

Quantitative testing of Google's Privacy Sandbox technologies – Additional CMA guidance to third parties on testing

26 October 2023

Introduction

1. On 29 June 2023, the CMA published a guidance note (the '**testing guidance**')¹ proposing two experimental designs market participants can use to test the effectiveness of the Privacy Sandbox and how they align with Google's planned launch of two testing modes in Chrome.²
2. To ensure it is clear to market participants how they should use Google's testing modes, this note provides further practical guidance on:
 - (a) How supply and demand-side platforms ('**SSPs**' and '**DSPs**' respectively) should utilise the experimental treatment and control groups and accompanying labels when testing; and
 - (b) How market participants on the demand side (i.e. advertisers and DSPs) can participate in experiments in a way that seeks to ensure advertisers/campaigns in the treatment and control groups are comparable.
3. For details on the experimental designs and how and when to submit results to the CMA, please refer to our testing guidance.³

¹ Quantitative testing of Google's Privacy Sandbox technologies - CMA guidance to third parties on testing. Our testing guidance was largely based on our November 2022 note on [Quantitative Testing of Google's Privacy Sandbox Technologies](#), in which we requested feedback from market participants on the two proposed experimental designs.

² See Google's original announcement on its Chrome-facilitated testing environments here: [The next stages of Privacy Sandbox: General availability and supporting scaled testing](#); and accompanying developer blog post here: [Preparing to ship the Privacy Sandbox relevance and measurement APIs - Chrome Developers](#). Since we published the testing guidance, Google has announced further details on how these testing modes will work in practice, including how traffic will be labelled so ad techs can identify experimental groups. You can find these publications here:^k

³ [CMA guidance to third parties on testing \(publishing.service.gov.uk\)](#)

How ad techs should use Google's experimental labels

4. The experimental designs outlined in the testing guidance rely on comparing outcomes of auctions in a group of treatment impressions with those in two groups of control impressions:
 - (a) **Treatment group**: impressions served without using data related to third-party cookies ('TPCs').
 - (b) **Control 1**: impressions served with data related to TPCs and removing data related to new Privacy Sandbox APIs (the '**status quo**').
 - (c) **Control 2**: impressions served using neither the Privacy Sandbox tools nor TPCs.

Experimental design 1

5. Chrome will provide a persistent set of labels for a subset of Chrome browsers, allowing for coordinated experiments across different ad techs on the same set of browsers. The labels do not affect any Chrome functionality, including the availability of TPCs; it is down to the participating parties to enforce the relevant parameters for the experiment.
6. These labels will be provided by Chrome under 'Mode A', its initiative to label browser groups to facilitate coordinated experiments without affecting any Chrome functionality.⁴ Chrome functionality to support this 'label only' testing is expected to launch on 31 October 2023 and will remain available through the first half of next year. Ad techs can use these labels to bid on impressions in each group using the appropriate technology (outlined in 4(a) to 4(c) above) and compare outcomes across the groups. This is in line with experimental design 1 in the testing guidance, and can be achieved by:
 - (a) SSPs using labels to identify traffic in each experimental group and suppressing/removing the signals that should not be used before issuing a bid request; or
 - (b) DSPs using the relevant signals to place bids in each of the experimental groups.
7. Results from experiments based on experimental design 1, if available earlier, might provide some useful information that could be incorporated into the design of the Privacy Sandbox APIs or the design of future experiments

⁴ <https://developer.chrome.com/en/docs/privacy-sandbox/chrome-testing/>

based on design 2. However, as experimental design 2 provides 'true suppression', we are likely to put more weight on results based on that design in our final assessment.

Experimental design 2

8. To facilitate more robust experiments along the lines of design 2 in the testing guidance, Google will also deprecate TPCs for 0.75% of randomly selected traffic to form the treatment group (comprising three 0.25% treatment labels).
9. It will also randomly suppress both the Privacy Sandbox and TPCs in a small fraction of traffic (roughly 0.25%) to form a new control group 2. These groups will be formed by Chrome under 'Mode B', its initiative to disable TPCs for a subset of browser to more closely approximate the conditions that these experiment groups are meant to represent.
10. Google will also provide a labelled 1% control group 1 (comprising four 0.25% control 1 labels) which still has TPCs. Note that Google will not suppress the Privacy Sandbox signals in control group 1. Because browsers in these groups will not have any of their functionalities altered, it is provided as part of Chrome's 'Mode A' initiative, however this group should be used as part of experimental design 2 rather than experimental design 1. The launch of these control 1 labels will align with the start of Mode B (in Q1 2024).
11. Having the Privacy Sandbox tools available in control group 1 will allow market participants to report conversion-based metrics in control group 1 and the treatment group that are measured using the same technology, based on Privacy Sandbox tools. If the Privacy Sandbox tools were not available in control group 1, any comparison of conversions between control group 1 and the treatment group could conflate true differences in the effectiveness of the technologies used in these two groups with 'noise' due to the different methods of measurement (ie Privacy Sandbox APIs necessary for measurement in the treatment group and TPCs in control group 1).
12. In addition, having the Privacy Sandbox tools and TPCs available in control group 1 will allow market participants to compare conversions as measured by different technologies for the same set of impressions (in control group 1). This will shed light on the effectiveness of the measurement solutions supported by the Privacy Sandbox tools.
13. From a practical perspective, it is also technically complex for Google to suppress all the Privacy Sandbox tools except the APIs necessary for measurement, meaning all the Privacy Sandbox APIs will be active control group 1 traffic.

14. However, because control group 1 is meant to represent the status quo, market participants should **refrain from using Privacy Sandbox-related signals when transacting ad requests in this group**. To the extent possible, SSPs should suppress these signals, and DSPs should refrain from using them if they have not been successfully suppressed (as described in paragraphs 4(a) and 4(b) above).⁵
15. The table below shows how we advise market participants to use Google's testing modes to make comparisons between treatment and control groups in each experimental design from the testing guidance. Generally, for design 2 market participants should use the treatment group, control group 2, and control group 1.⁶

Table: how market participants should use Google's Chrome-facilitated testing modes to run the two experiments outlined in the CMA's testing guidance

	Treatment versus control 1	Treatment versus control 2
Design 1	Labels chosen by ad tech, for example: Chrome's label_only_3 as treatment Chrome's label_only_1 as control 1	Labels chosen by ad tech, for example: Chrome's label_only_3 as treatment Chrome's label_only_2 as control 2
Design 2	Chrome treatment labels Chrome control 1 labels	Chrome treatment labels Chrome control 2 labels

Demand side participation in experiments and comparability between treatment and control groups

23. The testing guidance outlined how Chrome intends to randomly allocate users to treatment and control groups to ensure that the different groups include, on average, the same types of impression opportunities. This is an effective way to ensure comparability between experimental groups on the **supply side**.

⁵ Given TPCs will be available to use in control group 1, we consider there would be little incentive for market participants to use the Privacy Sandbox tools when bidding on these impressions in any case.

⁶ As discussed in paragraph 12, as part of design 2 market participants can also use both TPCs and the Privacy Sandbox APIs necessary for measurement within control group 1. This will allow them to compare measurement outcomes on the same set of impressions using the two different technologies to understand the effectiveness of the Privacy Sandbox tools,

24. For valid experimental comparisons, participants on the **demand side** (ie advertisers/campaigns) should also be comparable across the treatment and control groups.
25. Ideally, each campaign administered by a participating DSP should have an equal chance of participating in auctions within the treatment and control groups, irrespective of the campaign's characteristics or its expected valuation of bids in the different experiment groups. While this might not always be achievable, market participants (notably DSPs) should take care to ensure that their processes do not create any systematic differences in the mix of campaigns participating in the different experiment groups. This can be achieved in several ways:
- (a) **Advertiser experiment participation controls:** ad techs should not create new controls that directly allow advertisers to choose whether or not to participate in the treatment or control arms.
 - (b) **Advertiser-facing features:** if an ad tech plans to introduce new advertiser facing campaign types or features post-TPC deprecation and wants to include them in the experiment, they should ensure that the budgets allocated to those new features/campaign types cover both the treatment and the relevant campaign type in the control during the experiment. If an ad tech does include new campaign types or features in the experiment, they should submit details of these changes and their methods for ensuring comparability between the treatment and control groups to the CMA when submitting results. They should also consider additional breakdowns of results with and without these new campaigns/features.
 - (c) **Interaction between experiment groups:** where possible, ad techs should take steps to avoid outcomes in the treatment impacting outcomes in the control and vice versa. For example, ad techs should not introduce new logic to shift budgets from treatment to control depending on performance. Another example of an interaction between groups could be a bidding optimisation model that involves comparisons between multiple incoming bid requests (including those coming from both treatment and control groups) before deciding what impressions to bid on or the level of the bids. Where any interaction between groups or budget shifts might naturally arise, it is helpful to provide additional detail to the CMA when submitting results explaining the high-level mechanism for this interaction.

Next steps

26. This note seeks to provide market participants with further clarity on how they can test Google's Privacy Sandbox tools in a way that will be useful for our assessment. Market participants who are able to quantitatively test the Privacy Sandbox, should do so **before the end of Q2 2024** so we can use their test results in our assessment.⁷ If you would like to engage with testing but have questions on the testing guidance or this update, please contact us on privacysandbox@cma.gov.uk.

⁷ Please see our [testing guidance](#) for detailed information on how to submit results to the CMA, including templates for tables and the type of accompanying information that will help us understand market participants' tests.