# Carpe data: Regulating the market of personal information with present-biased users

Laura Abrardi[*], Carlo Cambini[†]

*Preliminary version*
January 2020

**Abstract**

We study the potential effects of regulation over the use of data by a monopolistic platform by means of a theoretical model. In digital markets, platforms are increasingly characterized by market power owing to strong network externalities. Their profit mainly originates from the sale of data that users freely and excessively provide due to their myopia on costs related to the loss of privacy. We thus consider a model in which a platform offers a service to users, but usage entails to users a privacy cost of which they might be partly unaware. As the platform sells the data to advertisers, it does not have an incentive to try to reduce the quantity of data that present-biased users provide. Regulating the platform by contracting over the price of data released to advertisers can improve welfare relative to the unregulated outcome. We study and compare two regulatory approaches: a price regulation on advertising; and an opt-in regulation, according to which consumers must actively confirm their consent for data collection.

**Keywords:** Antitrust, data, regulation, privacy.
**JEL codes:** D82, D83, D86, L12, L51

[*]Politecnico di Torino, DIGEP. Email: laura.abrardi@polito.it.
[†]Politecnico di Torino, DIGEP. Email: carlo.cambini@polito.it.

# 1 Introduction

In our economy, most of the data is collected by internet platforms, which hold considerable market power due to the presence of strong network externalities. As reported by *The Economist* (2017)[1], data are the "world's most valuable resources" and their exploitation is at the core of the new business models of digital platforms. At the same time, the massive and unprecedented scale of data is creating serious concerns to policy makers and the public for the large loss in terms of privacy. Competition commissions throughout the world are expressing concerns on the implications of the business models of digital platforms for competition, consumers, and society. Already in September 2018, the United Kingdom established the Digital Competition Expert Panel, with the aim to consider the potential opportunities and challenges the emerging digital economy may pose for competition. In a report for the European Commission, Cremer et al. (2019) also point out the risk for the market of the new digital giants, such as Google, Facebook and Amazon, invoking specific rules for data access and sharing as well as the creation of a new Authority for data regulation. Finally, the Australian competition and consumer commission also observed that the breadth and scale of the user data collected by platforms is relevant both for the assessment of their market power and for consumer concerns (ACCC, 2019).

Facing the challenges of digitalisation might require a revision of the current regulatory framework, and indeed many countries are considering policy changes in this area. For example, the UK government is currently establishing a new regulatory framework for digital markets, whereby companies that allow users to share user-generated content will be subject to an independent regulator[2].

If, on the one hand, the problem of controlling user-generated data is at the forefront of Antitrust authorities, on the other hand it is still unclear how the problem should be addressed. In this paper, we ask the following questions: How should policy makers regulate the data industry? Which kind of regulatory instruments could be implemented? And what impact would the regulation have on welfare?

[1] The Economist, May 6th 2017, "Regulating the internet giants. The world's most valuable resource is no longer oil, but data".

[2] Online Harms White Paper, Presented to Parliament on April 2019 and available at www.gov.uk/government/publications.

From a policymaking perspective, regulating people's data is challenging for two main reasons. First, it is difficult to balance out the value of information and the value of privacy. For example, ensuring the privacy of a consumer's purchases may protect her from price discrimination but also deny her the potential benefit of targeted offers and advertisement (Laufer and Wolfe, 1977).

Second, few consumers are fully informed of, or fully understand, the scale and scope of data collected and the main contents of current privacy regulation. While consumers claim to care about how their data will be used, in practice they often show little concern about it in their daily behavior (the so-called privacy paradox, Acquisti et al. (2015)). For instance, in one laboratory experiment, Tsai et al. (2011) demonstrate that people often do not pay any attention to the merchants' privacy policies, despite caring for them. Present-biased preferences can induce even the most privacy-conscious individuals to reveal personal information, if the immediate gratification from disclosure trumps the delayed, and hence discounted (and possibly uncertain), future consequences (Acquisti, 2004; Tucker, 2018). As a consequence of such inconsistency, consumers often relinquish considerable control over their uploaded content to digital platforms, thereby strengthening platforms' market power, and the potential of using the collected data in ways that harm competition and society.

From a policy perspective, the central question is thus what, if anything, should be done to regulate consumers data. So far, the efforts of regulators went as far as the definition of the limits for the processing of personal information through privacy regulations. In Europe, for example, the regulation of data management rests on the pillar of the recent General Data Protection Regulation (GDPR), which puts considerable constraints on the collection, aggregation and processing of data without the consent of the individual who generated this data. This kind of "opt-in" regulation thus requires the consumers to provide thier consent to collect and use the data generated. However, in many cases it is not clear what is "personal" and "non-personal" information. More importantly, a legislative gap emerges for the cases where the consent is freely, but inattentively and lightly granted. In other words, the GDPR can protect consumers only to the extent that it gives them the right to refuse their consent to the use of their personal data, but there is little it can do whenever they myopically consent and provide data.

If people lack the necessary lucidity for navigating the complex privacy trade-offs according to their best interest, traditional tools such as choice and consent no longer provide adequate protection (Solove, 2013). More invasive regulatory interventions may be needed to balance the interests of the subjects of data against the power of data brokers and limit the latter's market power.

In the attempt to construct a set of flexible instruments, actual privacy regulations focus on defining the rights for the use of data, and fail to intervene directly on data brokers and on their incentives for data generation. However, platforms have a powerful influence on consumers' privacy-related decisions. For example, it has been shown (Tsai et al., 2011) that when people are provided with salient information about the differences in privacy protection afforded by the various merchants on a search engine, they tend to switch (and share their credit card information with) more privacy-protecting (and more expensive) merchants.

Despite their dominant position, and despite their being a central element affecting the consumers' decision to disclose personal information, at the moment platforms are not directly subject to regulation, except that of privacy laws defining the property rights for personal information. Under this respect, the regulation of the market of platforms in still in its infancy, relative to other markets like the energy or telecommunications ones. Quite importantly, more direct regulation could have a significant impact on the outcomes in the market for data. In fact, a large part of the platforms' revenues originate from the sales of data to other markets, such as advertisers. Hence, platforms have little incentive to limit the amount of personal information that people unwittingly or lightly provide, despite the fact that their effort in raising the users' awareness could greatly impact the users' decisions in terms of data disclosure.

In this paper, we suggest alternative policy instruments that could complement actual privacy regulations and better help to regulate the data-collection behavior of a monopolistic platform in the presence of myopic users.

In our model, a platform –for example, a search engine– connects users to advertisers. The use of the website produces some traffic-generated data, so that users incur in a privacy cost due to the possible future (mis)use of their personal information. Users may be myopic and present time-inconsistence preference leading them to underestimate their future cost of privacy. Hence, when deciding the quantity of data to provide to the platforms, users might be present-biased on the future privacy cost. We assume that users are heterogeneous in their present-bias, which can take only two values, high or low. When users have a high present-bias, they disclose too much information.

The platform can develop an algorithm that make consumers more informed about their internet use. At the same time, the platform can monetize the data collected by selling them to advertisers. In this simple setup, the platform does not have any incentive to privately exert any effort and reduce the data subsequently sold to advertisers. Conversely, in the presence of a social planner, constraints can be imposed on the quantity of data sold, and on the price that advertisers should pay for getting data. In particular, the social planner wants to discourage data-generating traffic and imposes some remedies. We study and compare two alternative regulatory approaches: a price regulation on advertising; and an opt-in regulation, according to which consumers must actively confirm their consent for data collection, mimicking the current EU privacy regulation approach.

Our work is related to the recent and growing strand of literature on data acquisition, their impact on platform business (i.e. targetting) and on consumers' privacy. Recent papers have analysed the impact of competition on privacy (Casadesus-Masanell and Hervas-Drane, 2015), while others focus on the interplay between acquiring information to perform a better personalized price while accounting for the consumers' loss of privacy (Braulin and Valletti, 2016; Montes et al., 2019). Other papers study the impact of taxation on data collection (Bloch and Demange, 2018; Bourreau et al., 2018). Our research is also associated with the literature on data acquisition and pricing in a two-sided framework (Bergemann and Bonatti, 2015), which however do not address the problem of privacy costs.

Our paper instead is mostly focused on the potential role that regulation may have in the data acquisition market. To our knowledge, only two papers are dealing with this issue. The first one is the paper by Choi et al. (2019). The paper provides a theoretical model of privacy in which data collection requires consumers' consent and consumers are fully aware of the consequences of such consent. Their goal is to understand the potential limit of the current EU General Data Protection Regulation (GDPR). The second paper is done by Acemoglu et al. (2019) who model a data market in the presence of externalities among users, as in Choi et al. (2019), and (some) users value their privacy. This paper shows how the price of data is affected by data externalities leading to excessive data sharing. From a policy point of view, the authors propose a scheme based on mediated data-sharing that improves efficiency. Both these papers, however, though accounting for the potential privacy loss by consumers, do not consider the role of time-inconsistent preference that consumers may have and the role that myopia could have in inducing consumers to share their data, as we consider in this paper.

Indeed, our goal is to link the above mentioned literature on data acquisition and its impact o privacy with the literature pertaining to the inconsistency of preferences underlined by behavioral sciences (Kahneman and Tversky, 1979; Ainslie, 1992; Laibson, 1997). Because of this non-standard discount structure, a conflict arises between today's preferences, and the preferences that will be held in the future. As a consequence, in some contexts people make errors and fail to behave in their own best interests. In our model, people myopically discount their privacy cost that instead the social planner should account for when defining his policies. Moreover, we study two different regulatory instruments: the "opt-in" regulation, as in Choi et al. (2019), and its impact in presence of myopic users; and a more direct price regulation on the data value for advertisers.

Our contribution to the literature is thus twofold. First, we model the effort exerted by the platform to develop a value-enhancing algorithm. In our model, the platform is not merely a subject connecting two sides of a market, and internalizing the network externalities via the prices, but it is an active agent who can influence directly the quality of the service provided to users. Second, we introduce the element of the users' myopia. While platforms are quasi-monopolists, they present peculiar features in that they trade data, rather than goods. Data pose specific challenges both because of their public nature, and because they originate privacy costs which some users may myopically underestimate.

The next section of the paper presents our model. In Section 3, a social planner is introduced with the possibility to contract over the advertising revenue, under the assumption that the level of users' present-bias is the platform's private information. Section 4 analyzes optimal opt-in regulation. Section 5 concludes.

# 2 A simple model

We model the market of data by considering the interaction between a monopolist data-broker firm (a digital platform), users and advertisers.

A digital platform (e.g., a search engine) provides a service to a unit mass of consumers (the users). In choosing the level of their usage of the service, users seek to maximize their utility $u(x)$, with $u(0) = 0$, $u_x > 0$, $u_{xx} < 0$, where $x$ is the quantity of information obtained from the search. For simplicity sake, we adopt the functional form $u(x) = \frac{x^{1-\alpha}}{1-\alpha}$. The quantity of information $x$ depends on the quantity $q$ of usage of the service, but also by the effort $e$ exerted by the platform into developing a performing algorithm to process data and provide accurate search results. In particular, $x = q + e$: one additional unit of effort by the firm into developing a more effective algorithm allows consumers to obtain more relevant information, given the same quantity of searches. For example, $q$ can be interpreted as the number of searches on the platform, and $e$ as the ability of the algorithm to deliver more targeted search results. The platform incurs in the cost $\psi(e) = e^2/2$ for delivering effort $e$.

The traffic $q$ on the website reveals personal information about the user. In particular, the usage $q$ of the service produces data $dq$, with $d > 0$. The platform monetizes the data by selling them to advertisers at a fixed fee $R$.

The value generated by the data to advertisers is equal to $\gamma dq$, with $\gamma > 0$.

Users incur in a constant marginal cost $c$ for each unit of data produced. The cost $cdq$ can be thought of as the cost that consumers incur for their loss of privacy due to the future exploitation of their personal data. Users are present-biased in regards to this privacy cost, and apply a time-inconsistent discount factor $\beta_j$, with $j = \{L, H\}$, and $\beta_L < \beta_H \leq 1$.

Advertisers' profit is $v(q) = \gamma dq - R$. We assume that the platform is profit maximizer and has market power, so that it extracts all the profit from the advertisers, i.e. $R = \gamma dq$. Hence, the platform obtains the profit $\pi(q, e) = \gamma dq - e^2/2$. Note that, in this setup, advertisers are not allowed to decide the quantity of data to purchase. This assumption significantly simplifies the two-sided feature of the model, allowing to focus on the role of present-bias and platform's effort with no loss of generality[3].

[3] A two-sided framework would simply amplify the effects that we find in our simplified setup.

Users decide the quantity of usage by maximizing their utility, net of the privacy cost: $\max_q u(q+e) - \beta_j cdq$, obtaining

$$u_x(q+e) = \beta_j cd, \tag{1}$$

i.e. $\tilde{x}_j = u_x^{-1}(\beta_j cd)$ and $\tilde{q}_j(e) = u_x^{-1}(\beta_j cd) - e$, with $j \in \{L, H\}$. A better performing algorithm improves the efficiency of searches, thereby reducing the time spent on a search, $\tilde{q}_j$.

From a social planner's standpoint, the users' surplus is defined as the users' utility, net of the privacy cost: $S(q, e) = u(q+e) - cdq$.

The regulator maximizes the sum of the expected surplus to users plus the expected profit for the platform. Specifically, the regulator's objective is to maximize the welfare function $W = S(\tilde{q}, e) + \pi(\tilde{q}, e)$.

The timing is as follows. First, the platform develops its algorithm by choosing its effort $e$. Second, users purchase the quantity of usage $q$, usage generates data $dq$ and the platform sells the data to advertisers. Fourth, users pay the privacy cost.

We assume without loss of generality that the usual time-consistent discount factor is equal to 1, so as to focus exclusively on the role of the users' time-inconsistency in the privacy problem.

## 2.1 Welfare maximizing benchmark

In this section we characterize the welfare maximizing solution under perfect information, for a useful comparison both with the case of unregulated monopolist, and with the case of a regulated one under asymmetric information.

For a meaningful representation of the privacy problem, we make the following assumption.

**Assumption 1.** The marginal cost of data is larger than their marginal benefit, i.e. $c > \gamma$.

This assumption implies that data generate a social cost $c$ –in terms of loss of privacy– that is higher than their benefit $\gamma$ in terms of informational value.

The effort that maximizes the welfare solves the problem

$$\max_{e_j} u(\tilde{x}_j) - (c - \gamma) d\tilde{q}_j(e_j) - \psi(e_j). \tag{2}$$

Using the users' demand function (1), the welfare function can be rewritten as:

$$\max_{e_j} u(u_x^{-1}(\beta_j cd)) - (c-\gamma)d\left(u_x^{-1}(\beta_j cd) - e_j\right) - \psi(e_j). \tag{3}$$

Recalling that $c > \gamma$ from Assumption 1, the FOC of (3) yields

$$e_j^* : (c-\gamma)d = \psi'(e). \tag{4}$$

Note, from (4), that $e_L^* = e_H^*$. Hence, $q_j^* = u_x^{-1}(\beta_j cd) - e_j^*$, i.e. $q_L^* > q_H^*$. Moreover, $R_j^* = \gamma d q_j^*$. Then, $R_L^* > R_H^*$. The lower $\beta_j$ is, the higher is the users' propensity to generate data, as myopic users neglect part of the privacy cost.

The first best level of welfare is

$$W^* = u(u_x^{-1}(\beta_j cd)) - (c-\gamma)d\left(u_x^{-1}(\beta_j cd)\right) + (c-\gamma)de_j^* - \psi(e_j^*). \tag{5}$$

The platform's effort can improve the level of welfare, despite the users' overprovision of data.

## 2.2 Unregulated monopolist

In this Section we briefly describe the equilibrium in the case of an unregulated platform.

In the case the market for data is unregulated, the platform chooses the quantity $q_j$ and the quality of its algorithm $e_j$ so as to maximize its profit $\pi_j(q_j, e_j)$ subject to (1) and to the advertisers participation constraint ($v(q_j) \geq 0$).

As the participation constraint is binding in the monopolist's optimum, i.e. $v(q_j) = 0$, the platform's problem becomes

$$\max_{e_j} \gamma d\tilde{q}_j(e_j) - \frac{e_j^2}{2}$$

As the function $\tilde{q}_j(e_j)$ is decreasing in $e_j$, the platform chooses $e_j^U = 0$, where the superscript $U$ denotes the unregulated solution.

From (1), we obtain $q_j^U = u_x^{-1}(\beta_j cd)$. The platform makes a profit from the sale of users' data, hence it does not have the incentive to discourage data-generating traffic. In the absence of regulation, the platform curtails the costly and revenue-reducing effort. The privacy cost entirely falls on users, who disclose a large quantity of personal data as a consequence of the low effort of the platform. Users' surplus is ultimately be reduced, with detrimental effects on welfare.

# 3 Regulated monopolist with private information on the users' present-bias

We now assume that the value of $\beta_j$ is the platform's private information, although its probability distribution is common knowledge. Let us denote with $\lambda$ the probability that $j = H$, and with $1 - \lambda$ the probability that $j = L$. While the effort $e$ and the users' present-bias $\beta_j$ are unobservable by the regulator, the usage $q_j$ can be observed and verified.

In this setup, a platform catering to H users (for brevity, we will refer to it as an 'H platform') might prefer the contract $(R_L, q_L)$ designed for the L platform. In particular, the H users' demand is

$$\tilde{q}_H(\hat{e}_H) = u_x^{-1}(\beta_H cd) - \hat{e}_H.$$

The effort $\hat{e}_H$ that is necessary to have $\tilde{q}_H(\hat{e}_H)$ coincide with $\tilde{q}_L(e_L) = u_x^{-1}(\beta_L cd) - e_L$ is

$$\hat{e}_H(e_L) = e_L - u_x^{-1}(\beta_L cd) + u_x^{-1}(\beta_H cd), \tag{6}$$

which can also be expressed as

$$\hat{e}_H(e_L) = e_L - \Delta, \tag{7}$$

where $\Delta = u_x^{-1}(\beta_L cd) - u_x^{-1}(\beta_H cd)$. Given that $\beta_H > \beta_L$, then $\Delta > 0$ and $\hat{e}_H(e_L) < e_L$.

Suppose that the regulator offers the first best menu $\{(R_L^*, q_L^*); (R_H^*, q_H^*)\}$ and that the platform, after having privately observed that users are of type H, chooses the contract $(R_L^*, q_L^*)$. As $\hat{e}_H(e_L^*) < e_L^*$, a platform H can save on effort in the case it were to achieve the outcome $q_L^*$ of a platform L. Therefore, under asymmetric information, the platform H earns a rent by choosing a lower effort that delivers $x_L^*$. In particular, the profit $\hat{\pi}_H$ obtained by the platform H when it exerts effort $\hat{e}_H$ and mimics the behavior of a platform L in the market for data, given the first best contracts, is:

$$\hat{\pi}_H^* = R_L^* - \psi(\hat{e}_H(e_L^*)). \tag{8}$$

Given that $R_L^* > R_H^*$ and $\hat{e}_H(e_L^*) < e_L^* = e_H^*$, it holds $\hat{\pi}_H^* > \pi_H^* = R_H^* - \psi(e_H^*)$. The platform H earns a rent by failing to self-select properly withing the first best menu. The H type mimics the L one and selects contract $(R_L^*, q_L^*)$. This result is a direct consequence of the fact that, being H users more aware of their privacy costs, the platform H must exert less effort in developing its algorithm in order to induce its users to provide the same traffic-generated data than less aware ones. The complete information optimal contracts can no longer be implemented under asymmetric information.

The size of the informational rent depends on the effort $e_L$. In particular, $\hat{\pi}_H = \gamma d q_L(e_L) - \psi(e_L - \Delta)$: a higher effort $e_L$ reduces the deviation profit obtained by the H platform, both because it decreases the data provided by users –hence, the advertising revenue, and because is increases the cost of effort.

When firms have an informative advantage on the type $j$ of users, the regulator must offer a menu of incentive compatible contracts, which we denote with $\{(R_L^{SB}, q_L^{SB}); (R_H^{SB}, q_H^{SB})\}$. Each contract includes the specification of the quantity of traffic $q_j^{SB}$. In the optimal solution, the firm will self-select and choose contract $\left(R_H^{SB}, q_H^{SB}\right)$ if $j = H$, or $\left(R_L^{SB}, q_L^{SB}\right)$ if $j = L$. The correct self-selection of firms is ensured by the incentive compatibility of the menu of contracts.

By decreasing the effort $e_L^{SB}$, it is possible to reduce the rent that the platform H extracts because of its informative advantage. At the same time, a lower effort $e_L^{SB}$ entails higher privacy costs for the L users. Hence, the optimal second best menu entails the well-known trade-off between rent extraction and efficiency.

The regulator solves the following problem:

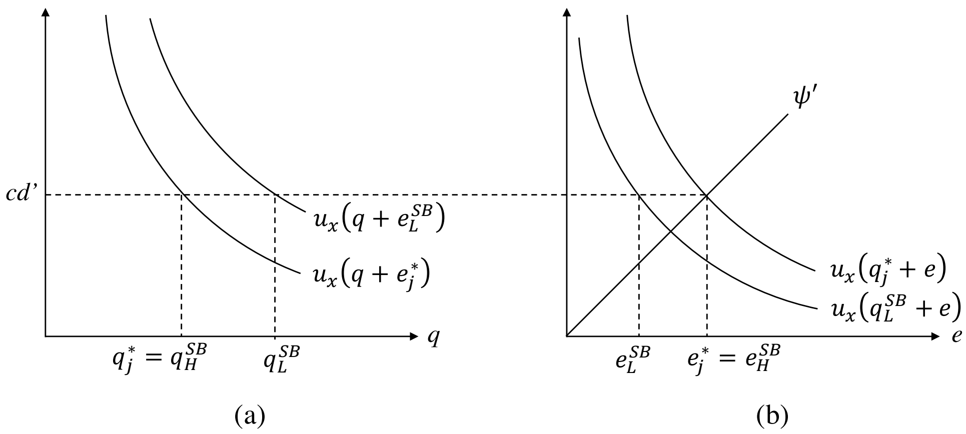
$$\begin{aligned}
&\max_{q_H, e_H, q_L, e_L} \lambda[u(x_H) - (c-\gamma)dq_H - \psi(e_H)] + \\
&+(1-\lambda)[u(x_L) - (c-\gamma)dq_L - \psi(e_L)] && (9) \\
&s.t. \\
&\gamma d q_j - \psi(e_j) \geq 0 \;\; \forall j && (10) \\
&\gamma d q_H - \psi(e_H) \geq \gamma d q_L - \psi(\hat{e}_H(e_L)) && (11) \\
&\gamma d q_L - \psi(e_L) \geq \gamma d q_H - \psi(\hat{e}_L(e_H)) && (12)
\end{aligned}$$

Constraint (10) constitutes the participation constraint of the platform $j$, while constraints (11) and (12) ensure the incentive compatibility of the contracts.

The solution of the regulator's problem is defined in Proposition 1.

**Proposition 1** *Optimal regulation under imperfect information is characterized by $e_L^{SB} > e_L^*$ and $e_H^{SB} < e_H^*$.*

Figure 1: Benchmark and second best solution of optimal regulation




Moreover, and $q_H^{SB} > q_H^*$. Finally, $R_H^{SB} > R_H^*$.

From Proposition 1, the L firm's effort is distorted upward. The overeffort of the L firm with respect to the first best decreases the data supplied by users in the L setting: $q_L^{SB} < q_L^*$. Lower data generate lower advertising revenue, i.e. $R_L^{SB} < R_L^*$.

Figure 1 provides a graphical representation of the results of Proposition 1. In panel (a) of Figure 1, the first and second best quantities are defined. Panel (b) of Figure 1 allows to identify the optimal and second best effort.

The incentive compatibility of the menu of contracts is ensured by the provision to the H firm of informational rents. These rents are decreasing with the effort exerted by the L firm. As a consequence, the reduction of rents requires the upward distortion of $e_L$.

# 4 Opt-in regulation

In the previous Section, the market inefficiency caused by the users' myopia is corrected by inducing the platform to exert effort. In this Section, we consider an alternative regulatory policy, which tries to rectify the market inefficiency by influencing the user's behavior rather than the platform's.

In order to prevent an excessive collection of personal data, "opt-in" regulations under the GDPR require that a customer must actively confirm her consent for data collection with explicit prior permission such as ticking an unchecked opt-in box. We assess the effectiveness of such regulations with myopic users.

The effect of regulating the usage of the platform is to reduce the immediate benefit of usage. For example, cookie banners interrupt the user's experience of the website.

We assume that the social planner can regulate the usage of the website by choosing the intensity $\rho \geq 0$ of the regulation, where $\rho = 0$ indicates an absence of regulation. The only cost of regulation is to reduce the user's utility by the factor $-\rho q$. Note that, in our setup, opt-in regulation has no beneficial effect on users' present-bias. Its only purpose is to annoy the user and reduce his traffic on the website.

Users decide the quantity of usage by maximizing their utility, net of the privacy cost:

$$\max_{q} u(q + e) - \beta_j cdq - \rho q, \tag{13}$$

obtaining

$$u_x(q + e) = \beta_j cd + \rho, \tag{14}$$

i.e. $\tilde{x}_j = u_x^{-1}(\beta_j cd + \rho)$ and $\tilde{q}_j(\rho, e) = u_x^{-1}(\beta_j cd + \rho) - e$, with $j \in \{L, H\}$. Note that $\tilde{q}_j$ is decreasing in $\rho$: more intense regulation decreases the marginal utility of usage, thereby reducing the time spent on the website.

The platform chooses the effort to maximize $\gamma d\tilde{q}_j(\rho, e) - \frac{e^2}{2}$. As the function $\tilde{q}_j(\rho, e)$ is decreasing in $e$, the platform chooses $e^U = 0$.

The social planner decides the intensity $\rho$ of regulation by maximizing the welfare

$$\max_{\rho_j} u(\tilde{q}_j + e^U) - (c - \gamma)d\tilde{q}_j - \rho\tilde{q}_j. \tag{15}$$

By comparing (13) with (15), it is possible to note that users' demand presents two distortions with respect to the first best. First, users underestimate the cost of privacy, i.e. $\beta < 1$. This induces overconsumption. Second, users neglect the positive externality $\gamma dq$ that their provision of data has on the advertising sector. This induces underconsumption. Which of the two distortions prevails depends on the severity of the present-bias. If $(c - \gamma)d > \beta_j cd$, i.e. $\beta_j < \bar{\beta} = \frac{c-\gamma}{c}$, then overconsumption occurs. Conversely, if $\beta_j > \bar{\beta}$, then there is underconsumption.

The intensity of regulation has two effects on welfare. On the one hand, a higher $\rho$ has a direct, negative effect on welfare by reducing the user's surplus. On the other hand, a higher $\rho$ has an indirect, positive effect on welfare as it influences the quantity $\tilde{q}_j$, hence it contributes to correct the distortions.

Using the users' demand function $\tilde{q}_j(\rho, e^U) = u_x^{-1}(\beta_j cd + \rho)$, the welfare function can be rewritten as:

$$\max_{\rho_j} u\left(u_x^{-1}(\beta_j cd + \rho_j)\right) - [(c-\gamma)d + \rho_j]u_x^{-1}(\beta_j cd + \rho_j). \tag{16}$$

The FOC of problem (16) allows to obtain

$$\rho^* = u_x\left(\frac{cd(\beta_j - \bar{\beta}_j)}{u_{xx}}\right) - \beta_j cd \tag{17}$$

If $\beta_j < \bar{\beta}_j$, the optimal level of regulation is decreasing in $\beta_j$. Intuitively, very myopic users offer a disproportionate amount of data. Only a invasive regulation can stop them from incurring in huge privacy costs. However, if $\beta_j > \bar{\beta}_j$, $\rho^* = 0$. Mildly myopic users are excessively parsimonious with the information they provide, and they fail to account for the positive externality on advertising. Hence, no regulation should be implemented.

When $\beta_j > \bar{\beta}_j$, welfare is $W^\rho = u\left(u_x^{-1}(\beta_j cd)\right) - (c-\gamma)du_x^{-1}(\beta_j cd)$. Then, $W^\rho > W^*$ from (5). When users are mildly myopic, a price regulation dominates a opt-in regulation, at least in the case of perfect information.

# 5 Conclusions

To be defined

# References


ACCC (2019). Digital platforms inquiry, final report. Report 06/19 1545, Australian Competition and Consumer Commission. 1

Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2019). Too much data: Prices and inefficiencies in data markets. NBER Working Paper 26296, NBER. 1

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th ACM Conference on Electronic Commerce*, EC '04, pages 21–29, New York, NY, USA. ACM. 1

Acquisti, A., Brandimarte, L., and Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221):509–514. 1

Ainslie, G. W. (1992). *Picoeconomics: the strategic interaction of successive motivational states within the person*. Cambridge University Press. 1

Bergemann, D. and Bonatti, A. (2015). Selling cookies. *American Economic Journal: Microeconomics*, 7(3):259–94. 1

Bloch, F. and Demange, G. (2018). Taxation and privacy protection on internet platforms. *Journal of Public Economic Theory*, 20(1):52–66. 1

Bourreau, M., Caillaud, B., and De Nijs, R. (2018). Taxation of a digital monopoly platform. *Journal of Public Economic Theory*, 20(1):40–51. 1

Braulin, F. C. and Valletti, T. (2016). Selling customer information to competing firms. *Economics Letters*, 149:10 – 14. 1

Casadesus-Masanell, R. and Hervas-Drane, A. (2015). Competing with privacy. *Management Science*, 61(1):229–246. 1

Choi, J. P., Jeon, D.-S., and Kim, B.-C. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173:113 – 124. 1

Cremer, J., De Montjoye, Y.-A., and Schweitzer, H. (2019). Competition policy for the digital era. Report B-1049, Directorate-General for Competition - European Commission. 1

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291. 1

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting*. *The Quarterly Journal of Economics*, 112(2):443–478. 1

Laufer, R. S. and Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues*, 33(3):22–42. 1

Montes, R., Sand-Zantman, W., and Valletti, T. (2019). The value of personal information in online markets with endogenous privacy. *Management Science*, 65(3):1342–1362. 1

Solove, D. J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, 126(7):1880–1903. 1

Tsai, J. Y., Egelman, S., Cranor, L., and Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, 22(2):254–268. 1

Tucker, C. (2018). Privacy, algorithms, and artificial intelligence. In Gans, A. and Goldfarb, editors, *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press. 1