Screening and the Welfare Trade-offs of Privacy

Anastasios Dosis * Casey Rothschild†

January 25, 2024

Abstract

A principal and agent interact in a class of models a la Mussa and Rosen (1978) with continuously divisible product quality. Prior to their interaction, a notional planner designs the information disclosure environment which determines the extent to which the principal can observe the willingness-to-pay of the agent. We contrast the class of Pareto optimal information disclosure policies when the willingness-to-pay of the consumer is exogenous and when, instead, the distribution of agent types is determined by an unobservable and costly investment by the agent. In either class, full privacy may or may not be Pareto optimal, and within the class of Pareto optima, there is always a strict tradeoff between agent welfare and total welfare (as measured by agent surplus plus principal surplus). In a two-type model, this tradeoff is fully parameterized by a one-dimensional measure of information disclosure. Full disclosure is always Pareto optimal with exogenous willingness-to-pay and never Pareto optimal with endogenous willingness-to-pay.

KEYWORDS: Screening, Privacy, Price discrimination, Welfare

JEL CLASSIFICATION: D61, D82, D83, D86, L12

*ESSEC Business School and THEMA, 3 Av. Bernard Hirsch, B.P. – 50105, Cergy, 95021, France; Email: dosis@essec.edu

†Wellesley College, 106 Central Street Wellesley, MA 02481, USA; Email: crothsch@wellesley.edu

1 Introduction

The widespread collection and processing of personal data has given rise to a heated debate regarding citizens' privacy. Opponents of privacy (such as the Chicago school, e.g., Stigler, 1980; Posner, 1981) argue that privacy regulations may hinder the free flow of information and potentially exacerbate allocative inefficiencies. They also suggest that such regulations could impede product and service improvements that enhance overall welfare. By contrast, advocates of privacy (e.g., Bennett, 2010; Zuboff, 2023) argue for limitations on data collection by large corporations and propose granting more property rights to individuals who generate this data.

The concerns regarding consumers' privacy have given rise to two important recent regulations: the General Data Protection regulation and the California Consumer Privacy Act. The scope of both regulations is to provide consumers with more rights over their personal data. The underlying principle of both regulations is that by giving more control to consumers over the use of their data, the latter can reap some of the benefits as well as limit exploitation by large firms.

Nonetheless, some fundamental questions remain unresolved. For example, is more privacy always beneficial for the economic agents who generate the data, such as consumers or workers? Moreover, if increased privacy indeed benefits consumers or workers, does it consistently lead to improved allocative efficiency and, consequently, higher social welfare? Answers to these questions hold significant policy implications. If strengthening privacy not only benefits consumers but also enhances allocative efficiency, it would provide further justification for more stringent privacy laws. Conversely, if complete privacy proves to be suboptimal, how do underlying economic principles and distributional goals affect the optimal level of privacy?

To answer these questions, we consider a class of settings in which an uninformed Principal with monopoly power interacts with an Agent who has private information. The particular example on which we focus is that in which the Principal is a monopolistic firm who sells a product to the Agent who is a consumer. Accordingly, we refer henceforth to the Principal as the firm and the Agent as the consumer. Prior to their interaction, a (notional) planner can commit to an information disclosure – or “privacy” – policy, whereby the firm may learn some of the consumer’s information. In Mussa and Rosen (1978) for instance, there is a consumer whose willingness-to-pay is unknown by the firm which can choose product quality.[1] Our primary focus is on whether, how, and how much privacy of the consumer should be reduced by the planner’s commitment to information disclosure.

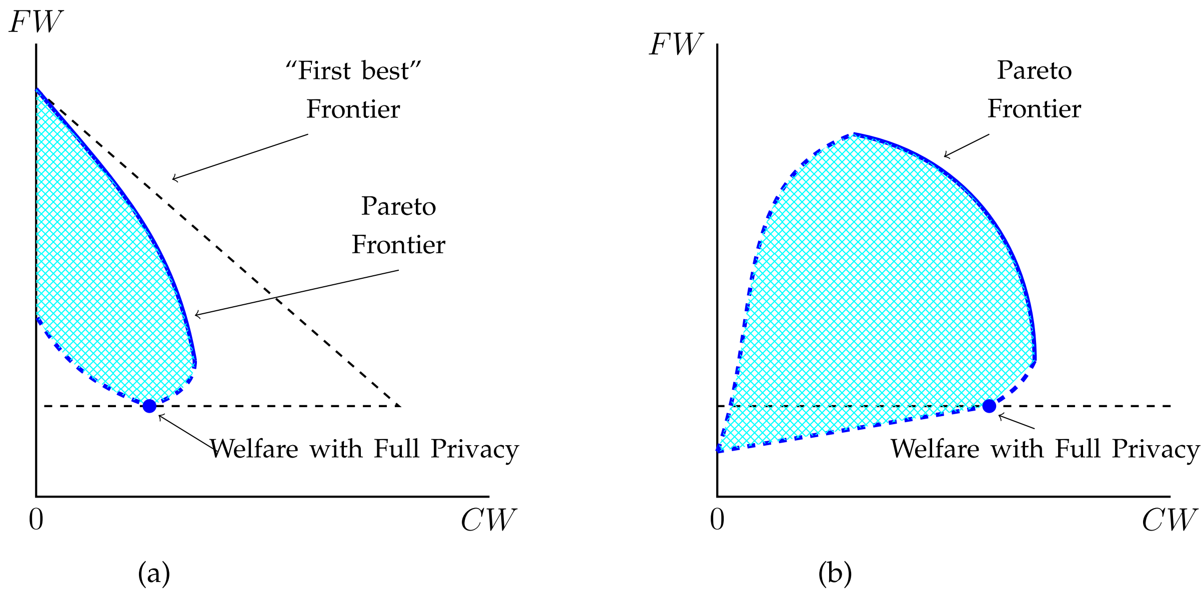

Figure 1: Panel (a) depicts the “First best” frontier as well as the Pareto frontier when the distribution of types is exogenously given. Panel (b) depicts the Pareto frontier when the distribution of types is endogenous and depends on the agent’s effort.

Our first main result (Proposition 1) is illustrated in Figure 1a. It states that, within the class of Pareto efficient information disclosure decisions, there is generically a “leaky bucket” a la Okun (1975): a tradeoff between consumer welfare and efficiency. Loosely speaking, information disclosure can yield Pareto improvements, but once Pareto gains have been exhausted, additional information disclosure can still always be used to improve efficiency – but only at the expense of consumer welfare. Or, put yet another way: points on the Pareto frontier that are better for consumers are farther away from the first-best frontier.

[1]Alternatively, the Principal may be a monopsonistic employer contracting with an employee of unknown productivity. Our analysis also applies to the design (a la Mirrlees, 1971 and Stiglitz, 1982) of a nonlinear tax polices by an extractive or Rawlsian state Principal for citizen Agents of unknown skill type.

This result stems from the trade-off between price discrimination and better product matching.[2] In particular, more information allows the firm to better tailor the products to the needs of the consumers but also increases its ability to extract consumer surplus. Therefore, although less refined information might be preferred by consumers, this comes at a cost in terms of total welfare.

The relationship between the trade-off between consumer and total welfare and the privacy regime is clearest when there are exactly two possible types. Our second main result (Proposition 2) characterizes the set of Pareto optimal information disclosure policies in this two-type case. It establishes that, in any Pareto optimal information structure (other than the one featuring zero privacy), the low-valuation type enjoys full privacy whereas the high-valuation type enjoys at least partial privacy. Along the Pareto frontier, more privacy (parameterized by the probability with which the high-valuation type becomes known to the principal) increases consumer welfare but decreases total welfare. This result uncovers a novel trade-off between consumer welfare and total welfare parameterized by the amount of privacy. Taken together, these two results (Propositions 1 and 2) imply that in environments in which firms can screen consumers, strengthening privacy might increase consumer welfare but will hurt allocative efficiency (and hence total welfare).

The fact that consumers might not prefer full privacy is a consequence of the trade-off between rent-extraction and efficiency: although more information disclosure decreases the probability with which the agent indeed receives an information rent, it increases the agent's information rent, conditional on receiving it. When the latter effect dominates, the resulting increase in consumer welfare arises because the firm is able to better tailor the quality of the product to the low-valuation type, thereby allows the high valuation type to earn a higher rent. Our analysis is thus related to the large literature that studies the trade-off between consumer identification that allows product matching (or more targeted advertising) and price discrimination.[3]

[2]A similar trade-off is also present in the recent contributions by Ichihashi (2020) and Hidir and Vellodi (2021). Both papers study consumers' optimal market segmentation in a multi-product monopoly market in which the monopolist is allowed to offer a single product in each segment.

[3]See Goldfarb and Tucker (2019) for a detailed survey on these topics.

We extend our model to accommodate for an endogenous distribution of types. In particular, after learning the information environment, we allow the agent to take an unobservable effort decision, with higher effort probabilistically increasing the value of the product to the consumer, and hence the potential gains from trade.

Consider, for instance, the example of a restaurant entrepreneur (the consumer) who is negotiating with a monopolistic realtor (the firm) to rent a new space for her business. The entrepreneur's valuation of location (quality) depends on many factors such as the restaurant's style, cuisine, etc. Moreover, and importantly, her valuation might crucially depend on whether her top-choice star chef has accepted to work in her restaurant—something that required past costly recruitment efforts via (potentially) private negotiations.

Introducing endogenous effort fundamentally changes the welfare effects of information disclosure. Figure 1b illustrates the qualitative insights of this change. In that Figure, full disclosure (zero privacy) is now Pareto dominated by full privacy and, indeed, yields 0 surplus for consumers and minimal surplus for the firm.[4] The basic intuition is simple: if consumers anticipate that the monopolist will observe their type, then they will anticipate that firms will extract all potential gains from their investment decision. As such, full information revelation will eliminate incentives to exert effort – hurting consumers and firms alike.[5]

The paper proceeds as follows. In Section 2, we describe both the baseline model that we study formally in the remainder of the paper and other, closely related models to which the same analysis would carry over directly. In Section 3, we first demonstrate the general tradeoff between consumer welfare and total welfare along the Pareto frontier and then fully characterize Pareto optimal information disclosure in a two-type version of the model. In Section 4, we characterize Pareto optimal information disclosure with endogenous effort, and show that many but not all of the insights from the exogenous-type model carry over directly; indeed, our characterization relies on a tool we develop for bootstrapping results from the exogenous to the endogenous case. Section 5 discusses

[4]That is, the surplus associated with selling to an agent who is known to be the low-willingness-to-pay type.

[5]Note that "first best efficiency" is the sense depicted in the earlier diagrams is a less useful concept in this context, as the total surplus achievable depends on effort.

implications and concludes.

2 Model

2.1 Focal Model: Firm and Consumer

Our focal model is a discrete-type version of Mussa and Rosen (1978). A monopolistic firm sells a customized product to a consumer, whose quality is indexed by $q \in [0, \infty]$. The firm is an expected utility maximizer with utility function

$$\Pi(q,t) = t - C(q). \tag{1}$$

The cost function $C(\cdot)$ is assumed to satisfy $C(0) = C'(0) = 0$, $C'(q) > 0$, $C''(q) > 0$ for every $q > 0$, and $C'''(q) \geq 0$ for every $q \geq 0$.

At the time the buyer contemplates a purchase, she has preferences

$$U_i(q,t) = \theta_i q - t. \tag{2}$$

Per Equation (2), θ_i measures consumer's (per unit) value for the product. We assume that there is a finite number of types $i \in \{1, ..., N\}$, where $N \geq 2$ and higher types have a higher value for the product $\theta_1 < \theta_2 < ... < \theta_N$. With some abuse of notation, we also use N to denote the set of types. The share of type i in the population is $\lambda_i > 0$, where $\sum_i \lambda_i = 1$.

We evaluate welfare of consumers and firms from an ex-ante perspective, so that consumer welfare CW and firm welfare FW are given respectively by the ex-ante expected value of $\theta_i q - t$ and $t - C(q)$ (as we elaborate on later).

2.2 Alternative Models

Although our formal analysis treats the Mussa-Rosen model described in the preceding subsection, all of our analysis applies equally well to the related but distinct model where preferences are instead:

$$\tilde{U}_i = t - q/\theta_i \tag{3}$$

$$\tilde{\Pi} = g(q) - t, \tag{4}$$

with g increasing and g and g' concave, or, equivalently (with an obvious re-parametrization of variables), as a model with

$$\hat{U}_i = t - h(q)/\theta_i \tag{5}$$

$$\tilde{\Pi} = q - t, \tag{6}$$

with $h(q)$ increasing and convex and h' weakly concave. The former has a natural interpretation as a model of a monopsonistic firm employing a worker of unknown productivity produce quality q in exchange for compensation t. As discussed below, the latter is interpretable as an optimal tax model.

2.3 Examples

The basic Mussa and Rosen (1978) model and the related models sketched above are standard, but we provide several descriptive examples of possible applications.

Software. An organization (the consumer) contemplating an investment in a new data management system, offered by a software enterprise (the firm) with monopoly power. The organization's valuation of the system's functionality (the quality) depends on the state of the company's legacy data which in its turn depends on how ready its employees are to learn and adopt a new system (i.e., the type of the consumer).

Patent Licences. An inventor (the consumer) has a new product that makes use of a patented idea. The inventor is negotiating a licensing deal with the patent-holder (the firm). Her valuation of the extent of the licence (the quality) depends on how well-developed her new product is (i.e., the type of the consumer).

Firm-Specific Human Capital. An employer (the firm) will assign a task q to their employee (the consumer) and compensate them with salary t. The ease of accomplishing the task is inversely proportional to the employee's type.

Taxation by an Extractive State. An extractive state (the firm) will raise revenue through a non-linear income tax $T(q)$ on observable earned income q, leaving $t = q - T(q)$ for consumers. The disutility of generating pre-tax income q depends inversely on the skill θ_i of the worker.

2.4 Information Disclosure and Privacy

Absent information disclosure – which we refer to as "full privacy" – the consumer's type is purely private information. Following the Bayesian persuasion and information design literatures (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019) we model *information* as follows: there is some finite set of signals S, with generic element s, and a set of probability distributions $\{\pi_i(s)\}_{i \in N}$, with the interpretation that $\pi_i(s)$ is the probability that firm will see signal s if the true type is i.[6] If $|S| = 1$, then there is *full privacy*. Conversely, the information structure with $S = \{s_1, ..., s_N\}$ and $\pi_i(s_j) = \delta_{ij}$, where δ_{ij} is the Kronecker δ (1 if $i = j$, 0 otherwise) implements *full disclosure* (i.e., no privacy). We refer to $\Sigma = (S, \{\pi_i(s)\}_{i \in N})$ as the *informational environment*, and we discuss its implications further after describing the timing of the model.

2.5 Timing and Equilibrium Basics

We now describe the structure of the game and its equilibrium. The timing is as follows:

0. The informational environment $\Sigma = (S, \{\pi_i(s)\}_{i \in N})$ is determined, and is observed by both the consumer and the firm.

1. The consumer's type i is realized.

2. The consumer's signal s is determined, with signal s sent with probability $\pi_i(s)$ for type i.

3. The firm observes s and chooses a price schedule $t(q; s)$. (Without loss of generality, we assume it must choose $t(0) = 0$).

[6] Although, only for exposition purposes, in the main text we focus on a finite set of signals, in the appendix we show that this is without loss of generality.

4. The consumer chooses a quality level q (and corresponding price $t(q;s)$) and utilities are realized.

We use *Perfect Bayesian Equilibrium (PBE)* as our solution concept for the game for any given Σ, and later characterize the set of Σ which are ex-ante Pareto optimal.

Fixing Σ and (without loss of generality) assuming that it satisfies $\sum_i \pi_i(s) > 0$ for all $s \in S$, an equilibrium is a set of posterior beliefs $\{\beta_i(s)\}_i$, a set of pricing menus $t(q;s)$ and a set of choices $\{q_i(s)\}_i$ such that: $q_i(s)$ is utility maximizing for type-i consumers given $t(q;s)$ for each s; the menus $t(q;s)$ are optimal for firms given beliefs and consumer choices $\{q_i(s)\}_i$; and beliefs are consistent with Bayes' rule, i.e.

$$\beta_i(s) = \frac{\pi_i(s)\lambda_i}{\sum_j \pi_j(s)\lambda_j}. \tag{7}$$

In the standard way, we appeal to the revelation principle to simplify our description of equilibrium. In particular, upon receiving signal s, we can think of the firm as choosing from among the incentive compatible and individually rational menus $\{(t_i(s), q_i(s))\}_i$.

2.6 Welfare

The informational environment Σ determines a distribution over signals s, and the equilibrium—and hence the ex-ante expected utilities of firms and consumers—is fully determined by the beliefs per (7) for each s.

The informational environment Σ implies ex-ante consumer welfare determined by the sum of the utilities (i.e., information rents) of all types:

$$CW(\Sigma) = \sum_i \lambda_i \sum_s \pi_i(s) U_i(q_i(s), t_i(s)) \tag{8}$$

and ex-ante consumer welfare plus firm welfare (total welfare):

$$TW(\Sigma) \equiv CW(\Sigma) + FW(\Sigma) = \sum_i \lambda_i \sum_s \pi_i(s) \left(\theta_i q_i(s) - C(q_i(s))\right). \tag{9}$$

Note that we omit the formula for $FW(\Sigma)$ because it is redundant with (8) and (9) (and $TW(\Sigma)$ turns out to be more useful in what follows than using $FW(\Sigma)$ directly).

We now give a formal definition of Pareto optimal informational environments.

Definition 1. *An informational environment* Σ *is Pareto optimal if there does not exist another informational environment* Σ' *with* $CW(\Sigma) \leq CW(\Sigma')$ *and* $FW(\Sigma) \leq FW(\Sigma')$*, and with at least one of the two inequalities strict.*

3 Results

Our goal in this section is to characterize the set of Pareto optimal informational environments. The following Proposition was illustrated in Figure 1a in the introduction. In it, we use TW^{max} to denote the first-best total welfare.

Proposition 1 (The Leaky Bucket)**.**

1. *If* $CW(\Sigma) > 0$, *then* $TW(\Sigma) < TW^{max}$.
2. *If* Σ *and* Σ' *are both Pareto optimal and* $CW(\Sigma') > CW(\Sigma)$, *then* $TW(\Sigma') < TW(\Sigma)$.

Proposition 1 has two important implications. First, any Pareto optimal informational environment that provides consumers with strictly positive surplus comes at a cost in terms of total welfare. In other words, only full disclosure allows the first-best total welfare to be achieved.[7] Second, any change in the (Pareto optimal) informational environment that increases the welfare of consumers must necessarily decrease total welfare. In light of Proposition 2 below, this uncovers a novel trade-off between consumer and total welfare that, as will become clear, is in fact parametrized by the level of privacy.

We omit the formalities of the proof because the intuition is relatively straightforward. Underlying the first part of Proposition 1 is the idea that any informational environment, other than full disclosure, entails a distortion for some of the types resulting from the desire of the firm to screen consumers. This distortion is socially costly and hence reduces total welfare relative to the first-best total welfare. The second part of Proposition 1 follows from the first part, the fact that the set of achievable (CW, TW) pairs is con-

[7]This is in contrast to the related model analyzed in Section IV of Bergemann et al. (2015) (see their Figure 8). The difference is attributable to the fact that their model features an exogenous bound on the highest quality level that is strictly lower than the unrestricted first best for both types.

vex[8], and the Pareto frontier therefore convex, the fact that $(0, TW^{max})$ is feasible (via full information revelation), and part 1.

The following immediate Corollary provides a test for Pareto optimality that will be useful in proving our second main result:

Corollary 1. *Consider two informational environments Σ and Σ'. If $CW(\Sigma') > CW(\Sigma)$, and $TW(\Sigma') \geq TW(\Sigma)$, then Σ is Pareto inefficient.*

3.1 Two Types

In this section, we focus on an environment with two possible types: $i = 1$ is the low (valuation) type and $i = 2$ is the high (valuation) type. Let $\Delta\theta \equiv \theta_2 - \theta_1$. Standard calculations yield the unique firm-optimal menu:

$$\begin{aligned} q_1(s) &= \max\left\{0, \xi\left(\theta_1 - \frac{\beta_2(s)}{\beta_1(s)}\Delta\theta\right)\right\} \\ t_1(s) &= \theta_1 q_1 \\ q_2(s) &= \xi(\theta_2) \\ t_2(s) &= \theta_2 q_2 - \Delta\theta q_1, \end{aligned} \tag{10}$$

where $\xi(x) \equiv C'^{-1}(x)$ is the inverse of the firm's marginal cost function.

Because the low type earns no information rents, ex-ante consumer welfare is given by:

$$CW(\Sigma) = \lambda_2 \Delta\theta \sum_s \pi_2(s) q_1(s) \tag{11}$$

and ex-ante consumer welfare plus firm welfare (total welfare) is given by:

$$\begin{aligned} TW(\Sigma) &\equiv CW(\Sigma) + FW(\Sigma) \\ &= \lambda_1 \sum_s \pi_1(s)\left(\theta_1 q_1(s) - C(q_1(s))\right) + \lambda_2(\theta_2 \xi(\theta_2) - C(\xi(\theta_2))). \end{aligned} \tag{12}$$

[8]This follows from a standard concavification argument: if there is an information structure $(S_1, \{\pi_{1i}(s)\}_{i\in N})$ yielding welfare (CW_1, TW_1) and a second, $(S_2, \{\pi_{2i}(s)\}_{i\in N})$, yielding welfare (CW_2, TW_2), then—without loss of generality taking S_1 and S_2 to be disjoint—the information structure $(S_1 \cup S_2, \{\pi_{\alpha i}(s)\}_{i\in N,})$ with $\pi_i \alpha i(s) = \alpha\pi_{1i}(s)$ if $s = S_1$ and $\pi_i \alpha i(s) = (1-\alpha)\pi_{2i}(s)$ if $s \in S_2$ yields welfare $\alpha(CW_1, TW_1) + (1-\alpha)(CW_2, TW_2)$.

The following Proposition characterises the set of Pareto optimal informational environments and highlights the trade-off between consumer and total welfare entailed by privacy.

Proposition 2 (Optimal Privacy with Exogenous e)**.** *There exists $\underline{\pi} \geq 0$ such that an informational environment is Pareto optimal if and only if it is equivalent to a two-signal information structure $S = \{s^{priv}, s^2\}$ with $\pi_1(s^{priv}) = 1$ and $\pi_2(s^2) \in [\underline{\pi}, 1]$. Within this class, equilibrium consumer welfare is decreasing and both firm welfare and total (firm plus consumer) welfare are increasing in $\pi_2(s^2)$.*

The firm-optimal informational environment always entails full information disclosure. The key content of Proposition 2 is to describe the information environments associated with the remainder of the Pareto frontier. In plain English, it can be understood as stating four distinct things about this remainder. First, any such Pareto optimum entails full privacy for the $i = 1$ types. Second, the consumer-optimal informational regime may or may not coincide with full privacy; that depends on whether $\underline{\pi} = 0$, in which case full privacy for both types is consumer-optimal, or if, instead, $\underline{\pi} > 0$, in which case the consumer optimum involves revealing type 2's type with probability $\underline{\pi}$.[9] Third, within the class of Pareto optimal information environments, the amount of privacy can be fully parameterized by the probability $\pi_2(s_2) \equiv 1 - p_2$ with which type $i = 2's$ type is revealed. Fourth, within the class of Pareto optima, consumer welfare increases with the amount of privacy (higher p_2) and while firm welfare—and hence, by Proposition 1 total welfare—decreases with the amount of privacy.

A detailed proof of the proposition is in the appendix, but much of the basic intuition can be understood via Figure 2. The solid red curve in that figure plots the consumer welfare, under full privacy, as a function of the prior fraction λ_2 of type 2s. (It can also be interpreted as the per-consumer utility as a function of the posterior beliefs of the firm.)

[9] Although it is not explicitly stated in the proposition, it is straightforward to construct examples of markets in which $\underline{\pi} > 0$. For instance, one can readily construct models such that, for λ_2 sufficiently high, the firm would optimally exclude type $i = 1$ (to avoid paying information rents to type $i = 2$) under full privacy, and CW would be zero under full privacy. In such a model, sufficiently decreasing privacy (without entirely eliminating it) will lead the firm to decide not to shut down type $i = 1$. As such, decreasing privacy which would increase CW—and, because more information is good for firms, will also increase FW.

For $\lambda_2 < \frac{\theta_1}{\theta_2}$, CW is a strictly positive and concave function of λ, while for $\lambda_2 \geq \frac{\theta_1}{\theta_2}$, the firm finds it optimal to exclude type 1 (i.e., chooses $q_1 = 0$), and CW is identically zero.

As a first step in the proof, note that it follows that full privacy is not Pareto optimal when $\lambda_2 \geq \frac{\theta_1}{\theta_2}$. To see why, start from such a prior, and then change the informational environment so as to reveal the identity of type 2's with sufficiently high probability that the firm's posterior conditional on non-revelation lies in the interval $\left(0, \frac{\theta_1}{\theta_2}\right)$. This will be Pareto improving, as more information is clearly good for firms, and this disclosure will lead to a positive probability of that the type 2 consumer's type will remain private and they will earn positive rents.

In fact, full privacy is not Pareto optimal whenever $\lambda_2 \geq \lambda_2^{max}$, where, as indicated in the figure λ_2^{max} is the unique point where the secant line connecting the red curve to the point $(1, 0)$ is tangent to the red curve. For priors $\lambda_2 \in (\lambda_2^{max}, 1)$, revealing type 2's identity with the probability $(1 - p_2^*)$ satisfying $\lambda_2^{max} = \frac{\lambda_2 p_2^*}{\lambda_2 p_2^* + (1-\lambda_2)}$ will mean that, with probability p_2^*, type 2s identity will remain private and the firm's posterior will be λ_2^{max}, while with probability $1 - p_2^*$, type 2's identity will be revealed. So consumer welfare will either be 0 (if the consumer is the L type or if the consumer's type is revealed) or will be given by the height of the red curve in Figure 2 at the point λ_2^{max}. In expectation, it will be given by the height of the dashed blue secant line in the figure at the prior λ_2. Full privacy is thus Pareto inefficient for any prior $\lambda > \lambda_2^{max}$: revealing type 2's identity with probability $1-p_2^*$ will increase expected consumer welfare from the red to the dashed blue line in Figure 2, and will simultaneously increase firm welfare. In fact, by the same basic logic, any information structure which puts positive probability on a posterior greater than λ_2^{max} is similarly Pareto inefficient. So any Pareto efficient allocation be a associated with a distribution of posterior beliefs with support on $[0, \lambda^{max}] \cup \{1\}$.[10]

As a second step, consider any information structure featuring multiple distinct signals that lead to posteriors in the interval $[0, \lambda^{max}]$. Because the red curve in Figure 2 is strictly concave over this range, "pooling" the consumers receiving these signals into a single signal—-with the single associated posterior in the range $[0, \lambda^{max}]$—will raise ex-

[10]The logic of Figure 2 can be understood as an application of the "concavication" approach pioneered by Kamenica and Gentzkow (2011).

pected consumer welfare. The appendix shows that it will also increase expected firm welfare. That means that any Pareto optimal info structure can be implemented with two signals: a signal s^{priv} that induces some $\lambda \in [0, \lambda^{max}]$, and a signal s^2 that induces posterior $\lambda = 1$ (and hence is only received with positive probability by the type 2 consumers).

As a third and final step, the appendix shows that consumer welfare is monotonically decreasing and firm welfare monotonically increasing in $\pi_2(s^2)$, so the Pareto frontier is traced out by varying $\pi_2(s^2)$. If the prior $\lambda_2 > \lambda^{max}$, $\pi_2(s^2) = 0$ is Pareto inefficient, and $\underline{\pi} = (1 - p_2^*) > 0$ is the minimum efficient information disclosure; otherwise $\pi_2(s^2) = 0$ is Pareto efficient (and $\underline{\pi} = 0$).

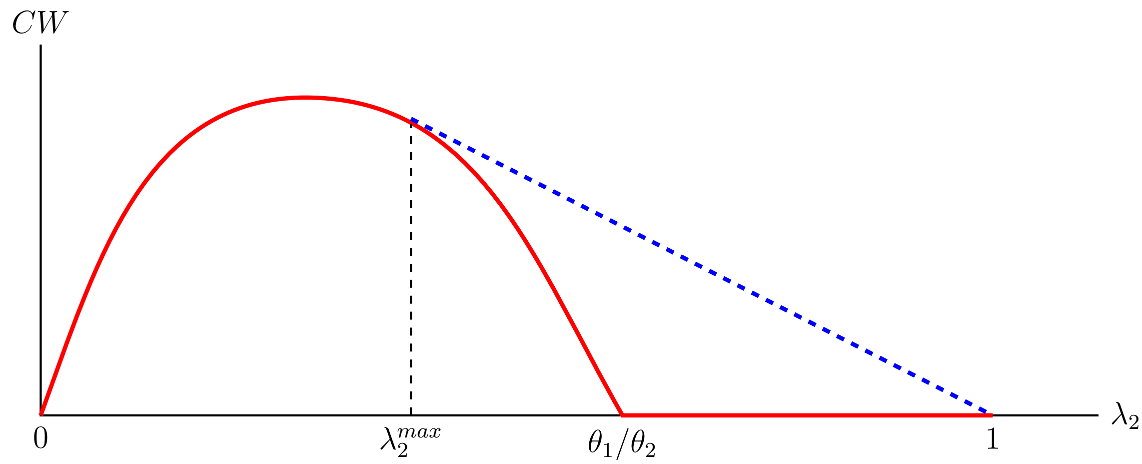

Figure 2: Maximum consumer welfare as a function of the prior λ_2. Depending on the prior, information disclosure can increase or decrease consumer welfare. For instance, for relatively low λ_2, full privacy maximizes consumer welfare but for high λ_2 information disclosure can increase consumer welfare.

3.2 Relationship to the Market Segmentation Literature

Our analysis is closely related to the so-called market segmentation literature following Bergemann et al. (2015; henceforth BBM). BBM study information disclosure policy in a setting with a finite set of consumer types who differ in their privately known willingness-to-pay for a fixed, single product produced by a monopolist.

Absent any information disclosure, the market may be inefficient (as depicted by the circle in Figure 3) insofar as the monopolist finds it optimal to set a price that excludes low-willingness-to-pay types. In such a situation, welfare can be Pareto improved via additional information disclosure. For instance, the upper left-hand point of the shaded triangle corresponds to the attainment of first-best total surplus by a perfectly price discriminating firm who is granted *full* information about consumer willingness to pay. More interestingly, BBM show that a partial information revelation system can be designed in such a way that all efficiency gains accrue to consumers, i.e., that the lower right point of the shaded triangle can also be achieved through information design.

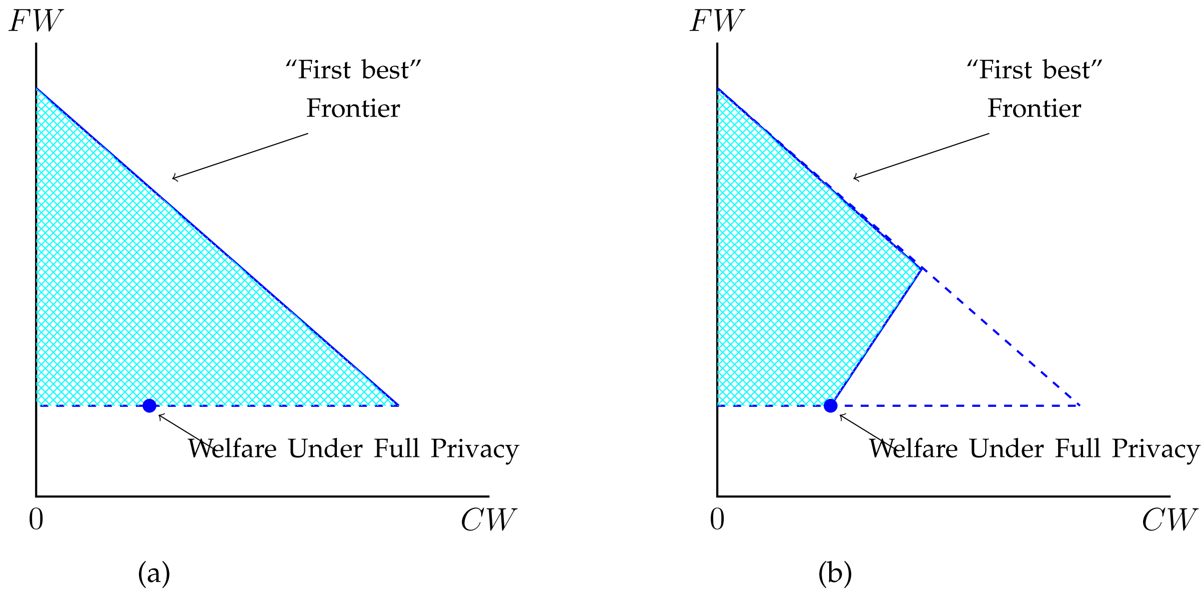

Figure 3: Panel (a) depicts The Surplus Triangle with a Single Product as in Bergemann, Brooks and Morris (2015). Panel (b) depicts The Surplus Area with Multiple Products as in Haghpanah and Siegel (2022,2023).

Haghpanah and Siegel (2022) and Haghpanah and Siegel (2023) extend BBM's analysis to multi-product monopolists. Their qualitative insights are illustrated in Figure 3b. Haghpanah and Siegel (2022) extend BBM by allowing the monopolist to "screen" by offering lower-quality products to some customers. If, under full privacy, the monopolist chooses to screen via such products, they show that the lower right-portion of the BBM triangle is no longer achievable, as depicted in Figure 3b. Nevertheless, and also as depicted in the figure, a significant portion of the first-best frontier is still achievable. Haghpanah and Siegel (2023) show that information disclosure can generically be employed to obtain Pareto improvements, as depicted in the upward-sloping segment of the shaded area in Figure 3b.

Unlike Haghpanah and Siegel (2022, 2023), we allow for a continuum of possible qualities. In particular, with a continuum of qualities, (i) the *only* achievable point on the first-best frontier corresponds to full disclosure and (ii) the existence of Pareto improvements via information disclosure, while possible, is non-generic.

Figure 1a illustrates point (i) directly: the shaded set of achievable surpluses lies strictly below the hypotenuse of the BBM triangle (except at the point on the vertical axis). Although the figure directly illustrates the case where information disclosure can be used to Pareto improve welfare (i.e., where full privacy is Pareto inefficient), it can still be used to illustrate point (ii). If, for example, the welfare under full privacy is on the Pareto frontier, there is no scope for *additional* Pareto-improving information disclosures. It is straightforward to construct analogous cases where the no-disclosure welfare is on the Pareto frontier – and remains so for all nearby models.

The basic intuition for the differences implied by a continuum of types is simple. First, with a continuum of products, screening will *always* be employed by the monopolist, except in the case of full information, in which case the firm will perfectly price discriminate and leave consumers with no welfare. So first-best efficiency is impossible except when firms get all the surplus. An implication is that there are many first-best inefficient informational environments that cannot be Pareto improved upon via information disclosure.

As mentioned in the introduction, also related are the papers by Ichihashi (2020) and Hidir and Vellodi (2021). Ichihashi (2020) considers a setting where the amount of information revelation is chosen by the consumer. He shows that, if given the opportunity, the monopolist prefers to commit to pricing prior to the consumers' information revelation choice, and that this commitment reduces consumer welfare. The key economic difference between our approaches is that we derive the entire Pareto frontier, and hence focus on the tradeoff between firm and consumer welfare that can be achieved by a planner who chooses the information environment, while Ichihashi (2020), in contrast, considers only outcomes that are (constrained) consumer optimal, and hence focuses instead on the tradeoff between firm and consumer welfare induced by firm commitment decisions under a "consumers choose the informational environment" constraint. His model is somewhat distinct than ours (e.g., does not consider continuous screening as we do), but in natural adaptations of his "firms commit to pricing before consumers choose the information environment" to our setting, the tradeoff involved in allowing such commitment is strictly worse than the tradeoff along the Pareto frontier that is traced by the an ex-ante choice of information regime in a no-commitment regime. That is, an appropriate choice (by a notional social planner, say) to partially reveal information is a strictly more efficient way to raise firm welfare than is permitting that firm to commit to pricing and then having consumers to choose the informational environment, a la Ichihashi (2020).

Like us, Hidir and Vellodi (2021) consider a model with continuum of products, but, unlike us, only allow the seller to offer a unique product (which may be conditional on their information). They find that the consumer-optimal market segmentation has an interval structure: types in the same interval are offered the same quality and pay the same price. The key driver of the differences, relative to our paper, in lessons about information disclosure, stems from their assumption of a unique product. Indeed, if, in the setting studied by Hidir and Vellodi (2021) (i.e., a pure horizontal differentiation model), the firm was allowed instead to "screen" via an arbitrary menu of products—as in our paper—it would be able to extract the entire surplus of the consumers, regardless of the information structure, and information disclosure would be irrelevant.

Pram (2021) and Ali et al. (2023) study information disclosure that is optimal for consumers. In particular, Pram (2021) studies an environment which encompasses our environment as a special case. Nonetheless, Pram (2021), as well as Ali et al. (2023), restricts attention to disclosure rules according to which, the consumers can declare in what subset they belong to (lying is not possible). Pram (2021) then shows that consumers can increase their welfare through disclosure if and only if the optimal mechanism without any disclosure excludes some of the types.[11] There are two main differences compared to these papers. First, we allow for any possible information environment and we characterise the entire Pareto frontier. By doing so, we uncover a trade-off between efficiency and consumer welfare not discussed in these papers. Second, if we were to restrict the set of informational environments to those imposed by Pram (2021) in our two-type model,

[11] Pram (2021) is closer to our framework than Ali et al. (2023) because in Ali et al. (2023) the firm offers a unique product and a price unlike our framework in which the firm offers multiple products.

optimal disclosure would take one of two forms: If the (full privacy) mechanism entailed no exclusion, the optimal disclosure would be nil; if, on the other hand, the (full privacy) mechanism entailed exclusion, the outcomes under full privacy and under full disclosure would coincide and information disclosure is irrelevant.

4 Extension to Endogenous Types

4.1 Extension of the Model

We continue to assume that there are only two possible types. In this section, however, we now suppose that the probability distribution over consumer types is endogenously and probabilistically determined by an *unobservable* action taken by the consumer, denoted as $e \in [0,1]$. Specifically, if the consumer takes action e, then her type will be $i = 2$ with probability e or $i = 1$ with probability $1 - e$, and she will incur an effort cost of $\psi(e)$.[12] We assume that $\psi'(0) = 0$ and $\psi'(e) > 0$ for all $e > 0$ and that $\psi''(e) > 0$ for every $e \geq 0$.

We now describe the structure of the game and its equilibrium. The timing is as follows:

0. The informational environment $\Sigma = \left(S, \{\pi_i(s)\}_{i \in \{1,2\}}\right)$ is determined, and is observed by both the consumer and the firm.

1. The consumer chooses her level of effort e.

2. The consumer's type i is realized: it is $i = 2$ with probability e and $i = 1$ otherwise.

3. The consumer's signal s is determined, with signal s sent with probability $\pi_i(s)$ for type i.

4. The firm observes s and chooses a price schedule $t(q; s)$. (Without loss of generality, we assume it must choose $t(0) = 0$).

5. The consumer chooses a quality level q (and corresponding price $t(q; s)$) and utilities are realized.

[12] Equating effort with probability is purely a notational choice: we *define* 'effort' as the probability, and then define $\psi(e)$ as the cost of achieving that probability.

We use *Perfect Bayesian Equilibrium (PBE)* as our solution concept for the game captured in steps 1-5 for any given Σ.

Fixing Σ and (without loss of generality) assuming that it satisfies $\pi_1(s)+\pi_2(s) > 0$ for all $s \in S$, an equilibrium is an effort e^*, a set of posterior beliefs $\beta_1(s)$ and $\beta_2(s)$, a set of pricing menus $t(q;s)$ and a set of choices $q_1(s)$ and $q_2(s)$ such that: $q_i(s)$ is utility maximizing for type-i consumers given $t(q;s)$ for each s; e^* maximizes consumers' expected utility given the menus $\{t(q;s)\}$ and the information structure Σ; the menus $t(q;s)$ are optimal for firms given beliefs and consumer choices $q_H(s)$ and $q_L(s)$; and beliefs are consistent with Bayes' rule and e^*, i.e.

$$\beta_1(s) = \frac{\pi_1(s)(1-e^*)}{\pi_2(s)e^* + \pi_L(s)(1-e^*)} \quad \beta_2(s) = \frac{\pi_2(s)e^*}{\pi_2(s)e^* + \pi_1(s)(1-e^*)}. \tag{13}$$

4.2 Results

The results from the exogenous e case will be helpful in characterizing the Pareto optimal information structures in the case where e is endogenous. Our first step will be to characterize equilibrium e^* for any fixed information structure Σ. To that end, observe that firm beliefs e and Σ together imply $q_1(s)$ and hence consumer welfare under any (potentially different) choice e' of effort:

$$CW(e'|e,\Sigma) = e'\Delta\theta \sum_s \pi_2(s) q_1(s|e,\Sigma) - \psi(e). \tag{14}$$

An equilibrium effort level must both be optimal for consumers and consistent with firm beliefs, so any equilibrium effort must solve

$$\overbrace{\Delta\theta \sum \pi_2(s) q_1(s|\Sigma,e^*)}^{MB(e^*|\Sigma)} - \psi'(e^*) = 0. \tag{15}$$

Since q_1 is non-increasing in e^* for each s, and ψ' strictly increasing in e^*, it follows that there is a unique equilibrium level of effort e^* (and, indeed, $e^* \in (0,1)$).

Together, (14) and (15) together have two strong implications. First, they can be combined to compute equilibrium consumer welfare:

$$CW^*(\Sigma) = e^*(\Sigma)\psi'(e^*(\Sigma)) - \psi(e^*(\Sigma)), \tag{16}$$

which is readily shown to be strictly increasing in $e^*(\Sigma)$. So equilibrium consumer welfare increases in response to a change in the information environment if and only if equilibrium effort increases.

Second, consider any two information environments Σ_1 and Σ_2, with associated equilibria e_1^* and e_2^* and signal probabilities π_i^1 and π_i^2. Then, by (15) – and the fact that its left-hand-side is decreasing in e, which implies that there is a unique equilibrium for any Σ – $MB(e_1^*|\Sigma_2) > MB(e_1^*|\Sigma_1)$ if and only if $e_2^* > e_1^*$, and hence, by the first implication, $MB(e_1^*|\Sigma_2) > MB(e_1^*|\Sigma_1)$ if and only if $CW^*(\Sigma_2) > CW^*(\Sigma_1)$. Moreover, a comparison of the first term in (14) with $MB(e|\Sigma)$ shows that $CW(e_1^*|e_1^*, \Sigma_2) > CW(e_1^*|e_1^*, \Sigma_1)$ if and only if $MB(e_1^*|\Sigma_2) > MB(e_1^*|\Sigma_1)$. Put together, these imply that $CW^*(\Sigma_2) > CW^*(\Sigma_1)$ if and only if $CW(e_1^*|e_1^*, \Sigma_2) > CW(e_1^*|e_1^*, \Sigma_1)$. In other words, equilibrium effort (and hence consumer welfare) increases if and only if a change from Σ_1 to Σ_2 would increase consumer well-being if e were *exogenously* fixed at e_1^*. That is, to check if a change in Σ increases consumer welfare in settings with endogenous effort, it suffices to check if the same change would increase consumer welfare in settings with effort exogenously fixed at the initial equilibrium level.

The preceding is summarized in the following Lemma:

Lemma 1.

1. *For any information structure Σ there is a unique equilibrium $e^*(\Sigma)$.*

2. *For any two information structures Σ_1 and Σ_2:*

$$CW^*(\Sigma_2) \geq CW^*(\Sigma_1) \Leftrightarrow e^*(\Sigma_2) \geq e^*(\Sigma_1) \Leftrightarrow CW(e_1^*|e_1^*, \Sigma_2) \geq CW(e_1^*|e_1^*, \Sigma_1) \quad (17)$$

Since firms will have the same optimal contracts (given by 10) for any given posterior beliefs, their expected welfare depends only on Σ and their beliefs e about effort: we can write $FW(e, \Sigma)$. Equilibrium welfare is $FW^*(\Sigma) = FW(e^*(\Sigma), \Sigma)$. The following Lemma provides a sufficient condition for a change in Σ to increase equilibrium firm welfare. The Lemma follows immediately from the readily established fact that $FW(e, \Sigma)$ is increasing in e (as firms get higher rents from H types).

Lemma 2. *Consider any two information environments Σ_1 and Σ_2. If $e^*(\Sigma_2) \geq e^*(\Sigma_1)$ and $FW(e^*(\Sigma_1), \Sigma_2) \geq FW^*(\Sigma_1)$, then $FW^*(\Sigma_2) \geq FW^*(\Sigma_1)$, and strictly so if the second inequality is strict.*

Together, Lemmas 1 and 2 straightforwardly imply the following Corollary – which will allow us to bootstrap results from the exogenous-e case to the endogenous-e environment.

Corollary 2. *Consider any informational environment Σ_1, with associated effort e_1^*. If Σ_1 is Pareto inefficient in the environment with exogenously given effort e_1^*, then it is also Pareto inefficient in the environment with endogenous effort.*

Towards stating our first main result in the endogenous e setting, we recall and briefly elaborate on our interpretation of Proposition 2: any Pareto optimum except the full-disclosure environment features full privacy to low types. More precisely, any Pareto optimum can be implemented with a two-signal information structure $S = \{s^{priv}, s^2\}$ with $\pi_1(s^{priv}) = 1$ and $\pi_2(s^2) \in [\underline{\pi}, 1]$ for some $\underline{\pi} \geq 0$. The class of Pareto efficient allocations can thus be parameterized via the probability $\pi_2(s^2)$ that the high type will be revealed to the firm. The range of Pareto optimal revelation probabilities is an interval with upper bound at 1, and firm welfare is increasing and consumer welfare decreasing in the revelation probability. The following Proposition shows that *almost* all of these result extend to the endogenous e case – the notable exception being the upper bound of the Pareto efficient revelation probabilities.

Proposition 3 (Optimal Privacy with Endogenous e)**.** *There exists $\underline{\pi}^C \geq 0$ and some $\underline{\pi}^F < 1$ such that the an information structure is Pareto optimal if and only if it is equivalent to a two-signal information structure $S = \{s^{priv}, s^2\}$ with $\pi_1(s^{priv}) = 1$ and $\pi_2(s^2) \in \left[\underline{\pi}^C, \underline{\pi}^F\right]$. Within this class, equilibrium consumer welfare is decreasing and both firm welfare and total (firm plus consumer) welfare are increasing in $\pi_2(s^2)$.*

It is straightforward to construct examples for which $\underline{\pi}^C = 0$ and other examples with $\underline{\pi}^C > 0$. Indeed, examples of both can be found within the class of models with $C(q) = \frac{1}{\gamma+1} q^{\gamma+1}$, $\gamma \geq 1$ and $\psi(e) = \frac{1}{\nu} k e^{\nu}$, $\nu > 1$, $k > 0$.

4.3 Comparative Statics

By Proposition 3, consumer-optimal information disclosure can be fully characterized by $\underline{\pi}^C$, we can think of $1 - \underline{\pi}^C$ as the consumer optimal level of *privacy*. The following Proposition shows that this consumer optimal privacy level has clean and intuitive comparative statics. It applies both with endogenous and exogenous e.

Proposition 4 (Comparative Statics)**.** *The consumer-optimal level of privacy $1 - \underline{\pi}^C$ increases if any one of the following changes occurs:*

1. *ψ' increases (everywhere)*

2. *C' increases and C''/C' weakly increases (everywhere)*

3. *$\Delta\theta$ decreases*

The proposition can be proved with the straightforward mechanical computation which formalizes the following intuition. By Proposition 2, within the class of Pareto optimal information structures, greater privacy is equivalent to greater privacy for type 2 consumers. Increasing the privacy of these types involves a tradeoff: on the one hand, greater privacy implies a higher probability of earning an information rent (a marginal benefit); on the other, greater privacy means greater incentives for the firm to distort q_1 down, hence reducing information rents conditional on privacy (a marginal cost). An increase in ψ' decreases the equilibrium e^* given any privacy level, which reduces the distortions and hence lowers the marginal cost of additional distortions associated with that privacy level (without affecting the marginal benefit of such reductions. Hence, an increase in ψ' makes marginal increases in privacy more desirable. Similarly, an increase in C' decreases e^* (by decreasing consumer's marginal benefit of effort), increasing the marginal cost of additional distortions without affecting the marginal benefit, and making reductions in privacy optimal at the margin. Finally, a greater $\Delta\theta$ increases the distortion at any privacy level, which again increases the marginal cost of additional distortions and reduces consumer-optimal privacy.

5 Discussion and Conclusion

This paper analyzes the effects of privacy on consumer welfare and efficiency in the context of canonical monopoly screening models. Full privacy might or might not be optimal for consumers. When types are exogenous, there is generically a trade-off between consumer welfare and total surplus. In particular, within the class of Pareto efficient allocations, any increase in consumer welfare comes at the expense of total social (firm plus consumer) welfare—a "leaky bucket" result a la Okun (1975). Moreover, this tradeoff is parameterized by the level of privacy: if full privacy is not consumer optimal, then reducing privacy is Pareto improving. Once gains to consumers from reductions in privacy have been exhausted, further decreases in privacy reduce consumer welfare while increasing total welfare. These results all extend to the case in which the consumer's type is endogenously determined by an ex ante action taken by the consumer. Moreover, when types are endogenous, privacy is more important: with exogenous types, total welfare is maximized when consumers have zero privacy; with endogenous types, even the firm-optimal privacy regime involves strictly positive privacy levels, because privacy improves consumers' incentives for undertake costly welfare improving effort.

Although the majority of our analysis focuses on the two-consumer-type case, most of the results are readily extended to many-type settings: our leaky bucket result is general, for instance, as is the fact that full privacy might or might not be consumer optimal and the fact that with endogenous types, Pareto efficiency requires some degree of privacy. The one result which hinges on two types is the simple parameterization of the Pareto frontier in terms of a single-dimensional measure of privacy.

Our paper contributes to the ongoing debate regarding privacy (see Acquisti et al. 2016 and Goldfarb and Que 2023 for surveys). For instance, Stigler (1980) and Posner (1981) argue that privacy inhibits free flow of information and can only be detrimental to efficiency. More recent contributions recognize that privacy bears a value when other market frictions are present. In a related contribution, Hermalin and Katz (2006) argue that withholding information can improve welfare because it may reduce distortions. In their model, privacy can either harm or benefit consumers and total welfare depending on parameter values.

Our analysis shows that optimal information disclosure typically features significant levels of privacy, particularly so when the informed agent can take an unobservable action which increases her willingness to pay for the monopolist's product. Indeed, in these unobservable action settings, our results indicate even firms find a commitment to *full* privacy for the lowest type to be optimal, and a commitment to some degree of privacy for the higher type to be desirable as well. To revisit this implication in some of the applications discussed in section 2.3:

- Monopolistic vendors (e.g., of software, or patent licences) may find it optimal to commit to limits on the research they do on their potential clients – or to cultivate a reputation for limiting such research.

- There may be significant welfare gains from laws which limit monopolists' ability to undertake such research.

- Firms may find it optimal to commit to significant privacy regarding their employees' work efficiency, e.g., by a policy which limits monitoring of their workers at-work behavior.

We conclude by observing that our endogenous type results can also be plausibly interpreted as providing some *normative* support for the epistemic foundations of a Mirrlesian approach to optimal taxation. Mirrleesian models assume–foundationally–that taxpayer types are unobservable to the planner. Insofar as earnings capacity is determined by risky investments in skill building, our results indicate that it may be in the best interests of society to commit to disallowing (or otherwise preventing) the government from observing skill at the interim stage, in order to provide incentives for skill building at the ex-ante stage. In other words, even if skill would in principle be observable to the government, it may be socially desirable to commit *not* to observe it in practice. Notably, this insight applies equally to a society featuring (or, at the ex-ante stage, anticipating) a Rawlsian planner and a society featuring an "extractive" planner – since, at the interim stage, both types of planner will design tax policies to be maximally extractive at high skill levels, where the returns from investments accrue.

A Proofs

A.1 Proof of Proposition 2.

The following simple observation–about a function that is closely related to consumer welfare–will be instrumental in the proof.

Lemma 3. *The function $g(r) \equiv r\xi(\theta_L - r\Delta\theta)$:*

1. *has $g(0) = g\left(\frac{\theta_L}{\Delta\theta}\right) = 0$;*

2. *is strictly concave on the interval $[0, \frac{\theta_L}{\Delta\theta}]$; and*

3. *has a unique maximum at some $r^{max} \in [0, \frac{\theta_L}{\Delta\theta}]$.*

Proof of Lemma 3. Recall that $\xi \equiv C'^{-1}$, $C'(0) = 0$, $C'' > 0$, and $C''' \geq 0$. So part 1 is immediate. So is $\xi' > 0$, $\xi'' < 0$, and hence

$$g''(r) = -2\Delta\theta\xi'(\theta_L - r\Delta\theta) + (\Delta\theta)^2 r\xi''(\theta_L - r\Delta\theta) < 0 \tag{18}$$

and part 2. Part 3 follows directly from parts 1 and 2. □

Towards proving the main Proposition, take any informational environment $\Sigma = \{S, \pi_1(s), \pi_2(s)\}$. Assume, without loss of generality, that $\pi_1(s) + \pi_2(s) > 0$ for all $s \in S$, and define $S^H \equiv \{s|\pi_1(s) = 0\}$ and $S^0 = \{s|\pi_1(s) > 0\}$. All $s \in S^H$ are outcome equivalent (since these signals fully reveal type 2s), so we can, without loss of generality assume that $\#S^H \in \{0, 1\}$. Similarly, any two states $s_A, s_B \in S^0$ with $\frac{\lambda_2\pi_2(s_A)}{(1-\lambda_2)\pi_1(s_A)} = \frac{\lambda_2\pi_2(s_B)}{(1-\lambda_2)\pi_1(s_B)}$ are outcome equivalent, and we can, without loss of generality, take $S^0 = \{s^0, \cdots, s^{n-1}\}$ with $\#S = n \geq 1$ and with $r(s^k) \equiv \frac{\lambda_2\pi_2(s^k)}{(1-\lambda_2)\pi_1(s^k)}$ strictly increasing in k.

Proving the first part of the Proposition amounts to establishing that any information structure with $\#S = n > 1$ is Pareto inefficient. We establish this below by first proving (Lemma 4) that Pareto efficiency requires $r(s) \leq r^{max}$ for all $s \in S^0$, and by second using the concavity of $g(r)$ to prove (Lemma 5) that it is Pareto improving to "collapse" all signals in S^0 into a single signal. Hence, any information structure outside of the class described by the first part of the Proposition is Pareto dominated by some information structure within it.

Having established the first part, the second part of the Proposition follows easily. Indeed, within the class described by the first part of the Proposition (and with $r(s) \leq r^{max}$ per Lemma 4), consumer welfare is $CW = \Delta\theta\lambda_2\pi_2(s^{priv})\xi(\theta_L - r\Delta\theta) = \Delta\theta(1-\lambda_2)g(r)$, which is strictly increasing in r for $r \leq r^{max}$, and hence strictly decreasing in $\pi_2(s^{priv})$. FW, on the other hand, is obviously increasing in $\pi_2(s^{priv})$. So all such structures are Pareto incomparable and hence all Pareto optimal. By Proposition 1, then, TW is increasing in $\pi_2(s^{priv})$. Proving Lemmas 4 and 5, which we turn to now, will thus complete the proof.

Lemma 4. *Any Pareto efficient information structure has* $r(s^{n-1}) \leq r^{max} \equiv \arg\max_r r\xi(\theta_L - r\Delta\theta)$.

Proof of Lemma 4. Consider any information structure with $r(s^{n-1}) > r^{max}$, and take a sufficiently small ε such that $\hat{r}(s^{n-1}) \equiv r(s^{n-1}) - \varepsilon > r^{\max}$ and, if $n > 1$, such that $\hat{r}(s^{n-1}) > r^{n-2}$. Now construct a new information structure $\hat{\Sigma} = \left\{\hat{S}, \hat{\pi}_1(s), \hat{\pi}_2(s)\right\}$ as follows:

1. $\hat{S} = S$ if $\#S^H = 1$ or $\hat{S} = S \cup s^H$ if $\#S^H = 0$,

2. $\hat{\pi}_1(s) = \pi_1(s)$ for all $s \in S^0$ (and hence $\hat{\pi}_1(s^H) = 0$ if $\#S^H = 0$)

3. $\hat{\pi}_2(s^k) = \pi_2(s^k)$ for all $k < n-1$, and $\hat{\pi}_2^{n-1} = \frac{1-\lambda_2}{\lambda_2}\left(r(s^{n-1}) - \varepsilon\right)\pi_1(s^{n-1}) < \pi_2^{n-1}$.

In words, this amounts to changing Σ by taking a small fraction of the H types who received the highest signal s^{n-1} and now revealing their type, so that the posterior type ratio r for that highest signal non-revealing signal decreases to $\frac{\lambda_2\hat{\pi}_2^{n-1}}{(1-\lambda_2)\hat{\pi}_1(s)} = \hat{r}(s^{n-1})$. Firms clearly prefer $\hat{\Sigma}$ to Σ (they have more information). To see that consumers also prefer $\hat{\Sigma}$, notice that the consumer welfare is only affected for the mass $\lambda_2\pi_2(s^{n-1})$ of consumers who have type 2 and would have received signal s^{n-1} under $\hat{\Sigma}$. Their total welfare under Σ is

$$CW = \Delta\theta\lambda_2\pi_2(s^{n-1})\xi(\theta_L - r(s^{n-1})\Delta\theta) = \Delta\theta(1-\lambda_2)\pi_1(s^{n-1})r(s^{n-1})\xi(\theta_L - r(s^{n-1})\Delta\theta). \quad (19)$$

Their total welfare under $\hat{\Sigma}$ is

$$\hat{CW} = \Delta\theta\lambda_2\hat{\pi}_2\hat{\pi}_2(s^{n-1})\xi(\theta_L - \hat{r}(s^{n-1}))\Delta\theta) = \Delta\theta(1-\lambda_2)\pi_1(s^{n-1})\hat{r}(s^{n-1})\xi(\theta_L - \hat{r}(s^{n-1})\Delta\theta). \quad (20)$$

The function $g(r) \equiv r\xi(\theta_L - r(s^{n-1})\Delta\theta)$ is strictly concave and, by definition, maximized at $r^{\max}$. Since $r^{max} < \hat{r}(s^{n-1}) < r(s^{n-1})$, it follows $\hat{CW} > CW$, completing the proof of the lemma.

□

Lemma 5. *Any Pareto efficient information structure has* $\#S^0 = n = 1$.

Proof of Lemma 5. Suppose that the information structure $\Sigma = \{S, \pi_1, \pi_2\}$ has $n = \#S^0 > 1$ and $\#S^H = \#\{s^H\} = 1$.[13] We need to show that it is not Pareto efficient. If $r(s^{n-1}) > r^{max}$ then, by Lemma 4, it is not Pareto efficient, so we need only to establish Pareto inefficiency for the $r(s^{n-1}) \leq r^{max}$ case. In that case, construct $\hat{\Sigma} = \{\hat{S}, \hat{\pi}_1, \hat{\pi}_2\}$ as follows:

- $\hat{S} = \{s^{priv}, s^H\}$
- $\hat{\pi}_1(s^{priv}) = \sum_{s \in S^0} \pi_1(s) = 1$; $\hat{\pi}_1(s^H) = \pi_1(s^H) = 0$
- $\hat{\pi}_2(s^{priv}) = \sum_{s \in S^0} \pi_2(s) = 1 - \pi_2(s^H)$; $\hat{\pi}_2(s^H) = \pi_2(s^H)$.

In words: construct $\hat{\Sigma}$ by collapsing the signals in S^0 into a single signal.

We will now show that $CW(\hat{\Sigma}) > CW(\Sigma)$ and $TW(\hat{\Sigma}) > TW(\Sigma)$. By Corollary 1, this will complete the proof that Σ is not Pareto inefficient. Before proceeding, observe that:

$$\begin{aligned}\mathbb{E}_{\pi_1}[r] &\equiv \sum_{s^k \in S^0} r(s^k)\pi_1(s^k) \\ &= \sum_{s^k \in S^0} \frac{\lambda_2 \pi_2(s^k)}{(1-\lambda_2)\pi_1(s^k)} \pi_1(s^k) = \frac{\lambda_2(1-\pi_2(s^H))}{1-\lambda_2} \equiv \hat{r}(s^{priv}). \end{aligned} \tag{21}$$

That is, the expectation of the posterior ratio r with respect to the distribution of type 1's under Σ coincides with the (unique) posterior ratio under $\hat{\Sigma}$.

Consumer welfare under the two information structures is given by:

$$\begin{aligned} CW(\hat{\Sigma}) &= \lambda_2 \Delta\theta \hat{\pi}_2(s^k) \xi(\theta_1 - \hat{r}(s^{priv})\Delta\theta) \\ &= (1-\lambda_2)\Delta\theta \hat{r}(s^{priv}) \xi(\theta_1 - \hat{r}(s^{priv})\Delta\theta) \end{aligned} \tag{22}$$

[13] The assumption that $\#S^H = 1$ is purely for expositional simplicity; the same exact logic would apply if $\#S^H = 0$.

and

$$
\begin{aligned}
CW(\Sigma) &= \lambda_2 \Delta\theta \sum_{s^k \in S^0} \pi_2(s^k)\xi(\theta_1 - r(s^k)\Delta\theta) \\
&= (1-\lambda_2)\Delta\theta \sum_{s^k \in S^0} \pi_1(s^k) r(s^k)\xi(\theta_1 - r(s^k)\Delta\theta) \\
&= (1-\lambda_2)\Delta\theta \mathbb{E}_{\pi_1}\left[r\xi(\theta_1 - r\Delta\theta)\right] \\
&< (1-\lambda_2)\Delta\theta \hat{r}(s^{priv})\xi(\theta_1 - \hat{r}(s^{priv})\Delta\theta) = CW(\hat{\Sigma}),
\end{aligned} \tag{23}
$$

where the key inequality follows from Lemma 3 and Jensen's inequality (and equation 21).

Total welfare under the two information structures is given by:

$$
TW(\hat{\Sigma}) = (1-\lambda_2)[\theta_1 \xi(\theta_1 - \hat{r}(s^{priv})\Delta\theta) - C(\xi(\theta_1 - \hat{r}(s^{priv})\Delta\theta))] + K, \tag{24}
$$

where $K \equiv \lambda_2 \left(\theta_2 \xi(\theta_2) - C(\xi(\theta_2))\right)$, and, defining $h(x) \equiv \theta_1 x - C(x)$,

$$
\begin{aligned}
TW(\Sigma) &= (1-\lambda_2) \sum_{s^k \in S^0} \pi_1(s^k)(\theta_1 \xi(\theta_1 - r(s^k)\Delta\theta) - C(\xi(\theta_1 - r(s^k)\Delta\theta)) + K \\
&= (1-\lambda_2)\mathbb{E}_{\pi_1} h(\xi(\theta_1 - r\Delta\theta)) + K \\
&< (1-\lambda_2) h\left(\mathbb{E}_{\pi_1}\left[\xi(\theta_1 - r\Delta\theta)\right]\right) + K \\
&< (1-\lambda_2) h\left(\xi(\theta_1 - \hat{r}(s^{priv})\Delta\theta)\right) + K = TW(\hat{\Sigma}),
\end{aligned} \tag{25}
$$

where the first inequality follows from Jensen's inequality and the strict concavity of the function $h(x) = \theta_1 x - C(x)$, and the second inequality follows from the fact that $h(x)$ is strictly increasing in x and from a second application of Jensen's inequality (using the strict concavity (in r) of $\xi(\theta_1 - r\Delta\theta)$). □

A.2 Proof of Corollary 2.

If Σ_1 is Pareto inefficient in the environment with exogenously given effort e^*, then there is another information structure Σ_2 such that $FW(e_1^*, \Sigma_2) \geq FW^*(\Sigma_1)$ and $CW(e_1^*|e_1^*, \Sigma_2) \geq CW(e_1^*|e_1^*, \Sigma_1)$, with at least one inequality strict. By Lemma 1, $e_2^* \geq e_1^*$. Hence, by Lemma 2, $FW^*(\Sigma_2) \geq FW^*(\Sigma_1)$, and strictly so if $FW(e_1^*, \Sigma_2) > FW^*(\Sigma_1)$. Also by Lemma 1, $CW^*(\Sigma_2) \geq CW^*(\Sigma_1)$, and strictly so if $CW(e_1^*|e_1^*, \Sigma_2) > CW(e_1^*|e_1^*, \Sigma_1)$.

A.3 Proof of Proposition 3.

Consider any information environment Σ_1 that is not equivalent to a two-signal information structure $S = \{s^{priv}, s^H\}$ with $\pi_L(s^{priv}) = 1$, and let e_1^* be the associated equilibrium effort. Then, by Proposition 2, Σ_1 would be Pareto dominated by some other structure Σ_2 in an environment with e exogenously equal to e_1^*. Hence, by Corollary 2, Σ_2 Pareto dominates Σ_1 with endogenous effort. So *only* information structures that are equivalent to a two-signal information structure can be Pareto optimal.

Moreover, within this class (and with some mild notational abuse) $CW(e|e, \pi_H(s^H) = 1) = 0$ for any e, and hence $e^*(\pi_H(s^H) = 1) = 0$, and hence $CW^*(\pi_H(s^H) = 1) = 0 < CW^*(\pi_H(s^H) = 0)$ and $FW^*(\pi_H(s^H) = 1) = \theta_L \xi(\theta_L) - C(\xi(\theta_L)) < FW^*(\pi_H(s^H) = 0)$. Hence $\pi_H(s^H) = 1$ is strictly Pareto dominated by, e.g., $\pi_H(s^H) = 0$, and by continuity, so is $\pi_H(s^H) = 1 - \varepsilon$ for sufficiently small ε. Thus, there is some supremal Pareto efficient $\pi_H(s^H)$ that is strictly less than 1.

We omit the straightforward (e.g. computational) proof that the set of Pareto optimal $\pi_H(s^H)$ takes an interval form and of the comparative statics of consumer welfare and firm welfare, and only note that the Leaky Bucket Lemma 1 applies and implies the comparative statics of total welfare.

References

ACQUISTI, A., C. TAYLOR, AND L. WAGMAN (2016): "The economics of privacy," *Journal of economic Literature*, 54, 442–492.

ALI, S. N., G. LEWIS, AND S. VASSERMAN (2023): "Voluntary Disclosure and Personalized Pricing," *The Review of Economic Studies*, 90, 538–571.

BENNETT, C. J. (2010): *The privacy advocates: Resisting the spread of surveillance*, Mit Press.

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The limits of price discrimination," *American Economic Review*, 105, 921–957.

BERGEMANN, D. AND S. MORRIS (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95.

GOLDFARB, A. AND V. F. QUE (2023): "The Economics of Digital Privacy," *Annual Review of Economics*, 15.

GOLDFARB, A. AND C. TUCKER (2019): "Digital economics," *Journal of economic literature*, 57, 3–43.

HAGHPANAH, N. AND R. SIEGEL (2022): "The Limits of Multiproduct Price Discrimination," *American Economic Review: Insights*, 4, 443–458.

——— (2023): "Pareto-Improving Segmentation of Multiproduct Markets," *Journal of Political Economy*, 131, 000–000.

HERMALIN, B. E. AND M. L. KATZ (2006): "Privacy, property rights and efficiency: The economics of privacy as secrecy," *Quantitative marketing and economics*, 4, 209–239.

HIDIR, S. AND N. VELLODI (2021): "Privacy, personalization, and price discrimination," *Journal of the European Economic Association*, 19, 1342–1363.

ICHIHASHI, S. (2020): "Online privacy and information disclosure by consumers," *American Economic Review*, 110, 569–595.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

MIRRLEES, J. A. (1971): "An exploration in the theory of optimum income taxation," *The review of economic studies*, 38, 175–208.

MUSSA, M. AND S. ROSEN (1978): "Monopoly and product quality," *Journal of Economic theory*, 18, 301–317.

OKUN, A. (1975): *Equality and Efficiency: The Big Tradeoff*, Washington DC: Brookings.

POSNER, R. A. (1981): "The economics of privacy," *The American economic review*, 71, 405–409.

PRAM, K. (2021): "Disclosure, welfare and adverse selection," *Journal of Economic Theory*, 197, 105327.

STIGLER, G. J. (1980): "An introduction to privacy in economics and politics," *The Journal of Legal Studies*, 9, 623–644.

STIGLITZ, J. E. (1982): "Self-selection and Pareto efficient taxation," *Journal of public economics*, 17, 213–240.

ZUBOFF, S. (2023): "The age of surveillance capitalism," in *Social Theory Re-Wired*, Routledge, 203–213.