User Data and Platform Competition

Philipp Dimakopoulos
Humboldt-University of Berlin[*]

Slobodan Sudaric
Humboldt University of Berlin[†]

January 10, 2017

Abstract

We analyze platform competition for users and advertisers where platforms collect user data which is used for improved ad-targeting. Considering that users incur privacy costs when providing data and face reduced nuisance costs when seeing more relevant ads, we show that the equilibrium amount of data provision is distorted when compared to the efficient level and can be inefficiently high or low. This distortion depends on the opposing cross-group externalities as well as the degree of competitiveness on each market side: if overall competition is weak or if targeting benefits are relatively low, too much data is collected, and vice-versa. Further, we find that softer competition on either side of the market increases the equilibrium level of data, which implies substitutability between competition policy measures on both market sides. Nevertheless, in a situation of over-provision of user data it is more effective to strengthen platform competition for advertisers.

JEL–classification numbers: D43, L13, L40, L86

Keywords: platform competition, user data, nuisance costs, ad targeting, privacy costs

This is a preliminary and unfinished draft. Please do not cite or distribute without the consent of the authors.

[*]philipp.dimakopoulos@hu-berlin.de

[†]sudarics@hu-berlin.de.

1 Introduction

Online platforms often do not charge monetary prices for consumers but monetize through an advertisement-based business model. The role of user data in this context is ambiguous. From the platform perspective user data is an input factor which can be used to gain insights about consumers and improve the targeting of advertisement resulting in a superior product for potential advertisers. This commodity-like attribute of data is mirrored to a lesser extent on the consumer side. Consumers typically accept some conditions to what extent personal data is collected and processed when using a platform service. In some cases the provision of personal data is necessary to make meaningful use of a platform service (e.g. social networks) while in other cases services do not require the collection of user data per se (e.g. search engines, mail providers). In both cases the provision of data from a consumer perspective can be interpreted as a price the consumer is willing to accept in exchange for the use of the platform including the display of ads. To put it in terms of platform economics, user data requirements exhibit price characteristics on the one hand, and affect directly the cross-group externalities of both market sides. This ambiguity makes it especially hard for policy makers as standard economic arguments might not be applicable. We therefore want to shed some light on the role of competition intensity in a two-sided market setting when users provide data and this data is monetized through advertising.

We analyze a setting of two competing ad-financed platforms in a two-sided market framework, where consumers agree to provide some specific level of data and platforms process user data to offer improved targeting to advertisers. Consumers incur disutility from providing data (privacy concerns, opportunity costs) but benefit from reduced nuisance due to seeing more relevant ads. Consumers and advertisers are assumed to single-home. We show that platforms will extract a distorted amount of data compared to the efficient benchmark. The distortion is induced through the one-sided monetization in a way that platforms do not perfectly balance the costs of data provision against the benefits on both sides of the market but put too much or too little weight on the targeting benefit. This distortion depends on the net effect of cross-group externalities as well as the degree of competition intensity on both market sides. If nuisance costs exceed the targeting benefit or platforms have significant market power, an inefficiently high level of data is collected. On the other hand, if competition is strong or targeting benefits sufficiently outweigh nuisance costs too little data is collected. From the point of view of consumers the competitive level of data provision is always too high, suggesting that applying a consumer standard to online platforms leads to underprovision of personal data. The competitive equilibrium level of data provision, however, is monotone in the the degree of competition intensity: the weaker the competition on either side of the market the higher the equilibrium amount of data provision. Our results indicate that first-best can be achieved by careful regulation and while competition policy measures on both market sides are substitutable there is a difference in effectiveness: in a situation of over-provision of user data, it is more effective to strengthen platform competition for advertisers than for users, and vice-versa. Lastly, we show that if platforms would introduce two-sided pricing, the resulting equilibrium data level is efficient and show that there is always overprovision of personal data if platforms collude.

Our research is closely related to the literature on platform competition in media markets. Anderson and Coate (2005) analyze competition for consumers and advertisers in a TV market setting and show that tv stations show too much or too little advertisments depending on the competitive structure of the market and the perceived nuisance from consumers. However, in their model advertisers multi-home as consumers differ in taste and different types of consumers can be reached by advertising on multiple TV stations while in our model advertisers single-home. Platform competition with two-sided single-homing has been analyzed by Armstrong (2006) in a more general framework which was later extended in Armstrong and Wright (2007). However, both papers consider a general setup where platforms engage in two-sided pricing and non-monetary aspects (as e.g. user data) are not modelled. A setting which is more closely related to our paper is presented in Anderson et al. (2012) where platforms are financed through ads and users incur nuisance costs, however the focus of their paper differs substantially. The concept of user attention is captured in Reisinger (2012) where consumers spend time using platform services and platforms translate this activity into better targeting and reduced nuisance. A similar setup is presented in Bourreau et al. (2016), however the research question differs substantially. Although our model is very closely related to the two previously mentioned papers, the fact that the amount of data is a strategic variable chosen by the platforms in our model, changes the competitive dynamics substantially. Regarding the competitive structure in media markets recent work focuses on multi-homing of consumers as in Athey et al. (2014) and Ambrus et al. (2014), which we do not allow by assumption. Our model can therefore be seen as a complement to the literature in a sense that it focuses on the role of competition intensity in a setting where cross-group externalities (targeting and nuisance) are directly affected by strategic platfrom behavior.

The remaining paper is structured as follows. Section 2 introduces the model. Section 3 characterizes the equilibrium for which we present comparative statics in section 4. Section 5 compares the equilibrium results to the welfare optimal benchmark and outlines policy implications. In section 6 we we extend the baseline model with respect to two-sided pricing and collusion. Section 7 concludes.

2 Model

We analyze a setting where two symmetric platforms $i = \{1, 2\}$ compete for advertisers and consumers. Advertisers and consumers are distributed uniformly on different Hotelling lines of unit length and are assumed to both single-home. This assumption is chosen to study a more competitive environment. Platforms are located at the ends of the respective Hotelling lines such that platform i is located at location $l_i = 0$ and platform $j \neq i$ at $l_j = 1$.

2.1 Users

Users obtain utility $u_i(x)$ from joining platform i where

$$u_i(x) = \underline{u} - \kappa(d_i) - \nu(d_i)a_i - t_u|l_i - x|. \tag{1}$$

The first term of the utility function is a fixed utility component $\underline{u}$ from using platform services, which is the same at both platforms. Second, $\kappa(d_i) \geq 0$ denotes the privacy costs (concerns) of providing user data d_i to the platform, whereby we assume non-concavity on these costs, i.e. $\kappa'(d_i) \geq 0$ and $\kappa''(d_i) \geq 0$. Third, users incur nuisance cost $\nu(d_i) \geq 0$ per advertisement a_i on the platform. These nuisance costs fall in provided data d_i as we assume that users dislike personalized data less, i.e. $\nu'(d_i) < 0$ and $\nu''(d_i) > 0$. The more personalized and relevant an ad, the higher the chance of value creation through a possible follow-up purchase.[1] Finally, users face transportation costs due to horizontal platform differentiation, whereby a consumer's location on the Hotelling line is denoted by $x \in [0, 1]$ and $t_u > 0$ is the associated transportation cost parameter.

Consumers in our model are not charged a monetary price explicitly, which makes our model comparable to e.g. Reisinger (2012). We follow the same line of reasoning as e.g. in Peitz and Reisinger (2016) and Waehrer (2015), that there are some exogeneous constraints preventing platforms from charging non-zero consumer prices. This restriction is, however, relaxed in section 6.1. In order to join a platform consumers have to provide some personal data d_i in our model. This is different to the setup in Reisinger (2012) or Bourreau et al. (2016) as in our model platforms can set the level of data which has to be provided by the consumers, whereas in their models consumers voluntarily provide a certain amount of time. The idea behind our setup is, that consumers accept terms and conditions when using a platform which requires them to accept a certain level of data provision or alternatively cases where users have to register for an account by providing personal information before they can use the platform service. This specification on the consumer side allows us to focus on user data d_i as primary strategic aspect for competition.

2.2 Advertisers

Advertisers are assumed to single-home when deciding where to advertise and obtain expected profit of $\pi_i(a)$ from posting a single ad with

$$\pi_i(a) = \int_{x \in x_i} \tau(d_i)(\alpha - p_i)dx - t_a|l_i - a|$$

and where an interaction with a consumer $x \in x_i$ generates expected revenue of α if the consumer decides to 'click on the ad' which happens with probability $\tau(d_i)$. The function $\tau(d_i) \geq 0$ can also be interpreted as the targeting ability of platforms: the more data d_i

[1] Note that our set-up allows for positive utility of advertisement as well, as long as this positive utility is again concave in the amount of provided data d_i, such that $\tilde{\nu}(d_i) = -\nu(d_i)$ and $\tilde{\nu}(d_i) \in \mathbb{R}$, while $\tilde{\nu}'(d_i) > 0$ and $\tilde{\nu}''(d_i) < 0$. However, for sake of clarity we stay with the notion of negative utility of nuisance in the subsequent text.

can be collected from users, the more effective the targeting and hence the higher the probability that a user clicks on this ad, i.e. we have that $\tau'(d_i) > 0$ and $\tau''(d_i) > 0$. For simplification we normalize α to one. At the same time we assume that advertisers only pay the platform if the ad has been clicked (cost-per-click) such that the expected revenue per consumer is given by $\tau(d_i)(1-pi)$, which is consistent with real-world pricing practices.

Profits can then be rewritten as

$$\pi_i(a) = \tau(d_i)(1-p_i)x_i - t_a|l_i - a| \tag{2}$$

where x_i denotes the number of consumers at platform i and $t_a|l_i - y|$ are the transportation cost when using platform i as opposed to platform $j \neq i$ while t_a denotes the transportation cost parameter on the advertiser side.

Note that on the advertiser and the user side we have different parameters of transportation costs, which we will later interpret as different competition intensities on each side. The competitive environment and hence the relevant market of platforms when competing for users may be different from the market when competing for advertisers. For example, different online platforms, like search engines, social networks, video streaming platforms or mail providers, may all compete for the same advertisers, however competition for users may occur separately and independently of the other segments.

2.3 Platforms

The business model of platforms in our model is purely data driven. While they offer (exogeneous) platform services ($\underline{u}$) to consumers, revenue is only generated through presenting ads to consumers, i.e.

$$\Pi_i\left(d_i, p_i\right) = a_i\tau(d_i)p_ix_i \tag{3}$$

where a_i denotes the number of advertisers at platform i and p_i is the per-click price advertisers have to pay if x_i consumers click with probability $\tau(d_i)$. The crucial novelty in our model is that we assume that besides charging prices platforms also extract data d_i from their consumers. While d_i shares some price characteristics from the point of view of consumers, data is a crucial input factor for the click-probability the advertisers are facing. At the same we assume that not only the click probability increases through better targeting possibilities but also the nuisance decreases.

2.4 Assumptions

We make the following assumption to assure full *advertiser* market coverage.

Assumption 1 *Competition for advertisers is sufficiently strong, i.e.* $t_a \leq \bar{t_a}$.
For this, it is necessary that competition for users is sufficiently weak and that there are gains of trade for all advertisers, even without data collection, i.e.

$$(a) \quad t_u > \nu(0)$$

$$(b) \quad t_a < \tau(0)$$

The upper bound on t_a is given by $\bar{t_a} := \frac{t_u\tau(0)-\nu(0)\tau(0)}{3t_u+\nu(0)}$. This assumption on the upper bound of t_a is merely technical, allowing us to isolate effects in a competitive environment. Intuitively, this constitutes a sufficient condition, such that for any level of (symmetric) data provision $d \geq 0$, it is assured that advertisers have enough market power to obtain positive profits such that their market is fully covered. Consequently, competition for advertisers in sufficiently strong.

The condition on the consumer nuisance function, i.e. the necessary condition (a) of assumption 1, can be motivated as follows: no platform will obtain the entire user market, even if all ads were placed on the rival platform. Technically, this is established by $t_u > \nu(0)$. Given any (symmetric) amount of data $d \geq 0$ collected by both platforms, even if all advertisers used platform j such that $a_i = 0$ and $a_j = 1$, at least the user most loyal to platform j, i.e. located directly at l_j, would rather stay at this platform j, even though it is full of ads. In other words, competition for users is sufficiently weak.

The condition on the targeting technology, i.e. the necessary condition (b) of assumption 1, states that even without collecting any data advertisers can still profitably join a platform. In particular we assume that there are gains of trade for all advertisers. Intuitively, this assumption states that there is a positive probability for consumers to click an ad even if the ad is not targeted at all. And this probability, $\tau(0)$, exceeds the transportation cost incurred by any advertiser t_a, so that we need not exclude any advertisers, even if too little data is collected.

To ensure full participation on the *user* side we have assumed that $\underline{u}$ is large enough, i.e. the platform service provides a high enough utility such that consumers are not deterred through the provision of personal data and the incurred nuisance from advertising.

The timing of the game is as follows. In the first stage platforms set prices and the required amount of data to join their platform simultaneously. In the second stage advertisers and consumers observe the platforms' choices and decide which platform to join simultaneously. [2] The solution concept is sub-game perfection and we solve the game by backward induction.

3 Equilibrium

The market shares on the consumer and advertiser side are given by the standard Hotelling procedure. Utilizing the unit length of the Hotelling line the number of consumers x_i joining platform i is then determined by the indifferent consumer $\hat{x} : u_i(\hat{x}, \theta) = u_j(\hat{x}, \theta)$ such that

$$x_i = \hat{x} \qquad = \frac{1}{2} + \frac{1}{2t_u}\left[\kappa(d_j) - \kappa(d_i) + \nu(d_j)a_j - \nu(d_i)a_i\right] \tag{4}$$

$$x_j = 1 - \hat{x} \qquad = \frac{1}{2} + \frac{1}{2t_u}\left[\kappa(d_i) - \kappa(d_j) + \nu(d_i)a_i - \nu(d_j)a_j\right] \tag{5}$$

Similarly, market shares on the advertiser side are given by the indifferent advertiser $\hat{a} : \pi_i(\hat{a}) = \pi_j(\hat{a})$. Note, that Assumption 1 assures market coverage gross of advertising prices. We for now ignore prices and check later that in equilibrium that market is covered even after financial transfers to the platforms. Market shares are then given by

$$a_i = \hat{a} \qquad = \frac{1}{2} + \frac{1}{2t_a}\left[\tau(d_i)(1-p_i)x_i - \tau(d_j)(1-p_j)x_j\right] \tag{6}$$

$$a_j = 1 - \hat{a} \qquad = \frac{1}{2} + \frac{1}{2t_a}\left[\tau(d_j)(1-p_j)x_j - \tau(d_i)(1-p_i)x_i\right] \tag{7}$$

[2] We could also consider an alternative timing where advertisers choose first and consumers last like in (examples, examples). The outcome is equivalent in our model.

Solving the system of equations given in (4) - (7) we obtain equilibrium market shares of $x_i(d_i, d_j, p_i, p_j), x_j(d_i, d_j, p_i, p_j)$ and $a_i(d_i, d_j, p_i, p_j), a_j(d_i, d_j, p_i, p_j)$.
Platforms then maximize their profits

$$\max_{p_i, d_i} \Pi_i(d_i, p_i) = a_i(d_i, d_j, p_i, p_j)\ \tau(d_i)\, p_i\, x_i(d_i, d_j, p_i, p_j)\ \ \forall i \in \{1, 2\} \tag{8}$$

and we obtain a symmetric solution $p_i^* = p_i^* = p^*$ and $d_i^* = d_i^* = d^*$ from the first order conditions. Regarding the curvature of the maximization problem we can roughly say that our solution[3] represents a maximum as long as the targeting technology $\tau(\cdot)$ is sufficiently concave, the nuisance cost $\nu(\cdot)$ is sufficiently convex or both. The details of this condition are given in the appendix.
The equilibrium amount of data collected from a single consumer is then implicitly given by

$$\kappa'(d^*) = \left(\frac{\nu(d^*) + t_u}{\tau(d^*) - t_a}\right)\frac{\tau'(d^*)}{2} - \frac{\nu'(d^*)}{2} \tag{9}$$

resulting in equilibrium advertiser prices of

$$p^* = 2\frac{t_a t_u + \nu(d^*)\tau(d^*)}{\tau(d^*)\left(t_u + \nu(d^*)\right)} \tag{10}$$

Then, in equilibrium, we get the following advertiser profit $\pi_i^*(a)$, user utility $u_i^*(x)$ and platform profits Π_i^*.

$$\pi_i^*(a) = \frac{\tau(d^*)}{2} - \frac{t_a t_u + \nu(d^*)\tau(d^*)}{t_u + \nu(d^*)} - t_a \min\{a, 1-a\} \tag{11}$$

$$u_i^*(x) = \underline{u} - \kappa(d^*) - \frac{\nu(d^*)}{2} - t_u \min\{x, 1-x\} \tag{12}$$

$$\Pi_i^* = \frac{t_a t_u + \nu(d^*)\tau(d^*)}{2\left(t_u + \nu(d^*)\right)} \tag{13}$$

Note that with a sufficiently high baseline utility $\underline{u}$, user utility will always be non-negative. Further note that the equilibrium price p^* does not exceed one and that advertiser profits

[3]Note that we here we prove equilibrium existence without claiming uniqueness of our symmetric solution.

as given by equation (11) will be positive for all advertisers due to assumption 1. The following Lemma summarizes this finding.

Lemma 1 *In equilibrium,* $p^* < 1$ *and* $\pi_i^*(a) \geq 0$.

Proof. Given equation (10), $p^* < 1$ if

$$2\frac{t_a t_u + \nu(d^*)\tau(d^*)}{\tau(d^*)t_u + \nu(d^*)\tau(d^*)} < 1 \iff t_a < \tau(d^*)\frac{(t_u - \nu(d^*))}{2t_u} < \tau(d^*) \tag{14}$$

By assumption 1 we have that $\tau(d) > t_a$ for all d and therefore in particular also $\tau(d^*) > t_a$. Further, we have that $0 < (t_u - \nu(d^*))/2t_u < 1$, hence the last inequality. Thus, assumption 1 is sufficient, that the expression above holds and $p^* < 1$.

Even the indifferent advertiser with highest transportation costs has positive profits in equilibrium because

$$\pi_i^*(\frac{1}{2}) = \frac{\tau(d^*)}{2} - \frac{t_a t_u + \nu(d^*)\tau(d^*)}{t_u + \nu(d^*)} - \frac{t_a}{2} \geq 0 \iff \tau(d^*)\frac{t_u - \nu(d^*)}{3t_u + \nu(d^*)} \geq t_a, \tag{15}$$

which is guaranteed by assumption 1 for all d and especially for d^*. For this note that the term on the left in the last inequality is increasing in d. ■

Before we continue, we state another lemma concerning the equilibrium effect of data provision on user utility.

Lemma 2 *In equilibrium,* $\kappa'(d^*) > -\nu'(d^*)/2$.

Proof. Rearranging terms in the first-order condition of platform profit maximization, given by equation (9), yields

$$2\kappa'(d^*) + \nu'(d^*) = \tau'(d^*)\frac{\nu(d^*) + t_c}{\tau(d^*) - t_a} > 0. \tag{16}$$

By assumption 1 we have $\tau(d^*) > t_a$. Hence the second term on the right hand side of equation (16) must be positive, such that $2\kappa'(d^*) + \nu\text{'}(d^*) > 0$. ■

Intuitively, Lemma 2 says that in equilibrium users' data provision is such, that the (negative) privacy concern effect on their utility is larger than the (positive) effect of reduced nuisance. Consequently, in the market outcome too much personal data compared to the user-optimal level is provided (compare section 5.2).

4 Comparative Statics

4.1 Effects on prices and collected data in equilibrium

In this subsection, we evaluate how the intensity of competition affects market outcomes, i.e. the price p^* and the amount of data collected d^* in the symmetric equilibrium. For this we have to distinguish between the platform competition intensity on the user side and on the advertiser side. Since we have horizontally differentiated platforms vis-a-vis users as well as advertisers, competition intensity on each side can be measured through the corresponding transportation cost parameter: The higher transportation costs are, the more monopolistic platforms can behave, and the lower competition intensity.

4.1.1 Competition for users

First, we evaluate the effects of user-side competition on data collection. Since we have only implicit solutions for d^*, we make use of the implicit function theorem by totally differentiating the first-order conditions from equations (9) and (10) w.r.t. t_u. Solving for $\mathrm{d}d^*/\mathrm{d}t_u$ yields

$$\frac{\mathrm{d}d}{\mathrm{d}t_u} = \frac{(\tau(d^*) - t_a)\,\tau'(d^*)}{Z(d^*)} > 0, \tag{17}$$

where we define the term in the denominator

$$Z(d^*) := \nu''(d^*)\,(\tau(d^*) - t_a)^2 - \nu'(d^*)\tau'(d^*)\,(\tau(d^*) - t_a) + (\nu(d^*) + t_u)\left[\tau'(d^*)^2 - (\tau(d^*) - t_a)\,\tau''(d^*)\right] > 0 \tag{18}$$

as $\tau'(d^*) > 0$ and $\tau''(d^*) < 0$ while $\nu'(d^*) < 0$ and $\nu''(d^*) > 0$ by construction, and $\tau(d^*) - t_a > 0$ by assumption 1.

Second, we analyze the effects of competition intensity for users on p^*. While we have an explicit solution for p^*, we still need to take into account the second-order effect of t_u on

p^* through d^*. From equation (10), we get for the derivative of p^* w.r.t. t_u

$$\frac{\mathrm{d}p^*}{\mathrm{d}t_u} = 2\frac{-\nu(d^*)\left[\tau(d^*) - t_a\right]\tau(d^*) - \frac{\mathrm{d}d^*}{\mathrm{d}t_u}\left[-\nu'(d^*)\tau(d^*)\left(\tau(d^*) - t_a\right) + \tau'(d^*)\left(t_u + \nu(d^*)\right)t_a\right]t_u}{\left[t_u + \nu(d^*)\right]^2\left(\tau(d^*)\right)^2} < 0, \tag{19}$$

since $t_a < \tau(d^*)$ by assumption 1 and $\mathrm{d}d^*/\mathrm{d}t_u > 0$ as established above.

The following propositions sums up comparative statics of user-side competition intensity.

Proposition 1 *When platform competition for users intensifies,*

- *less user data is collected*
- *the ad price-per-click for advertisers increases.*

Intuitively, lower user transportation costs, i.e. less sticky users, can be interpreted as less platform differentiation, and in other words, stronger platform competition for users. On the one hand, platforms care about the share of users on their platform because it increases their profits directly, but also indirectly through more attracted advertisers. On the other hand, platforms want to increase the amount of user data collected as it enhances targeting, attracts advertisers and hence increases profits. In equilibrium, stronger competition for users impacts the former effect of attracting users more than the latter of increasing targeting, therefore, platforms will collect less user data. Following the same intuition, platforms are willing to reduce advertiser shares in order to not repelling valuable users. Hence, advertiser prices can increase in equilibrium, where due to symmetry market shares are equalized nevertheless. These two results somewhat reflect the “standard” two-sided platform logic, where stronger competition on one side of the market reduces this side’s “price”, while it increases the other side’s.

4.1.2 Competition for advertisers

First, we consider the effects of advertiser-side competition on data collection. Since we have only implicit solutions for d^*, we make use of the implicit function theorem by totally differentiating the first-order conditions from equations (9) and (10) w.r.t. t_a. Solving for

$\mathrm{d}d^*/\mathrm{d}t_a$ yields

$$\frac{\mathrm{d}d^*}{\mathrm{d}t_a} = \frac{(\nu(d^*) + t_u)\,\tau'(d^*)}{Z(d^*)} > 0. \tag{20}$$

Second, we evaluate the effects of competition intensity for advertisers on p^*. While we have an explicit solution for p^*, we still need to take into account the second-order effect of t_a on p^* through d^*. From equation (10), we get for the derivative of p^* w.r.t. t_a

$$\frac{\mathrm{d}p^*}{\mathrm{d}t_a} = 2t_u \frac{[\nu(d^*) + t_u]\,\tau(d^*) - \frac{\mathrm{d}d^*}{\mathrm{d}t_a}\left[-\nu'(d^*)\,(\tau(d^*) - t_a)\,\tau(d^*) + \tau'(d^*)\,(t_u + \nu(d^*))\,t_a\right]}{[t_u + \nu(d^*)]^2\,(\tau(d^*))^2}. \tag{21}$$

Because $\nu'(d^*) < 0$ by construction, $t_a < \tau(d^*)$ by assumption 1 and $\mathrm{d}d^*/\mathrm{d}t_a > 0$ as established above, both terms in the numerator have opposing signs and we need further analysis. For this, insert dd^*/dt_a from equation (20) into $\mathrm{d}p^*/\mathrm{d}t_a$ from equation (21), which simplifies to

$$\frac{\mathrm{d}p^*}{\mathrm{d}t_a} = \frac{-2t_u\,(\tau(d^*) - t_a)\left[-\nu''(d^*)\tau(d^*)\,(\tau(d^*) - t_a) - (\nu(d^*) + t_u)\left(\tau'(d^*)^2 - \tau(d^*)\tau''(d^*)\right)\right]}{(\nu(d^*) + t_u)\,\tau(d^*)^2 Z(d^*)} > 0. \tag{22}$$

Proposition 2 *When platform competition for advertisers intensifies*

- *the ad price-per-click for advertisers falls and*
- *less user data is collected.*

Intuitively, lower advertiser transportation costs mean less sticky advertisers and hence increased platform competition for advertisers. Therefore, it is straightforward that advertiser prices fall. At the same time this would increase the share of advertisers on a platform, thereby repelling users. As a consequence, because in equilibrium user market share and hence advertiser attraction is even more important for platform profits than enhancing targeting, platforms will reduce the level of collected user data, such as not to shy away users. Contrary to the mechanics of comparative statics of user competition, this effect does not follow “standard” two-sided platform logic as here more competition for advertisers *reduces* users' data "payment".

4.2 Effects on platform profits, advertiser profits and user utility

In this subsection we provide further intuition on equilibrium profits and utility by presenting comparative statics.

Proposition 3 *The following table summarizes comparative statics of advertiser-side competition intensity t_a and user-side competition intensity t_u on equilibrium values of personal data provision d^*, ad-per-click price p^*, as well as platform profits Π_i^P, advertiser profits π_i^A and user utility u_i.*

z	$\mathrm{d}d^*/\mathrm{d}z$	$\mathrm{d}p^*/\mathrm{d}z$	$\mathrm{d}\Pi_i^P/\mathrm{d}z$	$\mathrm{d}\pi_i^A/\mathrm{d}z$	$\mathrm{d}u_i/\mathrm{d}z$
t_a	$+$	$+$	$+$	$-$	$-$
t_u	$+$	$-$	$-$	$+$	$-$

Proofs are contained in the following derivations and in Appendix A.2, whenever necessary.

4.2.1 Effects on platform profits

The effects on platform profits $\Pi_i^P = p^*\tau(d^*)\, x_i^*\, a_i^* = (1/4)\, p^*\tau(d^*)$ can be broken down to

$$\frac{\mathrm{d}\Pi_i^P}{\mathrm{d}z} = \frac{1}{4}\left[\frac{\mathrm{d}p^*}{\mathrm{d}z}\tau(d^*) + \tau'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}z}p^*\right]. \tag{23}$$

First, we look at the effects of advertiser competiton intensity. For $z = t_a$ both terms on the right-hand side are positive and hence $\mathrm{d}\Pi_i^P/\mathrm{d}t_a > 0$. Intuitively, when competition for advertisers becomes more intense (t_a decreases), then prices for ad-placing decrease. In turn, less data is collected from users, such that targeting becomes less effective, and less total revenue is made on the ad market. Both these effects decrease platform profits. Second, the effect of user-side competition is less straight-forward. For $z = t_u$, the first term on the right-hand side of (23) is negative, while the second term is positive. Intuitively, when competition for users becomes more intense (t_u decreases), less data can be collected from users, which leads to less effective ad targeting, hence the second term is negative. Nevertheless, the bottleneck position of platforms enables them to increase ad prices on the other side, since user market share is harder to achieve. This is the positive first-term effect, which is stronger in equilibrium[4]. Hence, overall, platforms benefit from harsher competition for users, i.e. $\mathrm{d}\Pi_i^P/\mathrm{d}t_u < 0$. This effect might seem counter-intuitive at first sight. However, it is important to note that platform revenues are exclusively made on the advertiser side and also the fact that platforms constitute a bottleneck for user access as well as personal data.

4.2.2 Effects on advertiser profits

The effects on advertiser profits (net of transportation costs) $\pi_i^A = (1 - p^*)\,\tau(d^*)\,x_i^*\,a_i^* = (1/4)\;(1 - p^*)\;\tau(d^*)$ are given by

$$\frac{\mathrm{d}\pi_i^A}{\mathrm{d}z} = \frac{1}{4}\left[-\frac{\mathrm{d}p^*}{\mathrm{d}z}\tau(d^*) + \tau'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}z}\,(1 - p^*)\right]. \tag{24}$$

First, stronger competition for advertiser (lower $z = t_a$) makes advertisers overall better off, i.e. $\mathrm{d}\pi_i^A/\mathrm{d}t_a < 0$. This is because, firstly, prices fall, such that the first term on the right hand side increases. Secondly, less personal data from users can be collected, which makes targeting less effective, therefore the second term is negative. Thirdly, also transportation costs decrease, which increases advertiser profits. Overall, the price and transportation cost reduction effects outweigh[5] decreased targeting effectiveness.

Second, stronger competition for user (increase $z = t_u$) hurts advertisers, hence $\mathrm{d}\pi_i^A/\mathrm{d}t_u > 0$. The platforms' bottleneck position allows them to increase prices (negative first term) and, further, less user data can be collected, such that targeting becomes less effective (negative second term). This effect is in line with the classic platform effect, that if one side becomes more price-elastic (here more competitve), then the other side has to pay more.

[4] See derivations in Appendix A.2.

[5] See derivations in Appendix A.2.

4.2.3 Effects on user utility

The effects on users' utility (net of transportation costs) $u_i = \underline{u} - \kappa(d^*) - \nu(d^*)a^* = \underline{u} - \kappa(d^*) - (1/2)\,\nu(d^*)$ are given by

$$\frac{\mathrm{d}u_i}{\mathrm{d}z} = -\frac{\mathrm{d}d^*}{\mathrm{d}z}\left[\kappa'(d^*) + \frac{\nu'(d^*)}{2}\right]. \tag{25}$$

Note that by Lemma 2 the term in brackets on the right-hand side is positive and that for $z \in \{t_a, t_u\}$ the first term is equal to zero.

First, more intense competition for advertisers (lower $z = t_a$) increases users' utility, i.e. $\mathrm{d}u_i/\mathrm{d}t_a < 0$. Intuitively, higher advertiser competition reduces the amount of data collected in equilibrium, which overall leaves users better off, as privacy concerns are reduced, although ads are less targeted and hence nuisance higher.

Second, stronger user competition (lower $z = t_u$) quite naturally increases users' utility, i.e. $\mathrm{d}u_i/\mathrm{d}t_u < 0$. Again, less data is collected, which on the one hand reduces privacy concerns and on the other hand increases nuisance costs. Overall, the positive effects prevail and are further strengthened by reduced transportation costs for users.

5 Welfare Analysis and Policy Implications

In this section we will derive welfare optimal and consumer optimal levels of data provision as benchmarks and then draw comparisons to the competitive outcome d^* established in section 3.

5.1 Welfare Optimum

Let us start with deriving the welfare efficient benchmark to draw first conclusions. It is easy to verify that the welfare optimum

$$arg\ max_{d_i,d_j,p_i,p_j}\ W = \int_0^{x_i} u_i dx + \int_{x_i}^1 u_i dx + \int_0^{y_i} \pi_i dy + \int_{y_i}^1 \pi_i dy + \Pi_i + \Pi_j \tag{26}$$

is given by the symmetric solution $d_i^o = d_j^o = d^o$ with d^o characterized by

$$\kappa'(d^o) = \frac{\tau'(d^o)}{2} - \frac{\nu'(d^o)}{2} \tag{27}$$

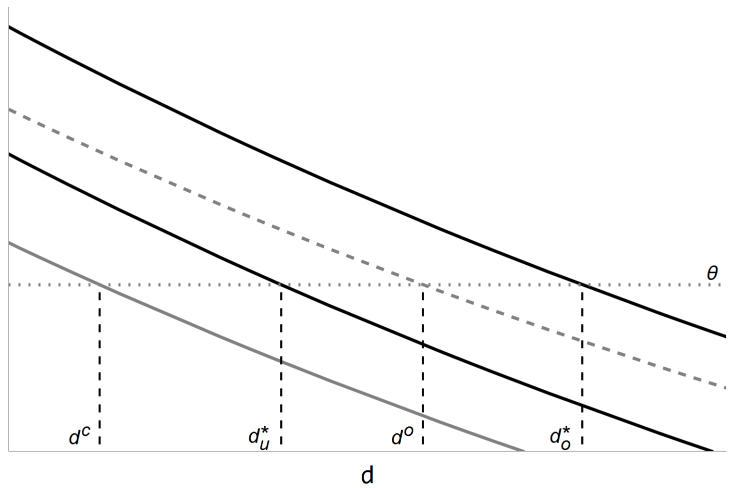
while prices $p_i^o = p_j^o = p^o$ can be freely chosen to split the rent between advertisers and platforms.

As we can see the welfare optimal level of data d^o is chosen in a way that the marginal cost of data provision $\kappa'(d^o)$ equals the sum of marginal benefits across both market sides, i.e. the marginal benefit of enhanced targeting $\tau'(d^o)/2$ and the marginal benefit of reduced nuisance $-\nu'(d^o)/2$. Furthermore, the optimal level of data provision is independent of transportion cost parameters t_a and t_u. Since prices are just transfers from advertisers to platforms they do not affect welfare and can be freely chosen to share the rent amongst both parties.

If we compare the RHS of the competitive level d^* in (9) and the efficient level d^o in (27) we can see that the comparison will crucially depend on the distortion induced by $\gamma(d^*) := \frac{\nu(d^*)+t_u}{\tau(d^*)-t_a}$ which gives more or less weight to the marginal benefit on the advertiser market side $\tau'(d^*)/2$. Note, that by assumption 1 the denominator of $\gamma(d^*)$ is positive, so that we have $\gamma(d^*) > 0$ in equilibrium. Whether the competitive level d^* of collected data is larger or smaller than the efficient level d^o will crucially depend on whether $\gamma(d^*)$ is smaller or larger than one. Note, that $\gamma(d^*)$ increases in t_a and t_u, putting more weight on $\tau'(d^*)/2$. Graphically this is represented in a right shift of the function pinning down the optimal level d^* as an intersection with $\kappa'(d)$. Figure 1 shows a graphical representation of the FOCs determining the respective optimal levels of d.

In fact we can see that depending on parameter values there can be under-provision ($d_u^* < d^o$) as well as over-provision ($d_o^* > d^o$) of personal data in the competitive equilibrium compared to the efficient benchmark. In particular we can infer from equations (9) and (27) that the competitive outcome leads to under-provision of personal data if $\gamma(d^*) < 1$ and to over-provision if $\gamma(d^*) > 1$. Note, for $\gamma(d^*) = 1$ expression (9) simplifies to (27), the efficient level of data provision. Using our definition of $\gamma(d^*)$ we can then see that

Figure 1: Provision of personal data



$d^* < d^o$ if

$$\gamma(d^*) < 1 \quad \Longleftrightarrow \quad \tau(d^*) - \nu(d^*) > t_a + t_u \tag{28}$$

and $d^* > d^o$ if

$$\gamma(d^*) > 1 \quad \Longleftrightarrow \quad \tau(d^*) - \nu(d^*) < t_a + t_u \tag{29}$$

These results are summarized in the two following propositions.

Proposition 4 *The competitive outcome leads to over-provision (under-provision) of personal data if competition on both market sides is weak (strong).*

We want to interpret this finding holding privacy concerns $\kappa(d)$ and the functions $\nu(d), \tau(d)$ fixed and ask the question which competitive environment leads to which scenario. Proposition 4 tells us if competition on both sides is strong, i.e. $t_a + t_u$ is small, platforms tend to collect and process an inefficiently small amount of data. Since consumers are likely to switch to more favorable data provision offers, the platforms' ability to gather data is limited $\frac{\mathrm{d}d^*}{\mathrm{d}t_c} > 0$. This increase in competition for consumers increases on the one hand the value of the collected data $\frac{\mathrm{d}p^*}{\mathrm{d}t_u} < 0$. However, if competition for advertisers is strong as well this effect might be offset by competitive pressure $\frac{\mathrm{d}p^*}{\mathrm{d}t_a} > 0$. A similar argument can be made if in turn competition on both sides is weak, i.e. $t_a + t_u$ is high. Consumers are likely to accept higher degrees of data collection due to difficulties switching to a competing platform. The resulting data inflation depresses prices on the advertiser market. However, if market power is sufficiently high the adverse effect can be offset and platforms have a monetary incentive to collect large amounts of data.

Proposition 5 *The competitive outcome leads to over-provision (under-provision) of personal data if net cross-group externalities are small (large).*

For this finding we hold the competitive environment on both sides fixed and analyze the effects of relatively strong or weak opposing cross-group externalities. On the one hand, an additional user imposes a positive externality on advertisers, which is equal to the targeting effect $\tau(d^*)$. On the other hand, an additional advertiser imposes a negative externality on users, which is equal to the nuisance costs $-\nu(d^*)$. If these effects together are relatively large (small), the LHS of equations (28) and (29) become relatively large (small) and hence we are in a situation of under-provision (over-provision).

5.2 Consumer Optimum

Let us now define a consumer optimal level of data provision. Obviously if consumers are free to decide on the amount of data provided, the consumer optimal level d^c is derived from consumers utility (1) and is given by

$$\kappa'(d^c) = -\frac{1}{2}\,\nu'(d_i^c) \tag{30}$$

Comparing the consumer optimal level d^c to the welfare optimal level d^o we immediately see that consumers provide an inefficiently low level of data. This result is summarized in the following proposition.

Proposition 6 *The consumer optimal level of data provision is inefficiently low.*

The reason for this result is straight forward. As consumer do not internalize the effect the data has on the advertiser market, they will provide data up to the point where the marginal decrease in nuisance equals marginal cost of data provision. Since from a welfare perspective the value creation aspect on the advertiser market is omitted, the resulting level of data provision is inefficiently low.

Furthermore, since $\gamma(d^*) > 0$ we also have $d^* > d^c$. In particular even in the case of under-provision, the competitive outcome is better from a welfare perspective than the consumer optimal choice. This situation is also depicted in graph 1. Depending on the competitive structure of the market, platform competition might lead to inefficiently low or high level of data collection. However, in a situation where a competitive setting would lead to under-provision, the resulting level of data provision is closer to the welfare optimum than the consumer optimal choice.

Proposition 7 *If the market outcome leads to under-provision of personal data it still outperforms the consumer optimal choice in terms of welfare.*

Unlike consumers platforms act as intermediaries and are able to internalize parts of the value creation on both sides of the market. Depending on the competitive structure of the market this might lead to putting too much or too little weight on the advertiser side of the market, resulting in a situation where the data collection is inefficiently low or high. However, if it turns out that there is an under-provision of personal data, the competitive outcome is closer to the welfare optimal level, since the additional positive effect on the advertiser market side is internalized. [6]

5.3 Policy Implications

In this section we would like to briefly discuss what conclusions can be drawn from our previous analysis when it comes to policy implications and regulation.

First-best regulation

An omnipotent regulator could obviously achieve the first-best by forcing $d_i = d_j = d^o$ and increasing competition on both sides of the market such that $t_u \to 0$ and $t_a \to 0$. In this case the efficient amount of data is provided while the total transportation costs approach zero.

[6] Note, even a scenario of over-provision is better in terms of welfare at least up to threshold. (Soll man das weiter ausarbeiten?)

Second-best regulation

Regulatory practices focus mainly on data / privacy regulation and measures to assure competitiveness on the consumer side. Examples...

Holding the competitive structure of the market fixed, the regulator could still improve upon the market outcome by enforcing $d_i = d_j = d^o$. However, a direct regulation of the amount of data in our model requires knowledge of the functions $\tau(d)$ and $\nu(d)$ as well as users' privacy concerns $\kappa(d)$.

An approach which is less demanding when it comes to information requirements is the regulation of t_u and t_a, i.e. affecting the competitiveness of the two market sides. Our results suggest that if competition is very weak on both sides ($t_u + t_a$ high) the amount of data collected is likely to be inefficiently high. Similarly, if competition is too strong ($t_u + t_a$ low) too little data is provided from a welfare point of view. While regulators still have to know whether there is over-provision or under-provision in the market in the first place, our results can still provide some guidance.

First of all, our comparative statics results suggest that increasing competition works in the same direction for both sides of the market. The equilibrium amount of data provision is a monotone function of the transportation cost parameters t_a and t_u and by altering either one of the parameters it is possible to push the competitive equilibrium amount of data d^* towards the welfare optimum d^o. However, even though both parameters work in the same direction they are not equally effective. Keeping in mind the implicit definition of d^* in equation (9) and going back to Figure 1 we can see that shifts in the transportation cost parameters correspond to shifts of the graph in a one-to-one relationship. Since the only source of distortion in our model is an inefficient amount of data, we can ask ourselves which parameter leads to a stronger reaction of d. We can therefore look at the reaction of the RHS of equation (9) $RHS(d^*) := \frac{1}{2}\left[\frac{\nu(d^*)+t_u}{\tau(d^*)-t_a}\tau'(d^*) - \nu'(d^*)\right]$ such that

$$\begin{aligned}\frac{\mathrm{d}RHS(d^*)}{\mathrm{d}t_a} &= \frac{1}{2}\frac{\nu(d^*)+tc}{(\tau(d^*)-ta)^2}\tau'(d^*)\\ \frac{\mathrm{d}RHS(d^*)}{\mathrm{d}t_u} &= \frac{1}{2}\frac{1}{\tau(d^*)-ta}\tau'(d^*)\end{aligned} \tag{31}$$

and can then see that the comparison $\frac{\mathrm{d}RHS(d^*)}{\mathrm{d}t_a} \lesseqgtr \frac{\mathrm{d}RHS(d^*)}{\mathrm{d}t_u}$ boils down to the same conditions as in (28) and (29) such that $\frac{\mathrm{d}RHS(d)}{\mathrm{d}t_a} >$ if $t_a + t_u > \tau(d^*) - \nu(d^*)$ and vice versa. This gives rise to the following proposition.

Proposition 8 *If the market exhibits over-provision (under-provision) regulation of the advertiser (consumer) side of the market is more effective than regulation of the consumer (advertiser) side.*

This result is particularly important in a scenario where the market exhibits under-provision and a regulator would have to reduce competition as this implies increasing transportation costs in the economy. Increasing transportation costs would then lead to more data collection in the subsequent market outcome. Whether we can increase total welfare by increasing transportation cost depends crucially on whether the benefit of higher and thus more efficient data provision (non linear) exceeds the increased costs of transportation (linear). This trade-off could call for a "second-best" regulation, where competition intensity is regulated, i.e. decreased, in such a way, that the amount of data provided in the subsequent market outcome balances the above mentioned benefits and costs at the margin. The resulting "second-best" level of provided data could be below the efficient "first-best" level.

Lastly, a regulator could also consider switching to a consumer standard and let consumer freely choose how much data they would like to provide. Our results show that the consumer optimal amount of data is always inefficiently low as consumers do not internalize the benefit on the advertiser side. In particular our results suggest that we can only improve in terms on welfare by switching to a consumer standard when there is extreme over-provision of data in the economy, i.e. platforms have significant market power on both sides of the market. If the market exhibits under-provision, switching to the consumer standard always reduces welfare.

6 Extension

In this chapter we sketch and briefly discuss different extensions of the baseline model presented in Section 2.

6.1 Consumer Prices

In this section we consider an alternative setup where platforms also take into account the possibility to charge prices on the consumer side of the market. Let p_i^c denote the price a consumer has to pay to join platform i. Consumer utility is then given by

$$u_i(x) = \underline{v}_i + \underline{d} - \kappa(d_i) - \nu(d_i)a_i - p_i^c - t_c|l_i - x| \tag{32}$$

while advertisers still face the same decision as in section 2. Market shares are obtained as before by pinning down indifferent consumers and advertisers and solve the resulting system of equations. Note, market shares on both sides of the market now additionally depend on p_i^c and p_j^c. The resulting profit maximization problem of platform i is then given by

$$\max_{p_i, d_i, p_i^c} = a_i \tau(d_i) p_i x_i + p_i^c x_i \;\forall i \in \{1, 2\} \tag{33}$$

and takes into account profits made from selling access to consumers. Following the same procedure as in our baseline model we obtain symmetric equilibrium values $p_i = p_j = \tilde{p}$, $p_i^c = p_j^c = \tilde{p}^c$ and $d_i = d_j = \tilde{d}$ where advertiser prices are given by

$$\tilde{p} = \frac{2\left(t_a + \nu(\tilde{d})\right)}{\tau(\tilde{d})} \tag{34}$$

consumer prices by

$$\tilde{p}^c = t_a + t_c + \nu(\tilde{d}) - \tau(\tilde{d}) \tag{35}$$

while the equilibrium amount of data is given by

$$\kappa'(\tilde{d}) = \frac{1}{2}\left[\tau'(\tilde{d}) - \nu'(\tilde{d})\right] \tag{36}$$

We immediately see from equations (27) and (36) that $\tilde{d} = d^o$. If platforms charge prices to consumers, the resulting amount of data collected is efficient. Since platforms can now extract rents from both sides of the market, they try to maximize the aggregate value, whereas in our baseline model platforms focused on the advertiser side of the market. This result, however, has very strong implications. If platforms demand prices from consumers and from advertisers, the resulting market outcome is welfare optimal. However, taking a closer look at equilibrium consumer prices in (35) we immediately see that negative, positive or zero prices are possible depending on parameter values and functional forms. There are two main conclusions we would like to draw from this result. Firstly, observing a consumer price $\tilde{p}^c = 0$ empirically is consistent with the equilibrium result above as well as with our baseline model presented in section 2. By observing zero prices we can not infer whether a price of zero is an optimal choice, making the model above the 'correct' model, or whether there are constraints which prevent platforms from setting consumer prices at all, making our baseline model more suitable. Secondly, since consumer prices depend on parameters of competition intensity and functional forms, observing zero prices in different markets, jurisdictions and industry sectors makes it unlikely that $\tilde{p}^c = 0$ is a profit maximizing choice in all cases. This strongly supports the argument made by Waehrer (2015), that consumer prices are not a variable of interest in the platforms maximization problems.

6.2 Collusion

Let us consider a collusive game where platforms agree on prices $p_i = p_j = p$ and data requirements $d_i = d_j = d$ such that joint profits are maximized. Since advertisers face transportation costs, the profit maximizing collusive price is such that the participation constraint of the indifferent advertiser is binding $p : \pi_i\left(\frac{1}{2}\right) \geq 0$ which yields

$$p = 1 - \frac{t_a}{\tau(d)} \tag{37}$$

Plugging the collusive price p in the platforms' profit functions (3) we obtain

$$\Pi_i = \frac{1}{4}(\tau(d) - t_a) \tag{38}$$

and immediately see that profits are increasing in d up to the point where the participation constraint of the indifferent consumer is binding $d : u_i(\frac{1}{2}) \geq 0$. Since we assumed $\underline{u}$ to be high enough to have interior solutions in the previous sections, we can infer that the collusive amount of data will be excessively high. This highlights the importance of competition among platforms.

7 Conclusion

We analyze the role of competition intensity in a two-sided market framework where platforms collect data from users and monetize through ad-sales. Our model predicts that the equilibrium amount of collected data will be distorted compared to the welfare efficient benchmark. Depending on the net effect of cross-group externalities and the competition intensity on both sides of the market, the distortion can lead to underprovision or overprovision of personal data. Since the level of collected data increases the more market power platforms have on either side of the market, side specific regulations are substitutes. However, our results suggest that regulation of the advertiser side is more effective if the competitive outcome exhibits overprovision of personal data. Also, a consumer standard would always lead to underprovision of data as consumers do not internalize improvements in the targeting capabilities. Lastly, we showed that two-sided pricing induces platforms to choose the efficient level of data provision and that collusion would always lead to overprovision.

While we think our model provides some useful insights we would also like to discuss some limitations. It would be interesting to explore the possibility of endogeneous multi-homing on the advertiser side as it would reduce the market power on the advertisers but propagate the bottleneck property vis--vis consumers. Secondly, one could alter the setting on the consumer side and consider heterogeneous consumers, while platforms engage in second degree discrimination by offering a menue of data choices. Those could be interesting aspects for future research.

Appendix

A Omitted Proofs

A.1 Second Order Conditions

In the following we derive sufficient conditions such that the equilibrium values $p^*,\ d^*$ derived from the maximization problem presented in section 2 characterize a global maximum. Let us consider the Hessian evaluated at equilibrium values. Starting with

$$\left.\frac{\partial^2 \Pi_i}{\partial p_i^2}\right|_{d^*,p^*} = -\frac{t_u^2\ \tau(d^*)^2(\nu(d^*)+t_u)}{4(t_u-\nu(d^*))^2(\nu(d^*)\tau(d^*)+t_a t_u)}$$

we immediately see that $\left.\frac{\partial^2 \Pi_i}{\partial p_i^2}\right|_{d^*,p^*} < 0$, a necessary condition for the Hessian to be negative definite. In the next steps we argue that we can always find functions $\tau(\cdot), \nu(\cdot)$ such that $det(H)|_{d^*,p^*} > 0$.

First, it is helpful to look at the numerator and the denominator of the Hessian seperately

$$det(H)|_{d^*,p^*} = \frac{H_{num}}{H_{den}}$$

where the numerator H_{num} and the denominator H_{den} are given by

$$\begin{aligned}
H_{num} = {} & \tau(d^*)^2 \big[-4t_u^2(t_a-\tau(d^*))(\nu(d^*)\tau(d^*)+t_a t_u)\left(\nu''(d^*)(t_a-\tau(d^*))+\tau''(d^*)(\nu(d^*)+t_u)\right) \\
& -t_u^2\nu'(d^*)^2(t_a-\tau(d^*))^3-\tau'(d^*)^2(\nu(d^*)+t_u)^2\left(\nu(d^*)(\nu(d^*)(t_a-\tau(d^*))+4t_c\tau(d^*))+4t_a t_u^2\right) \\
& +2t_u\nu(d^*)\nu'(d^*)\tau'(d^*)(t_a-\tau(d^*))^2(\nu(d^*)+t_u)\big] \\
H_{den} = {} & 64(t_a-\tau(d^*))(t_u-\nu(d^*))^2(\nu(d^*)\tau(d^*)+t_a t_u)^2
\end{aligned}$$

Note, that $H_{den} < 0$ as we have $(t_a - \tau(d^*)) < 0$ from Assumption 1. Rewriting H_{num} as

$$\begin{aligned}
H_{num} &= \tau(d^*)^2\left[H1_{num}\left(H2_{num}\nu''(d^*)+H3_{num}\tau''(d^*)\right)+H4_{num}+H5_{num}+H6_{num}\right] \\
H1_{num} &= -4t_u^2(t_a-\tau(d^*))(\nu(d^*)\tau(d^*)+t_a t_u) > 0 \\
H2_{num} &= (t_a-\tau(d^*)) < 0 \\
H3_{num} &= (\nu(d^*)+t_u) > 0 \\
H4_{num} &= -t_u^2\nu'(d^*)^2(t_a-\tau(d^*))^3 > 0 \\
H5_{num} &= -\tau'(d^*)^2(\nu(d^*)+t_u)^2\left(\nu(d^*)(\nu(d^*)(t_a-\tau(d^*))+4t_u\tau(d^*))+4t_a t_u^2\right) \lesseqgtr 0 \\
H6_{num} &= 2t_u\nu(d^*)\nu'(d^*)\tau'(d^*)(t_a-\tau(d^*))^2(\nu(d^*)+t_u) < 0
\end{aligned}$$

we can see that requiring $H_{num} < 0$ is equivalent to the condition

$$-\frac{1}{H1_{num}}\left(H4_{num} + H5_{num} + H6_{num}\right) > H2_{num}\nu''(d^*) + H3_{num}\tau''(d^*)$$

where $LHS \lesseqgtr 0$ while $RHS < 0$ due to our functional requirements on $\tau(\cdot)$ and $\nu(\cdot)$. The important thing to realize is that, firstly, the condition for negative definitness reduces to a condition which is linear in $\nu''(d^*)$ and $\tau''(d^*)$, the curvature information of the targeting and the nuisance functions, and secondly, is given by an upper bound. If the sign of the upper bound is positive then this condition is always fulfilled as we have $RHS < 0$. Only if the sign of the upper bound is negative, the condition is binding. But then we can assume that $\tau(\cdot)$ is sufficiently concave and/or $\nu(\cdot)$ is sufficiently convex such that this condition holds since for our results we only require $\tau''(\cdot) < 0$ and $\nu''(\cdot) > 0$ which is in line with this condition.

A.2 Proof of Compatative Statics

Proposition 3

Proof. To see that $\mathrm{d}\Pi_i^P/\mathrm{d}t_c < 0$, note that

$$\begin{aligned}
\frac{\mathrm{d}\Pi_i^P}{\mathrm{d}t_u} &= \frac{-\left[\tau(d^*) - t_a\right]\left[\nu(d^*) - t_u\nu'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}t_u}\right] + \frac{\mathrm{d}d^*}{\mathrm{d}t_u}\nu(d^*)\tau'(d^*)\left[t_u + \nu(d^*)\right]}{\left[t_c + \nu(d^*)\right]^2} \\
&= \frac{\left[\tau(d^*) - t_a\right]^2\left[\left(t_u + \nu(d^*)\right)\nu'(d^*)\tau'(d^*) - \nu(d^*)\left(\tau(d^*) - t_a\right)\nu''(d^*) + t_c\left(t_u + \nu(d^*)\right)\tau''(d^*)\right]}{\left[t_u + \nu(d^*)\right]^2 Z(d^*)} \\
&< 0, \qquad \text{(A.1)}
\end{aligned}$$

where $\mathrm{d}d^*/\mathrm{d}t_u$ is from equation (17), while $Z(d^*)$ is defined in equation (18).

To see that $\mathrm{d}\pi_i^A/\mathrm{d}t_a < 0$, note that

$$\begin{aligned}
\frac{\mathrm{d}\pi^A}{\mathrm{d}t_a} &= \frac{1}{4\left[t_u+\nu(d^*)\right]^2}\Bigg\{-6t_u\nu(d^*)-\nu(d^*)^2\left[1+2\tau'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}t_a}\right] \\
&\qquad + t_c\left[-4\nu'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}t_a}\left(\tau(d^*)-t_a\right)+t_c\left(-5+2\tau'(d^*)\frac{\mathrm{d}d^*}{\mathrm{d}t_a}\right)\right]\Bigg\} \\
&= -\frac{1}{4\left[t_u+\nu(d^*)\right]Z(d^*)}\Bigg\{-\nu'(d^*)\left(t_u+\nu(d^*)\right)\left(\tau(d^*)-t_a\right)\tau'(d^*)+3\left(t_u+\nu(d^*)\right)^2\tau'(d^*)^2 \\
&\qquad - \left(5t_u+\nu(d^*)\right)\left(\tau(d^*)-t_a\right)\left[-\nu''(d^*)\left(\tau(d^*)-t_a\right)+\left(t_u+\nu(d^*)\right)\tau''(d^*)\right]\Bigg\} \\
&< 0, \qquad \text{(A.2)}
\end{aligned}$$

where $\mathrm{d}d^*/\mathrm{d}t_a$ is from equation (20), while $Z(d^*)$ is defined in equation (18). ■

References

Ambrus, A., Calvano, E., and Reisinger, M. (2014). Either or both competition: A'two-sided'theory of advertising with overlapping viewerships. *Economic Research Initiatives at Duke (ERID) Working Paper*, (170).

Anderson, S. P. and Coate, S. (2005). Market provision of broadcasting: A welfare analysis. *The review of Economic studies*, 72(4):947–972.

Anderson, S. P., Foros, Ø., Kind, H. J., and Peitz, M. (2012). Media market concentration, advertising levels, and ad prices. *International Journal of Industrial Organization*, 30(3):321–325.

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691.

Armstrong, M. and Wright, J. (2007). Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory*, 32(2):353–380.

Athey, S., Calvano, E., and Gans, J. (2014). The impact of consumer multi-homing on advertising markets and media competition.

Bourreau, M., Caillaud, B., and De Nijs, R. (2016). Taxation of a digital monopoly platform.

Peitz, M. and Reisinger, M. (2016). The economics of internet media. In *Handbook of Media Economics, vol. 1A*. Elsevier.

Reisinger, M. (2012). Platform competition for advertisers and users in media markets. *International Journal of Industrial Organization*, 30(2):243–252.

Waehrer, K. (2015). Online services and the analysis of competitive merger effects in privacy protections and other quality dimensions. *Available at SSRN*.