The value of consumer data in online advertising∗

Marc Bourreau†  Bernard Caillaud‡  Romain De Nijs§

June 27, 2017

Abstract

In this paper we propose a model where personal data collected on consumers online have multidimensional characteristics and are used by platforms to offer ad slots with better targeting possibilities to a market of differentiated advertisers through real-time auctions. A platform controls the amount of information about consumers that it discloses to advertisers, thereby affecting the dispersion of advertisers’ valuations for the slot. We first show by way of simulations that the amount of consumer-specific information that is optimally revealed to advertisers increases with the degree of competition on the advertising market and decreases with the cost of information disclosure for a monopolistic platform, competing platforms or a welfare-maximizing platform, provided the advertising market is not highly concentrated. Second, we exhibit different properties between the welfare-maximizing situation and the imperfectly competitive market situations with respect to how the incremental value of information varies: there are decreasing social returns to consumers’ data while private returns may be increasing or decreasing locally.

∗This research has been financed by France Strategie, within the general framework of a Research Project on the “Evolution of the Value created by the Digital Economy and its Fiscal Consequences”, signed between France Strategie and a consortium consisting of PSE and Telecom ParisTech.
†Telecom ParisTech; Email: marc.bourreau@telecom-paristech.fr.
‡Paris School of Economics - Ecole des Ponts ParisTech; Email: caillaud@pse.ens.fr.
§Department of Economics, Ecole Polytechnique; E-mail: romain.de-nijs@polytechnique.edu.
1 Introduction

The increasing digitalization of the economy goes with an increasing concern by public authorities about the use of consumers’ personal data that are collected massively by many major online players. Whenever they surf on the internet, visit a website, login on an account, answer various requests about personal information, choose a product or some entertainment service and make payments, consumers leave footprints in the system that may be collected in various ways: tracking cookies, device fingerprinting, history sniffing methods, ...

Consumers’ personal data are a source of value in at least three different ways. First, they enable online firms to adapt their proposals to the consumer. Online retailers or service providers can propose more personalized goods or services that better match with the consumer’s characteristics and implicit tastes, and a better fit generates a larger surplus to grab. Along with more personalized services, personal data may be used to identify the consumer’s willingness to pay for a given (possibly personalized) product, opening the door to personalized pricing, i.e. third-degree or even first-degree price discrimination, by online sellers. Second, consumers’ personal data can be used to design personalized services that are external to the visited website: in this vein, the major use of data is through the more efficient targeting of advertising messages, i.e. through advertising messages that contain personalized elements increasing the probability of reaching the consumer, of raising his/her attention and of inducing him/her to look for the advertised product and ultimately make a purchase. Third and last, consumers’ data may be sold to intermediaries – the so-called data brokers – that accumulate data from various sources, consolidate and process them so that they can be sold back to other online players, websites or ad networks, in a useful format.

The mere fact of collecting, using and selling personal information about consumers is problematic as it may constitute a major violation of consumers’ privacy. Offering more adequate products or more relevant ad messages, absent any impact on prices, is presumably welfare improving and bring better value to the consumers, but the increasing capacity of online players to extract value from consumers through more sophisticated price discrimination may reduce or even reverse the direct effect of personalization of services and of advertising. A precise analysis of the impact of the availability of consumers’ personal data in online industries therefore calls for a more thorough study of how data can be turned into value, how a market for data works and how data is transferred or

---

1See e.g. the reports for the U.K. of OFT (2010) and (2013), for the US FTC (2014) or the French Autorité de la Concurrence and Bundeskartellamt’s report (2016).
2On the business models of data-driven platforms, see, for example, Lambrecht et al. (2014).
3For a recent survey on the economics of privacy, see Acquisti et al. (2016).
exchanged among players.

This paper addresses the second motive for the use of personal data, i.e. the fact that data can help ad networks or operators to improve ad targeting and therefore the efficiency of advertising campaigns. Typically, advertising slots that are consumer specific are sold in real time through auctions and bidders are able to use some of the personal data about the consumer that is attached to the slot so as to better assess the value of the slot for them and how much they are willing to bid for it. The platform that runs the auctions and that has collected information about the specific consumer can control the information revealed to bidders so as to extract maximal value from the auction. Of course, competition among platforms to access the consumer’s attention may mitigate this objective and lead the platform to modify its disclosure policy so as to fight for market shares on the advertising market.

In this setting, we investigate what is the value of information for data collecting platforms, depending on the market structure at the platform level and on the structure of the advertising market. More precisely, we analyze how the value of information and the incremental value of information, i.e. the value of an additional piece of information, vary as functions of the degree of competition in the advertising market, in three situations: when the platform is monopolistic, when there is competition between two platforms in a competitive bottleneck situation (see Armstrong, 2006), and finally when a central planner aims at maximizing social efficiency. The study of the incremental value of information relates to whether there are increasing or decreasing returns in the use of consumers’ personal information. Ultimately, these issues lead to the question of the incentives of monopolistic, competing or regulated platforms to collect information about consumers.

One important aspect of our setting is that we propose a model that simply formalizes what disclosure of information means (more or less) and that accounts for the fact that two differentiated advertisers may evaluate differently some revealed set of personal data. For this, we have to depart from models that rely on all or nothing disclosure that lead to prior information or perfect information ex post and we have to enrich the view of information as not being reducible to a one-dimensional variable, an increase in which would necessarily translate into a similar increase in the advertisers’ value. So, we develop a simple model in which information is multidimensional, additional pieces of information may increase or decrease advertisers’ value and information revelation is related to the dispersion of bidders’ interim evaluations in the advertising market. Although the model has many attractive features, its complete analytic analysis is cumbersome and in this paper we content ourselves with simulations to derive the main properties of the model.

Within this setting, we first show, that the amount of consumer-specific information
that is optimally revealed to advertisers in all situations increases with the degree of competition on the advertising market, i.e. with the number of advertisers, provided this market is not highly concentrated, i.e. provided there are at least 4 advertisers. The intuition is the following. Disclosing information increases the dispersion of valuations among bidders. When there are enough draws in this distribution, both the highest valuation and the second highest valuation increase with dispersion, as well as the difference between the two, as the probability of a higher value fit increases. Also, when there is a cost of revealing a piece of information, we find that the amount of consumer-specific information that is optimally revealed to advertisers in all situations decrease with the cost of information revelation. Secondly, we compare the amount of revealed information across situations: competing platforms fight to attract single-homing consumers and therefore reveal more information than a monopoly platform; a monopoly platform itself reveals too little information compared to the social optimum since it is interested in the second highest valuation and not the highest valuation; however, we find no systematic relationship between information disclosure by competing platforms and the socially efficient disclosure policy.

As for the incremental value of information about consumers, we find a contrast between the central planner situation and market situations. We find that there exists decreasing social returns to consumers’ data while private returns may be either decreasing or increasing. Similarly, the incremental social value of consumer information increases with the degree of competition in the advertising market (in the number of advertisers in the market), while the incremental private value can either increase or decrease with it.

2 Literature review

Our paper is related to the literature on information revelation in auctions. The basic setting corresponds to an auctioneer who wishes to sell an object to bidders with heterogeneous preferences. The auctioneer does not know the bidders’ preferences but has private information about the characteristics of the object that she can reveal to bidders, anticipating how this information affect bidders’ valuations. Ganuza (2004) and Ganuza and Penalva (2010) are the seminal contributions on this subject. Among other things, they show that a monopoly auctioneer reveals less information than what would be socially desirable. However, they do not consider the case of competing auctioneers nor discuss the incremental value of an additional piece of information as we do in this paper. Another related article is De Corniere and De Nijs (2016) which endogeneizes the value of additional information via the pricing strategies of bidders on a final downstream market. In the present article, we do not take into account this aspect. Moreover, contrary to
De Corniere and De Nijs (2016), we develop a model in which information is formalized as a discrete (and not a continuous) quantity which enables us to give more sense to the value of an additional piece of information. Last, our article is related to Board (2010) and Troncoso-Valverde (2017). Board (2010) shows that an auctioneer is always better off revealing information to all bidders in a second-price auction as long as there are at least three bidders. Troncoso-Valverde (2017) shows that when two auctioneers compete for the two bidders that are active on the market, they reveal information about their object, a result that is opposed to what is found for a monopolist auctioneer.

3 Model

We build a model where advertising-supported digital platforms sell their ad slots to advertisers via a second-price auction. An ad slot is specific to a consumer and a platform has to decide how much of the information about the consumer it should reveal to bidders. We present below the different players and their strategies. The market structure is represented in Figure 1.

Consumers

There is one (representative) consumer, defined by a vector of \( C \) characteristics. The consumer’s characteristics are independently and identically distributed: each characteristic takes value \(-1\) with probability \(1/2\) and \(1\) with probability \(1/2\). A characteristic can correspond to a demographic information (young vs. old), an interest (likes fishing vs. does not like fishing), etc.

Digital platforms

We consider two possible market structures: either the market is dominated by a monopoly platform or there are two competing platforms. Digital platforms act as intermediaries between the representative consumer and advertisers. They derive revenues from advertising. If the platform is a monopoly, we assume full consumer participation. In the duopoly case, we assume that the representative consumer multi-homes and that advertisers single-home.\(^4\)

A platform sells an ad slot displayed to the representative consumer to \( n \) potential advertisers via a second price auction, where \( n \) is assumed exogenous. Each platform has perfect information about the consumer (i.e., it knows all \( C \) characteristics) and has

\(^4\)For example, one could assume that the platform offers a basic service for free that attracts consumers, such as news, weather information, email, etc.
to decide how much of this information to reveal to advertisers, that is, the number of characteristics \( c \) it wishes to disclose, with \( 0 \leq c \leq C \). Note that once we know the number of characteristics that the platform optimally reveals to advertisers, we also have information on how much data about the consumer the platform has an incentive to collect.

We assume furthermore that the platform incurs a cost \( \delta \) for each characteristic revealed to the advertisers. Hence, if the platform reveals \( c \) characteristics, its cost of information revelation is \( \delta c \). For example, we can consider that the consumers dislike that their personal information be disclosed to advertisers and that the platform has to compensate them to ensure their participation, either through a lower price or a higher quality of service, both of which are costly.

**Advertisers**

There are \( n \geq 2 \) advertisers. Similar to the representative consumer, each advertiser is defined by a vector of \( C \) independently and identically distributed characteristics, where each characteristic takes either value \(-1\) or \(1\), with equal probability. The vector of characteristics of an advertiser is private information to this advertiser and in particular it is unknown to the platform.

The value of the match between a given advertiser and the representative consumer is additive across characteristics and is simply defined as the scalar product of their vectors of characteristics.

It is worth discussing at this stage the main effects at play in our model. If the platform does not reveal any information about the consumer, all advertisers have the same expected valuation for the ad slot. In contrast, by revealing information about the consumer, the platform creates dispersion among advertisers in their valuation for the slot. This dispersion increases the willingness to pay of some advertisers, which increases the platform’s expected revenues and the gain from trade between the consumer and the advertiser. However, increased dispersion also creates an informational rent for advertisers. Therefore, we can expect the platform to face a trade-off when it decides on the amount of information to disclose between increasing willingness to pay and reducing the informational rent of advertisers.

4 Second-price auction for ad slots

We start by calculating the relevant statistics of the second-price auction for the ad slot, for a given number \( c \leq C \) of consumer characteristics revealed by the platform to
Let $V(c)$ denote the random variable that corresponds to the value for an advertiser of displaying an ad to the consumer, given that $c$ characteristics have been disclosed. The value of a match for a given characteristic is 1, and the value of a non-match is $-1$. Note that the expected value for the characteristics that are not revealed is equal to 0 (since $1/2 \times (-1) + 1/2 \times 1 = 0$). The value of the ad slot for an advertiser, $V(c)$, therefore takes value in $\{-c, -c + 2, \ldots , c\}$. More precisely, if $j \leq c$ characteristics match between the advertiser and the consumer, we have $V(c) = x_j(c) \equiv (1) \times j + (-1) \times (c - j) = -c + 2j$ for $j = 0, 1, \ldots c$.

Let $p_j(c)$ denote the probability that the value of an ad for the consumer be equal to $V(c) = x_j(c)$ when $c$ characteristics have been disclosed. This probability is equal to the number of ways to choose $j$ characteristics that match among a total of $c$ characteristics, divided by the number of different vectors with $c$ characteristics. Therefore, we have:

$$p_j(c) = \left(\frac{1}{2}\right)^c \binom{c}{j}.$$

The random variable $V(c)$ therefore takes value $x_j(c) = -c + 2j$, for $j = 0, 1, \ldots c$ with
probability $p_j(c)$. The mean is null and the dispersion of this random variable increases with the number of characteristics $c$ that are revealed.

We model the auction for the ad slot as a second-price auction: the highest bidder wins the auction, and pays the second highest bid for the ad slot. We derive the density functions for the largest and second-largest bidder’s value for an ad to the consumer.

Define $P_j(c)$ with $P_0(c) = 0$ and $P_j(c) = p_1(c) + p_2(c) + \ldots + p_j(c)$ denote the discrete cumulative distribution function of $V(c)$. Omitting the dependency upon $c$, the discrete probability distribution of the second largest bidder is then given by:\(^5\)

\[
Prob\{V_{(n-1)} = x_j\} = \sum_{k=n-1}^{n} \binom{n}{k} \left[ P_j^k(1-P_j)^{n-k} - P_{j-1}^k(1-P_{j-1})^{n-k} \right],
\]

that is,

\[
Prob\{V_{(n-1)} = x_j\} = n \left[ P_j^{n-1}(1-P_j) - P_{j-1}^{n-1}(1-P_{j-1}) \right] + P_j^n - P_{j-1}^n . \tag{1}
\]

Similarly, omitting $c$, the discrete probability distribution for the largest bid is given by:

\[
Prob\{V_n = x_j\} = P_j^n - P_{j-1}^n . \tag{2}
\]

## 5 Information revelation by the platforms

In this section, we study the information revelation decision by the platforms. To begin with, we compute the surplus that the different parties derive from the sale of the ad slot. Then, we study how the market structure of the advertising market affects information disclosure, comparing the environment with a monopoly platform and the environment with competing platforms. We also analyze the value of an incremental piece of information disclosed to the advertisers, and its determinants. Finally, we discuss the impact of the cost of information on disclosure.

### 5.1 Gains from trade

The surplus that a monopoly platform expects to derive from the sale of its ad slot is equal to the expected price of the slot, which we denote by $v_m$, minus the information cost, $\delta c$. From the analysis of the second-price auction in the previous section, the expected price

---

of the ad slot is equal to the expected value of the second-highest bid, that is,

\[ v_m(c) = \sum_{j=0}^{c} x_j(c) \text{Prob}\{V_{n-1}(c) = x_j(c)\}, \]

where \( \text{Prob}\{V_{n-1}(c) = x_j(c)\} \) is defined in (1). The expected profit of the platform is then \( \pi(c) = v_m(c) - \delta c \).

Notice that the expected profit of the platform, \( \pi(c) \), depends on the number of consumer characteristics disclosed to the advertisers. In particular, if it reveals no consumer information to the advertisers (i.e., \( c = 0 \)), the platform makes zero profits (\( \pi(0) = v_m(0) = 0 \) as it can be easily checked).

The expected net surplus of the advertiser that wins the auction is equal to the difference between its value for the ad slot and the price it pays for it. Therefore, it is equal to the difference between the highest expected bid and the second-highest expected bid. Using (2), the expected value of the highest bid is equal to

\[ v_1(c) = \sum_{j=0}^{c} x_j(c) \text{Prob}\{V_{n}(c) = x_j(c)\}, \]

and the expected net surplus for the advertiser that wins the auction is then \( v_c(c) = v_1(c) - v_m(c) \). If the platform reveals no information about the consumer (i.e., \( c = 0 \)), advertisers all have the same expected value from the match, i.e., 0, and therefore, we have \( v_c(0) = 0 \).

Finally, for the social planner, the expected value from the sale of the ad slot is equal to the expected value of the highest bid, \( v_1(c) \), minus the information cost, \( \delta c \). That is, \( w(c) = v_1(c) - \delta c \).

In what follows, we interpret \( v_m \) as the private value of information and \( v_1 \) as the social value of information.

The monopoly platform chooses a number \( c \) of consumer characteristics to disclose to the advertisers in order to maximize its advertising profit, \( \pi(c) = v_m(c) - \delta c \). Under platform competition, the two platforms compete to attract single-homing advertisers. A platform \( s \) therefore chooses a number of consumer characteristics \( c_s \) that maximizes the winning advertiser’s expected surplus, that is, \( v_c \), subject to the constraint that its net profit should be positive, i.e., \( v_c \geq \delta c \). In equilibrium, the two platforms disclose the same amount of information, and we then assume that they each attract the single-homing winning advertiser with probability 1/2. The expected profit of a platform under platform competition is then \( (1/2)\pi(c_s) \).

Unfortunately, our combinatorial model is particularly complex, and we were not able
to derive analytical solutions on the amount of information disclosed to advertisers for our different scenarios. In the rest of the paper, we therefore revert to simulations to derive our results.

In our simulations, we make the following assumptions. First, we assume without loss of generality that the consumer has $C = 10$ characteristics. We also consider that the number of advertisers $n$ takes value in $\{2, 3, \ldots, 20\}$, and that the information cost parameter $\delta$ takes value in $\{0.1, 0.2, \ldots, 0.9\}$.

The two following results are useful for the rest of the analysis.

**Result 1** : The social value of information, $v_1$, is increasing in $c$ for all $n \geq 2$.

**Result 2** : The private value of information, $v_m$, decreases with the amount of information disclosed if $n = 2$, is constant with respect to $c$ and equal to 0 if $n = 3$, and is increasing in $c$ if $n \geq 4$.

Intuitively, the larger the number of consumer characteristics disclosed to the advertisers, the better the match between the consumer and the winning advertiser, and therefore the higher the social value of information.

From the point of view of the monopoly platform, if there are only two advertisers, revealing some information will decrease the second-highest bid below the mean valuation, i.e., zero. Therefore, the platform is better off revealing no information. If $n = 3$, the expected value for the second-highest bidder is just equal to the mean valuation (i.e., 0). Finally, if $n \geq 4$, disclosing more information increases the expected value for the second-highest bidder, and hence the advertising revenue.

### 5.2 Information revelation and the advertising market

We now study how the market structure of the advertising market (i.e., the number of advertisers, $n$) affects information revelation.

We begin by studying how the socially-optimal amount of information varies with the number of advertisers. We obtain the following result.

**Result 3** : The socially-optimal amount of information is increasing in the number of advertisers.

This result corresponds to Proposition 1 in Gauza (2004). The social planner cares about the value of the ad slot for the winning advertiser, $v_1(c)$, and the cost of information disclosure, $\delta c$. As shown by Result 1, the social value of information $v_1(c)$ increases with

---

6Note that the maximum value of a consumer characteristic is 1. So, we need to assume that $\delta < 1$, otherwise a platform would never disclose any information to the advertisers.
the number of characteristics. In addition, when the number of advertisers increases, the
expected value of the match also increases, since there is a higher probability of a good
match. This implies that the amount of information that the social planner would decide
to disclose is increasing in the number of advertisers.

We now turn to the information revelation decision of a monopoly platform. We
obtain the following result.

Result 4: A monopoly platform (i) does not reveal any consumer information if \( n = 2 \)
or \( n = 3 \); (ii) for \( n \geq 4 \), reveals an amount of information increasing in the number of
advertisers.

This result corresponds to Proposition 3 in Ganuza (2004). Though the objective
function of the monopoly platform is different from that of the social planner – the
former maximizes the second-highest bid and the latter the highest bid, all net of the
information costs – the intuition is similar to the one provided for the previous result.

Finally, we analyze the amount of information revealed under platform competition.
We obtain the following result.

Result 5: Under platform competition, (i) the platforms reveal no consumer informa-
tion if \( n = 2 \) or \( n = 3 \); (ii) for \( n \geq 4 \), the amount of information revealed in equilibrium
is increasing in the number of advertisers.

Under platform competition, platforms compete to attract single-homing advertisers.
Therefore, their incentive is to disclose as much information as possible to increase the
dispersion in valuations among the advertisers, and hence, to maximize the expected
surplus of the winning advertiser. However, the platforms have also to take into account
their information costs.

We find that this tension between maximizing information revelation to capture the
winning advertiser and information disclosure costs depends on the number of advertisers
in the market. This is because the expected revenue of a platform (gross of information
costs) increases with the amount of information disclosed, which implies that the number
of advertisers has to be sufficiently high for the revenue to cover the costs. This explains
why information revelation increases with the number of advertisers.

Finally, we compare the information disclosure with a monopoly platform and with
competing platforms to the social optimum.

Result 6: (i) the amount of information revealed by competing platforms is higher than
with a monopoly platform; (ii) a monopoly platform reveals less information than what is
socially optimal; (iii) competing platforms reveal less or more information than what is
socially optimal.
Under platform competition, the platforms compete in information disclosure to attract single-homing advertisers. It is therefore natural to find that information disclosure is higher with competing platforms compared to a monopoly platform.

The second result is due to the fact that the monopoly platform maximizes the value for the second-highest bidder, which is lower than the value for the highest bidder. Since the marginal cost of information disclosure is constant and the same for the social planner and the monopoly platform, the latter ends up disclosing less information than the former.

Finally, the third result shows that platform competition can reduce inefficiency in some cases, but can also generate another type of inefficiency when the competing platforms disclose too much information compared to the social optimum.

As an example, Figure 2 shows the socially-optimal amount of information to be disclosed (in brown on the figure) and the equilibrium amount for a monopoly platform (in red) and for competing platforms (in blue).

![Figure 2: Number of characteristics disclosed as a function of the number of advertisers, for δ = 0.3. Red bars: monopoly; Blue line: competition; Brown line: social planner.](image)

5.3 Incremental value of information

Now, we study how the incremental value of an additional piece of information, i.e., \( v(c + 1) - v(c) \), varies with (i) the amount of information already disclosed (i.e., \( c \)), and (ii) the number of advertisers in the market (\( n \)).
Figure 3: Incremental social value (vertical axis) as a function of the number of characteristics disclosed (horizontal axis). Red bars: \( n = 5 \); Blue line: \( n = 10 \); Green line: \( n = 15 \); Purple line: \( n = 20 \).

Question (i) corresponds to the question of whether we should expect increasing or decreasing returns in the value of data. Question (ii) asks whether the market structure in the advertising market can affect the incremental value of data.

We have the following results.

**Result 7**: The incremental social value of consumer information is decreasing in the number of characteristics disclosed.

Figure 3 provides an illustration of this result. The figure shows how the incremental social value, \( v_1(c+1) - v_1(c) \), varies with \( c \) for different numbers of advertisers (5, 10, 15 and 20). We observe that this incremental value is always decreasing in \( c \). This result therefore suggests decreasing social returns to consumer data.

**Result 8**: The incremental private value of consumer information varies non-monotonically with the number of characteristics disclosed.

Figure 4 shows how the incremental private value, \( v_m(c+1) - v_m(c) \), varies with \( c \) for the same numbers of advertisers as Figure 3 (i.e., 5, 10, 15 and 20). As we can see, the incremental private value can either increase or decrease in \( c \). This result therefore suggests that there is no clear increasing or decreasing private returns to consumer data. The relation between the incremental value of data and the amount of data already disclosed is more complex.

Finally, we analyze how the incremental value varies with the number of advertisers. We have the following result.
Result 9: The incremental social value of consumer information increases with the number of advertisers in the market. The incremental private value can either increase or decrease with it.

Figure 5 provides an example of the private and social incremental value of consumer information as a function of the number of advertisers, $n$, for $c = 3$. We observe that the incremental social is increasing for all $n$, whereas the incremental private value is first increasing then decreasing in $n$.

5.4 Cost of information

Finally, we study the impact of the information cost $\delta$ on the amount of information disclosed by the social planner, a monopoly platform, and competing platforms.

Result 10: The amount of information disclosed by the social planner, the monopoly platform, or competing platforms is decreasing in the cost of information.

This result is intuitive: when the cost of information increases, players react by reducing the amount of information disclosed to the advertisers. Figure 6 shows an example for $n = 20$. We observe that for low costs of information, the amount of information disclosed by the social planner, a monopoly platform and competing platforms coincide (and is maximum: $c = C = 10$). For larger costs of information disclosure, the amount of information is decreasing and the amount chosen by the social planner, the monopoly platforms and competing platforms diverge.
Figure 5: Incremental value (vertical axis) as a function of the number of advertisers (horizontal axis), at $c = 3$. Red bars: incremental private value; Blue line: incremental social value.

Figure 6: Impact of the cost of information (horizontal axis) on the amount of information disclosed (vertical axis), with $n = 20$. Red bars: monopoly platform; Blue line: competing platforms; Brown line: social planner.
6 Conclusion

Our analysis of the value of consumers’ personal information for platforms and advertisers has revealed some properties that are natural and that correspond to findings found elsewhere about the value of information disclosure in auction settings. Our results that platforms with strong market power are less prone to use consumers’ data than what is socially optimal suggest that intervention to limit data collection by platforms may be misplaced. On the other hand, competition among data generate stronger incentives for platforms to collect data and may even lead to too strong a use of personal data. So, in some sense, data collection by one strong dominant platform may be less of a concern that when the consumers has to choose one platform and only one within an oligopoly.

At the same time, profit maximizing platforms, irrespective of their market power, may face at least locally increasing returns to consumers’ data, which may lead to more concentration in the platform market.

Although it enables a rich treatment of information and consumers’ data, our model remains very simple on many aspects. Perhaps one important limitation is our assumption of a competitive bottleneck situation where consumers singlehome while advertisers multihome. Data can be collected from many different sources, on many different platforms, and these data can then be processed, consolidated and aggregated by data brokers. So, an obvious and relevant development of our research would allow for such a formalization of the market for data and would lead to a more realistic representation of the current evolution of the digital economy. This extension, however, is left for future research.
References


Autorité de la Concurrence and Bundeskartellamt (2016). “Competition law and data.”


