The Economics of Network Effects and User Data in the Provision of Search, Search Advertising, and Display Ad Intermediation

Andres V. Lerner

15 MAY 2019

Table of Contents

I. Introduction and Executive Summary

In the Preliminary Report of its Digital Platforms Inquiry (the Preliminary Report), the Australian Competition and Consumer Commission (ACCC) recognizes the significant benefits that digital platforms such as Google have created both for consumers and businesses. At the same time, the Preliminary Report suggests that "Google is insulated from dynamic competition by barriers to entry and expansion for search platforms…" arising from network effects, including network effects that supposedly flow from the collection and utilization of user data. The Preliminary Report also maintains that Google enjoys "strong advantages" in a cluster of services the Preliminary Report refers to as "intermediary services" (sometimes "ad tech intermediary services") in part due to Google's collection and utilization of user data.[1]

Network effects are an economic phenomenon by which the value of a certain product or service to a given user increases as the number of other users of the product or service grows. But as explained in this paper, one cannot simply assume that an online service (such as those offered by Google) exhibits network effects, let alone that network effects create a significant barrier to entry or expansion, without a rigorous fact-based examination of the relevant market(s). Real-world evidence indicates that network effects for search platforms are weak, at best. Search services are not characterized by "same side" network effects. Search services also are not characterized by strong "cross-side" network effects—while advertisers seek users, users generally ignore or may even be repelled by ads, which invalidates self-reinforcing feedback loops that could create durable barriers to entry.

Moreover, while user data are valuable inputs into the provision of search and search advertising services, there is no evidence indicating that a large volume of user data is a uniquely valuable input necessary to compete effectively; there are many other important inputs into the provision of search services, including engineering resources, crawling and indexing data, and specialized data feeds. In display ad intermediation, meanwhile, many ad tech firms have access to a variety of user data for targeting display ads. Access to user data, while valuable, therefore is not a meaningful barrier to entry or expansion in such services. Display ad intermediation also is characterized by a high degree of advertiser multi-homing, which further diminishes any network effects or other competitive advantage associated with the collection and use of data. As a result, Google competes with a large number of successful providers at each level of the ad tech stack.

Network effects in search and search advertising

Network effects benefit users and do not raise competitive concerns unless the network effects give rise to a positive "feedback loop," a self-reinforcing phenomenon where the large become larger and smaller providers and new entrants cannot catch up and compete effectively. However, network effects do not necessarily lead to such feedback loops and therefore to a market "tipping" to one large, dominant platform. There are several economic factors that limit the competitive significance of network effects in practice, and whether network effects have the potential to entrench a dominant platform depends on marketplace circumstances. In fact, in many if not most markets characterized by network effects,

[1] This cluster of services includes publisher and advertiser ad servers; ad exchanges and ad networks; supply side and demand side platforms; data management platforms; third party data; data analytics; and trading desks. Preliminary Report at 72-73. This paper refers to "display ad intermediation" when discussing this cluster of services.

multiple platforms can and do compete (*e.g.*, AirBnB competes with VRBO, and FlipKey, and others in the vacation rental services sector).

Search services are not characterized by the standard "same-side" network effects (sometimes called "direct" network effects) among users, wherein users of a particular platform or service value the participation of the same type of user (*e.g.*, users of a telephone network, who benefit from having more users that can be reached by telephone over the network). In contrast to telephone networks, search users do not place value on whether other users utilize the same search provider. The Preliminary Report suggests that users of search services like Google may benefit from improvements of quality resulting from increased user query volume, but this is not a network effect, which is a *demand-side* effect. It is, instead, a classic example of *supply side* economies of scale. Supply-side scale economies arise in many if not most businesses, and, for the reasons discussed below, do not give Google a unique or insurmountable competitive advantage.

"Cross-side" network effects (sometimes called "indirect" network effects) also do not give rise to entry barriers for search platforms. Cross-side network effects arise when the presence of members of one user group on the platform (*e.g.*, merchants in a shopping mall or online marketplace) attract members of *another* user group (shoppers). But such a dynamic does not exist between users and advertisers on a search platform. While advertisers seek users, user demand for a platform is not driven by the availability of ads. Many (if not most) users ignore ads, and many place negative value on having too many ads. One reason is that organic search results are generally superior to ads from users' perspective, even when those ads are well targeted. Ads make up only about 6% of user clicks on Google, with the other 94% going to organic (free) results.[2] The absence of any meaningful cross-side network effects for users invalidates the possibility of any self-reinforcing feedback loop, which requires *bidirectional* attraction. In the absence of such a mutually reinforcing feedback loop, if a smaller rival or new entrant offers a better service to users, cross-platform network effects do not inhibit users from switching and, once users switch, advertisers would follow.

Cross-side network effects are limited even for search advertisers. From an advertiser's perspective, the value of an online platform is that it allows the advertiser to reach users. But while an advertiser may derive more value in advertising on a platform with more users, advertising on such a platform also entails proportionately higher costs because advertisers pay on a per-click basis. As a result, advertising on a larger platform is not necessarily more profitable than advertising on a smaller one; in fact, the opposite may be true.

The Preliminary Report contends that cross-side network effects arise from "fixed costs" incurred by advertisers in using a particular platform and in running ad campaigns, which make larger platforms more attractive to advertisers because those fixed costs are spread over a larger amount of ad spend. However, real-world evidence indicates that fixed costs for advertisers to set up and run ad campaigns on a particular search platform are not significant. Advertisers can port data across different search platforms easily and at very low cost, and those costs are diminished by (1) compatible application programming interfaces (APIs), (2) tools offered by rival search platforms (such as Bing, Oath/Yahoo!, and Pinterest) that facilitate porting of ad campaigns, and (3) widely used software tools that allow advertisers to place ads across multiple search and other ad platforms (such as Marin, Adobe and Kenshoo).

---

[2] Search Engine Watch, Organic vs. Paid Search Results: Organic Wins 94% of Time, available at https://searchenginewatch.com/sew/news/2200730/organic-vs-paid-search-results-organic-wins-94-of-time.

User data in search and search advertising

The Preliminary Report maintains that, while search users do not *directly* derive value from other users utilizing the same search platform, they *indirectly* benefit because the collection and utilization of user data enables Google to offer higher-quality search services to users. According to the Preliminary Report, these "[s]ame-side network effects arising from data accumulation" protect Google's competitive position in search and search advertising. As discussed below, these benefits are not "network effects" but conventional supply-side scale effects. This distinction is not merely semantic, but goes directly towards (and undercuts) concerns that search is characterized by unique or distinctive entry barriers.

Google Search primarily relies on three types of user data: (1) the search query inputted by the individual user, (2) aggregate data on queries that users have entered and links they have clicked on in the past ("click-and-query" data), and (3) the location of the individual user. Unlike social networking sites that personalize their services to the individual user, Google's search results are the same or very similar for users in the same general location. References to user data in the context of discussing Google Search below refer to aggregate click-and-query data.

User data is a valuable input into the provision of many online services (including search), both in improving the quality of services offered to users and in monetizing those services effectively by targeting ads, which facilitates the provision of valuable services to users at subsidized prices, often for free. Because of the value of this data, firms of all types and sizes collect data to improve their products and services, and to monetize those services through targeted ads. These benefits of user data, however, do not mean that having access to a significant volume of user data gives large search platforms a significant competitive advantage, either in terms of providing higher quality services to users or better targeted ads to advertisers.

User data to improve search quality: User data is an important input into providing online services, including search (as well as many bricks and mortar providers, from airlines to credit card providers). But the history of the search industry shows that innovation and differentiation have driven competitive success.

That is so for several reasons. To begin with, there are also many other inputs that go into providing search services. A key determinant of search quality is the amount and caliber of engineering resources devoted to improving a search platform's algorithms. Another important input for search providers is the web crawling and indexing technology (for instance, it is estimated that Google's web index is many times the size of Bing's). Google also employs a large number of human raters (over 10,000 worldwide) to evaluate the quality and relevance of its search results, and uses those evaluations to improve its algorithms. There are also many other types and sources of data that are valuable inputs into the provision of search services. Both general and specialized search providers, for example, acquire from third parties "data feeds" that are used to answer particular user queries, including for consumer products, travel, sports scores, movie times, stock performance, weather forecasts, and currency exchange rates. These data feeds are not based on data collected from users. The fact that there are many other important inputs means that a provider without a large volume of user data can invest, for example, in innovation, engineering, web crawling, and/or other types of data to develop high-quality search services and, if successful, collect data from users.

Moreover, the value of user data is subject to diminishing returns, which means that the benefits from additional user data become progressively smaller as the provider's scale increases, and may disappear at some limited scale. Returns to user data scale diminish rapidly for queries entered frequently by users (often referred to as "head" queries), which include "navigational" queries (such as "Facebook"

or "Amazon"), popular products, personalities, or notable events. A large volume of user data is not necessary to provide search results pursuant to such queries. There can be value to having a greater amount of user data for some uncommon (or "tail") queries (which include, for example, misspelled queries, addresses, specific product descriptions or model numbers, and detailed queries composed of multiple terms). However, an increase in query volume has not reduced the portion of Google's traffic that is comprised of queries that it has never seen before (*i.e.*, for which Google has no data on what users have clicked on in response to the same search query), which has remained at about 15% since the early days of Google. As a result, providing relevant results in response to tail queries involves clever engineering and the use of other types of data. Empirical studies confirm that the value of user data in returning relevant search results is subject to diminishing returns—in particular, that the addition or reduction of large amounts of user data do not materially change the accuracy of search results, as described in more detail below.

Third, the collection of user data by incumbent online providers does not preclude smaller rivals from also collecting (and/or purchasing) the same or very similar data. Unlike most inputs, user data is non-rivalrous, and users commonly multi-home, which means that several online providers can collect data from the same users. As a result, many providers have access to large amounts of user data, and no one firm "controls" all, most, or even a significant share of user data. In search, various general and specialized search providers have access to a substantial volume of data on user activities. In fact, specialized data on a specific search category (*e.g.,* shopping or travel) is often more valuable than having lots of general user data, such as that collected by general search providers. For instance, as the largest product search provider in the world, Amazon has access to valuable data on users' product purchases.

For these reasons, the history of the search sector, both general and specialized, shows innovation and differentiation are key factors in search competition. In the early days of the online search industry, Yahoo!, Lycos, and AltaVista were all thought to be dominant search providers, but lost their competitive positions. Google did not become the most popular search platform as a result of access to large amounts of user data—in fact, it had no user data when it started—but through innovation (initially, its PageRank algorithm). And, innovation continues to drive competition in search. Google continues to innovate, and invest significant resources into research and development. Google's innovations (such as Voice Search, quick answers to common questions, Universal Search, Knowledge Graph, language translator, and RankBrain) are driven primarily by engineering resources and other investments, not access to a large scale of user data.

Another key aspect of competition in search has been driven through specialization in providing search services focused on specific areas of user activity, such as shopping (*e.g.*, eBay, Gumtree), travel (*e.g.*, TripAdvisor), restaurant recommendations (*e.g.*, Zomato), and real estate (*e.g.*, realestate.com.au, domain.com.au). Amazon, for example, has become a common "starting point" for online shopping activities (with roughly half of online shoppers starting their product search on Amazon in the U.S., and quickly ranking amongst the top Australian shopping sites). The search categories in which specialized providers compete are the most profitable, corresponding to the vast majority of ad revenues for Google. Specialized search providers can focus their efforts and investments on providing services in highly monetizable areas, as opposed to making investments in answering all types of user queries, many of which do not monetize at all.

User data for ad targeting: The Preliminary Report also theorizes that cross-side network effects for advertisers arise due to the collection and utilization of user data to improve ad targeting, which makes the platform more attractive to advertisers. However, the utilization of user data for purposes of ad targeting does not lead to barriers to entry and expansion or give Google a significant competitive

advantage. To a large extent, advertisers target their own search ads by selecting keywords, bids and budgets, and other criteria (such as geography). With respect to user data, the query entered by the individual user is the most valuable piece of information in targeting search ads, and targeting ads to a specific query does not require a large volume of user data. For instance, when a user enters a query such as "buy TV," that query reflects key information about the user for purposes of targeting ads. Search queries are reflective of a user's interests and intent, which is the essence of ad targeting. Advertiser targeting through selection of keywords and other criteria, and the user query, are the primary components of targeting search ads. Moreover, there are economic factors, such as the phenomenon of "congestion," discussed below, that lead to *dis*economies of scale in an online provider's ability to monetize its services.

Display ad intermediation

The Preliminary Report does not find that Google has market power in the supply of display ad intermediation services, but contends that it enjoys "strong advantages" in the provision of such services. As an initial matter, the Preliminary Report notes that display ad intermediation makes up a small share of Google's ad revenues, with less than 10% of Google's total ad revenue in Australia coming from display ad intermediation on third-party sites.

User data in display ad intermediation: Although the Preliminary Report acknowledges that data used in display ad targeting is collected by "numerous parties," it claims that Google's data is "extensive" and "likely to be unparalleled." However, for several types of display ads, user data does not play a significant role in targeting. For example, many display ads (*e.g.,* banner ads on a publisher home page) are sold *directly* by publishers to advertisers, and therefore do not rely on ad tech firms or ad platforms for targeting. In Australia, 37% of digital ads are estimated to be sold directly by publishers (and, because these ads are usually more valuable, they likely account for a higher share of ad revenues).[3] Similarly, many display ads are targeted *contextually*, meaning that they are primarily targeted based on the content of specific pages on a publisher's website (*e.g.*, the Automobile section of an online news publisher site) and are not individualized to specific users or targeted based on user data to a significant degree (other than, perhaps, by location). Even where targeting of display ads is based on user information, the volume of data collected does not create network effects or barriers to entry. For example, a common and popular type of display ad targeting is *remarketing*, which targets users who have had prior interaction with an advertiser's products or services. Remarketing is primarily based on being able to track an *individual user's* browsing or mobile app activities via widely available tracking technologies such as cookies and mobile advertising IDs. There are many successful players offering remarketing ads, including Criteo, Facebook, Amazon and AdRoll.

Moreover, many ad tech firms have access to a wide variety of user data to utilize in targeting display ads. The primary source of data that Google uses for targeting display ads on third-party sites or apps is user interaction with sites or apps collected through tags, pixels, and other signals, and then analyzed and aggregated for targeting purposes using cookies or other identifiers like mobile advertising IDs. But virtually all online service providers, and many third parties (including various ad tech firms) collect information on users' browsing and other activities and utilize cookies, mobile advertising IDs, and/or other tracking technologies to target ads. On average, around 20 third-party tracking firms place cookies on top websites in addition to the publisher, and some popular sites may have over 100 cookies.[4]

---

[3] "Two-Thirds of Australia's Digital Ads Bought Through Programmatic or Ad Networks in 2016," Nielsen, 3 February 2017, available at https://www.nielsen.com/au/en/insights/news/2017/australia-digital-ads-in-2016.html.

[4] Steven Englehardt and Arvind Narayanan, *Online Tracking: A 1-million-site Measurement and Analysis*, Princeton University, January 2016, available at

Many providers also collect other types of user data, such as locational information from mobile devices, "first-party" data collected when users visit a webpage (including basic user information as well as "behavioral" data), and "third-party" data provided by ad tech firms such as data management platforms (DMPs) and data brokers (including social, demographic, and behavioral data). Many ad tech firms combine these various types of user data in order to target display ads. Thus, it is unlikely that Google has "strong advantages" by virtue of having "unparalleled" access to user data.

Advertiser fixed costs and network effects: The Preliminary Report suggests that Google may derive a competitive advantage in display ad intermediation from its competitive position in search advertising and the existence of advertiser fixed costs, which makes the Google ad platform a "one-stop shop" for advertisers. However, as mentioned above, platform-specific fixed costs for advertisers of setting up and managing advertising campaigns are likely to be minimal. This is particularly true given the wide availability and use of software tools to manage advertising efforts across multiple platforms and channels (*e.g.*, Kenshoo, Marin, SEMrush, Centro Basis, and Adobe Advertising Cloud). Consistent with the lack of significant fixed costs, advertisers often multi-home—*i.e.*, use multiple ad intermediaries at each level of the ad tech stack. It is estimated that advertisers use on average four demand-side platforms (DSPs), and advertising agencies may use even more. Evidence of extensive multi-homing by advertisers across different ad intermediation providers undermines any theory that fixed costs for advertisers or network effects give Google "strong advantages" in the supply of ad intermediation services. This conclusion is supported by the fact that the display ad business is highly competitive, and Google competes with a large number of providers at each level of the ad tech stack.

II. Network Effects in the Provision of Search and Search Advertising

The Preliminary Report contends that "at least to a substantial extent, Google is insulated from dynamic competition by barriers to entry and expansion for search platforms…"[5] These barriers to entry and expansion are purported to arise from network effects, including "same-side" network effects and "cross-side" network effects between users and advertisers.[6] This section describes some economic concepts regarding the relevance of network effects to search and search advertising. It also explains why conclusions regarding network effects leading to barriers to entry and expansion, and to the entrenchment of a dominant platform, cannot be based on theoretical presumptions but rather must be based on a fact-specific inquiry of the relevant marketplace.

A. The economics of network effects

A product or service exhibits "network effects" when the value of the service to a user rises with the number of other users of that service. Because there are various types of network effects, and each

http://randomwalker.info/publications/OpenWPM_1_million_site_tracking_measurement.pdf; Timothy Libert and Rasmus Kleis Nielsen, "Third-Party Web Content on EU News Sites: Potential Challenges and Paths to Privacy Improvement," Reuters Institute and University of Oxford, May 2018, available at https://reutersinstitute.politics.ox.ac.uk/our-research/third-party-web-content-eu-news-sites.

[5] Preliminary Report at 35, 42.

[6] Preliminary Report at 42-43, 54, 57-58.

type can have distinct implications for competition, it is useful to begin with a brief description of the economics of network effects and their potential impact on competitive dynamics.[7]

Same-side (or direct) network effects refer to the economic phenomenon whereby members of a user group derive value from additional members of the *same* user group on a platform. The classic example is a telephone network—the greater the number of users that can be reached by telephone, the greater the value of the telephone network to each user. As the Preliminary Report describes, "[i]f there are only a few users of a telephone network, that network will be of relatively low value to any given user, and may, therefore, have difficulty attracting new users. Conversely, a network with many users will be of relatively high value to a user and thus, be attractive to new users."[8] Similarly, the greater the number of users on a social network, all else equal, the greater the value of that network to each user.

Cross-side (or indirect) network effects arise when members of one user group on the platform (say group *A*) derive value from the participation of members of *another* user group (group *B*). One example of a platform with cross-side network effects are shopping malls or online marketplaces, which bring together merchants and shoppers. All else equal, the greater the variety of merchants in the shopping mall or online marketplace, the greater the value to shoppers; and, greater consumer traffic increases the value to merchants participating in the shopping mall or online marketplace. Another example is ride-sharing services such as Uber, Taxify, or Ola. As more riders join, the more valuable the platform becomes to drivers because they have greater business opportunities; and, as more drivers join, the platform becomes more valuable to riders because they may have shorter wait times and more available ride locations. Yet another example are payment card networks, which are demanded by two distinct, inter-related groups of consumers—cardholders (who use the cards to make purchases) and merchants (who accept the cards as payment for purchases). All else equal, the greater the number of merchants that accept a payment card, the greater the value to cardholders; and, the greater the number of cardholders that carry the card, the greater the value to merchants.[9]

The potential competitive concern regarding network effects is that they can give rise to a positive "feedback loop," a self-reinforcing phenomenon where the large become larger and smaller providers and new entrants cannot catch up and compete effectively. For instance, there could be incentives for members of group *A* (*e.g.*, Uber riders) to use the platform with the greatest participation by members of group *B* (Uber drivers). The concentration of members of group *B* on that platform could then create incentives for members of group *A* to use that platform also. This process might continue until all members of group *A* and group *B* converged on a single platform.

However, network effects do not necessarily lead to a self-reinforcing feedback loop or to the entrenchment of a dominant platform. The competitive implications of network effects depend on the particular marketplace circumstances. In fact, in many if not most markets characterized by network effects, multiple platforms can and do compete.[10] For instance, Uber, Taxify and Ola (among others) compete vigorously in offering ride-sharing services; AirBnB competes with VRBO, and TripAdvisor's

---

[7] See Michael Katz & Carl Shapiro, *Systems Competition and Network Effects*, 8 J. ECON. PERS. 93 (1994); S. J. Liebowitz & Stephen E. Margolis, *Network Externality: An Uncommon Tragedy*, 8 J. ECON. PERS. 133 (1994).

[8] Preliminary Report at 43.

[9] See, *e.g.*, Benjamin Klein, Andres V. Lerner, Kevin M. Murphy & Lacey L. Plache, *Competition In Two-Sided Markets: The Antitrust Economics Of Payment Card Interchange Fees*, 73 ANTITRUST L.J. 571 (2006).

[10] Michael Katz & Carl Shapiro, *Systems Competition and Network Effects,* 8 J. ECON. PERS. 93, 106 (1994).

FlipKey (among others) in the vacation rental services sector;[11] and multiple payment card systems compete for merchant acceptance and bank issuance. There are several economic factors that limit the competitive impact of network effects in practice.

Diminishing returns: while the value of a network to users may increase with the number of other users on the network, the incremental value of more users may weaken or even disappear at some scale. For instance, a shopping mall may be more valuable to shoppers if there are more stores in it, but once a mall has a certain number of retailers in it, and covers most of the products and services that people typically look for in a mall, the appeal of adding more retailers may have limited incremental value to most shoppers. The diminishing value of additional retailers after some point limits the advantages of scale for a shopping mall. Network effects are less likely to lead to the marketplace tipping to a dominant platform when they diminish at fairly low scale.

Platform differentiation: the propensity of network effects to lead to the entrenchment of a dominant platform also is reduced by platform differentiation. Although some users may value a larger platform, smaller platforms and new entrants may be able to compete effectively by offering services that are preferred by some segment of consumers.[12] For example, Snapchat was able to differentiate itself from Facebook by creating a product aimed at a younger demographic, and in doing so was able to attract advertisers seeking to target that younger demographic.

Unidirectional network effects: network effects may be *unidirectional*, whereby one group in a multi-sided platform (group *A*) may value the participation of another group (group *B*), but group *B* may not value the presence of group *A*. This is true for many (if not most) ad-supported platforms. For instance, although newspaper advertisers value the ability to place ads that are seen by readers, consumer demand for a newspaper is generally not driven by the availability of ads. As discussed more extensively below, the unidirectional characteristic of network effects for some platforms, including online search platforms, diminishes or eliminates any "feedback loop" that can entrench a dominant platform.

Repulsion: relatedly, multi-sided platforms may be characterized by *repulsion*, whereby members of one group (group *A*) may value the participation of members of another group (group *B*), but participation by group *B* may lower the value of the platform to members of group *A*. Ad-supported platforms can exhibit this pattern of attraction and repulsion. While advertisers want to be where users are, users often want to avoid advertisers. This phenomenon helps explain why consumers are willing to pay for ad-free subscription services despite the availability of similar free, ad-funded services.

Congestion: multi-sided platforms also may be subject to *congestion*, wherein the presence of additional members one group (group *A*) makes that platform less valuable to members of the same group (group *A*). This is especially likely where participants on one side of the platform may compete with one another for the attention and patronage of users on the other side of the platform. Online marketplaces provide a good example. All else equal, the attractiveness of the marketplace to individual merchants tends to decrease with participation by more merchants because there is more competition to attract consumers. Similar effects can arise in the case of ad-supported platforms. From an advertiser's perspective, a greater number of advertisers may make that platform a less desirable place to advertise.

---

[11] Pamela Chan, "10 Airbnb Competitors That You Should Know About," Tripping.com, 1 March 2018, available at https://www.tripping.com/industry/rental-companies/9-airbnb-competitors-that-you-should-know-about.

[12] See, *e.g.*, Michael L. Katz & Carl Shapiro, *Systems Competition and Network Effects*, 8(2) J. ECON. PERS. 93, 106 (Spring, 1994): "[c]onsumer heterogeneity and product differentiation tend to limit tipping and sustain multiple networks. If the rival systems have distinct features sought by certain customers, two or more systems may be able to survive by catering to consumers who care more about product attributes than network size."

While advertisers want to be where users are, advertisers do not want to be where their rivals are because advertisers have to compete for users' attention, which may limit the effectiveness and/or raise the cost of ads. In addition, there are limited spaces available for search ads, and bidding competition between advertisers for scarce slots raises the cost of search ads.[13] Because of congestion, advertisers may prefer a small platform with fewer users, but also fewer rival advertisers.

Multi-homing: the competitive implications of network effects also are diminished due to *multi-homing* by platform participants. Multi-homing refers to the practice by users of participating on two or more platforms. For instance, retailers may locate in different shopping malls; users may participate in different dating, social, or messaging networks; video game developers produce games for various game consoles. Users can benefit from multi-homing when different networks offer different features, functions, or quality levels. Multi-homing diminishes the significance of cross-platform network effects because members of one group (group *A*) can access members of another group (group *B*) through two or more platforms, diminishing or eliminating the propensity for both groups to converge on a single platform. Moreover, the risks associated with trying a new platform are reduced by the possibility of multi-homing because users do not have to give up the value of a large platform in order to participate in a smaller one.[14] In contrast, if multi-homing costs were high, users could be confronted with an all-or-nothing choice between a large, established platform, and a small, nascent one, which could make it difficult for a smaller platform to get a foothold in the marketplace. Thus, user multi-homing facilitates entry and makes the coexistence of multiple competitors more likely.[15]

Because of these limitations to the competitive impact of network effects, it cannot be presumed that the existence of network effects is likely to lead to significant feedback loops, barriers to entry and expansion, or to the entrenchment of a dominant platform.

B. Network effects in search and search advertising

1. Search is not characterized by same-side network effects

The provision of online search services is not characterized by the standard same-side network effects among users. Search users do not place direct value on whether other users utilize the same search provider. The provision of search services is fundamentally distinct from platforms such as telephone or social networks, wherein users directly derive value from being able to interact with other users on the same network. Consistent with this premise, the Preliminary Report does not conclude that search users directly value the presence of other users on the search platform (as opposed to Facebook, which the Preliminary Report finds benefits from same-side network effects because users directly value that presence of other users on the social network).[16] However, the Preliminary Report maintains that Google's competitive position is protected by a different type of same-side network effects—"[s]ame-side network effects arising from data accumulation."[17] According to the Preliminary Report, Google collects

[13] Dakota Shane, "Facebook Ads Aren't Always the Answer. Here's What to Do Instead," Inc., 21 June 2018, available at https://www.inc.com/dakota-shane/why-its-time-to-start-advertising-on-platforms-other-than-facebook.html.

[14] See, *e.g.*, Carl Shapiro, *Exclusivity in Network Industries*, 7 GEORGE MASON LAW REV. 673, 677 (1999).

[15] See, *e.g.*, Aaron S. Edlin & Robert G. Harris, *The Role of Switching Costs in Antitrust Analysis: A Comparison of Microsoft and Google*, 15 YALE J. L. TECH. 169 (2013).

[16] Preliminary Report at 51.

[17] Preliminary Report at 42.

"a considerable quantity of data on its users and their use of the search platform" which enables Google to offer higher-quality services to users.[18]

However, the "relationship between the quantity of data and the quality of the search service" is not a network effect. Network effects are a *demand-side* phenomenon. A provider's product or service becomes more valuable as the number of users increases—that is, demand increases with user scale because users *directly* derive value from other users on the platform. For instance, users on a social network derive value from other users on the network, and demand for the network derives largely from enabling social connections between users. The utilization of user data to improve search quality, on the other hand, is a classic *supply-side* efficiency—user data is an input that allows a provider to increase the quality of the services it offers. Similarly, in classic supply-side economies of scale, average costs decrease with scale, and consumers may *indirectly* benefit from other consumers that use the product or service to the extent that greater scale leads to lower costs (and thus to potentially lower prices). But the fact that consumers may benefit from larger scale does not mean that such a product or service is subject to network effects.

The distinction between *demand-side* network effects and *supply-side* economies of scale is not only definitional or semantic. Because network effects imply that more users directly derive value from the participation of other users, they can lead to a "chicken-and-egg" problem, where users value a platform only if a sufficient number of other users also are on the platform. In some circumstances, this can make growth difficult for new entrants and nascent firms. For instance, it may be difficult for a nascent social network to attract users unless the platform already has a sufficient scale of users. In contrast, the fact that the utilization of user data to improve quality operates indirectly through supply-side efficiencies means that smaller rivals and new entrants can overcome a lack of user data scale through innovation and investments, and through using substitute inputs (such as engineering resources) to achieve a given level of quality, as discussed in Section III.B below.

2. Search is characterized by (at best) weak cross-side network effects

The Preliminary Report also contends that "[s]earch platforms and social media platforms are examples of multi-sided platforms characterised by cross-side network effects,"[19] which "may give rise to barriers to entry or expansion" and "ensure that Google has a competitive advantage in the supply of search advertising over a small-scale search platform."[20] However, as mentioned above and discussed more extensively in this section, cross-side network effects are limited (at best) for search platforms, and do not lead to significant barriers to entry or expansion.

a) Search users

According to the Preliminary Report, "an increase in the number of advertisers may increase the value of the platform to a user" because "the platform is able to serve ads that are more relevant to that user. For at least some users, being shown more relevant ads (as opposed to generic ads) improves the user experience."[21] However, many (if not most) users ignore paid results and/or prefer to be on a site

[18] Preliminary Report at 38, 42.

[19] Preliminary Report at 40.

[20] Preliminary Report at 43.

[21] Preliminary Report at 40.

with less ads. In fact, some users may place negative value on having too many ads on a search results page (an example of a phenomenon called *repulsion*, as discussed above).[22]

The theory put forth in the Preliminary Report seems to be that, although search users may not value ads, search users prefer ads that are well targeted to ads that are poorly targeted. Under this argument, additional advertisers allow a search engine to target ads better which, in turn, increases value to search users. While this will be correct for some users, in general user demand for a platform is not driven by the availability of ads, even targeted ads. One reason why demand for a platform is not driven by ads is that organic search results are often superior to ads from the users' perspective, even when those ads are targeted. This is suggested by the fact that ads make up only about 6% of user clicks on Google (even when those ads are targeted), with the other 94% going to organic (free) results.[23] Moreover, some users may place particular negative value on more targeted ads.[24]

The lack of strong network effects from advertisers to users weakens or invalidates any "feedback loop." The fact that cross-side network effects are essentially unidirectional (advertisers seek users but users do not seek advertising) fundamentally weakens or eliminates the possibility of a feedback loop that locks users and advertisers into a dominant platform. If a smaller rival or new entrant offers a better service to users, cross-side network effects do not inhibit users from switching and, once users switch, advertisers likely would follow. Simply, users do not continue patronizing a platform with inferior services for the access to ads, irrespective of whether those ads are well targeted.[25]

b) Search advertisers

Even for search advertisers, cross-side network effects are limited. From an advertiser's perspective, the value of an online platform is that it allows the advertiser to reach users. However, the competitive implications of cross-side network effects are fundamentally altered by the fact that most online providers price on a cost-per-click or cost-per-impression basis. Google prices its search ads on a cost-per-click basis, whereby an advertiser pays only when a user clicks on the advertiser's ad.[26] One implication of this pricing structure is that, while an advertiser may derive more value in advertising on a platform with more users because the advertiser's ads may obtain more user clicks, advertising on such a platform entails proportionately higher costs. For instance, all else equal, if a platform has 10 times as many users and an advertiser receives 10 times as many clicks, the advertiser will derive 10 times more value but also will pay 10 times more. Because of per-click pricing, an advertiser pays the same amount per user click whether the advertiser divides its campaign across multiple platforms or advertises on only

---

[22] For instance, a recent study by Pandora shows that having too many ads may encourage people to purchase an ad-free subscription and pushes many users to listen less or abandon the service. (Erin Griffith, "Pandora Learns the Cost of Ads, And of Subscriptions," Wired, 30 April 2018, available at https://www.wired.com/story/pandora-learns-the-cost-of-ads-and-of-subscriptions/.)

[23] Search Engine Watch, Organic vs. Paid Search Results: Organic Wins 94% of Time, available at https://searchenginewatch.com/sew/news/2200730/organic-vs-paid-search-results-organic-wins-94-of-time.

[24] The Preliminary Report posits that for some users, "the serving of targeted ads could decrease their user experience due to privacy concerns." (Preliminary Report at 40.)

[25] The Preliminary Report seems to recognize that a feedback loop requires that network effects operate in both directions—*i.e.*, that advertisers value more users, and that users value more advertisers. (Preliminary Report at 43.)

[26] Google Ads are generally priced on a cost-per-click (CPC) basis for search ads; and cost-per-thousand-impressions (CPM), or cost-per-action (CPA) for display ads.

one.[27] As a result, advertising on a larger platform is not necessarily more profitable than advertising on a smaller one. The cost to the advertiser per user click may be higher or lower on a large platform compared to a smaller platform.[28] Programmatic media mix optimisation software packages (discussed below) arbitrage across different platforms and help the advertiser allocate its budget based on the relative value and cost across platforms.

Fixed costs for advertisers: the Preliminary Report seems to recognize that the mere fact that advertisers can access more users on a larger platform, in and of itself, does not give the larger platform a competitive advantage over smaller ones.[29] It postulates, however, that cross-side network effects for advertisers arise because an advertiser may incur fixed set-up costs from using a particular platform and running a particular campaign. As a result, "if there are more users on a platform, an advertiser and a campaign obtain more traffic, which in turns [sic] reduces the average fixed costs," which gives a large platform like Google a competitive advantage.[30]

The Preliminary Report does not describe what these fixed costs are, or give any indication regarding the magnitude of such fixed costs. The sole citation in the Preliminary Report for the proposition that advertisers face significant platform-specific fixed costs is an article by David Evans (2011), which notes that "there are costs of setting up the platform, installing software, and learning how to use it."[31] However, the "activation fees" Evans pointed to in 2011 were generally trivial, and in any case no longer appear to be applicable.[32] Google and Bing, for example, currently have zero setup fees for advertisers using their search ad platform.[33] Evans theorizes that advertisers incur time and effort in setting up and monitoring particular campaigns, including "choosing keywords," "modifying bids," and "monitoring account statistics." But platform-specific fixed costs for advertisers to set up and run campaigns on a particular search platform such as Google are small. A key reason is that advertisers can port data across different search platforms easily, in part due to compatible interfaces and tools offered by search providers that make it easy to import advertising campaigns. For example, Bing introduced

---

[27] See, *e.g.*, Mark Armstrong, *Competition in 2-Sided Markets*, 37 RAND J. ECON 668, 669 (2006): "Cross-group externalities are weaker with per-transaction charges, since a fraction of the benefit of interacting with an extra agent on the other side is eroded by the extra payment incurred. If an agent has to pay a platform only in the event of a successful interaction, then that agent does not need to worry about how well that platform will do in its dealings with the other side. That is to say, to attract one side of the market, it is not so important that the platform first gets the other side 'on board.'"

[28] Search engines such as Google also allow advertisers to specify a maximum price per-click they are willing to pay, which diminishes concerns regarding not receiving value from the smaller platform.

[29] Preliminary Report at 40.

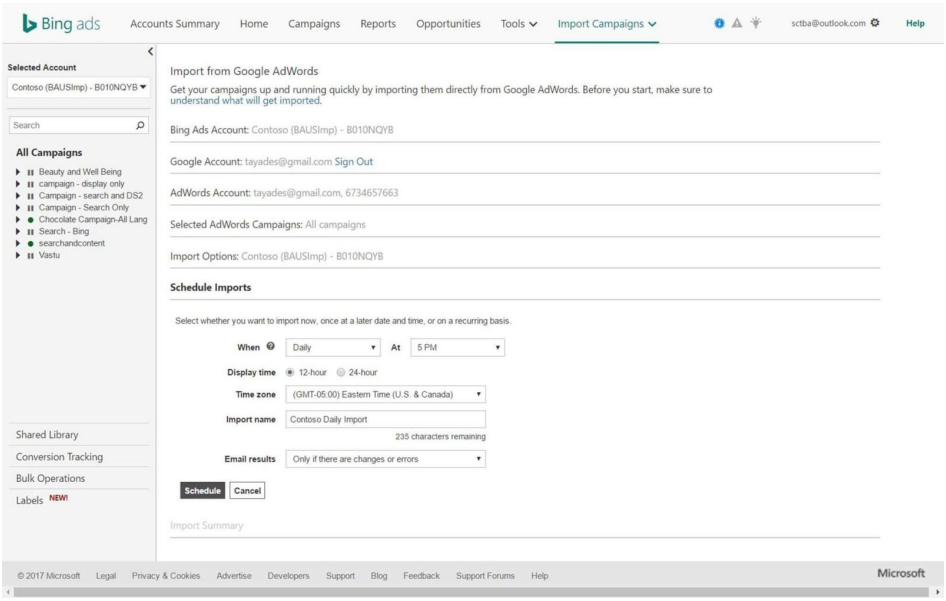
[30] Preliminary Report at 40.

[31] David Evans, *Economics of the Online Advertising Industry, in Platform Economics: Essays on Multi-sided Businesses*, December 2011 at 209.

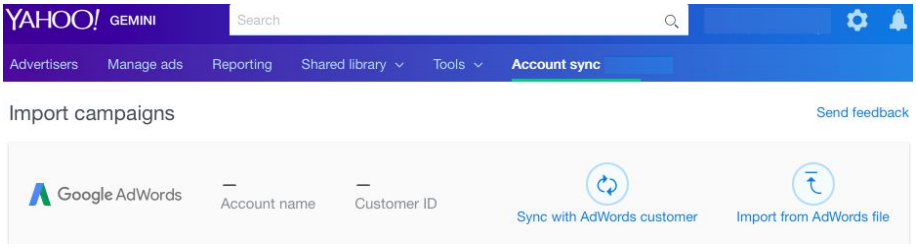
[32] Evans claims that "Google and Microsoft each charge $5; Yahoo has no fee for the 'Self Serve' version, but charges $199 for an assisted setup." (Evans (2011) at 209.)

[33] "Creating a Bing Ads account is free and easy to do with no setup costs or maintenance fees." ("Cost of Bing Ads," Bing Ads, available at https://advertise.bingads.microsoft.com/en-us/get-started/small-business/cost-of-bing-ads); "Only pay when people take action like clicking your ad…" ("Pricing," Google Ads, available at https://ads.google.com/home/pricing/.) "AdWords Express has no activation fee." ("Costs and payment," AdWords Express Help, available at https://support.google.com/adwords/express/answer/1184980?hl=en.)

Automated Imports in 2017, allowing advertisers to easily transfer existing Google Ads campaigns directly into Bing Ads.[34]



Similarly, Oath/Yahoo!'s ad platform allows an advertiser to "sync" their Google Ads account with their Yahoo! Gemini account, and then import Google Ads campaign data using a "super easy and straightforward" process.[35]

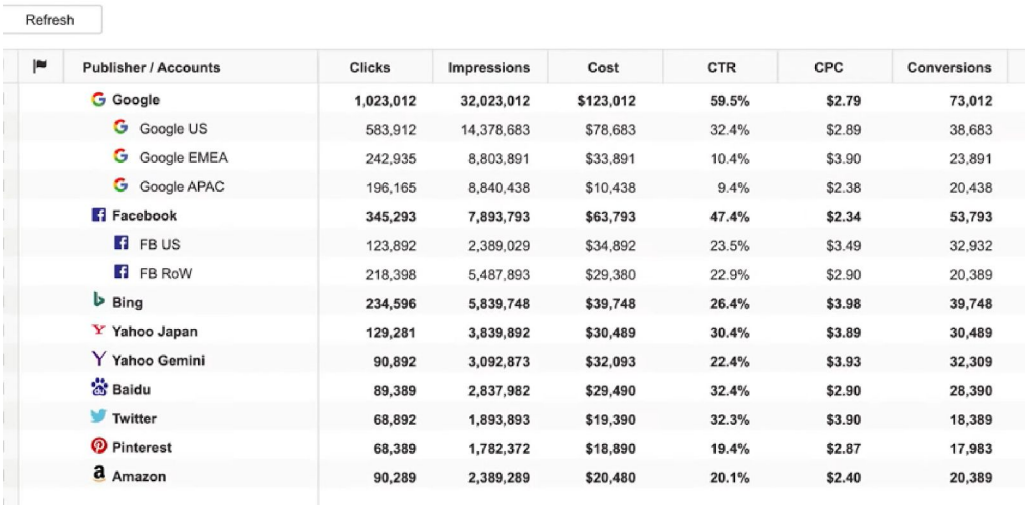


The widespread availability and use of software tools that help advertisers manage and seamlessly port ad campaigns (including search ads) across different search providers and advertising platforms also reduce or eliminate any platform-specific fixed costs. These software tools lower the costs of setting up ad campaign on each ad platform and/or monitoring the campaigns. For instance, Marin's software platform allows advertisers to manage ad campaigns across Google, Bing, Yahoo!, and Baidu (among many others). Marin's "cross-publisher sync" allows advertisers to "sync[] campaigns on Google,

[34] Ginny Marvin, "Bing Ads launches Automated Imports to automatically sync AdWords campaigns," Search Engine Land, 24 January 2017, available at https://searchengineland.com/bing-ads-automated-imports-adwords-campaigns-267996.

[35] "How To Import from Google Ads," Oath Ad Platforms, available at https://developer.yahoo.com/nativeandsearch/advertiser/guide/howtos/how-to-import-adwords/?guccounter=1; "Import Campaigns from Google Ads," Oath Ad Platforms, available at https://developer.yahoo.com/nativeandsearch/advertiser/guide/howtos/how-to-import-campaigns-adwords/?guccounter=1.

Facebook, and other publishers, driving incremental volume with minimal effort."[36] Marin's user interface is shown below.


Refresh

Publisher / Accounts	Clicks	Impressions	Cost	CTR	CPC	Conversions
Google	1,023,012	32,023,012	$123,012	59.5%	$2.79	73,012
Google US	583,912	14,378,683	$78,683	32.4%	$2.89	38,683
Google EMEA	242,935	8,803,891	$33,891	10.4%	$3.90	23,891
Google APAC	196,165	8,840,438	$10,438	9.4%	$2.38	20,438
Facebook	345,293	7,893,793	$63,793	47.4%	$2.34	53,793
FB US	123,892	2,389,029	$34,892	23.5%	$3.49	32,932
FB RoW	218,398	5,487,893	$29,380	22.9%	$2.90	20,389
Bing	234,596	5,839,748	$39,748	26.4%	$3.98	39,748
Yahoo Japan	129,281	3,839,892	$30,489	30.4%	$3.89	30,489
Yahoo Gemini	90,892	3,092,873	$32,093	22.4%	$3.93	32,309
Baidu	89,389	2,837,982	$29,490	32.4%	$2.90	28,390
Twitter	68,892	1,893,893	$19,390	32.3%	$3.90	18,389
Pinterest	68,389	1,782,372	$18,890	19.4%	$2.87	17,983
Amazon	90,289	2,389,289	$20,480	20.1%	$2.40	20,389

Other examples of ad management software tools include Kenshoo (which can help advertisers manage search ad campaigns on Google, Bing, Baidu, and other platforms),[37] Adobe Advertising Cloud (formerly Adobe Media Optimizer), Google's Search Ads 360, and SEMrush (a platform specializing in managing search ad campaigns).[38]

The role of user data to improve ad targeting: the Preliminary Report also maintains that cross-side network effects for advertisers arise because a "platform with more users has access to more data which can improve the relevance of ads presented to users. All else equal, an advertiser may prefer a larger platform, because its ads will tend to be more targeted."[39] However, analogous to the utilization of user data to improve the quality of search results, the utilization of user data for purposes of ad targeting is a supply-side efficiency, not a network effect. Equivalent to classic supply-side economies of scale, wherein average costs decrease with scale, the utilization of user data to improve ad targeting implies that quality (for advertisers) increases with scale. But, as discussed below, the role of user data scale in targeting search ads is limited because advertisers largely target their own search ads by selecting keywords, bids and budgets, and other criteria (such as geography). With respect to user data, the user query itself is the most valuable piece of information in targeting search ads. Targeting ads to a specific query is based on aggregate user click-and-query data, not data on the individual users (other than location), and does not require a large volume of user data.

User single-homing: the Preliminary Report points to another effect that it argues compels advertisers to use Google—that "a large fraction of Australian users of general search services effectively 'single home' on Google" and thus to reach these users advertisers must use Google's advertising services.[40] However, since a user typically uses multiple online services, websites, and mobile

[36] "Marin One," Marin Software, available at https://www.marinsoftware.com/.

[37] "Search Marketing Magnified," Kenshoo, available at https://www.kenshoo.com.au/products/kenshoo-search/.

[38] "Organic Research," SEMrush, available at https://www.semrush.com/features/organic-research/.

[39] Preliminary Report at 40. See also Preliminary Report at 57.

[40] Preliminary Report at 57.

applications on a given day, advertisers have many different ways to target that user beyond Google's search ads, even if that user only relies on Google for search service (which, as discussed below, is not usually the case, since there are different ways that most users search for information online).

The Preliminary Report states that Google has a 94% search "market share" in Australia, and claims that from this "it follows that most Australian users of general search almost always use Google, and thus effectively single home."[41] However, the market shares cited in the Preliminary Report are based on pageviews, not users.[42] Thus it cannot be inferred that most users single-home. For instance, if every user conducted 94 searches on Google and six searches on Bing or Baidu, Google's share of pageviews (*i*.e., search results clicked on) would be 94%, but no users would single-home. Moreover, the 94% estimate pertains to only *general* search providers, and does not include *specialized* search services. As discussed in Section III.D., most users do multi-home across different search providers in conducting online activities. As a result, advertisers can reach the same user (and even users conducting the same type of activity) by advertising on different online platforms, including on general and specialized search platforms. In fact, the Preliminary Report finds that advertising on a specialized search platform is a substitute to placing ads on a general search platform from the perspective of advertisers, defining a relevant market for search advertising that includes ads on both specialized and general search services.[43]

Moreover, even if a large share of Australian users "single-homed" on Google search, this would not create network effects or a self-reinforcing feedback loop that can entrench a dominant platform. Assuming for purposes of illustration that only a limited share of users utilized the services of Google's rivals (*e.g*., 10%), nothing would inhibit advertisers from allocating a proportionate share of their budget (10%) to rival platforms. And, the fact that advertisers can multi-home (and have a variety of tools to facilitate multi-homing, as discussed above) would prevent the market from "tipping" to a platform that has a large share of users.[44]

c) Rivals' ability and incentive to invest

The Preliminary Report also seems to hypothesize that because the collection of user data allows a search platform to more effectively target ads, it leads to greater monetization which enables the platform to make investments and thereby better compete for users. Even if rivals were not able to monetize as effectively as larger incumbents, it would not imply the existence of network effects or barriers to entry and/or expansion. Lower monetization would not impede rivals' ability to compete, either by inhibiting their *ability* or *incentive* to invest.

With regard to the financial *ability* to invest, very few tech firms fund investment out of current revenue. Investment is driven by expectations of future revenue funded either by current cash or through external fundraising. There are many online providers that possess ample financial resources to invest in improving their products and services, and many outside means through which online providers can obtain funding to grow their businesses, including venture capital firms, "angel" investors, and IPOs. More generally, the claim that barriers to entry and scale economies exist because an incumbent earns

[41] Preliminary Report at 57.

[42] See StatCounter Global Stats FAQs at http://gs.statcounter.com/faq#methodology: "For our search engine stats, we analyze every page view referred by a search engine."

[43] Preliminary Report at 57.

[44] In fact, the Preliminary Report does not seem to contend that network effects result from users purportedly "single-homing" on Google, but only notes that this works to limit the competitive constraint Google faces from other search providers. (Preliminary Report at 57.)

greater revenues and thus has more capital to invest is not based on sound economics. Absent capital market imperfections, small firms and new entrants can raise the necessary capital to expand or enter. The argument that established or larger firms can invest more because they earn more revenues would imply, nonsensically, that firms in all industries are subject to feedback loops resulting in significant entry and expansion barriers, since it would imply that larger firms earn more revenues and therefore can invest more in attracting customers.

Neither does greater scale confer greater *incentives* to invest. Consider the marginal effects of an increase in the quality of services offered to users. Because the price of most online services to users is zero, the benefits of increased quality to a platform come in the form of increased user activity, which has the potential to increase the amount of ads available for sale by the platform.[45] There is no sound economic basis for assuming that the incremental user activity due to a given quality investment is greater for a larger platform than a smaller one. For example, a smaller platform with less users may gain more users from investments in increasing the quality of its services than would a larger platform. As a general matter, a smaller platform could well have greater incentives to invest in quality improvements than a larger platform.

Thus, although user data may enable an online platform to monetize more effectively, greater monetization does not necessarily lead to greater ability or incentive to invest in improving the quality of services in order to compete for users. The absence of such "feedback" weakens or negates the competitive impact of enhanced monetization from user data, and makes it unlikely that the benefits of user data in monetizing online services would lead to barriers to entry or the entrenchment of dominant platforms.

III. User Data in Search and Search Advertising

The Preliminary Report posits that the collection and utilization of user data by online platforms protects Google's competitive position in search and search advertising. As discussed in this section, user queries and clicks are the most important data Google utilizes in the provision of search and search advertising, not personal or demographic information about users. And, although user data is a valuable input into the provision of search services, the competitive success of search platforms is not determined by how much user data they can collect. The collection of user data does not lead to entry barriers and expansion barriers, or to the entrenchment of dominant online platforms.

A. User data facilitates the provision of online services

1. Utilization of user data to improve the quality of services

a) User data in the provision of online services

User data is a valuable input into the provision of online services. Online providers utilize data collected from users to improve their services in several ways, including to improve the user experience, develop new products, and provide better user support.[46] Because of the value of this data, firms of all

[45] In the context of search, quality improvements could also reduce the volume of advertising because users may substitute clicking on unsponsored (organic) search results for clicking on ads.

[46] See, *e.g.*, Exec. Office of the President, *Big Data: Seizing Opportunities, Preserving Values*, at 48 (2014): "To a greater degree than ever before, this data is being harnessed by businesses, governments, and entrepreneurs to improve the services they deliver and enhance how people live and work."

types and sizes collect data to improve their products and services, and many firms have access to large amounts of user data.

For instance, online music providers, such as Spotify and Soundcloud, utilize usage information to provide personalized services.[47] Netflix also utilizes usage data to offer personalized content recommendations.[48] Similarly, online news publishers personalize the user experience through the utilization of user data, for example, by tailoring the types of stories, editorial content, and video news clips shown to users,[49] as well as to recommend articles to users.[50] Some online providers also personalize their services through the use of "social" data, which may include user-provided information on a customer's interests, hobbies, and lifestyle, and similar information from other users in the customer's "social circle." This type of data allows online providers to identify products, content, or services that a user is most likely to find useful.[51] (As mentioned above, unlike other types of online services that personalize their services to the individual user, Google's search results are the same or very similar for users in the same general location.)

These are just a few examples of the way that online platforms utilize user data to improve the services offered to users. The benefits created for consumers by the use of such data are widely recognized. The Preliminary Report explains that "[t]he collection and analysis of user data by digital platforms can benefit consumers, such as when user data is used to fix problems, to improve products and services, or create new products."[52]

Because of these competitive benefits, essentially all online providers track user activity on their sites and collect demographic, behavioral, and other data from users in order to improve and monetize their services. Virtually all online firms collect, process, analyze, learn from, make decisions, and provide services based on user data and other information. The collection of user data is conducted by firms of all

---

[47] Spotify's Discover Weekly feature gives users a customized playlist of new songs based on personal listening history and the listening history of other users with similar tastes. ("Discover Weekly," Spotify, available at https://support.spotify.com/us/using_spotify/playlists/discover-weekly/.)

[48] Ashok Chandrashekar, "Artwork Personalization at Netflix," Netflix Technology Blog, 7 December 2017, available at https://medium.com/netflix-techblog/artwork-personalization-c589f074ad76.

[49] "The Boston Globe has begun using personalization technology to tailor both marketing and editorial content. Hearst Newspapers, which publishers [sic] papers including the San Francisco Chronicle, used Google's natural language processing tool to categorize its content to personalize stories for readers." (Max Willens, "News Publishers Are Giving Personalization a Fresh Look," Digiday, 22 March 2018, available at https://digiday.com/media/news-publishers-giving-personalization-fresh-look/.) Also, "companies like Iris TV and the AOL-owned Gravity[] tailor the queues of videos shown to users based on data that publishers have about their viewers." (Max Willens, "Why Publishers Are Getting More Serious About Personalization," Digiday, 10 October 2016, available at https://digiday.com/media/publishers-getting-serious-personalization/.)

[50] See, *e.g.*, Federica Cherubini and Rasmus Kleis Nielsen, *Editorial Analytics: How News Media Are Developing and Using Audience Data and Metrics*, Reuters Institute for the Study of Journalism, 2016, available at http://www.digitalnewsPreliminary Report.org/publications/2016/editorial-analytics-2016.

[51] For example, Expedia uses social data in order to provide personalized offers to its users. (Nikki Gilliland, "How six travel & hospitality brands use personalisation to enhance the customer experience," Econsultancy, 26 June 2017, available at https://econsultancy.com/how-six-travel-hospitality-brands-use-personalisation-to-enhance-the-customer-experience/.)

[52] Preliminary Report at 167.

sizes, by firms offering a wide array of online services and content to users, by firms offering online services as well as dozens of firms that provide data analytics, and by both new entrants and established players. Indeed, even the most basic sites log extensive information about each user visit, and software tools to analyze those logs and other data are available to even the smallest publishers.

The collection and use of customer data is commonplace not only for online firms, but also for traditional offline ("brick-and-mortar") businesses. Banks, credit card companies, airlines, telephone companies, retailers, and many other types of brick-and-mortar firms collect customer data to improve their businesses. Brick-and-mortar firms collect data from customers in various ways, including public records, data purchased from third-party brokers, retailer loyalty cards, credit/debit card payments, and retailer receipts. The collection of data from customers serves a similar purpose for offline firms as it does for online providers. For instance, the collection and analysis of customer data helps retailers target those who are most likely to purchase their products.[53] Retailers also use data collected from customers to improve their products, product placement, and store locations.[54] Some brick-and-mortar retailers also track customer movements with user location data accessed through a customer's Wi-Fi-connected smartphone, which automatically connects to sensors placed in stores.[55] This can, for example, yield insights similar to those obtained by online retailers that track the products users viewed or put in their online "cart" even without purchasing them.

b) User data in the provision of search

Search providers, including general search engines (*e.g.*, Google and Bing) and specialized search sites (*e.g.*, Amazon and Expedia), utilize data from users to deliver relevant search results. Google generally does not personalize search results based on demographic, personal, or other data about users. Rather, user queries (which evidence user intent) and user click data (which helps measure the quality of the results) are the most important data Google utilizes in the provision of search. By collecting and analyzing user search terms and clicks on search results (known as "click-and-query" data), search providers are able to refine and improve their services. Generally, search results that other users have clicked on in response to the same or similar queries can be a useful "signal" of the relevance of a particular search result. Google and other search providers also frequently experiment with changes to search result pages and collect data on how users interact with new or changed elements in order to assess the value of the modifications to users.[56]

---

[53] See, *e.g.*, Bernard Marr, "Big Data: A Game Changer In The Retail Sector," Forbes, 10 November 2015, available at https://www.forbes.com/sites/bernardmarr/2015/11/10/big-data-a-game-changer-in-the-retail-sector/#6c09e08b9f37. Also see "6 Big Data Use Cases in Retail," Ingram Micro Advisor, 17 June 2017, available at http://www.ingrammicroadvisor.com/data-center/6-big-data-use-cases-in-retail.

[54] See, *e.g.*, Donna Ferguson, "How supermarkets get your data – and what they do with it," The Guardian, 7 June 2013, available at http://www.theguardian.com/money/2013/jun/08/supermarkets-get-your-data.

[55] Elizabeth Dwoskin, "What Secrets Your Phone Is Sharing About You," The Wall Street Journal, 13 January 2014, available at https://www.wsj.com/articles/what-secrets-your-phone-is-sharing-about-you-1389663783.

[56] See, *e.g.*, Barry Schwartz, "Google tests sub-images in search result snippets for sitelinks," Search Engine Land, 17 September 2018, available at https://searchengineland.com/google-tests-sub-images-in-search-result-snippets-for-sitelinks-305461; Matt Southern, "Google Tests New Design for Desktop Search Results," Search Engine Journal, 23 August 2018, available at https://www.searchenginejournal.com/google-tests-new-design-for-desktop-search-results/266669/.

User data also enhances the ability of search providers to suggest "related searches" or recommend particular products or services in which the user is likely to be interested. For example, shoppers on shopping sites such as Kogan and Amazon are shown recommended products based both on their own data and data collected from other users.[57] Zomato recommends restaurants based on a user's preferences for specific cuisines, venues, locations, and price ranges.[58] Services such as YouTube, Netflix, iTunes, and Spotify similarly use "recommendation engines" to suggest content in which users are likely to be interested (*e.g.*, "if you like this, you might also like this").[59]

2. Utilization of user data to improve monetization

Online providers also utilize user data in order to monetize their services effectively. Many, if not most, online providers monetize their services through the showing targeted ads to users.

Benefits to advertisers: advertisers want to reach potential customers as efficiently as possible. An effective mechanism for achieving these efficiencies is through *targeted* advertising, whereby ads are shown to consumers that are most likely to be interested in the products or services offered by the advertiser and/or most likely to respond to the ad by making a purchase, becoming aware of a product or brand, or obtaining information about a product.[60] Ads can be targeted based on demographics, consumer interests, or consumer behavior (such as a user's web browsing behavior or product purchase history). Targeting gives advertisers the ability to reach potential customers more efficiently. The Preliminary Report describes the benefits of user data for ad targeting, stating that "user data can be a key input to the supply of targeted advertising services, as digital platforms can use user data to create segmented user profiles that are sold to advertisers wishing to target ads to an audience with particular characteristics. Greater collection of user data is also likely to lead to more efficient targeting of ads."[61]

Targeted advertising may reduce the tendency for advertisers to favor larger platforms. Without the ability to target ads, an advertiser would have a greater chance of finding an interested consumer on the platform with the most users. Ad targeting technology allows an advertiser to effectively target users regardless of platform size, based on specific information about the user collected from a variety of available sources.

Search ads are targeted to search queries entered by users. Although advertisers mainly target their own ads to user queries by choosing keywords, bids, and other criteria (*e.g.,* keyword match types, time of day, day of the week), search platforms also employ systems that enhance the targeting of ads (see discussion Section III.C below). Data collected from users, including aggregate click-and-query data, is valuable in determining the relevance and usefulness of particular ads in response to user search queries.[62]

---

[57] For example, Kogan features a "Recommended for you" section for its members. (Kogan homepage, available at https://www.kogan.com/au/.)

[58] Richa Bhatia, "The Amazing Way Zomato Uses Data Science for Success," Analytics India Magazine, 11 June 2018, available at https://www.analyticsindiamag.com/the-amazing-way-zomato-uses-data-science-for-success/.

[59] Wayne Thompson, "How do recommendation systems know what you might like?" SAS, available at https://www.sas.com/en_us/insights/articles/big-data/recommendation-systems.html.

[60] "Ad Targeting," Techopedia, available at https://www.techopedia.com/definition/30295/ad-targeting.

[61] Preliminary Report at 167. Also see Preliminary Report at 79.

[62] Other types of online providers similarly target ads to user searches. For instance, Kayak's proprietary advertising platform allows an airline to only show ads to users searching for routes where the airline offers service, or for a

Benefits to users: the ability of online providers to monetize their services effectively through the use of targeted ads generates significant benefits not only to advertisers, but also indirectly to users. In particular, the ability to monetize effectively creates incentives for online providers to attract users by (1) offering valuable services, often for free, and (2) investing in the quality of services offered to users.[63] This is because the greater the advertising revenues that a provider can earn from each user, the greater the benefits of attracting and retaining users. Thus, the ability to earn greater advertising revenues enhances competition for users, creating incentives for providers to invest in improving the quality of services offered and to offer those services to users at low or zero prices.[64] The targeting of ads, including through the collection and utilization of user data, therefore generates huge consumer benefits.[65]

Targeted advertising is widespread in all industries, both online and offline, and by firms of all types and sizes.[66] Offline, television networks and advertising agencies track the viewing activity and interests of consumers, using data from third-party market research companies such as Nielsen.[67] This allows firms to place ads during television shows that are viewed by the subset of consumers most likely to respond to the ad.

B. Competitive success in Search is not determined by the amount of user data

Although there is agreement that the collection of user data leads to the provision of better services and lower (often zero) prices, the Preliminary Report maintains that because of these competitive benefits, having access to a significant volume of user data gives large online platforms a competitive advantage, both in terms of providing higher quality services to users and better targeted ads to

---

hotel to only show ads to users making travel plans during times the hotel has low occupancy. (Kayak Software Corporation 2012 10-K, p. 3.)

[63] See, *e.g.*, Adam Thierer, "Relax and Learn to Love Big Data," U.S. News and World Preliminary Report, 13 September 2013, available at http://www.usnews.com/opinion/blogs/economic-intelligence/2013/09/16/big-data-collection-has-many-benefits-for-internet-users ("[D]ata collection means all consumers enjoy a fuller range or goods and services, usually at a very low price.").

[64] This interconnection between the different sides of a multi-sided platform—in this case, the advertising and the user sides—and the fact that improving monetization on one side tends to decrease prices and increase quality on the other, has been discussed extensively in the economics literature on "two-sided" markets. See, *e.g.*, Benjamin Klein, Andres V. Lerner, Kevin M. Murphy & Lacey L. Plache, *Competition In Two-Sided Markets: The Antitrust Economics Of Payment Card Interchange Fees*, 73 ANTITRUST L.J. 571 (2006).

[65] A study by McKinsey & Company estimated that the consumer surplus from advertising-funded online services in the U.S. and Europe was €100 billion already in 2010. ("Consumers driving the digital uptake: The economic value of online advertising-based services for consumers," IAB Europe, September 2010, available at http://www.youronlinechoices.com/white_paper_consumers_driving_the_digital_uptake.pdf.)

[66] The ability to offer targeted ads is not limited to large providers—for instance, the Digital Advertising Alliance noted that small online providers derive 60% of their revenue from targeted ads. (Katy Bachman, "Advertisers Pay 3 Times More for Cookie-Based Ads," AdWeek, 10 February 2014, available at http://www.adweek.com/news/technology/study-interest-based-ads-are-workhorse-internet-155616.)

[67] Nielsen has a long history of collecting user data on television viewing habits; new entrants like comScore and Rentrak also offer audience tracking data. (Amol Sharma and Christopher Stewart, "Nielsen Feels Digital Heat From Rivals," The Wall Street Journal, 12 February 2014, available at http://online.wsj.com/news/articles/SB10001424052702304703804579378973829456660.)

advertisers.[68] In short, the Preliminary Report posits that since having substantial user data can be beneficial in producing a high-quality service for users, lack of access to such data puts other firms at a competitive disadvantage.

As discussed in this section, competitive success of search platforms, including search platforms, is not determined by how much user data they can collect. Rather, competition in search is characterized by innovation and product differentiation. Moreover, while data collected from users can be important, there are many other inputs into providing high-quality search services including, perhaps most importantly, engineering resources, technological investments, and other types and sources of data. As a result, new entrants and less established search providers without access to a large volume of user data can (and do) compete through innovation and differentiation, and through the use of other inputs to provide high quality search services and monetize effectively. And, if successful, those smaller platforms can collect data from users. Thus, it is incorrect to assert that a general or specialized search platform lacking user data scale cannot attain scale. Marketplace examples demonstrate that user data is not a necessary precondition for achieving success as platform. In fact, causality likely flows in the opposite direction—successful platforms have a lot of users and therefore have access to a lot of data, but the opposite is not necessarily true. The fact that search providers can gain user scale in ways that do not involve user data weakens any theorized feedback loop that could create barriers to entry and expansion.

1. Product innovation and differentiation

a) Dynamic competition in search

Innovation and platform differentiation are key aspects of dynamic competition in search. In the early days of the online search industry, Yahoo!, Lycos, and AltaVista were all in turn thought to be dominant, entrenched search providers, but each lost its position as the most successful search engine.[69] Google did not become the most popular search engine as a result of access to large amounts of user data. In fact, when Google first started, it had no user data. Google's main initial innovation was the PageRank algorithm, devised by its founders, which analyzes the number of webpages which link to potentially relevant webpages (and the webpages that link to those webpages), as a way to determine the quality and relevancy of webpages.[70] This innovation, which did not require user data, was one of the key factors that helped Google succeed as a new entrant that faced multiple established search engines. Google earned its current competitive position through innovation and clever engineering, not as a result of having a large volume of user data.

Innovation continues to be the primary driver of competition in search. Google continues to innovate, and invest significant resources into research and development. The Preliminary Report

---

[68] Preliminary Report at 42-43, 57.

[69] See, *e.g.*, Stacey Kimmel, *Robot-generated Databases on the World Wide Web*, 19(1) DATABASE 40 (February 1996); Mark Veverka, "Merger Mania Has Finally Penetrated the Internet Space," The San Francisco Chronicle, 24 June 1998; Paul Cook, Portfolio Manager Munder Netner Fund, speaking on CNN MoneyWeek, Transcript # 98061400V35, 14 June 1998; Jack Schofield, "Computing and the Net: The new seekers," The Guardian, 6 November 1997; Danny Sullivan, "A Eulogy For AltaVista, The Google Of Its Time," Search Engine Land, 28 June 2013, available at http://searchengineland.com/altavista-eulogy-165366.

[70] See, *e.g.*, Danny Sullivan, "What Is Google PageRank? A Guide For Searchers & Webmasters," Search Engine Land, 26 April 2007, available at http://searchengineland.com/what-is-google-pagerank-a-guide-for-searchers-webmasters-11068.

describes that, in 2017, Google spent over $16 billion globally on R&D, approximately equal to 15% of its total revenue.[71] Google's technological innovations include:

> Voice Search: Google Voice Search, launched in 2008, allows a user to verbally speak queries, rather than entering them in a search box using a keyboard.[72]
>
> Integrated maps and local enhancements: Google has continually invested in improving the features and functionality of Google Maps, including integrating maps with local information (*e.g.*, Zagat), rerouting around traffic accidents, and the ability to "explore" nearby locations corresponding to various categories (where to eat, where to shop, etc.).[73]
>
> Quick answers to common questions: users often seek specific, discrete information, such as weather, sports scores, answers to math problems, or "how-to" questions. Google uses numerous resources to answer these questions, including its own "knowledge base" for history questions (*e.g.*, who was the first prime minister of Australia).[74]
>
> Universal Search: launched in May 2007, Universal Search draws results from multiple specialized search areas (*e.g.*, for images, videos, local, news) and "blends" those results into the general result page, via a "comparative ranking" algorithm.[75]
>
> Knowledge Graph: launched in May 2012, Google's Knowledge Graph provides a unified "panel" with key facts about people, places and things. Google assembles these results through "graphing" (a technical term for relating connected ideas and concepts) of various kinds of information from a fact database that Google has collected from various sources, including by licensing data from other entities.[76]

---

[71] Preliminary Report at 47.

[72] To substantially expand the language recognition ability in Voice Search, Google distributed phones to hundreds of participants in various countries for the purpose of recording their natural speech patterns, and integrated this data into Voice Search. "Google Voice Search – Research at Google," Harvard Business School Digital Initiative, 26 March 2018, available at https://digit.hbs.org/submission/google-voice-search/. Among other things, this example demonstrates how a service can acquire necessary data without first operating a popular service.

[73] Dan Moren, "Google Maps gets stunning makeover for Web, enhancements to mobile," PCWorld, 15 May 2013, available at https://www.pcworld.com/article/2038823/google-maps-gets-stunning-makeover-for-web-enhancements-to-mobile.html.

[74] See, *e.g.*, Ann Smarty, "The Evolution and Expansion of Google's Featured Snippets," Search Engine People, 17 July 2018, available at https://www.searchenginepeople.com/blog/evolution-expansion-googles-featured-snippets.html.

[75] Danny Sullivan, "Google Launches 'Universal Search & Blended Results," Search Engine Land, 16 May 2007, available at https://searchengineland.com/google-20-google-universal-search-11232; Danny Sullivan, "Google Universal Search Expands," Search Engine Land, 30 January 2008, available at https://searchengineland.com/google-universal-search-2008-edition-13256; Bret C. Taylor, Marissa Mayer, Orkut Buyukkokten, "Interface for a Universal search engine," Patent EP1700239A2, available at https://patents.google.com/patent/EP1700239A2/zh.

[76] Amit Singhal, "Building the search engine of the future, one baby step at a time," Google Blog, 8 August 2012, available at https://www.blog.google/products/search/building-search-engine-of-future-one/. When Knowledge Graph launched in 2012, Google had already compiled over 3.5 billion facts in its database, including for: Actors, Directors, Movies; Art Works & Museums; Cities & Countries; Islands, Lakes, Lighthouses; Music Albums & Music Groups; Planets & Spacecraft; Roller Coasters & Skyscrapers; and Sports Teams. (Danny Sullivan, "Google

Language translator: launched in 2006, Google Translate translates over 100 languages.[77] Google has continued to improve this service, for instance by adding optical character recognition which allows users of the Google Translate mobile app to take a picture of text (*e.g.*, street signs, menus) which will then be translated to the language of the user's choice.[78]

Search refinement options: Google has engineered multiple ways that users can filter search results. For instance, users can select "All," "Images," "Videos," or "News" to receive results from particular categories. Refinements also allow users to put symbols or words in the Google search query box to make results more precise (*e.g.*, a user can put "-" in front of terms to exclude them from results), or include "wildcard" characters ("*") to allow for broader matching of queries.[79]

Predictive search: Google Now was launched in 2012, and uses "predictive search" technology to suggest personalized facts and results to users. For example, Google Now can "anticipate" that a user typically searches for movie listings on the weekend, and will find and display local listings for weekend movies on Friday night.[80]

RankBrain: RankBrain is a machine-learning artificial intelligence system that Google rolled out around 2016. RankBrain, which is one "signal" in Google's overall search algorithm, attempts to understand the "intent" of a user's search, even when the literal search terms do not match the keywords that best correspond to the intent of the search.[81]

These innovations are driven by engineering resources and other investments, not having a large scale of user data. Moreover, the significant investments and innovation by Google are at odds with the contention that dynamic competition has been inhibited as a result of network effects and user-data-related barriers to entry and expansion.

The marketplace evidence also shows successful entry in general search through innovation and differentiation. For instance, DuckDuckGo is a general search engine that emphasizes privacy as its key differentiator to users. DuckDuckGo only has a tiny fraction of Google's scale, but has been growing rapidly and profitably.[82] In February 2018 it launched its live servers in Australia.[83] The growth of

---

Launches Knowledge Graph To Provide Answers, Not Just Links," Search Engine Land, 16 May 2012, available at https://searchengineland.com/google-launches-knowledge-graph-121585.)

[77] Barak Turovsky, "Ten years of Google Translate," Google Blog, 28 April 2016, available at https://blog.google/products/translate/ten-years-of-google-translate/.

[78] Alice Bryant, "Eight Things You Didn't Know Google Translate Could Do," VOA Learning English," 8 November 2017, available at https://learningenglish.voanews.com/a/eight-things-you-didnt-know-google-translate-could-do/4097446.html.

[79] Oliver Cragg, "Insightful tips to help you use Google Search the right way!" Android Authority, 4 July 2018, available at https://www.androidauthority.com/google-search-tips-879151/.

[80] Danny Sullivan, "The Amazing 'Google Now' — When Google Searches Before You Think To," Search Engine Land, 29 April 2013, available at https://searchengineland.com/amazing-google-now-157223.

[81] Danny Sullivan, "FAQ: All about the Google RankBrain algorithm," Search Engine Land, 23 June 2016, available at https://searchengineland.com/faq-all-about-the-new-google-rankbrain-algorithm-234440.

[82] See, *e.g.*, Adam Dorfman, "Why DuckDuckGo matters," Search Engine Land, 5 March 2019, available at https://searchengineland.com/why-duckduckgo-matters-313408; Natasha Lomas, "Pro-privacy search engine DuckDuckGo hits 30M daily searches, up 50% in a year," Tech Crunch, 11 October 2018, available at

DuckDuckGo shows that network effects and user data—or, in fact, any other hypothesized source of scale—are not a prerequisite for successful entry and expansion.

c) Competition from specialized search providers

A key aspect of competition in search is differentiation by specializing in providing search services focused on specific areas of user activity. Specialized search providers (sometimes referred to as "vertical" or "targeted" search providers) focus on tasks such as shopping (*e.g.*, eBay, Gumtree), travel (*e.g.*, TripAdvisor), restaurant recommendations (*e.g.*, Zomato), and real estate (*e.g.*, realestate.com.au, domain.com.au).

Amazon, for example, is becoming increasingly important as a "starting point" for online shopping. In the U.S., roughly half of online shoppers start their product search on Amazon, with Google accounting for about a third.[84] Amazon has also increased its U.S. share of total product search queries from roughly 45% to 55% between 2015 and 2018, while Google's share of product searches has declined.[85] Amazon has grown quickly in Australia and, as of January 2019, amazon.com ranked third amongst Australian shopping sites behind eBay and Gumtree.[86] Amazon also recently launched its dedicated amazon.com.au website in Australia.[87] Amazon has moved beyond its traditional activity as a retailer or marketplace and started offering targeted ads, competing with Google, Facebook, and many other online platforms.[88] Amazon seems to have similar plans in Australia.[89]

The Preliminary Report contends that specialized search providers do not provide a competitive constraint on general search providers such as Google, stating that specialized search is differentiated from general search in that "specialised search is restricted to providing information regarding its area of

---

https://techcrunch.com/2018/10/11/pro-privacy-search-engine-duckduckgo-hits-30m-daily-searches-up-50-in-a-year/.

[83] "Welcome to the Duck Side, Australia! DuckDuckGo servers are now live in Australia, providing faster results for searches in the region." (DuckDuckGo on Twitter, available at https://twitter.com/duckduckgo/status/960162786172133387?lang=en.)

[84] Krista Garcia, "More Product Searches Start on Amazon," eMarketer, 6 September 2018, available at https://retail.emarketer.com/article/more-product-searches-start-on-amazon/5b92c0e0ebd40005bc4dc7ae.

[85] Krista Garcia, "More Product Searches Start on Amazon," eMarketer, 6 September 2018, available at https://retail.emarketer.com/article/more-product-searches-start-on-amazon/5b92c0e0ebd40005bc4dc7ae.

[86] Rankings based on total Australian user browsing traffic in Similar Web's "Shopping" category.

[87] Amazon.com.au is ranked 7th by total Australian user browsing traffic in Similar Web's "Shopping" category as of January 2019.

[88] Spencer Soper, "Amazon Increases Ad Market Share at Expense of Google, Facebook," Bloomberg, 19 September 2018, available at https://www.bloomberg.com/news/articles/2018-09-19/amazon-increases-ad-market-share-at-expense-of-google-facebook; "Frequently Asked Questions," Amazon Advertising, available at https://advertising.amazon.com/resources/faq/?ref_=a20m_us_sp2_re_faq#display-ads.

[89] Amazon Media Group recently appointed an Australia-based managing director. (Max Mason, "Amazon hires advertising boss Henry Tajer in Australia," Australian Financial Review, 29 August 2018, available at: https://www.afr.com/business/media-and-marketing/amazon-hires-advertising-boss-henry-tajer-in-australia-20180829-h14nj9.)

specialisation."[90] But the fact that rivals offer differentiated products does not mean that they do not compete with each other—rather, differentiation is a primary dimension of competition.

Moreover, the search categories (sometimes referred to as "verticals") in which specialized providers compete are the most profitable. In fact, a large share of queries entered on Google are not monetized at all—CONFIDENTIAL [redacted] Google revenues are even more concentrated on a small share of queries—CONFIDENTIAL [redacted] The vast majority of queries (such as general factual questions about history, celebrities, or current events) are not well-suited to being monetized. Thus, specialized search providers can focus their efforts and investments on providing services in highly monetizable areas, as opposed to making investments in answering all types of user searches, many of which do not monetize at all. Indeed, this has been the nature of entry and competition in search, with specialized providers focusing on commercial queries.

Neither does the fact that specialized search services "provide certain features that are unavailable on generalised search services" mean that there is "limited substitutability" with general search services.[91] As mentioned, product differentiation, including innovation in new and differentiated features, is a key dimension of competition, not a sign of an absence of competition. Moreover, the fact that specialized providers offer "certain features that are unavailable on generalised search services" means that such providers may be a more attractive alternative for some consumers, not less. Specialized search providers also have access to specialized user data in their area of search, as well as from "feeds" related to particular business areas (*e.g.*, for financial products, real estate, consumer products, or travel), which also makes them strong competitors in those areas.[92]

The Preliminary Report also contends that there is limited substitutability between specialized and general search providers because users "may also often access a specialised search service via a more general online search such as Google."[93] However, the empirical evidence shows that while some users do access specialised search providers via Google, the vast majority do not. In fact, mobile apps and direct navigation are generally the predominant means of accessing specialized search providers. For example, in 2018, 59% of the total traffic to the top four shopping platforms in Australia (eBay, Gumtree, OzBargain, and Amazon) was via mobile apps and 23% of traffic came from users directly navigating to the provider website, while only 10% came from Google organic search referrals.[94] Furthermore, a

[90] Preliminary Report at 41.

[91] Preliminary Report at 41.

[92] Even within a particular category of specialized search, online providers offer differentiated services and products, but this does not mean that such services do not compete against one another. For instance, top online providers of travel services in Australia range from "metasearch" sites like Kayak (which show offers and itineraries from multiple other travel sites) to "travel agent" sites like Expedia (which provide both planning and booking services), to "review" sites like Tripadvisor (which allow users to read and post reviews of travel destinations via a "social network"-type forum), to the sites of airline and hotel brands. The fact that these travel sites provide distinct functionality does not mean that they do not compete against each other. Similarly, among top Australian online shopping sites, Amazon (a full-fledged e-commerce site) competes with Gumtree (a "classifieds" site focused on connecting individual buyers and sellers), and OzBargain (a "social" price comparison site where a community of users posts information on bargains and deals).

[93] Preliminary Report at 41.

[94] SimilarWeb data for Shopping category.

substantial share of Google organic search referrals to these providers was pursuant to "branded" (or "navigational") search queries, where the user searched for the provider's name. Excluding this traffic, Google organic referrals accounted for just 6% of traffic.

2. Other inputs used in the provision of search

There are also many inputs other than user data that go into providing online services including, perhaps most importantly, investments in engineering and technology. Even the processing and analysis of data requires investments in computers, software, and engineers. Other inputs, including engineering talent, are not only complementary with user data, but often can be substitutable. A provider without large volumes of user data can, for example, invest more in engineering and innovation to develop high-quality online services.

In the provision of search in particular, there are many inputs other than user data that go into the provision of high-quality services. Google has invested significant resources into other inputs, and those investments have been instrumental in Google's success:

> Engineering resources: a key determinant of search quality is the amount and caliber of engineering resources devoted to improving a search engine's algorithms. Search providers, including Google, need high-caliber engineers that can use data to improve the search algorithms. A blog post by a Microsoft engineer stated that "Google's engineers are amazing."[95]
>
> Crawling and indexing technology: another important input for search providers is the provider's web crawling and indexing technology, which is the process of gathering information from across hundreds of billions of webpages and organizing that information in a search index.[96] Recent estimates suggest that Google's web index is around 100 times the size of Bing's (measured by the number of webpages crawled).[97] Another important dimension of the crawling and indexing technology is how quickly the provider can crawl and index new content (which relies on engineering ability, not user data), which is vital in identifying results that are relevant to "fresh" queries—*i.e.*, queries regarding recent events and information. Google's improvements in providing results to "fresh" queries is in large part due to a web indexing system it introduced in June 2010 called Caffeine. Google identifies fresh queries primarily based on rapid increases in *content* on a particular topic, which can be observed from documents recently crawled and large increases in relevant documents recently indexed (*e.g.*, recent articles on a current event).[98]
>
> Human raters: Google employs a large number of human raters (over 10,000 worldwide) to evaluate the quality and relevance of its search results, and uses those evaluations to improve its algorithms. As mentioned above, several of Google's products such as Voice and Translate also utilize input of human evaluators and contributors. Google also expends substantial resources to

---

[95] Alex Clemmer, "What 'viable search engine competition' really looks like," 4 January 2014, available at http://blog.nullspace.io/building-search-engines.html.

[96] See, *e.g.*, "How Search organizes information," Google.com, available at https://www.google.com/search/howsearchworks/crawling-indexing/.

[97] See, *e.g.*, "The size of the World Wide Web: Estimated size of Google's index" and "The size of the World Wide Web: Estimated size of Bing index," available at https://www.worldwidewebsize.com/.

[98] Saul Hansell, "Google Keeps Tweaking Its Search Engine," The New York Times, 3 June 2007, available at http://www.nytimes.com/2007/06/03/business/yourmoney/03google.html?ei=5088&en=f003a2b328ec0a72&ex=1338523200&partner=rssnyt&emc=rss&pagewanted=all.

fight "spam" in its search results (and other services such as Gmail), including manual review and algorithmic screening.[99]

There are also other dimensions of quality and means of competing for users that do not involve user data. Innovative features, an attractive user interface, complementary services, speed and ease of use are all important quality attributes for search providers.[100] Developing these features does not require having access to a significant volume of user data, but is based on innovation and investments in engineering and other inputs.

3. Other types and sources of data used in the provision of search

There are also many types and sources of data, other than user data, that can be valuable inputs into the provision of search services. For example, Google's Page Rank algorithm relies upon data collected while crawling websites (*i.e.*, data regarding frequency of links between sites), not on user data. Search engines also receive many misspelled queries from users, and using various types and sources of data to "revise" these queries and return the desired result is an important aspect of search quality. While user data plays a role in spelling revisions, search engines also analyze "crawled" data from Internet websites to gauge common spellings and misspellings of words.

Search engines, including Google, also acquire data from third parties that are used to answer particular user queries, including for consumer products, tourist activities, sports scores, movie times, stock performance, weather forecasts, currency exchange rates, and other financial information. These "data feeds" help search engines to better respond to user queries, and allow the results to be presented in more user-friendly formats. For example, product merchants submit standardized feeds to search engines that contain product images, prices, and dimensions (among other things). These standardized data feeds give the search engine enhanced ability to develop a more attractive and useful interface for searchers to browse product results.[101]

Similarly, specialized search providers that offer search services on specific topics such as shopping, travel, and local interests also have many sources of data that they can use to improve their services. Many of these online providers offer aggregation and comparison services (*e.g.*, for financial

---

[99] See, *e.g.*, Christopher Ratcliff, "How Google fights webspam and what you need to learn from this," Search Engine Watch, 4 May 2016, available at https://searchenginewatch.com/2016/05/04/how-google-fights-webspam-and-what-you-need-to-learn-from-this/. Google uses machine learning technology it has developed to filter spam email and phishing attempts in Gmail. (Frederic Lardinois, "Google says its machine learning tech now blocks 99.9% of Gmail spam and phishing messages," TechCrunch, 31 May 2017, available at https://techcrunch.com/2017/05/31/google-says-its-machine-learning-tech-now-blocks-99-9-of-gmail-spam-and-phishing-messages/.)

[100] For example, in 2016, Gumtree unveiled a major rebranding of their mobile and desktop site to improve user experience. The improvements included an easier way to search items based on location, and an updated app allowing users to post ads more quickly and manage responses more easily, and a 24-hour customer service live chat to assist users with their support queries. ("The Start of Something Special," Gumtree, 7 January 2016, available at https://blog.gumtree.com/the-start-of-something-special/.)

[101] See, *e.g.*, Christi Olson, "Move over, text ads: How data feeds are driving new search experiences," Search Engine Land, 19 May 2016, available at https://searchengineland.com/move-text-ads-data-feeds-driving-new-search-experiences-249716.

products, real estate, or consumer products), for which the primary source of the data used are "feeds" from firms offering these products and services (*e.g.*, banks, realtors, retailers).

The Preliminary Report contends that Google's market power in search is protected because Google enjoys an "important advantage of scope" due to Google's "accumulation" of user data from multiple sources,[102] which weakens the constraint on Google from dynamic competition.[103] The Preliminary Report points to "two principal data sources" that give rise to the purported "advantage of scope": 1) Google's owned and operated sites (*e.g.*, Google.com, Gmail, YouTube); and 2) third-party sites.[104] However, other than the user query itself, the primary sources of data in the provision of Google search services are user click-and-query data, web crawled data, and other sources of data such as "data feeds" used to answer particular categories of queries (*e.g.*, consumer products, travel, sports scores, movie times, stock performance, weather forecasts, currency exchange rates, and other financial information), as discussed above. These types of data, which are not derived from users of third-party sites or from Google's owned and operated properties (other than Google.com), are far more important for providing search services.

Moreover, the claim that Google has a competitive advantage due to its access to a significant "accumulation" of data fails to recognize that data is highly diverse, both in terms of quality and uses. It is not the sheer amount of user data that is important to search providers, but how useful the data is in improving the services offered. For instance, data on a specific type of user activity (*e.g.*, user purchases) is often more valuable in providing a particular service (*e.g.*, comparison shopping) than having lots of data across a wide range of activities, such as data on general searches conducted by users. Specialized sites collect user data that is valuable in supporting the products or services they offer.

For example, as the largest product search provider in the world, Amazon has access to significant data on product purchases. Amazon collects and analyzes many types of user data, including purchase history, products viewed or searched for, length of visits to certain pages, and page interaction information (such as scrolling, clicks, and mouse-overs). Amazon uses such data to recommend products based on a user's personal data, inform users of similar products that other customers have purchased, and provide product reviews from other customers.[105] Other shopping sites also collect similar types of information. And, many other specialized search providers also have large volumes of user data that are valuable for providing particular online services and targeting ads.

---

[102] Preliminary Report at 48.

[103] Preliminary Report at 48.

[104] Preliminary Report at 48.

[105] An article in Fortune described how Amazon's "recommendation system is based on a number of simple elements: what a user has bought in the past, which items they have in their virtual shopping cart, items they've rated and liked, and what other customers have viewed and purchased." (JP Mangalindan, "Amazon's recommendation secret," Fortune, 30 July 2012, available at http://tech.fortune.cnn.com/2012/07/30/amazon-5/.) Amazon also offers its shopper data to brand advertisers, as it continues to invest in its advertising platform, Amazon Advertising Platform. (Kiri Masters, "Amazon News Roundup: Ad Revenue to Reach $20B By 2020, And A Sales Tax Kerfuffle," Forbes, 3 April 2018, available at https://www.forbes.com/sites/kirimasters/2018/04/03/amazon-ad-revenue-to-reach-20b-by-2020-and-other-amazon-news/#4c7112404bb3.)

C. Diminishing returns to user data scale

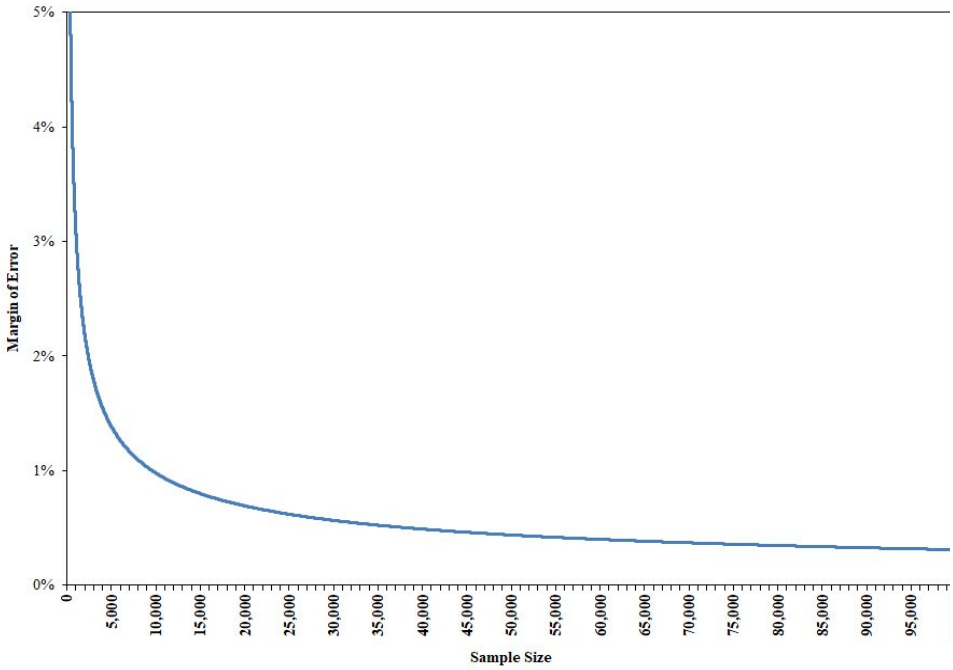
The foregoing discussion does not mean that user data is not an important input into providing online services, including search. As discussed above, user data enhances search providers' ability to improve the quality of services offered to users and to monetize those services through targeted ads. However, the value of such data in providing online services, including search, is subject to diminishing returns, which weakens the propensity for the collection and utilization of user data to lead to barriers to entry and expansion.

1. User data scale in search

In the provision of many online services, as in most businesses, there are economies of scale. But these economies are often subject to rapidly diminishing returns, meaning that any advantages of scale weaken or even disappear. Diminishing returns to scale imply that the benefits from additional user data scale (either in attaining lower costs, higher quality, or higher monetization) become progressively smaller as a provider's scale of users or customers increases. As a result, any competitive advantage that a provider with more data has over smaller rivals also diminishes rather than strengthens.

In the context of the use of data, diminishing returns can be illustrated by statistical sampling error when performing a survey. The survey sampling error decreases as more people are surveyed, but the rate of decrease slows as the sample size gets larger. A survey with a random sample of 1,000 respondents has a margin of sampling error of approximately 3 percentage points at a 95% confidence level. This means that 95% of the time, the survey would give rise to reported results within 3 percentage points of the true value. Increasing the scale of the sample by 1,000 respondents would reduce the margin of error by 0.9 percentage points (to about 2.2%). However, for a survey with 10,000 respondents, increasing the sample size by 1,000 yields a decline in the margin of error of only 0.05 percentage points. For a survey with 100,000 respondents, increasing the sample size by 1,000 yields a decline in the margin of error of only 0.002 percentage points. Figure 1 below illustrates that the margin of sampling error (measured on the vertical axis) declines with sample size (measured on the horizontal axis), but the rate of the decline diminishes rapidly. After a certain sample size, the reduction in the margin of error as the sample increases is very small.

Figure 1: Diminishing Returns to Scale—Sampling Error and Sample Size



A similar concept of diminishing returns generally applies to the utilization of user data collected by search providers. Most user data collected and utilized by search providers involves information about the search terms that users have entered, and the results on which users have clicked. This click-and-query data allows search providers to gain insight into what results may be relevant in response to particular search terms, and thereby return more relevant results. Click-and-query data is an important input into search algorithms, but the value of incremental data in providing relevant search results decreases as the amount of data available to those algorithms increases. One of the signals that is used to determine relevant search results (and their ranking) is the average click-through rate—*i.e.*, the fraction of users that have clicked on a result based on the same or similar query. As illustrated by the above example, the accuracy of statistical estimation of an average is improved by having more data, but the incremental value of that data (*i.e.*, the marginal improvement in accuracy) decreases as more and more data is added.

Returns to user data scale diminish rapidly for common queries, often referred to as "head" queries, which are queries entered frequently by users. A large share of head queries is made up of navigational queries, which are queries that users enter on search engines with the intent to navigate to a particular website (such as "Facebook" or "Amazon").[106] Other types of head queries include popular products, personalities, or notable events (*e.g.*, "Galaxy S9," "Meghan Markle," or "World Cup"). The marginal value of additional head queries is likely to be minimal, or even zero, at a very low scale for a search provider. For instance, a search provider that has data on user clicks for 10,000 previous queries for "Facebook" or "Amazon" is likely to be able to determine the most relevant search results equally as

[106] CONFIDENTIAL [redacted]

well as a provider that has seen that same query 100,000 times. Thus, it is generally accepted that no significant user scale is necessary to provide relevant search results for more popular queries.[107]

The Preliminary Report, however, postulates that "large quantities of [user] data improves [sic] the ability for the algorithm to generate reliable relevance rankings for queries that are uncommon."[108] Queries that are infrequently entered by users, referred to as "tail" queries, include misspelled queries, addresses, specific product descriptions or model numbers, and detailed queries composed of multiple terms. For some tail queries, there can be value to having a greater amount of user click-and-query data, since this information may contain data for the same tail queries. However, an increase in scale does not generally eliminate or materially reduce the need to deal with tail queries in other ways. For instance, an increase in query volume has not materially reduced the portion of Google's traffic that is comprised of queries that it has never seen before—"On a daily basis, 15 percent of queries submitted -- 500 million -- have never been seen before by Google's search engine, and that has continued for the nearly 15 years the company has existed…"[109]

Thus, having greater query volume does not eliminate a platform's need to return relevant results to queries that it has never seen before. As a result, providing relevant results in response to tail queries, including queries never seen before, involves clever engineering and the use of other types of data. For example, a search engine's ability to return the best results for a query comprised of a specific product model number may have nothing to do with whether other users previously had conducted the same search query. Instead, it would most likely depend on whether the search engine had "crawled" web pages containing the exact model number queried by the user, or obtained product data feeds from manufacturers or retailers, inputs which are independent of the search provider's volume of user data. Search engines also receive many misspelled queries from users, and use various types and sources of data to "revise" these queries and return the desired result. While user data plays a role in spelling revisions, search engines also analyze "crawled" data from websites to gauge common spellings and misspellings of words.[110] Researchers are also studying innovative engineering techniques to use head queries to better answer tail queries.[111]

Empirical studies show that the use of data is subject to diminishing returns to scale. For example, a recent study of Amazon's product inventory forecasting system found that there were diminishing returns to additional data in improving inventory forecast accuracy.[112] This finding of

---

[107] See, *e.g.*, Aleksandr Chuklin and Maarten de Rijke, *Incorporating Clicks, Attention and Satisfaction into a Search Engine Result Page Evaluation Model*, CIKM '16 Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, October 2016: "While a small number of head queries represent a big part of a search engine's traffic, all modern search engines can answer these queries quite well."

[108] Preliminary Report at 43.

[109] Dan Farber, "Google Search scratches its brain 500 million times a day," CNET, 13 May 2013, available at http://www.cnet.com/news/google-search-scratches-its-brain-500-million-times-a-day/.

[110] Dr. Jim Kleban, "Bing Search Quality Insights: Building a State-of-the-Art Speller," Bing Blogs, 3 January 2013, available at http://blogs.bing.com/search/2013/01/03/bing-search-quality-insights-building-a-state-of-the-art-speller/.

[111] See, *e.g.*, Yangqiu Song, Haixun Wang, Shusen Wang, and Weizhu Chen, *Transfer Understanding from Head Queries to Tail Queries*, CIKM'14 Proceedings of the 23rd ACM International on Conference on Information and Knowledge Management, November 2014.

[112] Patrick Bajari, Victor Chernozhukov, Ali Hortaçsu, and Junichi Suzuki, *The Impact of Big Data on Firm Performance: an Empirical Investigation*, Working Paper 24334, NBER, February 2018.

diminishing returns applied both to having sales data on more products within a category as well as from having a longer history of sales data for any particular product.

Similarly, empirical evidence indicates that the value of user data in returning relevant search results also is subject to diminishing returns. Chiou and Tucker (2017) studied whether the accuracy of Bing and Yahoo! search results was affected when the amount of historical click-and-query data these search engines retained for use in their algorithms was increased or decreased (as a result of European regulatory changes).[113] The authors found that none of these changes affected the accuracy of search results in a statistically significant way. Chiou and Tucker conclude that their study finds "little empirical evidence that reducing the length of storage of past search engine searches affected the accuracy of search."[114] These findings suggest that the addition or removal of large amounts of user data do not materially change the accuracy of search engine results. The findings also suggest that the value of user data in providing high-quality search results depreciates fairly rapidly, such that the accumulation of user data over time does not give established platforms such as Google a material competitive advantage over more recent entrants.

2. User data scale in search advertising

As discussed above, the Preliminary Report maintains that barriers to entry and expansion also arise because of the benefit of having large amounts of user data for purposes of targeting search ads.[115] However, as discussed in this section, although user data has some value for purposes of targeting search (and other types) of ads, search ads are primarily targeted by the advertisers themselves by choosing keywords and other criteria; advertiser keywords and the user query are the primary components of targeting search ads, which does not require significant scale of user data.

To a large extent, advertisers target their own search ads. Advertisers placing search ads select keywords, enter bids (referred to as the maximum cost-per-click) and budgets, and choose other criteria such as keyword "match type" (*i.e.*, how closely user queries must match keywords). The advertiser chooses these inputs based on its own analysis regarding which search keywords are likely to reflect an interest in its products or services.[116] Geographic location and time scheduling also are used by advertisers to more narrowly target ads to users. The images below show part of the Google Ads Editor, which enables advertisers to choose keywords, bids, keyword match types, and other criteria (such as geographic targeting and the time ads are shown).

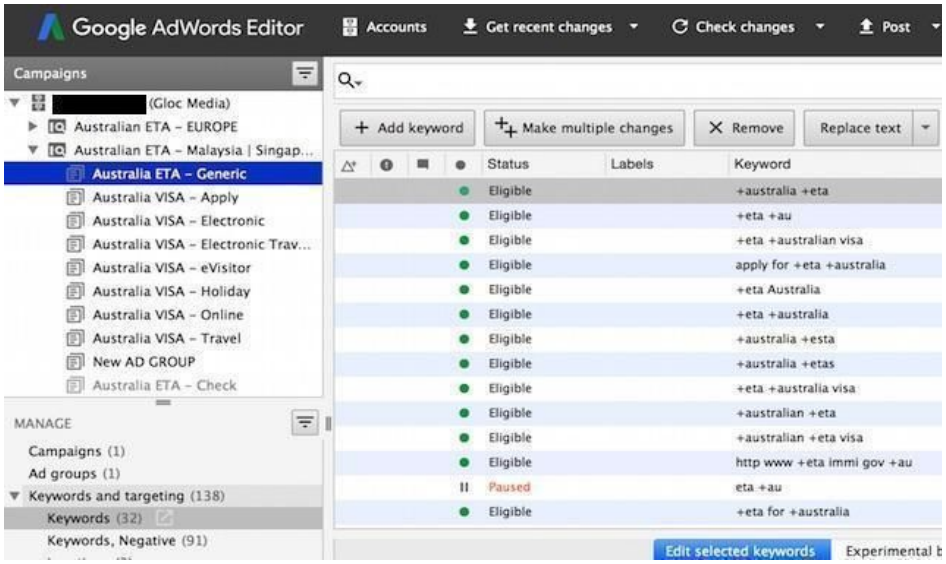
[113] Lesley Chiou and Catherine Tucker, *Search Engines and Data Retention: Implications for Privacy and Antitrust*, NBER Working Paper 23815, September 2017.

[114] *Id.*

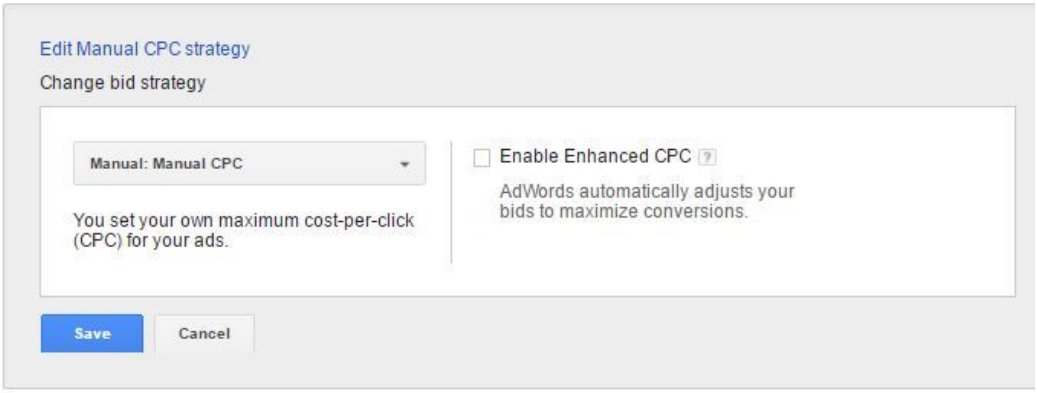
[115] Preliminary Report at 57.

[116] Advertisers may use Google's "Keyword Planner" tool to assist in choosing relevant keywords.

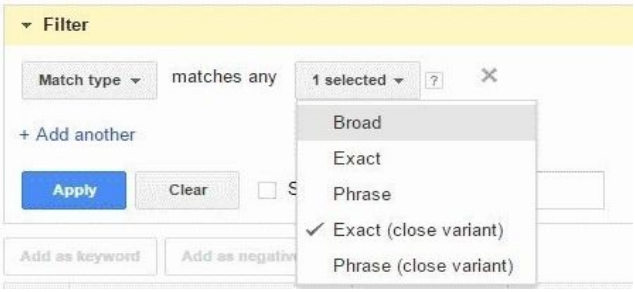
Google Ads Editor: Keywords



Google Ads Editor: Bids and Bidding Strategy



Google Ads Editor: Keyword Match Type



With respect to user data, the user query is the most valuable piece of information in targeting search ads. When a user enters a query such as "buy TV," that query reflects key information about the user for purposes of targeting ads. Search queries are reflective of a user's interests and intent, which is the essence of ad targeting. As the Preliminary Report notes, user queries "provide a particular form of

strong evidence about the user's purchasing intentions."[117] The advertiser's selection of keywords and the user query are the primary components of targeting search ads. Targeting ads based on user queries and advertiser-chosen keywords does not require significant scale of user data.

User data does help train Google's "ad quality" models used in ranking ads, which is another component of which ads are shown to users and their ranking. Google determines if and where to show each ad on the search results page based on the "Ad Rank," a measure which incorporates the advertiser's bid and the ad quality. The higher the ad quality, the higher that ad will rank (and the higher the probability that it will be shown), all else equal. Ad quality is based on Google's prediction of the relevance and usefulness of the ad, keywords chosen by the advertiser, and the advertiser's landing page.[118] However, similar to the use of data for purposes of improving the quality of services offered to users, there are diminishing returns to user data for purposes of estimating ad quality. Ad quality models also are based on many factors other than user data, including engineering efforts and human "rater" input to evaluate ad quality.

Moreover, counteracting any positive impact of scale on ad targeting, greater user scale can lead to lower monetization. In particular, an increase in the number of users corresponds to an increase in the supply of ad slots that the platform can offer advertisers which tends to *reduce* monetization as a platform's volume of users increases. Fundamental economic principles indicate that, all else equal, an increase in supply leads to lower price when demand is downward-sloping. Advertisers are likely to have diminishing value of incremental ads shown on a given platform—*i.e.*, a downward-sloping demand.[119] As the number of users increases, the incremental value of an ad slot to advertisers, and their willingness-to-pay for the ad slot, also is likely to decrease. The lower willingness-to-pay means that advertisers will bid less as the number of ad slots increases, which can lead to a lower average monetization per user. This does not suggest that there are no economies of scale in monetizing at *any* scale. When a platform is small, the positive effects of user data in targeting ads can overcome the diminishing marginal value to advertisers of incremental ad slots. But diminishing returns to user data in targeting search ads, combined with disadvantages of scale in monetization, means that any scale economies related to ad targeting are likely to diminish or even decrease after some scale.[120]

A "natural experiment"—in which Bing's query volume increased substantially—strongly suggests that user data scale does not determine search ad monetization. In August 2010, Microsoft's deal to have its Bing subsidiary serve search ads on the Yahoo! platform essentially doubled Bing's query volume. Advertisers could reach Bing and Yahoo! users by advertising on one ad platform (adCenter, later rebranded "Bing Ads"), and Bing had access to Yahoo! click-and-query data for ad targeting. If

---

[117] Preliminary Report at 56.

[118] "Ad Rank," Google Ads Help, available at https://support.google.com/google-ads/answer/1752122?hl=en.

[119] The diminishing value of incremental users is particularly pronounced for advertisers with limited budgets—as the number of users increases those limited budgets are spread out over more ads.

[120] The effect of improved ad targeting in improving monetization by attracting more advertisers also is subject to diminishing returns. This is so because, particularly for platforms with a significant number of advertisers, additional advertisers "crowd out" other advertisers on the platform. For instance, in the context of search ads, because there is a maximum number of ad slots on a search result page, for "fully-sold" auctions (*i.e.,* auctions for which all ad slots on a search result page are filled by ads) successful bids by additional advertisers will displace the ads of other advertisers. Even in auctions that are not "fully-sold," the presence of additional successful bidders decreases the probability that users will click on other ads on the page (*i.e.,* decreases the expected click-through-rate of other ads).

claims regarding the effects of scale on search revenue were correct, Bing should have increased its own monetization, as well as Yahoo!'s. However, Yahoo!'s monetization is reported to have *worsened* after Bing began serving ads on its site. Yahoo!'s then-CEO, Carol Bartz, attributed Yahoo!'s decline in search revenue to poor performance by Microsoft, stating that "adCenter isn't yet producing the RPS [revenue per search] we hoped for and are confident as [sic] possible… technical limitations in the current adCenter platform mean the click volumes just isn't [sic] there yet."[121] This natural experiment demonstrates that user data scale does not determine monetization, and highlights the importance of not confusing superior execution and engineering (or poor execution by rivals) with scale effects.

D. The collection of user data does not preclude rivals from also collecting data

Another important economic factor that weakens the potential for user scale to create barriers to entry and expansion is the fact that the collection of data by incumbent online providers, including search providers, does not preclude smaller rivals from also collecting the same or very similar user data. Standard economic theories of "vertical foreclosure"—*i.e.*, of foreclosing rivals through control of an input—are based on a dominant firm depriving rivals of access to an essential input through the purchase of exclusionary rights, or other arrangements that deprive rivals of access to the essential input.[122] Pursuant to the utilization of user data in the provision of online services, including search, (1) there are no exclusive contracts with users; (2) user multi-homing is ubiquitous; and (3) data is non-rivalrous. These characteristics significantly diminish any data-driven feedback loops and the possibility that the collection of user data entrenches large online platforms. If an incumbent online provider collects data from users, such collection does not deprive smaller rivals and new entrants of an essential input.

Lack of exclusivity: there is no exclusivity in the collection of user data by online providers, including search providers. Search platforms do not require no exclusive contracts with users. There is also generally no pricing structure, such as lump-sum fees or loyalty/volume discounts, that makes it uneconomical to multi-home or that locks-in users to a particular platform. In fact, many large online platforms, including Google, provide services to users for free, without any fees, long-term contracts, or other requirements regarding the use of the platform or of rival platforms. Simply, Google and other online providers do not have exclusive access to any user.

User multi-homing: in practice, the lack of exclusivity is reinforced by the fact that users tend to utilize, and share their information with, a variety of different online platforms. Users often utilize multiple online services, even for the same type of task, which gives multiple providers the ability to collect data on the same user. Users can benefit from multi-homing when different networks offer different features, functions, or quality levels. And, in online markets, costs to users of switching platforms or multi-homing are very low.[123] There is essentially no incremental cost for users to interact with multiple digital platforms. Even when users are required to sign up for an online service, it is

[121] Danny Sullivan, "The Yahoo Search Revenue Disaster," Search Engine Land, 20 April 2011, available at http://searchengineland.com/the-yahoo-search-revenue-disaster-73868. Two years later, in 2013, it appeared that Microsoft's contribution to Yahoo!'s search ad revenue continued to lag, according to Yahoo! CFO Ken Goldman: "There was still a gap in monetization and we will work with Microsoft to improve our search monetization." (Danny Sullivan, "Why Yahoo Will Never Reach The 'Revenue Per Search' That Microsoft Promised," Marketing Land, 7 May 2013, available at http://marketingland.com/yahoo-microsoft-rps-guarantee-42680.)

[122] See, *e.g.*, Eric B. Rasmusen, J. Mark Ramseyer & John S. Wiley, Jr., *Naked Exclusion*, 81(5) AMER. ECON. REV. 1137 (1991).

[123] See, *e.g.*, Aaron S. Edlin & Robert G. Harris, *The Role of Switching Costs in Antitrust Analysis: A Comparison of Microsoft and Google*, 15 YALE J. L. TECH. 169 (2013).

relatively easy for users to sign up multiple times and to enter similar personal details across login profiles.

As a result, there is extensive user multi-homing whether one looks at search providers that provide similar services or differentiated providers that nevertheless may compete for users with regard to a particular activity. For instance, in searching for products, a recent study found that consumers used multiple sites to conduct product research, including social platforms, product review sites, and the manufacturer's website.[124] Similarly, for travel-related search, surveys indicate that consumers utilize multiple sites to research and book travel, including general search engines and vertical search sites specializing in travel.[125] The ubiquity of user multi-homing for online activities reflects the fact that the costs of doing so are very low, and there is often value to users of utilizing multiple platforms that may provide different information or services related to a particular user task. All these online providers can and do collect user information in the process of providing services.

The non-rivalrous nature of user data: in contrast to most inputs, user data also is "non-rivalrous," meaning that collection and use by one provider does not detract from collection and use by others.[126] A rivalrous good is a good whose consumption by one consumer or firm prevents simultaneous consumption by others. A non-rivalrous good is one for which consumption by one consumer or firm does not reduce the resources available to other consumers or firms.

In combination with extensive multi-homing, the non-rivalrous nature of user data means that collecting data from a given user does not prevent or inhibit rivals from collecting data from that same user. User multi-homing and the non-rivalrous nature of user data weaken any claimed data-driven feedback loops. User multi-homing means that rivals, even smaller ones, can compete for users and, if they do so successfully, can collect data from those users. The implication of the non-rivalrous nature of data and user multi-homing is that many providers have access to significant volumes of user data. As a result, no one firm controls all, most, or even a significant share of user data. Many online providers collect data from users (including from the same users), and many have access to significant amounts of user data from various sources, including for the same users.

One example of this is the fact that many parties collect user locational information from mobile devices. Another example is the fact that, virtually all online providers, and many third parties (such as data analytics firms), use web cookies to collect information from online users, even from the same user's activities on the same website (see Section IV.A). More generally, many firms compete to attract users and collect data from users. This includes large established firms as well as many other online platforms.

---

[124] "Which Touchpoints Matter? Decoding the Path to Purchase," American Marketing Association, 27 February 2018, available at https://www.ama.org/partners/content/Pages/which-touchpoints-matter-decoding-the-path-to-purchase.aspx.

[125] See, *e.g.*, "The Travel Marketer's Guide to the U.S. Digital Travel Landscape," PhocusWright, December 2017, available at https://www.phocuswright.com/Free-Travel-Research/The-Travel-Marketers-Guide-to-the-US-Digital-Travel-Landscape.

[126] "The Age of Analytics: Competing in a Data-Driven World," McKinsey & Company, December 2016, available at https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/The%20age%20of%20analytics%20Competing%20in%20a%20data%20driven%20world/MGI-The-Age-of-Analytics-Full-Preliminary Report.ashx.

The collection of user data is conducted by firms of all sizes, by firms offering a wide array of services to users, and by both new entrants and established players.

In sum, the collection of user data by established online platforms, including search providers like Google, does not lead to barriers to entry and expansion or to the entrenchment of a dominant platform. Smaller rivals, including new entrants, can (and often do) compete for users, and through that competition gain user scale.

IV. User Scale, Network Effects, and Advertiser Fixed Costs in Display Ad Intermediation

The Preliminary Report does not find that Google has market power in display ad intermediation services, but contends that Google enjoys "strong advantages" in the supply of such services.[127] As I discuss in this section, Google does not have a significant competitive advantage in display ad intermediation services. In fact, the display ad intermediation marketplace is highly competitive, and many providers have access to significant scale and scope of user data.

A. User data in the provision of display ad intermediation

The Preliminary Report contends that, according to some third-party submissions, Google collects user data that is "not otherwise accessible by advertisers" not using Google's ad platform, including from Google's owned and operated properties (*e.g.*, Google.com and YouTube) and from third-party websites.[128] According to the Preliminary Report, Google's data is "extensive" and "likely to be unparalleled."[129] Yet however "extensive" Google's data is, it does not prevent many other firms from also collecting data and competing with Google to provide display ad intermediation services.

Display ads are ads that appear on publisher websites or in mobile applications. Publishers own online "real estate" which they sell to advertisers for placement of display ads; advertisers purchase display ad space in which to show ads in order to promote their products or services. Display ads can take various forms, including static ads (such as banner ads) and rich media ads (such as video ads). Publishers may negotiate and sell display ad space on their sites directly to advertisers (or ad agencies). Advertisers and publishers also use a variety of ad tech and intermediation services to buy, sell, and place ad space programmatically. On the demand side, there are various services that advertisers use, such as ad servers, demand-side platforms (DSPs), ad networks, and ad exchanges. The lines between most of these demand-side ad tech and intermediation services have blurred. Most providers offer similar capabilities to help advertisers buy ads and manage and optimize their ad campaigns across different channels and formats. Similarly, on the supply side, publishers use a variety of ad technologies and intermediation services to sell and manage their ad inventory, including ad servers, supply-side platforms (SSPs), ad networks, and ad exchanges. As with demand side services, most of these services offer overlapping capabilities. Notably, many ad tech players in practice provide multiple such services. For example, SSPs typically offer an integrated ad exchange as part of that service.[130]

---

[127] Preliminary Report at 82.

[128] Preliminary Report at 82.

[129] Preliminary Report at 83.

[130] As noted in one article from 2014, "The distinction between an ad exchange and a supply-side platform (SSP) has become muddled as the once disparate but complementary technologies have merged." (Ryan Joe, "Defining SSPs, Ad Exchanges And Rubicon Project," Ad Exchanger, 7 February 2014, available at

It is worth noting that display ad intermediation makes up a relatively small share of Google's ad revenues. First, only about 17% of Google ad revenue is from display ads.[131] Moreover, it is estimated that two-thirds of Google display ad revenue is from YouTube, which Google owns and operates.[132] And, as the Preliminary Report acknowledges, less than 10% of Google's total ad revenue in Australia is from display ad intermediation on third-party sites.[133] Google's display ad revenue is split over a number of different intermediary services, so it is not necessarily directly comparable to the revenue of competitors who offer fewer solutions.

In assessing the role of user data in the targeting of display ads, it is important to distinguish between user data *scale* and user data *scope*. User data scale pertains to having access to data from a lot of users; user data scope refers to having multiple data points on specific users—*i.e.*, different types and/or sources of data for a specific user. As discussed below, for several types of display ads, user data scale is not a prerequisite to effective targeting. And, although it can be valuable to have multiple data points on specific users, many other ad tech firms (as well as publishers and advertisers) have access to a substantial scope of user data. Thus, Google does not have a significant competitive advantage in display ad intermediation services by virtue of having access to any claimed "unparalleled" scale or scope of user data. In fact, the display ad business is highly competitive, with many providers at each level of the ad tech stack.

1. User scale and scope plays a limited role in various forms of display ad targeting

Direct sale of display ad space: many display ads (*e.g.,* banner ads on a publisher home page) are sold directly by publishers to advertisers. For these directly-sold display ads, publishers do not rely on ad intermediation service firms to target ads. Rather, publishers interact directly with advertisers to negotiate volume, payment, and placement terms for advertising that is placed on publisher sites.[134] Direct sales account for a large share of digital display ads sold. Worldwide, it is estimated that about 40% of digital display ads are sold directly by publishers.[135] In Australia, 37% of digital ads are estimated to be sold

---

https://adexchanger.com/yield-management-tools/defining-ssps-ad-exchanges-and-rubicon-project/.) For example, aerServ, a top SSP, has an integrated ad exchange called aerMarket. ("Best Supply Side Platform (SSP) Software," G2 Crowd, available at https://www.g2crowd.com/categories/supply-side-platform-ssp; "aerMarket Ad Exchange," AerServ, available at https://www.aerserv.com/aermarket/.) OpenX offers both ad exchange and SSP services for publishers. ("OpenX Ad Exchange," OpenX, available at https://www.openx.com/publishers/adexchange/; "OpenX - Topic: SSP," OpenX, available at https://www.openx.com/tag/ssp/.)

[131] "Google Search Ad Revenues," eMarketer, 1 September 2018, available at https://forecasts-na1.emarketer.com/5a0a236aba8ffd0d34173452/5a09e843ba8ffd0d341733fe; "Google Display Ad Revenues," eMarketer, 1 September 2018, available at https://forecasts-na1.emarketer.com/5a0a236aba8ffd0d34173452/5a09dbfeba8ffd0d341733fc.

[132] "YouTube Ad Revenues," eMarketer, 1 September 2018, available at https://forecasts-na1.emarketer.com/59e8c7a4bfce890eb411ee86/59e7bf7ebfce890eb411ec54.

[133] Preliminary Report at 38.

[134] Claudia Barbiero, "Key Differences between Direct-Sold and Programmatic Advertising," Ad-Juster, 6 November 2017, available at https://blog.ad-juster.com/differences-between-direct-sold-and-programmatic-advertising/.

[135] "Programmatic Ad Spending Worldwide, 2012-2019," eMarketer, 20 November 2017, available at http://totalaccess.emarketer.com/chart.aspx?r=214016, citing "Programmatic Marketing Forecasts 2017," Zenith.

directly.[136] Moreover, because directly-sold ads are generally the most profitable, the share of display ad revenues made up by directly-sold ads is likely to be even higher.[137] For some publishers, direct ad sales make up the vast majority of their ad revenue.

Contextually-targeted display ads: many display ads, including most directly-sold ads, are targeted contextually, meaning that they are targeted to the content on a publisher website, or specific pages on the publisher website.[138] The content of a webpage can be valuable in targeting groups of users that are likely to be interested in an advertiser's products or services. For example, a carmaker might want to show ads on the "Automotive" section of an online news publisher site; a restaurant may target readers of the "Food & Dining" section.

Contextual targeting does not require a large scale or scope of user data. Contextually-targeted ads are generally not individualized to specific users (although the ads often differ by geography and/or time slot). Likewise, the accuracy and effectiveness of contextual ad targeting does not typically depend on the scale of users involved, or on the amount of data available on individual users. Rather, many contextual ads depend primarily on how well the content of the webpage the user visits matches an advertiser's targeting goals. To the extent there is differentiation among ad tech vendors in terms of their ability to implement contextual targeting, it is more likely to stem from innovative methods of analyzing the content of a publisher's page (*e.g.*, superior natural language processing) than the application of user data.

Audience-targeted display ads: in contrast to contextual targeting, audience-targeted display ads are targeted to individual users based on user data. There are three main mechanisms for audience-based targeting: (1) remarketing (sometimes called "retargeting"), which targets a particular user based on prior interaction with the advertiser; (2) demographic/interest segmentation, which targets a group of users based on known attributes; and (3) behavioral targeting, which attempts to target the users most likely to interact with a display ad (*e.g.*, click on an ad or purchase a product).

Remarketing: this type of display ad targets users who have had prior interaction with an advertiser's products or services.[139] The concept of remarketing is based on identifying a particular consumer's prior online activity and show the user a relevant ad based on that information. An example of remarketing is a user that visits an online retailer's website, browses a particular product, and then is shown an ad for that product on a different website. Another example is a user that clicks on an ad and is subsequently shown an ad for the same or similar product on another website.

Remarketing of display ads is based on tracking a particular user's prior interaction with a specific website or app. On desktop computers, user tracking is generally accomplished through the

---

[136] "Two-Thirds of Australia's Digital Ads Bought Through Programmatic or Ad Networks in 2016," Nielsen, 3 February 2017, available at https://www.nielsen.com/au/en/insights/news/2017/australia-digital-ads-in-2016.html.

[137] Ratko Vidakovic, "Display Ads: How Direct Buys & RTB Interact," Marketing Land, 4 March 2013, available at https://marketingland.com/display-ads-how-direct-buys-and-rtb-interact-34483.

[138] For example, news publisher sites are typically divided into sections (such as Food and Dining, Arts and Entertainment, Events, Kids, Automotive). See, *e.g.,* Kenny Katzgrau, "10 Advantages That Small Publishers Have Over Tech Giants in Selling Ads," MediaShift, 6 April 2018, available at http://mediashift.org/2018/04/10-advantages-publishers-silicon-valley/.

[139] See, *e.g.,* Amber Usmani, "Amazon's New Display Retargeting Ad Product to Rival Criteo," Elite SEM, 3 July 2018, available at https://elitesem.com/blog/amazons-display-retargeting-ad/; "What is ReTargeting and How Does it Work?" ReTargeter, available at https://retargeter.com/what-is-retargeting-and-how-does-it-work/.

collection of user browsing information, typically through the use of "cookies." Cookies are small text files that are placed on users' browsers. The firm placing the cookie can assign an ID to the cookie, and "read" the cookie ID on subsequent website visits to identify the same browser.[140] For example, an advertiser may place a cookie on a user's browser when it serves an ad on that browser for the first time. Later, the same user (or browser) may visit a different website where the same advertiser serves ads. The advertiser receives cookie information as part of the ad request, and may serve a remarketing ad to the user.

On mobile devices, tracking of user activities is accomplished through mobile advertising IDs, which are similar to cookies.[141] Mobile advertising IDs provide the same functionality as cookies do in desktop settings, namely to identify a particular user.[142] Each mobile device has a unique mobile advertising ID, which is an anonymous identification number that can be used by app developers and mobile advertisers to identify a mobile app user. In addition to facilitating the tracking of users' mobile browsing activity, mobile advertising IDs can also be used to identify users' app activity to target in-app ads. Ad tech firms that specialize in in-app advertising use mobile advertising IDs to target display ads on mobile devices.

A company can target remarketing ads to a user if it can identify a user's interaction with a product and then track that user's subsequent browsing, often via cookies or mobile advertising IDs.[143] Multiple remarketing companies can set cookies or access mobile advertising IDs at the same time and on the same sites and devices. Any of those companies can therefore target a particular user with a remarketing ad. As a result, many companies offer remarketing services, including Criteo, Facebook, Amazon, and AdRoll.

While additional data points can be beneficial for remarketing, remarketing does not require building extensive interest profiles from users' browsing activity. In addition, the availability of user data over a long period of time (*i.e.*, a long browsing history) is unlikely to be useful for this type of ad targeting for most products because the intent signaled by browsing activity is likely to be short-lived. The longer ago that a user looked at a pair of shoes, for example, the less likely that a remarketing ad will lead to the user taking a desired action (*e.g.*, a conversion).

Demographic/interest segmentation: another type of display ad targeting is referred to as segmentation. Segmentation targeting involves grouping users ("segments") according to some shared set of characteristics.[144] The two main dimensions of segmenting users are (1) demographic segmentation and (2) interest segmentation.

---

[140] See, *e.g.*, Google, Google Privacy & Terms, available at https://policies.google.com/privacy/key-terms#key-terms.

[141] Paulina, "Mobile Advertising IDs Or Finding The Right Mobile Users," AdSquare, 14 June 2017, available at https://www.adsquare.com/mobile-advertising-ids-or-finding-the-right-mobile-users/.

[142] Mitchell Reichgut, "Advertiser ID Tracking And What It Means For You," Forbes, 16 May 2016, available at https://www.forbes.com/sites/onmarketing/2016/05/16/advertiser-id-tracking-and-what-it-means-for-you/#bc6997218bf0.

[143] In some instances, additional user data signals may be employed to determine the optimal frequency and timing to show the retargeted display ads to a particular user. But many firms have access to such data, as discussed below.

[144] "Customer Segmentation & Targeting – A Guide," Digital Marketing Institute, 24 April 2018, available at https://digitalmarketinginstitute.com/en-us/blog/2018-04-24-customer-segmentation-targeting-a-guide.

Demographic segmentation involves targeting a group of users by demographic attributes such as age, gender, and income.[145] For instance, a particular advertiser may want to target males aged 18-25. There is a limited role of user data scale or scope for purposes of demographic segmentation because determining the demographic segment of a particular user does not require having data on *other* users. For example, targeting the audience segment of 25-year-old males merely requires knowing, for each user, their age and gender. As discussed in Section IV.A below, many ad firms have access to such data.

Interest segmentation involves targeting groups of users by topics or subjects of interest. For instance, an advertiser may want to target newlyweds, or auto enthusiasts. Interest segmentation does not require scale across users, but rather is based on identifying whether a particular user fits a specific interest group. For instance, targeting the "auto enthusiast" segment merely requires some means of classifying each user as an auto enthusiast or non-auto enthusiast. The primary source of user data for interest-based targeting is a user's web browsing or app activities—*e.g.*, the type of websites or ads with which the user interacts.[146]

With regard to data scope, having multiple data points on a particular user may be valuable in order to better infer interest segments, because having a variety of signals about a particular user may allow for more effective predictions about their interests or intent. As a simple example, knowing a user visits car manufacturer websites is useful in classifying the user as an auto enthusiast, but knowing that the user also browsed car review websites might result in a "stronger" classification of the user as an auto enthusiast. However, as discussed below, many firms have access to and utilize substantial user browsing data, as well as other types of data, for targeting display ads based on interest segmentation.

Behavioral targeting: the goal of behavioral targeting is to identify the set of users that are most likely to interact with a display ad or undertake a desired action (*e.g.*, click on an ad, sign up for a newsletter, or purchase a product).[147] There is some value to having user data scale and scope in predicting user responses to being shown an ad, but there are diminishing returns to scale and scope from such additional data.[148] Moreover, as discussed below, many firms have substantial data scale and scope of user online behavior data for purposes of targeting display ads, including for behavioral targeting.

2. Many ad tech firms have access to a large scale and scope of user data

Many ad tech firms have access to a wide variety of user data to use in targeting display ads. As discussed, user data is not exclusive to any one firm, and many different firms have access to data on the same users. Thus, the fact that one firm collects and utilizes user data does not preclude any other firms from doing the same. The wide availability of user data is also enhanced by user multi-homing, as well as the non-rivalrous nature of user data.

---

[145] "Customer Segmentation & Targeting – A Guide," Digital Marketing Institute, 24 April 2018, available at https://digitalmarketinginstitute.com/en-us/blog/2018-04-24-customer-segmentation-targeting-a-guide.

[146] Phil Frost, "How to Show Your Ads Only To Your Ideal Prospects," Main Street ROI, 25 April 2012, available at https://www.mainstreetroi.com/interest-based-advertising/.

[147] See, *e.g.,* Yandong Liu et al., *Finding the Right Consumer: Optimizing for Conversion in Display Advertising Campaigns*, WSDM'12, 8 - 12 February 2012, available at https://www.cs.cmu.edu/afs/cs/Web/People/yandongl/papers/wsdm2012.pdf.

[148] Ye Chen et al., *Large-Scale Behavioral Targeting*, KDD'09, 28 June - 1 July 2009, available at https://www.cc.gatech.edu/~zha/CSE8801/ad/p209-chen.pdf.

User browsing and app activities data: the primary source of data that Google uses for targeting display ads on third-party sites or apps is user interaction with sites or apps collected through tags, pixels, and other signals, and then analyzed and aggregated for targeting purposes using cookies or other identifiers like mobile advertising IDs. Many ad tech firms (and publishers) have access to the same type of data. Virtually all online service providers, and many third parties (including ad intermediation service providers), collect information on users' browsing and other activities and utilize cookies, mobile advertising IDs, and/or other tracking technologies to target ads.

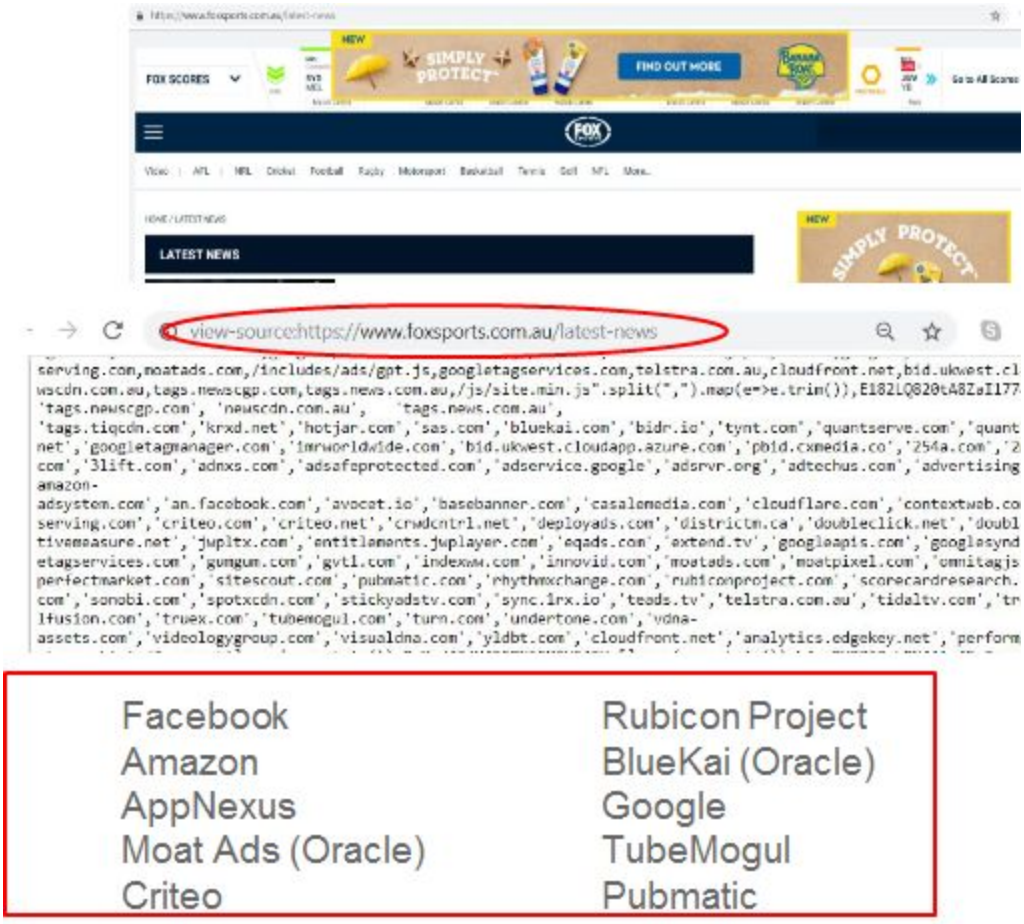
The ubiquitous nature of ad tech firms collecting user browsing data can be seen in the large number of cookies present on many websites. On average, top websites have around 20 third-party tracking firms placing cookies (in addition to the publisher also likely using cookies to track the user),[149] with popular news sites in some countries having more than 100 cookies per page.[150] As an example, Figure 2 below shows a portion of HTML code of a webpage in *FoxSports* Australia. The code shows the names of parties tracking activity of a user—including Facebook, Amazon, AppNexus, Rubicon Project and BlueKai (an ad tech firm owned by Oracle), as well as several others. Such firms may place cookies and/or read cookie data either by showing ads or placing "pixels" which are tiny "tags" embedded in the webpage or other firms' ads.[151]

[149] Steven Englehardt and Arvind Narayanan, *Online Tracking: A 1-million-site Measurement and Analysis*, Princeton University, January 2016, available at http://randomwalker.info/publications/OpenWPM_1_million_site_tracking_measurement.pdf.

[150] Timothy Libert and Rasmus Kleis Nielsen, "Third-Party Web Content on EU News Sites: Potential Challenges and Paths to Privacy Improvement," Reuters Institute and University of Oxford, May 2018, available at https://reutersinstitute.politics.ox.ac.uk/our-research/third-party-web-content-eu-news-sites.

[151] Preliminary Report at 83.

Figure 2: FoxSports Australia– Third-Party Tracking Firms



Various ad tech firms also utilize other methods of tracking users. One method is device "fingerprinting," which is a technique of identifying a device by its unique technical configuration (which may include information such as the browser version, time zone settings, the operating system, and plug-ins) in conjunction with an IP address.[152] By combining these device attributes, it is possible to create a unique "fingerprint" for each device to identify a user and user actions without the use of cookies or mobile advertising IDs.[153] Device fingerprinting raises privacy concerns, and Google does not use that approach for targeting ads. Many other ad tech firms do, however, and it illustrates the many ways firms can obtain targeting data.

Location data: many online providers and app developers collect and use locational information from Android devices, as well as devices for other mobile platforms such as iOS. This information allows ad tech firms and app developers to track user location, and serve ads based on this data (or sell data to

[152] See, *e.g.*, Katarzyna Szymielewicz and Bill Budington, "The GDPR and Browser Fingerprinting: How It Changes the Game for the Sneakiest Web Trackers," Electronic Frontier Foundation, 19 June 2018, available at https://www.eff.org/deeplinks/2018/06/gdpr-and-browser-fingerprinting-how-it-changes-game-sneakiest-web-trackers.

[153] See, *e.g.*, MediaMath, Privacy Policy, available at http://www.mediamath.com/privacy-policy/.

other ad tech firms).[154] For instance, companies such as the Weather Channel collect location data from users of their app, and sell the data to ad tech firms and advertising agencies, which may use this data to target ads based on the location of the users (*e.g.,* users that are near a retailer's location).[155] Thus, Google is not unique in having access to mobile device location data.

Other user data: ad tech firms (and publishers) have access to other sources of user data used in targeting display ads. Such data may be collected directly by the website publisher from its users, by the advertiser, or by another ad tech firm. The combination of all these data sources provides large scale and scope of user data for many ad tech firms that compete with Google.

Many publishers collect so-called "first-party" data from their users, which may be shared with ad tech firms (or advertisers). When a user visits a webpage, some basic information such as the browser type, IP address, operating system, and webpage URL are available to the website publisher.[156] Publishers also may collect other information from users for targeting purposes. For example, users may provide their age or birthday when registering for a login on a publisher site. Publishers may also collect information about a user's behavior, actions, or interests through the user's activity on a publisher site.[157] Publishers may sell or trade this information to third parties.[158] Social media sites collect demographic and behavioral data from their users through user profiles and user interactions with the website, including information on a user's personal background (*e.g.,* family and employment status) and a user's personal interests.[159]

In some instances, data is not collected from users directly by the firm utilizing it, but instead is provided by other ad tech firms such as data management platforms (DMPs) and "data brokers." This data is sometimes called "third-party" data. DMPs and data brokers are used by various participants of the online advertising industry, including advertisers and publishers, as well as DSPs and SSPs. Data brokers collect data on individual users from a variety of sources, and resell this information to other firms for the purposes of ad targeting.[160] DMPs can help firms collect and analyze data from their own users, and can also supply additional user data (often obtained through data brokers), to aid in ad

---

[154] Christopher Mims, "Your Location Data Is Being Sold—Often Without Your Knowledge," The Wall Street Journal, 4 March 2018, available at https://www.wsj.com/articles/your-location-data-is-being-soldoften-without-your-knowledge-1520168400.

[155] Elizabeth Dwoskin, "In Digital Ads, It's Location, Location, Location," The Wall Street Journal, 6 January 2014, available at http://blogs.wsj.com/digits/2014/01/06/in-digital-ads-its-location-location-location/.

[156] Timothy Libert and Rasmus Kleis Nielsen, "Third-Party Web Content on EU News Sites: Potential Challenges and Paths to Privacy Improvement," Reuters Institute at the University of Oxford, 25 May 2018, available at https://timlibert.me/pdf/Libert_Nielsen-2018-Third_Party_Content_EU_News_GDPR.pdf.

[157] Lexie Pike, "What Is First Party Publisher Data?" SpotX, 14 April 2016, available at https://www.spotx.tv/resources/blog/product-pulse/what-is-first-party-publisher-data/.

[158] Adam Uzialko, "How Businesses Are Collecting Data (And What They're Doing With It)," Business News Daily, 3 August 2018, available at https://www.businessnewsdaily.com/10625-businesses-collecting-data.html.

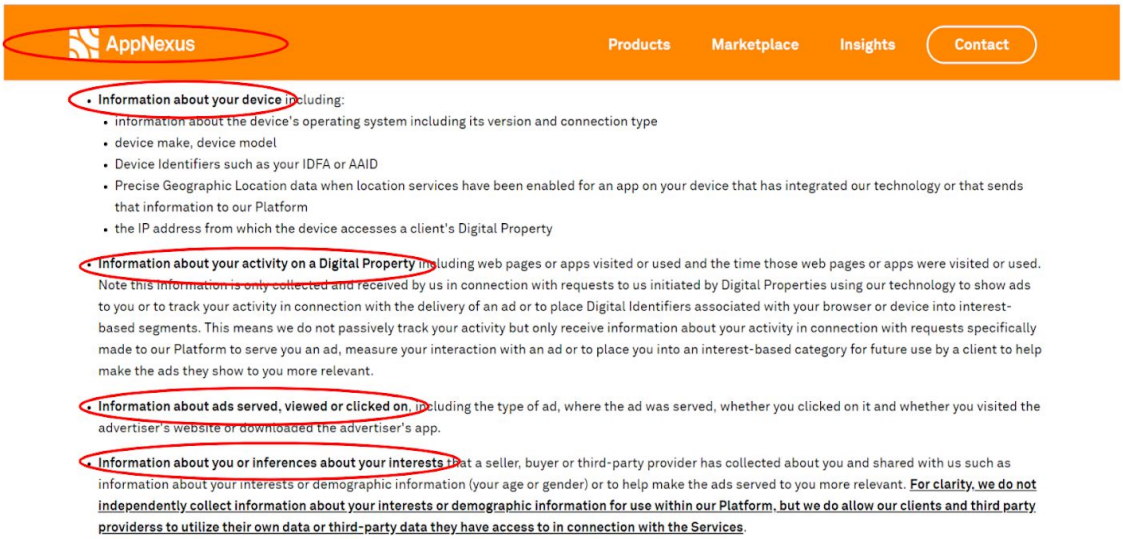
[159] Hayley Tsukayama, "Facebook updates data use policy," The Washington Post, 29 August 2013, available at http://www.washingtonpost.com/business/technology/facebook-updates-data-use-policy/2013/08/29/3f0faa08-10cd-11e3-b4cb-fd7ce041d814_story.html.

[160] Yael Grauer, "What Are 'Data Brokers,' and Why Are They Scooping Up Information About You?" Vice, 27 March 2018, available at https://motherboard.vice.com/en_us/article/bjpx3w/what-are-data-brokers-and-how-to-stop-my-private-data-collection.

targeting. The types of information include lists of names, email addresses, online browsing history, interests and offline activity, purchase histories, loyalty card data, and demographic data (*e.g.*, age, ethnicity, education level, income, number of children).[161] DMPs help advertisers pull together all the available data in the most effective way for use in ad targeting, for instance by collecting and combining both first-party and third-party data that a DSP can use to create a particular audience profile desired by the advertiser. As discussed below in Section IV.C, there are many DMPs providing services in the display ad tech stack.[162]

Ad tech firms commonly combine these various types of user data (browsing and app activity, location, and other data) from various sources to use in targeting display ads. Thus, even if a firm (*e.g.*, an advertiser or publisher) does not have a large amount of its own "first-party" data to use for ad targeting, the firm can access third-party data via data brokers and/or DMPs, and even combine it with its own first-party data, to improve ad targeting capabilities. For instance, an automobile manufacturer can use its own list of consumers that have purchased its cars in the past, and combine this with third-party data on browsing activity (obtained via a DMP and/or data broker) to determine which customers on that list are actively researching a new car purchase. As described in a recent article: "companies are using a cornucopia of sources to capture and process customer data on metrics, from demographic data to behavioral data."[163] For example, AppNexus combines data on user devices, browsing activity and app activity, ads seen and clicked on, demographic data, and inferences about user interests (see Figure 3 below).[164]

Figure 3: AppNexus – Types of User Data Utilized



[161] *Id*.

[162] Google's DMP used to be called Google Audience Center, and is now integrated into Google Marketing Platform.

[163] Adam Uzialko, "How Businesses Are Collecting Data (And What They're Doing With It)," Business News Daily, 3 August 2018, available at https://www.businessnewsdaily.com/10625-businesses-collecting-data.html.

[164] AppNexus, Platform Privacy Policy, available at https://www.appnexus.com/platform-privacy-policy.

Similarly, AT&T's recently rebranded ad tech service Xandr will combine user data from AT&T's phone, Internet and TV subscribers with Time Warner's subscriber data, and the data AT&T will have access to by virtue of its acquisition of AppNexus.[165] News Corp's *The Australian* can collect its users' IP addresses, browser types, email addresses, and browsing activity.[166] News Corp may then combine the user data it collected with "information collected from other trusted businesses with whom [the user] also ha[s] a relationship or through public sources."[167] Firms also may choose to enhance their access to user data through acquisition--several large advertising companies have spent billions on data-focused transactions in recent years. Publicis announced plans in April of this year to acquire Epsilon for $4.4 billion,[168] Interpublic spent $2.3 billion to purchase Acxiom's data capabilities in 2018,[169] and Dentsu Aegis Network invested $1.5 billion to take a majority stake in Merkle.[170] Other examples of ad tech firms besides Google that combine various sources of user data to achieve larger user data scope are shown in Table 1 below.

Table 1: Examples of Ad Tech Firms with Substantial User Data Scope

Ad Tech Firm	Data Scope
Oracle	40,000 data attributes 40 third-party data providers
Criteo	10,000 websites share data with Criteo
Oath	Yahoo!/AOL websites At least 15 third-party data partners (including BlueKai, Experian, Equifax, Nielsen)
AppNexus	34,000 publishers and 177,000 brands transact in marketplace
Rubicon Project	Processes ads on over 1 million websites, across 50% of top websites and apps
MediaMath	Over 60 data providers
LiveRamp	500 partner integrations 150 data providers

Source: Ad tech firms' websites and privacy policies.

[165] Ronan Shields, "AT&T Unveils Xandr, Its Newly Rebranded Ad-Tech Unit," Adweek, 25 September 2018, available at https://www.adweek.com/programmatic/att-unveils-xandr-its-newly-rebranded-ad-tech-unit/.

[166] News Corp Australia, Data Usage Policy, available at https://preferences.news.com.au/data; News Corp Australia, Privacy Policy, available at https://preferences.news.com.au/privacy.

[167] News Corp Australia, Data Usage Policy, available at https://preferences.news.com.au/data.

[168] Megan Graham, "Publicis announces $4.4 billion deal to acquire data marketing company Epsilon", CNBC, 14 April 2019, available at https://www.cnbc.com/2019/04/14/publicis-to-buy-epsilon-for-4point4-billion.html.

[169]" Interpublic to buy Acxiom's marketing solutions division for $2.3 billion," Reuters, 2 July 2018, available at https://www.reuters.com/article/us-acxiom-divestiture-interpublic-grp/interpublic-to-buy-acxioms-marketing-solutions-division-for-2-3-billion-idUSKBN1JS2IQ.

[170]Suzanne Vranica, "Dentsu To Buy Majority Stake in Data Marketing Firm Merkle," Wall Street Journal, 8 August 2016, available at https://www.wsj.com/articles/dentsu-to-buy-majority-stake-in-data-marketing-firm-merkle-1470650402.

Many ad tech firms besides Google also have access to data from a very large number of users (*e.g.*, large user data *scale*). For instance, Criteo claims that is has built "the world's largest open shopper data set" covering "72% of online shoppers globally."[171] As shown by the examples in Table 2 below, a number of ad tech firms have user data scale of over one billion users.

Table 2: Examples of Ad Tech Firms with Substantial User Data Scale

Ad Tech Firm	Data Scale
Oracle	2 billion user profiles worldwide
Criteo	1.4 billion online shoppers
Oath	1.3 billion users
MediaMath	1 billion consumer records
AdRoll	1.2 billion online shoppers
LiveRamp	800 million recently matched cookies

Sources: Ad tech firms' websites and press releases; Forbes.com; MMAglobal.com.

Lastly, the effectiveness of display ad targeting depends largely on the ability to utilize data, not simply the scale or scope of data to which the provider has access.[172] Online advertising firms are investing substantial engineering and technological resources to research and develop improved ways of utilizing user data and other information to offer more effective display ad targeting solutions. For example, DSP MediaMath recently announced a partnership with IBM in order to utilize IBM's "Watson" artificial intelligence (AI) technology to optimize ad targeting and personalization.[173]

In summary, the limited role of user data scale and scope in display ad targeting, and the widespread availability and utilization of user data by ad tech firms besides Google, demonstrate that Google does not have a significant competitive advantage in display ad intermediation. This conclusion is supported by the fact that Google competes with a large number of ad tech firms for display advertising services, as I discuss in Section IV.C below.

[171] "Explained: Data in the Criteo Engine: Introduction," Criteo, available at https://www.criteo.com/insights/explained-data-in-the-criteo-engine/?slide=2.

[172] See, *e.g.,* Gary Eastwood, "Big data, algorithms and the future of advertising," Network World, 4 May 2017, available at https://www.networkworld.com/article/3194585/big-data/big-data-algorithms-and-the-future-of-advertising.html: "big data is less important than smart data and strategy."

[173] MediaMath, "Delivering on the Promise to Re-Architect Programmatic Advertising: MediaMath and IBM Watson Marketing Announce Media Optimizer, a New Solution Connecting AdTech and MarTech and Infusing 'AI Everywhere'," October 2, 2018, available at http://www.mediamath.com/blog/mediamath-and-ibm-watson-marketing-announce-media-optimizer/.

B. Network effects and advertiser fixed costs in the provision of display ad intermediation

The Preliminary Report also contends that Google enjoys "strong advantages" in the supply of ad intermediation services due to its "substantial" market power in search advertising, and the existence of fixed costs for advertisers.[174] The theory put forth in the Preliminary Report appears to be that because of Google's position as a supplier of search ads, advertisers must use the Google Ads platform, and "once an advertiser has incurred the fixed costs of purchasing one service through Google Ads, it may choose to purchase another service through Google Ads, in order to avoid incurring additional set-up costs."[175]

The implementation of other forms of advertising (*e.g.*, display, video, native, in-app), however, is quite different from search ad campaigns. This diminishes any economies in using Google for these other ad formats. For example, an online video ad campaign may involve different audience targeting, different ad creative, and different ROI goals than a search ad campaign. The advertiser's incremental effort to set up the video ad campaign likely would exist whether or not the advertiser uses Google's ad platform.

Moreover, as discussed above in the context of search, platform-specific fixed costs for advertisers of "setting up and operating" advertising campaigns are likely to be minimal, as are the costs of multi-homing across different ad platforms. This is particularly true given the wide availability and use of a variety of software tools which allow advertisers to manage advertising efforts across multiple platforms and channels, lowering any costs of multi-homing associated with fixed costs of setting up and operating on each ad platform. These products often incorporate marketing mix optimization tools that help advertisers allocate advertising budgets most effectively across channels and formats (including across different general and specialized search platforms, social networking platforms, and display, video, and other ad formats), and across different DSPs.[176] Some examples of these tools include Kenshoo,[177] SEMrush,[178] Centro Basis,[179] and Adobe Advertising Cloud.[180]

Consistent with the lack of significant fixed costs, advertisers often multi-home—*i.e.*, use multiple ad intermediaries at each level of the ad tech stack. In the U.S., it is estimated that advertisers use on average four DSPs to purchase publisher ad space.[181] Advertising agencies may use even more, according to some sources, and many agencies also buy ads programmatically through their own trading

---

[174] Preliminary Report at 83-84.

[175] Preliminary Report at 48.

[176] See, *e.g.*, Cesar Brea and Laura Beaudin,"The Future of Marketing Mix Optimization is Here," Forbes, 28 February 2018, available at https://www.forbes.com/sites/baininsights/2018/02/28/the-future-of-marketing-mix-optimization-is-here/#585a3af71387.

[177] "Search Marketing Magnified," Kenshoo, available at https://www.kenshoo.com.au/.

[178] "Organic Research," SEMrush, available at https://www.semrush.com/features/organic-research/.

[179] "Basis by Centro," Centro, available at http://www2.centro.net/basis-by-centro.

[180] "Adobe Advertising Cloud Features," Adobe, available at https://www.adobe.com/advertising/adobe-advertising-cloud-features.html.

[181] "Average Number of DSPs Used by US Advertisers, Jan 2016-April 2018," eMarketer, 29 May 2018, available at https://www.emarketer.com/Chart/Average-Number-of-DSPs-Used-by-US-Advertisers-Jan-2016-April-2018-among-largest-100-advertisers-on-Pathmatics-platform/219189.

desks in addition to doing so through third party DSPs.[182] Similarly, on the supply side, publishers also typically multi-home across various SSPs, with the average publisher using around six SSPs.[183] As discussed above, advertisers generally multi-home because ad tech firms offer differentiated services and targeting capabilities.[184] Advertisers choose a set of ad tech firms whose service offerings best meet their specialized needs. For example, some DSPs specialize in desktop, others in mobile.[185] This multihoming diminishes any network effects.

The Preliminary Report explains that the "ACCC has not yet reached a concluded view of these effects because it requires: more information about the extent to which advertisers 'multi-home' or use multiples of the same type of intermediaries, and more evidence about the extent to which advertisers could switch to rival intermediaries…"[186] Indeed, evidence of multi-homing by advertisers across different ad intermediation providers fundamentally undermines the theory suggested in the Preliminary Report that fixed costs for advertisers give Google "strong advantages" in the supply of ad intermediation services.

Notably, there is also interoperability between different components of the display ad tech stack. Interoperability refers to the ability of two different software programs to share information and operational functions.[187] As noted in the Preliminary Report, "ad tech tools can be configured to interact with each other in different ways and are generally interoperable."[188] Interoperability between different components of the display ad tech stack enhances the ability of advertisers (and publishers) to multi-home, since advertisers (and publishers) can "mix and match" ad tech stack components from unaffiliated providers depending on their needs. Because of interoperability between different software tools in the ad tech stack, ad tech firms can specialize in providing services for a particular level of the ad tech stack—*i.e.*, advertisers or publishers. For example, Centro is a DSP that focuses on machine learning and artificial intelligence, and has been rated highly by industry observers for its "hyper-local" targeting technology and cross-device tracking capability.[189] This interoperability, like multi-homing, diminishes any network effects.

Although there is substantial interoperability between different ad tech platforms, many advertisers and publishers may choose to use Google because it (along with other ad tech firms such as

---

[182] Agencies such as PMG, US Interactive Media, and IMM use five or more DSPs. (Tobi Elkin, "How Many DSPs Does A Brand Need?" MediaPost, 24 February 2016, available at https://www.mediapost.com/publications/article/269704/how-many-dsps-does-a-brand-need.html.)

[183] Ross Benes, "Publishers Are Using Fewer Sell-Side Vendors," eMarketer, 17 July 2018, available at https://www.emarketer.com/content/publishers-purged-one-fourth-of-their-ssps-over-two-years.

[184] Karen Moked, "Programmatic Media Buying 101: What's The Difference Between DSPs & Ad Networks," Digilant, 20 June 2018, available at https://www.digilant.com/programmatic-buying-101-whats-the-difference-between-dsps-ad-networks/.

[185] Arvind Kesh, "What is a Demand-Side Platform and How to Choose One," Adbeat, available at https://blog.adbeat.com/programmatic-advertising/demand-side-platform.

[186] Preliminary Report at 84.

[187] Preliminary Report at 72.

[188] Preliminary Report at 72. News Corp also acknowledges the interoperability of Google's AdX with other ad networks: "AdX draws from networks beyond just AdSense, meaning a greater number of potential buyers." (News Corp Submission at 36.)

[189] "Demand-Side Platform," Centro, available at https://www.centro.net/solutions/basis/dsp/.

Oath) offers a set of complementary products that includes functionality throughout the ad tech stack. The Preliminary Report notes that "Google likely also has a competitive advantage because it offers intermediary services across all functional levels of the programmatic supply chain."[190] The Preliminary Report considers that there is "potential" for such "bundles" to "lessen competition in certain advertising markets."[191] This paper does not directly address the "bundling and tying" concerns in the Preliminary Report, including the contention that Google's vertical integration in the ad tech stack is "likely" to provide Google a competitive advantage.[192] It is important to note, however, that the efficiencies of purchasing several products or services from one provider, either because of convenience or superior integration, does not amount to anticompetitive "bundling and tying," wherein consumers are compelled to purchase multiple products from one firm. For instance, consumers typically purchase various products at a supermarket due to the convenience, but one would not characterize that as "bundling or tying."

Nor does Google's provision of a full range of services necessarily translate to substantial advantages in terms of the data that it can collect. The types of user data that a DSP collects, for example, are not necessarily the types of data that enhance the ability of a supply-side service to better target ads to users. And, of course, Google faces a range of full-line competitors that offer a similar set of services and can thus also capture whatever data advantages might hypothetically exist.

Moreover, Google's ability to offer a single platform for ad intermediary services is an example of the economic efficiencies of vertical integration, which are widely recognized and acknowledged to be procompetitive.[193] In fact, the Preliminary Report recognizes potential efficiencies of using an all-Google (*i.e.,* "vertically-integrated") tech stack. Such efficiencies do not imply that such benefits confer a significant competitive advantage or entrench a dominant platform. The significant competition between ad tech providers on both the demand and supply sides, as discussed below, highlights that any such efficiencies are not significant enough to inhibit competition from providers that do not supply all the components of the ad tech stack.

For these reasons, barriers to entry and expansion and network effects remain insignificant in the business of display ad intermediation.

C. The display ad tech stack is highly competitive at each level

Consistent with the fact that the Preliminary Report does not find that Google has market power in display ad intermediation, the economic evidence indicates that the display ad tech stack in Australia is highly competitive at each level. Multiple display advertising firms compete vigorously with Google at each level of the ad tech stack. This competition is evidenced not only by the large number of firms, but also by the differentiated business models offered by these ad tech firms.

---

[190] Preliminary Report at 84.

[191] Preliminary Report at 82.

[192] Preliminary Report at 84.

[193] For instance, in the AT&T/Time-Warner decision, Judge Leon noted "the recognition among academics, courts, and antitrust enforcement authorities alike that 'many vertical mergers create vertical integration efficiencies between purchasers and sellers.'" (*Memorandum Opinion in United States of America v. AT&T, Inc.*, United States District Court for the District of Columbia, 12 June 2018, citing Michael H. Riordan & Steven C. Salop, *Evaluating Vertical Mergers: A Post-Chicago Approach*, 63 ANTITRUST L.J. 513, 519 (1995)).

Demand-side Tools: many highly successful players compete to offer demand-side services to advertisers and ad agencies, including Amazon, Verizon Oath, Facebook, The Trade Desk, MediaMath, and Criteo, among many others. Some compete as "point players" offering a standalone DSP or ad network while others are vertically integrated. Criteo, for example, is a point player focused on remarketing (some in fact consider Criteo to be the inventor of this technology)[194] and was recently ranked as the "#1 player in the 'advertising software market.'"[195]

Supply-side Tools: many firms, including AppNexus, OpenX, Rubicon Project, and PubMatic, compete to provide supply-side ad placement services to publishers. As is the case on the demand side, these firms offer a range of high-quality services and have grown rapidly. For example, PubMatic was named by Forrester as a "leader" in providing programmatic advertising services to publishers, in part for its "comprehensive" ability to integrate with other ad tech platforms and offer a large number of innovative selling options.[196]

DMPs: there are a growing number of DMPs that offer data collection, analysis, and targeting services in the ad tech stack. These include long-established data broker firms such as Adobe (Audience Manager) and Oracle (BlueKai), as well as more recent entrants including LiveRamp, MediaMath, and Lotame. Lotame won an award for "Global DMP of the Year," and is an industry leader in developing innovative tools to use television viewership data for online ad targeting.[197] LiveRamp specializes in cross-device targeting via its "IdentityLink" service.[198] These and other DMPs offer their customers access to large amounts of user data. For example, MediaMath combines data from over 60 providers.[199] MediaMath offers integrated DSP services as well.[200]

These basic realities of the ad tech marketplace show that Google's access to user data has not given it a significant competitive advantage in targeting display ads. Instead, Google faces a large number of competitors.

---

[194] "Personal Retargeting in Practice. A Guest Lecture on E-Commerce by Criteo's Managing Director for the German language area," WWU Münster, available at https://www.marketingcenter.de/en/study/guest-lectures/personal-retargeting-practice-guest-lecture-e-commerce-criteos-managing.

[195] Ludovic Leforestier, "[REPORT] IDC: Criteo #1 for market share in advertising software," Criteo, 18 September 2018, available at https://www.criteo.com/insights/criteo-number-1-in-adtech-idc-Preliminary Report/.

[196] "New Preliminary Report Names PubMatic a Leader Among Sell-Side Platforms and Exchanges for Publishers," Nexus Venture Partners, 30 April 2014, available at https://nexusvp.com/press-releases/new-Preliminary Report-names-pubmatic-a-leader-among-sell-side-platforms-and-exchanges-for-publishers/.

[197] "Lotame Named 'Global DMP of the Year'," Lotame, 2 November 2017, available at https://www.lotame.com/lotame-global-dmp-year-award/.

[198] "LiveRamp Partners with Thunder Experience Cloud," MarTech Advisor, 27 September 2018, available at https://www.martechadvisor.com/news/data-management/liveramp-partners-with-thunder-experience-cloud/.

[199] "Partner Audiences," MediaMath, available at http://www.mediamath.com/audience/partner-audiences/.

[200] "MediaMath Named a Leader in Data Management Platform Preliminary Report by Independent Research Firm," MediaMath, 1 June 2017, available at http://www.mediamath.com/news/mediamath-named-a-leader-in-data-management-platform-report-by-independent-research-firm/.