Optimal non-linear pricing with data-sensitive consumers

Daniel Krähmer[†] Roland Strausz[‡]

February 15, 2021

Abstract

We introduce consumers with intrinsic privacy preferences into the monopolistic non-linear pricing model. Next to classical consumers, there is a share of data-sensitive consumers who refrain from buying if their purchase reveals information about their valuation to the monopolist. When the monopolist observes consumers' privacy preferences, data-sensitive consumers obtain a pooling schedule, while classical consumers obtain the standard non-linear pricing schedule. Data-sensitive consumers with a low valuation obtain a strictly higher utility than classical consumers with the same valuation. By contrast, when privacy preferences are consumers' private information, classical consumers obtain a higher utility than data-sensitive consumers with the same valuation. Data-sensitive consumers and the monopolist are worse off when privacy preferences are private information, whereas classical consumers are better off. The results are relevant for policy measures that target the data-awareness of consumers, such as the European GDPR.

Keywords: Screening, privacy.

Preliminary version and work in progress. We thank Martin Peitz and Volker Nocke. We gratefully acknowledge financial support from the German Research Foundation (DFG) through Germany's Excellence Strategy EXC 2126/1–390838866 and CRC TR 224 (Krähmer) and CRC TR 190 (Strausz).

[†]University of Bonn. Email:daniel.kraehmer@hcm.uni-bonn.de.
[‡]Humboldt Universität Berlin. Email: roland.strausz@hu-berlin.de.

1 Introduction

As a result of privacy awareness initiatives and regulatory measures such as the European GDPR, consumers have become more and more aware that firms store and use the data that they voluntary provide.[1] While these initiatives have made consumers more data-sensitive, most consumers have however no intimate knowledge about how their data is being used.[2] In particular, consumers often do not know which information worsens their dealings with a firm (and which the consumer had better withhold), and which information improves a firm's offer (and which the consumer had better reveal). Moreover, news stories of data breaches and hackers have made consumers wary about providing information even to trustworthy organizations, as they could be hacked and abused by third parties.[3] Given the many and often unpredictable ways in which the provided data affects consumers, consumers face stark ambiguities, leading to a demand for privacy and a perceived utility loss associated with the act of revealing information.[4]

Classical economic models that analyze firms dealing with privately informed consumers abstract from the presence of such data-sensitive consumers. The increasing importance of such consumers calls for a reexamination of these models and their results. In particular, the question arises to what extent the presence of data-sensitive consumers affects the insights of these models and how firms optimally respond when facing such consumers.

In this paper, we consider how the presence of data-sensitive consumers affects a monopolist's optimal non-linear pricing policy in the seminal screening model of Mussa and Rosen (1978). To illustrate the effect of such consumers most clearly, we consider two extreme classes of consumers: Classical consumers, who have no preference for privacy at all, and data-sensitive consumers, who have a strong preference for privacy. More precisely, data-sensitive consumers in our model incur a prohibitively large loss in utility whenever they engage in an "informative purchase", i.e., a purchase from which the monopolist deduces something about their valuation. While extreme, this modeling of data-sensitivity yields a tractable setup in which we can analytically identify the key effects of consumers' privacy preferences. In particular, the model allows us to fully characterize the optimal non-linear pricing policy in the natural case where privacy concerns are private information, even though the resulting screening problem is multi-dimensional.

Our main insight is that the impact of consumers' privacy preferences depends crucially on whether privacy preferences are private information or not. If they are not, i.e., if the monopolist can distinguish between the two different classes of consumers, then the monopolist screens the classical privacy-unconcerned consumers using a standard non-linear pricing schedule. By contrast, she offers data-sensitive consumers a pooling contract consisting of a single quantity that matches the first-best quantity of the lowest valuation type for a price at which the lowest valuation type is just willing to buy this quantity. Hence, the optimal contract for classical consumers exhibits the usual feature of "no distortion at the top", whereas the optimal contract for data-sensitive consumers exhibits the opposite feature of "no distortion at

[1] EU (2018, p.105) reports from a EU-wide consumer survey that 67% of respondents said to be aware of targeted advertising, 62% were aware of personalized offers, and 44% were aware of personalized pricing by firms.

[2] In an experiment studying the ability of consumers to recognize the direction of price changes of personalised pricing, EU (2018, p.119) reports "Relatively few participants in the price discrimination scenario correctly identified whether and how price personalisation had occurred.".

[3] See `https://privacyrights.org/data-breaches` and `https://en.wikipedia.org/wiki/List_of_data_breaches` for a list of recent data breaches and the many different methods leading to such breaches.

[4] The crude tool of regularly deleting browser cookies as recommended by some privacy websites, e.g. `https://www.allaboutcookies.org/`, is one expression of such general privacy concerns.

the bottom". Moreover, a data-sensitive consumer with a relatively low valuation gets a better deal than a classical consumer with the same valuation. The intuition is that a data-sensitive consumer has a stronger bargaining position, as he credibly rejects offers which a classical consumer would accept.

Our crucial insight is that this latter result is overturned if, by contrast, consumers' privacy preferences are their private information so that the monopolist cannot distinguish the two different classes of consumers. In this case, a classical consumer gets a better deal than a data-sensitive with the same valuation. The reason is that a classical consumer can costlessly mimic, and thus obtain at least the same utility as a data-sensitive consumer. In contrast, the reverse incentive constraint, dissuading data-sensitive consumers from mimicking a classical consumer, is non-binding, because the monopolist can simply require classical consumers to provide information about their valuation when signing the contract. As a consequence, when privacy preferences are private information the monopolist faces the additional incentive constraint that she has to prevent classical consumers from mimicking data-sensitive consumers. The monopolist meets this constraint by improving her offer to classical consumers, while worsening it to data-sensitive consumers.

Hence, when privacy preferences are private information, classical consumers benefit from the presence of data-sensitive consumers at the expense of the monopolist. In this case, the monopolist's optimal contract pools classical consumers with valuations below a cutoff valuation with data-sensitive consumers who, for all valuations, are offered the same single quantity. Classical consumers with a valuation above the cutoff obtain the same quantity as with the standard non-linear schedule but at a reduced price; they are therefore strictly better off than in the absence of data-sensitive consumers. Moreover, a classical consumer with a high valuation is also strictly better off than a data-sensitive consumer with the same valuation.

Directly comparing the cases with observable and unobservable privacy preferences reveals that data-sensitive consumers gain when privacy preferences become publicly observable, whereas classical consumers lose. Unsurprisingly, the monopolist does better when privacy preferences are publicly observable.

2 Related Literature

Spurred by the rise of online markets and Big Data, the literature on privacy has in recent years become vast and extensive. Athey (2014) and Acquisti et al. (2016) provide surveys highlighting the main economic perspectives on privacy in these markets. Our specific contribution to this literature is the study of optimal pricing when some consumers have intrinsic preferences for privacy.

While already Becker (1980) makes the point that consumers may value privacy intrinsically, the economics literature that studies consumers with such intrinsic preferences is small but growing. There is an empirical literature that points to the existence of an intrinsic value of privacy and seeks to quantify it (see Tsai (2011), and more recently, Lin (2020) and Tang (2020)). On the theoretical side, Choi et al (2019) and Acemoglu et al. (2020) model consumers with intrinsic privacy concerns in order to study the role of data externalities. While these papers model privacy concerns as a fixed cost that a consumer incurs when he discloses sensitive data, in our model, this cost occurs if the seller's belief about the consumer is affected by the transaction. In this respect, our approach is related to Eilat et al. (2021) who capture privacy losses as the (Kullback-Leibler) divergence of the seller's beliefs about the consumer before and after the sale. These authors, however, model privacy restrictions as an explicit constraint that the privacy loss associated with the monopolist's selling mechanism must not exceed a certain level. In contrast, we follow the literature on intrinsic privacy preferences, where privacy concerns affect consumers' utility functions directly. This enables us to address the question how the seller optimally screens between consumers with and without privacy concerns and to analyze the resulting welfare implications. To the extent that the seller in our model has to take into account how much information her selling contract reveals to her, our paper is also related to Dworczak (2020) who considers the optimal design of an auction when bidders care about the information the auction reveals about the winner. In contrast to us, in Dworczak (2020) it is a third party's, not the seller's, beliefs that bidders care about, and the seller can explicitly design the information released by the auction.

A different strand of the literature studies privacy concerns that emerge for instrumental, rather than intrinsic, reasons in contexts in which sellers can engage in behavior-based price discrimination (see, e.g, Taylor (2004), Fudenberg and Villas-Boas (2006), Calzolari and Pavan (2006), Ishihashi (2020)). Consumers may then fear that, due to potential ratchet effects, a lack of privacy will harm them in future interactions with the same or other sellers, but they may also benefit, as more data allows a seller to better tailor the offer to consumers' needs. Linking the two strands of the literature, the intrinsic privacy preferences that we consider in this paper can be broadly seen as a reduced form for an instrumental privacy preference of an ambiguity averse consumer who adopts a worst-case scenario towards evaluating the costs and benefits of releasing information to the seller.

Finally, our paper also makes a conceptual contribution to the screening literature which is of independent interest. Formally, the presence of data-sensitive consumers renders the monopolist's non-linear pricing problem into a screening problem with a type-dependent outside option (Lewis and Sappington (1989), Jullien (2000)). The reason is that the outside option for consumers without privacy concerns is effectively given by the terms of trade offered to data-sensitive consumers as the former can mimic the latter. Thus, the novelty that emerges in our setting is that outside option for consumers without privacy concerns is endogenous because the terms offered to data-sensitive consumers are chosen by the monopolist.

3 The Model

We consider a monopolist who can produce a quantity $q \in [0, \bar{q}]$ at a cost $c(q)$, where $c(\cdot)$ exhibits $c'(\cdot), c''(\cdot) > 0$ and satisfies the Inada conditions $c(0) = c'(0) = 0$ and $\lim_{q \to \bar{q}} c(q) = \lim_{q \to \bar{q}} c'(q) = \infty$. The monopolist's profit is $p - c(q)$, when she sells a quantity q at price a p. She receives an outside option of 0 if no sale takes place.

There is a unit mass of consumers indexed by their (marginal) valuation of consumption θ which is distributed on the support $[\underline{\theta}, \overline{\theta}]$ with cdf F and pdf f, where $0 \leq \underline{\theta} < \overline{\theta}$. While θ is a consumer's private information, the monopolist only knows its distribution. We impose the regularity conditions that

$$\alpha(\theta) \equiv \theta + \frac{F(\theta)}{f(\theta)} \text{ and } \beta(\theta) \equiv \theta - \frac{1 - F(\theta)}{f(\theta)} \tag{1}$$

are well-defined, differentiable, and strictly increasing: $\alpha'(\theta), \beta'(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$.

Hence, the surplus maximizing quantity for a consumer with valuation θ equals $q^*(\theta)$, implicitly defined by $c'(q^*(\theta)) = \theta$. For future reference, we also define the schedules $q_\alpha(\theta)$ and $q_\beta(\theta)$ implicitly as the maximizers of $\alpha(\theta)q - c(q)$ and $\beta(\theta)q - c(q)$ respectively:

$$c'(q_\alpha(\theta)) = \alpha(\theta) \text{ and } c'(q_\beta(\theta)) = \max\{\beta(\theta), 0\}. \tag{2}$$

We also note that there is a unique $\theta_0 \in (\underline{\theta}, \overline{\theta})$ so that $\beta(\theta_0) = 0$ if and only if $\beta(\underline{\theta}) < 0$. Hence, only for the case $\beta(\underline{\theta}) < 0$, the maximum-operator in (2) has relevance. Given our assumptions, schedules $q_\alpha(\theta)$ and $q_\beta(\theta)$ are well-defined, increasing in θ, and exhibit

$$q_\beta(\theta) \leq q^*(\theta) \leq q_\alpha(\theta), \tag{3}$$

for all $\theta \in [\underline{\theta}, \overline{\theta}]$. That is, $q_\beta(\cdot)$ exhibits a downward distortion, whereas $q_\alpha(\cdot)$ exhibits an upward distortion.

The novelty of our framework is that we assume a share μ of consumers to be "data-sensitive" and they care about the information that their purchasing decision reveals to the monopolist, as described below. The remaining share, $1-\mu$, of consumers is "classical". We index a classical consumer by the privacy type $\rho = 0$. Such a consumer obtains a utility of

$$u^0(q, p|\theta) \equiv \theta q - p, \tag{4}$$

when he buys a quantity q at a price p and his valuation is θ.

We index consumers who are data-sensitive by the privacy type $\rho = 1$. As motivated in the introduction, we capture the preferences of such a consumer by the reduced form assumption that he incurs a utility loss, K, when the monopolist learns something about his valuation from a sale. Denoting F^p as the cdf that represents the principal's belief after a transaction has taken place, a data-sensitive consumer with valuation θ obtains the utility

$$u^1(q, p, F^p|\theta) \equiv \begin{cases} u^0(q, p|\theta) & \text{if } F^p = F \\ u^0(q, p|\theta) - K & \text{otherwise,} \end{cases} \tag{5}$$

where $K > 0$.[5]

Finally, we assume that a consumer's privacy type ρ and his valuation θ are stochastically independent. This means that the monopolist would learn nothing about a consumer's valuation if he learnt that a consumer was classical or data-sensitive.

Considering the benchmark where valuations are public information so that privacy plays no role, the surplus maximizing quantity for a consumer with valuation θ equals $q^*(\theta)$. Our objective is to identify a profit-maximizing selling mechanism and to identify the distortions relative to the benchmark, given the following timing:

t=0: Consumers observe their types $(\theta, \rho) \in \Theta \times \{0, 1\}$.

[5]Hence, our assumption that all consumers' outside options are zero implicitly means that if a consumer does not buy, he is unconcerned about whether the principal learns anything about his valuation from not buying. While this assumption is appropriate for some but not all economic environments, it guarantees that the consumer's outside option is independent of both his valuation type θ and his privacy type ρ.

t=1: Monopolist chooses a message set M and a mechanism $\gamma = (q, p) : M \to \mathbb{R}_+ \times \mathbb{R}$, specifying terms of trade $\gamma(m) = (q(m), p(m))$ contingent on a message $m \in M$.

t=2: Consumer accepts or rejects the mechanism.

t=3: Upon acceptance, a consumer chooses a message m.

t=4: Observing a message m, the monopolist updates her beliefs to $F^p(m)$.

t=5: Payoffs accrue.

Given a mechanism γ, a *strategy* $\sigma = (a, s)$ for a consumer specifies for each type (θ, ρ) an acceptance decision $a(\theta, \rho) \in \{0, 1\}$ whether to accept ($a = 1$) or reject ($a = 0$) the mechanism, together with a reporting decision, $s(\theta, \rho) \in M$, which message to send.[6] If the consumer accepts and sends a message m, the monopolist updates her beliefs from F to $F^p(m)$. We call the monopolist's collection of updated beliefs, $F^p = \{F^p(m)\}_{m\in M}$, an *assessment*. An equilibrium associated with the mechanism γ is a pair (σ, F^p) such that the strategy σ is optimal given the assessment F^p, and the assessment F^p is Bayes' consistent given the strategy σ.[7]

We say that a mechanism γ *respects (data-sensitive consumers') privacy* if the mechanism sustains an equilibrium (a, s, F^P) in which the monopolist does not learn anything from a message m that some data-sensitive consumer sends in equilibrium:

$$\text{For all } m, \theta \text{ such that } a(\theta, 1) = 1 \text{ and } s(\theta, 1) = m: \quad F^p(m) = F. \tag{6}$$

Under a mechanism that respects privacy, equilibrium payoffs do not depend on the principal's equilibrium beliefs. Hence, the economic outcome associated with a mechanism that respects privacy, is fully described by its induced pair (q, p), yielding the monopolist and consumer the payoffs $\Pi = p - c(q)$ and $u = \theta q - p$, respectively.

Therefore, we can capture the set of outcomes that are attainable by some mechanism that respects privacy by the class of *social choice functions* of the form $\varphi : \Theta \times \{0, 1\} \to \mathbb{R}_+ \times \mathbb{R}$, which assign to any consumer type (θ, ρ) an outcome $(q(\theta, \rho), p(\theta, \rho))$.[8] In particular, we say that a mechanism γ that respects privacy *implements* the social choice function $\varphi : \Theta \times \{0, 1\} \to \mathbb{R}_+ \times \mathbb{R}$, if the mechanism has an equilibrium satisfying (6) with consumer strategies (a, s) such that $\varphi(\theta, \rho) = a(\theta, \rho)\gamma(s(\theta, \rho))$ for all (θ, ρ).

Throughout, we assume that a consumer's valuation θ is his private information. We begin our analysis with the benchmark case that a consumer's privacy type ρ is publicly observable, and then move on to the case in which ρ is the consumer's private information.

[6] As is usual in applications of mechanism design, we focus on deterministic mechanisms that induce pure strategy equilibria both for tractability reasons and because stochastic mechanisms are difficult to implement in practice.

[7] It is without loss to assume that any message in M is sent with positive probability: $\forall m \in M : \{(\theta, \rho) | a(\theta, \rho) = 1 \wedge s(\theta, \rho) = m\} \neq \emptyset$. Otherwise, we could obtain an outcome-equivalent mechanism with this property by removing all zero-probability messages from the message set.

[8] As we discuss in more detail below, a more involved notion of a social choice function (one that also includes the monopolist's assessment F^p) is needed for describing economic outcomes of mechanisms that do not respect privacy

4 Publicly observable privacy types

We first study the benchmark case where a consumer's privacy type ρ is public information. In this case, the monopolist conditions her mechanism directly on ρ and treats the two consumer groups independently.

When consumers are classical ($\rho = 0$), the monopolist faces the standard non-linear pricing problem as in Mussa and Rosen (1978). Consequently, the monopolist optimally offers a direct revelation mechanism where a consumer obtains the terms of trade $(q^0(\theta), p^0(\theta))$ upon reporting valuation θ. The profit maximizing schedule $\bar{q}^0(\theta)$ equals $q_\beta(\theta)$ as defined in (2).

A consumer of type θ obtains the utility

$$\bar{u}^0(\theta) = \int_{\underline{\theta}}^{\theta} q_\beta(t)\, dt,$$

and the optimal price schedule is

$$\bar{p}^0(\theta) = \theta q_\beta(\theta) - \bar{u}^0(\theta). \tag{7}$$

If the monopolist faces a data-sensitive consumer ($\rho = 1$) she can respect privacy only by offering a mechanism where all consumer types θ who accept the mechanism end up with the same pooling outcome (q^1, p^1). To see this, note that if, in equilibrium, some data-sensitive consumers end up buying the bundle $(q^1(\hat{m}), p^1(\hat{m}))$ associated with the message $\hat{m}$, while some other data-sensitive consumers end up buying the different bundle $(q^1(\tilde{m}), p^1(\tilde{m}))$ associated with the message $\tilde{m}$, then the principal's belief $F^p(\hat{m})$ and $F^p(\tilde{m})$ differ so that they cannot both coincide with the ex ante belief F.

These considerations imply that a mechanism that respects the privacy of data-sensitive consumers simply offers a quantity q^1 at a price p^1.

Moreover, all data-sensitive consumers either all accept or reject the mechanism, because otherwise, the monopolist would infer a consumer's valuation from his acceptance decision, thus not respecting privacy. In particular, profit maximization implies that the monopolist keeps the lowest type $\underline{\theta}$ just indifferent between accepting and rejecting, i.e. $p^1 = \underline{\theta} q^1$. This price gives all other types a strict incentive to buy so that in equilibrium they accept, yielding the monopolist a profit $\underline{\theta} q^1 - c(q^1)$. Hence, the profit-maximizing quantity level, $\bar{q}^1$, satisfies the first order condition $c'(\bar{q}^1) = \underline{\theta}$ and corresponds to the first best level, $q^*(\underline{\theta})$, for type $\underline{\theta}$, characterized by $c'(q^*(\underline{\theta})) = \underline{\theta}$. The utility of a data-sensitive consumer with valuation θ is

$$\bar{u}^1(\theta) = \theta q^*(\underline{\theta}) - \underline{\theta} q^*(\underline{\theta}) = q^*(\underline{\theta})(\theta - \underline{\theta}). \tag{8}$$

Having characterized the profit-maximizing selling mechanisms for consumers with different privacy types, we next compare them. Since the Mussa-Rosen schedule $\bar{q}^0$ is strictly distorted downwards for all except for the highest type $\overline{\theta}$, it follows

$$\frac{d\bar{u}^0(\underline{\theta})}{d\theta} = q^*(\underline{\theta}) > \bar{q}^0(\underline{\theta}) = \frac{d\bar{u}^1(\underline{\theta})}{d\theta} \quad \text{for all } \theta \in [\underline{\theta}, \overline{\theta}). \tag{9}$$

Together with $\bar{u}^0(\underline{\theta}) = \bar{u}^1(\underline{\theta}) = 0$, this implies that $\bar{u}^0(\theta) > \bar{u}^1(\theta)$ for θ close to $\underline{\theta}$. Hence, with publicly observable privacy types, a data-sensitive consumer with a low valuation type is strictly better off than a

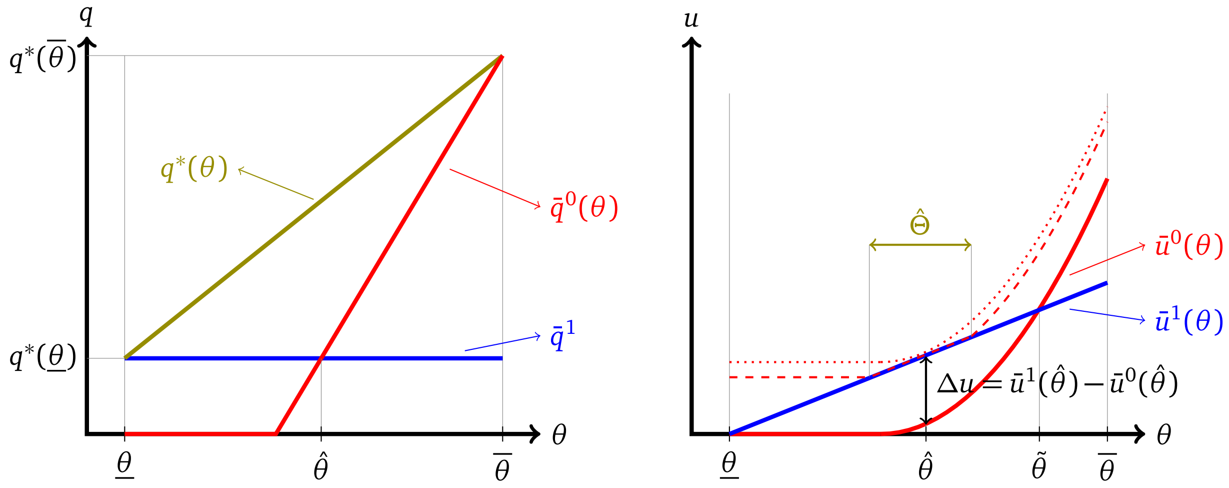

Figure 1: Optimal schedules with public privacy concerns

classical consumer of the same valuation type: The next proposition summarizes these results.

Proposition 1 *Suppose privacy types ρ are publicly observable. Then a classical consumer with valuation θ obtains the Mussa-Rosen schedule $\bar{q}^0(\theta) = q_\beta(\theta)$ at a price $\bar{p}^0(\theta)$ as given by (2) and (7). All data-sensitive consumers obtain the same quantity $\bar{q}^1 = q^*(\underline{\theta})$ at a price $\bar{p}^1 = \underline{\theta}\bar{q}^1$. Data-sensitive consumers with a low valuation are strictly better off than classical consumers with the same valuation.*

Figure 1 illustrates the proposition for the uniform-quadratic case, where valuations are uniformly distributed and the cost function is quadratic.[9] The left panel illustrates the optimal quantity schedule, while the right panel shows the consumer's payoffs. Data-sensitive consumers all receive the same quantity level $\bar{q}^1$, which equals the first best level, $q^*(\underline{\theta})$, of the consumer with the lowest valuation $\underline{\theta}$. The utility of data-sensitive consumers, $\bar{u}^1(\theta)$, increases therefore linearly in the valuation θ, as illustrated in the second panel. Classical consumers get screened through an increasing quantity schedule $\bar{q}^0(\theta)$, which is downward distorted compared to the surplus maximizing quantity $q^*(\theta)$. Low valuation consumers do not receive the good, whereas higher valuation consumers receive a quantity level that is strictly increasing in valuation. Their utility level $\bar{u}^0(\theta)$ is therefore convex in valuation, as illustrated in the second panel.

The second panel illustrates that data-sensitive consumers are strictly better off than their classical counterparts for any valuation in the interval $(\underline{\theta}, \tilde{\theta})$. Hence, any such consumer would have an incentive to claim that he is data-sensitive if he could do so. This shows that the optimal selling mechanism from Proposition 1 is no longer feasible if consumers' privacy types are their private information.

In the next section, we show how we can adapt standard tools from mechanism design to study the case when consumers' privacy types are their private information. We will show that in this case the main economic insight of Proposition 1 changes in three respects. First, data-sensitive consumers are harmed by the fact that privacy types are private information. Second, overturning the last result of Proposition 1, all data-sensitive consumers are strictly worse off than classical consumers with the same valuation. Third,

[9] We present this case as an example in Appendix B.

classical consumers are better off in the presence of data-sensitive consumers. Thus, classical consumers benefit at the expense of data-sensitive consumers when privacy types are private information.

5 Unobservable privacy types

We next turn to the case where privacy types are a consumer's private information. To provide intuition for this case, we first ask whether the monopolist is able to implement the quantity schedules of Proposition 1 through some other mechanism if she cannot observe consumers' privacy types. The dotted upward shifted curve, $\hat{u}^0(\theta)$, in the second panel in Figure 1 shows a partial answer to this question. By lowering the prices to classical consumers by the amount Δu, the monopolist can ensure that at least the classical consumers have no incentive to "misrepresent" their privacy types by picking $(\bar{q}^1, \bar{p}^1)$. As illustrated in Figure 1, the minimal price decrease Δu that achieves this satisfies $\Delta u \equiv \bar{u}^1(\hat{\theta}) - \bar{u}^0(\hat{\theta})$ with

$$\hat{\theta} \equiv \arg\max_{\theta} \{\bar{u}^1(\theta) - \bar{u}^0(\theta)\}. \tag{10}$$

If the privacy cost K is large enough, also a data-sensitive consumer has no incentive to misrepresent his type by picking an option $(q^0(\theta), p^0(\theta))$ that is intended for classical consumers. The reason is that this would induce the monopolist to believe that the consumer's valuation is θ and thus cause a loss of K to the data-sensitive consumer. Thus, this decrease in prices does indeed enable the principal to implement the quantity schedules of Proposition 1. Because under the adapted scheme, the monopolist sells the good for Δu less to a classical consumer, the adapted scheme lowers the principal's profits by $\mu\Delta u$. Hence, $\mu\Delta u$ represents an upper bound on the principal's loss from not being able to observe the privacy types.

However, next to lowering prices to classical consumers, the principal has two additional tools by which to reduce her loss. First, she can lower the quantity $\bar{q}^1$ that she offers data-sensitive consumers. Because $\bar{q}^1$ was chosen optimally, this has no first order effect on the monopolist's profits, but has a first order effect on the amount by which she has to lower prices for classical consumers so as to dissuade them from picking $(\bar{q}^1, \bar{p}^1)$. Second, the monopolist can reduce the adaptation costs by lowering prices to classical consumers by less than Δu and offering the pair $(\bar{q}^1, \bar{p}^1)$ rather than the Mussa-Rosen schedule to classical consumers with valuations θ in a small interval $\hat{\Theta}$ around $\hat{\theta}$. As illustrated by the dashed curve, this corresponds to shifting the dotted curve in Figure 1 downward for valuations θ outside of $\hat{\Theta}$ and pooling it with $\bar{u}^1$ on $\hat{\Theta}$. This is profitable, because offering the pair $(\bar{q}^1, \bar{p}^1)$ to the small set of classical consumers with valuations in $\hat{\Theta}$ has no first order effect on profits, as the dotted line is tangent to $\bar{u}^1$. But because all classical consumers with θ outside of $\hat{\Theta}$ now obtain a smaller price decrease than Δu, the monopolist makes a first order profit gain.

In order to determine the principal's profit-maximizing mechanism with data-sensitive consumers, we study the structure of mechanisms that respect privacy. Note first that the usual class of incentive compatible, direct mechanisms, by definition, does not respect the consumer's privacy. Since privacy matters only to data-sensitive consumers, a natural class of mechanisms to consider instead is one that offers an option that can be chosen without sending a message to the monopolist. We first confirm that this intuition is indeed correct.

To do so, we define a *privacy mechanism* $\gamma = ((q^0(\theta), p^0(\theta)), (q^1, p^1))$ as a combination of a direct mechanism

$$(q^0, p^0) : \Theta \to \mathbb{R}_+ \times \mathbb{R}, \tag{11}$$

that asks consumers to report a valuation, and a single "privacy option"

$$(q^1, p^1) \in \mathbb{R}_+ \times \mathbb{R}, \tag{12}$$

that a consumer can pick without sending a message. The next lemma establishes a revelation principle for mechanisms that respect privacy and says that any mechanism that respects privacy can be replicated by a privacy mechanism in which all classical consumers reveal their type truthfully and all data-sensitive consumers pick the privacy option.

Lemma 1 *Let ρ be consumers' private information, and let γ be a mechanism that respects privacy and implements the social choice function φ. Then there is a privacy mechanism $\tilde{\gamma} = ((q^0(\theta), p^0(\theta)), (q^1, p^1))$ that also implements φ and respects privacy. Moreover, it induces an equilibrium in which (i) all consumer types accept the mechanism, and (ii) all classical consumers truthfully reveal their valuation and (almost) all data-sensitive consumers pick the privacy option (q^1, p^1).*

Property (ii) of Lemma 1 shows that we can restrict attention to privacy mechanisms that satisfy the incentive compatibility conditions

$$\theta q^0(\theta) - p^0(\theta) \geq \theta q^0(\theta') - p^0(\theta') \tag{IC_θ}$$

$$\theta q^0(\theta) - p^0(\theta) \geq \theta q^1 - p^1 \tag{IC^0}$$

$$\theta q^1 - p^1 \geq \theta q^0(\theta') - p^0(\theta') - K \tag{IC^1}$$

for all $\theta, \theta' \in \Theta$.

Here, (IC_θ) and (IC^0) ensure that a classical consumer does not misrepresent his valuation and does not pretend to be data-sensitive and pick the privacy option, while (IC^1) ensures that a data-sensitive consumer does not pretend to be classical by announcing a valuation θ'. In the latter case, the monopolist would believe the consumer's valuation to be θ' so that the (data-sensitive) consumer would suffer the privacy cost K.

Moreover, part (i) of Lemma 1 implies that we can focus on mechanisms that satisfy the individual rationality conditions

$$\underline{\theta} q^1 - p^1 \geq 0. \tag{IR^1}$$

$$\theta q^0(\theta) - p^0(\theta) \geq 0. \tag{IR^0}$$

for all $\theta \in \Theta$.

Note that (IR^1) only imposes individual rationality for a data-sensitive consumer with the lowest valuation, because this is equivalent to the requirement that the mechanism is individually rational for a data-sensitive consumer with any valuation θ. Also note that (IC^0) and (IR^1) imply (IR^0) so that we can disregard it.

Hence, we call a privacy mechanism γ *feasible* if it satisfies (IC_θ), (IC^0), (IC^1), and (IR^1).

Our motivation for focusing on mechanisms that respect privacy is to capture situations in which data-sensitive consumers suffer a large cost K from a privacy violation. Note that a large K ensures that (IC^1) is satisfied. We will therefore disregard (IC^1) and later state a lower bound on the size of K such that (IC^1) is indeed satisfied.

Disregarding (IC^1), all constraints except for (IR^1) are relaxed when increasing p^1. Because raising p^1 also increases the principal's objective, the constraint (IR^1) is binding at an optimum, implying $p^1 = \underline{\theta} q^1$.

Together, these considerations imply that the monopolist's problem boils down to[10]

$$\mathcal{Q}: \qquad \max_{q^0(\cdot),p^0(\cdot),q^1} W = (1-\mu)\left\{\int_{\underline{\theta}}^{\overline{\theta}} p^0(\theta) - c(q^0(\theta))\, dF(\theta)\right\} + \mu\{\underline{\theta}q^1 - c(q^1)\} \qquad s.t. \tag{13}$$

$$\theta q^0(\theta) - p^0(\theta) \geq \theta q^0(\theta') - p^0(\theta') \qquad \forall \theta, \theta' \in \Theta. \tag{IC_θ}$$

$$\theta q^0(\theta) - p^0(\theta) \geq (\theta - \underline{\theta}) q^1 \qquad \forall \theta \in \Theta. \tag{IC^0}$$

Note that (IC^0) requires a classical consumer to obtain at least the utility $(\theta - \underline{\theta})q^1$ from the mechanism. Effectively, it is as if a classical consumer has a type-dependent outside option of $(\theta - \underline{\theta})q^1$. In contrast to standard screening problems with type-dependent outside options, the novelty here is that the value of the outside option is endogenous, as it depends, in addition, on the monopolist's choice of q^1.

To solve the monopolist's problem, we first re-write the constraints. Note that (IC^0) must hold with equality for some θ at an optimum. The reason is that otherwise the monopolist could uniformly increase the price $p^0(\cdot)$ by a small amount. This would maintain (IC_θ) but increase profits.

Further, let $u^0(\theta) = \theta q^0(\theta) - p^0(\theta)$ denote a classical consumer's indirect utility function. As is well-known from standard screening problems, (IC_θ) is then satisfied if and only if $q^0(\cdot)$ is increasing and $du^0(\theta)/d\theta = q^0(\theta)$. In particular, the indirect utility function $u^0(\cdot)$ is convex. Because the function on the right hand side of (IC^0) is linear, convexity of $u^0(\cdot)$ implies that the set of valuations where (IC^0) holds with equality is an interval $[\theta_a, \theta_b]$, that is, $u^0(\theta) = (\theta - \underline{\theta})q^1$ if and only if $\theta \in [\theta_a, \theta_b]$. Together with the fact that $du^0(\theta)/d\theta = q^0(\theta)$, this also means that and $q^0(\theta) = q^1$ if and only if $\theta \in [\theta_a, \theta_b]$ The next lemma summarizes these considerations.

Lemma 2 *Problem $\mathcal{Q}$ is equivalent to the problem where (IC_θ) and (IC^0) are replaced by the following constraints: There are $\theta_a, \theta_b \in \Theta$ with $\theta_a \leq \theta_b$ so that*
(i) $q^0(\cdot)$ is increasing, and $q^0(\theta) = q^1$ if and only if $\theta \in [\theta_a, \theta_b]$.
(ii) $u^0(\cdot)$ is convex with $du^0(\theta)/d\theta = q^0(\theta)$ for almost all θ, and $u^0(\theta) = (\theta - \underline{\theta})q^1$ if and only if $\theta \in [\theta_a, \theta_b]$.

Lemma 2 confirms our graphical intuition as expressed by the dashed curve in the second panel of Figure 1 that it is optimal to pool classical and data-sensitive consumers in some interval $[\theta_a, \theta_b]$.

Following the literature on screening with type-dependent outside options, we use part (ii) of Lemma 2 to recover $u^0(\cdot)$ from integration as follows:

$$u^0(\theta) = u^0(\theta_a) - \int_{\theta}^{\theta_a} q^0(t)\, dt \quad \text{if } \theta \in [\underline{\theta}, \theta_a), \quad u^0(\theta) = u^0(\theta_b) + \int_{\theta_b}^{\theta} q^0(t)\, dt \quad \text{if } \theta \in (\theta_b, \overline{\theta}]. \tag{14}$$

[10]Note that because $p^1 = \underline{\theta}q^1$ at an optimum, the monopolist's profit from a data-sensitive consumer is $\underline{\theta}q^1 - c(q^1)$. Similarly, the right hand side of (IC^0) becomes $p^1 - \theta q^1 = (\theta - \underline{\theta})q^1$.

As shown in Appendix A, we can use these expression to re-write the monopolist's problem in terms of maximizing the virtual surplus subject to the monotonicity constraints from part (i) of Lemma 2:

$$\mathscr{P}: \max_{q^0(\cdot),\theta_a,\theta_b,q^1} (1-\mu)\left\{\int_{\underline{\theta}}^{\theta_a} \alpha(\theta)q^0(\theta)-c(q^0(\theta))-(\theta_a-\underline{\theta})q^1\,dF(\theta)\right. \tag{15}$$

$$+[F(\theta_b)-F(\theta_a)](\underline{\theta}q^1-c(q^1))$$

$$\left.+\int_{\theta_b}^{\overline{\theta}} \beta(\theta)q^0(\theta)-c(q^0(\theta))-(\theta_b-\underline{\theta})q^1\,dF(\theta)\right\}$$

$$+\mu\{\underline{\theta}q^1-c(q^1)\} \qquad s.t.$$

$$q^0(\cdot)\text{ increasing;}\quad q^0(\theta)=q^1\text{ if and only if }\theta\in[\theta_a,\theta_b]. \tag{MON}$$

The first line of the virtual surplus corresponds to the total surplus generated by classical consumers with valuations in $[\underline{\theta},\theta_a)$ minus the information rent the monopolist has to leave to these consumers for not mimicking a marginally higher valuation. The second line corresponds to the total surplus generated by classical consumers in $[\theta_a,\theta_b]$ minus the information rent the monopolist has to leave to these consumers for not mimicking a data-sensitive consumer and the third line corresponds to the total surplus generated by classical consumers with valuations in $(\theta_a,\overline{\theta}]$ minus the information rent the monopolist has to leave to these consumers for not mimicking a marginally lower valuation. Finally, the fourth line corresponds to the profit the monopolist obtains from data-sensitive consumers.

The next lemma derives the optimal quantity schedule $\hat{q}^0(\theta)$ offered to classical consumers for a given q^1. To characterize this schedule, we first define for any q^1 the two cutoffs θ_α and θ_β as the valuations where $q_\alpha(\cdot)$ and $q_\beta(\cdot)$ intersect with q^1. To cover cases where this intersection does not exist, we formally define

$$\theta_\alpha(q^1)=\arg\min_{\theta\in[\underline{\theta},\overline{\theta}]}|q_\alpha(\theta)-q^1|,\quad \theta_\beta(q^1)=\arg\min_{\theta\in[\underline{\theta},\overline{\theta}]}|q_\beta(\theta)-q^1|. \tag{16}$$

Note that for $q^1=0$, and $\beta(\underline{\theta})<0$, we have that $\arg\min_{\theta\in[\underline{\theta},\overline{\theta}]}|q_\beta(\theta)-q^1|=[\underline{\theta},\theta_0]$. For this case, we define $\theta_\beta(q^1)=\theta_0$.

We now show that at an optimum, marginal costs equal marginal virtual valuations outside of the interval $[\theta_\alpha,\theta_\beta]$ where (IC^0) is not binding. In particular, for valuations larger than θ_b, the schedule corresponds to the Mussa-Rosen schedule.

Lemma 3 *A solution* $\{\hat{q}^0(\cdot),\hat{\theta}_a,\hat{\theta}_b,\hat{q}^1\}$ *to* $\mathscr{P}$ *exhibits* $\hat{\theta}_a=\theta_\alpha(\hat{q}^1)$, $\hat{\theta}_b=\theta_\beta(\hat{q}^1)$ *and*

$$\hat{q}^0(\theta)=\begin{cases} q_\alpha(\theta) & if \quad \theta\in[\underline{\theta},\theta_\alpha) \\ \hat{q}^1 & if \quad \theta\in[\theta_\alpha,\theta_\beta] \\ q_\beta(\theta) & if \quad \theta\in(\theta_\beta,\overline{\theta}]. \end{cases} \tag{17}$$

Lemma 3 characterizes a solution to $\mathscr{P}$ up to q^1. Therefore, if we insert $\hat{q}^0(\cdot)$ into the objective, the monopolist's problem boils down to the choice of q^1. We next show that the optimal q^1 lies below the efficient level of the lowest value type, implying that $\hat{\theta}_a = \theta_\alpha = \underline{\theta}$.

Lemma 4 *A solution* $\{\hat{q}^0(\cdot), \hat{\theta}_a, \hat{\theta}_b, \hat{q}^1\}$ *to* $\mathscr{P}$ *exhibits* $\hat{q}^1 \le q^*(\underline{\theta}) = \bar{q}^1$. *In particular,* $\hat{\theta}_a = \underline{\theta}$.

The lemma states that when privacy types are unobservable, data-sensitive consumers end up with a lower quantity than when the monopolist can distinguish them from classical consumers. Moreover, recalling that the optimal price for a data-sensitive consumer is $p^1 = \underline{\theta}\hat{q}^1$, the lemma implies that when privacy types are unobservable, then data-sensitive consumers are worse off compared to when the monopolist cannot recognize them.

Note that, together with part (ii) of Lemma 2, Lemma 4 confirms the intuition from the beginning of the section that the monopolist optimally uses two tools to reduce the cost of maintaining incentive compatibility: reducing the quantity q^1 offered to data-sensitive consumers (Lemma 4) and pooling some classical consumers with data-sensitive consumers (Lemma 2). The second part of Lemma 4 that $\hat{\theta}_a = \underline{\theta}$, means that the monopolist optimally uses the second tool in an extreme sense in that she pools all classical consumers with a low valuation with data-sensitive consumers.

It remains to characterize the optimal quantity $\hat{q}^1$ consumed by data-sensitive consumers and the resulting highest valuation $\hat{\theta}_b$ of the pooling interval. In order to take care of possible corner solutions, we define the auxiliary function $\phi : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ by

$$\phi(\theta) \equiv \underline{\theta} - \beta(\theta) - (1-\mu)\frac{(1-F(\theta))^2}{f(\theta)}.$$

Recall that for $\beta(\underline{\theta}) < 0$ there is a unique θ_0 so that $\beta(\theta_0) = 0$. The next lemma completes the characterization of the profit-maximizing mechanism when privacy types are private information.

Lemma 5 *Let* $\{\hat{q}^0(\cdot), \hat{\theta}_a, \hat{\theta}_b, \hat{q}^1\}$ *be a solution to* $\mathscr{P}$.
If $\beta(\underline{\theta}) < 0$ *and* $\phi(\theta_0) \le 0$, *then* $\hat{\theta}_b = \theta_0$ *and* $\hat{q}^1 = 0$;
Otherwise, $\hat{\theta}_b$ *is given as the unique solution to* $\phi(\hat{\theta}_b) = 0$, *and* $\hat{q}^1 = q_\beta(\hat{\theta}_b)$, *with* $\hat{q}^1 > 0$.

Taken together, the previous lemmata fully characterize the monopolist's profit-maximizing privacy mechanism $\hat{\gamma} = \{\hat{q}^0(\cdot), \hat{\theta}_a, \hat{\theta}_b, \hat{q}^1\}$ under the assumption that the incentive constraint of data-sensitive consumers, (IC^1), is not binding. Using this characterization, we obtain a lower bound on the privacy costs K such that the characterized solution does indeed satisfy the neglected incentive constraint (IC^1). Let

$$\bar{K} \equiv \hat{u}^0(\overline{\theta}) - \hat{u}^1(\overline{\theta}) = \int_{\theta_\beta}^{1} \hat{q}^0(\theta) - \hat{q}^1 \, d\theta, \tag{18}$$

where $\hat{u}^0(\cdot)$ and $\hat{u}^1(\cdot)$ are the utilities of a classical and data-sensitive consumers under $\hat{\gamma}$. It follows that, for $K \ge \bar{K}$, the mechanism $\hat{\gamma}$ satisfies the neglected constraint (IC^1) so that we obtain the following proposition.

Proposition 2 *Suppose privacy types are private information and* $K \ge \bar{K}$. *Then the profit-maximizing selling mechanism is characterized by Lemmata 3, 4, and 5. In particular, we have:*

1. *Classical consumers with a valuation* $\theta \ge \hat{\theta}_b$ *obtain the Mussa-Rosen quantity* $\hat{q}^0(\theta) = \bar{q}^0(\theta)$ *at a price below the Mussa-Rosen price:* $\hat{p}^0(\theta) \le \bar{p}(\theta)$.

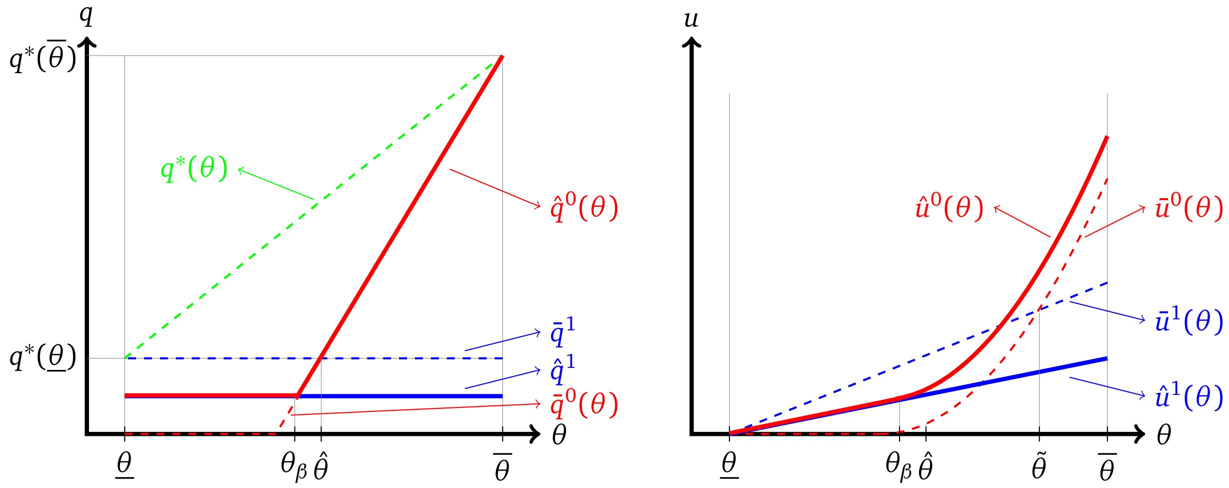

Figure 2: Optimal schedules with private vs. public privacy concerns

2. *For any valuation θ, classical consumers obtain a larger utility than data-sensitive consumers:* $\hat{u}^0(\theta) \geq \hat{u}^1(\theta)$.

3. *Compared to when privacy types are public, i) Classical consumers obtain a higher utility:* $\hat{u}^0(\theta) \geq \bar{u}^0(\theta)$*; ii) Data-sensitive consumers obtain a lower utility:* $\hat{u}^1(\theta) \leq \bar{u}^1(\theta)$.

4. *With unobservable privacy types full exclusion of data-sensitive consumers might be optimal, that is,* $\hat{q}^1 = 0$*. In particular, this exclusion is optimal if and only if* $\beta(\underline{\theta}) < 0$ *and* $\phi(\theta_0) \leq 0$.

Figure 2 illustrates the proposition by depicting the optimal schedules and the associated utilities for the cases where privacy types are unobservable (dashed curves) and observable (solid curves). In line with the second statement of the proposition, it shows that when privacy types are unobservable our earlier result from the case with observable privacy types is overturned: when privacy types are unobservable a consumer is better off if she is not data-sensitive.

Figure 2, moreover, illustrates the third result of the proposition that, except for the type $\underline{\theta}$, classical consumers strictly benefit from the private information at the expense of data-sensitive consumers, who all obtain a strictly lower utility when the principal cannot distinguish them from classical consumers.

This insight has policy relevance if one interprets privacy policies initiatives such as the European GDPR as enabling consumers to keep their privacy attitudes private. In particular, it implies that classical consumers benefit from such privacy policies, whereas data-sensitive consumers lose. The intuition behind this result is that the consumers' mimicking abilities are asymmetric: The monopolist can costlessly prevent data-sensitive consumers from mimicking classical consumers, whereas it is costly for her to prevent classical consumers from mimicking data-sensitive consumers. Hence, when privacy types are unobservable, the monopolist has to pay extra information rents to classical consumers, and these extra rents explain why classical consumers are better off when privacy types are unobservable. In order to reduce these rents, the monopolist makes the offer to data-sensitive consumers less attractive, and this explains why data-sensitive lose out when privacy types are unobservable. The insight also implies that, in general, we cannot Pareto-rank the outcomes of such policies among consumers.

The final result of Proposition 2 shows that, as indicated at the beginning of this section, the monopolist may use her first tool by which to reduce information rents to classical consumers in an extreme sense; lowering the quantity q^1 all the way to zero, thereby effectively excluding data-sensitive consumers from consumption. This allows her to offer the classical consumers the Mussa-Rosen schedule without adapting prices. While the proposition states the exact conditions under which this occurs, intuitively exclusion is optimal when the share of data-sensitive consumers and the lowest valuation type, $\underline{\theta}$, are relatively small. In this case, the monopolist does not lose much from excluding data-sensitive consumers and optimally serves only classical consumers. This also illustrates in a stark way why raising data-awareness of consumers lowers their utility. Rather than reducing their consumption, they stop buying all together. Note that if excluding data-sensitive consumers is profit-maximizing, the outcome with observable privacy types Pareto dominates the outcome with unobservable privacy types. In this case, all parties lose from policy measures that prevent the monopolist from identifying and conditioning its offer on the data-sensitivity of consumers. This result is similar to the well-known result in the literature that prohibiting (3rd degree) price discrimination yields a Pareto inferior outcome when it leads to the exclusion of certain consumer groups.

For the case, where full exclusion of data-sensitive consumers is too costly for the monopolist, the next lemma derives various comparative statics results with respect to μ, the share of data-sensitive consumers. To the extent that μ is likely to differ across markets, this gives rise to testable predictions of our model. Moreover, as pointed out in the introduction, we may interpret μ also as a policy variable to the extent that governmental data-awareness initiatives can be seen as raising the share of data-sensitive consumers. In what follows, we denote the monopolist's profits under the optimal contract by $\hat{W}$.

Lemma 6 (Comparative statics) *As the share of data-sensitive consumers increases, we have:*
(i) The monopolist's profit decreases: $d\hat{W}/d\mu \leq 0$*;*
(ii) The quantity offered to classical consumers increases: $d\hat{\theta}_b/d\mu \geq 0$*;* $d\hat{q}^1/d\mu \geq 0$*;* $d\hat{q}^0(\theta)/d\mu \geq 0$*;*
(iii) Consumer utility for given privacy types increases: $d\hat{u}^0(\theta)/d\mu, d\hat{u}^1(\theta)/d\mu \geq 0$*;*

The first result states that the monopolist's profits fall when the share of data-sensitive consumers rises. The second result implies that a higher share of data-sensitive consumers translates into more classical consumers being pooled with data-sensitive ones and a less distorted quantity $\hat{q}^1$. The third result states that all consumers, whose privacy type remains unchanged, benefit from an increase in the share of data-sensitive consumers. This holds for both data-sensitive and classical consumers. The result does, however, not imply that a raise in μ benefits all consumers, since a raise in μ means that some classical consumers are turned into data-sensitive consumers. Indeed, the second statement in Proposition 1 implies that classical consumers with a valuation exceeding θ_β lose when they "switch" from being a classical to a data-sensitive type. Hence, the third result of the lemma implies that all consumers benefit from an increase in the share of data-sensitive consumers, except for classical high valuation types who become data-sensitive. This last insight illustrates the somewhat counter-intuitive effect that raising the data-awareness of consumers may actually lower their utility.

To conclude this section, we finally compare the aggregate surplus generated under the optimal contract $\bar{\gamma}$ with observable privacy types and the optimal contract $\hat{\gamma}$ with unobservable privacy types. According to Proposition 1, classical consumers with a valuation exceeding θ_β obtain the same quantities under $\bar{\gamma}$ and

$\hat{\gamma}$. Hence, there is no difference in the surplus generated by these consumer types. On the other hand, for classical consumers with valuations below θ_β, the contract $\hat{\gamma}$ yields a higher surplus than the contract $\bar{\gamma}$, because $\hat{q}^0(\theta)$ is less (downward) distorted than $\bar{q}^0(\theta)$. Finally, with respect to data-sensitive consumers, $\bar{\gamma}$ generates a higher surplus, since $\bar{q}^1$ is less (downward) distorted than $\hat{q}^1$. Hence, in general the comparison of aggregate surplus when privacy types are observable with when they are unobservable is ambiguous.

6 Conclusion

A natural extension of our analysis is to study mechanisms that do not respect the privacy of data-sensitive consumers. Indeed, such mechanisms will be optimal when the privacy cost K that data-sensitive consumer incur is relatively small. We leave such an analysis for future research, since it involves multiple complications. First, a study of mechanisms that intrude on the privacy of data-sensitive consumers requires a more extended concept of a social choice function. In particular, social choice functions have to specify also the principal's belief function $F^p(m)$, since it affects the equilibrium payoffs of some data-sensitive consumers explicitly. Hence, focusing on mechanisms that respect the privacy of data-sensitive consumers allows us to disregard beliefs in the social choice function. Second and as a result of the more complicated notion of a social choice function, the analysis of profit-maximizing mechanisms that intrude on the privacy of data-sensitive consumers turns out to be less tractable and solving it requires the development of new analytical techniques. Finally, when mechanisms violate privacy the use of allocative efficiency as a proper welfare criterium is put into question, because it disregards any welfare losses of those data-sensitive consumers who incur the privacy cost K.

Appendix A: Proofs

Proof of Lemma 1: Consider some mechanism $\gamma = (q, p) : M \to \mathbb{R}_+ \times \mathbb{R}$ inducing an equilibrium with consumers' strategies $a(\theta, \rho)$ and $s(\theta, \rho)$, the monopolist's belief $F^p(m) = F$ for all $m \in M$, and implementing φ. Let $\Theta^r = \{\theta | a(\theta, 1) = 0\}$ represent the data-sensitive consumers who reject the mechanism, and $\Theta^a = \{\theta | a(\theta, 1) = 1\} = \Theta \backslash \Theta^r$ represent the data-sensitive consumers who accept the mechanism. We distinguish between three cases: i) (almost) all data-sensitive consumers reject; Θ^r has measure 1, ii) (almost) all data-sensitive consumer accept; Θ^a has measure 1, iii) some data-sensitive consumers accept and some reject; both Θ^a and Θ^r have positive measure.

Case i): If $a(\theta, 1) = 0$ for (almost) all θ, then we can replace the mechanism γ with a privacy mechanism $\tilde{\gamma} = ((\tilde{q}^0(\theta), \tilde{p}^0(\theta)), (\tilde{q}^1, \tilde{p}^1))$ with $\tilde{q}^0(\theta) = a^0(s(\theta, 0))q^0(s(\theta, 0))$, $\tilde{p}^0(\theta) = a^0(s(\theta, 0))p^0(s(\theta, 0))$ and $\tilde{q}^1 = \tilde{p}^1 = 0$. By construction, $\tilde{\gamma}$ respects consumer privacy, induces an equilibrium in which all consumer types accept the mechanism, all classical consumers truthfully reveal their valuation and (almost) all data-sensitive consumers pick the privacy option (q^1, p^1), and also implements φ.

Case ii): Let $a(\theta, 1) = 1$ for (almost) all θ. We first show that $\gamma(s(\hat{\theta}, 1)) = \gamma(s(\tilde{\theta}, 1))$ for (almost) all $\hat{\theta}, \tilde{\theta} \in \Theta$. Indeed, suppose this is not the case, implying that there are sets $\hat{\Theta}$ and $\tilde{\Theta}$, both with positive mass, such that

$$s(\hat{\theta}, 1) = \hat{m} \neq \tilde{m} = s(\tilde{\theta}, 1) \quad \forall \hat{\theta} \in \hat{\Theta}, \tilde{\theta} \in \tilde{\Theta}. \tag{19}$$

Since γ respects privacy, we also have that $F^p(\hat{m}) = F^p(\tilde{m}) = F$, but this can only be the case if (almost) all classical consumers with valuation not in $\hat{\Theta}$ send message $\hat{m}$ and (almost) all classical consumers with a valuation not in $\tilde{\Theta}$ send message $\tilde{m}$. In particular:

$$s(\hat{\theta}, 0) = \tilde{m}, \quad and \quad \hat{m} = s(\tilde{\theta}, 0) \quad \text{for (almost) all } \hat{\theta} \in \hat{\Theta}, \tilde{\theta} \in \tilde{\Theta}. \tag{20}$$

Recall however that under a mechanism that respects privacy, data-sensitive and classical consumers with the same valuation receive the same utility from a message. Therefore, (19) and (20) imply that all consumers with valuations in $\hat{\Theta} \cup \tilde{\Theta}$ are indifferent between $\hat{m}$ and $\tilde{m}$ and thus between $(q(\hat{m}), p(\hat{m}))$ and $(q(\tilde{m}), p(\tilde{m}))$. However, due to single crossing, consumers with different valuations cannot be indifferent between two different outcomes, and it follows that $(q(\hat{m}), p(\hat{m})) = (q(\tilde{m}), p(\tilde{m}))$, a contradiction to the assumption that $\gamma(s(\hat{\theta}, 1)) \neq \gamma(s(\tilde{\theta}, 1))$.

Consequently, there is $(\hat{q}, \hat{p})$ so that $\gamma(s(\theta, 1)) = (\hat{q}, \hat{p})$ for (almost) all $\theta \in \Theta$. Thus, we can replace γ with the privacy mechanism $\tilde{\gamma} = ((\tilde{q}^0(\theta), \tilde{p}^0(\theta)), (\tilde{q}^1, \tilde{p}^1))$ with $\tilde{q}^0(\theta) = a^0(s(\theta, 0))q^0(s(\theta, 0))$, $\tilde{p}^0(\theta) = a^0(s(\theta, 0))p^0(s(\theta, 0))$ and $(\tilde{q}^1, \tilde{p}^1) = (\hat{q}, \hat{p})$. By construction, $\tilde{\gamma}$ respects consumer privacy, induces an equilibrium in which all consumer types accept the mechanism, all classical consumers truthfully reveal their valuation and (almost) all data-sensitive consumers pick the privacy option (q^1, p^1), and also implements φ.

Case iii): Suppose some data-sensitive consumers accept and some reject; both Θ^a and Θ^r have positive measure. Consider a data-sensitive consumer with valuation $\theta^a \in \Theta^a$, who sends the message $m^a = s(\theta^a, 1)$ and obtains $(q(m^a), p(m^a))$. Since γ respects privacy, it holds $F^p(m^a) = F$, which can only be the case if (almost) all classical consumers with valuation $\theta^r \in \Theta^r$ also send the message m^a. Hence, (almost) all classical consumers with values in Θ^r (weakly) prefer $(q(m^a), p(m^a))$ to rejecting the mechanism, while all data-sensitive consumers with values in Θ^r, (weakly) prefer rejecting the mechanism to $(q(m^a), p(m^a))$.

Because data-sensitive and classical consumers with the same valuation have the same reservation utility 0, and under a mechanism that respects privacy, they receive the same utility from a message, it follows that (almost) all classical consumers with valuations in Θ^r must be indifferent between rejecting the mechanism and $(q(m^a), p(m^a))$. This is, however, only the case if $(q(m^a), p(m^a)) = (0,0)$, implying that for any $\theta^a \in \Theta^a$, we have $(q(s(\theta^a, 1)), p(s(\theta^a, 1)) = (0,0)$. Hence, any data-sensitive consumer obtains the outcome $(q, p) = (0,0)$ either because he rejects the mechanism or accepts and submits $s(\theta^a, 1)$. As a result, we can replace γ with the privacy mechanism $\tilde{\gamma} = ((\tilde{q}^0(\theta), \tilde{p}^0(\theta)), (\tilde{q}^1, \tilde{p}^1))$ with $\tilde{q}^0(\theta) = a^0(s(\theta, 0))q^0(s(\theta, 0))$, $\tilde{p}^0(\theta) = a^0(s(\theta, 0))p^0(s(\theta, 0))$ and $\tilde{q}^1 = \tilde{p}^1 = 0$. By construction, $\tilde{\gamma}$ respects consumer privacy, induces an equilibrium in which all consumer types accept the mechanism, all classical consumers truthfully reveal their valuation and (almost) all data-sensitive consumers pick the privacy option (q^1, p^1), and also implements φ. □

Proof of Lemma 2: The proof is in the main text. □

Rewriting of the principal's maximization problem to obtain $\mathscr{P}$:
By definition, $u^0(\theta) = \theta q^0(\theta) - p^0(\theta)$, so that we have $p^0(\theta) - c(q^0(\theta)) = \theta q^0(\theta) - c(q^0(\theta)) - u^0(\theta)$ and can rewrite the monopolist's objective (13) of problem $\mathscr{Q}$ as:

$$W = (1-\mu)\int_{\underline{\theta}}^{\overline{\theta}} \theta q^0(\theta) - c(q^0(\theta)) - u^0(\theta)\, dF(\theta) + \mu(\underline{\theta} q^1 - c(q^1)). \tag{21}$$

We calculate the integral by dividing the interval $[\underline{\theta}, \overline{\theta}]$ into three subintervals $[\underline{\theta}, \theta_a]$, $[\theta_a, \theta_b]$, and $[\theta_b, \overline{\theta}]$. By (14),

$$\int_{\underline{\theta}}^{\theta_a} u^0(\theta)\, dF(\theta) = F(\theta_a)u^0(\theta_a) - \int_{\underline{\theta}}^{\theta_a}\int_{\theta}^{\theta_a} q^0(t)\, dt\, dF(\theta) \tag{22}$$

$$= F(\theta_a)(\theta_a - \underline{\theta})q^1 - \int_{\underline{\theta}}^{\theta_a} \frac{F(\theta)}{f(\theta)} q^0(\theta)\, dF(\theta), \tag{23}$$

where the second line follows from integration by parts and since $u^0(\theta_a) = (\theta_a - \underline{\theta})q^1$ by part (ii) of Lemma 2. Thus,

$$\int_{\underline{\theta}}^{\theta_a} \theta q^0(\theta) - c(q^0(\theta)) - u^0(\theta)\, dF(\theta) = \int_{\underline{\theta}}^{\theta_a} \alpha(\theta) q^0(\theta) - c(q^0(\theta)) - (\theta_a - \underline{\theta})q^1\, dF(\theta). \tag{24}$$

With analogous steps, we obtain:

$$\int_{\theta_b}^{\overline{\theta}} \theta q^0(\theta) - c(q^0(\theta)) - u^0(\theta)\, dF(\theta) = \int_{\theta_b}^{\overline{\theta}} \beta(\theta) q^0(\theta) - c(q^0(\theta)) - (\theta_b - \underline{\theta})q^1\, dF(\theta). \tag{25}$$

Moreover, because $u^0(\theta) = (\theta - \underline{\theta})q^1$ for all $\theta \in [\theta_a, \theta_b]$ by part (ii) of Lemma 2:

$$\int_{\theta_a}^{\theta_b} \theta q^0(\theta) - c(q^0(\theta)) - u^0(\theta)\, dF(\theta) = \int_{\theta_a}^{\theta_b} \underline{\theta} q^1 - c(q^1)\, dF(\theta) = (F(\theta_b) - F(\theta_a))[\underline{\theta} q^1 - c(q^1)]. \quad (26)$$

Inserting these expressions in W and re-arranging terms yields the objective (15) of problem $\mathscr{P}$.

Finally, note that the constraints of problem $\mathscr{P}$ reduce to the constraints of part (i) of Lemma 2, because we have inserted the constraints from part (ii) into the objective. □

Proof of Lemma 3: With straightforward algebra, we can rewrite the objective (15) as

$$\begin{aligned} W = (1-\mu)\Bigg\{ &\int_{\underline{\theta}}^{\theta_a} \alpha(\theta) q^0(\theta) - c(q^0(\theta)) - [\theta_a q^1 - c(q^1)]\, dF(\theta) \\ &+ \int_{\theta_b}^{\overline{\theta}} \beta(\theta) q^0(\theta) - c(q^0(\theta)) - [\theta_b q^1 - c(q^1)]\, dF(\theta) \Bigg\} \\ &+ \underline{\theta} q^1 - c(q^1). \end{aligned} \quad (27)$$

By integration by parts,

$$\begin{aligned} &\int_{\underline{\theta}}^{\theta_a} \theta_a\, dF(\theta) = F(\theta_a)\theta_a = \int_{\underline{\theta}}^{\theta_a} \theta\, dF(\theta) + \int_{\underline{\theta}}^{\theta_a} F(\theta)\, d\theta = \int_{\underline{\theta}}^{\theta_a} \alpha(\theta)\, dF(\theta), \\ &\int_{\theta_b}^{\overline{\theta}} \theta_b\, dF(\theta) = (1 - F(\theta_b))\theta_b = \int_{\theta_b}^{\overline{\theta}} \theta\, dF(\theta) - \int_{\theta_b}^{\overline{\theta}} 1 - F(\theta)\, d\theta = \int_{\theta_b}^{\overline{\theta}} \beta(\theta)\, dF(\theta). \end{aligned}$$

Thus, (27) rewrites as

$$\begin{aligned} W = &(1-\mu)\Bigg\{ \int_{\underline{\theta}}^{\theta_a} \alpha(\theta) q^0(\theta) - c(q^0(\theta)) - [\alpha(\theta) q^1 - c(q^1)]\, dF(\theta) \\ &+ \int_{\theta_b}^{\overline{\theta}} \beta(\theta) q^0(\theta) - c(q^0(\theta)) - [\beta(\theta) q^1 - c(q^1)]\, dF(\theta) \Bigg\} \\ &+ \underline{\theta} q^1 - c(q^1). \end{aligned}$$

For a given q^1, denote by $\mathscr{P}(q^1)$ the monopolist's problem to choose $\eta = (\theta_a, \theta_b, q^0(\cdot))$ so as to maximize W subject to MON. Consider an auxiliary problem, denoted $\mathscr{A}$, where the monopolist chooses θ_a, θ_b and two increasing schedules $q_a^0 : [\underline{\theta}, \theta_a) \to [0, q^1)$ and $q_b^0 : (\theta_b, \overline{\theta}] \to (q^1, \infty)$ so as to maximize

$$\begin{aligned} \tilde{W} = (1-\mu)\Bigg\{ &\int_{\underline{\theta}}^{\theta_a} \alpha(\theta) q_a^0(\theta) - c(q_a^0(\theta)) - [\alpha(\theta) q^1 - c(q^1)]\, dF(\theta) \\ &+ \int_{\theta_b}^{\overline{\theta}} \beta(\theta) q_b^0(\theta) - c(q_b^0(\theta)) - [\beta(\theta) q^1 - c(q^1)]\, dF(\theta) \Bigg\} \\ &+ \underline{\theta} q^1 - c(q^1). \end{aligned}$$

Contrary to problem $\mathscr{P}(q^1)$, θ_a in problem $\mathscr{A}$ does not need to be smaller than θ_b. Therefore, for every $\eta = (\theta_a, \theta_b, q^0(\cdot))$ that is feasible in $\mathscr{P}(q^1)$, we obtain a combination $\lambda = (\theta_a, \theta_b, q_a^0(\cdot), q_b^0(\cdot))$ which is feasible in $\mathscr{A}$ when we set $q_a^0(\cdot) = q^0(\cdot)$ on $[\underline{\theta}, \theta_a)$ and $q_b^0(\cdot) = q^0(\cdot)$ on $(\theta_b, \overline{\theta}]$. By construction, $\tilde{W}(\eta) = W(\lambda)$. In particular, this implies that $\max \tilde{W} \geq \max W$.

Reversely, this also implies that if there is a solution λ of $\mathscr{A}$ with the property that $\theta_a \leq \theta_b$, then this corresponds to a solution η of $\mathscr{P}(q^1)$ when we set

$$q^0(\theta) = q_a^0(\theta)\mathbb{I}_{[\underline{\theta},\theta_a)}(\theta) + q^1\mathbb{I}_{[\theta_a,\theta_b]}(\theta) + q_b^0(\cdot)\mathbb{I}_{(\theta_b,\overline{\theta}]}(\theta). \tag{28}$$

We now argue that $\hat{\lambda} = (\theta_\alpha(q^1), \theta_\beta(q^1), q_\alpha(\cdot), q_\beta(\cdot))$ is a solution to $\mathscr{A}$. Note first that our regularity assumptions ensure that $q_a(.)$ and $q_b(.)$ are increasing. Together with the definitions of $\theta_\alpha(\cdot)$ and $\theta_\beta(\cdot)$ it then follows that $\hat{\lambda}$ is feasible in $\mathscr{A}$. Consider some $\lambda = (\theta_a, \theta_b, q_a^0(\cdot), q_b^0(\cdot))$ which is feasible in $\mathscr{A}$. We show that $\tilde{W}(\hat{\lambda}) \geq \tilde{W}(\lambda)$. We consider the case that $\theta_a < \theta_\alpha(q^1)$ and $\theta_b < \theta_\beta(q^1)$. The other cases can be dealt with analogously. Then we have

$$\begin{aligned}
\frac{\tilde{W}(\hat{\lambda}) - \tilde{W}(\lambda)}{1-\mu} \quad = \quad & \int_{\underline{\theta}}^{\theta_a} \alpha(\theta)q_\alpha(\theta) - c(q_\alpha(\theta)) - [\alpha(\theta)q_a^0(\theta) - c(q_a^0(\theta))]\, dF(\theta) \\
& + \int_{\theta_a}^{\theta_\alpha(q^1)} \alpha(\theta)q_\alpha(\theta) - c(q_\alpha(\theta)) - [\alpha(\theta)q^1 - c(q^1)]\, dF(\theta) \\
& + \int_{\theta_b}^{\theta_\beta(q^1)} \beta(\theta)q^1 - c(q^1) - [\beta(\theta)q_b^0(\theta) - c(q_b^0(\theta))]\, dF(\theta) \\
& + \int_{\theta_\beta(q^1)}^{\overline{\theta}} \beta(\theta)q_\beta(\theta) - c(q_\beta(\theta)) - [\beta(\theta)q_b^0(\theta) - c(q_b^0(\theta))]\, dF(\theta).
\end{aligned}$$

Because $q_\alpha(\theta)$ maximizes $\alpha(\theta)q - c(q)$, the first and the second integral are positive. To see that the third integral is positive, note that for all $\theta \in (\theta_b, \theta_\beta(q^1))$, we have

$$q_\beta(\theta) < q^1 < q_b^0(\theta),$$

where the second inequality holds since, by assumption, λ is feasible in $\mathscr{A}$. Hence, because $\beta(\theta)q - c(q)$ is concave in q and maximized by $q_\beta(\theta)$, it follows that $\beta(\theta)q^1 - c(q^1) \geq \beta(\theta)q_b^0(\theta) - c(q_b^0(\theta))$.. Thus the third integral is positive. Finally, the fourth integral is positive because $q_\beta(\theta)$ is a maximizer of $\beta(\theta)q - c(q)$.

Thus, we have established that $\hat{\lambda}$ is a solution to $\mathscr{A}$. Because $\theta_\alpha(q^1) \leq \theta_\beta(q^1)$, the combination $\hat{\eta} = (\theta_\alpha(q^1), \theta_\beta(q^1), \hat{q}_0(\cdot))$, with $\hat{q}_0(\cdot)$ defined by (28), is a solution to $\mathscr{P}(q^1)$. This solution has all the properties stated in the lemma. $\square$

Proof of Lemma 4: By contradiction, suppose that $\hat{q}^1 > q^*(\underline{\theta})$ is optimal. By Lemma 3 and the definition of $\theta_\alpha(\cdot)$, it then follows that $\hat{\theta}_a = \theta_\alpha(\hat{q}^1) > \underline{\theta}$.

Consider an alternative mechanism $(\tilde{q}^1, \tilde{\theta}_a, \tilde{\theta}_b, \tilde{q}^0(\theta))$ with $\tilde{q}^1 = q^*(\underline{\theta})$ so that, by Lemma 3, $\tilde{\theta}_a =$

$\theta_\alpha(\tilde{q}^1) = \underline{\theta}$, $\tilde{\theta}_b = \theta_\beta(\tilde{q}^1) \in (\underline{\theta}, \overline{\theta})$, and

$$\tilde{q}^0(\theta) = \begin{cases} \tilde{q}^1 & if \quad \theta \in [\underline{\theta}, \theta_\beta(\tilde{q}^1)] \\ q_\beta(\theta) & if \quad \theta \in (\theta_\beta(\tilde{q}^1), \overline{\theta}]. \end{cases}$$

By concavity of $c(\cdot)$ and by construction of the mechanism:

$$c'(q^0(\theta)) \geq c'(\tilde{q}^0(\theta)) \geq \beta(\theta), \quad \text{and} \quad u^0(\underline{\theta}) \geq \tilde{u}^0(\underline{\theta}) = 0. \tag{29}$$

In order to see that this alternative mechanism yields strictly more, note that we can use the second expression in (14) together with integration by parts to express the principal's profits from an arbitrary mechanism as

$$W = (1-\mu)\left\{\int_{\underline{\theta}}^{\overline{\theta}} \beta(\theta)q^0(\theta) - c(q^0(\theta))\, dF(\theta) - u^0(\underline{\theta})\right\} + \mu\{\underline{\theta}q^1 - c(q^1)\}. \tag{30}$$

It then follows that, for three reasons, our alternative mechanism based on $\tilde{q}^1 = q^*(\underline{\theta})$ yields more than the original mechanism based on $q^1 > q^*(\underline{\theta})$. First, by (29), the integral is larger for the alternative than the original mechanism (since, due to $c''(q) > 0$, the virtual surplus , $\beta(\theta)q - c(q)$, is concave in q). Second, by (29) $u^0(\underline{\theta}) \geq \tilde{u}^0(\underline{\theta})$. Finally, the last term, $\mu\{\underline{\theta}q^1 - c(q^1)\}$, is larger for the alternative than the original mechanism, because the alternative mechanism specifies the first best quantity at $\underline{\theta}$. This proves that at an optimal mechanism, we have $\hat{q}^1 \leq q^*(\underline{\theta})$. By definition of $\theta_\alpha(\cdot)$, Lemma 3 then implies that $\hat{\theta}_a = \underline{\theta}$ □

Proof of Lemma 5: Because $\hat{\theta}_a = \underline{\theta}$ by Lemma 4, Lemma 3 implies that the optimal schedule $\hat{q}^0(\cdot)$ is such that $\hat{q}^0(\theta) = q^1 \leq q^*(\underline{\theta})$ for valuations $\theta \in [\underline{\theta}, \hat{\theta}_b]$ and $\hat{q}^0(\theta) = q_\beta(\theta)$ for valuations $\theta \in (\hat{\theta}_b, \overline{\theta}]$. Moreover, $\hat{\theta}_b = \theta_\beta(q^1)$, or, equivalently $q^1 = q_\beta(\theta_\beta)$. Because of the one-to-one relationship between q^1 and θ_β, we can write the monopolist's problem as a problem of choosing θ_β (rather than q^1). Recall that if $\beta(\underline{\theta}) \geq 0$, then the range of θ_β is $[\underline{\theta}, \overline{\theta}]$, while when $\beta(\underline{\theta}) < 0$ the range of θ_β is $[\theta_0, \overline{\theta}]$, where θ_0 is defined by $\beta(\theta_0) = 0$.

By (27), the objective as a function of θ_β can be written as

$$\begin{aligned} W(\theta_\beta) =& (1-\mu)\left\{\int_{\theta_\beta}^{\overline{\theta}} \beta(\theta)q_\beta(\theta) - c(q_\beta(\theta))dF(\theta) - [1-F(\theta_\beta)][\theta_\beta q_\beta(\theta_\beta) - c(q_\beta(\theta_\beta))]\right\} \\ &+ \underline{\theta}q_\beta(\theta_\beta) - c(q_\beta(\theta_\beta)). \end{aligned} \tag{31}$$

The derivative is

$$
\begin{aligned}
W'(\theta_\beta) &= (1-\mu)\{-f(\theta_\beta)[\beta(\theta_\beta)q_\beta(\theta_\beta)-c(q_\beta(\theta_\beta))] \\
&\qquad +f(\theta_\beta)[\theta_\beta q_\beta(\theta_\beta)-c(q_\beta(\theta_\beta))]-[1-F(\theta_\beta)][q_\beta(\theta_\beta)+(\theta_\beta-c'(q_\beta(\theta_\beta)))q'_\beta(\theta_\beta)]\} \\
&\quad +[\underline{\theta}-c'(q_\beta(\theta_\beta))]q'_\beta(\theta_\beta) \\
&= (1-\mu)\{f(\theta_\beta)[\theta_\beta-\beta(\theta_\beta)]q_\beta(\theta_\beta)-[1-F(\theta_\beta)][q_\beta(\theta_\beta)+(\theta_\beta-\beta(\theta_\beta))q'_\beta(\theta_\beta)]\} \\
&\quad +[\underline{\theta}-\beta(\theta_\beta)]q'_\beta(\theta_\beta) \\
&= -(1-\mu)[1-F(\theta_\beta)][(\theta_\beta-\beta(\theta_\beta))q'_\beta(\theta_\beta)]+[\underline{\theta}-\beta(\theta_\beta)]q'_\beta(\theta_\beta) \\
&= \left\{\underline{\theta}-\beta(\theta_\beta)-(1-\mu)\frac{(1-F(\theta_\beta))^2}{f(\theta_\beta)}\right\}q'_\beta(\theta_\beta)=\phi(\theta_\beta)q'_\beta(\theta_\beta),
\end{aligned}
$$

where the second equality uses $c'(q_\beta(\theta_\beta)) = \beta(\theta_\beta)$, the third equality follows after using $f(\theta_\beta)[\theta_\beta - \beta(\theta_\beta)] = 1-F(\theta_\beta)$ from the definition of $\beta(\theta)$ and cancelling terms, and the third equality uses the definition of $\beta(\theta)$ after factoring out the term $q'_\beta(\theta_\beta)$.

We next argue that $W(\theta_\beta)$ is quasi-concave in θ_β. To see this, write $h(\theta) = (1-F(\theta))/f(\theta)$ as the inverse hazard rate so that $\beta'(\theta) = 1-h'(\theta)$. It then follows

$$
\begin{aligned}
\phi'(\theta) &= -(1-h'(\theta))-(1-\mu)h'(\theta)(1-F(\theta))+(1-\mu)h(\theta)f(\theta) \\
&= -(1-h'(\theta))+(1-\mu)(1-F(\theta))(1-h'(\theta)) \\
&= [(1-\mu)(1-F(\theta))-1](1-h'(\theta))=[(1-\mu)(1-F(\theta))-1]\beta'(\theta) \\
&< 0,
\end{aligned}
$$

where the inequality follows since the term in square brackets is negative, whereas $\beta'(\theta)$ is positive. Moreover, it holds

$$
\phi(\underline{\theta}) = \mu\frac{1}{f(\underline{\theta})} > 0, \quad \phi(\overline{\theta}) = \underline{\theta}-\overline{\theta} < 0.
$$

Together with $\phi'(0) < 0$, this implies that there is a unique θ_ϕ such that $\phi(\theta_\phi) = 0$; and together with $q'_\beta(\theta_\beta) \geq 0$, it also implies that $W(\theta_\beta)$ is quasi-concave in θ_β.

For the case $\beta(\underline{\theta}) \geq 0$, where the monopolist can choose any $\theta_\beta \in [\underline{\theta}, \overline{\theta}]$, this implies that θ_ϕ is a unique maximizer of $W(\theta_\beta)$.

For the case $\beta(\underline{\theta}) < 0$, where the monopolist can choose any $\theta_\beta \in [\theta_0, \overline{\theta}]$, this implies that θ_ϕ is a unique maximizer of $W(\theta_\beta)$ if $\theta_\phi \geq \theta_0$, whereas the corner solution θ_0 maximizes $W(\theta_\beta)$ if $\theta_\phi < \theta_0$.

Finally, note that $\hat{\theta}_b = \theta_\phi$ implies $\hat{q}^1 = q_\beta(\theta_\phi)$, in which case $\hat{q}^1 > 0$ because $\beta(\theta_\phi) > 0$ and thus $q_\beta(\theta_\phi) > 0$ by (2). Moreover, $\hat{\theta}_b = \theta_0$ implies $\hat{q}^1 = q_\beta(\theta_0) = 0$. This completes the proof. □

Proof of Proposition 2: The first observation follows directly from Lemmata 3, 4, and 5, and the fact that (IC^1) is slack whenever $K \geq \bar{K}$, and using the price according to (7). Part 2 of the proposition follows from incentive compatibility (IC_0). Part 3 of the proposition follows from a direct comparison of the optimal schedules when privacy types are private and public. Part 4 follows from Lemma 5. □

Proof of Lemma 6: To see (i), fix μ and the associated optimal privacy contract $\hat{\gamma}$. Denote by $\hat{W}^0$ the monopolist's expected profit conditional on facing a classical consumer and by $\hat{W}^1$ the monopolist's expected profit conditional on facing a data-sensitive consumer. Hence, we have $\hat{W} = \mu\hat{W}^1 + (1-\mu)\hat{W}^0$. Note that instead of offering $\hat{\gamma}$, it is feasible to only offer the privacy option $(\hat{q}^1, \hat{p}^1)$ to all consumer types. Under this alternative contract the monopolist would obtain profits $\hat{W}^1$ also from classical consumers. Hence, optimality of $\hat{\gamma}$ implies that $\hat{W}^1 \leq \hat{W}^0$ Note moreover that if we decrease μ to $\mu' = \mu - \Delta\mu$ with $\Delta\mu > 0$, the privacy contract $\hat{\gamma}$ remains feasible (even though no longer optimal) and yields the expected profit of $\hat{W} - \Delta\mu(\hat{W}^1 - \hat{W}^0)$, which is larger than $\hat{W}$, because $\hat{W}^1 \leq \hat{W}^0$. The optimal privacy contract under μ' must therefore yield at least $\hat{W}$. Thus, $d\hat{W}/d\mu \leq 0$.

To see (ii), recall from Lemma 5 and its proof that $\phi(\hat{\theta}_b) = 0$, $\partial\phi/\partial\theta \leq 0$. Moreover, from inspection, $\partial\phi/\partial\mu \geq 0$. By the implicit function theorem, it then follows that $d\hat{\theta}_b/d\mu \geq 0$. To see the second statement in (ii), recall that $\hat{q}^1 = q_\beta(\hat{\theta}_b)$. Thus, $d\hat{q}^1/d\mu = q'_\beta(\hat{\theta}_b) \cdot d\hat{\theta}_b/d\mu$, which is positive since both factors are positive. Finally, recall from Proposition 2 that $\hat{q}^0(\theta) = \hat{q}^1$ for $\theta \in [\underline{\theta}, \theta_\beta]$ and $\hat{q}^0(\theta) = \bar{q}^0(\theta)$ for $\theta \in (\theta_\beta, \overline{\theta}]$. Since $d\hat{q}^1/d\mu \geq 0$ and since $\bar{q}^0(\cdot)$ is independent of μ, it follows $d\hat{q}^0(\theta)/d\mu \geq 0$.

To see (iii), note that

$$\hat{u}^1(\theta) = (\theta - \underline{\theta})\hat{q}^1, \quad \text{and} \quad \hat{u}^0(\theta) = \begin{cases} (\theta - \underline{\theta})\hat{q}^1 & \text{if } \theta \leq \theta_\beta \\ (\theta_\beta - \underline{\theta})\hat{q}^1 + \int_{\theta_\beta}^{\theta} \hat{q}^0(t)dt & \text{otherwise.} \end{cases} \tag{32}$$

Taking derivatives with respect to μ yields

$$\frac{d\hat{u}_1(\theta)}{d\mu} = (\theta - \underline{\theta})\frac{d\hat{q}^1}{d\mu}, \quad \text{and} \quad \frac{d\hat{u}_0(\theta)}{d\mu} = \begin{cases} (\theta - \underline{\theta}) \cdot d\hat{q}^1/d\mu & \text{if } \theta \leq \theta_\beta \\ (\theta_\beta - \underline{\theta}) \cdot d\hat{q}^1/d\mu & \text{otherwise,} \end{cases} \tag{33}$$

where the last line follows, because

$$\frac{d\hat{u}^0(\theta)}{d\mu} = (\theta_\beta - \underline{\theta})\frac{d\hat{q}^1}{d\mu} + \hat{q}^1\frac{d\hat{\theta}_b}{d\mu} - \hat{q}^0(\theta_\beta)\frac{d\hat{\theta}_b}{d\mu} \tag{34}$$

and because $\hat{q}^0(\theta_\beta) = \hat{q}^1$

Statement (iii) then directly follows as $d\hat{q}^1/d\mu \geq 0$ by part (ii). □

References

Acquisti, A., Taylor, C.R., Wagman, L., 2016. "The economics of privacy." Journal of Economic Literature 52 (2), 442-492.

Athey, S., 2014. Information, Privacy and the Internet: An Economic Perspective. Central Planning Bureau (CPB) Netherlands Bureau for Economic Policy Analysis.

Becker, G. S. (1980), "Privacy and malfeasance: A comment", Journal of Legal Studies 9(4), 823-826.

Calzolari, G., A. Pavan (2006). "On the Optimality of Privacy in Sequential Contracting," Journal of Economic Theory, 130(1), 168-204.

Choi, J., D. Jeon, B. Kim (2019). "Privacy and Personal Data Collection with Information Externalities," Journal of Public Economics, 173, 113-124.

Dworczak, P. (2020) "Mechanism Design with Aftermarkets: Cutoff Mechanisms", Econometrica 88 (6), 2629-2661.

Eilat, R., K. Eliaz, and X, Mu (2021) "Bayesian Privacy". Forthcoming in Theoretical Economics.

EU (2018) Consumer market study on online market segmentation through personalised pricing/offers in the European Union Request for Specific Services 2016 85 02 for the implementation of Framework Contract EAHC/2013/CP/04 Final report

Fudenberg, D. and Villas-Boas, J. M. (2006). "Behavior-based price discrimination and customer recognition", Handbook on Economics and Information Systems 1, 377–436.

Ichihashi, S. (2020), "Online Privacy and Information Disclosure by Consumers", American Economic Review 110, 569-595.

Jullien, B. (2000), "Participation Constraints in Adverse Selection Models". Journal of Economic Theory 93, 1-47.

Lewis, T. and D. Sappington (1989), "Countervailing incentives in agency problems." Journal of Economic Theory 49, 294-313

Lin, T. (2020), "Valuing Intrinsic and Instrumental Preferences for Privacy", mimeo Boston University Questrom School of Business.

Mussa, M and S. Rosen (1978), "Monopoly and Product Quantity", Journal of Economic Theory 18, 301-317.

Tang, H. (2020), "The Value of Privacy: Evidence from Online Borrowers.", mimeo HEC Paris.

Taylor, R., C. (2004). "Consumer Privacy and the Market for Customer Information," RAND Journal of Economics, 631-650.

Tsai, J., S. Egelman, L. Cranor and A. Acquisti (2011), "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study", Information Systems Research 22 (2), 254-268.