select an ad creative. Sec-

tion 4.1.1 discusses usage trends while Section 4.1.2 presents the vendor market structure. Appendix D discusses PA API component auctions and Appendix E examines PA API's raw input data.

4.1.1 PA API: Aggregate use trends

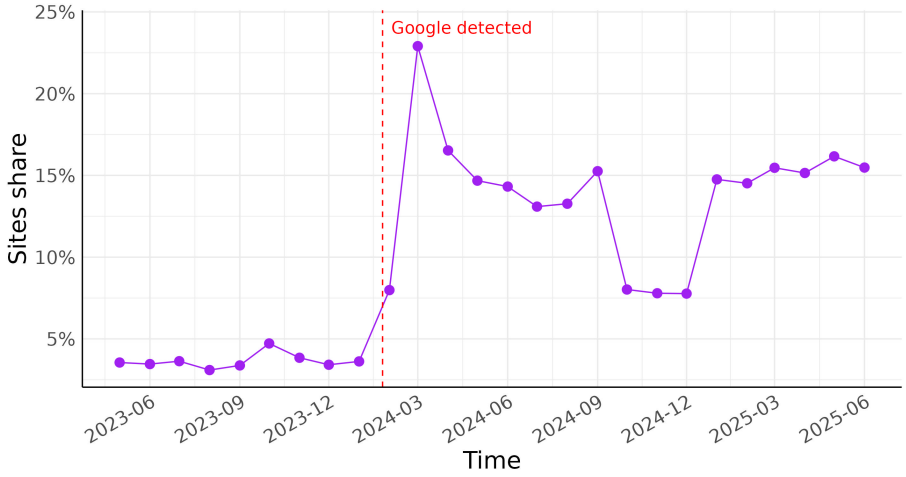
Figure 4 shows aggregate PA API adoption, distinguishing IG creation (Figure 4a) and auctions (Figure 4b). Observed PA API use was low prior to February 2024, partly because our measurement system could not detect Google's participation until that month (see Section 3.1). Before February 2024, we observe that fewer than 5% of sites created IGs, and at most 1% ran auctions. Once we detect Google's participation, observed adoption jumps to 8.0% of sites for IGs and 13.3% for auctions in February 2024.

PA API's IG adoption peaked at 22.9% in March 2024, driven by both Google and rapid growth from the second-largest vendor, Audigent (Figure 5a). A sharp decline in Q4 2024 was largely due to a temporary seven-fold drop in Google's reach between September and November. IG use then stabilized in 2025 between 14.5% and 16.2% of sites. Note that we linearly interpolate the missing Criteo IG data between November 2023 and March 2024 due to Criteo blocking data-center traffic (see Section 3.1).

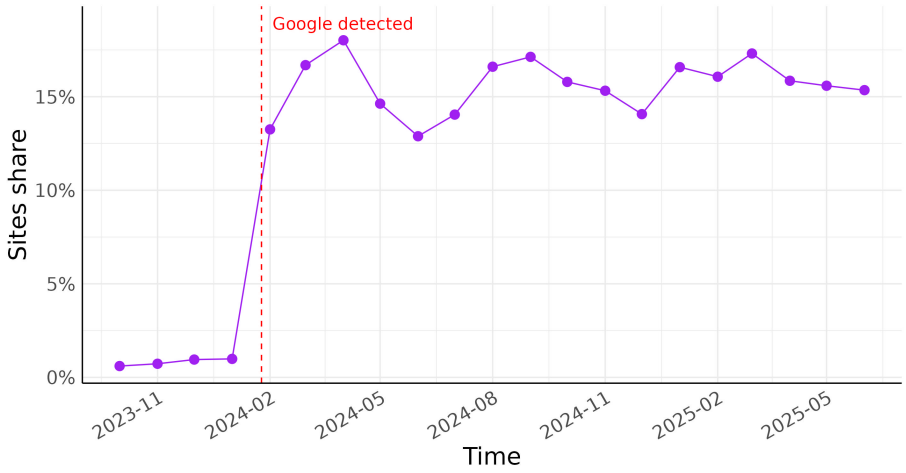
Auction adoption followed a similar pattern: steady levels after Google's February 2024 entry, a more modest dip in late 2024 linked to Google's temporary reach reduction, and stabilization in 2025 at 15.4–17.3% of sites. Neither the July 2024 announcement of Chrome's shift to user cookie consent nor the April 2025 decision to abandon cookie deprecation produces noticeable aggregate changes in PA API's IG or auction use.

4.1.2 PA API: Market structure

By intercepting data passed through PA API, we can identify the ad tech vendor that a site uses to operate the API. Vendor reach is again defined as the number of sites engaging that vendor, and we use this to calculate market shares and concentration.



(a) PA API interest groups



(b) PA API auctions

Note: As discussed in Section 3.1 and indicated by the dotted vertical red line, we do not observe sites that use Google for PA API until February 2024 for technical reasons. We also lack data on Criteo's IG use between November 2023 and March 2024 for technical reasons discussed in Section 3.1. We linearly interpolate the number of sites that only use Criteo for those dates and add this to the total sites that we observe using PA IGs for those months in Figure 2a.

Figure 4: Evolution of Protected Audience API use by websites

Interest groups (IGs). Figure 5 graphs several features of PA API's IG market structure. Figure 5a plots the monthly evolution of market shares for the top ten IG vendors. In early 2023, two of the three major initial adopters are visible in our data: Criteo, with a share exceeding 90%, and RTB House. Other studies, however, report Google's earlier participation, with notable activity in August 2022 (Rumiński et al., 2022) and Q3 2023 (Calderonio et al., 2024). Subsequent entry expanded the vendor set: Raptive (September 2023), Teads (October 2023), AdRoll and Audigent (January 2024), Taboola (May 2024), and Microsoft (December 2024). During Google's temporary retreat in Q4 2024, Audigent briefly led the market. Notable exits include Teads (January 2025), Taboola (April 2025), and Raptive (December 2024–January 2025).

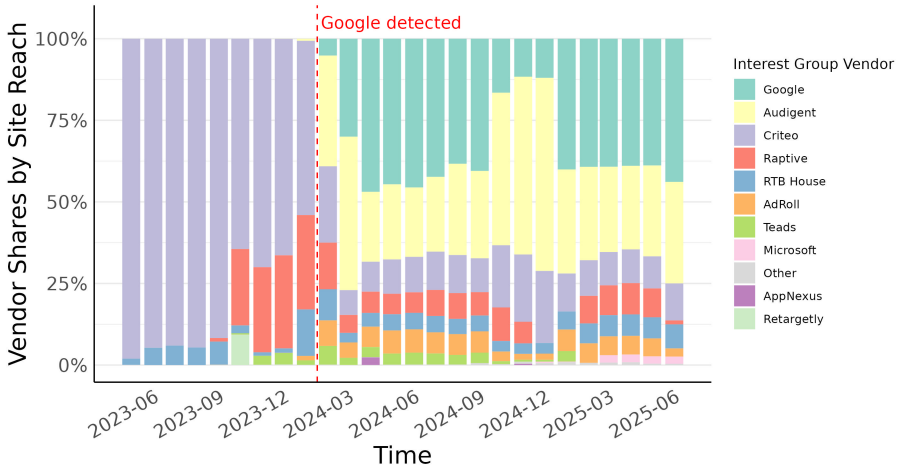
Figure 5b aggregates market shares over the post-February 2024 period when we observe Google in the market. Google (36.3%) and Audigent (31.2%) together control about two-thirds of the IG market. Criteo (11.8%), Raptive (6.8%), AdRoll (5.5%), and RTB House (5.3%) bring the combined share of the top six to 97%. The remaining vendors collectively hold under 0.1% each; these include AdForm, Amazon, Conversant, Dynalyst, Optable, Retargetly, Sportradar, and Yahoo!. Nonetheless, Amazon's actual economic role may be larger if its IG additions cover a significant share of its user traffic.

Figure 5c presents HHI values generally between 2,500 and 3,500—levels considered highly concentrated under competition guidelines[22]—with a spike in Q4 2024 during Google's temporary retreat.

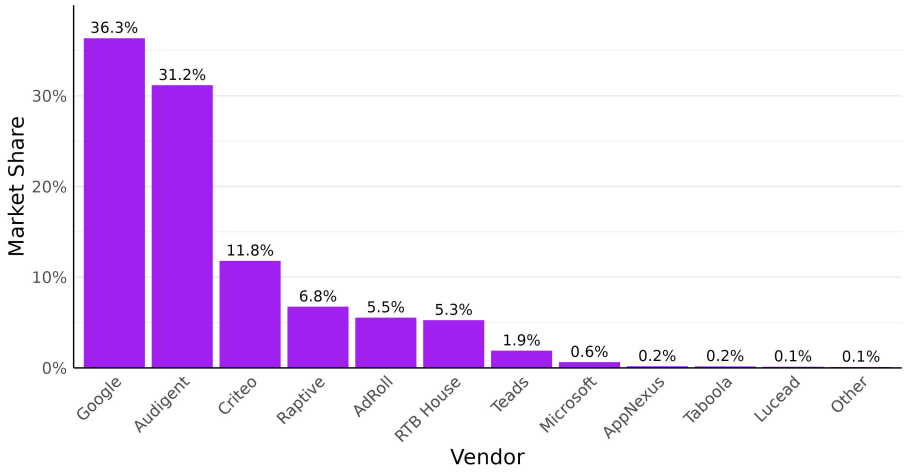
Auctions. The top-level seller is the vendor initiating the on-device auction and determining the ad shown in a publisher's slot (Section 2.1). PA API auctions can be single or meta-auctions composed of multiple "component" auctions. Figure 6 describes the market structure for the top-level seller in either case.

Figure 6a shows the monthly evolution of market shares among the four largest top-level sellers. The earliest adopters were Seedtag and Google, although Google is again

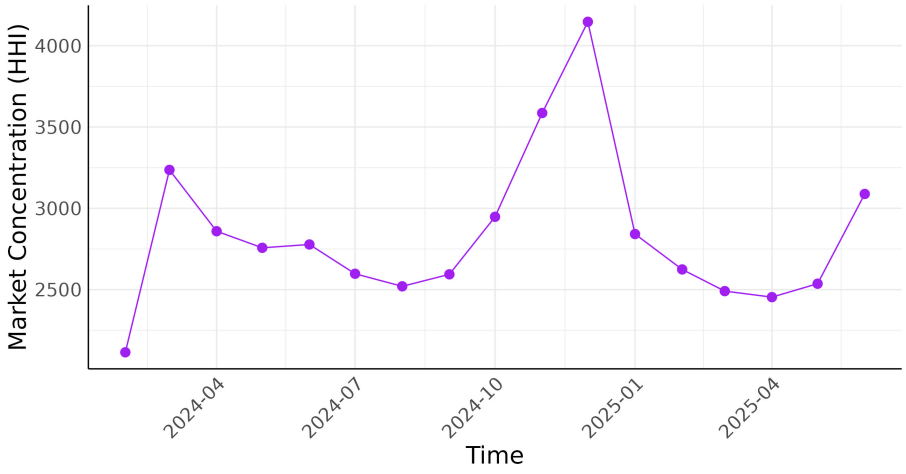
[22]The FTC & DOJ Horizontal Merger Guidelines (2010) define high concentration as HHI $\geq$ 2,500; the revised guidelines (2023) lower the threshold to 1,800.



(a) Evolution of top 10 vendor market shares



(b) Aggregated vendor market shares (Feb. '24-June '25)



(c) Evolution of market concentration (Feb. '24-June '25)

Note: As discussed in Section 3.1 and indicated by the dotted vertical red line, we do not observe sites that use Google for PA API until February 2024 for technical reasons. Given that Google is largest vendor in PA API, our aggregate market shares (Figure 5b) aggregate over vendor site-reach by month and omit data prior to February 2024. Figure 5c's HHI figures also omit data prior to February 2024. We also lack data on Criteo's IG use between November 2023 and March 2024 for technical reasons discussed in Section 3.1. We linearly interpolate Criteo's site reach over this period.

Figure 5: Protected Audience API market structure: Interest group vendors

only observed from February 2024 onward. Upon Google's observed entry, Seedtag's share falls to 6.7%. Google's share, by contrast, never drops below 80% during its observed period. Subsequent entrants include Taboola (April 2024) and Lucead (September 2024), but Taboola exited in May 2025.

Figure 6b also presents aggregated market shares from February 2024 onwards. Google dominates with 88.0% of the market, followed by Taboola (6.4%) and Seedtag (5.3%). All other vendors combined account for only 0.3%, most of which is Lucead. Remaining vendors have negligible reach and include: AppNexus, Bounce Exchange, Coupang, Microsoft, `Onet.pl`, Teads as well as pre-February usage by Onetag and the domain `ayads.io`.

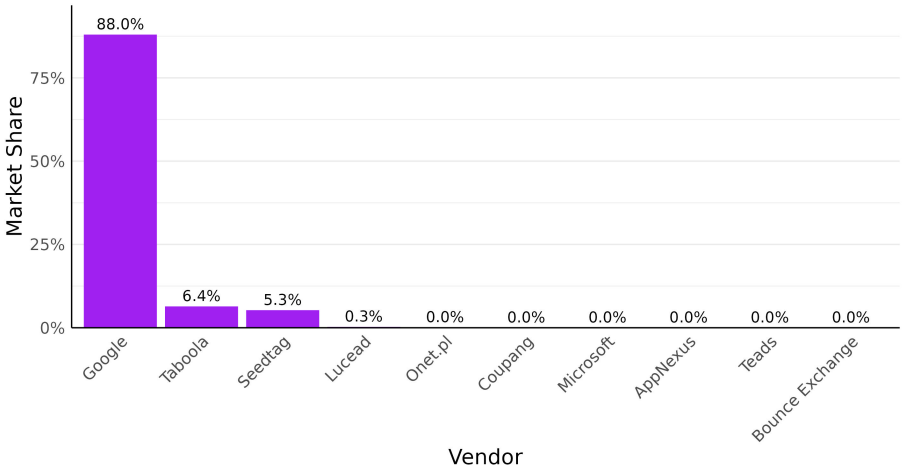
Figure 6c shows concentration levels consistently above 7,000 except for October 2024. This market therefore approaches monopoly levels and spiked to about 8,750 after Taboola's exit in May 2025. While Google's share is dominant, it often hosts component auctions in which other vendors participate as sellers, meaning these top-level HHI figures may somewhat overstate Google's competitive position. We explore this possibility in Appendix D where we find that Google also has a dominant share as a component-auction seller. We see more vendors enter as component-auction sellers than top-level sellers (Figure 6b); nevertheless, the component-auction seller market exhibits high concentration as well. Future work will examine the market structure for ad vendors that are listed as buyers in PA auctions.

4.2 Topics API

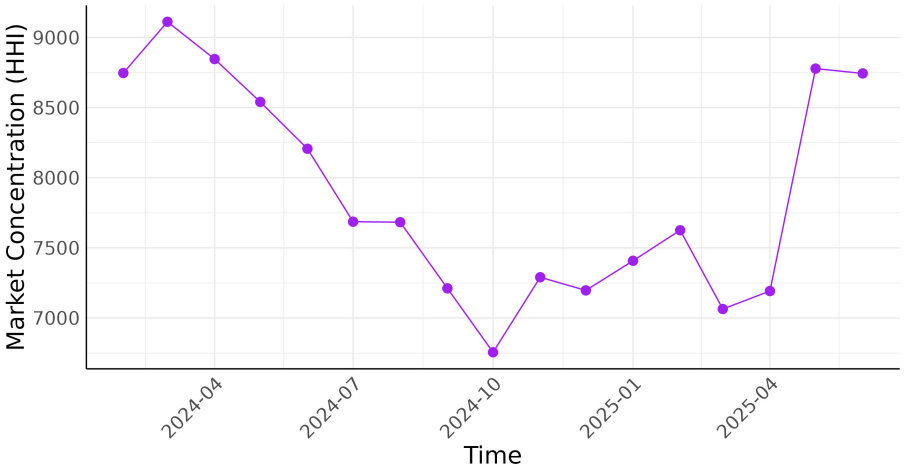
As described in Section 2, Topics API allows ad tech vendors to request a user's interest segments from the browser. Section 4.2.1 reports adoption trends as well as the share of sites that opt out of Topics. Section 4.2.2 examines vendor market structure. Finally, Section 4.2.3 analyzes how Topics API classifies the domains in our sample.



(a) Evolution of vendor market shares



(b) Aggregated vendor market shares (Feb. '24-June '25)



(c) Evolution of market concentration (Feb. '24-June '25)

Note: As discussed in Section 3.1 and indicated by the dotted vertical red line, we do not observe sites that use Google for PA API until February 2024 for technical reasons. Given that Google is largest vendor in PA API, our aggregate market shares (Figure 6b) aggregate over vendor site-reach by month and omit data prior to February 2024. Figure 6c's HHI figures also omit data prior to February 2024.

Figure 6: Protected Audience API market structure: Auction vendors (top-level seller)

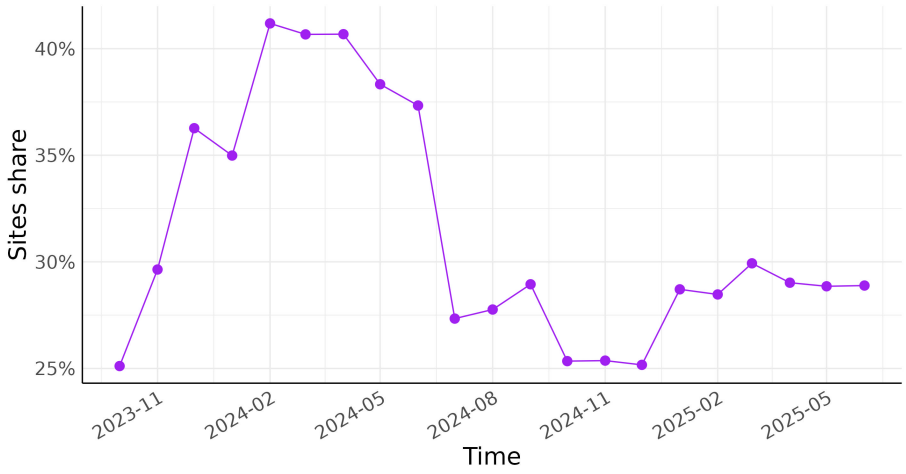
4.2.1 Topics API: Aggregate use trends

We record calls to the browser's Topics API that request the user's behavioral segment information. Figure 7a shows that adoption grows from 25.1% of sites in October to a peak of 41.2% of sites in February 2024. The share fell slightly in Q2 2024, then dropped by about 10 percentage points in July 2024. This drop may reflect industry reaction to multiple developments: 1) Google's announcement that Chrome would move to a user-consent approach for cookies, and 2) the API's muted performance results from industry testing in early 2024 (CMA 2025). Some of the drop may also reflect substitution toward API call methods we do not observe, as we only track JavaScript calls in our data (see Section 3.1).[23] Despite these shifts—and Google's April 2025 decision to abandon cookie deprecation—Topics API usage stabilized around 29% of sites through 2025. This notably exceeds the share of sites using either PA API on the buy or sell sides. Topics API's relative simplicity likely contributes to its broader uptake: any ad tech vendor present on a webpage can call the API without the website's explicit permission, and the API's segment output can enrich traditional programmatic ad transactions.

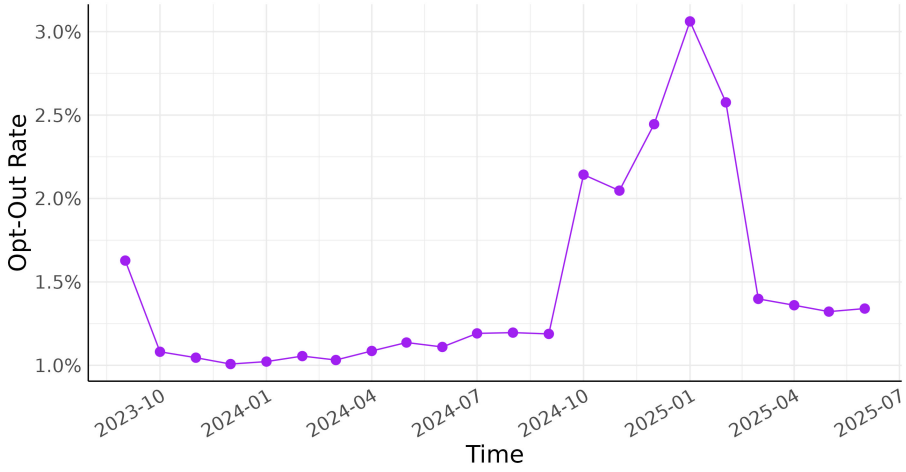
Websites can opt out of Topics API by including an opt-out flag in their HTTP header.[24] Opting out prevents both the inclusion of the site's domain in the API's segment assignment and the receipt of Topics segment data. Opt-out sites therefore forgo both the benefits and costs of participation. Figure 7b shows the opt-out share falls from 1.6% in September and remains around 1.0–1.2%, except for a brief spike above 2% in early 2025. We do not know the cause of the 2025 jump, though we speculate that a network of sites is

[23]Chrome's page load statistics show a rapid increase in `TopicsAPI_BrowsingTopics_Method` (see `https://chromestatus.com/metrics/feature/timeline/popularity/4182`) from July (0.5%) to December 2023 (11.3%) and a gradual downward trend through June 2025 (9.7%). The steeper decline in our data could result from a combination of our unit of analysis (site versus page load), site sample, and emphasis on the minority of users in the Chrome-CMA experiment groups (CMA 2023). However, the Chrome statistics show a modest increase in the share of page loads with HTTP header calls to the Topics API (which we do not observe): 6.8% of page loads in February 2024 to 9.7% in June 2025 (see `https://chromestatus.com/metrics/feature/timeline/popularity/4460`).

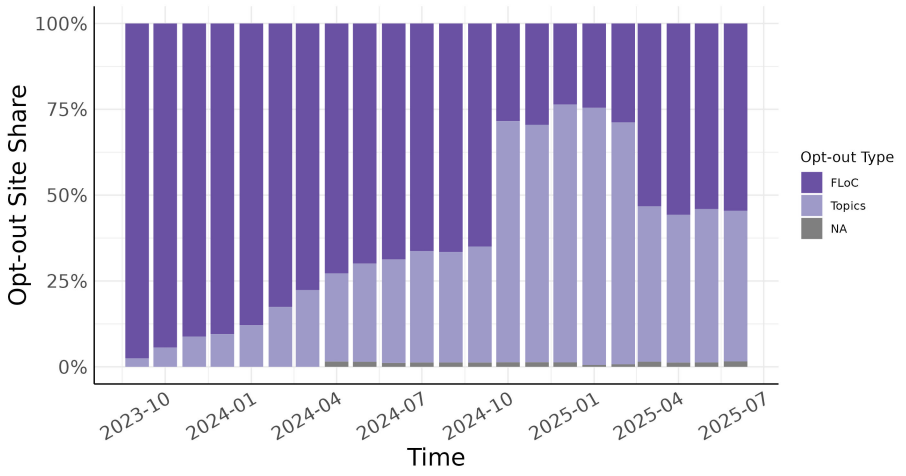
[24]Specifically, websites declare this in their "Permissions Policy," where web developers can specify which features are allowed or blocked on a website (see `https://developer.mozilla.org/en-US/docs/Web/HTTP/Permissions_Policy`).



(a) Calls to Topics API



(b) Opt-out share



(c) Opt-out type share

Figure 7: Evolution of Topics API calls & opt-out choices on websites

responsible. Earlier data from July 2022 bounds the opt-out rate between 2.7% and 3.9%.[25] The decline between 2022 and 2023 rates may reflect the shift from the more controversial Federated Learning of Cohorts (FLoC) to Topics API in January 2022. While similar in assigning behavioral segments, FLoC supported thousands of unlabeled segments and drew heavier privacy criticism (Berke & Calacci, 2022). In particular, FLoC carried greater fingerprinting risk both because the segments were more granular, and because Chrome would share a consistent segment label with participating vendor domains rather than randomizing labels at the vendor and site level. Moreover, the unlabeled segments were inscrutable to users, raising concerns that FLoC segment labels would reveal sensitive user information despite built-in protections (Medina et al., 2021).

Google retained the same opt-out mechanism for both FLoC and Topics, and sites could specify either or both. Sites did not need to update this flag, however, as Google announced that Topics API would respect both FLoC and Topics opt-out signals. Nevertheless, Figure 7c shows that in September 2023, only 2.5% of opt-out flags referred to Topics, but this share grew to 43.9% by the end of the sample. These patterns suggest that websites were updating their privacy settings in response to Google's evolving proposals.

4.2.2 Topics API: Market structure

From the domains making Topics API calls, we infer the calling ad tech vendor. Unlike PA API, vendors do not transmit data to the Topics API, only receive it.[26] Figure 8 summarizes market structure statistics, showing a broader and more diverse vendor base than

[25] As mentioned in Section 3.2, Sincera's historical data is limiting because it is not structured as panel data. We therefore bound the opt-out rate in July 2022 from below by the total number of sites that Sincera only ever recorded as opting out and from above by all those sites that Sincera observed ever opting out until that date.

[26] The Topics API call source consists of a URL that often passes additional parameters to that domain such as the publisher ID and consent-related data: e.g., `https://gum.criteo.com/syncframe?origin=publishertag&topUrl=www.domain.com&us_privacy=1YN-`. However, these parameters are not required by the Topics API, so vendors choose which data to include and how to format this for their own purposes. In July 2023 and across all sites scanned by Sincera, Criteo in particular includes a lot of data in its call source APIs, often including lengthy consent-related strings: a median of 542 characters and a mean of 633.3 characters in the URL after its domain name. In comparison, the other vendors include a median of 36 characters and mean of 42.0 characters after their domain name in the call source URLs.

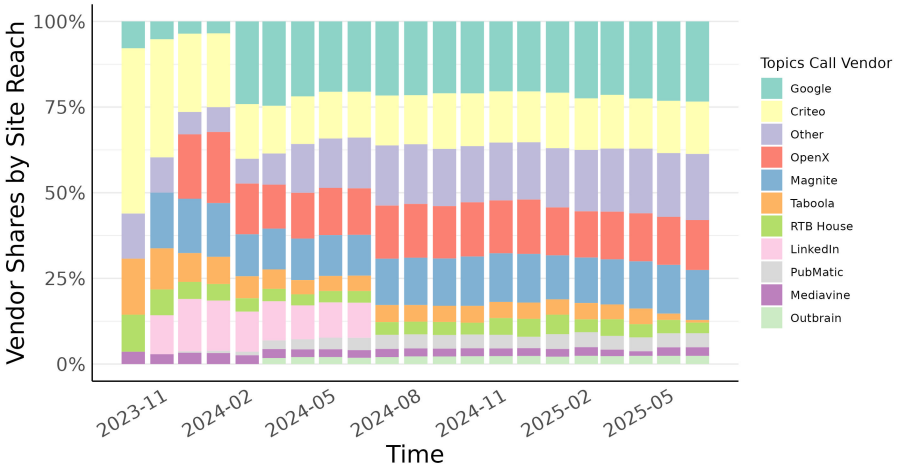
PA API.

Figure 8a shows monthly market shares for the top ten vendors. Early adopters span multiple verticals: demand-side platforms (Criteo, RTB House), publisher ad management (Mediavine), native advertising platforms (Taboola), and full-service providers (Google). Subsequent entrants include the supply-side platform Magnite and social network LinkedIn (November 2023), supply-side platform PubMatic and ad exchange OpenX (December 2023), and native ad platform Outbrain (March 2024). The only exit among the top ten is LinkedIn in June 2024, whose substantial share (10.2%) may have contributed to the July 2024 decline in Topics calls (Figure 7a). Taboola's activity also tapered toward the end of the sample. Google's reach expanded sharply in February 2024, possibly reflecting the same detection limitations noted for PA API (see Section 3.1).

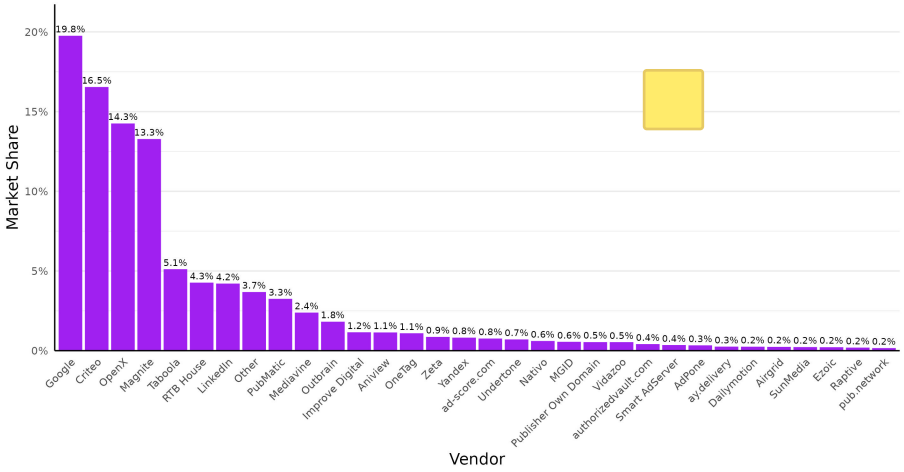
Figure 8b aggregates market shares over the full sample period (October 2023 to June 2025). Google leads with 19.8% share, followed by Criteo (16.5%), OpenX (14.3%), and Magnite (13.3%). In total, Figure 8b lists 31 vendors with shares above 0.1%, while the remaining 13 account for 3.7%. Some websites (e.g., `dailymail.co.uk`, `usatoday.com`) experimenting with calling Topics API from their own domains, grouped here as "Publisher Own Domain" (0.5% of site–vendor pairs). However, since Topics only returns data for domains observed on that category's page, such experiments may return little user segment data if any.[27]

Figure 8c shows market concentration falling below 1,500 points in December 2023 and remaining there through June 2025, indicating an unconcentrated market under DOJ & FTC (2010; 2023) thresholds. As of June 2025, HHI stands at 1,273 points.

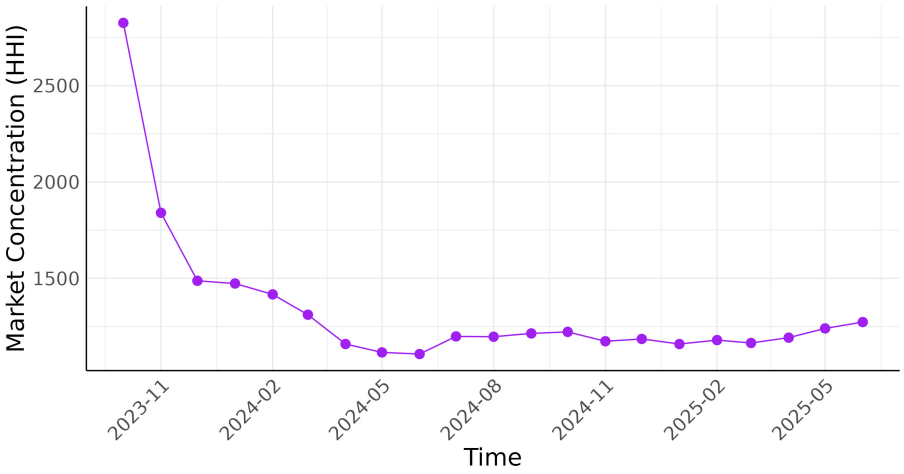
[27] Google's Topics API team indicated that some of these may arise from a Prebid header bidding configurations (see Appendix B) that erroneously tried to call the API from the publisher's own domain. This error has been corrected and the publisher own-domain share has fallen over time.



(a) Evolution of top 10 vendor market shares



(b) Aggregated vendor market shares (full sample)



(c) Evolution of market concentration

Figure 8: Topics API market structure: API calls by vendor

4.2.3 Topics API: Website sample category labels

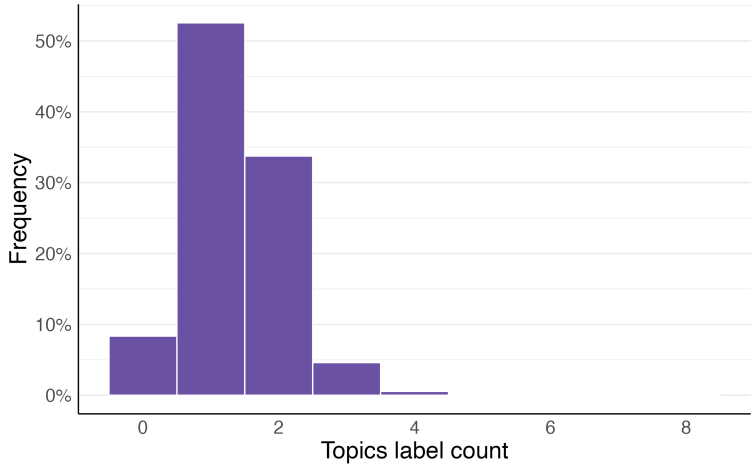
Here, we examine the category labels that Google's Topics API assigns to the 59,186 top site domains in our panel. Note that we do not examine the segment data returned by the Topics API for a group of real users, though this has been studied elsewhere (Dick et al., 2025; Król, 2023; Selman, 2022). Instead, we show how the Topics API classifies websites at the domain level using Google's deterministic classifier (see `chrome://topics-internals`). This exercise sheds light on which categories advertisers are most likely to encounter in practice.

Figure 9 summarizes the distribution of segment labels in our site panel. First, we note that the Topics API's deterministic classifier may assign more than one label to a website or none at all. Figure 9a shows that the median domain receives a single label, the average domain receives 1.37 labels, and 8.35% of sites are not classified at all.[28]

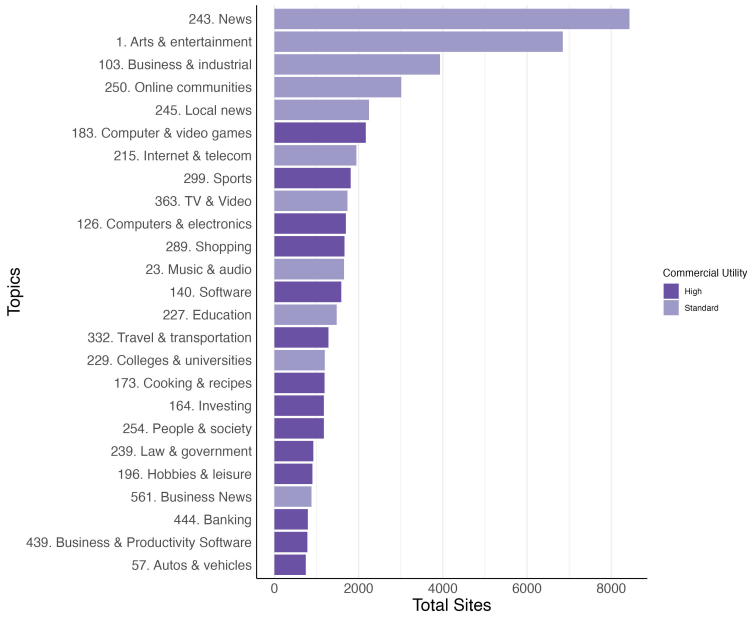
Figure 9b presents the 25 most frequently appearing categories across our domain sample.[29] As mentioned in Section 2.2, Google splits the Topics API taxonomy by whether the category is of high commercial value (e.g., "Investing" and "Shopping" versus "News" and "Online communities") so that Topics API can prioritize commercially valuable segments in its reports to ad tech vendors. Although 14 of the top 25 labels have high commercial utility, none rank among the five most common categories. This highlights the role of Google's pre-sorting mechanism in ensuring that Topics surfaces valuable categories despite their lower raw frequency. Together, these patterns illuminate how Google's Topics API categorizes domains.

[28]The site with the most labels in our sample, `huawei.com`, has eight labels.

[29]Criteo and RTB House instead report the top Topics API (version 1) segments they observe in user data (Król, 2023; Selman, 2022). This list is dominated by generic categories like "News" and "Arts & Entertainment" as the data precede changes to Topics API: the updated taxonomy as well as pre-sorting segments by commercial use.



(a) Number of labels per site domain



(b) Top 25 Topics API labels in site sample

Note: We use Chrome's Topics API classifier (see `chrome://topics-internals`) to assign segment labels to 59,186 domains of the top sites in our panel.

Figure 9: Topics API labels assigned to domains in our top site panel

5 Descriptive results: Adopting website characteristics

We examine the types of websites that adopt Sandbox technologies using a simple regression framework.

Our dependent variables indicate whether we ever detect a site using a given technology during the 2023–2025 panel. For PA API, adoption includes adding a user to an interest group or running an auction. For Topics API, we consider measures of both adoption (making API calls) and refusal (sending an opt-out signal). We focus on our "ever detect" measures for two reasons: (1) the single cross-section matches our cross-sectional covariates, and (2) Sincera's tracking provides only a lower bound on use (see Section 3.1), so the measure mitigates bias from incomplete monthly observations.

We estimate the following linear probability model for the adoption indicator y_i and domains indexed by i:

$$y_i = \alpha + \beta_1 log(site\,rank_i) + \beta_2 log(ad\,units_i + 1) + \beta_3 US_i + \beta_4 EU_i + \varepsilon_i$$

Here, site rank (in logs) comes from the Tranco Alexa-Majestic list (Appendix A). Ad units (log(x + 1)) refer to the average number of ad slots per page that Sincera observed across the pages it scanned for a given domain. US and EU are indicators for whether the *majority* of a site's traffic originates from each region, based on third-party traffic data and top-level domains (e.g., .fr for France). Variable construction details appear in Appendix A.3.

Results in Table 1 indicate that Sandbox adoption is higher among higher-ranked sites, though Topics opt-out rates also increase with rank. Sites with more ad slots adopt Sandbox more frequently and are less likely to opt out of Topics. US sites exhibit greater adoption across all measures, with the Topics opt-out coefficient being statistically insignificant. EU sites show a mixed pattern: higher adoption of PA Auctions, lower PA Interest Groups and Topics use, and higher Topics opt-out rates. Unless noted otherwise,

	Ever detected			
	PA API		Topics API	
	Interest Group	Auction	Calls	Opt Out
log(site rank)	-0.035***	-0.026***	-0.026***	-0.007***
	(0.002)	(0.001)	(0.002)	(0.001)
log(ad units + 1)	0.093***	0.300***	0.260***	-0.002**
	(0.002)	(0.001)	(0.001)	(0.001)
US site	0.155***	0.038***	0.041***	-0.001
	(0.004)	(0.002)	(0.004)	(0.002)
EU site	-0.045***	0.061***	-0.016***	0.011***
	(0.005)	(0.004)	(0.005)	(0.003)
Intercept	0.591***	0.249***	0.594***	0.116***
	(0.020)	(0.011)	(0.017)	(0.010)
Num. Obs.	59,060	59,060	59,060	59,060
R^2	0.081	0.589	0.328	0.002

Notes: Outcome variables are indicators for whether websites ever adopted a Privacy Sandbox technology during the full observation period (2023-2025). See Appendix A.3 for additional details on covariate construction. Heteroskedastic standard errors in parentheses. Marker p-value thresholds: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1: Predicting site adoption of Privacy Sandbox: linear probability model results

all estimates are highly statistically significant at $p < 0.01$. Robustness checks using a logit specification yield qualitatively similar results.[30]

6 Discussion

Our descriptive findings complement other findings in the literature while raising new questions. We discuss our findings and our public dashboard below.

6.1 Sandbox adoption & market structure

Our results show meaningful adoption of Sandbox technologies across websites and dozens of vendors across the ad industry, despite multiple challenges and mixed stakeholder reactions. Consumer surveys suggest muted responses to PEATs (Jerath et al., 2025). Privacy advocates criticized the Sandbox as insufficient,[31] while privacy regulators neither

[30] Results available upon request.

[31] See `https://www.eff.org/deeplinks/2019/08/dont-play-googles-privacy-sandbox-1` and `https://noyb.eu/en/google-sandbox-online-tracking-instead-privacy`.

endorsed it nor relaxed consent requirements. Some vendors doubted Google's commitment to cookie deprecation after repeated delays. Others, such as The Trade Desk, openly opposed the initiative. Adoption required significant engineering investment,[32] leading some firms to defer participation until incentives were stronger. Reflecting these challenges, Google subsidized vendors to support uptake and participation in the CMA's experiment.

Adoption patterns differ across technologies and geographies. PA API is technically demanding, yet more than a dozen firms deployed it at scale. Topics API, by contrast, reached broader adoption thanks to its simpler integration, even though performance tests suggested limited value to advertisers (CMA 2025). We also find evidence of retreat—particularly for Topics—after Google stepped back from full cookie deprecation. Nonetheless, Sandbox use persisted even after Google abandoned deprecation, with little evidence of vendor exit at least as of June 2025. US sites adopted Sandbox more readily despite the absence of federal privacy regulation, while EU sites showed uneven adoption. This divergence is puzzling given Europe's stronger privacy environment, though again regulators' limited support for PEATs as compliance tools may partly explain it.

Our findings complement research on Sandbox performance in the CMA (2023; 2025) experiment. Our findings support the hypothesis that limited adoption constrained performance: advertisers faced scarce PA API inventory and publishers earned little revenue from the few buyers using it (Gu et al., 2025; Kobayashi et al., 2024). Together, this evidence suggests that Sandbox performance has significant room to improve with greater industry adoption. Google's implementation of PA API also increased ad-loading times, causing some impressions to fail to render and reducing publisher revenue (Gu et al., 2025). As we have shown that Google dominates as a PA API seller, this implies these latency effects have wide impact.

Market structure differs across Sandbox technologies. Topics API is unconcentrated.

[32] See `https://www.adexchanger.com/programmatic/most-chrome-privacy-sandbox-testers-see-a-silver-lining-in-their-investment`.

PA API interest groups are highly concentrated. PA API auctions—at the top-level seller—approach monopoly levels. Although vendor entry initially expanded participation, momentum slowed once Google abandoned full cookie deprecation in 2024. In today's mixed regime, with Sandbox and cookies coexisting, competition concerns may be less acute, and the CMA has proposed closing its investigation.

Google nevertheless holds the largest share across all three markets we track. This dominance underscores tensions between privacy, competition, and platform power (Dubé et al., 2025; Johnson, 2024). Critics view Google's role as self-serving, noting that Google Ad Manager only supported Google as a top-level seller in PA API, and that Google allegedly retained an informational advantage as a top-level seller (Google, 2023b, 2024a). At the same time, Google's dominance reflects in part its broad presence across the ad tech ecosystem (CMA 2020). Furthermore, Google's significant investment can be seen as both a necessary signal of commitment and as underwriting a public good: advancing privacy-enhancing ad technologies that other firms can use and build upon.

6.2 Public dashboard

To support transparency and accessibility, we developed and maintained a public dashboard tracking website adoption of Privacy Sandbox technologies on a dedicated webpage: `sincera.io/privacy-sandbox`. Updated monthly, the dashboard reported adoption rates and vendor market shares, drawing on several terabytes of raw data. This effort continued until Sincera shuttered its tracking service in June 2025 after its acquisition by The Trade Desk though the dashboard remained online until August 2025 (see Figure 10).

This paper not only develops new data collection methods to study emerging privacy-centric approaches to the future of marketing, but also embraces a current research paradigm in which interactive tools are central to the work itself.[33] Interactive tools like an online

[33]See recent calls for research on new methods for the future of marketing (`https://www.ama.org/2024/12/09/call-for-papers-journal-of-marketing-research-new-methods-for-the-future-of-marketing/`) and research-driven apps (`https://www.ama.org/journal-of-marketing-call-for-papers-mar`

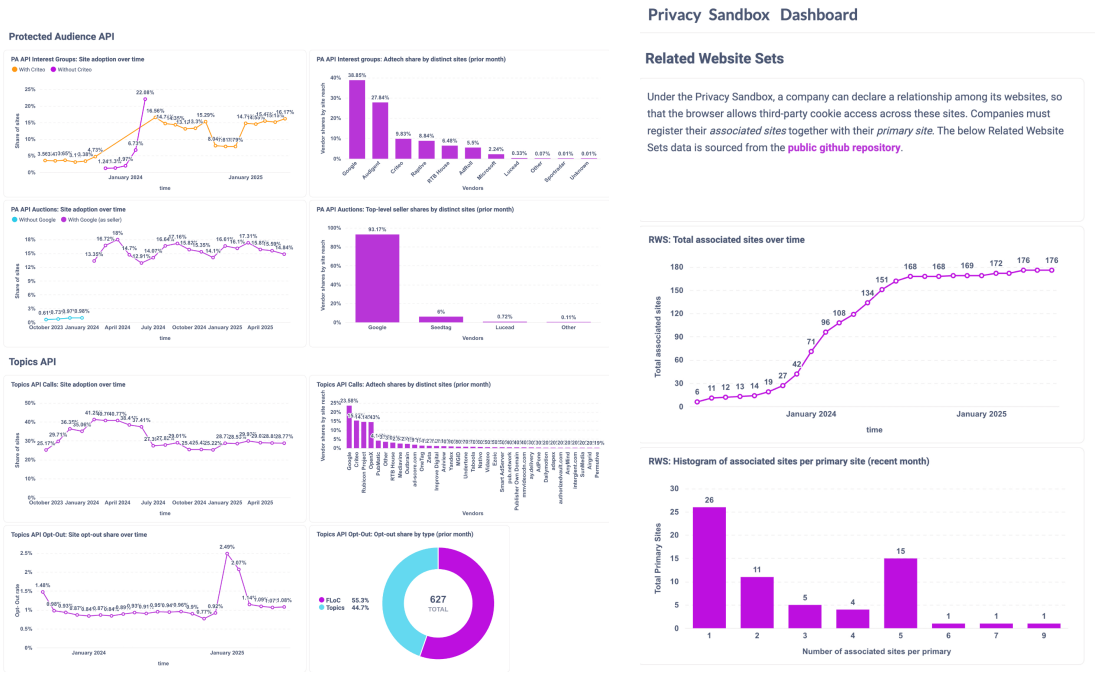


Figure 10: Screenshots of Sandbox website use dashboard

Notes: Dashboard was hosted at `https://app.sincera.io/privacysandbox` until August 2025.

dashboard can help make research accessible for a wider audience, including practitioners and regulators. Prior marketing research has used interactive tools to improve transparency around model assumptions (Shapiro et al., 2021), translate knowledge into practical tools (Warren et al., 2021), and provide flexible visualizations (Matthe et al., 2023; Ringel, 2023).

The Sandbox dashboard was designed for a broad audience—academics, practitioners, journalists, and regulators—helping them navigate the technical complexity of PEATs. This complexity alone may engender skepticism from consumers and other stakeholders.[34] By providing accessible and timely insights, the dashboard informed the public debate and supported the CMA's ongoing evaluation of Privacy Sandbox. Its findings and updates attracted coverage from journalists and engagement from industry stakeholders.[35]

keting-impact-with-research-driven-apps).

[34]See e.g., `https://mobiledevmemo.com/the-privacy-sandbox-paradox/`.

[35]Media coverage includes `https://www.adexchanger.com/data-privacy-roundup/the-kids-arent-playing-in-the-privacy-sandbox/` and `https://www.adexchanger.com/data-privacy-roundup/n`

7 Conclusion

This paper presents a novel empirical study of the online advertising industry's adoption of privacy-enhancing technologies, focusing on Google's Protected Audience and Topics APIs within its Privacy Sandbox initiative. We develop and implement a method for detecting these technologies on a website and intercepting their associated data, enabling the first large-scale panel measurement of their use in the wild. Our results show that Topics API stands out for its broader adoption among both websites and ad tech vendors—perhaps reflecting its greater ease of use—while Protected Audience API adoption remains more limited. We identify several ad tech vendors that participate in Privacy Sandbox: vendor markets for Topics API show low concentration, while those for Protected Audience API are highly concentrated.

Our research has important limitations. In particular, we provide conservative estimates of industry adoption. For technical reasons, we do not observe two major ad tech vendors (Criteo and Google) in our PA API data for certain periods, and we capture Topics API calls via JavaScript but not via HTTP headers.

Future research can investigate how Sandbox use and market structure evolve after June 2025 in response to Google's shifting roadmap as well as changes in privacy regulation, guidance, and enforcement. Future research could also extend our approach to measure adoption of other Sandbox technologies—both advertising-related and non-advertising-related—as well as similar privacy-enhancing technologies introduced by other platforms and intermediaries.

---

o-shade-but-does-anyone-still-care-about-the-privacy-sandbox/. References by firms include https://www.linkedin.com/pulse/half-baked-reality-testing-googles-privacy-sandbox-nextroll-seqkc/, and https://www.rtbhouse.com/blog/cookieless-newsroom-another-delay-in-third-party-cookies-phase-out-cma-disconcerted.

References

Alcobendas, M., Kobayashi, S., Shi, K., & Shum, M. (2023). The impact of privacy measures on online advertising markets. Available at SSRN 3782889.

Alvim, M. S., Fernandes, N., McIver, A., & Nunes, G. H. (2023). A quantitative information flow analysis of the Topics API. In *Proceedings of the 22nd Workshop on Privacy in the Electronic Society*, WPES '23 (pp. 123–127). New York, NY, USA: Association for Computing Machinery.

Aridor, G., Che, Y.-K., Hollenbeck, B., Kaiser, M., & McCarthy, D. (2024). Evaluating the impact of privacy regulation on e-commerce firms: Evidence from Apple's App Tracking Transparency. Available at SSRN.

Bashir, M. A., Arshad, S., Kirda, E., Robertson, W., & Wilson, C. (2019). A longitudinal analysis of the ads.txt standard. In *Proceedings of the Internet Measurement Conference*, IMC '19 (pp. 294–307). New York, NY, USA: Association for Computing Machinery.

Berke, A. & Calacci, D. (2022). Privacy limitations of interest-based advertising on the web: A post-mortem empirical analysis of Google's FLoC. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security* (pp. 337–349).

Beugin, Y. & McDaniel, P. (2024). A public and reproducible assessment of the Topics API on real data. In *2024 IEEE Security and Privacy Workshops (SPW)* (pp. 1–8). Los Alamitos, CA, USA: IEEE Computer Society.

Bleier, A., Goldfarb, A., & Tucker, C. (2020). Consumer privacy and the future of data-based innovation and marketing. *International Journal of Research in Marketing*, 37(3), 466–480.

Calderonio, G., Ali, M. M., & Polakis, J. (2024). Fledging will continue until privacy improves: Empirical analysis of google's Privacy-Preserving targeted advertising. In *33rd USENIX Security Symposium (USENIX Security 24)* (pp. 4121–4138). Philadelphia, PA: USENIX Association.

Carey, C., Dick, T., Epasto, A., Javanmard, A., Karlin, J., Kumar, S., Muñoz Medina, A., Mirrokni, V., Nunes, G. H., Vassilvitskii, S., et al. (2023). Measuring re-identification risk. *Proceedings of the ACM on Management of Data*, 1(2), 1–26.

Choi, H., Mela, C., Balseiro, S., & Leary, A. (2020). Online display advertising markets: A literature review and future directions. *Information Systems Research*, 31(2), 556–575.

Choi, W. J. & Jerath, K. (2022). Privacy and consumer empowerment in online advertising. *Foundations and Trends® in Marketing*, 15(3), 153–212.

Competition and Markets Authority (2020). *Online platforms and digital advertising market study*. Technical report, Competition and Markets Authority.

Competition and Markets Authority (2023). *Quantitative testing of Google's Privacy Sandbox technologies – CMA guidance to third parties on testing*. Technical report, Competition and Markets Authority.

Competition and Markets Authority (2025). *Testing of Google's Privacy Sandbox proposals: CMA summary*. Technical report, Competition and Markets Authority.

Cooper, D. A., Yalcin, T., Nistor, C., Macrini, M., & Pehlivan, E. (2023). Privacy considerations for online advertising: A stakeholder's perspective to programmatic advertising. *Journal of Consumer Marketing*, 40(2), 235–247.

Dawson, M., Ghazi, B., Kamath, P., Kumar, K., Kumar, R., Luan, B., Manurangsi, P., Mundru, N., Nair, H., Sealfon, A., et al. (2023). Optimizing hierarchical queries for the Attribution Reporting API. In *AdKDD at ACM SIGKDD Conference on Knowledge Discovery and Data Mining*.

Deisenroth, D., Manjeer, U., Sohail, Z., Tadelis, S., & Wernerfelt, N. (2024). *Digital Advertising and Market Structure: Implications for Privacy Regulation*. Technical report, National Bureau of Economic Research.

Dick, T., Epasto, A., Javanmard, A., Karlin, J., Medina, A. M., Mirrokni, V., Vassilvitskii, S., & Zhong, P. (2025). Differentially private synthetic data release for Topics API outputs. *arXiv preprint arXiv:2506.23855*.

Domagała, B. (2024). Is it possible to use first-party data in a cookieless environment? RTB House Blog.

Dubé, J.-P., Lynch, J. G., Bergemann, D., Demirer, M., Goldfarb, A., Johnson, G., Lambrecht, A., Lin, T., Tuchman, A., & Tucker, C. (2025). Frontiers: The intended and unintended consequences of privacy regulation for consumer marketing. *Marketing Science*, forthcoming.

Englehardt, S. & Narayanan, A. (2016). Online tracking: A 1-million-site measurement and analysis. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1388–1401).: ACM.

Epasto, A., Medina, A. M., Ilvento, C., & Karlin, J. (2022). Measures of cross-site re-identification risk: An analysis of the Topics API proposal.

Geng, T., Dawson, M., & Nair, H. (2023). *Effectively Combining the Event and Aggregate Summary Reports from the Privacy Sandbox Aribution Reporting API for Improving Ad-Measurement Fidelity*. Technical report, Google Ads.

Geradin, D., Katsifis, D., & Karanikioti, T. (2021). Google as a de facto privacy regulator: analysing the Privacy Sandbox from an antitrust perspective. *European Competition Journal*, 17(3), 617–681.

Goldfarb, A. (2013). What is different about online advertising? *Review of Industrial Organization*, (pp. 1–15).

Goldfarb, A. & Que, V. F. (2023). The economics of digital privacy. *Annual Review of Economics*, 15(1), 267–286.

Google (2022). The Privacy Sandbox: technology for a more private web. https://privacysandbox.com.

Google (2023a). *Google Ads Test: Testing Interest Based Audience Solutions with Privacy Preserving Signals*. Technical report, Google.

Google (2023b). *Privacy Sandbox Progress Report: Q3 Reporting Period - July to September 2023. Prepared for the CMA, 24 October 2023*. Technical report, Google.

Google (2024a). *Privacy Sandbox Progress Report: Q1 Reporting Period - January to March 2024. Prepared for the CMA, 25 April 2024*. Technical report, Google.

Google (2024b). *Testing privacy preserving signals in the absence of third-party cookies on Google's display ads platforms*. Technical report, Google.

Gordon, B. R., Jerath, K., Katona, Z., Narayanan, S., Shin, J., & Wilbur, K. C. (2021). Inefficiencies in digital advertising markets. *Journal of Marketing*, 85(1), 7–25.

Gu, Z., Johnson, G., & Kobayashi, S. (2025). Can privacy technologies replace cookies? Ad revenue in a field experiment. Available on SSRN.

Hard, A., Kiddon, C. M., Ramage, D., Beaufays, F., Eichner, H., Rao, K., Mathews, R., & Augenstein, S. (2018). Federated learning for mobile keyboard prediction.

Höring, F. (2022). An update on FLEDGE Chrome testing. Criteo R&D Blog.

Index Exchange product team (2024). Insights from Privacy Sandbox testing. Index Exchange Blog.

Information Commissioner's Office (2022). Privacy-enhancing technologies (PETs). In *Draft anonymisation, pseudonymisation and privacy enhancing technologies guidance* chapter 5. Information Commissioner's Office.

Jerath, K., Miller, K., & Sokol, D. D. (2025). Towards developing an understanding of consumers' perceived privacy violations in online advertising. Available at SSRN.

Johnson, G. A. (2024). Economic research on privacy regulation: Lessons from the GDPR and beyond. In A. Goldfarb & C. E. Tucker (Eds.), *Economics of Privacy* chapter 4. University of Chicago Press.

Johnson, G. A., Runge, J., & Seufert, E. (2022). Privacy-centric digital advertising: Implications for research. *Customer Needs and Solutions*, 9(1), 49–54.

Johnson, G. A., Shriver, S., & Goldberg, S. (2023). Privacy & market concentration: Intended & unintended consequences of the GDPR. *Management Science*, 69(10), 5695–6415.

Johnson, G. A., Shriver, S. K., & Du, S. (2020). Consumer privacy choice in online advertising: Who opts out and at what cost to industry? *Marketing Science*, 39(1), 33–51.

Karaj, A., Macbeth, S., Berson, R., & Pujol, J. M. (2018). Whotracks.me: Monitoring the online tracking landscape at scale. *CoRR*, abs/1804.08959.

Kenny, C. T., McCartan, C., Kuriwaki, S., Simko, T., & Imai, K. (2024). Evaluating bias and noise induced by the U.S. Census Bureau's privacy protection methods. *Science Advances*, 10(18), eadl2524.

Kobayashi, S. J., Johnson, G. A., & Gu, Z. (2024). Privacy-enhanced versus traditional retargeting: Ad effectiveness in an industry-wide field experiment. Available at SSRN.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from A/B testing. *Management Science*, 68(9), 6434–6453.

Król, M. (2023). Topics API today (december 2023): Assessing its scale and usability. RTB House Blog.

Lachner, F., Chen, Y., & Olsauskas-Warren, T. (2023). User attitudes towards controls for ad interests estimated on-device by the browser. In *Symposium on Usable Security and Privacy (USEC)*.

Le Pochat, V., Van Goethem, T., Tajalizadehkhoob, S., Korczyński, M., & Joosen, W. (2019). Tranco: A research-oriented top sites ranking hardened against manipulation. In *Proceedings of the 26th Annual Network and Distributed System Security Symposium*, NDSS 2019.

Lefrere, V., Warberg, L., Cheyre, C., Marotta, V., & Acquisti, A. (2022). Does privacy regulation harm content providers? A longitudinal analysis of the impact of the GDPR. Available at SSRN.

Lerner, A., Simpson, A. K., Kohno, T., & Roesner, F. (2016). Internet Jones and the raiders of the lost trackers: An archaeological study of web tracking from 1996 to 2016. In *25th USENIX Security Symposium (USENIX Security 16)*.

Libert, T. (2015). Exposing the hidden web: An analysis of third-party HTTP requests on one million websites. *International Journal of Communication*, 9, 18.

Martin, K., Nissenbaum, H., & Shmatikov, V. (2025). No cookies for you!: Evaluating the promises of big tech's 'privacy-enhancing' techniques. *Georgetown Law Technology Review*, 9(1).

Matthe, M., Ringel, D. M., & Skiera, B. (2023). Mapping market structure evolution. *Marketing Science*, 42(3), 589–613.

McGuigan, L., Sivan-Sevilla, I., Parham, P., & Shvartzshnaider, Y. (2024). Private attributes: The meanings and mechanisms of "privacy-preserving" adtech. *New Media & Society*, forthcoming, 14614448231213267.

Medina, A. M., Kleber, M., Karlin, J., & Vale, M. (2021). Measuring sensitivity of cohorts generated by the FLoC API. Google White Paper.

Messing, S., DeGregorio, C., Hillenbrand, B., King, G., Mahanti, S., Mukerjee, Z., Nayak, C., Persily, N., State, B., & Wilkins, A. (2020). Facebook Privacy-Protected Full URLs Data Set.

Miller, K. M., Lukic, K., & Skiera, B. (2025). The impact of the General Data Protection Regulation (GDPR) on online tracking. *International Journal of Research in Marketing*, forthcoming.

Miller, K. M. & Skiera, B. (2023). Economic consequences of online tracking restrictions: Evidence from cookies. *International Journal of Research in Marketing*.

Nikiforakis, N., Kapravelos, A., Joosen, W., Kruegel, C., Piessens, F., & Vigna, G. (2013). Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. In *Security and privacy (SP), 2013 IEEE symposium on* (pp. 541–555).: IEEE.

Nottingham, M. (2021). Playing fair in the Privacy Sandbox: Competition, privacy and interoperability standards. Available at SSRN.

Olejnik, L. (2023). *Handbook on the Politics and Governance of Big Data and Artificial Intelligence*, chapter On the governance of privacy-preserving systems for the web: should Privacy Sandbox be governed? Elgar Handbooks in Political Science. Edward Elgar Publishing.

Olejnik, L. (2024). Reconciling Privacy Sandbox initiatives with EU data protection laws. LL.M. Thesis.

Parsons, T. (2024). Privacy Sandbox testing results show shortfalls to meet CMA requirements. Criteo Blog.

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory spillovers and data governance: Evidence from the GDPR. *Marketing Science*, 41(4), 318–340.

Philipse, M. (2024). Post-third-party cookies: Analyzing Google's Protected Audience API. Master's thesis, Radboud University.

Quéré, R. (2022). On the topic of Topics. Xandr Tech Blog.

Ringel, D. M. (2023). Multimarket membership mapping. *Journal of Marketing Research*, 60(2), 237–262.

Rumiński, M., Iwańczak, P., & Włodarczyk, Ł. (2022). Findings from the early FLEDGE experiments. *Available at SSRN 4219796*.

Sanchez-Rola, I., Dell'Amico, M., Kotzias, P., Balzarotti, D., Bilge, L., Verviker, P.-A., & Santos, I. (2019). Can I opt out yet? GDPR and the global illusion of cookie control. In *ACM ASIACCS 2019*.

Selman, E. (2022). Topics API: Criteo's first look at Google's interest-based advertising solution. Criteo R&D Blog.

Shapiro, B. T., Hitsch, G. J., & Tuchman, A. E. (2021). TV advertising effectiveness and profitability: Generalizable results from 288 brands. *Econometrica*, 89(4), 1855–1879.

Shchetkina, A. (2025). Blind targeting: Personalization under third-party privacy constraints. *Available at SSRN 5342212*.

Sousa, T. (2024). Competition and targeting in auctions for online ads. Working paper.

Stroube, B. & Dushnitsky, G. (2023). Built with BuiltWith: A technical note. *Available at SSRN*.

Technology Analysis Division of the Office of the Privacy Commissioner of Canada (2017). *Privacy Enhancing Technologies – A Review of Tools and Techniques*. Technical report, Office of the Privacy Commissioner of Canada.

Tholoniat, P., Caulfield, A., Cavicchioli, G., Chen, M., Goutzoulias, N., Case, B., Cidon, A., Geambasu, R., Lécuyer, M., & Thomson, M. (2025). Big Bird: Privacy budget management for W3C's Privacy-Preserving Attribution API.

Tholoniat, P., Kostopoulou, K., McNeely, P., Sodhi, P. S., Varanasi, A., Case, B., Cidon, A., Geambasu, R., & Lécuyer, M. (2024). Cookie Monster: Efficient on-device budgeting for differentially-private ad-measurement systems. In *Proceedings of the ACM SIGOPS 30th Symposium on Operating Systems Principles*, SOSP '24 (pp. 693–708). New York, NY, USA: Association for Computing Machinery.

Thomson, M. (2023). A privacy analysis of Google's Topics proposal.

Tucker, C. E. (2012). The economics of advertising and privacy. *International Journal of Industrial Organization*, 30(3), 326 – 329. Selected Papers, European Association for Research in Industrial Economics 38th Annual Conference, Stockholm, Sweden, September 1-3, 2011.

US Department of Justice and the Federal Trade Commission (2010). Horizontal merger guidelines. http://www.ftc.gov/os/2010/08/100819hmg.pdf.

US Department of Justice and the Federal Trade Commission (2023). Merger guidelines.

Vekaria, Y., Standaert, B., Ostapenko, M., Amjad, A. H., Dimova, Y., Munir, S., Böttger, C., & Iqbal, U. (2024). Privacy. In B. Pollard (Ed.), *The 2024 Web Almanac* chapter 10. HTTP Archive.

Warnat-Herresthal, S., Schultze, H., et al. (2021). Swarm learning for decentralized and confidential clinical machine learning. *Nature*, 594(7862), 265–270.

Warren, N. L., Farmer, M., Gu, T., & Warren, C. (2021). Marketing ideas: How to write research articles that readers understand and cite. *Journal of Marketing*, 85(5), 42–57.

Wernerfelt, N., Tuchman, A., Shapiro, B. T., & Moakler, R. (2025). Estimating the value of offsite tracking data to advertisers: Evidence from Meta. *Marketing Science*, 44(2), 268–286.

Appendix

A Website panel construction

We want to track a consistent and explainable list of higher-quality, commercial websites over time. Constructing our target website list was an iterative process as some of Sincera's Privacy Sandbox data collection predates these efforts. Once we finalized our target list, Sincera agreed to add the remaining sites to its roster of regularly scanned sites, so our sample of scanned sites grows over time.

We begin with a Tranco list of top 100,000 sites as of October 2022, which is designed and constructed using the methodology from Le Pochat et al. (2019). Our chosen Tranco list aggregates independent rankings by Alexa and Majestic: the list is available at `https://tranco-list.eu/list/4K8VX/1000000`.[36] We chose the 100,000 top sites threshold in order to balance our site coverage and the overlap with Sincera's regularly scanned sites as of mid-2023.

Section A.1 describes our *ex ante* website domain filtering criteria to focus on higher quality, commercial domains. Section A.2 describes our *ex post* list of scannable websites that removes unreachable, redirecting, or web-vendor domains. Section A.2 concludes by describing how our data collector's scanning coverage of our target list grows over time.

A.1 Domain filtering criteria

Sincera began as a startup that regularly scans a list of websites to collect data on the site's content as well as the vendors and technologies that the site employs. Sincera scans hundreds of thousands of websites every week and this list has grown over time. Sincera

[36]The primary Tranco list aggregates independent rankings from Umbrella and Farsight in addition to Alexa and Majestic (Le Pochat et al., 2019). In practice, this list places a greater emphasis on network traffic than consumer traffic alone. For instance, its top 10 list include domains that would be unfamiliar to many consumers like `gtld-servers.net` (#2), `akamaiedge.net` (#5), `amazonaws.com` (#10). We therefore prefer to use our (Alexa & Majestic) Tranco list, which also includes more sites that Sincera tracked in mid-2023.

considers several factors in selecting the sites on its scan list, however, its most important criterion is customer demand. Sincera is a commercial service whose customers can request scans of sites for a list that the customer provides. In addition to ourselves, Sincera's customers include ad tech companies, web publishers, and even journalists. As a startup, Sincera applied some initial criteria to filter domains opting to prioritize commercially-relevant websites and cost efficiency over full coverage. Nevertheless, customer demands override Sincera's domain filtering criteria.

We filter our Tranco list of domains using three broad criteria that closely align with Sincera's initial filtering criteria. First, we exclude domains whose top-level domain indicates that the site is located in certain countries (e.g., `.ru`, `.cn`). Sincera's excluded country list is *ad hoc*, but predominantly includes large countries in Asia, Africa, and the Middle East. The full country list is: Angola, Azerbaijan, China, India, Iran, Japan, Kazakhstan, Kenya, Laos, Morocco, North Macedonia, Russia, Saudi Arabia, Singapore, South Korea, Taiwan, Tanzania, Turkey, Ukraine, Uzbekistan, and Vietnam. Relatedly, we exclude country-specific instances of `google.com` (e.g., `google.ro`) since these often redirect to `google.com` when scanned from outside that country.

Second, Sincera excludes certain non-commercial, top-level domains including: `.edu`, `.gov`, `.land` (i.e., real estate), `.mil` (i.e., military), `.org`, and `.win` (i.e., Microsoft Windows).

Third, Sincera seeks to exclude low-quality sites by excluding domains containing certain terms (including the top-level domain). These terms fall into the following broad categories:

- Adult content.[37]
- Spam content including the terms: casino, cash, free, game, mp3, pharma, torrent, and yeezy.
- Marketing content including the terms: ads, advertising, agency, cdn, digital, dns,

[37]The excluded adult-rated content terms include: adult, sugardaddy, viagra, brides, nsfw, fetish, tranny, slut, porn, fuck, sex, and xxx.

group, marketing, media, pub, publisher, seo, solutions, and yield.

- Domains containing "blogspot": this removes sites from Google's popular blogging service `blogger.com`. Advertising on these sites is optional, and is restricted to Google AdSense ads.

- Domains containing "desi": this predominantly avoids content associated with the countries India, Pakistan, or Bangladesh, including websites with the top-level domain "`.desi`".

We acknowledge that some of these terms are broad; for instance, the "desi" criterion also excludes domains containing the term "design." Nonetheless, our filtered domain list balances Sincera's scanning coverage with our goals of an explainable list of quality, commercially-relevant websites. These domain filtering criteria eliminate 24.2% of sites in our Tranco list, leaving an interim list of 75,797 sites.

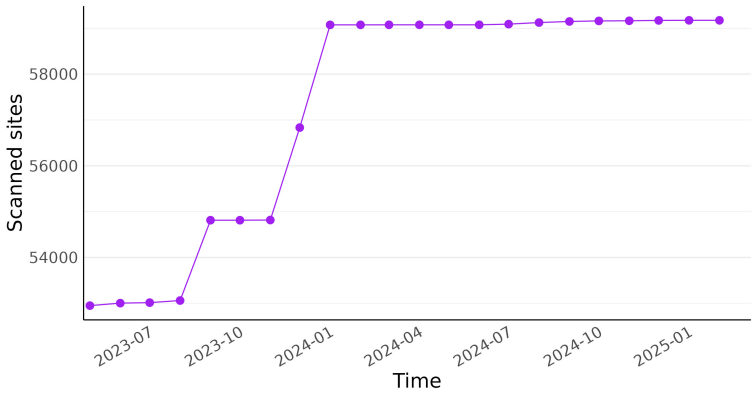
A.2 Scannable websites

From the interim list of 75,797 sites, we remove sites that are unreachable, that generate HTTP errors, that are web vendor domains, or that redirect to another domain.

We built our own site-scanning algorithm to evaluate domains on our list that Sincera did not scan. Surprisingly, our Tranco list includes many domains that are unreachable or redirect to other domains. Our site scanning algorithm evaluates whether a site is reachable, records the HTTP response code returned by the domain, and records any redirect domain.

We classify unscannable websites using several criteria, which we apply cumulatively. First, our site scanner identifies 11,423 websites that are unreachable. We expect that many of these sites are not consumer-facing sites: e.g., web technology vendors employ these domains to provide advertising and content services for websites. Second, another 482 websites are unreachable in that they return HTTP codes that indicate either a client-

Figure A.1: Number of Sincera scanned sites from our target list over time



Note: This plots the number of sites that Sincera scanned as of the end of a given month from the 59,186 websites in our panel.

or server-side error. Third, we eliminate 3,488 sites that redirect to different domains; however, many of these redirect domains are included in our list (e.g., `a.co` redirects to `amazon.com`). Fourth, we eliminate 403 domains that are identified by Sincera as associated with web technology vendors, and that Sincera does not scan. After these steps, we have a target list of 60,001 filtered and potentially scannable websites.

Sincera ultimately scanned 59,186 sites in this list, which form our panel. Some of remaining sites were not scannable: for instance, Sincera identified another 361 sites that it could not scan for various reasons.[38]

Figure A.1 shows that Sincera's site scanning coverage for our 59,186 site panel grew over time. Sincera's initial coverage of our filtered list was high: 52,948 sites (89.5%) in May 2023. The missing sites are disproportionately lower ranked.[39] By January 2024, Sincera was scanning almost all of our panel sites (i.e., 59,175).

[38]Sincera encountered HTTP errors on 155 sites, Captchas on 130 (i.e., blocking automated scanning), failed page loads for 69, general errors for 6, and the scan timed out for 1 site.

[39]Sincera uses the primary Tranco site list (see footnote 36) to guide its scanning decisions. Half of the unscanned sites as of December 2023 are ranked outside of the primary Tranco top 250,000 list, which contributed to Sincera ignoring these sites.

A.3 Site covariate definition

This appendix provides additional detail on the construction of the website-level covariates used in Section 5.

Average Ad Units. This variable measures the average number of ad units per page, as recorded by Sincera across the pages it scans for a given site. When the measure is missing, we impute a value of zero. Among the 24,286 sites in our sample with ads.txt files—industry-standard public declarations of the ad exchanges and networks authorized to sell a publisher's inventory—3,007 sites (12.4%) lack ad unit data. Because ads.txt files are a strong signal that a site monetizes via programmatic display advertising, their absence suggests the site is unlikely to host display ads, making zero a reasonable imputation for such cases. Our main results are qualitatively unchanged when dropping sites that lack ad unit data despite having ads.txt files.[40]

US and EU Indicators. These binary variables indicate whether a site primarily serves users in the United States or in the European Union, respectively. We construct this using a combination of third-party traffic data and inference from country top-level domains. Country assignment applies the following criteria in descending order of priority: (1) SimilarWeb data indicating at least 50% of traffic from a given country; (2) Amazon Web Information Services (AWIS) data indicating at least 50% of traffic from a given country; (3) the site's country-specific top-level domain (TLD).

SimilarWeb data, available from January 2024, is incomplete across both sites and countries. We use a targeted sampling strategy to maximize coverage within a budget limit, focusing on the top 75,000 Tranco-ranked sites after excluding those with country-specific TLDs.[41] We collect data for 13 countries: the United States, the United Kingdom, Israel, and the ten most populous EU member states (which together account for 82% of the EU population). Sites with country-specific TLDs are excluded from the SimilarWeb

[40] Robustness results available upon request.

[41] We are most grateful to Klaus Miller for obtaining this data on our behalf.

queries because traffic to these sites generally originates from the corresponding country. Note that our main results are qualitatively unchanged when restricting to the top 75,000 sites where country assignments are most reliable.

We also use AWIS data, from April 2020, which provides country-level traffic shares for top sites as of Spring 2018. Like SimilarWeb, it is incomplete for our site list, though includes all user countries.

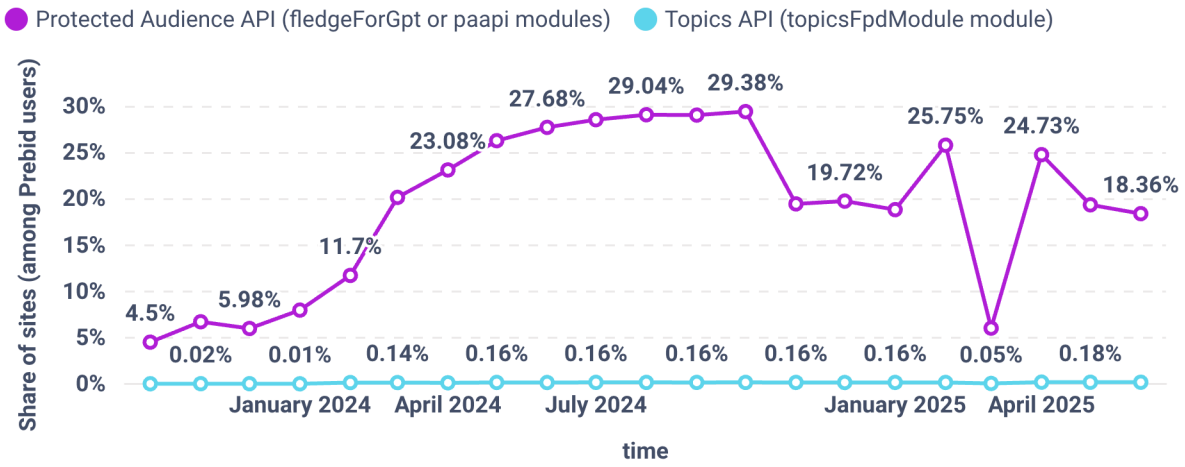
B Alternate Privacy Sandbox adoption measures

We explore alternative measures of Privacy Sandbox adoption.

B.1 Header bidding (Prebid) modules

First, we examine website adoption of Privacy Sandbox technologies via header bidding. Many websites sell advertising through a mechanism called header bidding, which allows websites to obtain multiple bids from ad buyers in parallel. To do so, websites use a free and open-source standard called "Prebid." Prebid allows websites to integrate PA API auctions by adding a module called `fledgeForGpt` or its successor `paapi` then enabling Protected Audience API in their Prebid JavaScript configuration. Similarly, Prebid allows websites to integrate Topics API into header bidding using the module `topicsFpdModule`. Sincera scans websites and records their Prebid configuration, which allows us to identify sites that have enabled the relevant modules. Note that websites do not need Prebid to access the Sandbox technologies, but this provides another measurement of publisher's deliberate engagement with the Sandbox.

Figure A.2 shows the evolution of Privacy Sandbox module use in Prebid settings. Since this measurement is specific to Prebid, we focus on the subsample of Prebid users in our site sample. Figure A.2 shows steady growth in the PA API-related modules from 4.5% of Prebid users in October 2023 to 28.5% in July 2024. Usage falls to 19.4% in Novem-



Note: This data above consists of the subsample of sites in our panel that use Prebid.

Figure A.2: Privacy Sandbox module use in site's header bidding (Prebid) settings

ber 2024 from 29.4% in the prior month and becomes more erratic thereafter. We do not know the cause of the steep drop in March 2025, though we speculate this is a data quality issue. This evidence complements our site scanning evidence in Figure 4b: the growth period is similar though the decline happens several months after Google retreated from cookie deprecation in July 2024.

The Topics API module adoption is negligible throughout (<0.2%). This contrasts with our site scan evidence (Figure 7a) showing that 25-41% of sites contain a Topics API call over that period. Again, this does not contradict our earlier evidence as sites do not need to use Prebid to use the Sandbox technologies. We speculate that the ease of calling the Topics API directly reduces the demand for the related Prebid module.

B.2 Privacy Sandbox vendor attestations

To use the Privacy Sandbox APIs, vendors must attest that they will not abuse the APIs to track users across sites. Google has a formal registration process whereby vendors make attestations for specific APIs and do so separately for the Chrome browser and the Android mobile operating system. Vendors also log these attestations in a standardized format on their own websites (in the `/.well-known` directory). We regularly crawl the



Note: Data includes vendors that registered with Google as of February 2025, as listed at `https://github.com/privacysandbox/attestation`.

Figure A.3: Vendor Privacy Sandbox attestations

Tranco top 1 million websites to log these attestations. However, Google released the universe of vendors that registered with it on GitHub in June 2024 and updated the list in February 2025.[42] Below, we present data from the universe of registered vendors from Google's list as of February 2025.

Figure A.3 shows the share of attestations by technology and device for the 326 sites that registered attestations. We see that attestations are much higher for Chrome than Android. In Chrome, the highest attestation rates are in the targeting APIs: 76.4% for Topics API and 72.7% for PA API. Though Topics API is higher, the difference is not as great as in our scanning data. This is followed by the measurement API (Attribution Reporting) at 69.9%. The Shared Storage API (62.9%) and Private Aggregation API (59.2%) are lowest. Though these two APIs are seldom discussed in media coverage of Privacy Sandbox, they have multiple uses including measurement and ad frequency management.

[42] See `https://github.com/privacysandbox/attestation`.

C Adoption of other Privacy Sandbox technologies

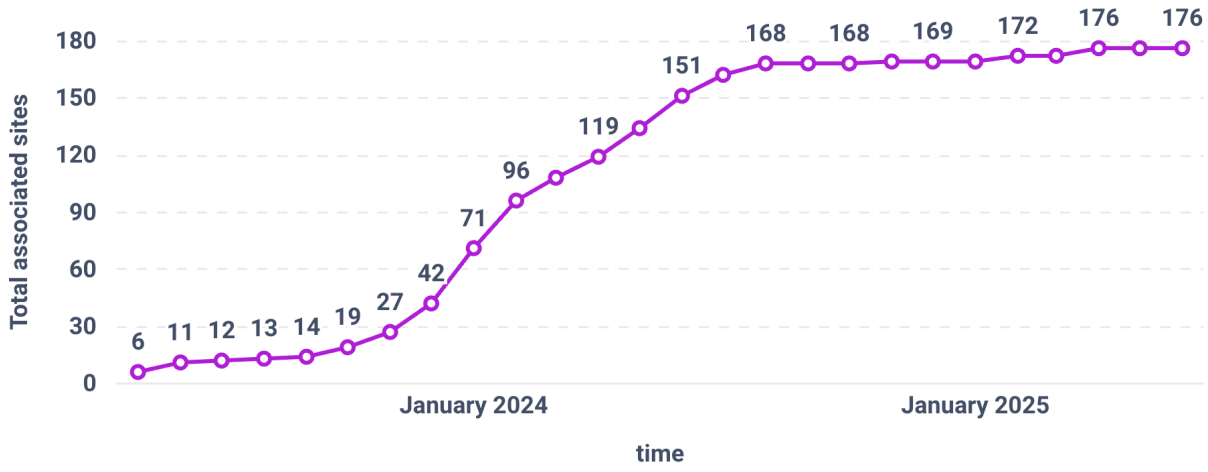
C.1 Related Website Sets

Google's Privacy Sandbox provides a mechanism called Related Website Sets (RWS) for third-party cookies to still function within a narrow network of related sites. This could be relevant for several example use cases, particularly websites that are owned by a single company. For instance, Meta owns multiple domains associated with its social media platforms: `facebook.com`, `instagram.com`, and `whatsapp.com`. Also, a website like `google.com` could operate several country-specific domains like `google.ca`, `google.in`, and `google.fr`. Finally, a company could have several domains relating to a single brand, such as Apple's `apple.com`, `icloud.com`, and `applemusic.com`.

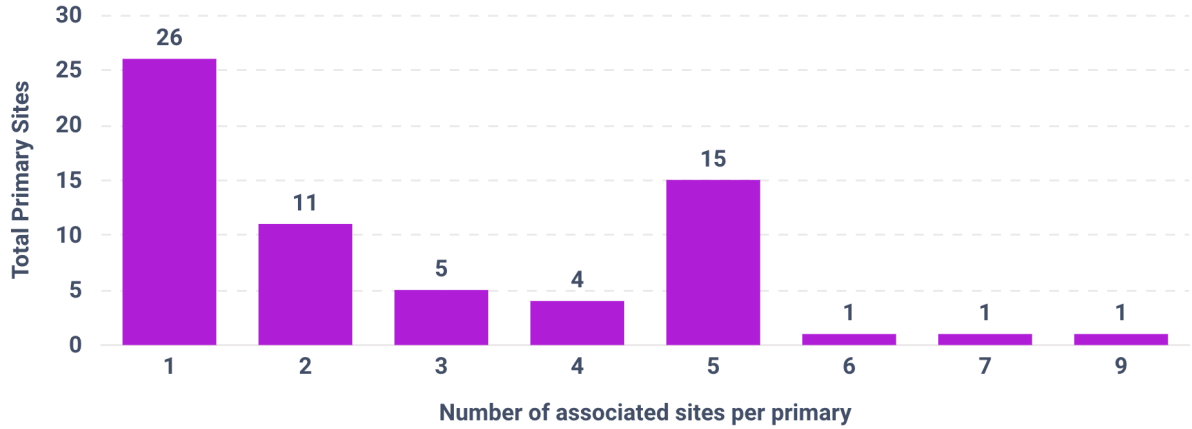
Under the Privacy Sandbox, a company can declare a relationship among its websites, so that the browser allows third-party cookie access across these sites. To do so, companies must register their associated sites along with their primary site. The RWS list is stored on a public GitHub repository: `https://github.com/GoogleChrome/related-website-sets`. This allows us to document the universe of websites in the RWS list.

Figure A.4a below shows that the number of associated sites in the RWS list grows from 6 sites in the initial May 2023 list to 162 sites in July 2024. Growth then tapered off to 176 sites in the data by June 2025. The RWS list has an average of 2.75 associated sites per primary domain as of June 2025. Figure A.4b presents a histogram of the number of associated sites per primary site as of June 2025. Most of these are below 5 sites, which is the threshold that Google automatically grants in its RWS guidelines.[43] In sum, the RWS data conforms with the broader pattern of Privacy Sandbox growth then stalling after Google retreated from cookie deprecation in July 2024.

[43] See `https://github.com/GoogleChrome/related-website-sets/blob/main/RWS-Submission_Guidelines.md`.



(a) Associated sites over time



(b) Histogram of associated sites per primary site

Figure A.4: Related Web Sites (universe of sites)

D Protected Audience API component auctions

PA API auctions can be structured either as single-seller auctions or as meta-auctions with multiple 'component' sellers. In component auctions, each component seller runs an independent auction and the top-level seller selects among their winners. Because a component auction may involve only one seller—sometimes identical to the top-level seller—some configurations are operationally equivalent to single-seller auctions.

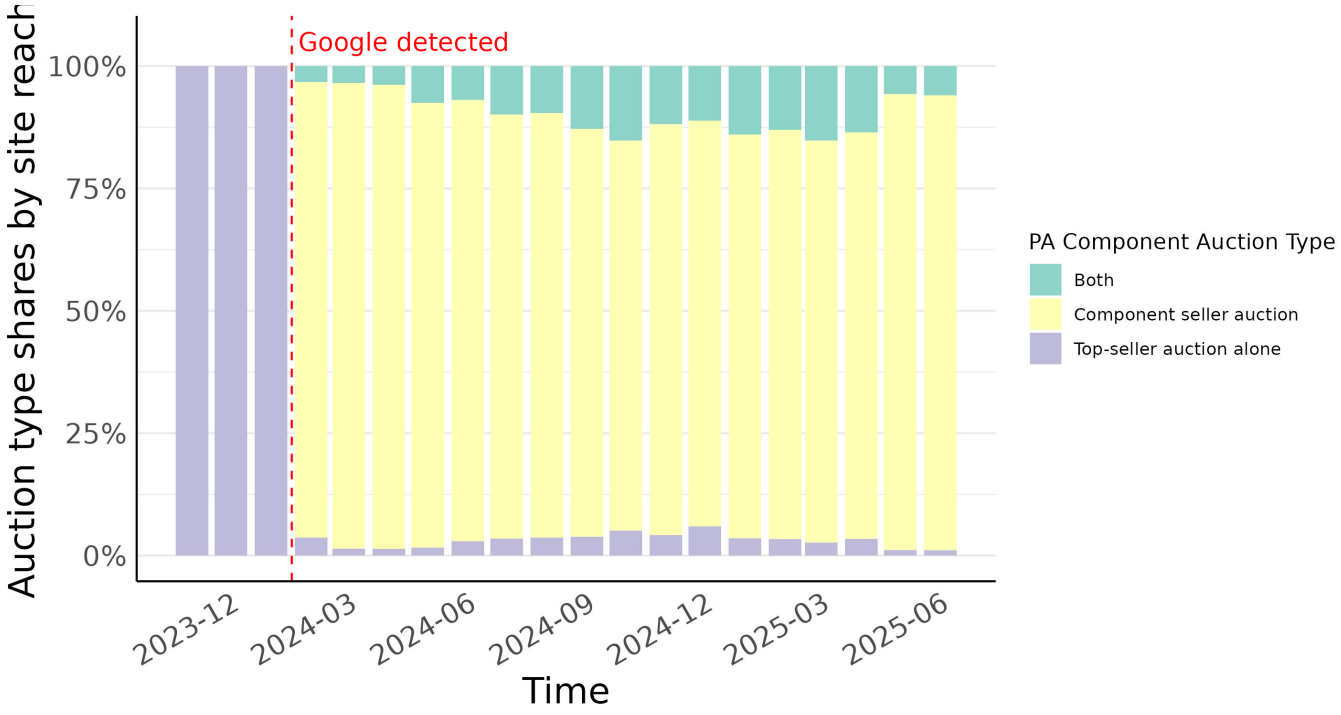
D.1 PA API component auctions use trends

Figure A.5 plots the share of PA API component auctions over time. We categorize those sites that employ PA API auctions (see Figure 2b) by whether they use a component auction, a top-level seller auction alone, or both during a given month. Prior to February 2024, we only detect auctions from the top-level seller Seedtag, which did not use component auctions during that period. After February 2024, the vast majority of sites rely on PA API component auctions, with top-seller-only configurations falling below 6%. Google accounts for nearly all top-level seller roles in component auctions.[44]

D.2 PA API component-auction seller market structure

Figure A.6 summarizes the market structure for component auction sellers. This analysis conditions on the set of auctions that employ component auctions and therefore runs from February 2024 to June 2025 (see Figure A.5). Figure A.6a shows that Google began with a 90.4% share in February 2024, with Raptive, Index Exchange, and OpenX accounting for most of the remainder. Pubmatic, Magnite, and TripleLift all enter the next month in March 2024. Supply-side platform Media.net's participation peaked in mid-2024. Magnite, OpenX, Pubmatic, Raptive, and TripleLift all but exit by June 2025.

[44]We observe small scale tests where other vendors experiment with component auctions. In the largest such test, Seedtag serves as both top-level seller and sole component auction seller. We see Onetag do this as well. In another case, we see Pubmatic serve as top-level seller where Bounce Exchange serves as the



Note: As discussed in Section 3.1 and indicated by the dotted vertical red line, we do not observe sites that use Google for PA API until February 2024 for technical reasons. Note that component seller auctions can employ a single component seller, and that component seller can also be identical to the top-level seller.

Figure A.5: Evolution of Protected Auction API by use component seller auctions

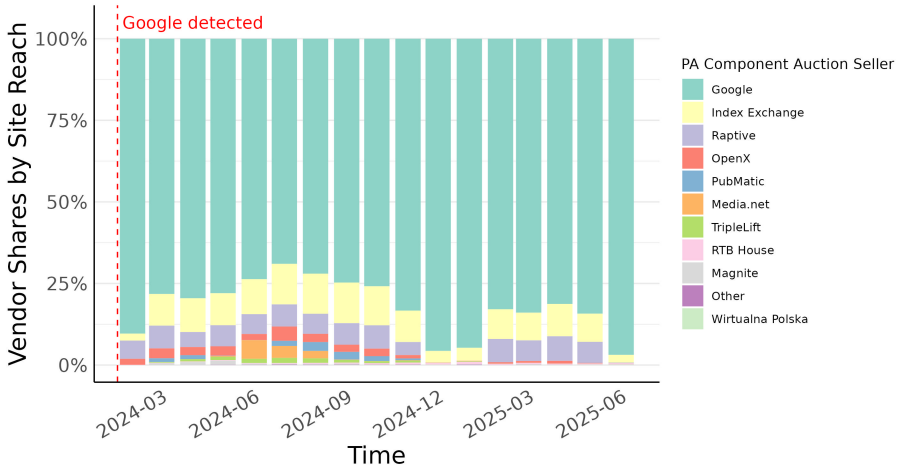
The aggregate shares over the sample period (Figure A.6b) confirm Google's dominance at 81.2%, with Index Exchange (9.0%) and Raptive (5.4%) the next largest participants. Smaller vendors collectively account for under 5%.

Market concentration, measured by HHI (Figure A.6c), frequently exceeds 9,000 points—indicating extreme concentration. It dipped to just under 5,000 in July 2024, showing the market was at times less concentrated than the top-level seller market.

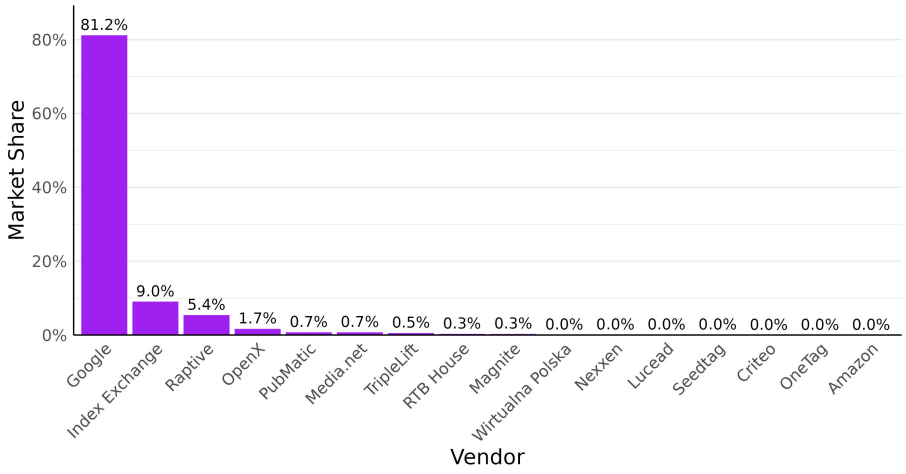
E Protected Audience API input data

Here, we examine the raw JSON data that is being passed through PA API (see Figure 2) on both the demand- and supply-sides using early data from Sincera.

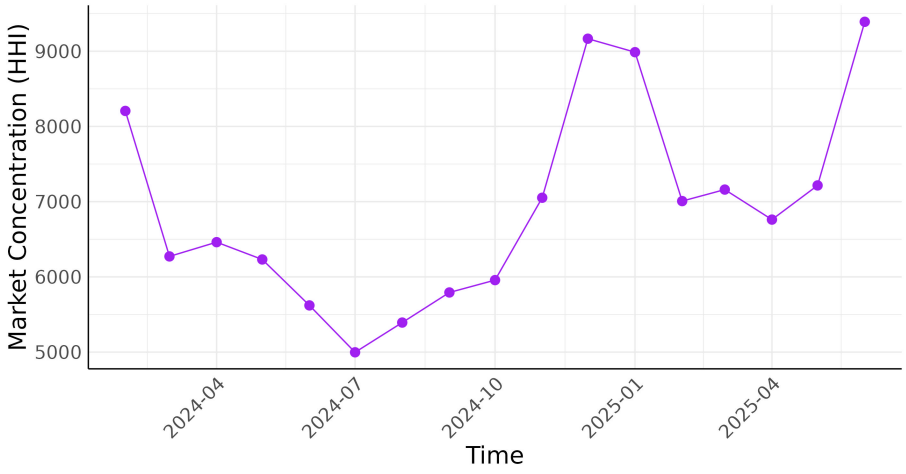
sole component auction seller.



(a) Evolution of top 10 vendor market shares



(b) Aggregated vendor market shares (Feb. '24-June '25)



(c) Evolution of market concentration (Feb. '24-June '25)

Note: As discussed in Section 3.1 and indicated by the dotted vertical red line, we do not observe sites that use Google for PA API until February 2024 for technical reasons. All figures cover the February 2024 to June 2025 period as we do not observe component auctions in the data beforehand.

Figure A.6: Protected Audience API market structure: Component auction sellers

E.1 Interest group input data

Below, we discuss the PA interest group data in greater depth using data through July 2023, covering the wider sample of all sites tracked by Sincera at that time. We intercept PA interest group data, which provides some insight into ad buyers' preferences and how they use the PA API. Nevertheless, these signals are often masked or encrypted to protect the advertiser's and ad tech vendor's business secrets. Notably, however, the "RenderURL" parameter can reveal the ad creative. We have not systematically investigated this data, but manual inspection suggests these are predominantly retargeting ads, where the creative often features multiple product images. At the time, Sincera detected 58,151 different PA interest groups being created on 4,042 publishers. This included 4,258 advertisers (according to the "name" field used by the API) yielding 4,521 unique publisher–advertiser combinations.

Only Criteo and RTB House are observed as interest group owners at this time, and we use the detailed interest group data to compare how the two vendors use PA API. RTB House explains that interest groups can contain the advertiser's first-party user data to personalize the ad creative (Domagała, 2024). On average, RTB House interest groups included 9.9 ad creative entries, but Criteo interest groups included only 0.8. Though vendors must specify an ad creative URL in this field, Criteo did not in 49.5% of cases. Criteo explained that this was a deliberate choice that reflected testing behavior (Höring, 2022). Through the "adComponents" field, PA allows vendors to create ad creatives with subcomponents: e.g., dynamic retargeting creatives can use "adComponents" to feature different product images. Criteo did not use this functionality, whereas RTB House included an average of 194 such components (minimum of 9 and maximum of 1,400). RTB House always set the interest group duration of Chrome's maximum of 30 days, whereas Criteo always set this to 43.2 minutes (perhaps erroneously).[45]

[45]The duration must be set in milliseconds, but Criteo has specified the duration to be the total seconds in 30 days.

In terms of bidding, PA allows vendors to specify a numerical priority for the interest groups in cases where the seller restricts the number of interest groups that it considers per vendor in the PA auction. RTB House is specifying this parameter statically, while Criteo is not. Neither company used PA's optional parameters for dynamically calculating priority .[46] RTB House alone is using PA's option to run bidding functions in WebAssembly, which is potentially faster than the JavaScript default. Finally, Criteo is passing a user ID as part of its userBiddingSignals parameters: presumably this is used for testing and debugging. In sum, Criteo created more PA interest groups, whereas RTB House seemingly used more of the technology's functionality in 2023.

E.2 Auction group input data

From early data that we collected until June 2022, we detected PA API auctions on 1,363 distinct sites among the wider sample that Sincera scanned at the time (0.5% of the total). As with the buy side, the raw JSON data provides some insight into this ad market while masking and encrypting several business-sensitive inputs. We note that the data identify partners that can be included in the PA auction. However, ad tech vendors need not be identified in the PA auction as buyers to participate in the auction, so this undercounts vendor participation. Nevertheless, the PA auction API data most often lists Google alone (67.7% of cases), RTB House alone (30.5%) as buyers,[47] both in 1.0% of cases, or none in 0.7% of cases.

The auction data also identify specific data that are passed to individual vendor participants under the `perBuyerSignals` input to the PA API. We find that Google's sell-side initiated auctions pass less volume of data on average to RTB House (3.7 kilobytes) than to Google's buy side (10.0 kilobytes). However, these may not be directly comparable due to differences in encoding and data structures. RTB House uses a JSON structure with 18

[46] These are the `priorityVector`, `prioritySignalsOverrides`, and `enableBiddingSignalsPrioritization` parameters.

[47] RTB House appears as three domains, reflecting servers in different locations throughout the world.

data elements, some of which may be interpretable (e.g., time stamps, ad size format, URLs). Google instead uses a nested array consisting of numbers.[48]

[48]Note that this data structure includes a large number of space characters for formatting, which appears to artificially increase the data size (e.g., by a factor of 4 in one case).