Screening, Investment Incentives and the Welfare Trade-offs of Privacy

Anastasios Dosis *   Casey Rothschild†

February 25, 2025

Abstract

This paper explores how information disclosure influences incentives to innovate and overall welfare. Using a class of principal-agent models, we characterize Pareto-optimal privacy regimes that balance efficiency gains with strategic incentives. Our findings reveal a fundamental trade-off: while information disclosure can enhance efficiency, it may also reduce R&D investment due to surplus appropriation risks. We delineate conditions under which full privacy, full disclosure, or intermediate privacy are Pareto optimal. These insights inform intellectual property and innovation policies, highlighting when privacy fosters or hinders economic efficiency and innovation.

*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

How does information disclosure affect incentives to innovate? Consider, for example, a university holding a patent on a groundbreaking discovery. The university, as the patent holder, can license this innovation to a company for commercialization. However, the success of the commercialized product depends on the company's investment in R&D—greater investment increases the likelihood of success and, consequently, potential profits. The patent holder, exercising monopoly power over the breakthrough, must determine the scope of the licensing agreement (extensive vs. limited) and set appropriate royalty fees.[1]

A key question that arises in this setting is: how much access to information about the product's commercial value should the patent holder have? Under full privacy, only the company knows the product's commercial potential, which may prevent the patent holder from fully appropriating the surplus but could also exacerbate ex post inefficiencies that distort incentives. Conversely, under full disclosure, the patent holder has perfect information, mitigating these inefficiencies but potentially reducing the company's incentives to invest in R&D due to the risk of surplus appropriation.

This paper examines this tradeoff by characterizing Pareto-optimal privacy regimes in a broad class of settings. In these settings, there is both an efficiency rationale for revealing information and a strategic incentive for innovators to maintain privacy—namely, to prevent exploitation based on disclosed information.[2] Our framework allows us to: (i) identify contexts where additional information revelation improves efficiency and others where it worsens it; (ii) determine when trade-offs arise between economic efficiency and equity; and (iii) derive insights into the types of information disclosure policies that yield constrained-efficient outcomes.

[1]One real-world example that fits this story is the classic case study of the University of California (UC) and Genentech, a biotechnology firm. In 1976, the University of California (UC) licensed its groundbreaking recombinant DNA technology—pioneered by researchers Herbert Boyer and Stanley Cohen at UCSF—to Genentech, a newly founded biotechnology firm. This agreement was innovative for its time. It granted Genentech exclusive rights to commercialize the technology while UC retained royalty rights on future products. For more details see

[2]We abstract from intrinsic preferences for privacy; in our analysis, privacy improves outcomes through its economic effects rather than providing inherent utility to agents.

We analyze a broad class of settings in which a (potentially) uninformed principal with monopoly power interacts with an informed agent. Although our framework is quite general, for expository clarity we focus on a concrete example: a monopolistic firm that sells a product to a consumer who privately knows her own willingness-to-pay for quality. In our motivating example, the principal is the patent holder, and the agent is the company responsible for commercializing the invention.

Prior to the interaction, a planner determines the extent of "privacy" in the market. Specifically, the planner sets the degree to which the firm can observe the consumer's willingness-to-pay type. Once the information regime is established and observed by both parties, the agent undertakes a costly action that probabilistically determines her willingness to pay. The principal then offers a menu of contracts that is contingent on the information disclosed, while remaining unaware of the agent's specific action at the time of the offer.

To build intuition and address our central questions—namely, how information affects both the incentives to innovate and overall welfare, we begin with a baseline case in which effort is exogenously given and known to the principal. This simpler setting allows us to later bootstrap our insights to the more complex case in which effort is endogenous and privately observed by the agent.

Our first result (Lemma 2) establishes a "leaky bucket" phenomenon reminiscent of Okun (1975): there is an inherent tradeoff between consumer welfare and efficiency. Information disclosure can initially yield Pareto improvements; however, once these gains are exhausted, any additional disclosure can only improve efficiency at the expense of consumer welfare. This tradeoff emerges from the dual effects of information: while it enables the firm to better tailor its products to consumer needs (thereby improving matching), it simultaneously enhances the firm's ability to extract consumer surplus.[3] In effect, although consumers might prefer less refined information, such a regime can reduce overall welfare by constraining total efficiency.

Building on Lemma 2, we show (Lemma 3) that any Pareto optimal information en-

[3]A similar trade-off is also highlighted in the contributions by Ichihashi (2020) and Hidir and Vellodi (2021), who study optimal market segmentation in a multi-product monopoly setting.

vironment exhibits a straightforward structure: the low-valuation type enjoys full privacy, while the high-valuation type benefits from partial privacy (Lemma 5). We further demonstrate that full privacy may or may not be Pareto optimal. When the agent obtains high rents under full privacy, increasing disclosure raises the potential ex post rents (if the agent's type remains private) but also reduces the likelihood of capturing those rents. In certain parameter configurations, the latter effect dominates, making full privacy Pareto optimal. Conversely, under alternative parameter configurations, full privacy is not Pareto optimal: relinquishing some privacy can improve outcomes for both the firm and the consumer by mitigating quality distortions.[4]

Finally, we establish that along the Pareto frontier, increasing privacy (parameterized by the probability with which the high-valuation type remains unknown to the principal) enhances consumer welfare but simultaneously reduces total welfare.

We then examine how our results change when effort is endogenized. First, we characterize the unique equilibrium level of effort (Lemma 6). Building on this, a series of intermediate results (Lemmas 7 and 8) allow us to extend our findings from the exogenously fixed effort case to the endogenous setting. In particular, Proposition 1 demonstrates that if an information environment σ with its associated equilibrium effort $e^*(\sigma)$ is not Pareto optimal under fixed effort, then it cannot be Pareto optimal when effort is allowed to vary. Consequently, we establish (Corollary 2) that any Pareto optimal information environment retains the simple structure identified in the exogenous case.

Finally, one of our main results (Proposition 2) shows that among Pareto optimal allocations, an information environment is preferable to the agent (and less so to the principal) if and only if it involves greater privacy. This finding has significant implications for the trade-offs involved in strengthening privacy.

Our last exploration is on whether full privacy and full disclosure are Pareto optimal. We identify conditions under which full privacy is not Pareto optimal. Proposition 3 states that when the consumer already has a high probability of being high type or when the cost of exerting effort to increase her value is sufficiently low, full privacy is not Pareto

[4]This analysis is related to the broader literature on the trade-off between consumer identification for improved product matching (or targeted advertising) and the risk of price discrimination; see Goldfarb and Tucker (2019) for an extensive survey.

optimal. Indeed, when the cost of exerting effort is relatively low, the consumer will exert high effort, which tends to reduce the rents she receives, thereby lowering her incentive to exert effort. Consequently, moving away from full privacy provides the consumer with higher rents, which in turn increases her effort. Given that the firm prefers both higher effort and more disclosure, an environment with less than full privacy becomes Pareto optimal.

We also look whether full disclosure is Pareto optimal. Proposition 4 establishes that, regardless of the initial effort level, if the marginal cost of effort is zero at the original point is zero, full disclosure cannot be Pareto optimal. 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.

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 fully characterize Pareto optimal information disclosure in a two-type version of the model. In Section **??**, 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 implications and concludes.

2 Model

2.1 Baseline Model

Consider a firm (the principal) selling a customized product to a consumer (the agent). The product's quality is indexed by $q \in [0, \infty]$. The firm maximizes its expected profit, with a profit function given by:

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

where $t \in \mathbb{R}$ represents a transfer. The cost function $C(\cdot)$ satisfies the following assumptions: $C(0) = C'(0) = 0$, $C'(q) > 0$, $C''(q) > 0$ for $q > 0$, and $C'''(q) \geq 0$ for all $q \geq 0$.[5] The firm has monopoly power in that it can make a take-it-or-leave-it offer to the consumer as described below.

At the time the consumer contemplates a purchase, she has preferences given by

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

Per Equation (2), θ_i measures the consumer's (per unit) valuation of the product. We assume that there are two possible types, $i \in \{H, L\}$. The high type ($i = H$) has a higher (per unit) valuation for the product than the low type ($i = L$), so

$$\Delta\theta \triangleq \theta_H - \theta_L > 0.$$

For simplicity, we assume that both types have an outside option that provides them with zero payoff at any stage of the game (described below).

The probability distribution over the two consumer types is endogenously determined by an *unobservable* action taken by the consumer, denoted by $e \in [e_0, 1]$, where $e_0 \in [0, 1)$. Specifically, if the consumer takes action e, her type will be $i = H$ with probability $l_H(e) \equiv e$ or $i = L$ with probability $l_L(e) \equiv 1 - e$. Effort is costly and the consumer incurs an effort cost of exerting effort e equal to $\psi(e)$.[6] We assume that $\psi'(0) = 0$, $\psi'(e) > 0$ for all $e > 0$, and $\psi''(e) > 0$ for every $e \geq 0$.

In the example mentioned in the introduction, the seller is the patent holder who offers a licensing contract to the buyer (the company) seeking to commercialize an invention. In this example, q is the extent of licensing: higher q implies a broader license. More extensive licensing comes at a higher cost for the patent holder—reflecting lost exclusivity and foregone future rents—while it provides the buyer with greater commercial value through enhanced integration opportunities. Crucially, the buyer's commercial success

[5]The condition $C'''(q) \geq 0$ ensures that the optimal quality is a concave function of the parameters and is satisfied by commonly used cost functions such as $C(q) = \frac{1}{\gamma}q^{\gamma}$ for $\gamma \geq 2$.

[6]Equating effort with probability is purely a notational choice. Alternatively, we could allow the consumer to take action $a \in [0, \infty)$ at a unit per-unit cost. By taking action a, her type would be $i = H$ with probability $p(a)$, and $i = L$ with probability $1 - p(a)$, where $p'(a) > 0$ and $p''(a) < 0$, with $p(0) = p_0 \in [0, 1)$ and $p(\infty) = 1$.

depends not only on q but also on its unobservable R&D effort. Increased effort improves the per-unit value of the invention by enabling better adaptation and utilization of the technology.

2.2 Information, Timing and Equilibrium

We follow the Bayesian persuasion Kamenica and Gentzkow (2011) and information design Bergemann and Morris (2019) literatures and model an *informational environment* as a pair $\sigma = (\Omega, \boldsymbol{\mu})$, where Ω is a finite set of states and $\boldsymbol{\mu} = \{\mu_i\}_{i \in \{H,L\}}$ is a pair of mappings, with $\mu_i : \Omega \to [0,1]$ satisfying $\sum_\omega \mu_i(\omega) = 1$ for each i and (without loss of generality) $\sum_i \mu_i(\omega) > 0$ for all $\omega \in \Omega$.[7]

This modeling approach allows for two key interpretations. In the classic Bayesian persuasion and information design framework, Ω represents a set of signals, and $\mu_i(\omega)$ denotes the conditional probability that type i sends signal ω. Alternatively, following Bergemann et al. (2015), Ω can be interpreted as a set of market segments, where

$$\nu_i(\omega) \triangleq l_i(e)\mu_i(\omega) \tag{3}$$

represents the share of type i in segment ω. Regardless of the interpretation, knowing e and using Bayes' rule, the conditional probability[8] that the type is i after observing signal ω is:

$$\lambda_i(\omega) \triangleq \frac{\nu_i(\omega)}{\nu_H(\omega) + \nu_L(\omega)}. \tag{4}$$

It is also useful to define the likelihood ratio of high-type to low-type consumers as:

$$\kappa(\omega) \triangleq \frac{\nu_H(\omega)}{\nu_L(\omega)}.^9 \tag{5}$$

Two important informational environments are *full privacy* and *full disclosure*. Under full privacy, denoted by σ_{priv}, every signal is fully uninformative, whereas under full

[7] Although, for exposition purposes, we focus on a finite set Ω in the main text, in the appendix we show that this assumption is without loss of generality. The assumption that $\sum_i \mu_i(\omega) > 0$ for all $\omega \in \Omega$ is also without loss of generality because signals that are sent with probability zero by both types are never realized and are therefore irrelevant.

[8] Note that our assumption that $\sum_i \mu_i(\omega) > 0$ for all $\omega \in \Omega$ ensures that this probability is well defined.

[9] Although clearly $\nu_i(\omega)$, $\lambda_i(\omega)$ and $\kappa(\omega)$, apart from ω, all depend on σ and e, to ameliorate the notational burden, we refrain from adding the arguments.

disclosure, denoted by σ_{disc}, every signal is fully informative.[10]

We now describe the structure of the game and its equilibrium. The timing of events is presented in Figure **??**.

- **Stage 0:** The informational environment σ is determined and observed by both the consumer and the firm.
- **Stage 1:** The consumer (privately) chooses her level of effort e.
- **Stage 2:** The consumer's type i is realized with probability $l_i(e)$.
- **Stage 3:** Signal ω is realized with probability $\mu_i(\omega)$ if the realized type is i.
- **Stage 4:** The firm observes ω and chooses a menu of contracts $\{(t_i(\omega), q_i(\omega))\}_{i\in\{H,L\}}$.
- **Stage 5:** The consumer chooses a contract from the menu or her outside option, and the payoffs are realized.

Figure 1: The timing of events.

For any given informational environment, we apply *Perfect Bayesian Equilibrium (PBE)* as our solution concept for the game. Fixing the informational environment, an equilibrium consists of: (i) a level of effort; (ii) a set of posterior beliefs; (iii) a menu of contracts; and (iv) a pair of contract choices (one for each type). These satisfy the following conditions: the equilibrium level of effort maximizes the consumer's ex-ante payoff, the set of contracts is incentive-compatible, individually rational for each type, and maximizes the firm's ex-ante profit given its beliefs for each realized signal, and the firm's beliefs are consistent with Bayes' rule and the equilibrium level of effort.

[10]Full privacy can be implemented by any information structure in which $\mu_i(\omega) = \mu_j(\omega)$ for every $i, j \in \{H, L\}$ and every $\omega \in \Omega$. Conversely, full disclosure can be implemented by an information structure in which Ω consists of two elements ω and ω' such that $\mu_H(\omega) = \mu_L(\omega') = 1$ and $\mu_H(\omega') = \mu_L(\omega) = 0$.

3 Results

In presenting our results, we proceed as follows. First, we characterize the optimal contracts offered by the firm for any informational environment and every realized signal. We then examine the case in which the effort level is exogenously fixed and hence known by the firm. In particular, for any fixed level of effort, we characterize the set of Pareto-optimal informational environments. Next, we demonstrate that many of the results established when effort is exogenously given can be bootstrapped to the case in which the effort level is endogenous. Specifically, we characterize the set of Pareto-optimal informational environments and provide conditions under which full privacy, partial privacy, or full disclosure are Pareto optimal.

3.1 Preliminaries

Suppose that the informational environment is given by σ and the effort exerted by the consumer is given by e. For a realized signal ω, the firm updates its beliefs as per (4). Given that the firm has monopoly power, it offers a menu of contracts that solves the following standard maximization program:

$$\begin{aligned}
\max_{\{(q_H,t_H),(q_L,t_L)\}} \quad & \lambda_H(\omega)\pi(q_H,t_H) + \lambda_L(\omega)\pi(q_L,t_L) \\
s.t. \quad & u_H(q_H,t_H) \geq u_H(q_L,t_L) \\
& u_L(q_L,t_L) \geq u_L(q_H,t_H) \\
& u_H(q_H,t_H) \geq 0 \\
& u_L(q_L,t_L) \geq 0,
\end{aligned}$$

where $\pi(\cdot,\cdot)$ and $u_i(\cdot,\cdot)$ are defined in (1) and (2) respectively.

The following lemma summarizes the firm's profit-maximizing contracts and the payoffs of the two players.

Lemma 1. *For any informational environment σ, effort level e, and realized signal ω, the firm* *offers profit-maximizing qualities* $\{q_i(\omega)\}_{i\in\{H,L\}}$ *such that*

$$q_H(\omega) \triangleq \xi(\theta_H)$$
$$q_L(\omega) \triangleq \max\{0, \xi(\theta_L - k(\omega)\Delta\theta)\}, \tag{6}$$

where $\xi(x) \equiv C'^{-1}(x)$ *is the inverse of the firm's marginal cost function, and* $\kappa(\omega)$ *is the likelihood ratio as per (5).*

The ex-post payoffs of the two types are given by

$$u_H(\omega) \triangleq \Delta\theta q_L(\omega)$$
$$u_L(\omega) \triangleq 0. \tag{7}$$

Lemma 1 follows standard results: in the firm's profit-maximizing contracts, there is no distortion in the quality offered to type H, while type L faces a downward distortion (or even a potential shutdown). Type H earns a positive information rent, whereas type L earns no rent.

The ex-ante payoff of the consumer is obtained by taking the expectation over ω. Since the consumer earns information rents only when she is of type H, her ex-ante payoff is given by:

$$U(\sigma|e) \triangleq l_H(e) \sum_{\omega} \mu_H(\omega) u_H(\omega) - \psi(e).\text{[11]} \tag{8}$$

Ex post, the firm earns the surplus generated by each type minus the rent paid to that type. For notational simplicity, let

$$S_i(\omega) \triangleq \theta_i q_i(\omega) - C(q_i(\omega)) \tag{9}$$

denote the surplus generated by type i when the signal is ω. Then the firm's ex-ante profit is found by taking the expectation over ω and i:

$$\Pi(\sigma|e) \triangleq \sum_i l_i(e) \sum_{\omega} \mu_i(\omega) \left(S_i(\omega) - u_i(\omega)\right). \tag{10}$$

Finally, ex-ante total welfare is found by summing up (8) and (10):

$$W(\sigma|e) \triangleq U(\sigma|e) + \Pi(\sigma|e) = \sum_i l_i(e) \sum_{\omega} \mu_i(\omega) S_i(\omega) - \psi(e). \tag{11}$$

3.2 Exogenous Effort

Although our goal is to study how information affects the consumer's incentives to exert effort, and more precisely to characterize the set of Pareto optimal informational environments, we begin by assuming that the effort level is exogenously given and known to the firm. This assumption helps build intuition. Moreover, we later show how many of the results derived under exogenous effort can be bootstrapped to the case where effort is endogenous.

We first provide a formal definition of Pareto optimal informational environments when effort is exogenously given.

Definition 1 (Pareto Optimality Under Exogenous e)**.** *Suppose that the effort level is exogenously given and equal to e. An informational environment σ is Pareto optimal if there does not exist another informational environment σ' such that*

$$U(\sigma'|e) \geq U(\sigma|e) \quad \textit{and} \quad \Pi(\sigma'|e) \geq \Pi(\sigma|e),$$

with at least one of the two inequalities being strict.

For any exogenously given effort, the maximum total welfare is achieved under full disclosure. Therefore, full disclosure is Pareto optimal.[12] Indeed, in this case, there are no distortions. Any informational environment other than full disclosure introduces a distortion for type L, which arises from the firm's incentive to screen consumers. This distortion is socially costly and, therefore, reduces total welfare. Given this observation and the convexity of the set of achievable payoffs for both the firm and the consumer (i.e., the Pareto frontier is convex), increasing the consumer's payoff necessarily decreases overall welfare, as formally stated in the following lemma.

Lemma 2 (The Leaky Bucket)**.** *For any exogenously given e:*

1. *If $U(\sigma|e) > 0$, then $W(\sigma|e) < W(\sigma_{disc}|e)$.*

2. *If σ and σ' are both Pareto optimal and $U(\sigma'|e) > U(\sigma|e)$, then $W(\sigma'|e) < W(\sigma|e)$.*

[12] Note that this is not necessarily true when effort is endogenous as full disclosure provides no information rents to the consumer and hence no incentives to increase e.

Lemma 2 has two important implications. First, under exogenous effort, any Pareto optimal informational environment that provides consumers with strictly positive surplus necessarily reduces total welfare. In other words, only full disclosure achieves first-best total welfare. Second, any Pareto optimal informational environment that increases consumer welfare must necessarily decrease total welfare. In Section 5, we discuss the implications of this result in the market segmentation literature (Bergemann et al. 2015).

Next, we analyze the structure of Pareto optimal informational environments in greater detail. Although we initially allow for general information structures, the following lemma establishes that Pareto optimal informational environments have a particularly simple form.

Lemma 3 (Optimal Information Structures)**.** *An informational environment is Pareto optimal if and only if it is equivalent to a two-signal informational environment* $\Omega = \{\omega_H, \omega_{LH}\}$*, where* $\mu_L(\omega_H) = 0$ *and* $\mu_H(\omega_{LH}) \equiv \mu \in [0, \overline{\mu}]$ *, for some* $\overline{\mu} \in (0, 1]$.

This result follows in two steps. First, note that any informational environment induces a distribution over posterior beliefs $\lambda_H(\omega)$. For each such belief, there exists a unique firm-optimal quality $q_L(\omega)$ and, consequently, a conditional surplus $u_H(\omega)$ for type H when they receive signal ω. If $\lambda_H(\omega)$ is sufficiently high, the firm sets $q_L(\omega) = 0$, meaning that the consumer earns no surplus upon receiving signal ω. Such signals cannot belong to the support of the posterior beliefs induced by a Pareto optimal informational environment. To see why, consider "splitting" such a signal into two new signals, ω_A and ω_B, by revealing the identity of a portion of type H individuals who originally received signal ω. This implies $\lambda_H(\omega_B) = 1$ and $0 < \lambda_H(\omega_A) < \lambda_H(\omega)$. Since revealing more information cannot be detrimental to the firm, and if $\lambda_H(\omega_A)$ is low enough for the firm to set $q_L(\omega_A) > 0$, then type H consumers who receive ω_A obtain positive surplus. This suggests that any informational environment containing signals with $\lambda_H(\omega) < 1$ and $q_L(\omega) = 0$ is not Pareto optimal.

The second step leverages concavity properties of $U(\sigma|e)$ and $W(\sigma|e)$ to show that merging all signals for which $q_L(\omega) \neq 0$ into a single signal improves both consumer and total welfare. Therefore, because an immediate corollary of Lemma 2 is that if for any informational environment there is another one increases the consumer's payoff as well as total welfare the former cannot be Pareto optimal, an information structure inducing a non-degenerate distribution over such signals is not Pareto optimal.

Taken together, these two steps imply that any Pareto optimal informational environment must consist of exactly two signals: one sent exclusively by type H, and another in which type H and type L are pooled such that the firm sets a positive $q_L(\omega)$.

Lemma 3 allows us to parametrize privacy (equivalently, disclosure) in terms of the fraction of type H consumers who are pooled with type L in equilibrium. Therefore, privacy can be parametrized by μ with higher μ implying higher privacy (less disclosure), $\mu = 0$ implying full disclosure and $\mu = 1$ implying full privacy.[13]

In light of Lemma 3, the main question now becomes how the payoffs of the firm and the consumer depend on μ. It is not difficult to show that, when effort is exogenously given, more privacy hurts the firm. Indeed, consider any level of privacy and suppose that this decreases slightly. The firm can still offer the same contracts upon observing signal ω_{LH} and have more types sending signal ω_H. Therefore, it will have a higher profit.[14] This is stated in the following lemma.

Lemma 4 (Firm's Optimal Privacy)**.** *For any exogenously given e, firm's profit is decreasing in the level of privacy.*

Regarding the consumers, the level of privacy depends on the level of effort (which recall that in this section is exogenously given). The easiest way to wit is to consider the case in which under full privacy type L is shut down. This happens when e is sufficiently high (i.e., higher than $\frac{\theta_L}{\theta_H}$). In this case, slightly moving away from full privacy will give consumers some information rents. This will be Pareto improving, as in light of Lemma 4 less privacy is clearly good for the firm when effort is exogenous, and also for the consumer. The following result characterizes the consumer-optimal and Pareto optimal levels of privacy for any e.

[13]Note also here that given that, as we argued above, full disclosure is Pareto optimal, the lower bound of μ is zero which corresponds to full disclosure.

[14]As we show in the next section, this is not necessarily true when effort is endogenous as more privacy can increase effort and hence benefit the firm.

Lemma 5 (Consumer's Optimal Privacy)**.** *There exists* $\hat{e} \in \left(0, \frac{\theta_L}{\theta_H}\right)$ *such that*

1. *For* $e \leq \hat{e}$*, full privacy is optimal for the consumer and hence Pareto optimal.*

2. *For* $e > \hat{e}$*, full privacy is not optimal for the consumer and hence not Pareto optimal.*

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 e. (It can also be interpreted as the per-consumer utility as a function of the posterior beliefs of the firm.) For $e < \frac{\theta_L}{\theta_H}$, the expected payoff of the consumer under full privacy is a strictly positive and concave function of e, while for $e \geq \frac{\theta_L}{\theta_H}$, the firm finds it optimal to exclude type L, and the payoff of the consumer drops to zero.

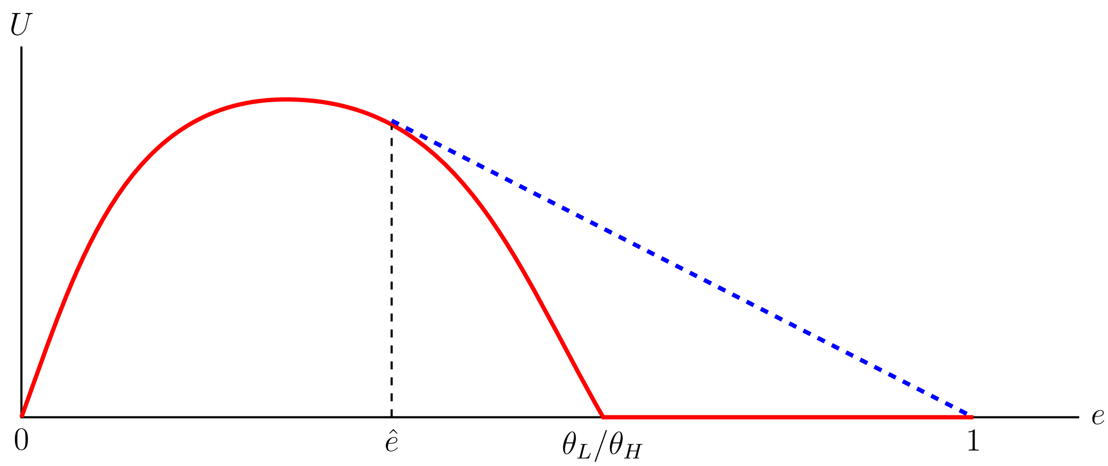


Figure 2: Maximum consumer welfare as a function of the prior e. Depending on the prior, information disclosure can increase or decrease consumer welfare. For instance, for relatively low e, full privacy maximizes consumer welfare but for high e information disclosure can increase consumer welfare.

Full privacy is optimal for the consumer, and hence Pareto optimal, for sufficiently low e. This is because for low e, the consumer earns considerable rents already. A decrease in privacy increases these rents if the consumer turns out to be type H and send signal ω_{LH} but decreases the probability with which these rents are obtained. For sufficiently low e the latter effect dominates and hence any decrease in privacy will hurt the consumer.

Conversely, full privacy is not optimal for the consumer, and hence not Pareto optimal, whenever $e > \hat{e}$, where, as indicated in the figure $\hat{e}$ 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 $e \in (\hat{e}, 1)$, revealing type H's identity with the probability $1-\mu^*$ satisfying $\hat{e} = \frac{e\mu^*}{e\mu^*+(1-e)}$ will mean that, with probability μ^*, type H's identity will remain private and the firm's posterior will be $\hat{e}$, while with probability $1-\mu^*$, type H'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 $\hat{e}$. In expectation, it will be given by the height of the dashed blue secant line in the figure at the prior e. Full privacy is thus Pareto inefficient for any prior $e > \hat{e}$: revealing type H's identity with probability $1-\mu^*$ 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 $\hat{e}$ is similarly Pareto inefficient. So any Pareto efficient allocation be a associated with a distribution of posterior beliefs with support on $[0,\hat{e}] \cup \{1\}$.[15]

Lemma 4 establishes that the firm prefers less privacy and Lemma 5 establishes conditions under which the consumer prefers full or less than full privacy. Taken together these two results establish the following corollary.

Corollary 1 (Trade-off of Privacy for an Exogenous e)**.** *Suppose that effort is exogenously given and equal to e. Within the set of Pareto optimal allocations, an informational environment is preferable to the consumer (firm) if and only if it involves more (less) privacy.*

3.3 Endogenous Effort

Having established several useful results in the exogenous effort case, we now turn our attention to the case in which effort is endogenous. Our first step is to characterize the equilibrium effort for any information structure.

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

The same graph can also be used to explain Lemma 3. Consider any informational environment featuring multiple distinct signals that lead to posteriors in the interval $[0,\hat{e}]$. 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 $[0,\hat{e}]$—will raise expected 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 as described in Lemma 3.

To that end, fix any information environment σ. Recall that for any anticipated effort by the consumer, the firm will offer contracts as characterized by Lemma 1. The ex-post payoffs of the consumer in state ω depend on the anticipated effort e through (3) and (5). The consumer maximizes (8) by taking the ex-post utilities as given; however, in equilibrium the anticipated effort must match the actual effort. Hence, the equilibrium effort is characterized by a fixed point. For any σ, define the mapping $e \mapsto \Gamma(e|\sigma)$, where

$$\Gamma(e|\sigma) \triangleq \max_{e' \in [e_0, 1]} \underbrace{e' \, \Delta\theta \sum_{\omega \in \Omega} \mu_H(\omega) \max\left\{0, \xi\left(\theta_L - \frac{e}{1-e}\frac{\mu_H(\omega)}{\mu_L(\omega)}\Delta\theta\right)\right\}}_{MB(e|\sigma)} - \psi(e'). \tag{12}$$

An equilibrium is a fixed point of the mapping $e \mapsto \Gamma(e|\sigma)$.

Lemma 6 (Equilibrium Effort). *For any σ, there exists a unique equilibrium effort $e^*(\sigma)$ that is a fixed point of $e \mapsto \Gamma(e|\sigma)$.*

Figure 3 depicts the mapping $e \mapsto \Gamma(e|\sigma)$. There are two cases: Panel (4a) depicts the case in which there is an interior equilibrium, whereas Panel (4b) depicts a corner solution. To illustrate, suppose that e_0 is relatively low and $\psi'(e_0) = 0$. Moreover, if σ allows the consumer to earn strictly positive rents for sufficiently low anticipated e, then the consumer has an incentive to exert effort when the anticipated e is low, so an interior fixed point exists. Conversely, if, for instance, e_0 is relatively high and σ does not allow the consumer to earn strictly positive rents, the fixed point will occur at the corner (i.e., the consumer has no incentive to exert additional effort).

The payoffs of the consumer and the firm, as well as total welfare, are readily obtained by substituting the equilibrium effort $e^*(\sigma)$ in (8), (10), and (11).

Our goal now is to characterize the set of Pareto optimal informational environments when effort is endogenous. We first define Pareto optimal informational environments under endogenous effort.

Definition 2 (Pareto Optimality Under Endogenous e). *Suppose that the effort level is endogenous. An informational environment σ is Pareto optimal if there does not exist another informational environment σ' such that*

$$U(\sigma'|e^*(\sigma')) \geq U(\sigma|e^*(\sigma)) \quad \textit{and} \quad \Pi(\sigma'|e^*(\sigma')) \geq \Pi(\sigma|e^*(\sigma)),$$

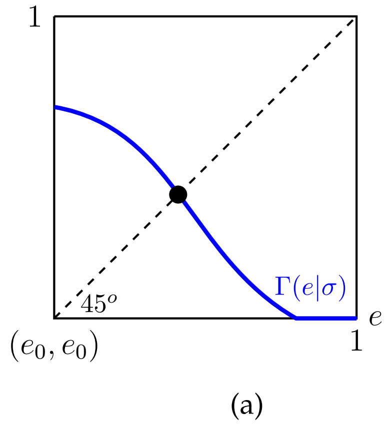


Figure 3: The mapping $e \to \Gamma(e|\sigma)$ is decreasing. Panel (a) depicts an interior fixed point whereas Panel (b) depicts a corner fixed point.

with at least one of the two inequalities being strict.

To start, note that from the first-order condition of equilibrium effort, we can establish the following inequality:

$$U(\sigma|e^*(\sigma)) \leq e^*(\sigma)\psi'(e^*(\sigma)) - \psi(e^*(\sigma)), \tag{13}$$

with equality holding for an interior equilibrium level of effort. Provided that $\psi'' \geq 0$, the right-hand side is readily shown to be strictly increasing in $e^*(\sigma)$. This observation leads to the following useful result:

Lemma 7. *$U(\sigma'|e^*(\sigma')) \geq U(\sigma|e^*(\sigma))$ if and only if $e^*(\sigma') \geq e^*(\sigma)$.*

Thus, equilibrium consumer welfare increases in response to a change in the information environment if and only if equilibrium effort increases.

Now consider any two informational environments σ_1 and σ_2. If effort were exogenously fixed at $e^*(\sigma_1)$, then σ_2 would be preferable for the consumer over σ_1 if and only if it increased the expected information rents. Note that the expected information rents are given by $MB(e|\sigma)$ as defined in (12). Nonetheless, any change in $MB(e|\sigma)$ would necessarily alter $e^*(\sigma)$. In particular, an increase in the expected information rents would increase equilibrium effort. This implies that equilibrium effort (and hence consumer welfare) increases if and only if a change from σ_1 to σ_2 would increase the consumer's payoff if e were *exogenously* fixed at $e^*(\sigma_1)$. 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 increases consumer welfare in settings with effort exogenously fixed at the initial equilibrium level.

Let us now rank the profit of the firm for any two distinct informational environments σ_1 and σ_2. Suppose that the firm prefers σ_2 over σ_1 when effort is exogenously fixed at $e^*(\sigma_1)$. Note that the firm naturally prefers higher effort because it earns a higher ex-post profit when dealing with type H. Therefore, a sufficient condition for a change from σ_1 to σ_2 to increase equilibrium firm profit is that equilibrium effort increases.

The preceding is summarized in the following lemma:

Lemma 8. *Consider any two informational environments σ_1 and σ_2. Then:*

1. $U(\sigma_2|e^*(\sigma_2)) \geq U(\sigma_1|e^*(\sigma_1)) \Leftrightarrow e^*(\sigma_2) \geq e^*(\sigma_1) \Leftrightarrow U(\sigma_2|e^*(\sigma_1)) \geq U(\sigma_1|e^*(\sigma_1))$.

2. *If $e^*(\sigma_2) \geq e^*(\sigma_1)$ and $\Pi(\sigma_2|e^*(\sigma_1)) \geq \Pi(\sigma_1|e^*(\sigma_1))$, then $\Pi(\sigma_2|e^*(\sigma_2)) \geq \Pi(\sigma_1|e^*(\sigma_1))$, with strict inequality if the second inequality is strict.*

Lemma 8 straightforwardly implies the following result—which is the most useful result, as it allows us to bootstrap results from the exogenous-effort environment to the endogenous-effort environment.

Proposition 1 (Bootstrapping)**.** *Consider any informational environment σ_1 with associated effort $e^*(\sigma_1)$. If σ_1 is Pareto inefficient when effort is exogenously given and equal to $e^*(\sigma_1)$, then it is also Pareto inefficient when effort is endogenous.*

In light of Lemma 3, an immediate corollary of Proposition 1 is the following.

Corollary 2 (Optimal Information Structures Under Endogenous e)**.** *An informational environment is Pareto optimal if and only if it is equivalent to a two-signal informational environment $\Omega = \{\omega_H, \omega_{LH}\}$, where $\mu_L(\omega_H) = 0$ and $\mu_H(\omega_{LH}) \equiv \mu \in \left[\underline{\mu}, \overline{\mu}\right]$, for some $0 \leq \underline{\mu} \leq \overline{\mu} \leq 1$.*

As in the previous section, μ can be interpreted as the level of privacy. Nonetheless, unlike the case in which effort is exogenously given, under endogenous effort full disclosure is not necessarily Pareto optimal. The reason is simple: full disclosure provides no rents to the consumer and hence no incentive to increase effort. Given that higher effort is preferred by both the firm and the consumer, full disclosure is not necessarily Pareto optimal (i.e., $\underline{\mu}$ can be bounded away from zero).

Note that, as established above, higher effort is always preferred by the consumer. Therefore, if an informational environment that perhaps entails more privacy increases effort, it necessarily increases the consumer's payoff. On the other hand, while the firm prefers higher effort, it simultaneously prefers more disclosure. This implies that, along the Pareto frontier, an informational environment is preferred by the consumer if and only if it entails more privacy. This is formally stated in the following result, which resembles the result established under exogenous effort.

Proposition 2 (Trade-off of Privacy Under Endogenous e)**.** *Suppose that effort is endogenous. Within the set of Pareto optimal allocations, an informational environment is preferable to the consumer (and hence less preferable to the firm) if and only if it involves more privacy.*

An interesting question is whether full privacy and full disclosure are ever Pareto optimal.

Recall, for example, that under exogenous e we established that if $e > \hat{e}$, full privacy is not Pareto optimal. Based on our bootstrapping result, to see if full privacy is not Pareto optimal under endogenous effort it suffices to check whether

$$e^*(\sigma_{\text{priv}}) > \hat{e}.$$

Indeed, whenever this is true, full privacy is not Pareto optimal under an exogenously given effort equal to $e^*(\sigma_{\text{priv}})$, which implies that it cannot be Pareto optimal when effort is endogenous. Therefore, characterizing conditions under which full privacy is not Pareto optimal boils down to characterizing conditions under which the equilibrium effort under full privacy exceeds $\hat{e}$. Such conditions are provided in the following proposition.

Proposition 3 (Full Privacy)**.** *If $e_0 > \hat{e}$ or $e_0 \leq \hat{e}$ but $\psi'(e)$ is not very steep, then full privacy is not Pareto optimal.*

This result has several implications. It states that when the consumer already has a high probability of being high type or when the cost of exerting effort to increase her value is sufficiently low, full privacy is not Pareto optimal. The intuition is as follows. When the cost of exerting effort is relatively low, the consumer will exert high effort, which tends to reduce the rents she receives, thereby lowering her incentive to exert effort. Consequently, moving away from full privacy provides the consumer with higher rents, which in turn increases her effort. Given that the firm prefers both higher effort and more disclosure, an environment with less than full privacy becomes Pareto optimal.

In further characterizing the Pareto frontier, note that when full privacy is not Pareto optimal, the Pareto frontier consists of more than one point. For instance, if the consumer prefers less than full privacy, then slightly decreasing privacy will have no first-order effect on effort (and hence on the consumer's payoff) but will strictly increase the firm's payoff. Hence, by continuity and the convexity of the Pareto frontier, the Pareto frontier will consist of multiple points in this case.

Similarly, it is not difficult to construct examples in which the Pareto frontier has more than one point even if full privacy is Pareto optimal. For example, if $\psi'(e)$ is sufficiently steep, reducing privacy will have a negligible impact on equilibrium effort but a significant direct positive effect on the firm's profit by reducing the rents it must pay.

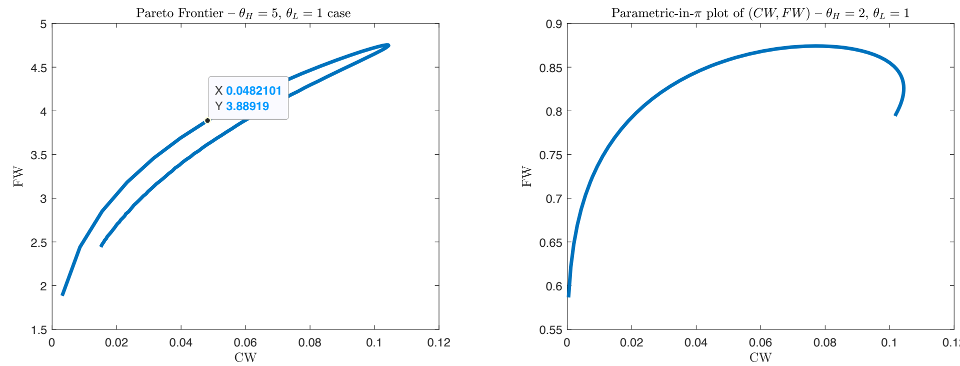
It is also interesting to examine whether full privacy is, in fact, the only point on the Pareto frontier.[16] This is the case in numerical examples, as presented in Figure 4.

We now examine whether full disclosure is Pareto optimal. When effort is exogenously given, we showed that the maximum total welfare is achieved under full disclosure—hence, full disclosure is Pareto optimal in that setting and is associated with maximum profit for the firm. Nonetheless, under endogenous effort this is not necessarily true, as full disclosure entirely suppresses the consumer's incentive to exert effort. If effort is always valuable, even the firm may prefer less than full disclosure.

Proposition 4 (Full Disclosure). *If $\psi'(e_0) = \psi''(e_0) = 0$, full disclosure is not Pareto optimal.*

By Proposition 2, consumer-optimal privacy can be fully characterized by $\overline{\mu}$. The following Proposition shows that this consumer optimal privacy level has clean and intuitive comparative statics. It applies both with endogenous and exogenous e.

[16] When the Pareto frontier consists of a single point, $\underline{\mu} = \overline{\mu}$.



(a) (b)

Figure 4: In the left panel the Pareto frontier consists of a single point: full privacy is the unique Pareto optimal informational environment. In the right panel, the Pareto frontier consists of many points. As per Proposition 2, points on the Pareto frontier that are preferred by the consumer entail more privacy.

Proposition 5 (Comparative Statics)**.** *The consumer-optimal level of privacy $\overline{\mu}$ 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 a straightforward mechanical computation formalizing the following intuition. By Proposition **??**, 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; on the other, greater privacy means greater incentives for the firm to distort q_1 down, hence reducing information rents conditional on privacy. An increase in ψ' decreases the equilibrium e^* given any privacy level, which reduces the distortions and hence the marginal cost of additional distortions, making additional privacy desirable at the margin. Similarly, an increase in C' decreases e^* (by decreasing consumer's marginal benefit of effort), increasing the marginal cost of additional distortions and making less 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.

4 Extensions

4.1 Alternative Models

Although our baseline model described in the preceding subsection considers a seller-buyer relationship, all of our analysis applies equally well to the related but distinct models.

Related Models. Reconsider the model we described above, but now suppose that the principal's payoff is given by

$$\pi(q,t) = -t + S(q), \tag{14}$$

where $S'(q) > 0$, $S''(q) < 0$, and $S'''(q) \leq 0$. The agent's payoff is given by

$$u_i(q,t) = t - \frac{1}{\theta_i} q. \tag{15}$$

This specification encapsulates two distinct and relevant applications. In one example, an *employer* (the principal) enters into a contract with an *employee* (the agent). The employer assigns a task q to the employee that provides him a (gross) payoff $S(q)$ and compensates the employee with salary t. The ease of accomplishing the task is inversely proportional to the employee's type—the high type has a lower marginal cost of performing the task than the low type. The employee's type is determined probabilistically by the worker's prior investment e in firm-specific human capital, with higher e raising the probability of being a higher (lower-cost) type.

Similarly, in another example, a *retailer* (the principal) contracts with a *manufacturer* (the agent). In this case, $S(q)$ refers to the profit the retailer earns in the downstream market and t is the payment of the retailer to the manufacturer. The manufacturer might be efficient and have a low cost of production or inefficient and have a high cost of production. Nonetheless, the manufacturer's type is determined probabilistically by a prior investment e in new technology, with higher e raising the probability of being a higher (lower-cost) type.

As a concrete example of our setting consider a firm that is a leading customer relationship management (CRM) software provider (such as Salesforce). The provider offers a highly customizable platform that allows businesses to tailor its tools to their specific needs. Suppose that the consumer is a medium-sized retail company that wishes to adopt provider's CRM software to automate marketing workflows, and integrate analytics into its operations. The ultimate value the retail company derives from the CRM software depends on the state of its legacy systems (e.g., existing customer databases and internal workflows) and the efforts of its employees to fully adopt and utilize the platform.

Employees must invest significant effort to learn the new system, migrate data from outdated formats, and adjust their processes to align with the new features. If the employees dedicate sufficient effort to training and system adoption, the CRM software will integrate seamlessly, enabling better customer insights and streamlined operations (i.e., high-type valuation). Conversely, if the employees invest minimal effort, the system may not integrate effectively, leading to incomplete data migration and suboptimal use of the CRM software (i.e., low-type valuation). The effort invested by the employees is analogous to the unobservable action in our model, which probabilistically determines the realized type of the consumer.

Note also that we could further extend the last model by allowing the payoff of the principal to depend on the type of the agent as in **?**. For example, suppose that

$$\pi_i(q,t) = -t + S_i(q), \tag{16}$$

where $S'_H(q) > S'_L(q) > 0$, $S''_i(q) < 0$, and $S'''_i(q) \leq 0$. Therefore, the principal gets a higher payoff by trading with type H rather than with a type L. This is an example of common values and our results extend to this environment as well.

Optimal Taxation. Alternatively, by a straightforward reinterpretation of the variables in (14) and (15), we can denote the utility of the firm as

$$\pi(q,t) = -t + q, \tag{17}$$

and the utility of the consumer as

$$u_i(q,t) = t - \frac{1}{\theta_i}H(q), \tag{18}$$

where $H'(q) > 0$, $H''(q) > 0$, and $H'''(q) \leq 0$.

In this example, an *extractive state* (the firm) raises revenue through a non-linear income tax $T(q)$ on observable earned income q, leaving $t = q - T(q)$ for citizens (consumers). The disutility of generating pre-tax income q depends inversely on the skill θ_i of the citizen. Moreover, the skill is determined probabilistically by an unobservable educational investment e.

Based on this setting, our results can also be plausibly interpreted as providing some *normative* support for the epistemic foundations of a Mirrlesian approach (Mirrlees 1971) 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.

5 Discussion and Conclusion

5.1 Relationship to the Market Segmentation Literature

Our analysis contributes to the so-called market segmentation literature following Bergemann et al. (2015; henceforth BBM). BBM study information disclosure 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 5) insofar as the monopolist finds it optimal to set a price that excludes low-willingness-to-pay types. In such a situation, efficiency – as measured by the ex-ante consumer plus producer surplus – can be 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.

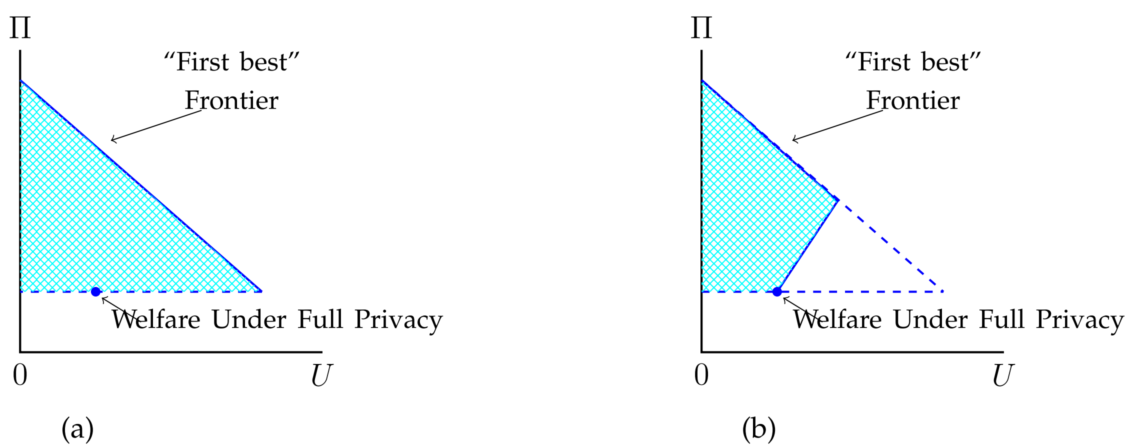


Figure 5: Panel (a) depicts The Surplus Triangle with a Single Product as in Bergemann, Brooks and Morris (2015). Panel (b) depicts The Surplus Area with Multiple Products as in Haghpanah and Seigel (2022,2023).

Haghpanah and Siegel (2022) and Haghpanah and Siegel (2023) extend BBM's analysis to multi-product monopolists. Their qualitative insights are illustrated in Figure 5b.

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 5b. 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 5b.

Lemma 2 and Lemma 5 contribute to this literature by showing that if, effort is exogenously given and, one allows for a continuum of possible 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.

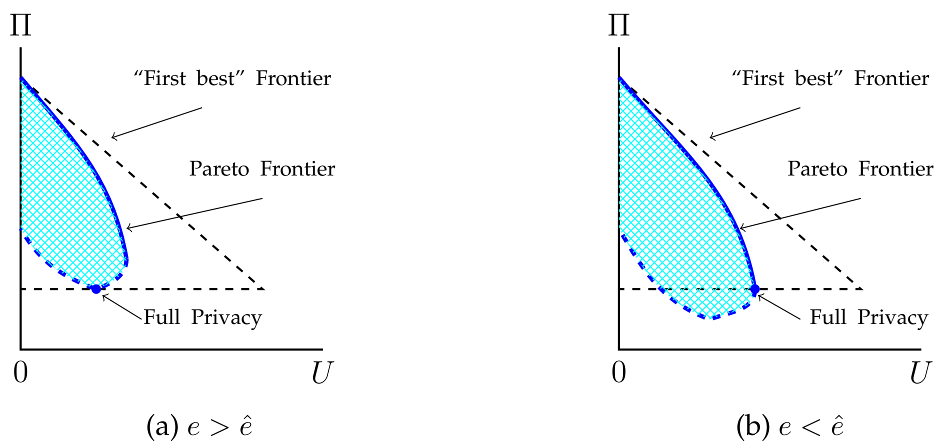


Figure 6: Panel (a) depicts the "First best" frontier as well as the Pareto frontier when $e > \hat{e}$ and hence full privacy is not Pareto optimal. Panel (b) depicts the "First best" frontier as well as the Pareto frontier when $e < \hat{e}$ and hence full privacy is Pareto optimal.

Figure 6a 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). Figure 6a illustrates 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, related are also the papers by Ichihashi (2020) and Hidir and Vellodi (2021). Ichihashi (2020) considers a finite set of products and compares discriminatory and non-discriminatory pricing by the seller. He shows that the seller would prefer to commit to a non-discriminatory pricing but this makes consumers worse off. Unlike Ichihashi (2020), we allow for a continuum of products and unlike both papers we allow the seller to offer any number of products to any segment.

Hidir and Vellodi (2021) consider a continuum of products but allow the seller to offer a unique product. 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. In fact, in the setting studied by Hidir and Vellodi (2021) (i.e., a pure horizontal differentiation model), if the firm were allowed to offer any number of products, as we do, it would be able to extract the entire surplus of the consumers for any information structure. Therefore, this setting is not appropriate for an exercise like that we perform in this paper.

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.[17] 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

[17] 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.

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), for two possible types, the optimal disclosure would entail either full privacy (i.e., no disclosure) whenever the (full privacy) mechanism entailed no shut down or full privacy and full disclosure would yield the same outcome if the (full privacy) mechanism entailed shut down.

5.2 Old conclusion

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.

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 high levels privacy, particularly so when the informed agent can take an unobservable action which increases her type. Indeed, in these unobservable action settings, our results indicate even firms find a commitment to *full* privacy to be optimal for the lowest type and a commitment to significant privacy to be optimal for the higher type. To revisit this implication in some of the examples discussed in section **??**:

- 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 it.
- 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.

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.

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.

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