Quantitative testing of Google's Privacy Sandbox technologies – seeking input from affected firms and others on the CMA's proposals

November 2022

Introduction

1. Google's Privacy Sandbox is a set of technologies that allows the removal of third-party cookies ('**TPCs**') on Chrome, with the aim of protecting people's privacy while allowing companies to personalise adverts and enabling other functionality. It will lead to a significant change in how online advertising functions, globally. To make sure advertising will function well and competitively after the Privacy Sandbox's introduction, the Competition and Markets Authority ('**CMA**') agreed legally binding Commitments with Google in February 2022 about how the technology will function (the '**Commitments**').

2. The Commitments include testing and trialling the new technology. The best way to conduct testing is through having many different parties involved. We are now looking to hear the views of many of the firms who will be affected and who might consider being part of the testing programme, as well as others interested in the Privacy Sandbox. In particular, we are keen for market participants to:

   (a) Respond to this note and provide their views on the proposals, including ways in which the testing could be made more effective;

   (b) Consider engaging in trials and to highlight any barriers that might prevent them from doing this.

3. This note sets out our initial thinking on how quantitative testing might inform our assessment of the impacts of the Privacy Sandbox changes. We see quantitative testing and trialling as being a key part of the Commitments agreed with Google. While it will not be possible to test all elements of the Privacy Sandbox tools, and our ultimate assessment will need to take account of wider evidence and views on qualitative impacts, we think that effectiveness testing, including by third-party market participants, can provide an important source of evidence on possible impacts on competition outcomes.

4. We expect to work with Google to develop the initial ideas in this note further and to design experiments and trials which Google can carry out itself. As set out in the Commitments, we will require Google to be transparent in reporting the results of these tests, including the underlying data where appropriate, so that they can be properly scrutinised.

5. At the same time, we also want to encourage other market participants to get involved in testing themselves. Google's advertising business represents only one part of the market, and it is very important that we understand likely impacts of the Privacy Sandbox changes on wider market participants and competition. The note therefore includes some initial suggestions on approaches to coordinate testing across different market participants.

6. We are very keen to get feedback on the ideas in this note. While we have already discussed possible approaches with Google and a small number of other market participants, we want to ensure we get the broadest possible input from interested parties and industry experts, to help us develop a robust methodology. We have set out a series of questions throughout the note on which we would be particularly keen to hear views. **We would welcome responses by 1 December 2022.**

7. The remainder of the note is structured as follows:

   (a) We first provide some context on Google's Commitments, the current implementation timeline, and the role of testing and trialling;

   (b) Second, we describe our main proposals for quantitative experiments, including experimental designs, proposed metric and the potential role of third parties;

   (c) Third, we describe alternative testing approaches where full trials are not realistic;

   (d) Fourth, we ask some broader questions for general feedback;

   (e) Finally, we set out our proposed next steps.

Context: The Commitments, implementation of the Privacy Sandbox and role of testing and trialling

8. Under its Privacy Sandbox proposals, Google plans to deprecate the use of TPCs in Chrome by the third quarter of 2024 and replace many use cases they currently serve with alternative technologies.[1]

9. In February 2022, Google agreed formal Commitments to address the CMA's competition concerns about potential impacts of the Privacy Sandbox proposals. The Commitments establish a set of Development and Implementation Criteria against which we will assess the effectiveness of the Privacy Sandbox changes before Google deprecates TPCs. These criteria include impacts on publishers, advertisers, and competition.[2]

10. We anticipate that our assessment of effectiveness at the Standstill Period before removal of TPCs will draw on a wide range of quantitative and qualitative evidence from both Google and third parties. One element of the quantitative evidence we gather will be results from tests Google carries out to measure the performance of the Privacy Sandbox technologies. As recognised in our Commitments Decision, we understand that third-party participation in quantitative testing will also be a key factor in our ability to carry out a well-rounded assessment of the Privacy Sandbox.[3]

11. However, we understand that including third-party market participants in testing with this aim is challenging due to the complexity of the market for digital advertising and Google's Privacy Sandbox changes; and the significant resources market participants might be required to invest in testing. One aim of this note is to provide greater certainty for market participants on possible approaches to effectiveness testing to help them plan future trials.

12. Our concerns regarding the impact on competition and impact on publishers are primarily related to the *effectiveness* of the targeting and measurement APIs in replicating the functionality currently supported by TPCs.[4] For example, the CMA has previously outlined concerns that, if designed in a way that restricts functionality for third parties but not for Google, Topics, FLEDGE,

---

[1] See Building a more private web, 22 August 2019, and Building a more private web: A path towards making third-party cookies obsolete, 14 January 2020.
[2] The Commitments, paragraphs 8(a) to 8(e) . There are three further Development and Implementation Criteria: impact on privacy outcomes, impact on user experience, and technical feasibility that we are not seeking to evaluate through quantitative testing.
[3] Commitments Decision, Appendix 4, paragraphs 10 and 17.
[4] See Commitments Decision, Appendix 4, paragraph 13 for the difference between *functional* and *effectiveness* testing; and paragraph 12 for the three broad use cases Google has committed to consider in its quantitative testing.

or the Measurement APIs could distort competition in the supply of ad inventory or ad tech services.[5] Such distortions might result in higher demand for Google's owned and operated inventory (and so disproportionately higher prices relative to other publishers) or a higher volume of traffic flowing through Google's ad tech services.

13. As described in the Commitments Decision, concerns of this nature are best suited to testing through experiments.[6] These experiments will involve measuring market outcomes for samples of impressions auctioned using the new Privacy Sandbox APIs and comparing them against the same metrics for samples of impressions marketed in some counterfactual settings (see further below for a discussion of the role of counterfactuals in the experiments).

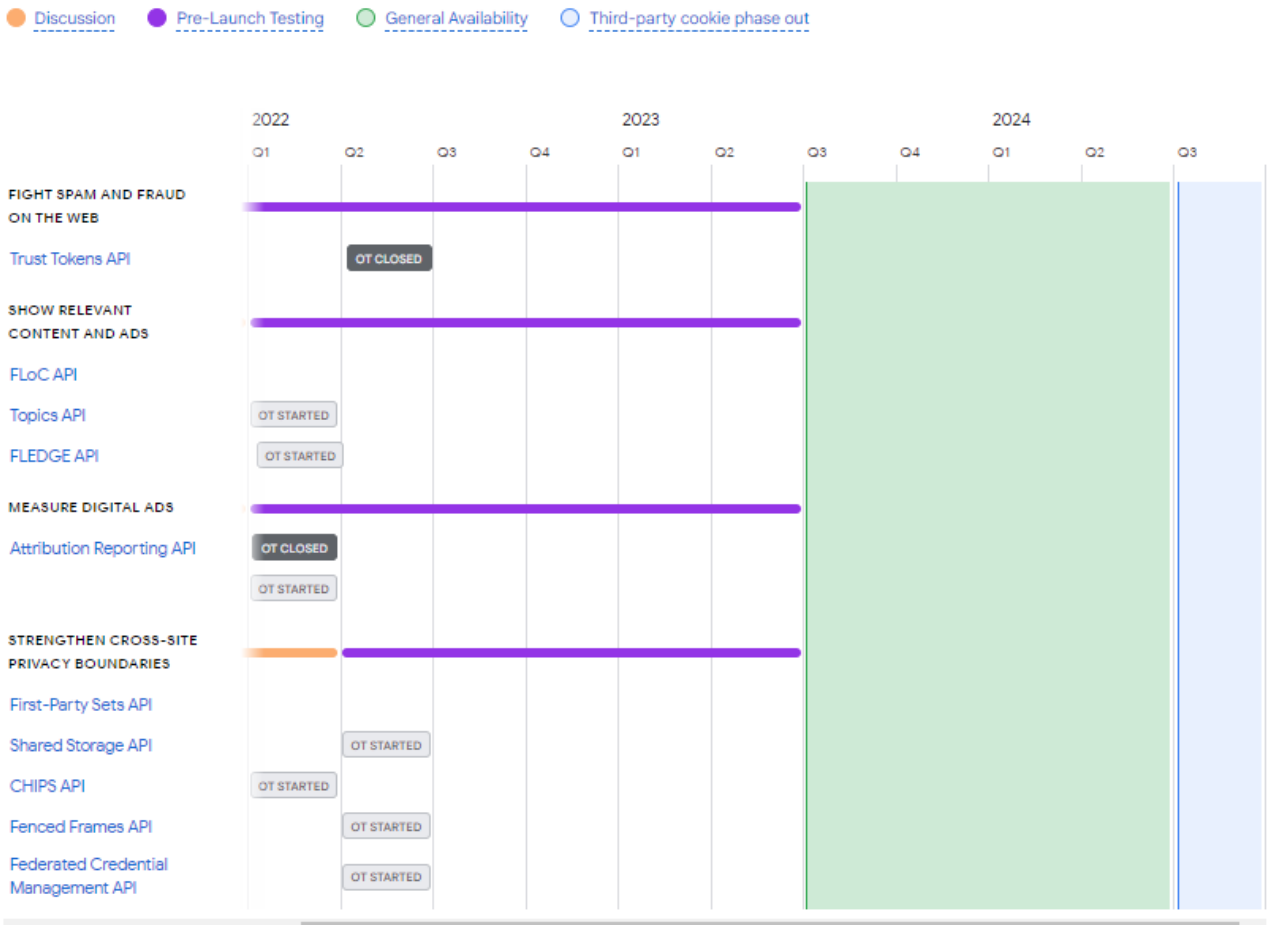
14. Experiments are not the only means of quantitative testing we intend to use in its assessment. There are other features of the Privacy Sandbox technologies for which experiments might not be informative or feasible, and so are better assessed through alternative quantitative tests. For example, we understand that some aspects of technical performance, such as latency and accuracy/consistency of consumer categorisation or attribution reports might be more comprehensively measured through simulation.[7] In addition, we intend to use the results of both experiments and any alternative quantitative tests alongside a range of evidence we will collect as part of our assessment.

15. The Stable Origin Trials ('**OT**') for Topics, FLEDGE, and the Attribution Reporting API started in Q1 of 2022 and are scheduled to run until the end of Q2 2023. During this phase, Google is carrying out functional testing. From Q3 2023, Google will move the APIs to General Availability until its planned date of TPCs deprecation in Q3 of 2024.[8] Our current understanding is that Google intends to conduct its quantitative testing during General Availability, between Q3 2023 and Q3 2024, to, where possible, assess the effectiveness of the Privacy Sandbox against the Development and Implementation Criteria.[9] We intend to update stakeholders with a better idea of timing in the coming months. Figure 1 below shows the most recent version of Google's timeline for the Privacy Sandbox.

---

[5] Commitments, paragraph 7. See also Commitments Decision, paragraphs 3.29 to 3.83.
[6] Commitments Decision, Appendix 4, paragraph 15(a). The experiments we describe discuss in this not can also be referred to as A/B tests.
[7] Commitments Decision, Appendix 4, paragraphs 15(b).
[8] See 'The Privacy Sandbox Timeline for the Web'.
[9] Google's development of the APIs will continue into the General Availability phase. See Commitments Decision, Appendix 4, paragraph 13 for the difference between *functional* and *effectiveness* testing; and paragraph 12 for the three broad use cases Google has committed to consider in its quantitative testing.

Figure 1: Google's timeline for the deprecation of TPCs and testing



Source: 'The Privacy Sandbox Timeline for the Web'

Quantitative testing through experiments

16. This section sets out our initial thinking on methodologies and approaches to quantitative trials, based on observing real-world outcomes where the Privacy Sandbox changes are tested on a subset of market participants.

Defining a counterfactual

17. Experiments will involve comparing the market outcomes for subsamples of ads served using the Privacy Sandbox technologies against a subsample of ads served under some **counterfactual** scenario. To make this comparison, we propose allocating ad requests that are part of experiments to the following three experimental arms:

    (a) **control group 1:** keeping data related to TPCs and removing data related to new APIs before issuing the request for bids;

(b) **control group 2**: removing data related to both TPCs and the new APIs; and

(c) **the treatment group**: removing data related to TPCs and keeping data related to new APIs.[10]

18. The treatment group is designed to serve ads in a way that represents all the various signals and technologies that will actually be available after the deprecation of TPCs. The two control groups can be viewed as counterfactual scenarios in which either the current use of TPCs continues (control group 1) or there are no replacement technologies after TPCs are deprecated (control group 2).

19. Ordinarily, counterfactual scenarios would represent what is known to be the most likely state of the world in the absence of the treatment. Neither of the two control groups described above represents a true counterfactual in this sense. As regards control group 2, we are aware that a counterfactual representing no alternative technologies is not necessarily realistic. However, it is not currently possible to define a counterfactual that might represent some unknown future technology. We see control group 2 as providing a 'bound' on market outcomes without ad-supporting technologies. As regards control group 1, it may be that some practices or processing of personal data based on TPCs are unlawful. As such, this comparison would not imply that their continuation is a legitimate counterfactual for the purpose of measuring the effectiveness of the Privacy Sandbox technologies. Nevertheless, it will be useful to understand how the Privacy Sandbox technologies perform against the status quo.

20. As a result, we do not intend to use comparisons between the treatment group and the two control groups in isolation in the assessment. Rather, we intend to use comparisons across the groups alongside wider quantitative and qualitative evidence to understand what impact the deprecation of TPCs and introduction of the Privacy Sandbox will have on competition.

Experimental designs

21. A key preliminary step in experiments will be specifying how ad requests are allocated to the treatment and control groups defined above. We are currently considering two ways in which this might be done. The first involves individual **market participants creating their own treatment and control groups** by

[10] We understand that in the treatment group market participants might also use other replacement technologies available at the time of the experiments alongside the Privacy Sandbox APIs.

selectively opting to suppress the use of TPCs and using the Privacy Sandbox APIs to create treatment groups. In the second, these groups would be **created by Chrome**. Below, we outline these two designs and discuss their main justifications and/or limitations.

Design 1 – Google Ads or other market participants creating their own experimental groups

22. This approach works as follows:

(a) **Step 1:** from the universe of all ad requests for traffic made available by Chrome for testing for which Privacy Sandbox APIs have been enabled by the user, a market participant randomly selects a subsample (size to be determined);

(b) **Step 2:** starting from the subsample selected at step 1, market participants randomly allocate each ad request to one of the three experimental groups based on 17(a) to 17(c):

(i) keeping data related to TPCs and removing data related to new APIs to construct **control group 1**;

(ii) removing data related to both TPCs and the new APIs to construct **control group 2**; and

(iii) removing data related to TPCs and using their own solutions integrated alongside any use of the Privacy Sandbox APIs to construct the **treatment group**.

(c) **Step 3:** bids are placed on the three categories of impressions, the ads are served, and outcomes (discussed more below) are measured and reported.

23. In this design, it is not possible to guarantee that, for a given market participant and treatment ad request, all participants in the auction do not have access to (or opt out of using) TPCs.[11] If some participants do use TPCs in their bidding, comparing the outcomes of the treatment and control impressions will not be fully informative of the potential impact of the Privacy Sandbox changes. However, Design 1 does have the advantage that it can be implemented by market participants at an earlier stage to, for example, experiment with individual Privacy Sandbox APIs. We are interested in

[11] For example, in the presence of header bidding.

understanding the results of market participants' testing at various stages of Google's testing and trialling timeline (Figure 1). We would therefore encourage those who can adopt the approach in Design 1 to do so and would be interested to understand their results.

Design 2 – Chrome creating the experimental groups

24. As it is not possible to rule out the use of TPCs for all parties participating in the auction consistently in Design 1, the second approach relies on **Google defining treatment and control groups within Chrome** for a slice of traffic. We are discussing with Google whether this approach would be feasible in practice. If so, we envisage that the design could proceed as follows:

    (a) **Step 1:** starting from the universe of all users for whom the Privacy Sandbox APIs are available for use in their browser and who have consented to be part of Google experiments,[12] Chrome would randomly allocate each to the control and treatment groups.

    (b) **Step 2:** for the duration of the experiment and for a slice of traffic (size to be determined),[13] Chrome would:

        (a) keep data related to TPCs and remove data related to new APIs before issuing the request for bids for users in **control group 1**;

        (b) remove data related to both TPCs and the new APIs for users in **control group 2**; and

        (c) remove data related to TPCs and keep data related to new APIs for users to incorporate into their existing solutions in the **treatment group**.

    (c) **Step 3:** market participants would observe control and treatment ad requests and bids would be placed on the three categories of impressions, the ads served, and outcomes measured.

25. The removal of TPCs might be achieved by mimicking the effect of users blocking all TPCs in Chrome. As a result, websites including ad tech providers

[12] During General Availability, the Privacy Sandbox APIs will only be available for users who have updated their browser. It is possible there will be other criteria used to select the sample of users that will be included in the experiment, for example excluding all those who have actively disabled TPCs in Chrome. This aspect of the design is still under consideration.
[13] Some users might drop out of the experiment if, for example, they change their consent. Allocation to experimental groups might also be re-randomised over time. These aspects of the design are still under consideration.

would be prevented from accessing TPC information, and so any DSPs participating in the auction for a treatment impression would not be able to place bids based on user information associated with TPCs.

26. We have discussed this design with Google, which is currently exploring its practical and legal feasibility, together with its user experience implications. These aspects may affect the content of the test design, and the geography in which it is implemented. For example, to ensure market participants can identify bid requests originating from traffic allocated to each experimental arm (ie control 1 or 2, treatment), Google is considering the feasibility of developing a mechanism to inform participants of the allocation of each ad request. The primary purpose of this feature would be to enable market participants to use their experimental resources appropriately. We are also considering the data protection implications of using this mechanism in experiments.

27. Although there remain challenges with this approach (described in more detail in the next subsection), it would ensure that impressions allocated to the treatment group would not contain a TPC ID that DSPs could use in their bidding logic. Google is currently exploring the technical feasibility of blocking TPCs in Chrome for this particular purpose. It is important to recognise that the test design would seeks to minimise any disruption from a potential Chrome-side block of TPCs. Consequently, the sample size for any such experiment would be chosen carefully in order to balance statistical validity with the possibility of adverse impact on the ad tech ecosystem. Google is carefully considering how best to minimise any such disruption, and we welcome feedback on this too.

The key role of third-party engagement in experiments

28. As part of its Commitments, Google will design and carry out quantitative tests in agreement with the CMA as it develops the Privacy Sandbox technologies.[14] However, using only Google-run tests would not provide a full understanding of the impact of the Privacy Sandbox on the wider advertising and ad tech ecosystems.

29. Google is considering experiments based on both Design 1 (to test individual APIs) and Design 2 (to test the Privacy Sandbox APIs together during General Availability). In both cases, this would involve Google SSPs measuring outcomes for subsamples of impressions allocated to the

[14] Commitments, paragraph 17(c)(i)-(vi).

treatment and control groups. Third-party engagement in this programme of experiments is important for two reasons.

30. First, it is desirable that third-party DSPs and the advertisers they represent bid on the experimental impressions marketed by Google SSPs using the appropriate technology/models and in good faith, so that outcome prices and quantity provide a reasonable proxy for the true equilibrium value of these impressions in the different treatment and control groups. Second, it would be informative if other SSPs (and participants on the demand side) could run their own experiments along similar lines so that the CMA and the market has information on experimental outcomes for impressions that are not marketed by Google SSPs (which might differ from the subset of impressions marketed by Google SSPs).

31. However, there are several aspects of the digital advertising market that mean it is difficult to make economically meaningful and representative comparisons through the experiments described above. Both designs would require market participants to conduct their own experiment at whatever level of the ad tech stack they operate. However, we are seeking to measure differences in overall market outcomes.

32. Given this, we are seeking to:

    (a) Encourage a wide range of market participants, particularly SSPs, who plan to experiment with and eventually incorporate the Privacy Sandbox technologies into their own solutions, conduct their own experiments and report the results to us. This would enable us to draw on a wide range of results/feedback from various segments of the market in its assessment.

    (b) Where appropriate, facilitate the coordination of experimentation among market participants, including Google, to ensure auctions for experimental ad requests are being bid on using the appropriate technologies/models.

33. Such coordination could be achieved by, for example, SSPs advertising their participation in experiments with market participants on the demand side, or where a group of publishers, SSPs, DSPs, and/or advertisers are able to organise an experiment among themselves. This could result in the outcomes of auctions being a reasonable proxy for the true equilibrium outcomes in the different treatment and control groups. We are particularly interested to hear feedback from market participants on the feasibility of and potential approaches to coordinated experimentation (see the questions below).

34. We understand that it is not possible to fully capture long-run effects of the Privacy Sandbox changes due to adoption and learning. However, by engaging with market participants to understand adoption, in-house experiments, and the development of their models for auctions in Chrome post-TPCs deprecation, we can consider quantitative tests in context and make a rounded assessment of the Privacy Sandbox.

Potential metrics

35. To understand the potential impact of the Privacy Sandbox on different aspects of competition, there are a range of outcomes we would like to assess through experiments. These could include:[15]

    (a) Revenue per impression (**impact on publishers/competition**)

    (b) Clicks per dollar (**impact on advertisers/competition**)

    (c) Conversions per dollar (**impact on advertisers/competition**)

    (d) Clicks per impression (**targeting effectiveness/competition**)

    (e) Latency (**impact on publishers/competition**)

    (f) Client crashes per 1 million page loads (**impact on publishers/competition)**

    (g) Share of users closing ads (**targeting effectiveness/competition**)

    (h) Volume of activity through supply/demand-side platforms (eg total unique bid requests or bids, **ad tech competition**)

36. The primary focus of experiments will be to compare these (or any other relevant) outcomes across the control and treatment groups. However, we are also interested in how the impact of the Privacy Sandbox might vary across certain dimensions, for example publisher/advertiser/ad tech size, or Google versus third-party services.

37. Google can record the above metrics over time. We propose to also collect them from market participants who engage in experiments over several months, as opposed to at one point in time. As adoption and testing of the APIs increases over time, comparing these outcomes will allow us to measure adoption/and or learning.

[15] We understand that some of these metrics can inform multiple aspects of our assessment. For example, latency and client crashes are an important part of user experience. Given the focus of this note is on assessing the impact on publishers and competition, we list metrics here as they relate to those criteria.

38. It might not be possible to measure some of the metrics we have listed in 35(a) to 35(h) reliably, or it might be that some are not informative of the aspects of competition we are seeking to measure. In future engagement, we intend to understand differences in how these metrics might be recorded across market participants. It will also fully consider any caveats around what they might reveal.

39. As a result, we would be grateful for feedback on the metrics we have proposed, the extent to which they will or will not be informative of the outcomes we have indicated they measure, and whether there are any additional metrics that would allow us to learn about the impact of the Privacy Sandbox technologies. We would also welcome feedback on the most suitable means of providing any information – for example in the form of data or results.

Potential limitations

40. We are conscious that any approach to measuring the effectiveness of the Privacy Sandbox technologies through experimentation has inherent limitations. If TPCs are removed and replaced by the Privacy Sandbox technologies, the price and allocation of impressions will depend on the equilibrium behaviour of all market participants. It is not clear that we can truly replicate such behaviour in the context of these experiments. Not all market participants will participate in the experiments, and those that do will likely allocate a fraction of their budget and volumes to the treatment and control impressions. The existing technologies will be available in parallel and will likely continue to support the bulk of activity in the market. It is also evident that market participants will need time to learn the value of the new technologies and embed their usage in their processes, and it is not clear that this learning process can be completed with the timelines and volumes envisaged for the experiments. For all these reasons, it is not clear that the prices and allocations that will be measured in these experiments will reflect the outcomes that would truly obtain if changes were implemented at scale in the market.

41. Nevertheless, we believe that there is value in using experiments to test how market participants respond to the new technologies. As stated above, we intend to consider a broad range of metrics. We will also interpret these results alongside more qualitative feedback from participants.

CMA questions on third-party experiments and metrics

42. We welcome any feedback on the experimental designs, third-party engagement, and metrics we have described above. However, feedback on the following themes, where possible, would be particularly helpful:

(a) Have you begun any experimentation (or A/B testing) with the Privacy Sandbox targeting or measurement APIs?

(b) Independent of our plans, do you plan to experiment with the APIs going forward? Do you plan to test APIs in isolation or combined? And when would you plan to carry out this type of experimentation?

(c) Please describe any barriers to experimentation you perceive or face (eg financial incentives or resources).

(d) Do you have the capability to conduct experiments as in Design 1 (ie creating your own treatment and control groups)? If not, please explain why. In either case, please describe any practical difficulties you can identify in this approach.

(e) Do you have the capability to conduct experiments as in Design 2 (ie using Chrome's assignment of users to treatment and control groups)? Would you be able to incorporate the use of a signal of experimental status into your models? If not, please explain why. In either case, please describe any practical difficulties you can identify in this approach. If possible, please also explain any parameters, for example volume of traffic allocated to experiments, that are important in running an experiment based on Design 2.

(f) In your view, is it necessary to coordinate certain aspects of these experiments with other market participants? Please explain.

(g) If your view is that coordination is necessary, are you able to coordinate to carry out an experiment? Please explain any views on how such a coordinated approach to experimenting might be achieved.

(h) If you plan to experiment, what metrics do you plan to measure? Do they coincide with the metrics we have described above? Please provide any other feedback on how informative the metrics we have described might be of competitive outcomes.

(i) Please provide any information on the length of time or volume of traffic your models might need for initial training and, eventually, initial maturity for targeting.

Potential areas to be explored through alternative testing approaches

43. As outlined in paragraph 14, there are aspects of the Privacy Sandbox technologies that might be assessed through tests that do not involve market participants trialling the Privacy Sandbox technologies on subsamples of 'real-world' impressions. For example:

    (a) Latency;

    (b) Accuracy/consistency of measurement/attribution reports;

    (c) Accuracy/consistency of consumer categorisation;

    (d) Conversion rates;

    (e) Total conversion value; and

    (f) Entropy and information content of Privacy Sandbox signals.

44. One way in which these outcomes might be assessed is through simulations: modelling and simulating hypothetical scenarios in order to understand the impact of the Privacy Sandbox changes. However, the design and purpose of alternative tests is currently flexible. This is primarily due to the breadth of possible assessments that could be made, for example, through simulations. We are therefore particularly interested in market participants' views on alternative means of testing.

45. Our current understanding is that Google intends to carry out alternative testing over time. This will provide a picture of how the Privacy Sandbox changes are affecting the Google ecosystem in ways not captured by the experiments we have outlined above. Where possible, we are also proposing to collect them over several months from third-party market participants (see paragraphs 48 to 53 below for timelines).

CMA questions on alternative testing

46. We welcome any feedback on alternative testing and metrics. Feedback on the following themes would be particularly helpful:

(a) Independent of our plans for testing, do you plan to conduct simulations or any other non-experimental tests to record how using the Privacy Sandbox technologies is impacting your business?

(b) If you have the capability to run simulations, please describe what you might be able to simulate, for example latency or accuracy of attribution reports.

(c) If you plan to use any alternative quantitative tests, please provide a description.

(d) Are there any other metrics you think could be measured through alternative tests that would be informative of the impact of the Privacy Sandbox technologies on market participants (and competition)? Please provide a description.

General qualitative feedback

47. We would be grateful for any general qualitative feedback on experiences testing and trialling the Privacy Sandbox tools that stakeholders consider appropriate for our assessment. For example, as illustrative topics, we would welcome feedback on:

(a) How easy it is to test the Privacy Sandbox technologies or whether there are any barriers to doing so.

(b) Any plans for testing, including when and how stakeholders plan to test the various Privacy Sandbox technologies.

(c) Initial thoughts from any use of the technologies on:

(a) technical performance of the APIs and comparability with alternatives (including the status quo);

(b) the impact of the APIs on stakeholders' quality of product/service; and

(c) the impact of the APIs on stakeholders' revenues.

Timeline and next steps

48. We are seeking to engage with market participants on the design of testing between now and at least the beginning of General Availability in Q3 2023. As we have outlined in this note, this pre-experimentation engagement will inform our approach to quantitative testing. Because we do not know with certainty

when Google's experimenting might begin and given also the issues that are currently being explored in relation to Design 2 (paragraphs 24(a) to 24(c)), we cannot currently say whether or how far into General Availability this phase of engagement might extend.

49. Of course, we would encourage any market participants who plan to experiment using, for example, Design 1 (paragraphs 22(a) to 22(c)) to begin experimenting whenever they wish and would be grateful to hear about their results. As we outlined in paragraph 37 above, if possible we would like to collect experimental metrics from market participants at several stages during their experimentation.

50. This is also true for any coordinated testing, for example based on Design 2. In addition, we also intend to allow for some iteration of this experimental design in order to refine the methodology where necessary. As such, it would be beneficial to start experimenting in some capacity, for example by piloting an experiment based on Design 2, as early as possible in 2023. The exact timing of course depends on, among other factors, results of Google's functional testing and the capacity for Google and third-party market participants to engage in testing.

51. As a first step in our pre-experimentation engagement, we intend to use market participants' feedback on the proposals and questions in this note to inform, first, Google's design of quantitative tests and, second, our approach to quantitatively testing and assessing the Privacy Sandbox technologies. We envisage that future engagement will involve obtaining information on the results of third-party testing – this could be either quantitative (eg experiment results and underlying data) and/or qualitative (observations from the market) – and using this information in our final assessment of the Privacy Sandbox technologies, either alongside results from Google's tests or to provide context for such results.

52. We intend to plan future engagement in the coming months, prior to the General Availability phase, based on the feedback we receive. At that point, we will be better placed to provide more detail on how quantitative tests, and in particular experiments, can work in practice.

53. We also intend to continue direct engagement with stakeholders who can run experiments, and quantitative tests more generally, in order to ensure they are fully informed on the methodologies and aims of our testing. At any point, we encourage market participants to inform us of any plans to experiment with the Privacy Sandbox technologies and provide qualitative feedback on their experiences.