CESifo WORKING PAPERS

9525
2022

January 2022

Value Creation by Ad-Funded Platforms

Gregor Langus, Vilen Lipatov

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9525*

Value Creation by Ad-Funded Platforms

Abstract

We identify features of interactions on online platforms that make an ad-funded business model attractive for the platform, but also for consumers. We then show that ad-funded platforms heavily rely on data for their ability to create value for their users. Formally, we show that data restrictions may trigger a switch away from ad-funded to fee-funded model, resulting in a loss of consumer welfare. We also argue that restricting the effort to increase data quality weakens competition to the detriment of consumers.

JEL-Codes: K210, L220, L400, M370.

Keywords: ad-funded business model, data aggregation restrictions, targeted advertising, platform competition, merchant competition, transaction costs.

Gregor Langus
CompetitionSphere
Avenue Louis Thevenet 8
Belgium – 1180 Brussels
gregor.langus@competitionsphere.com

Vilen Lipatov
CompetitionSphere
Lindenstrasse 19
Germany – 35647 Waldsolms
vilen.lipatov@competitionsphere.com

Value creation by ad-funded platforms

Gregor Langus* and Vilen Lipatov†

1 Introduction

Much of the debate in antitrust circles in recent years has been around the collection and use of personal data by online platforms. A few concerns have been put forward, among them that access to data may provide incumbent platforms with unsurmountable competitive advantage and allow them to deter rivals from entering a market. There have also been concerns that access to data may help platforms to engage in harmful price discrimination or steering of consumers, as well as that access to data may enable platforms to systematically violate users' privacy.

Economic theory has not been able to provide a universal policy advice in relation to the antitrust concerns with the use of data by platforms. In general, access to and use of better data may either harm or benefit platform users and the competitive process. And while this ambiguity suggests a case-by-case approach to antitrust concerns, the policy makers' instincts have been that a systemic solution must be found. Accordingly, various antitrust authorities have announced plans to limit the ability of large online platforms to collect and employ user data across various services. Because the business of ad-funded platforms is built around targeted advertising that relies on the collection and aggregation of user data across various services, the spotlight has been on such platforms.[1]

Our contribution in this paper is twofold. First, we contribute to the ongoing policy discussion on restrictions on data collection and aggregation. We do so in two steps by: (i) identifying

[1] See, e.g., Caffarra et al (2020), Mazzucato et al (2021), Economides and Lianos (2021) for contributions that pay special attention to the ad-funded platforms and their use of data.

the specific features of interactions facilitated by a platform that likely lead to the choice of ad-funded business models and explaining how, due to these features, abandoning this monetization strategy would result in substantial loss of welfare; (ii) showing that ad-funded platforms heavily rely on data for their ability to create value for their users.

Second, we contribute to the economic literature on targeted advertising in the presence of network effects by formally modelling the endogenous choice of a business model in various settings. With a standalone platform, we formulate conditions under which data restrictions trigger a switch away from ad-funded to fee-funded model, accompanied by a substantial loss of consumer welfare. In a platform competition setting, we identify the conditions under which data used in targeted advertising is pro-competitive, so restricting the effort to increase its quality weakens competition to the detriment of consumers.

In the following, we summarize our analysis in more detail. We start with the broader policy discussion and then focus on specific aspects of the problem that we modelled formally.

It is true, as some commentators claim,[2] that proposed restrictions on data aggregation will limit the potential for harmful use of data by large online platforms. However, these same restrictions will also limit the ability of these platforms to use data to create value and share it with consumers. In particular, the restrictions will likely especially harm ad-funded (rather than fee-funded) platforms by hindering their ability and incentive to create value and compete for users. This is because ad-funded platforms arise to facilitate interactions that are particularly prone to generating positive externalities, where a single interaction increases value for non-participating users as well. Harnessing these externalities requires user data: the better the data available, the better these externalities can be generated and mobilized for value. An example is when a user posts a review, reacts to a content, or starts an interest group. Such actions provide value for users beyond those directly involved in that particular interaction. Not all users derive equal value from these engagements and data helps to match value with users.[3]

Ad-funded platforms are multi-sided platforms that subsidize the consumer side, while adopting advertising as their primary monetizing strategy. Consumers primarily use these platforms to access a service or content that is not related to advertisement, or only indirectly so (e.g., social or professional networking, AV content, or internet search). Advertisers typically pay per advert click and are willing to pay more when the conversion rate (the share of clicking users that end up purchasing the advertised good or service) increases.

Aggregation restrictions will affect the ability and incentive for ad-funded platforms to create value for users in at least two ways. First, ad-funded platforms, as also some fee-funded platforms, use data to improve services to their participants in ways not related to advertising. Insofar as

[2] See, for instance, Economides and Lianos (2021)

[3] There is a number of different externalities that result from interactions on a platform, both on the same platform sides and across sides. We think of the activities (e.g., posts) of one user that are directly enjoyed by other users (e.g., by watching/reading the post), but also the reaction of these other users (e.g., likes) that may bring satisfaction to the activity-generating user as important examples of same-side externalities. An interested user getting access to the information about wider range of products through higher participation of merchants on the other side of a platform is a classic example of cross-side externality. In our formal modelling, we restrict externalities to another classic example: users directly value participation of other users.

they use aggregated data for direct service improvements, it is straightforward that restrictions on data aggregation will harm the platforms' ability to create value for their stakeholders including advertisers, content providers, and consumers. Platforms also will not be able to internalize network effects as effectively as without restrictions, as the lower expected value of single interaction will be amplified by the lesser ability to orchestrate value from these interactions for the group of users as a whole.

Second, with restrictions, the ad-funded platforms will not be able to create as much value for advertisers as without restrictions. The data that ad-funded platforms aggregate across various services help them to target ads to individual users, thereby increasing conversion rates. Higher conversion rates translate into higher per-user revenues. This, in turn, means that the ad-funded platform has stronger incentives to attract additional users by offering them higher quality services.

Data aggregation restrictions, therefore, will, in some circumstances, render an ad-funded business model not viable. When the ad-funded model is a preferred one in the first place, a shift to a fee-funded mode will reduce consumer welfare by raising prices for platform service, compromising its quality by lowering the quality of many of its interactions, reducing the user coverage and the welfare-enhancing network effects with it, and relaxing competition.

When the term 'ad-funded platforms' is understood in the narrower sense of charging strictly zero price for the platform's services (as we do in our formal modelling), the change of business model away from ad-funded causes yet another loss to social welfare. Free-to-use platforms eliminate transaction costs related to charging user fees. Switching to positive fees will therefore result in the loss of savings in transaction costs for users who choose to stay on the platform. Moreover, many users who would have still participated at low prices in the absence of transaction costs will actually choose not to use the platform services when these costs are present, further reducing the scope for network effects and decreasing welfare.

For a recap, consumers enjoying ad-funded platforms' services may be harmed by data aggregation/collection restrictions in the following ways: (1) the price for each user may be increased and the quality of service decreased; (2) some users will not find it attractive to use the service at all; (3) transaction costs will be increased for each participating user; (4) some utility will be lost due to less scope for network effects, and (5) competition among platforms may be weakened.

To confirm our intuition, we consider a specific setting for formal modelling, making a number of simplifying (and limiting) assumptions. First, we assume away any privacy issues that availability of better data to a platform may raise. We think that privacy belongs to a different debate and any demonstrable harm arising from data collection and aggregation would have to be incorporated to the analysis for a calculation of a net welfare effect. Second, we do not model advertisers explicitly. Instead, we assume that higher data quality results in more precise targeting for which a platform can charge a higher per-click advertisement fee. Third, we do not take into account the effect of data on the quality of service and, with that, user experience that platforms provide.

The latter two assumptions make our analysis conservative in the sense of the result that consumers are likely to suffer from data-related restrictions. More data through better targeting could encourage merchant competition by making entry of niche firms easier, as we formally illustrate in section 6. Also, more data would likely mean better core services on the platform regardless of any improvements in targeted advertising.

To be clear, however, our aim is not to show that more data collection is good for consumers in general. Instead, it is to shed some light on a particular channel through which consumers are harmed when targeting is restricted. This channel consists in restriction of the ability and incentive of platforms to create value for their users due to the limitation of their monetization options (a decrease in the ad fee, in particular). Additionally, we also shed some light on what factors make platforms choose one or another monetisation strategy in the first place and why forcing a change in such strategy can be welfare-decreasing.

We show, in a formal setting with a single platform, that three factors play an important role in the decision to set zero price (and therefore choose the ad-funded business model): (i) strength of network externalities; (ii) targeting precision with derived ability to gain advertising revenue; and (iii) user transaction costs associated with non-zero subscription fees.[4] An increase in any of the three factors results in higher likelihood of choosing zero fee for platform services provided to consumers.

We also show, in this setting, that reducing the ability of an ad-funded platform to aggregate data (or some other restriction that degrades the platform's user data quality) may trigger a change in the business model from ad-funded to subscription-funded, that, in turn, will cause a sharp drop in consumer surplus. This drop is caused by four forces that are at play simultaneously. First, some (maybe many) users choose not to participate in the platform activities at a positive fee. These users lose utility from the platform services and, potentially, from matches with relevant advertisers. Second, the reduction in the number of participating users reduces the utility of users who remain on the platform. This works via network effects which are weakened by the egress of a share of users. Third, the users that remain must also pay a positive subscription fee. Finally, these users bear transaction costs related to the payment of a positive fee.[5]

The same mechanisms are at play in the competitive setting.[6] In addition, a data quality degradation will lead to weakening of competition. This is because a decrease in ad fee triggers an optimal increase in subscription fee to which a competitor platform reacts by optimally increasing subscription fee as well.

The rest of the paper is structured as follows. In section 2, we discuss the contribution of ad-funded business models to consumer welfare and how it could be severed with data-related restrictions. In section 3, we provide a brief overview of the economic literature on targeted advertising. Section 4 contains a formal setup and results of our analysis for standalone platforms.

[4]While allowing for positive subscription prices in our analysis, we assume that prices below zero are not feasible. This can be motivated, for example, by considering that transaction costs involved in paying customers for the use of the platform are prohibitive; or by the necessity to exclude opportunities for disingenuous use of the platform that paying users for participation would create.

[5]This may be the cost of providing payment details such as credit card which may be prohibitive for certain categories of users; there are also psychological costs associated with paying positive price.

[6]Unless all the users leaving the platform that introduces a positive price join competitors' platforms. In this case, the first and the second points in the list of how restrictions harm consumers are not relevant.

In section 5, we formally analyse platform competition and, in section 6, we present a way for targeting to increase merchant competition. All derivations are left to appendix.

2 Platforms choose ads as monetization business model when substantial value would otherwise have been lost for society

2.1 Platforms choose functionalities and monetisation mode to maximise value creation

To succeed in attracting users, online platforms must offer them a valuable service. Some platforms do that directly, by acquiring, organizing, and delivering information or content to users. Examples are internet search, e-books, and curated AV streaming platforms, such as Netflix. Alternatively, a platform can mediate social interactions or commercial transactions between various classes of platform users. These online intermediaries leverage the internet and other technologies to substantially reduce the cost of interactions between third party participants.[7]

To minimize negative externalities associated with its service or enhance positive ones, platforms provide additional functionalities depending on the nature of interactions between the participants including user-to-user or user-to-content matching, content curation, customer reviews, identity verification, quality verification, payment, and dispute settlement mechanisms. Whether the platform will attract users and keep them engaged on the platform thus depends in large part on how well it can manage network effects. Accordingly, the platforms' choices of the functions it offers and the pricing structure it adopts—i.e. its choice of the business model—reflect the concern with internalization of network effects.

For many online platforms—like Google Search, Youtube, Facebook, LinkedIn, Snapchat, or Twitter—monetization primarily consists in charging advertisers for their ads, either per click or by impressions. The ads are positioned on a web page or in a mobile app. In contrast, fee-funded platforms monetize their primary services either directly via a subscription or transaction fee charged to the end user, e.g., AV streaming platforms like Netflix, Google Play Movies, or via a commission to a merchant (e.g., online marketplaces like Amazon, Alibaba, eBay).[8] Whatever the choice of the business model, it is not a random one; instead, it is driven by the nature of the service that the platform provides to its users.

[7] In interactions on platforms at least two classes of users are involved. On social networks, one class of users post messages or AV content while the other class receives messages and engages with posted content; the roles can be reversed when a user from the latter group replies with a message or by posting responsive content. On platforms that mediate commercial transactions—such as online auction platforms, marketplaces, and app stores—users can be divided into sellers and buyers; also here a single user can play both roles, for example, when buying and selling products on eBay.

[8] While in the context of online marketplaces the sellers also advertise their products or services, those advertisements are part of the platform's core value proposition to mediate in the search and sale/purchase of products and services and are thus integral to the participation decision. In contrast, in the case of ad-funded platforms users are not there primarily to buy a product or browse through ads in order to get a sense of what is on offer.

2.2 When interactions on the platform possess certain characteristics, charging users a positive fee would result in a material loss of value

Some characteristics of interactions on a platform lead to the choice of an ad-funded business model. These are

1. a large number of interactions, most of which generate a very small value;
2. uncertainty and heterogeneity of value across interactions;
3. asymmetry of the parties to an interaction in terms of the value they get from the interaction; and
4. externalities that interactions have on users that are not directly involved in them.

In addition to the characteristics listed above, positive externalities from participation (even if participants do not engage in interactions), or the so-called audience effect, may also play a role in the choice of business model.

In the following, we explain how each feature makes setting a positive fee destroy significant value for the users. In their totality, the features may render a fee-funded model not viable at all.

2.2.1 Many low-value interactions make pay-per-transaction prohibitively costly

The low value of each interaction (feature 1 above) means that most users are not willing to pay much to initiate one. The relatively low value of each interaction to individual users makes it impractical to charge a fee per transaction as even small transaction costs could prevent individuals from initiating an interaction. This could result in a significant loss of value because of the potentially vast amount of these interactions.

Think, for example, of an internet search query. As the value of the search is often low, the user initiating a casual search is not willing to pay much for it. But even when the search does not result in much value, the overall value of search is significant over many search interactions.

2.2.2 Uncertainty and heterogeneity of value across users make 'one-size-fits-all' membership fee inefficient

A platform facing a large number of low-value interactions could, in principle, charge a membership fee, a larger amount on a monthly or yearly basis, so that transaction costs would not stop many users from joining. However, because the value is heterogenous across interactions and interacting parties (features 2 and 3), such membership fee would need to vary across users. This is often not possible due to informational and technical constraints, but also because people have a distaste for such price discrimination.

Moreover, even tailor-made (but positive) membership fees would still deter some users from joining. For some, because of transaction costs that a positive fee introduces, for others, because of the valuation below the price, resulting in both direct and indirect via network effects) loss of value, as we discuss in more detail in section 4.

2.2.3 Introducing fees for certain interactions would stifle positive externalities and may bring significant loss of value

In the context of social networks, externalities (feature 4) originate from the wish of users to keep up with their friends and acquaintances by observing their online activities, without necessarily actively posting messages or other content. Externalities generated in this way result in direct positive network effects.[9] In the case of internet search, externalities arise because the totality of searches performed by all users may improve the relevance of search results to an individual user.

Think, for example, about a low willingness to pay for posting a photo from the last weekend's family trip on a social network and, relatedly, the willingness of the followers of this family to pay for seeing the photo or posting a comment to it. At the same time, if a large number of users engage with the initial post, the overall value of the post is substantial.

Even if only a fraction of all interactions on the platform were shut off by a positive fee, this could create a disproportionate loss of value because of the externalities these forgone interactions have on other users of the service. Similarly, if participation of passive users is reduced because of fees, the lost value may be significant because of the 'audience effect'.

2.2.4 When charging users directly does not involve substantial loss of value, platforms do so

When the interactions on the platform do not possess the listed features, charging users is often a good option. For example, on platforms that mediate commercial transactions, an individual interaction/transaction often does not have strong external effects (violation of feature 4); each interaction potentially represents a significant value (e.g., purchasing consumer goods, violation of feature 1); the value for the seller is certain (price of the good, violation of feature 2). Moreover, additional transaction costs of payment to the platform by a seller are unlikely to deter any purchase as such payments can be aggregated across transactions.

2.2.5 Ad-funded platforms face a dilemma: forego the value creation or find a way to monetize without charging users

From our discussion, it is apparent that platforms that enable interactions characterized by features listed above cannot charge consumers positive price without significantly compromising value creation. A possible way out could be to attract another side to the platform that would be able to charge consumers without causing a substantial loss in welfare.

This is exactly what ad-platforms do. On platforms that choose this business model, users do not have to pay fees for membership or transactions and thus their participation and engagement are not subdued. The platform services are paid for by the advertisers who sell their goods/services to platform users. This way, the decision to engage in interactions mediated by the platform remain undistorted and the value created on the platform is maximized.

[9] However, as long as a user can obtain access to his core group of connections, further expansions of network size may provide a small additional value only. See, e.g., Engels (2016).

2.3 Ad-funded business model preserves the value of a certain type of consumer interactions

2.3.1 Merchants do and consumers may benefit from advertising

The ad-funded business model not only allows maximisation of user interactions on the platform and associated with it value creation. Users of ad-funded platforms also interact with advertisers through ads. Sometimes, both parties benefit from this interaction. The value to the advertiser (seller) is straightforward. This value is created when an interaction results in either a direct or delayed purchase (the latter is often labeled as creating product or brand awareness).

On the user side, things are a little bit different. When a user clicks on an ad, or otherwise actively engages with it, it is reasonable to assume that the ad brings positive, even if small, expected utility to the user. This can be motivated by a revealed preference argument, which, in this context, is that the consumer must have perceived the ad as potentially valuable (e.g., in terms of information that he or she is interested in) or else he or she would not have clicked on it. The positive utility to the user is even more straightforward if the user follows through with the purchase. The value of the advertising service is then split between consumers, for whom targeted advertising reduces search costs, and sellers.[10]

2.3.2 Not all advertisement is good, but ad-funded model creates incentive for the platform to preserve the value creation

For ads that do not involve a user's active engagement (a click), it is reasonable to assume that they may bring no value or even a certain disutility (take away webpage space, distraction) to the user. When many and not relevant, ads may also present a significant nuisance to private platform users and potentially decrease the utility that consumer obtains from the platform experience.

In fact, the common wisdom is that users get negative externality from an increase in the number of advertisers on the platform, due to the associated increase in nuisance from advertising. Although this needs to be qualified, as with targeting more advertisers may mean covering more potential needs rather than more nuisance, it is apparent that externalities between users and advertisers are largely asymmetric: advertisers likely value additional users more than users value additional advertisers.[11]

Because of this asymmetry, it is often optimal for the platform to charge the advertisers and let users enjoy the service for free. That is how incentives of a profit-maximizing platform get aligned with maximization of consumer welfare. By charging zero usage fees, the platform can compensate users for any negative externalities generated by advertisers. In turn, zero usage fees, and the positive value they get from participation, will compensate users (to the extent

[10] See Ackerberg (2003); de Corniere (2016); Burguet and Petrikaite (2017); Lau (2020).

[11] Schelanski et al. (2017) classify internet platforms into service-based and subsidy-based (where latter category includes ad-funded). While these authors do not study the question why a platform chooses to be of a certain type, they clearly note a different structure of cross-side externalities that the two types feature. For service-based platforms (Airbnb, Uber, Amazon, etc.), externalities are symmetric; for subsidy-based platforms (Twitter, Youtube, Snapchat, etc.), they are asymmetric.

possible) for the positive externalities that they generate on each other.

Ad-funded business models therefore appear to be a solution to preserving the value creation on platforms that facilitate the types of interactions discussed. In this way, the welfare-superior outcomes can be achieved without costly government intervention.[12]

2.4 Platforms rely on data to provide core functionality; more data facilitates relying on advertisement

2.4.1 User data is indispensable for providing discovery and matching functions

Besides providing the infrastructure that reduces the costs of interactions (or transactions) between users, platforms provide other functionalities, depending on the nature of interactions between the participants. These functionalities, as already mentioned in the previous subsection, include content or information discovery, user-to-user matching, content curation, customer reviews, identity verification, quality verification, payment, and dispute settlement mechanisms. Especially the discovery and matching functionalities heavily rely on personal data.

Discovery of information or content is clearly the primary function of internet search. It also plays a central role for social networks which feature a vast variety and quantity of content and many opportunities to engage in interactions with other users.

While the internet has drastically decreased the transaction cost of obtaining information and accessing content, the human ability to process this information has not increased much if at all.[13] When facing an overwhelming amount of information or options for interactions with other users or content, people need tools to help them make better decisions. Information on users' location, interests, and past behaviour, for example search history, fed into recommendations systems can help a platform improve the relevance of the content that the user is presented with. Indeed, as the CMA 2020 market study noted "[s]uccessful social media platforms feature a vast quantity of content that may be shown to consumers. To prevent congestion and maintain consumer attention, platforms must determine the most relevant content for a given consumer, and the order in which to present it. They do this using an algorithm, which makes decisions based on a range of data about the consumer and the content."[14]

The matching of users on social networks is distinct from the discovery function in the sense that both parties to an interaction need to be interested in it for the match to work. In this function, social networks, like Facebook, Snapchat or Twitter, use personal information to suggest new connections between users and may suggest that a user engages in certain ongoing interactions among other users.

In the discovery and matching functions, the platforms are reducing the users' cost of the negative network effects that may arise due to congestion from overload of information or from the many

[12]Indeed, without ad-funding option available, the platforms of interest would either have to engage in inefficient pricing or mandated to charge zero price and be subsidized by the government.

[13]See Diehl (2003).

[14]CMA, 2020, "Online platforms and digital advertising" Market study final report.

options for connection and interactions among users of social networks.[15]

User data seem less important for many platforms that mediate commercial transactions, like online marketplaces. This is because of the narrower scope of the interactions that they aim at enabling. While both ad-funded and some fee-funded platforms rely on personal data for discovery and matching, ad-funded platforms also heavily rely on such data for targeting ads.

2.4.2 Ad targeting and learning heavily rely on user data

Advertisers get value when consumers make a purchase because they saw an ad. This value depends on two factors: how many potential consumers see the ad, and how many of those will have followed through with the purchase. While it may be possible to generate increased sales by increasing advertising intensity, i.e., by showing more ads to a larger number of platform users, the potential to grow sales in this way is likely very limited. Ads can quickly overwhelm the webpage and become a nuisance to users who are on the platforms primarily for reasons other than to engage in ads.

The second factor—embodied in conversion rates—is likely to have stronger potential for enhancing advertisement value. With better targeting, increasing the effectiveness of advertising is possible for the same nuisance level.[16] This requires the use of various types of personal data. When social media platforms use targeted advertising, advertisers typically pay per click. And their willingness to pay increases with the increase in conversion rate.[17]

Finally, platforms also use data to learn how users utilize the functionalities of the platform in order to improve them.[18] The learning takes place across and within individual user data. Both aspects of this 'data-enabled learning' are pertinent to the efficient functioning of ad-funded platforms, in particular in relation to ad-targeting.[19]

[15] On the role of personal data in discovery and matching algorithms see, e.g., 2015 report by the Monopolies Commission: *Competition Policy: The challenge of digital markets*, Special Report No 68, page 72

[16] or vice versa, reducing the nuissance for the same effectiveness of advertising.

[17] As mentioned earlier, the conversion rate stands for the share of ad-clicking users that end up purchasing the advertised product. CMA study: Understanding preferences, purchasing intent and behaviour: understanding the wants and needs of specific consumers at any point in time is valuable to advertisers as they can target their adverts towards those individuals that they suspect are most likely to make a purchase. This targeting – whether it is based on contextual information such as the subject of a web page, or on personal data such as the individual's age or recent purchases – can result in a higher return on investment for advertisers, and a willingness to pay more. Similarly, advertisers are more likely to be willing to pay more in the future if they are given evidence that consumers exposed to adverts on a platform went on to make a purchase. Platforms are therefore rewarded by advertisers for having extensive and up-to-date knowledge of their consumers' characteristics, preferences and behaviour. The key input to this knowledge is data.

[18] For a discussion on how firms can improve their products through learning from customer data, see, e.g., Hagiu and Wright (2021).

[19] see, e.g., Esteban, Gil and Hernandez; Iyer, Soberman and Villas-Boas (2005).

3 Review of economic literature shows that the role of user transaction costs in the endogenous choice of monetisation has not been modelled so far

Our formal analysis of the following sections contributes to the economic literature on targeted advertising in the presence of network effects. The early literature (Rochet and Rochet 2003, Caillaud and Jullien 2003, Armstrong 2006) on two-sided platforms explained how externalities from participation within and across various groups of platform participants make the platform's pricing to the different groups interdependent. Although this literature does not consider the role of data explicitly, it provides a straightforward mechanism through which users might benefit from data aggregation: when data-enabled ad-targeting makes users more valuable to advertisers, the platform may have an incentive to expend more (implicit) subsidies to attract them.

Ample theoretical literature analyses the potential impact of targeted advertising on consumers. One strand finds that targeting can soften competition and lead to higher prices by allowing sellers to segment the market. Examples include Roy (2000); Iyer, Soberman, and Villas-Boas (2005); and Gaelotti and Moraga-Gonzalez (2008). In contrast, de Cornière (2016) shows, in the context of search-advertising, that when consumers actively search for products, targeting may lead to more intense competition. Consumers know that, when targeting is in place, they draw firms from a pool that better matches their preferences and the costs of additional search is therefore reduced. This means that searchers will sample a greater pool of sellers, so that each seller will face a larger number of competitors, which may intensify competition.

In a different setting, De Cornière and Nijs (2016) study an online platform auctioning one or more advertising slots. When only one advertising slot is available, product prices tend to be higher. However, when the platform makes two slots available for advertising—such that sellers compete—better targeting can lead to higher or lower prices, depending on demand conditions; and when the prices are higher with better targeting, consumers can be partly or fully compensated by better matches.

De Cornière and Taylor (2020) consider how data collection affects competition between platforms. They show how additional data may induce a platform to compete for users more aggressively ('unilaterally pro-competitive' data) or less aggressively ('unilaterally anti-competitive' data), depending on the modelling setup. If data is unilaterally anti-competitive, a restriction on data collection may improve consumer welfare. The authors show that data may be pro- or anti-competitive under plausible sets of assumptions, leading to contrasting conclusions about the desirability of restrictions on data gathering and about that of mandatory data-sharing. Our paper is related in that we consider how restrictions on data aggregation may affect the welfare of private users, including in a competitive setting. In particular, in our setting with data affecting the advertisement fee, the data is pro-competitive.

Hagiu and Wright (2021) consider how mandatory data sharing may affect platforms' entry, pricing decisions, and user welfare. They find that if platforms experience 'data-enabled learning', entrants with a superior technology can subsidize consumers initially to overcome any shortage of data vis-à-vis the incumbent.[20] This result holds regardless of whether consumers are forward-looking or myopic. In Hagiu and Wright's baseline model, data sharing does not change the identity of the winning firm. It does, however, eliminate competition for data that induces firms to subsidize users. Data sharing thus raises prices and harms user welfare. Our paper is related in that we consider the effect of data aggregation on prices and user welfare. We do not consider entry decisions or mandatory data sharing.[21]

Etro (2021) sets up a model of competition between device-funded and ad-funded platforms with differentiated ecosystems supporting apps that are provided under monopolistic competition. In Etro's model, the incentives of a device-funded platform in investing in app curation, introducing its own apps, and setting commissions on in-app purchases of external aps, are largely aligned with those of consumers. This is less likely to be the case for ad-funded platforms. In Etro's model, the consumers thus gain from a positive commission set by the device-funded platform because it implies a lower price of the device; consumers on ad-funded platforms do not receive similar benefits from a higher commission because they already enjoy zero price. Unlike Etro, we consider the business model choice (ad-funded vs fee-funded) to be endogenous. This is key to a proper assessment of welfare implications of different policies. Our focus on data aggregation restrictions and our modelling choices are also different from Etro's.

Prüfer and Schottmüller (2021) argue that data collection can lead to market tipping. They set up a simple dynamic model of competition between two firms that invest in product quality. Data on users serves to reduce the cost of product improvement and exhibits 'data-driven indirect network effects': the amount of user information is a function of a firm's past sales. The firm that initially has a superior product, and thus higher sales, acquires more user information in the first period. In subsequent periods, this firm continues to extend its quality advantage until it has obtained a monopoly position, where incentives for further quality improvement are muted. Consumers may therefore ultimately be harmed. This mechanism is very similar to tipping in the presence of standard network effects. A serious limitation of this model is that it does not allow firms to adopt dynamic pricing strategies, unlike in, e.g., Hagiu and Wright (2021), who reach a very different conclusion.[22] Aside from the fact that we also study the effect of data collection on consumer welfare, our paper is not closely related to Prüfer and Schottmüller (2021).

Leonard (2019) discusses the ad-funded platforms explicitly. He argues that data aggregation allows ads to be better targeted to users, benefiting users directly through more relevant ads and indirectly through the enhanced 'incentives for the service to innovate with respect to user

[20] See also insulated pricing discussed by White and Weyl (2016)

[21] Condorelli and Padilla (2020) present a model where a firm dominant in a "primary" market enters a data-rich "secondary" market. The dominant firm engages in predatory pricing in the secondary market and privacy-policy tying. This confers the dominant firm an advantage in the primary market and, under some conditions, entry is deterred and consumers are harmed. We consider the limitations of Condorelli and Padilla's analysis in Langus and Lipatov (2020). This paper is distantly related to Condorelli and Padilla (2020) in that we also study effects of aggregation of data across adjacent markets on welfare. Our model is, however, different and we focus on the choice of business model instead of choice of entry.

[22] Moreover, in the real world, platforms do not exclusively rely on user data for product improvement, and collecting data from users is costly, so that the stark conclusions of Prüfer and Schottmüller (2021) that mandatory data sharing would improve consumer welfare should be discounted

experience'. Similarly, we focus on ad-funded platforms and the effect of potential restrictions to data aggregation. However, unlike Leonard (2019), we focus on pricing and study the platforms' incentives and effect on users formally.

Calvano and Polo (2020) analyse endogenous differentiation in business models in the context of competing broadcasters. They show that when revenue potential of the two sides is sufficiently balanced, it may be optimal for each broadcasting station to focus monetisation on the different group of agents (viewers vs advertisers) in order to relax competition on each side of the market. Our results complement the findings of Calvano and Polo (2020) as we focus on the heterogeneity of platform users and the positive externality their participation has on the same-side platform users while not modelling advertisers explicitly.

We are not aware of a prior paper that would formally study the role that transaction costs play in the choice of an ad-funded business model; and, specifically, how the presence of these costs impacts the welfare effects of a potential restriction on data collection.

4 Stronger network effects and better data may lead to creation of ad-funded platforms; data restrictions may force a platform to charge its users a fee

4.1 A formal setting

In the following, we provide a formal description of certain mechanisms that play a role in the choice of an ad-funded business model. We abstract from the fact that data aggregation likely allows the platform to directly improve the users' experience, independently of the effect of data on the level of advertisement targeting.[23]

We take a reduced-form approach to modelling advertising process.[24] In particular, we assume that merchants are characterized by a demand-for-clicks function, which is decreasing in price-per-click e and increasing in the conversion ratio ρ (probability of purchase given a click). The platform faces inverse demand $e(a, \rho)$ where a denotes the number of clicks. We first consider perfectly elastic demand such that the number of clicks is determined by the supply; the price per click can then be written as a function of conversion probability only, $e(\rho)$.

Consumers are independently distributed on a unit interval along two attributes: θ describes how valuable platform services are to the user; η denotes the user's cost of a click on the link associated with an ad. This cost may be associated with the expected time waste and effort from clicking, waiting for the advertiser's webpage to load and making sense of the content of that webpage—which may or may not end up being relevant.

The utility derived from platform services is $u(\theta, .)$, where the second argument stands for the measure of other users present on the platform. This is social interaction effect or, in

[23]At the same time, we also abstract from privacy issues and any effect that better data could have on the competition between merchants.

[24]See, e.g., Gomes (2014), for a different approach—a model of optimal auction design in advertising context

economic terms, a within-group network effect; a positive (under our assumptions) externality that participation of users has on the users of the same type. Under a simplifying assumption that users only take into account the utility from platform services in their participation decision, the measure can be conveniently recast as mass of consumers opting out of the platform θ_0, with $u(\theta, \theta_0)$ increasing in its first argument and decreasing in its second argument. θ_0 is then defined by $u(\theta_0, \theta_0) = p$, where $p \geq 0$ is the subscription fee and the demand for platform services can be described by $1 - F(\theta_0)$, where F is the cdf of θ.

We express the cost of a click η in money metric so that the net utility of the user of type (θ, η) who joins the platform and clicks on the ad is $u(\theta, \theta_0) - p + \rho v - \eta$. Here, v is the net gain of the user in case of purchase of advertised product that was triggered by clicking the ad. Users with $\eta < \rho v$ who are on the platform will click the ad link, so the total measure of clicks will be $(1 - F(\theta_0)) G(\rho v)$, where G is the cdf of η.

We model data aggregation at a very high level by assuming that data aggregation increases the conversion rate ρ. This is a direct implication of the assumption that better data (e.g. associated with aggregation across various sources) results in a higher relevance of the ad. In other words, users will be shown the most relevant ads such that they are more likely to click on the ad and follow-through with the product purchase. We assume that advertisement does not result in any nuisance costs for users that do not click on it.

4.2 A platform is more likely to choose ad-funded business when externality is strong and data is ample

To understand the factors driving the choice of user fee in this setting, consider platform profits $p(1 - F(\theta_0)) + e(\rho) a$ that consist of subscription revenue $p(1 - F(\theta_0))$ and advertising revenue $e(\rho) a$. The platform is maximizing these profits by choosing an optimal subscription fee.

To illustrate the general result that we present below, we turn to a simple example.

Consider $u = \theta - \alpha\theta_0$ with θ distributed uniformly. $\alpha < 1$ is a parameter that characterises the strength of network externality. The indifferent between joining and not joining user can be characterised as $\theta_0 = \frac{p}{1-\alpha}$ and the optimal subscription fee is

$$p = \max\left\{\frac{1}{2}(1 - \alpha - eG(\rho v)), 0\right\}$$

which turns to zero when $\alpha + eG(\rho v) \geq 1$, i.e. exactly when the value of other users participating is high and/or when advertisers are willing to pay a relatively high fee. Note that this also implies that the more important the network effect is for user value, the less the necessary contribution of the advertising revenue to the choice of ad-funded model—and vice versa.

At the same time, ad profitability builds on both the net value of ads to users and the value of ads to merchants. The former is necessary for the ad-funded business model (where the platform charges advertisers per click) to work because the users would simply never click on an ad if this did not create any net value to them.

The following assumptions are helpful in deriving more general results.

Assumption 1: $u_1 > -u_2$.

This is a regularity condition which has a straightforward interpretation: the 'direct' effect of own propensity to enjoy the platform services is stronger than 'indirect' effect of deriving utility form marginally higher participation of others.

Assumption 2: $1 - F(\theta_0(p)) - (eG(\rho v) + p) f(\theta_0(p)) \frac{d\theta_0}{dp}$ is decreasing in p.

This is a technical condition that implies some regularity assumptions on F and u that we specify in Appendix A.

We can now formulate the following proposition:

Proposition 1: Under assumptions 1 and 2, the platform is 'more likely' to choose the ad-funded business model when e is high, i.e., it can charge higher fee per click; and when u_2 is high by absolute value, i.e., the network externality is high.

The proof is presented in Appendix A.

Recall that e—the fee per click—increases in the conversion probability ρ, which is increasing in the quality of data that platform has at its disposal. This allows us to formulate the following corollary to Proposition 1.

Corollary: Availability of better data increases attractiveness of the ad-funded business model for the platform.

In other words, the better the quality of the data (e.g. due to availability of data aggregation), the larger the set of parameter values for which the platform chooses to charge zero subscription fee from its users.

We now turn to welfare assessment, focusing on consumer surplus (CS). In our setting, CS can be written as a sum of two components: the net aggregate utility from platform services $\int_{\theta_0}^{1} (u(x, \theta_0) - p) \, dF(x)$ and the aggregate utility from purchases triggered by advertising net of cost of clicking $(1 - F(\theta_0)) \int_0^{\rho v} (\rho v - y) \, dG(y)$. [25]

The direct effect of changes in conversion probability is on the latter component as it increases expected net utility from clicks for users. The marginal utility from increasing conversion probability is maximized at $p = 0$. This is intuitive, as with ad-funded business the number of consumers who join the platform is maximized, and so the positive effect of higher ad relevance is enjoyed by the largest possible group of users.

The indirect effect of conversion probability works through the price of service. In Appendix B, we show that an increase in conversion probability decreases the price unless it is already zero. Intuitively, the higher revenue per click induces the platform to decrease subscription fee in order to attract more users—eventually hitting the zero price constraint.[26] The indirect welfare

[25] Recall that we are abstracting from privacy concerns; users do not experience discomfort in relation to the collection of data.

[26] In a more general formulation, where the ad-funded platform can invest in increase in quality, these investments would continue to increase with ρ

effect is therefore positive.

For ad-funded platforms, the indirect effect is zero as long as the change in the matching quality is marginal. For infra-marginal reduction of conversion probability, such as presumably would be triggered by inability to aggregate data or by limitations on data collection, an ad-funded platform may be pushed to charging positive subscription fee. We consider this situation in more detail in the next section.

4.3 With transaction costs, data degradation may force a platform to abandon ad-funded model with harm to consumers

We now introduce transaction costs by assuming that users joining a platform that charges positive subscription fees incur one time cost t (money metric). In reality, transaction costs likely vary across potential platform users, but we make a simplifying assumption that they are uniform. With transaction costs, the net utility of the user (from the platform services) has a discontinuity with $u(\theta, \theta_0(0))$ at $p = 0$ and $u(\theta, \theta_0(\varepsilon)) - \varepsilon - t$ at $p = \varepsilon > 0$.

There will therefore exist a user indifferent between joining and not joining the platform at a positive price located at inframarginal distance from $\theta_0(0)$, denoted by θ_0 in this subsection. Specifically, while the location of the "zero-price indifferent" user will be determined by $u(\theta_0, \theta_0) = 0$, the location of the indifferent user at strictly positive price, will be determined by $u(\theta_1, \theta_1) = p + t$. The measure of users who will only be served at zero price is then $F(\theta_1) - F(\theta_0)$.

In deciding on its business model, the platform will compare the profits from setting optimally a price $p > 0$ with the profit from setting $p = 0$. Straightforwardly, with positive transaction costs the set of parameters for which ad-funded business model is preferred over fee-funded one, expands relative to the situation with no transaction costs. This is because the platform profits with zero price are independent of transaction costs, while profits with optimal non-zero price strictly decrease in transaction costs.

To find the level of the restriction on data aggregation (directly translated into deterioration of the ad relevance and ultimately into conversion probability ρ) where the platform is indifferent between choosing the two business models, consider the following equality:

$$e(\rho) G(\rho v) (F(\theta_1(\rho)) - F(\theta_0(\rho))) = p(\rho) (1 - F(\theta_1(\rho))).$$

Here, the left-hand side is the cost of switching to subscription-based model in terms of advertising profits lost, whereas the right-hand side is the benefit of such switching in terms of subscription revenue. In other words, at this value of ρ, the advertising revenue from the demand lost due to transaction cost $F(\theta_1(\rho)) - F(\theta_0(\rho))$ is exactly equal to the subscription revenue from the "reduced" population (demand at optimal non-zero price $1 - F(\theta_1(\rho))$).

A slight decrease from this level of ρ will result in a decrease of consumer surplus that consists of two parts. The first part of the decrease includes utility from platform services and from purchasing advertised goods for the set of consumers who choose not to participate at positive fee. The second part of the decrease is the lost utility from (i) underuse of network effects, (ii) paying positive subscription fee, (iii) bearing transaction cost for all consumers who choose to stay on the platform even with the positive fee.

This allows us to formulate the following proposition.

> Proposition 2. A marginal degradation of data quality used by an ad-funded platform may result in a change of the business model by this platform and associated with it inframarginal loss of consumer surplus.

The formal proof of proposition 2 is left to Appendix C.

5 Better data may stimulate platform competition

Here we extend the model used to analyse the effects of data quality in Section 3 by introducing competition between two platforms that choose whether they will adopt an ad- or a fee-funded business model. In this model, the indirect value for platform users comes exclusively from same-group network effects (the more users, the higher the value of the service for each user).

The main results confirm the findings from our analysis of standalone platforms. In particular, we find that the restriction of data aggregation (quality of data available to platforms) may induce a change in the business models that are adopted in equilibrium: the higher the data quality, or the value of additional user in terms of advertising revenue more generally, the more likely it is that the equilibrium will feature both platforms charging zero price. Conversely, the lower the data quality, the more likely the market is not fully covered in equilibrium.[27] In this way, Proposition 2 and the corollary to Proposition 1are both corroborated in the setting with platform competition and single-homing.

5.1 Formal setting and pro-competitiveness of data

When we introduce competition to the baseline model considered in Section 3, it may generally affect both user engagement characterised by the set of indifferent users θ_0 and the inverse demand for ads $e(\rho)$. Regarding the former, we assume that θ is a two-dimensional attribute uniformly distributed in the unit square. Coordinates of a point then represent the value of each platform's service to the user situated in this point. With single-homing, each user chooses the platform that she expects to bring her higher net utility (in case it is non-negative). With multi-homing, the user splits her time between the two platforms proportionally to the expected net utility in case the net utilities from both platforms are positive. Regarding the latter, we adopt the simplest formulation by keeping the assumption that $e(\rho)$ is the exogenous price that the market (advertisers) is ready to pay per click. The conversion probability ρ_i may generally differ across platforms.

[27]The market is fully covered in our setting as long as at least one of the platforms is ad-funded. Only when both platforms are subscription-based may there be partial coverage in equilibrium

We consider the single-homing case.[28] Each user compares 3 options: go with the first platform; go with the second platform; or stay out of platform services. Our baseline setting is asymmetric: while users of one of the platforms enjoy a network externality of the type described in Section 3 with its strength reflected in parameter α, the other platform does not. In other words, users of one platform do not value the participation of other users, while the users of another platform do so.

The formal details of our setting are left to Appendix D. While the demand for a platform's service is determined by participation constraint of the users when this platform is standalone, the demand is also affected by the price and offer of the rival when there are competing platform services. Each platform, therefore, considers the effect of a (subscription) price decrease on its demand along two margins: participation (how many users will decide to join the platform afresh) and competition (how many users will decide to switch away from the competitor platform).

At low (close to zero) prices, the participation margin is likely not relevant, so the residual demand is less elastic than at higher price.

Depending on the relative strength of network effects and the quality of data (via associated with it advertising fee), various types of equilibria are possible. These include:

- both platforms choose an ad-funded business model
- one platform chooses an ad-funded business model and another chooses a subscription-fee-based model
- both platforms choose to charge a subscription fee and the whole market is covered
- both platforms choose to charge a subscription fee and some users decide not to purchase either of the services.

Our results in this competition setting corroborate our findings in the setting with a standalone platform. In particular, the following Proposition 1C is an analogue of Proposition 1 in the setting with platform competition.

> Proposition 1C. In the competitive setting formulated above, the equilibrium with both platforms choosing ad-funded business only exists when advertising fee e is sufficiently high and when network externality α is sufficiently high (but not too high).

The proof is presented in Appendix D2 (Subsection Both platforms are ad-funded). The first part of the proposition (regarding advertising fee e) also holds in the multi-homing setting (proof to the first-oder effect is in Appendix E2).

Furthermore, the restrictions on the data quality may be detrimental to consumer welfare as they change the type of equilibrium that characterizes the market interaction. As long as there is full coverage, consumers will only suffer from higher price and transaction cost. Once the data quality is so low that there is partial coverage in equilibriums, all three channels that may harm

[28]The multihoming case is more cumbersome and is left to Appendix E. In the following, whenever we are able to confirm a single-homing result in multi-homing setting, we explicitly mention that

consumers (lost participation value and lost network effects in addition to the transaction cost) come into action and Proposition 2 becomes fully relevant also for platform competition.

Finally, we obtain two new propositions in the setting with platform competition, both indicating that data is 'pro-competitive' in the sense that users enjoy higher consumer surplus in equilibrium if the quality of data increases. This is reminiscent of the notion of 'unilaterally pro-competitive' introduced by de Cornière and Taylor (2020) and obtained by them in a variety of applications. In targeted advertising context, these authors analyse a model that is different from ours as it features an auction among advertisers, a single source of heterogeneity among users, and no direct same-side network effects for users. In their setting, data is unilaterally pro-competitive if and only if the demand for ads becomes less elastic with better targeting.

> Proposition 3. In the competitive setting formulated above, the data is 'pro-competitive' in the equilibrium when one of the platforms chooses to be ad-funded and another does not. In particular, the higher the data quality is, the lower the price that not-ad-funded platform chooses to charge its users.

The proof is presented in Appendix D2 (Subsection Only one platform is ad-funded). This proposition also holds in the multi-homing setting as shown in Appendix E3.

> Proposition 3A. In the competitive setting formulated above, the data is 'pro-competitive' in the equilibrium with full coverage and none of the platforms choosing to be ad-funded. In particular, the higher the data quality is, the lower the sum of prices that the two platforms charge to their users.

The proof is presented in Appendix D2 (Subsection No platform is ad-funded). A similar result can be obtained in the multi-homing setting if only one of the platforms changes the data quality (see Appendix E4).

5.2 The role of transaction costs

Transaction costs discussed at length in the previous section play a distinct role in the competitive environment considered above. In particular, they make the equilibrium with both firms choosing to be ad-funded more likely. Intuitively, with transaction cost for users associated with positive subscription fee, any platform that wants to deviate from ad-funded equilibrium has to compensate its users for such transaction costs. This makes such deviation less profitable for the platform.

Formally, we obtain this result in Proposition 4. The following lemma, valid only when profitable deviation exists at least without transaction cost, is useful for the proof of this proposition.

> Lemma 1. For the same parameter values, the optimal deviation profit is larger when transaction cost is zero, $\max_{p_1}(p_1 + e(\rho))D_1(p_1, t = 0) > \max_{p_1}(p_1 + e(\rho))D_1(p_1, t > 0)$.

> Proof. Suppose $\exists p_1^* : (p_1^* + e(\rho))D_1(p_1^*, t > 0) \geq \max_{p_1}(p_1 + e(\rho))D_1(p_1, t = 0)$. Because $D_1(p_1, t = 0) > D_1(p_1, t > 0)$ for any p_1 in the interior, it must be that

$(p_1^* + e(\rho))D_1(p_1^*, t > 0) < (p_1^* + e(\rho))D_1(p_1^*, t = 0)$. But by definition of maximum, $\max_{p_1}(p_1 + e(\rho))D_1(p_1, t = 0) \geq (p_1^* + e(\rho))D_1(p_1^*, t = 0)$. We must therefore have that $\max_{p_1}(p_1 + e(\rho))D_1(p_1, t = 0) > (p_1^* + e(\rho))D_1(p_1^*, t > 0)$, a contradiction with our supposition, Q.E.D.

Proposition 4. In the competitive setting formulated in Appendix D3, the presence of transaction cost ($t > 0$) extends the set of parameter values for which the all ad-funded equilibrium exists, as compared to the situation with no transaction cost ($t = 0$).

Proof. Suppose the 'all ad-funded' equilibrium is less stable when transaction costs are present, i.e, there is a range of parameter values for which it is an equilibrium without transaction cost, but a deviation is profitable with transaction cost. Consider, under these parameter values, a platform deviating, in the situation without transaction cost, to a positive price, say p_1^d, that is optimal in the situation with transaction cost. By our supposition, such deviation is profitable in the situation with transaction cost. But $\pi(p_1^d, t = 0) = (p_1^d + e(\rho))D_1(p_1^d, t = 0) > (p_1^d + e(\rho))D_1(p_1^d, t > 0) = \pi(p_1^d, t = 0)$, i.e., the deviation profit is higher in the situation with no transaction cost. The equilibrium profit is however the same in the two situations, $\pi(0, t = 0) = e(\rho)D_1(0, t = 0) = e(\rho)D_1(0, t > 0) = \pi(0, t > 0)$, because, in 'all ad-funded' equilibrium, $D_1 = \frac{1}{\alpha^2}(\alpha + \sqrt{1 - 2\alpha + 2\alpha^2} - 1)$ is independent of t. We have therefore reached a contradiction and it is not possible that all ad-funded equilibrium is less stable when transaction costs are present. To complete the proof, we need to show that the set of parameter values for which 'all ad-funded' equilibrium obtains is not the same regardless of whether transaction cost is present or not. Consider the set of parameter values at the border of the set for which 'all ad-funded' equilibrium obtains with transaction cost. The optimal deviation, say p_1^0, is characterised by equality of equilibrium and deviation profits at such a border, $(p_1 + e(\rho))D_1(p_1^0, t > 0) = e(\rho)D_1(0, t > 0)$. By lemma 1, we can always find a profitable deviation at these parameter values when transaction cost is absent. Indeed, $\max_{p_1}(p_1 + e(\rho))D_1(p_1, t = 0) > (p_1 + e(\rho))D_1(p_1^0, t > 0) = e(\rho)D_1(0, t > 0) = e(\rho)D_1(0, t = 0)$. Therefore, the border of the set of parameter values for which 'all ad-funded' equilibrium obtains with transaction cost lies outside of the set of parameter values for which 'all ad-funded' equilibrium obtains without transaction cost, Q.E.D.

6 Better data may increase competition among merchants

In a straightforward way, more data aggregation and better targeted advertising can enhance competition by enabling niche firms to efficiently reach consumers and compete for sales with generalist firms. A niche firm is characterized by the ability to serve a relatively small group of consumers, whereas a generalist firm can sell across various groups of consumers.[29]

[29]The CMA's report states: "Platforms such as Google and Facebook have made it substantially easier for businesses to reach and serve adverts to consumers all around the world, in a way that was only previously possible for large

Such situations are common and may arise for a number of reasons. One reason is that a firm may have a narrow geographic coverage. An example is a local burger restaurant competing in its home town with Big Mac. Another reason may be that the firm offers a narrow range of products, while competing with another firm that offers a broad range of products. One can think of a start-up offering specialized electric bikes competing with an established producer that offers all kinds of bikes and maybe also other sports gear.

When advertising is not targeted, a niche firm will offer lower bids for advertisement compared to a generalist firm. This is because it will expect lower click and conversion rates.[30] Accordingly, the platform will show the generalist's ads to users that would consider—even prefer—the niche firm, such that they may end up addressing their demand to the generalist.

To see how this may occur in a formal setting, consider a consumer characterized by two attributes: θ is relevance of a product to consumer (horizontal differentiation, with cdf F over $[0, 1]$; and v_i is the value of product i with cdf G_i. Both θ and v are unobserved without targeted advertising.

In a simplest version, we consider two firms offering one product each. For convenience, we can assume that F is uniform (almost without loss of generality). Furthermore, for extreme simplification, consider G_1 being uniform on $[0, 1]$ for any θ, and G_2 being uniform on $[0, k]$ for $\theta \in \left[\theta_0 - \frac{\varepsilon}{2}, \theta_0 + \frac{\varepsilon}{2}\right]$ and zero otherwise, with $k > 1$.

Abstracting from cost of production, without targeting, firm 1 is willing to bid $\frac{1}{4}$ for the ad, whereas firm 2 is willing to bid $\frac{k}{4}\varepsilon$ (we assume that $k\varepsilon < 1$ so that the value generated by firm 2 to a small fraction of consumers cannot compensate the value generated by firm 1 to all consumers). Firm 1 then wins the auction and pays $\frac{k}{4}\varepsilon$ in a second-prize auction.

With targeting, firm 1 wins any slot that is offered to consumers other than on $\left[\theta_0 - \frac{\varepsilon}{2}, \theta_0 + \frac{\varepsilon}{2}\right]$ and pays zero to the platform (firm 2 has no value from posting advertisement to such consumers). For $\left[\theta_0 - \frac{\varepsilon}{2}, \theta_0 + \frac{\varepsilon}{2}\right]$, firm 2 wins and pays $\frac{\varepsilon}{4}$ to the platform (this is what firm 1 would get from advertising to these consumers).

Consumer surplus is $\frac{1}{8}$ without targeting. It is $\frac{k^2}{8}\varepsilon + \frac{1-\varepsilon}{8}$ with targeting, which is higher. Note that this is a pure efficiency gain as we abstract from direct competition between generalist and niche firms.

The price discrimination channel is shut down[31] in this formalisation by the assumption that targeting does not concern willingness to pay, but only preference about certain product relative to other products. In such a setting, we indeed lose consumer surplus if targeted advertising is prohibited. Total surplus is also harmed without targeted advertisement as the second firm is out of the market and does not provide its customers with the high value of its product.

Note, however, that the modelling approach is conservative to the potential of targeted advertising

companies. This has opened up greater advertising possibilities for a long tail of small businesses, and enabled large numbers of predominantly online businesses to thrive that may otherwise not have been viable."

[30]This needs a further qualification.

[31]For a more general setting where targeting can be used to reallocate more of the consumer surplus, see, e.g., Padilla et al (2021)

to enhance consumer welfare in the sense of assuming that the goods offered by a large generalist firm and a small niche producer are independent. As a result, winning an advertisement auction essentially grants a monopoly on the segment of users that are targeted. When goods are imperfect substitutes, targeted advertising also contributes to consumer welfare by increasing competitive pressure on the large generalist firm. Additionally, it opens possibilities for entry as startups may find it feasible to reach their customers unlike in a situation without targeted ads.

7 Concluding remarks

In current policy debates, there have been voices raised implying that ad-funded platforms need to be restricted in their ability to collect and aggregate user data. We show that, while not without merit, such proposals have a significant potential to harm consumers. We explain how ad-funded business models allow platforms to facilitate value generation that would otherwise be lost to society. User data is an integral part of this function, and restricting collection or aggregation of this data could impair the viability of ad-funded platforms, with sad consequences for welfare.

With help of formal modelling, we identify network effects, ability of platforms to charge a higher advertisement fee for better targeting, and transaction costs involved in charging users a positive fee for platform services as crucial mediating factors in the mechanism by which data aggregation restrictions may harm consumers.[32]

Our model can be extended along several dimensions. A few examples are as follows. First, the impact of better data on the quality of the core service can be added to that on conversion rate. Second, the structure of industry in which advertisers compete may be modelled to study how it could affect the mechanism described in our paper. Finally, alternative preference structures may be modelled to better reflect the environment in which conglomerate platforms compete in real life.

8 Acknowledgements

We are grateful to Yassine Lefouili for helpful comments and suggestions. We also thank Jad Benmoussa, Raphael Poncet, and Taco Prins for excellent research assistance. In a significant part, the authors have carried out the research for this paper while with E.CA Economics. Financial support from Meta inc. is gratefully acknowledged. The opinions and judgements expressed in the paper are the authors' views and do not necessarily represent the views of Meta. All errors are the authors'.

[32]The abandonment of ad-based monetisation strategy is often the conductor through which data aggregation restrictions harms consumers.

9 References

Ackerberg, D. (2003): Advertising, learning, and consumer choice in experience good markets: an empirical examination, International Economic Review 44, 1007-1040.

Armstrong, M. (2006): Competition in two-sided markets, RAND Journal of Economics 37, 668–691.

Burguet, R. and V. Petrikaite (2017): Targeted Advertising and Costly Consumer Search, Working Papers 971, Barcelona Graduate School of Economics.

Caffarra, C., F. Etro, O. Latham, and F. Scott Morton (2020): Designing regulation for digital platforms: Why economists need to work on business models, VoxEU 04 June 2020, available at https://voxeu.org/article/designing-regulation-digital-platforms , last accessed on 29 October 2021.

Caillaud and Jullien (2003): Chicken & Egg: competition among intermediation service providers, RAND Journal of Economics 34, 309-328.

Condorelli, D. and J. Padilla (2020): Data-driven Envelopment with Privacy-Policy Tying, available at SSRN: https://ssrn.com/abstract=3600725 or http://dx.doi.org/10.2139/ssrn.3600725

Calvano, E. and M. Polo (2020): Strategic Differentiation by Business Models: Free-to-air and Pay-TV, The Economic Journal 130, 50-64.

de Corniere, A. (2016): Search Advertising, American Economic Journal: Microeconomics 8, 156–188.

de Cornière, A., and R. Nijs (2016): Online advertising and privacy, RAND Journal of Economics 47, 48–72.

de Cornière, A., and G. Taylor (2020): Data and Competition: a General Framework with Applications to Mergers, Market Structure, and Privacy Policy, TSE Working Paper 20-1076.

Diehl, K. (2003): Personalization and Decision Support Tools: Effects on Search and Consumer Decision Making, Advances in Consumer Research 30, 166-169.

Economides, N. and I. Lianos (2021): Privacy and Antitrust in Digital Platforms, Journal of Competition Law and Economics forthcoming, https://doi.org/10.1093/joclec/nhab007

Engels, B. (2016): Data portability among online platforms, Internet Policy Review 5, https://doi.org/10.14763/2016.2.408.

Esteban, Gil and Hernandez (2003): Informative Advertising and Optimal Targeting in a Monopoly, Journal of Industrial Economics 49, 161-180.

Etro, F. (2021): Device-funded vs ad-funded platforms, International Journal of Industrial Organization 75, 102711.

Galeotti A. and J. Moraga-Gonzalez (2008): Segmentation, advertising and prices, Internatioal Journal od Industrial Organization 26, 1106-1119.

Gomes, R. (2014): Optimal Auction Design in Two-Sided Markets, RAND Journal of Economics 45, 248-272.

Hagiu, A. and J. Wright (2021): Data-enabled learning, network effects and competitive advantage, Working Paper.

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

Langus, G., and V. Lipatov (2021): Does Envelopment through Data Advantage Call for New Regulation? CESifo Working Paper 8932, March 2021.

Lau, Y. (2020): A Brief Primer on the Economics of Targeted Advertising, FTC Economic Issues Paper.

Leonard, G. (2019): The Economics of Online Ad-Supported Services. Available at SSRN: https://ssrn.com/abstract=3663317 or http://dx.doi.org/10.2139/ssrn.3663317

Mazzucato, M., R. Kattel, T. O'Reilly, and J. Entsminger (2021): Reimagining the Platform Economy, Project Syndicate 5February 2021, available at https://www.project-syndicate.org/onpoint/platform-economy-data-generation-and-value-extraction-by-mariana-mazzucato-et-al-2021-02?referral=3db0f9 , last accessed on 29 October 2021.

Padilla, J., S. Piccolo, and H. Vasconcelos (2021): Should Vertically Integrated Platforms be Mandated to Share Information with their Rivals? Economics Letters 203, 109849.

Prüfer, J. and C. Schottmüller (2021): Competing with Big Data, Journal of Industrial Economics, forthcoming.

Rochet, J. and J. Tirole (2003): Platform Competition in Two-Sided Markets, Journal of the European Economic Association 1, 990-1029.

Roy, S. (2000): Strategic segmentation of a market, International Journal of Industrial Organization 18, 1279-1290.

Shelanski, K., S. Knox and A. Dhilla (2017): Network Effects and Efficiencies in Multisided Markets, available at https://one.oecd.org/document/DAF/COMP/WD(2017)40/FINAL/en/pdf, last accessed on 27 October 2021.

White, A. and E. G. Weyl (2016): Insulated Platform Competition. Available at SSRN: https://ssrn.com/abstract=1694317

Appendix A - proof of Proposition 1

The platform is maximizing its profits by choosing an optimal subscription fee:

$$\max_p \left(p + e\left(\rho\right) G\left(\rho v\right)\right)\left(1 - F\left(\theta_0\left(p\right)\right)\right). \tag{1}$$

The optimal fee will solve

$$1 - F\left(\theta_0\right) - \left(p + e\left(\rho\right) G\left(\rho v\right)\right) f\left(\theta_0\right) \frac{d\theta_0}{dp} \leq 0, \tag{2}$$

where the effect of price on the indifferent user can be found from totally differentiating $u\left(\theta_0, \theta_0\right) = p$:

$$\frac{d\theta_0}{dp} = \frac{1}{u_1\left(\theta_0, \theta_0\right) + u_2\left(\theta_0, \theta_0\right)}. \tag{3}$$

Since a platform can be defined as ad-funded if it monetizes its services on advertisement revenue and charges consumers zero fees, the condition for the platform to choose ad-funded business model is the condition for problem (1) to have a corner solution:

$$1 - F\left(\theta_0\left(0\right)\right) \leq eG\left(\rho v\right) f\left(\theta_0\left(0\right)\right) \frac{d\theta_0}{dp}|_{p=0}.$$

The platform is 'more likely' to choose ad-funded business when the lhs is small and the rhs is large. Because, under assumption 1, $\frac{d\theta_0}{dp}|_{p=0} > 0$ from (3), the rhs will be large if e is high, i.e., when the platform can charge higher fee per click; and when u_2 is high by absolute value, i.e., the network externality is high, so that $u_1 + u_2$ is small, Q.E.D.

Assumption 2 guaranties that the condition (2) is also sufficient for the corresponding fee to be optimal (profit is a concave function of subscription fee). It can be written as

$$-f\left(\theta_0\right) \frac{d\theta_0}{dp} - f\left(\theta_0\right) \frac{d\theta_0}{dp} - \left(p + e\left(\rho\right) G\left(\rho v\right)\right) f'\left(\theta_0\right) \left(\frac{d\theta_0}{dp}\right)^2 - \left(p + e\left(\rho\right) G\left(\rho v\right)\right) f\left(\theta_0\right) \frac{d^2\theta_0}{dp^2} < 0,$$

and rewritten as

$$2f\left(\theta_0\right) + \frac{p + e\left(\rho\right) G\left(\rho v\right)}{u_1 + u_2} \left(f'\left(\theta_0\right) - f\left(\theta_0\right) \frac{u_{11} + u_{12} + u_{21} + u_{22}}{\left(u_1 + u_2\right)^3}\right) > 0,$$

and thus a (very strong) sufficient condition on functions F and u would be

$$\frac{f'\left(\theta_0\right)}{f\left(\theta_0\right)} > \frac{u_{11} + u_{12} + u_{21} + u_{22}}{\left(u_1 + u_2\right)^3}.$$

Appendix B - welfare assessment

In our setting, CS can be written as

$$\int_{\theta_0}^{1}\left(u\left(x,\theta_0\right)-p+\int_{0}^{\rho v}\left(\rho v-y\right)dG(y)\right)dF\left(x\right),$$

which has two components: the net aggregate utility from platform services $\int_{\theta_0}^{1}\left(u\left(x,\theta_0\right)-p\right)dF\left(x\right)$ and the aggregate utility from purchases triggered by advertising net of cost of clicking $\left(1-F\left(\theta_0\right)\right)\int_{0}^{\rho v}\left(\rho v-y\right)dG(y)$. The direct effect of changes in ρ is on the latter component as it increases expected net utility from clicks. The corresponding derivative with respect to ρ is

$$vG(\rho v)\left(1-F\left(\theta_0\right)\right),$$

which attains its maximum at $p = 0$. This is intuitive, as with ad-funded business the number of consumers who join the platform is maximized, and so the effect of a better match between advertiser and user is enjoyed by the largest possible group of users.

The indirect effect goes through price of service–the corresponding derivative is

$$\int_{\theta_0}^{1}\left(u_2\left(x,\theta_0\right)\frac{d\theta_0}{dp}-1\right)\frac{dp}{d\rho}dF\left(x\right)-\left(u\left(\theta_0,\theta_0\right)-p\right)f\left(\theta_0\right)\frac{d\theta_0}{dp}\frac{dp}{d\rho}$$

Below, we show that $\frac{dp}{d\rho}$ is negative at the interior. Intuitively, 'cheaper' to get advertising revenues induce the platform to decrease subscription fee in order to ensure higher coverage of potential users. The indirect effect is therefore positive.

The effect of conversion probability on price can be found from the condition for interior optimum (clearly, at the interior of the set $\{\rho : p\left(\rho\right)=0\}$, we have $\frac{dp}{d\rho}=0$)

$$\left(1-F\left(\theta_0\right)\right)\left(u_1\left(\theta_0,\theta_0\right)+u_2\left(\theta_0,\theta_0\right)\right)=\left(p+e\left(\rho\right)G\left(\rho v\right)\right)f\left(\theta_0\right)$$

Totally differentiating this condition and rewriting, we get

$$\begin{aligned}&\left(\left(\begin{array}{c}\left(1-F\left(\theta_0\right)\right)\left(u_{11}+u_{12}+u_{21}+u_{22}\right)-f\left(\theta_0\right)\left(u_1+u_2\right)\\-\left(p+e\left(\rho\right)G\left(\rho v\right)\right)f'\left(\theta_0\right)\end{array}\right)\frac{d\theta_0}{dp}-f\left(\theta_0\right)\right)dp\\=\;&\left(e'\left(\rho\right)G\left(\rho v\right)+ve\left(\rho\right)g\left(\rho v\right)\right)f\left(\theta_0\right)d\rho\end{aligned}$$

or

$$\begin{aligned}&\left(\left(\begin{array}{c}\left(1-F\left(\theta_0\right)\right)\left(u_{11}+u_{12}+u_{21}+u_{22}\right)\\-\left(p+e\left(\rho\right)G\left(\rho v\right)\right)f'\left(\theta_0\right)\end{array}\right)\frac{d\theta_0}{dp}-2f\left(\theta_0\right)\right)dp\\=\;&\left(e'\left(\rho\right)G\left(\rho v\right)+ve\left(\rho\right)g\left(\rho v\right)\right)f\left(\theta_0\right)d\rho\end{aligned}$$

The lhs is negative by SOC, and rhs is positive by inspection, therefore $\frac{dp}{d\rho}<0$.

Appendix C - proof of Proposition 2

To find the level of the ability to aggregate data (directly translated into deterioration of the ad relevance and ultimately into conversion probability ρ) that makes the platform indifferent between choosing the two business models, we need to compare the platform's profit from setting $p = 0$,

$$e(\rho) G(\rho v)(1 - F(\theta_0)),$$

with the profit from adopting a subscription-based business model and setting a positive p,

$$(p(\rho) + e(\rho) G(\rho v))(1 - F(\theta_1(\rho))).$$

Rearranging, we get

$$e(\rho) G(\rho v)(F(\theta_1(\rho)) - F(\theta_0(\rho))) = p(\rho)(1 - F(\theta_1(\rho))),$$

an equality that determines the indifference level of ρ and is described in the main text.

A slight decrease from this level will result in consumer surplus going (down) to

$$\int_{\theta_1}^{1} \left(u(x, \theta_1) - \varepsilon - t + \int_0^{\rho v} (\rho v - y)\, dG(y) \right) dF(x)$$

from

$$\int_{\theta_0}^{1} \left(u(x, \theta_0) + \int_0^{\rho v} (\rho v - y)\, dG(y) \right) dF(x).$$

Because $\theta_0 < \theta_1$, the drop in CS is

$$\int_{\theta_0}^{\theta_1} \left(u(x, \theta_0) + \int_0^{\rho v} (\rho v - y)\, dG(y) \right) dF(x)$$
$$+ \int_{\theta_1}^{1} (u(x, \theta_0) - u(x, \theta_1) + \varepsilon + t)\, dF(x),$$

each component of which is intuitively described in the main text. Both components are strictly positive and inframarginal, QED.

Appendix D - platform competition with single-homing

D1.The setting and best responses

Given the joint distribution F of the attribute θ, for each price vector p, the set of indifferent between the two platforms users will be $(\theta_1, \theta_2) : u_1(\theta_1, D_1) - p_1 = u_2(\theta_2, D_2) - p_2$, where D_i now denotes consumer demand served by platform i. The set of the indifferent between joining one of the platforms users will be

$(\theta_i, \theta_{-i}) : u_i(\theta_i, D_i) - p_i = 0$. The demands will therefore be generally defined by

$$D_1 = \int_{\max\left\{u_1^{-1}(p_1)|D_1,0\right\}}^{1} \int_0^{\max\left\{\min\left\{u_2^{-1}(p_2-p_1+u_1(x,D_1))|D_2,1\right\},0\right\}} dF(x,y),$$

and analogously D_2.

Since this is obviously an extremely complicated setting because of the way network externalities are defined, we consider a very straightforward example with $u_1 = \theta_1 + \alpha D_1 - p_1$ and $u_2 = \theta_2 - p_2$, i.e. when the intragroup network externalities are relevant for one of the platforms only. The demand becomes

$$D_1 = \int_{\max\{p_1-\alpha D_1,0\}}^{1} F_2^x(\max\{\min\{p_2 - p_1 + \alpha D_1 + x, 1\}, 0\})\, dF_1(x),$$

where F_2^x stands for the distribution of the second attribute conditional on $\theta_1 = x$.

For the uniform distribution, this further becomes

$$D_1 = \int_{\max\{p_1-\alpha D_1,0\}}^{1} (\max\{\min\{p_2 - p_1 + \alpha D_1 + x, 1\}, 0\})\, dx.$$

Going through different possibilities for the parameter combinations and subscription fees of the second platform, we arrive at the following demand function:

$$D_1 = \begin{cases} 1, & \text{if } \quad p_1 \le p_2 + \alpha - 1; \\ \frac{1}{\alpha^2}\left(\begin{array}{c} \alpha + \sqrt{-2\alpha + 2\alpha^2 - 2\alpha p_1 + 2\alpha p_2 + 1} \\ +\alpha p_1 - \alpha p_2 - 1 \end{array}\right), & \text{if } \quad 2(p_2 + \alpha - 1) < 2p_1 < \alpha\left(2p_2 - p_2^2 + 1\right); \\ \frac{1}{2\alpha-2}\left(2p_1 - 2p_2 + p_2^2 - 1\right), & \text{if } \quad \alpha\left(2p_2 - p_2^2 + 1\right) < 2p_1 < 2p_2 - p_2^2 + 1; \\ 0, & \text{if } \quad 2p_1 > 2p_2 - p_2^2 + 1. \end{cases}$$

Platform maximizes the sum of subscription fees and advertising revenues

$$(p_i + e(\rho))\, D_i(p_i, p_j, \alpha),$$

where we for simplicity assume that $G(0) = 1$.

Performing the necessary optimization, we arrive at the following best response p_1:

$$\begin{array}{ll} p_2 + \alpha - 1, & \text{if } \quad p_2 + \alpha \ge 1; \\ \frac{1}{\alpha}\left(\alpha + \sqrt{-2\alpha + 2\alpha^2 - 2\alpha p_1 + 2\alpha p_2 + 1} + \alpha p_1 - \alpha p_2 - 1\right) & \\ \quad + (p_1 + e(\rho))\left(\frac{-1}{\sqrt{-2\alpha+2\alpha^2-2\alpha p_1+2\alpha p_2+1}} + 1\right) \le 0, & \text{if } \quad 2(p_2 + \alpha - 1) \le 2p_1 \le \alpha\left(2p_2 - p_2^2 + 1\right); \\ p_2 - \frac{p_2^2-1}{2} - e(\rho), & \text{if } \quad \alpha\left(2p_2 - p_2^2 + 1\right) \le 2p_2 - p_2^2 + 1 - 2e(\rho); \\ \mathit{any}, & \text{if } \quad 2p_1 > 2p_2 - p_2^2 + 1. \end{array}$$

Going through the same procedure for the second platform, we get the following best response with p_2 (implicitly) defined as

$$\begin{cases} -\frac{1}{\alpha}\begin{pmatrix} \alpha+\sqrt{2\alpha^2-2\alpha-2\alpha p_1+2\alpha p_2+1} \\ -\alpha^2+\alpha p_1-\alpha p_2-1 \end{pmatrix} \\ \leq (p_2+e(\rho))\left(\frac{1}{\sqrt{2\alpha^2-2\alpha-2\alpha p_1+2\alpha p_2+1}}-1\right), & \text{if} \quad \begin{matrix} 1-\alpha+p_1 \geq p_2, \\ 2p_2-p_2^2 \geq \frac{2p_1}{\alpha}-1; \end{matrix} \\ 2\alpha-2p_1+p_2+\alpha p_2-\alpha p_2^2-1=(p_2+e(\rho))\cdot \\ \cdot\begin{pmatrix} \frac{1}{1-p_2}\left(2\alpha-2p_1+p_2+\alpha p_2-\alpha p_2^2-1\right) \\ -1-\alpha+2\alpha p_2 \end{pmatrix}, & \text{if} \quad \begin{matrix} 2p_2-p_2^2 \leq \frac{2p_1}{\alpha}-1, \\ 2p_1-1 \leq 2p_2-p_2^2, \\ 2p_2-\alpha p_2^2 \geq 2p_1-\alpha; \end{matrix} \\ -\begin{pmatrix} \alpha^2 p_2^4-4\alpha^2 p_2^3+6\alpha^2 p_2^2+4\alpha^2 p_2-7\alpha^2 \\ +4\alpha p_1 p_2^2-8\alpha p_1 p_2+4\alpha p_1 \\ -4\alpha p_2^2-8\alpha p_2+12\alpha+4p_1^2-8p_1+8p_2-4 \end{pmatrix} \\ =(p_2+e(\rho))\begin{pmatrix} 4\alpha^2 p_2^3-12\alpha^2 p_2^2+12\alpha^2 p_2+4\alpha^2 \\ +8\alpha p_1 p_2-8\alpha p_1-8\alpha p_2-8\alpha \end{pmatrix}, & \text{if} \quad \begin{matrix} 2p_2-\alpha p_2^2 \leq 2p_1-\alpha, \\ 2p_2-p_2^2 \leq \frac{2p_1}{\alpha}-1, \\ 2p_1-1 \leq 2p_2-p_2^2; \end{matrix} \\ \frac{1-e(\rho)}{2}, & \text{if} \quad p_1 > 1+\frac{e(\rho)}{2}-\frac{(1+e(\rho))^2}{8}. \end{cases}$$

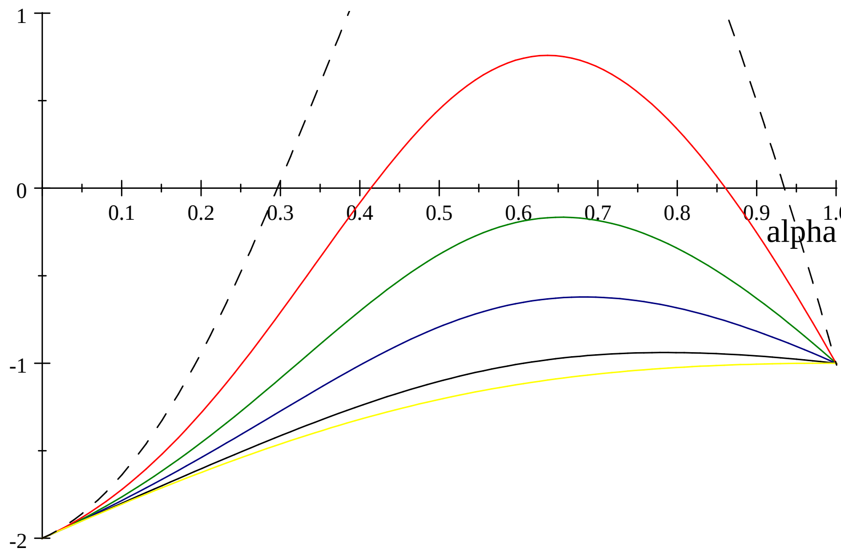
D2. Equilibria

Both platforms are ad-funded

The equilibrium with both firms charging zero price to their subscribers results when the following condition is satisfied

$$b-(1-\alpha) \leq \alpha e(\rho)\left(\frac{1}{b}-1\right), \tag{4}$$

where $b := \sqrt{1-2\alpha(1-\alpha)}$. The condition is obtained from intersection of best responses of the two platforms at $p_1 = 0, p_2 = 0$ under the observation that.$b-(1-\alpha) > \alpha^2 - b + (1-\alpha), \alpha \in (0,1)$. From simple inspection of the condition it is clear that higher values of ρ imply larger set of values for α for which the condition is satisfied. The following figure also illustrates that this equilibrium is 'more likely' to result when α is large. Here, the positive value on y-axis means that the equilibrium exists. Curves of different colours stand for different values of advertisement fee. Yellow is for no advertising, black for $e(\rho) = \frac{1}{2}$; blue for $e(\rho) = 2$; green for $e(\rho) = 4$; red for $e(\rho) = 8$; dashed black for $e(\rho) = 16$.



We observe that for $e(\rho) \leq 4$ there is no value of α that would make this equilibrium possible. More precisely, the condition for the existence of equilibrium at least for some value of α can be found by simultaneously putting to zero the expression (4) and its derivative to get $e(\rho) > 4.7218$. This effectively completes the proof of the statement of Proposition 1C.

Only one platform is ad-funded

An equilibrium with $\{p_1 > 0, p_2 = 0\}$ exists whenever $\exists p_1 < \frac{\alpha}{2}$such that it simultaneously satisfies the two restrictions below:

$$\begin{aligned} \frac{1}{\alpha}\left(\alpha + b\left(p_1\right) + \alpha p_1 - 1\right) + \left(p_1 + e\left(\rho\right)\right)\left(1 - \frac{1}{b\left(p_1\right)}\right) &= 0, \\ -\frac{1}{\alpha}\left(\alpha + b\left(p_1\right) - \alpha^2 + \alpha p_1 - 1\right) - e\left(\rho\right)\left(\frac{1}{b\left(p_1\right)} - 1\right) &\leq 0; \end{aligned}$$

with $b(p_1) := \sqrt{1 - 2\alpha(1 - \alpha + p_1)}$. Intuitively, the first condition is for the first platform to prefer positive price; the second condition is for the second platform to prefer zero price. By rearranging these conditions, we can establish a necessary condition on α and $e(\rho)$ for the two restrictions to be satisfied

$$\frac{1}{\frac{1}{b} - 1} > \frac{2e\left(\rho\right)}{\alpha} > \frac{1}{\frac{1}{b\left(\frac{\alpha}{2}\right)} - 1} - \frac{1}{2}.$$

For illustration, we consider $\alpha = 0.3$, for which the all-ad-funded equilibrium does not exist for $e(\rho) = 8$ or below. In this case, there exists an equilibrium with $\{p_1 = 0.05, p_2 = 0\}$ when $e(\rho) = 0.49171$.

Another possibility for the existence of $\{p_1 > 0, p_2 = 0\}$ equilibrium is when $e(\rho) < \frac{1-\alpha}{2}$ and $p_1 = \frac{1}{2} - e(\rho)$. Continuing the example above, that implies $e(\rho) < 0.35$ and $p_1 > 0.15$. The tendency is apparent: lower advertisement fee (resulting, for example, from the lower data quality) results in less intense competition between the two platforms and therefore higher subscription fee for one of the platforms.

Formally, for the first case ($e(\rho) \geq \frac{1-\alpha}{2}$), we can show that

$$\frac{dp_1}{de(\rho)} = \frac{\frac{1}{b(p_1)} - 1}{2\left(1 - \frac{1}{b(p_1)}\right) - \frac{\alpha}{(b(p_1))^3}(p_1 + m)} < 0,$$

whereas for the second case ($e(\rho) < \frac{1-\alpha}{2}$), it is straightforward that $\frac{dp_1}{de(\rho)} = -1$. This effectively completes the proof of Proposition 3.

No platform is ad-funded

An equilibrium with $\{p_1 > 0, p_2 > 0\}$ exists whenever $\exists p_1, p_2 : 2(p_2 + \alpha - 1) < 2p_1 < \alpha\left(2p_2 - p_2^2 + 1\right)$ and the two restrictions below are satisfied:

$$\begin{aligned} \frac{1}{\alpha}\left(\alpha + b(p_1,p_2) + \alpha p_1 - \alpha p_2 - 1\right) &= (p_1 + e(\rho))\left(\frac{1}{b(p_1,p_2)} - 1\right), \\ -\frac{1}{\alpha}\left(\alpha + b(p_1,p_2) - \alpha^2 + \alpha p_1 - \alpha p_2 - 1\right) &= (p_2 + e(\rho))\left(\frac{1}{b(p_1,p_2)} - 1\right); \end{aligned}$$

with $b(p_1,p_2) = \sqrt{2\alpha^2 - 2\alpha - 2\alpha p_1 + 2\alpha p_2 + 1}$. Intuitively, this is an intersection of the two best responses when both platforms decide to charge their users positive price and when there is a full coverage. (no participation constraint is binding),

Denoting $p = p_1 + p_2$, and $q = p_1 - p_2$, we can rewrite the restrictions above as

$$\begin{aligned} \alpha &= (p + 2e(\rho))\left(\frac{1}{b(q)} - 1\right) \\ \frac{2}{\alpha}\left(\alpha + b(q) + \alpha q - 1\right) &= \alpha + q\left(\frac{1}{b(q)} - 1\right) \end{aligned}$$

where $b(q) = \sqrt{2\alpha^2 - 2\alpha - 2\alpha q + 1}$. Note that we can first find q from the second equality above and it does not depend on advertising fee $e(\rho)$. This fact allows for a simple comparative static from the first equality above. In particular, $\frac{dp}{de(\rho)} = -2$ since $b(q)$ does not depend on $e(\rho)$. This has a very intuitive meaning: since there is a full coverage, a change in advertising fee does not have an effect on total demand, and it changes the some of prices by twice as much as it changes each price one-to-one.

D3. The role of transaction costs

Consider the setting described in Appendix D1, but with transaction cost t that each user has to incur in case she pays a positive subscription fee. Starting from the candidate equilibrium in which both platforms are ad-funded, the interior part of the demand for the services of the first platform deviating from this equilibrium can be written as follows:

$$D_1 = \begin{cases} \frac{1}{\alpha^2}\left(\alpha a + \sqrt{1-2\alpha(a-\alpha)} - 1\right), & \text{if} \quad p_1 \le \frac{\alpha}{2} - t; \\ \frac{1}{\alpha^2}\left(\alpha b - \sqrt{1+2\alpha b} + 1\right), & \text{if} \quad p_1 \ge \frac{\alpha}{2} - t. \end{cases}$$

Here, $a := 1 + t + p_1$ and $b = t + p_1 - 1$.

The optimal (interior) deviation can then be described as

$$\begin{cases} \begin{array}{c} \frac{1}{\alpha}\left(\alpha a + \sqrt{1-2\alpha(a-\alpha)} - 1\right) \\ = (p_1 + e(\rho))\left(\frac{1}{\sqrt{1-2\alpha(a-\alpha)}} - 1\right), \end{array} & \text{if} \quad p_1 \le \frac{\alpha}{2} - t; \\ \begin{array}{c} \frac{1}{\alpha}\left(\alpha b - \sqrt{1+2\alpha b} + 1\right) \\ = (p_1 + e(\rho))\left(\frac{1}{\sqrt{1+2\alpha b}} - 1\right), \end{array} & \text{if} \quad p_1 \ge \frac{\alpha}{2} - t. \end{cases}$$

In the main body of the paper, we formulate a proposition that proves that the deviation described above is profitable for a smaller set of parameter values when transaction costs are positive compared to the case when they are zero.

Appendix E - platform competition with multi-homing

E1. Derivation of demand

Given the joint distribution F of the attribute θ, for each price vector p, the share of users of the type (θ_1, θ_2) that is served by platform 1 is

$$s_1(\theta, p) \begin{cases} 0, & \text{if} \quad p_1 > u_1(\theta_1, D_1) \\ \frac{u_1(\theta_1,D_1)-p_1}{u_1(\theta_1,D_1)-p_1+u_2(\theta_2,D_2)-p_2}, & \text{if} \quad \min\{u_1(\theta_1, D_1) - p_1, u_2(\theta_2, D_2) - p_2\} > 0 \\ 1, & \text{if} \quad p_2 > u_2(\theta_2, D_2), p_1 \le u_1(\theta_1, D_1) \end{cases}$$

and the corresponding demand is $D_1 = \int_0^1 \int_0^1 s_1(x, y, p)\, dF(x, y)$. With our simple utility functions and uniform distribution, and when both platforms charge zero price, the demand for the services of the first platform is $D_1 = \int_0^1 \left(\int_0^1 \frac{x+\alpha D_1}{x+y+\alpha D_1} dy\right) dx$. Computing the integral, we get the demand implicitly defined by

$$D_1 = \frac{1}{2}\alpha^2 D_1^2 \ln \frac{\alpha D_1}{(\alpha D_1 + 1)^2}(\alpha D_1 + 2) + \alpha D_1 \ln \frac{1}{\alpha D_1 + 1}(\alpha D_1 + 2) + \frac{1}{2}.$$

One can establish that $D_1 > \frac{1}{2}$, for example, for $\alpha = 0.1$, $D_1 = 0.532\,25$. Because this candidate equilibrium is characterised by full coverage, we also have $D_2 = 1 - D_1$.

E2. To proposition 1C

When a platform deviates from zero price equilibrium, two things happen. First, it loses completely all the users with $p_i > \theta_i + \alpha D_i \mathbb{I}_{i=1}$. Second, it loses part of the attention of all other users by the amount $\frac{\theta_i + \alpha D_i \mathbb{I}_{i=1}}{\theta_1 + \alpha D_1 + \theta_2} - \frac{\theta_i + \alpha D_i \mathbb{I}_{i=1} - p_i}{\theta_1 + \alpha D_1 + \theta_2 - p_1}$. On both accounts, the associated loss of profit is $e(\rho)\,\Delta D_i$. The gain in profit from charging the remaining users a positive price is $p_i D_i(p_i, 0)$. The candidate equilibrium 'all ad-funded' is therefore deviation-proof if $e(\rho)\,\Delta D_i > p_i D_i(p_i, 0)$. In terms of first-order effect of data quality, it is clear that higher quality makes deviation less attractive as it only enters the loss term. We therefore confirm Proposition 1C to the first-order in its part related to the advertising fee, but do not make any statement about the effect of α.

E3. To proposition 3

In an equilibrium where two platforms choose different business model, the price of the fee-funded one is found from FOC:

$$D_i(p_i, 0) + (p_i + e(\rho))\frac{\partial D_i}{\partial p_i} = 0$$

A standard comparative statics exercise tells us that the sign of $\frac{dp_i}{de}$ is the same as that of $\frac{\partial D_i}{\partial p_i}$, and so proposition 3 is confirmed.

E4. To proposition 3A

In equilibrium with no ad-funded platform, we have

$$\begin{aligned} D_1(p_1, p_2) + (p_1 + e(\rho))\frac{\partial D_1}{\partial p_1} &= 0 \\ D_2(p_1, p_2) + (p_2 + e(\rho))\frac{\partial D_2}{\partial p_2} &= 0 \end{aligned}$$

The comparative statics exercise then leads us to the following system:

$$\begin{aligned} \left(\frac{\partial D_1}{\partial p_1} + (1 + p_1 + e) D_{11}\right)\frac{dp_1}{de} &= -\frac{\partial D_1}{\partial p_1} - (1 + p_1 + e) D_{12}\frac{dp_2}{de} \\ \left(\frac{\partial D_2}{\partial p_2} + (1 + p_2 + e) D_{22}\right)\frac{dp_2}{de} &= -\frac{\partial D_2}{\partial p_2} - (1 + p_2 + e) D_{21}\frac{dp_1}{de} \end{aligned}$$

Solving for $\frac{dp_1}{de}$, $\frac{dp_2}{de}$, we see that their sign is not unambiguous even after involving optimality condition $2\frac{\partial D_i}{\partial p_i} + (p_i + e)\, D_{ii} < 0$.

$$
\begin{aligned}
&\left(\frac{\partial D_1}{\partial p_1} + (1 + p_1 + e)\, D_{11} - \frac{(1 + p_1 + e)\,(1 + p_2 + e)\, D_{21} D_{12}}{\frac{\partial D_2}{\partial p_2} + (1 + p_2 + e)\, D_{22}} \right) \frac{dp_1}{de} \\
= \quad & -\frac{\partial D_1}{\partial p_1} + \frac{(1 + p_1 + e)\, D_{12} \frac{\partial D_2}{\partial p_2}}{\frac{\partial D_2}{\partial p_2} + (1 + p_2 + e)\, D_{22}}
\end{aligned}
$$

However, considering a change in data quality for a single firm i, we get that $\frac{dp_i}{de}$ still has the same sign as $\frac{\partial D_i}{\partial p_i}$ and proposition 3A can be confirmed in this sense.