Privacy and the Value of Data

By Simone Galperti and Jacopo Perego*

January 13, 2023

Personal data has become an essential input of the modern economy. Companies and platforms routinely collect it and use it to improve their products and services, and to make pricing decisions. This widespread practice has caught the attention of policymakers and has renewed interest in the issue of data privacy. Across the globe, new legislation has been introduced to give consumers more control over how their personal data is collected and used by firms.[1] These new laws have the potential to impact consumers' welfare, the functioning of data markets, and the multi-billion dollar businesses that rely on them. To understand these effects, economists have employed theoretical, empirical, and experimental methods, leading to a growing body of literature.[2]

This article contributes to this literature by examining how data-privacy laws can affect the value of personal data for firms and, in turn, can impact consumers' welfare. At first glance, it is natural to expect that giving consumers more control will make their data less valuable to firms. However, we argue that the effects of privacy laws are more complex. In particular, they can be redistributive: The data of some types of consumers may become more valuable at the expense of others. This is due to a particular externality that is created by how firms—for example, e-commerce and matching platforms—use consumers' data to mediate interactions between agents with conflicting interests. This externality has been studied by Galperti, Levkun and Perego (2022) (henceforth, GLP).

We illustrate these effects in the context of a stylized model inspired by Bergemann et al. (2015). We formalize privacy protection by introducing elicitation constraints to an otherwise standard information-design problem. In the model, an e-commerce platform intermediates the interaction between a monopolistic seller and a population of heterogeneous buyers. The platform aims to maximize buyers' surplus. To do so, it influences the seller's price by providing information about the buyers. How effective the platform is at influencing the seller ultimately depends on what it knows about the buyers, which is endogenous in the model. We analyze three cases: A benchmark case where the platform can directly observe buyers' data—perhaps because it has collected them without their consent; a case where each buyer controls her data and decides whether to verifiably disclose it to the platform; a case where such disclosure is unverifiable. Our goal is to understand how valuable each buyer's data is for the platform and how these values change when we change how buyers' privacy is protected.

Our analysis yields three main insights. First, protecting buyers' privacy can affect the value of data in complex ways, as it can increase or decrease the value of some buyers' data while not changing that of others' data. Second, privacy protection can impact how data is used by the platform and, therefore, buyers' payoffs. We show that privacy protection can benefit some buyers but harm others, particularly those who have no reason to withhold their data. Third, protecting buyers' privacy increases the average transaction price but also limits trade. Overall, this leaves the seller indifferent but has a negative impact on the platform.

I. The Environment

An e-commerce platform (*it*) mediates the interactions between a population of heterogeneous buyers (*she*) and a single seller (*he*).

The seller sets the price of his product, denoted

* Galperti: UC San Diego (email: sgalperti@ucsd.edu). Perego: Columbia Business School and CEPR (email: jacopo.perego@columbia.edu). This research is supported by grants from the NSF (Galperti: SES-2149289; Perego: SES-2149315). Galperti also gratefully acknowledges financial support from the UPenn's CTIC and the Warren Center for Network & Data Sciences.

[1] In 2016, the European Union passed a groundbreaking legislation—the General Data Protection Regulation—that requires, among other things, that firms request consumers' consent before collecting their data. Since then, similar privacy laws have been enacted in the United States, for example in California, Colorado, Connecticut, and Virginia.

[2] For a review, see Acquisti et al. (2016). On the theoretical side, recent papers include Choi et al. (2019), Hidir and Vellodi (2021), Acemoglu et al. (2022), Ichihashi (2020), Ali et al. (2022), and Galperti et al. (2022). On the empirical side, see, for example, Aridor et al. (2022).

by $a \in A$ where A is finite. The platform is used by a population of buyers, each with a unit demand for the seller's product. A buyer's willingness to pay (WTP) is $\theta \in \{1,2\}$. Together, a and θ determine the buyer's final purchase decision. The buyer purchases the product if $\theta \geq a$; otherwise, she chooses an outside option whose value is zero. Thus, her surplus is $\hat{g}(a,\theta) = \max\{\theta - a, 0\}$. The seller's profit, instead, is $\hat{\pi}(a,\theta) = a\mathbb{1}(\theta \geq a)$, where his marginal cost is assumed to be zero. Finally, the platform's payoff is a combination of the seller's profit and the buyer's surplus $\hat{u}(a,\theta) = r\hat{\pi}(a,\theta) + (1-r)\hat{g}(a,\theta)$. We focus on the case of $r = 0$, where the platform's objective consists of maximizing buyers' surplus. In Section IV, we will explain how things change when $r > 0$.

For each buyer, there is a corresponding *data record* that provides information about her WTP θ. This record could include, for example, her age, gender, and other demographic information. We model this record as the realization of an exogenous signal, denoted by $\omega \in \Omega$. We refer to ω as the *type* of the buyer's record. We assume there are three types of records, $\Omega = \{1,2,\varnothing\}$, with the following properties: $\omega = \theta$ fully reveals to the platform that the corresponding buyer has WTP θ; instead, $\omega = \varnothing$ reveals nothing about the buyer's WTP. In this case, the buyer's θ is 2 with probability $p > 1/2$ and 1 with probability $1-p$. Given action a and an ω record, we denote by $u(a,\omega)$ the *expected* platform's payoff, by $\pi(a,\omega)$ the expected seller's profit, and by $g(a,\omega)$ the expected buyer's surplus.

The collection of all buyers' data records forms the platform's *database*, denoted by $q = (q_1, q_2, q_\varnothing)$, where q_ω is the quantity of ω records. The primitives A, Ω, u, π, p, and q are common knowledge. We focus on the case where $q_1 \leq (2p-1)q_\varnothing$. In this case, when the seller is uninformed, he will charge a price of 2, rendering our problem interesting. We comment on the other cases in Section IV.

The platform is an information designer that mediates each buyer-seller interaction by conveying information about the buyers to the seller to influence his price a. How effective the platform is depends on what it knows about the buyers. We study three cases. In Section II, the platform can observe all buyers' records—perhaps because it has collected them without the buyers' consent. In Section III, each buyer decides whether to verifiably disclose her record to the platform. In Section IV, buyers' disclosure is unverifiable. Our goal is to compute *the value* of each data record *for the platform* and how it depends on the way buyers' privacy is protected.

II. The Value of Data Under No Privacy

We begin with the case where the platform can observe and use the record of each buyer without their consent. We interpret this case as a situation where the platform collects data records with no concern about the buyer's privacy. In this case, the platform faces a standard information-design problem. By usual arguments (see Bergemann and Morris, 2016), we can express this problem as choosing a recommendation mechanism $x : A \times \Omega \to \mathbb{R}_+$ subject to incentive compatibility (IC) and feasibility constraints:

$$\begin{aligned} \mathcal{P}: \max_{x: A\times\Omega\to\mathbb{R}_+} & \sum_{\omega,a} u(a,\omega)x(a,\omega) \\ \text{s.t.} & \sum_{\omega} \big(\pi(a,\omega) - \pi(\hat{a},\omega)\big)x(a,\omega) \geq 0 \quad \forall a,\hat{a} \\ & \sum_{a} x(a,\omega) = q_\omega \qquad\qquad\qquad\quad \forall \omega \end{aligned}$$

We denote an optimal solution of $\mathcal{P}$ by x^* and the resulting *total* payoff of the platform by U^*.

Solution of $\mathcal{P}$. First, note that we can let $A = \{1,2\}$ w.l.o.g. Since $r = 0$, the platform's objective simplifies to $x(1,2) + px(1,\varnothing)$, namely the mass of high-WTP buyers who are charged price 1. The IC constraints in $\mathcal{P}$ are

$$\begin{aligned} x(1,1) - x(1,2) - (2p-1)x(1,\varnothing) &\geq 0, \\ -x(2,1) + x(2,2) + (2p-1)x(2,\varnothing) &\geq 0. \end{aligned}$$

Given this, it is optimal to set $x^*(1,1) = q_1$ and $x^*(2,1) = 0$, as it maximally relaxes the constraints without affecting the platform's payoff. The second IC constraint then always holds and can be ignored. The first constraint must bind. We claim that $x^*(1,2) = 0$ and $x^*(1,\varnothing) = q_1/(2p-1)$. If not, the platform can increase $x^*(1,\varnothing)$ by some ε, which increases expected surplus by $p\varepsilon$. To satisfy the first constraint, it also has to reduce $x^*(1,2)$ by $(2p-1)\varepsilon$, which decreases expected surplus by $(2p-1)\varepsilon$. Overall, this change is beneficial since $2p-1 < p$.▲

Intuitively, in the solution x^*, the platform uses the buyers' records to create two market segments (see Table 1.a). The first contains all buyers with type-1 records and as many buyers with type-$\varnothing$ records as

(a) Solution x^* to $\mathcal{P}$

$x^*(a,\omega)$	$a=1$	$a=2$
$\omega=1$	q_1	0
$\omega=2$	0	q_2
$\omega=\varnothing$	$\frac{1}{2p-1}q_1$	$q_\varnothing-\frac{1}{2p-1}q_1$

(b) Values of Records

	v^*_ω
$\omega=1$	$\frac{p}{2p-1}$
$\omega=2$	0
$\omega=\varnothing$	0

(c) Payoffs

Platform:	$\frac{p}{2p-1}q_1$
$(\omega=1)$-Buyers:	0
$(\omega=2)$-Buyers:	0
$(\omega=\varnothing)$-Buyers:	$\frac{p}{2p-1}q_1$
Seller:	$2(q_2+pq_\varnothing)$

Table 1—: The Problem Without Privacy

possible until the seller is indifferent between charging price 1 or 2 to this segment. The second segment contains all remaining buyers, and this induces the seller to charge price 2.

What is the value of an ω record for the platform? To answer this question, GLP study the dual of $\mathcal{P}$. This is a linear program that selects $v=(v_1,v_2,v_\varnothing)\in\mathbb{R}^3$ and $\lambda_1,\lambda_2\geq 0$ to solve

$$\begin{aligned}\mathcal{D}:\ &\min_{v,\lambda}\sum_\omega v_\omega q_\omega\\ &\text{s.t. for all } \omega\in\{1,2,\varnothing\}\\ &\quad v_\omega\geq u(1,\omega)+(\pi(1,\omega)-\pi(2,\omega))\,\lambda_1\\ &\quad v_\omega\geq u(2,\omega)+(\pi(2,\omega)-\pi(1,\omega))\,\lambda_2,\end{aligned}$$

where v_ω is the multiplier of the feasibility constraint of $\mathcal{P}$, and λ_1 is the multiplier for the IC constraint of $\mathcal{P}$ where the seller is recommended $a=1$ and considers deviating to $\hat{a}=2$ (similarly for λ_2).

We denote a solution to $\mathcal{D}$ by (v^*,λ^*). In it, v^*_ω captures the value of data records of type ω. For every ω, v^*_ω is equal to the marginal change in the platform's total payoff U^* when adding a new ω record to the database. In addition, since by strong duality $\sum_\omega v^*_\omega q_\omega=U^*$, problem $\mathcal{D}$ can be viewed as an internal accounting exercise: It assigns a share (namely, v^*_ω) of U^* to each ω record that reflects its actual contribution to it. For more details, we refer to GLP.

Solution of $\mathcal{D}$. The constraints of $\mathcal{D}$ can be written as

$$\begin{aligned}v_1&=\max\{\lambda_1,-\lambda_2\}=\lambda_1,\\ v_2&=\max\{1-\lambda_1,\lambda_2\},\\ v_\varnothing&=\max\{p-(2p-1)\lambda_1,(2p-1)\lambda_2\}.\end{aligned}$$

Since $2p-1>0$, it is optimal to set $\lambda^*_2=0$ to relax the problem as much as possible. To find λ^*_1, note that the objective of $\mathcal{D}$ becomes

$$\begin{aligned}q_1\lambda_1\ &+\ q_2\max\{1-\lambda_1,0\}\\ &+\ q_\varnothing\max\{p-(2p-1)\lambda_1,0\}.\end{aligned}$$

Since $q_1\leq(2p-1)q_\varnothing$ by assumption, we obtain $\lambda^*_1=\frac{p}{2p-1}$. Hence, $v^*_1=\frac{p}{2p-1}$ and $v^*_2=v^*_\varnothing=0$. ▲

Intuitively, the platform manages to achieve a positive surplus for some high-WTP buyers only because it withholds information about their θ by pooling them with low-WTP buyers in the same market segment. Therefore, v^*_1 reflects the surplus that records of type 1 help the platform achieve with high-WTP buyers who are pooled with low-WTP buyers. By contrast, $v^*_\varnothing$ and v^*_2 reflect the fact that, by themselves, records of type $\varnothing$ and 2 cannot lead to any positive surplus (given $p>1/2$).

III. The Value of Data Under Privacy

We now consider the case where buyers control their data records. While the platform still knows q, it cannot tell buyers apart. Each buyer can disclose the type of her record to the platform. As in many privacy laws, buyers need to give consent for their data to be collected and used by the platform. In this section, we assume that buyers' disclosure is verifiable. That is, each buyer can either disclose her record's type as is or conceal it entirely. From the viewpoint of the platform, a concealed record cannot be distinguished from a record that is of type $\varnothing$ to begin with.[3]

The buyers' incentive to disclose their data depends on how the platform will use it. We can formalize the platform's problem by adding *disclosure constraints* to problem $\mathcal{P}$: for all $\omega\in\{1,2\}$, the rec-

[3]This model of disclosure is akin to Dye (1985). It also relates to Ali et al. (2022), who model buyer's privacy as the voluntary disclosure of verifiable information.

(a) Solution $\hat{x}^*$ to $\hat{\mathcal{P}}$

$\hat{x}^*(a,\omega)$	$a=1$	$a=2$
$\omega=1$	q_1	0
$\omega=2$	$\frac{q_1 q_2}{(2p-1)q_\varnothing+q_2}$	$q_2-\frac{q_1 q_2}{(2p-1)q_\varnothing+q_2}$
$\omega=\varnothing$	$\frac{q_1 q_\varnothing}{(2p-1)q_\varnothing+q_2}$	$q_\varnothing-\frac{q_1 q_\varnothing}{(2p-1)q_\varnothing+q_2}$

(b) Values of Records

	$\hat{v}^*_\omega$
$\omega=1$	$\frac{pq_\varnothing+q_2}{(2p-1)q_\varnothing+q_2}$
$\omega=2$	0
$\omega=\varnothing$	0

(c) Payoffs

Platform:	$\frac{q_1(pq_\varnothing+q_2)}{(2p-1)q_\varnothing+q_2}$
$(\omega=1)$-Buyers:	0
$(\omega=2)$-Buyers:	$\frac{q_1 q_2}{(2p-1)q_\varnothing+q_2}$
$(\omega=\varnothing)$-Buyers:	$\frac{pq_1 q_\varnothing}{(2p-1)q_\varnothing+q_2}$
Seller:	$2(q_2+pq_\varnothing)$

Table 2—: The Problem With Verifiable Privacy

ommendation mechanism x needs to satisfy

$$(1) \qquad \sum_a g(a,\omega)\Big(\frac{x(a,\omega)}{q_\omega}-\frac{x(a,\bar{\omega})}{q_\varnothing}\Big)\geq 0.$$

Note that $x(a,\omega)/q_\omega$ is the probability that the platform recommends price a to the seller *conditional* on the record being of type ω. Therefore, the constraint means that, from the ex-ante viewpoint, a buyer with an ω record has to prefer disclosing her record to withholding it. In the context of our example, constraint (1) is trivially satisfied for $\omega=1$: The seller always charges at least price 1, so the surplus of a buyer with $\omega=1$ is always zero independently of her data disclosure. Therefore, we will add only constraint (1) for $\omega=2$ to problem $\mathcal{P}$ and call this augmented problem $\hat{\mathcal{P}}$.[4]

Constraint (1) captures a common way in which buyers' privacy is protected in practice: Individuals can withhold their data or force the platform to delete it. This formulation involves a convenient separation between disclosure constraints and obedience constraints: The buyer has private information, while the seller has the ability to act. This avoids the concerns of "double deviations" and makes the problem rather tractable. This separation need not hold generally, but may emerge elsewhere in privacy applications. Investigating this aspect seems fruitful for future research.

With this formulation of $\hat{\mathcal{P}}$, we can interpret the platform's problem under privacy as follows. As in $\mathcal{P}$, it is as if the platform still owns all buyers' records. Unlike in $\mathcal{P}$, it is now more constrained in how it can use them—in the sense that it also has to satisfy constraint (1).[5] Given this and that $\hat{\mathcal{P}}$ is still a linear program with the same feasibility constraint as $\mathcal{P}$, we can apply the principles developed in Galperti et al. (2022) to study the value of records under privacy, namely through the dual of $\hat{\mathcal{P}}$, denoted by $\hat{\mathcal{D}}$. Letting $\hat{\mu}\geq 0$ be the multiplier of constraint (1) for $\omega=2$ and defining $\hat{v}$, $\hat{\lambda}_1$, and $\hat{\lambda}_2$ as before, we have

$$\begin{aligned}
\hat{\mathcal{D}}:\ &\min_{\hat{v},\hat{\lambda}_1,\hat{\lambda}_2,\hat{\mu}} \sum_\omega \hat{v}_\omega q_\omega \\
\text{s.t.: } &\hat{v}_1=\max\{\hat{\lambda}_1,-\hat{\lambda}_2\}=\hat{\lambda}_1 \\
&\hat{v}_2=\max\{1-\hat{\lambda}_1+\tfrac{\hat{\mu}}{q_2},\hat{\lambda}_2\} \\
&\hat{v}_\varnothing=\max\{p-(2p-1)\hat{\lambda}_1-\tfrac{\hat{\mu}}{q_\varnothing},(2p-1)\hat{\lambda}_2\}.
\end{aligned}$$

Solution of $\hat{\mathcal{D}}$ and $\hat{\mathcal{P}}$. We begin with $\hat{\mathcal{D}}$. First, since $p>1/2$, it is optimal to set $\hat{\lambda}_2^*=0$. Next, we claim that $\hat{\mu}^*>0$. Suppose not. Then, fixing $\mu=0$, the optimal $\hat{\lambda}_1$ would equal $\frac{p}{2p-1}>1$, as in the previous section. Therefore, $1-\hat{\lambda}_1<0$. It is possible to do strictly better by increasing μ by some ε and decrease $\hat{\lambda}_1$ so that $1-\hat{\lambda}_1+\frac{\varepsilon}{q_2}<0$ and $p-(2p-1)\hat{\lambda}_1-\frac{\varepsilon}{q_\varnothing}=0$. Therefore, the optimal $\hat{\mu}$ and $\hat{\lambda}_1$ must satisfy

$$\begin{cases} -\hat{\lambda}_1+\frac{\hat{\mu}}{q_2}=0 \\ -(2p-1)\hat{\lambda}_1-\frac{\hat{\mu}}{q_\varnothing}=0. \end{cases}$$

With the solution $\hat{\lambda}_1^*$, we obtain

$$\hat{v}_1^*=\frac{pq_\varnothing+q_2}{(2p-1)q_\varnothing+q_2}$$

[4] Since a buyer' disclosure decision is binary (i.e., disclose all or nothing about her record), this formulation of the disclosure constraints is without loss of generality. In more general versions of this problem, the literature (starting from Green and Laffont, 1986) has investigated conditions under which a form of revelation principle holds by which the disclosure constraints can be formulated in a similar fashion.

[5] In fact, we can also interpret (1) as a privacy protection that grants buyers the right to request that the platform delete their data, thereby turning their record type into $\omega=\varnothing$.

and $\hat{v}_2^* = \hat{v}_\varnothing^* = 0$ (see also Table 2).

We now solve $\hat{\mathcal{P}}$. Denote the recommendation mechanism by $\hat{x}$, so as to distinguish it from the one of the previous section. As before, it is optimal to set $\hat{x}^*(1,1) = q_1$ and $\hat{x}^*(2,1) = 0$. Since $\hat{\mu}^* > 0$, type-2 buyers must be indifferent between disclosing their data or not. This requires that they be treated as type-$\varnothing$ buyers. That is, $\hat{x}^*(a|2) = \hat{x}^*(a|\varnothing)$ for all a, where $\hat{x}^*(a|\omega) = \hat{x}^*(a,\omega)/q_\omega$. To see why, note that constraint (1) requires $\hat{x}^*(1|2) \geq \hat{x}^*(1|\varnothing)$ and the aforementioned indifference requires $\hat{x}^*(1|2) = \hat{x}^*(1|\varnothing)$. Next, since $\hat{\lambda}_1^* > 0$, the seller must be indifferent between price 1 and 2 when recommended 1:

$$q_1 + \hat{x}^*(1,2) + \hat{x}^*(1,\varnothing) = 2\hat{x}^*(1,2) + 2p\hat{x}^*(1,\varnothing).$$

Using $\hat{x}^*(1,2) = q_2 \frac{\hat{x}^*(1,\varnothing)}{q_\varnothing}$ and solving, we obtain $\hat{x}^*$ in Table 2. ▲

The solutions of $\hat{\mathcal{D}}$ and $\hat{\mathcal{P}}$ (summarized in Table 2) have three noteworthy implications. First, privacy *lowers* the value of type-1 records (i.e., $\hat{v}_1^* < v_1^*$), even though type-1 buyers are not those who have to be incentivized to disclose their data. In other words, this happens even though privacy does not constrain how the platform can use type-1 records. By contrast, the value of the other records remains unchanged at zero. This interdependence between the value of some records and the use of other records is a hallmark of the fact that the platform faces a non-trivial intermediation problem (see GLP). It principle, it may also affect the market price of data records (Galperti and Perego, 2022).

Second, privacy increases the expected payoff of buyers with type-2 records, but decreases that of buyers with type-$\varnothing$ records. This is true even though the latter do not care about privacy per se. Comparing x^* and $\hat{x}^*$, it is easy to see that buyers with type-2 (resp. type-$\varnothing$) records are more (resp. less) likely to be charged price 1. That buyers with a type-2 record can benefit from privacy is intuitive, as they can conceal their records in the hope of being charged a low price. It is far less immediate that buyers with records of type $\varnothing$ are hurt: Since under privacy the platform must treat buyers with type-$\varnothing$ and type-2 records in the same way, it cannot pool as many type-$\varnothing$ records with type-1 records in the low market segment as it did under no privacy (otherwise, with the additional type-2 buyers, the seller will strictly prefer to charge price 2 for this segment). Therefore, more buyers with type-$\varnothing$ records will be offered price 2 and enjoy lower surplus. Overall, the effect on the platform's total payoff is negative (see Table 1 and 2).

Third, the quantity of trades that happen at price 2 is higher than it was without privacy. Under $\hat{x}^*$, this quantity is

$$(pq_\varnothing + q_2)\Big(1 - \frac{q_1}{(2p-1)q_\varnothing + q_2}\Big);$$

under x^*, it is

$$q_2 + p\Big(q_\varnothing - \frac{1}{2p-1}q_1\Big),$$

which is lower because $p < 1$. Intuitively, since type-2 and type-$\varnothing$ buyers have to be treated equally, the platform cannot pool as many of the latter buyers as before with type-1 buyers in the low-WTP market segment. Thus, more buyers in the 2–$\varnothing$ group are assigned to the high-WTP segment. This has two effects: on one hand, conditional on a trade happening, the seller makes more profits because it is more likely he can charge a high price; on the other, fewer trades happen because it is more likely that a low-WTP buyer whose record is of type $\varnothing$ is charged a high price. Overall, these two effects cancel out and the seller profits are the same as under no privacy (see Table 1 and 2).

IV. Further Analysis

Privacy Can Increase the Value of Data. So far, we have assumed that q satisfies $q_1 \leq (2p-1)q_\varnothing$. There are two other cases to consider. When $(2p-1)q_\varnothing < q_1 \leq (2p-1)q_\varnothing + q_2$, the uninformed seller still charges a price of 2. Thus, the platform's problem remains interesting. Following the same analysis of the previous sections, we find that the values of data records under no privacy are $v_1^* = 1$, $v_2^* = 0$, and $v_\varnothing^* = 1 - p$ (for derivations, see Appendix E in GLP). By contrast, with verifiable privacy, the values $\hat{v}^*$ are identical to those reported in Table 2.b. Specifically, the value of type-1 records *increases* (from 1 to $\hat{v}_1^* > 1$), while that of type-$\varnothing$ records decreases (from $1-p$ to 0). Finally, the last case to consider is $(2p-1)q_\varnothing + q_2 < q_1$. In this case, price 1 is strictly optimal for the seller under no information. Then, the platform never reveals anything about the buyers to the seller, rendering this case uninteresting.

Unverifiable Privacy. Personal data often involves

hard information (e.g., cookies, SSNs, addresses, etc.), which led us to model privacy protections as verifiable disclosure constraints. In some cases, however, personal information may be soft and individuals have to self-report it. In this case, a privacy protection may grant them significant flexibility in choosing what to report. An extreme way of modeling this form of privacy is through classic truthtelling constraints in mechanism design (what Bergemann and Morris (2019) call "information design with elicitation"): Each type in Ω can report to be *any* other type in Ω.[6] Let us briefly discuss how this affects our setting. For any mechanism x, all buyers will report any ω that leads to the highest probability of trading at price 1. Thus, the platform's problem becomes a linear program where x cannot depend on ω. In other words, the platform provides no information to the seller, who then always charges price 2 (since $q_1 \leq (2p-1)q_\varnothing$). This implies $v_1 = v_2 = v_\varnothing = 0$. The intuition is that, if the seller always charges 2, adding any record to the database will lead to zero additional surplus. In sum, introducing unverifiable privacy wipes out the value of all records: Type-1 records retained some value under verifiable privacy because they could separate themselves from type-2 records, which is now no longer possible.

Changing the Platform's Objective. We briefly consider other objectives of the platform. One can show that $\mathcal{P}$ and $\mathcal{D}$ (resp. $\hat{\mathcal{P}}$ and $\hat{\mathcal{D}}$) have the same solutions if $r \in [0,1/2]$. If $r > 1/2$, it is always optimal for the platform to fully inform the seller about what it knows regarding the buyers (i.e., ω). The value of each record is then equal to the payoff the platform directly obtains from using that record (see GLP, Proposition 2). As a result, privacy regulations can never have cross effects on the values of records across record types.

[6] This relates to Hidir and Vellodi (2021), who model buyer's disclosure as a cheap-talk message.

REFERENCES

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar, "Too Much Data: Prices and Inefficiencies in Data Markets," *Americal Economic Journal: Microeconomics*, 2022, *14 (4)*, 218–56.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman, "The Economics of Privacy," *Journal of Economic Literature*, June 2016, *54* (2), 442–92.

Ali, S. Nageeb, Greg Lewis, and Shoshana Vasserman, "Voluntary Disclosure and Personalized Pricing," *Review of Economic Studies*, 2022, *Forthcoming.*

Aridor, Guy, Yeon-Koo Che, and Tobias Salz, "The Effect of Privacy Regulation on the Data Industry: Empirical Evidence from GDPR," *RAND Journal of Economics*, 2022, *Forthcoming.*

Bergemann, Dirk and Stephen Morris, "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 2016, *11*, 487–522.

_ and _, "Information Design: A Unified Perspective," *Journal of Economic Literature*, March 2019, *57* (1), 44–95.

_ , Benjamin Brooks, and Stephen Morris, "The Limits of Price Discrimination," *American Economic Review*, 2015, *105 (3).*

Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim, "Privacy and personal data collection with information externalities," *Journal of Public Economics*, 2019, *173*, 113–124.

Dye, Ronald A., "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 1985, *23* (1), 123–145.

Galperti, Simone, Aleksendr Levkun, and Jacopo Perego, "The Value of Data Records," *Review of Economic Studies*, 2022, *Forthcoming.*

_ and Jacopo Perego, "Competitive Markets for Personal Data," *Working Paper*, 2022.

Green, Jerry R. and Jean-Jacques Laffont, "Partially Verifiable Information and Mechanism Design," *The Review of Economic Studies*, 1986, *53* (3), 447–456.

Hidir, S. and N. Vellodi, "Personalization, Discrimination and Information Revelation," *Journal of European Economic Association*, 2021, *119* (3), 1342–1363.

Ichihashi, Shota, "Online Privacy and Information Disclosure by Consumers," *American Economic Review*, February 2020, *110* (2), 569–95.