What Should a Firm Know? Protecting Consumers' Privacy Rents*

Daniel Bird and Zvika Neeman†

March 9, 2020

Abstract

A monopolistic firm observes a signal about the state of the world and then makes a take-it-or-leave-it offer to an uninformed consumer who has recourse to some outside option. We provide a geometric characterization of the firm's information structure that maximizes the consumer's surplus: the optimal regime partitions the space of payoff states into polyhedral cones with disjoint interiors. We interpret our results in terms of the maximization of the consumer's "privacy rents." We illustrate and motivate our approach through an example of the regulation of the privacy of genetic information.

JEL Codes: D42, D82, D83, L51.

Keywords: Information Design, Privacy Design, Privacy Rents, Persuasion, Genetic Non-Discrimination.

*Acknowledgments to be added.

†Eitan Berglas School of Economics, Tel Aviv University (e-mails: `dbird@tauex.tau.ac.il`, `zvika@tauex.tau.ac.il`).

1 Introduction

We consider a model in which a monopolistic firm observes a signal about the state of the world and then makes a take-it-or-leave-it offer to an uninformed consumer who has recourse to some outside option. We focus our attention on situations in which the state of the world reflects the consumer's personal characteristics. In such situations, we ask how it is possible to design the firm's information structure so as to maximize the consumer's surplus. In other words, we solve for the policy that maximizes the consumer's information rents or "privacy rents."

Our study is motivated by the following example. Recent technological developments in personalized medicine allow health providers to tailor medical treatment to patients' personal genetic structure. Genetic bio-markers indicate susceptibility to disease, and so genetic testing can be used to predict the efficacy of various medical treatments. This type of information, although personal, is nevertheless proprietary to the health provider whose research and expertise generated it (in some cases, it is also protected by patent). No one else is able to perform the required diagnostic tests or interpret their results. In particular, the patient does not know or would not even be able to understand this information if it were presented to her without proper explanation. This raises the concern that once the health provider learns a patient's genetic profile, it may bias the treatment selection it offers the patient in order to reduce its cost at the expense of the patient's welfare, or altogether deny coverage to the patient.[1]

The medical provider may be viewed as a firm that observes a signal about the state of the world, which is given by a patient/consumer's genetic makeup, and offers the patient a choice among several insurance or treatment options. We are interested in the question of how a regulator should restrict medical providers' use of information to allow patients to reap the full benefits of personalized medicine. Put more generally, we address the problem of how to balance firms' use of information so that they have enough information to facilitate efficient transactions, while limiting their ability to extract surplus from consumers through exploitation of their informational advantages.

More concretely, suppose that a firm sells a standard medical insurance policy that covers treatment for some medical condition. It is able to discern patients/consumers' types through genetic testing. Patients do not know their types. Suppose that a patient may be one of the following three, equally likely, types: *immune* to the relevant medical condition, *responsive* to treatment for this condition, or *non-responsive* to treatment.[2]

[1]For example, insurance firms can, and do, deny life insurance to carriers of cancer genes such as BRCA. The Genetic Literacy Project website (https://geneticliteracyproject.org) lists many such examples.

[2]Genetic tests are typically divided into two classes: predictive and prognostic. Predictive tests predict susceptibility to disease (e.g., the BRCA gene is a predictive bio-marker for breast cancer, and polymorphisms

We describe patients' payoffs and the firm's profit from this insurance policy in the table below. Payoffs of both patients and the firm from not offering the policy are normalized to zero.

Patient's Type	Patient's payoff	Firm's profit
immune	-1	1
responsive	2	-1
non-responsive	-2	-1

An immune patient has no need for medical insurance, and so her payoff from the medical insurance policy is -1, a patient who is responsive to the treatment that is provided through the policy has a payoff of 2, and a non-responsive patient has a payoff of -2 because on top of the insurance premium, she also incurs the costs associated with receiving treatment. A firm that offers this policy to consumers obtains a profit of 1 from an immune patient (because such a patient has no need for costly treatment), a profit of -1 from a responsive patient (because such a patient requires costly treatment), and a profit of -1 from a non-responsive patient (because such a patient requires the same costly treatment as a responsive patient).

If the firm has no information about the patients' types, possibly because genetic testing is prohibited by a genetic nondiscrimination law, then the firm would not offer patients any coverage because the expected profit from the standard insurance policy is negative. Moreover, even if the standard insurance policy were offered, then the patient would reject it because it would provide her with a negative expected payoff.[3]

Suppose instead that genetic testing is unregulated and the firm knows the patients' types. If the firm offers insurance only to immune patients, then patients would deduce from being offered the insurance that they are immune and would reject the firm's offer. The strategy that maximizes the firm's profit subject to the constraint that patients are willing to accept it is to offer the insurance policy to all of the immune patients and to half of the responsive patients. Such a strategy generates an expected payoff of $\frac{2}{3}\cdot(-1)+\frac{1}{3}\cdot(2) = 0$ for a patient who is offered the policy, and an ex-ante expected profit of $\frac{1}{3}\cdot 1+\frac{1}{3}\cdot\frac{1}{2}\cdot(-1)+\frac{1}{3}\cdot\frac{1}{2}\cdot 0+\frac{1}{3}\cdot 0 = \frac{1}{6} > 0$ for the firm. Hence, with no regulatory restrictions patients do not benefit from the possibility of genetic testing.

in the gene encoding the enzyme MTHFR are predictive bio-markers for acute lymphocytic leukemia). Prognostic tests predict the efficacy of treatment if such is needed (e.g., the TPMT genetic test screens for a gene variant of thiopurine methyltransferase to assess the effectiveness of the thiopurines class of drugs that are prescribed to patients suffering from leukemias and autoimmune disorders, and the Prometheus IBD Serology 7 test kit identifies a subset of inflammatory bowel disease patients who will benefit from budesonide).

[3]The fact that for *some price* both the firm and the patient prefer no coverage to coverage implies that at least one of them would prefer no coverage for *any price*.

Suppose, however, that genetic privacy regulation implies that the firm can only learn if a patient is non-responsive or not. In this case, the firm would not offer the policy to non-responsive patients, but it would be willing to offer it to patients who are not non-responsive. Such a strategy generates an ex-ante expected profit of $\frac{2}{3}(\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1)) + \frac{1}{3} \cdot 0 = 0$ for the firm, and an ex-ante expected payoff of $\frac{2}{3}(\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 2) + \frac{1}{3} \cdot 0 = \frac{1}{3} > 0$ for patients.

This example demonstrates how judicious regulation of what firms may know (or how they can use their knowledge) can increase consumer surplus. We have illustrated our argument through the example of genetic or medical testing. But, the same argument applies to any "credence good" that is sold in a monopolistic market. Examples include medical treatment, repair services, other types of personal services, and expert advice.[4]

The popularity and success of products and services that require consumers' to voluntarily share some of their personal information with firms, such as iPhone and Netflix, indicates that the voluntary sharing of private information can obviously benefit consumers in much broader settings than the one described above. Moreover, it is also widely appreciated that the information that is shared involuntarily through consumers' credit scores facilitates the smooth operation of credit markets. However, it is also at least as obvious that voluntary or involuntary sharing of some types of personal information may harm consumers, because it may permit price and other forms of discrimination, not to mention the scandalous manipulation of behavior such as the one recently allegedly performed by Cambridge Analytica. This basic trade-off is illustrated by Varian's (2009) well-known example: a consumer who likes apples would like the apple seller to know whether she prefers Jonathan to Macintosh apples, but not her willingness to pay for her preferred type of apple. For a recent survey on the economics of privacy, see Acquisti, Taylor and Wagman (2016).

Lawmakers and regulators have appreciated this insight at least since the enacting of the Fair Credit Reporting Act of 1970. The act facilitates the sharing of individuals' recent financial history, but clearly states that certain financial events in a person's past cannot impact credit scores. The recent proliferation of data-gathering by firms has heightened the importance of such regulation and triggered a wave of new laws intended to protect the privacy of consumers.[5]

[4]A credence good is characterized by the fact that consumers can observe the utility they derive from the good ex post, but cannot tell if the type or quality of the good they have received is the ex-ante needed one; however, an expert seller is able to identify the type or quality that fits the consumer's needs by performing a diagnostic test (Darby and Karni 1973). For a survey of the economics of credence goods, see Dulleck and Kerschbamer (2006).

[5]Examples include the European Union's General Data Protection Law of 2016 and India's proposed Personal Data Protection Bill of 2019, which aims to regulate privacy and data protection, or the United States suggested Deceptive Experiences To Online Users Reduction Act which aims to restrict large Internet platforms from manipulating consumers.

Two laws that are especially relevant to this paper are The Genetic Information Nondiscrimination Act (GINA) of 2008 and the Preserving Employee Wellness Programs Act of 2017. GINA protects Americans from discrimination based on their genetic information in both health insurance and employment. In particular, GINA prohibits health insurers from discrimination based on the genetic information of enrollees. That is, health insurers may not use genetic information to make eligibility, coverage, underwriting, or premium-setting decisions. Moreover, health insurers may not request or require individuals (or their family members) to undergo genetic testing or to provide genetic information. The analysis in this paper (as in the example presented above) suggests that such a sweeping nondiscrimination law may be too blunt to be effective, and describes how a more nuanced approach may serve to promote social welfare. This exact concern is reflected by the Preserving Employee Wellness Programs Act of 2017 that limits the scope of GINA and allows employers to collect genetic data of employees and their families (and impose sanctions for noncompliance) to be used in employee wellness programs.

We describe a general model that generalizes the example presented above. We describe the problem as one of how to structure the information of a monopolistic firm in a way that maximizes the consumer's expected payoff, subject to participation and incentive compatibility constraints for both the firm and consumer.[6] In this sense, we describe a model for the optimal regulation of privacy.

The case in which the firm can make only two offers to the consumer (offer/not offer a standard contract) admits a particularly elegant formulation and solution. In this case, the payoff space for the problem can be reduced to a two-dimensional Euclidean space: payoffs and profits from the default offer are normalized to zero, and payoffs and profits from the other offer are measured relative to the origin. We show that optimal privacy requires that this two-dimensional space be divided into two half-spaces by a (straight) line that passes through the origin. Each half-space corresponds to an offer. The firm is only allowed to learn to which half-space the state of the world belongs, and the half-spaces are designed so that in each half-space, the firm is induced to make the offer that is associated with this half-space, and this suggestion is in turn accepted by the consumer.

If there are three or more offers that the firm can make, the analysis is more involved. First, the offers that are incentive compatible for the firm in different states of the world are more difficult to characterize geometrically. Intuitively, with more possible offers, the set of deviations is much richer: the firm may deviate after observing several rather than just one signal in order to make this deviation more palatable for the consumer. Second, the

[6]This formulation is equivalent to a formulation in which the firm knows the state of the world and what is regulated is the firm's ability to condition its offer to the consumer on various characteristics of the consumer.

payoff space must be divided into more than two regions. Nevertheless, we show that the aforementioned result generalizes in the following sense.

An alternative (but mathematically equivalent) description of the optimal solution for the case with two offers is that the payoff space is partitioned by a hyperplane into two sets, each represented by a polyhedral cone.[7] We show that this interpretation of the result continues to hold also in the case where the firm can make more than two offers. That is, optimal privacy is attained by partitioning the payoff space into polyhedral cones with disjoint interiors that are each associated with a specific offer. This result provides a guideline for evaluating privacy policies through the emphasis it places on the dimensions of the consumers' type that should be considered. Namely, it shows that under optimal privacy regulation, firms should not be allowed to distinguish among consumer types for which the *ratio* of payoffs from *any* two offers is similar. This simple insight suggests that there may exist two consumer types for which a given offer is best for both the consumer and the firm, but the regulator may nevertheless require the firm to distinguish between them (because they have different ratios of payoffs).

If the firm's preferences are independent of the state of nature, our analysis provides sharper results with respect to the structure of optimal regulation. If there are only two possible offers, then we show that optimal regulation confers no benefit to consumers relative to no regulation. More generally, if there are more than two offers, we find that regulation can help the buyer, albeit in a very limited way. In particular, there is at most one offer that allows the consumers to benefit relative to the case of no regulation. The role of regulation is limited to increasing the expected payoff to the consumer from this offer and the probability with which this offer is made.

Related Literature

As explained above, we are interested in understanding how a firm's information about a consumer's type affects the offer it will make to the consumer and the resulting consumer welfare. Accordingly, our paper is related to the literature on information design (see Bergemann and Morris 2019, for a recent survey) and, in particular, the impact of information design on monopoly pricing problems (Bergemann, Brooks and Morris 2015; Roesler and Szentes 2017; Ichihashi 2020). The key difference between the aforementioned papers and ours is that in those papers the players do not transmit any information to one another. Bergemann, Brooks and Morris (2015) and Ichihashi (2020) assume that the buyer knows her valuation before deciding whether to purchase a good, and focus on designing the monopolist's information about the buyer's type. By contrast, Roesler and Szentes (2017) assume

[7]Each payoff vector is associated with the ray that emanates from the origin and passes through it. A polyhedral cone is the convex hull that contains a collection of such rays.

that the firm has no information about the buyer's type, and focus on the design of the buyer's information structure about her own valuation. A second distinction between the aforementioned papers and our own is that we allow for general preferences, while in those papers the firm's payoff is given by the price paid, and the buyer's utility is given by the difference between her valuation and the price.

In our model the firm's choice of an offer conveys information to the consumer about her type. Hence, our paper is also related to the literature on Bayesian persuasion; see Kamenica and Gentzkow (2011) and the recent survey by Kamenica (2019).[8] Especially close are those papers that study the impact of exogenous restrictions on the sender's ability to persuade others, including multiple receivers (e.g., Alonso and Camara 2016 and Galperti and Perego 2019), multiple senders (e.g., Gentzkow and Kamenica 2017 and Li and Norman 2019), or a privately informed receiver (e.g., Kolotilin et al. 2017 and Matyskova 2018). The paper that is perhaps most closely related to ours is Ichihashi (2019) who analyzes the impact of coarsening the sender's information about the state of nature. Our work differs from his in two key aspects. First, whereas Ichihashi (2019) considers a model with only two possible actions, we allow for any finite number of actions or offers. As will become clear below, this difference has significant implications on the type of incentive constraints that the information designer must respect. Second, Ichihashi (2019) characterizes the set of payoffs that can be attained under some information structure, but does not consider how to attain each such a payoff/profit pair. By contrast, our main result describes the geometry of the firm's information structure that maximizes the consumer's expected payoff. Thus, our analysis provides insights into the regulation of the firm's information that cannot be obtained from Ichihashi's work.

The rest of the paper proceeds as follows. Section 2 describes the model. The case with two offers is analyzed in Section 3, and the general case in Section 4. The effectiveness of regulation is discussed in Section 5 under the special assumption that the firm's preferences are state-independent. Section 6 concludes. All proofs are relegated to the appendix.

2 Model

We consider the problem of a designer who attempts to protect uninformed consumers from being manipulated by an interested monopolistic firm. Formally, the firm makes a take-it-or-leave-it offer to the consumer from a finite set $A = \{a_0, \ldots, a_N\}$, where a_0 is the default

[8]Since we allow the firm to restrict the actions that are available to the consumer, our model is not a "pure persuasion" model.

offer.[9] That is, a_0 denotes the outcome that prevails if the consumer rejects a firm's offer of a_n for $n \in \{1, \dots, N\}$. The offers in the set A may be interpreted as different possible binding contracts, or terms of trade between the consumer and the firm that have been pre-approved by the designer.

The payoff from each offer depends on the state of the world $\theta \in \Theta = \{\theta_1, \dots, \theta_M\}$ and is given by $u(a, \theta)$ and $v(a, \theta)$ for the consumer and firm, respectively. The prior distribution of the states is given by $\pi(\theta_i) > 0$. We denote by $\overline{\theta}$ the vector $(\theta_1, \dots, \theta_M)$, and by $\pi(\overline{\theta})$ the vector $(\pi(\theta_1), \dots, \pi(\theta_M))$. The vectors $(u(a_i, \theta_1), \dots, u(a_i, \theta_M))$ and $(v(a_i, \theta_1), \dots, v(a_i, \theta_M))$ are denoted by $u(a_i, \overline{\theta})$ and $v(a_i, \overline{\theta})$, respectively.

Before making its offer to the consumer, the firm observes a signal $s \in S$, upon which it can condition its offer. We consider a situation in which the firm employs the same strategy with respect to all of its consumers. This effectively implies that before observing the signal, the firm commits to an offer strategy $q(a|s)$ that specifies the probability that the firm offers a after observing signal s.

We depart from the literature by assuming that the information structure that gives rise to the signal observed by the firm is chosen strategically by the designer in order to maximize the consumer's expected payoff. We assume that the designer selects a finite set of signals S and a set of conditional distributions $\{p(s|\theta)\}_{\theta \in \Theta}$ that describe the probability that the firm observes signal $s \in S$ in state θ. We refer to the pair $\langle p(s|\theta), S \rangle$ as a privacy regime.

The timing of the game is as follows. First, the designer selects the privacy regime. Then, before observing the realized signal, the firm selects its offer strategy $q(a|s)$. Finally, the state of the world and signal are realized, the firm makes an offer to the consumer, and the consumer decides whether or not to accept the firm's offer. Note that even though the firm *can* commit to the strategy it employs to influence the consumer under the privacy regime imposed by the designer, it *cannot* commit to the strategy it would employ under other possible privacy regimes. That is, the firm does not have commitment power with respect to the designer and so it chooses an optimal strategy $q(\cdot|\cdot)$ for each privacy regime chosen by the designer. We use a perfect Bayesian equilibrium as our solution concept, and so the designer (correctly) believes that the firm will respond optimally to any privacy regime that the designer might choose.

3 The Case with Two Offers

We start by analyzing the simplest case of our model where the firm can make only two offers, the default offer a_0 and another offer a_1. It is instructive to present the consumer's

[9] Since the consumer is uninformed, offering her a menu of alternatives from which to choose confers no advantage to the firm.

payoff and the firm's profit as in Figure 1 below, which provides a geometric representation of the "payoff space" of the designer's problem.

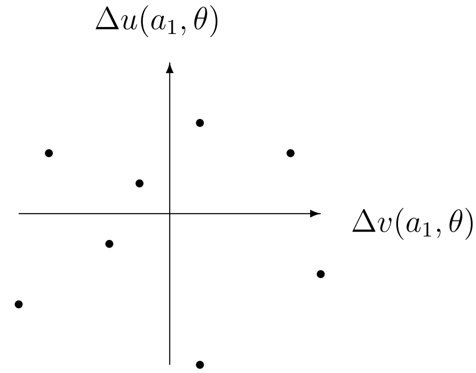

Figure 1: Geometric representation of $\Delta u(a_1,\theta)$ and $\Delta v(a_1,\theta)$

Each point in Figure 1 represents a state of the world. A point's X-coordinate depicts the difference between the firm's payoff from offer a_1 and the firm's profit from the default offer a_0, $\Delta v(a_1,\theta) \equiv v(a_1,\theta) - v(a_0,\theta)$, and a point's Y-coordinate depicts the difference between the consumer's payoff from offer a_1 and the consumer's payoff from the default offer a_0, $\Delta u(a_1,\theta) \equiv u(a_1,\theta) - u(a_0,\theta)$. Note that Figure 1 does not contain any information about the probabilities of the different states.

When there are only two offers, it is assumed without loss of generality that the designer selects an information structure with at most two signals: one that induces the firm to offer a_1 and one that induces the default offer[10] a_0. Note that the designer's signals may be interpreted as the designer's instructions to the firm: "offer a_1" and "offer a_0," respectively. We refer to the states in which the designer instructs the firm to offer a_1 as the "acceptance region" and to the states in which the designer instructs the firm to offer the default offer a_0 as the "rejection region." Notice that the designer may send both signals in some states, and hence a state may belong to both regions.

Observe that if a state belongs to the first (top right) quadrant in Figure 1, then both the consumer and the firm prefer that this state belong to the acceptance region; if a state belongs to the third (bottom left) quadrant, then both the consumer and the firm prefer that this state belong to the rejection region; and if a state belongs to the second (top left) or fourth (bottom right) quadrants, then the consumer and firm have different preferences.

If it were possible, the designer (and consumer) would have liked the acceptance region to coincide with area that lies above the X-axis in Figure 1. However, in this case, following the designer's instruction to offer a_1 is not necessarily a best response for the firm. For example, if states in the second quadrant are very likely, under the buyer-preferred partition

[10]We establish this "revelation principle" formally in Theorem 2 below.

of the payoff space the firm may strictly prefer to offer a_0 rather than a_1 in the acceptance region.[11]

In general, when assigning states between the two regions, the designer must ensure that the firm is willing to forward the designer's suggestion to the consumer. That is, the firm must find it optimal to commit to offer a_1 if and only if it receives the signal that indicates that the state belongs to the acceptance region.

In this simple version of the model with two offers, this implies that the designer's choice of the optimal privacy regime must satisfy two constraints:[12]

1. In the acceptance region the firm must weakly prefer a_1 to a_0.

2. In the rejection region the firm must either weakly prefer a_0 to a_1 or, if it prefers a_1 to a_0, then it must be the case that the consumer is unwilling to accept a_1 conditional on any convex combination of the buyer's induced beliefs in the acceptance and rejection regions.

The first point follows immediately from the fact that the firm can force the consumer to accept the default offer. The second point is more subtle. If the firm prefers a_0 to a_1 in the rejection region, than obeying the designer's instruction to offer a_0 in the rejection region is obviously incentive compatible for the firm. But incentive compatibility is also possible if the firm prefers a_1 to a_0 in the rejection region. For this to be the case, it must be that a deviation to a recommendation function, where $q(a_1|rejection) = \epsilon$, $q(a_0|rejection) = 1 - \epsilon$ and $q(a_1|acceptance) = 1$, induces the consumer to reject the firm's offer of a_1; that is, it violates the consumer's incentive compatibility constraint, for any $\epsilon > 0$. In other words, the firm is willing to offer a_0 with certainty in the rejection region only if it is its preferred action there, or if it cannot induce the consumer to accept a_1 in this region with some small probability.

If a privacy regime is not incentive compatible for the firm, then the designer may "sweeten" it for the firm in one of two ways:

1. The designer can move states in the second quadrant, which give a positive payoff to the consumer but a negative payoff to the firm, from the acceptance to the rejection region.

2. The designer can move states in the fourth quadrant, which give a positive payoff to the firm but a negative payoff to the consumer, from the rejection to the acceptance region.

[11] We slightly abuse terminology and refer to "partitions" of the payoff space, even though in some states more than one offer may be made.

[12] In Section 4.1 we formally derive the firm's incentive compatibility constraints for the general model.

Note that both changes strengthen the firm's incentives to follow *both* of the signals. For each state θ in the 2nd and 4th quadrants, define the "slope of state θ" by the ratio

$$\rho(\theta) = \left| \frac{u(a_1, \theta) - u(a_0, \theta)}{v(a_1, \theta) - v(a_0, \theta)} \right|.$$

That is, $\rho(\theta)$ measures the ratio of the buyer's gain (loss) to the firm's loss (gain) from offer a_1 relative to the default offer.

To construct the optimal solution, the designer begins with the buyer-preferred partition and orders the states in both the second and fourth quadrants by their slopes $\rho(\theta)$. The designer then moves states with a small slope from one region to the other (from the acceptance to the rejection region in the second quadrant and from the rejection to the acceptance region in the fourth quadrant), until the firm is willing to follow the suggested actions in both regions. Intuitively, states with a small slope $\rho(\theta)$ are more efficient sweeteners than states with a large slope because they generate a relatively large increase in the payoff to the firm at the expense of a relatively small loss in the payoff to the consumer. For example, assigning a state with a small slope in the second quadrant to the acceptance region provides the consumer with a small benefit relative to assigning this state to the rejection region. However, such an assignment induces a high cost to the firm.

We depict this idea graphically in Figure 2 below. Start with an acceptance region that coincides with the area above the X-axis. To make the problem nontrivial, suppose that the firm prefers a_0 to a_1 in both regions.[13] Then progressively tilt the boundary of the acceptance region through the origin until the firm is indifferent between the two actions in the acceptance region, i.e., until the firm is induced to follow the designer's instruction in the acceptance region.

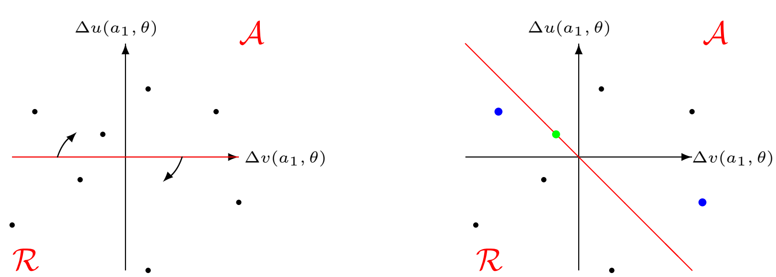

Figure 2: The Optimal Partition

In Figure 2, the states marked in blue are the states with the smallest slopes. These blue states are moved between the regions in order to sweeten the acceptance region for the firm. The green state is the state in which the designer sends both instructions with positive probability.

[13]The cases where the firm prefers a_1 in the rejection region can be handled in a similar way.

We say that a privacy regime is a "half-plane regime" if it is characterized by a line with a negative slope that passes through the origin (as in Figure 2) such that states that lie above this line belong to the acceptance region and states that lie below this line belong to the rejection region. Generically, at most one state lies on the line, and this state may be split between the acceptance and rejection regions.

The previous illustration assumes that both offers are made with positive probability. For some parameters, this need not be the case. For example, if there are two equally likely states represented by $(-2, 1), (1, -2)$, then offer a_1 cannot be made and accepted in equilibrium.[14] Hence, the optimal privacy regime will induce only the default offer. The following sufficient condition ensures that the two offers are made under the optimal regime.

Lemma 1. *Suppose that the firm can make two offers and the first and third quadrants of the payoff space are nonempty. There exists an optimal privacy regime in which both offers are made with positive probability.*

Under the conditions of Lemma 1 the optimal privacy regime will induce both offers and the above reasoning implies the following result.

Theorem 1. *Suppose that the firm can make two offers and the first and third quadrants of the payoff space are nonempty. Every optimal privacy regime is a half-plane regime.*

Note that if no state is split between the acceptance and rejection regions in the optimal privacy regime, then there are many lines that represent that regime (in terms of which states are assigned to which regions). Moreover, the above reasoning implies that the optimal privacy regime is (generically) unique in terms of which states are assigned to which regions.[15]

4 The General Case

We now return to the general case. We begin by deriving a "revelation principle" that provides a characterization of the set of incentive-compatible "direct privacy regimes" from which the designer may choose. That is, we look at the privacy regimes in which the signal that is observed by the firm may in fact be interpreted as an instruction given by the designer to the firm regarding which offer the firm should make to the consumer. We then rely on this revelation principle to provide a geometric characterization of the optimal privacy regime for the consumer.

[14] To see this, observe that for any conditional belief on the two states, at least one player strictly prefers the default offer to a_1.

[15] It is generic because if two states lie on the same line that passes through the origin (have the same slope), then only the weighted sum of the probabilities with which they are assigned to each region is determined.

4.1 A Revelation Principle

An incentive compatible "direct privacy regime" for the designer is given by a privacy regime $\langle p(s|\theta), S\rangle$ that satisfies the following constraints.

1. $S \subseteq A$.
2. $\{p(s=a|\theta)\}_{\theta\in\Theta}$ is such that whenever the designer instructs the firm to offer $a \in S$ to the consumer, it is in the interest of the firm to do so. Namely, setting $q(a|s=a)=1$ for every "instruction" $a \in S$ is an optimal strategy for the firm.

It is convenient to describe the incentive-compatibility constraints through the restrictions they impose on the beliefs that a direct privacy regime induces on the firm and consumer.

Suppose that the designer uses a direct privacy regime $\langle p(s|\theta), S\rangle$. Denote by μ_{a_i} the firm's conditional beliefs (over Θ) when the firm receives the signal $a_i \in S$. The beliefs μ_{a_i} are derived via Bayesian updating and are given by

$$\mu_{a_i}(\theta) = \frac{\pi(\theta)p(s=a_i|\theta)}{\sum_{\theta'}\pi(\theta')p(s=a_i|\theta')}.$$

Note that in a direct privacy regime, if the firm follows the designer's instructions, then the consumer's beliefs upon being offered a_i by the firm are also given by μ_{a_i}. Moreover, our assumption that the firm commits to its strategy implies that the firm's strategy is a general rule of behavior that is observable by the consumer. Therefore, were the firm to decide not to follow the instructions it receives, the consumer would be aware that the firm had done so. Thus, for example, if the firm were to decide to offer a_j (with some probability) when it was instructed to offer a_i, then the consumer would realize that a_j is being offered in the event where it should be offered and also (with some probability) in the event where the firm is instructed by the designer to offer a_i.

Let $p_{a_i} = \sum_\theta \pi(\theta)p(s=a_i|\theta)$ denote the probability that signal $a_i \in S$ is sent by the designer. The system of beliefs $\{\mu_{a_i}\}_{a_i\in S}$ is said to be Bayes-plausible if the beliefs $\{\mu_{a_i}\}_{a_i\in S}$ satisfy the following equation:

$$\sum_{a_i\in S} p_{a_i}\mu_{a_i} = \pi(\bar{\theta}). \tag{1}$$

A well-known result in the literature on Bayesian persuasion (Kamenica and Gentzkow 2011) is that the designer can induce a system of beliefs $\{\mu_{a_i}\}_{a_i\in S}$ if and only if these beliefs are Bayes-plausible.

Suppose that the designer uses a direct privacy regime $\langle p(s=a|\theta), S\rangle$ in which the firm obeys the designer's instructions. As explained above, the consumer's beliefs conditional on receiving the offer a_i are given by μ_{a_i}. The consumer will accept the firm's offer if her payoff from doing so is no less than her payoff from the default offer. Denote the vector of payoff differences between offer a_i and the default offer by $\Delta u(a_i, \overline{\theta}) \equiv (\Delta u(a_i, \theta_1), \ldots, \Delta u(a_i, \theta_M))$. The consumer would accept offer a_i if and only if

$$\mu_{a_i} \cdot \Delta u(a_i, \overline{\theta}) \geq 0. \tag{2}$$

Denote the set of beliefs under which the consumer accepts offer a_i by

$$\Xi_{a_i} = \{\mu \in \Pi : \mu \cdot \Delta u(a_i, \overline{\theta}) \geq 0\},$$

where Π is the space of probability distributions over Θ.

The consumer's incentive-compatibility constraint (2) can therefore also be equivalently written as

$$\mu_{a_i} \in \Xi_{a_i} \quad \forall a_i \in S. \tag{3}$$

The simplest incentive-compatibility constraint for the firm captures the idea that when instructed to offer the consumer $a_i \in S$, there is no other offer that the firm (strictly) prefers to a_i that it can induce the consumer to accept.[16] That is, if there is any offer a_j that the firm prefers to a_i under belief μ_{a_i}, it must be the case that under μ_{a_j} the buyer: (i) is indifferent between accepting a_j and rejecting it, and (ii) would reject a_i. Intuitively, if this were not the case, the firm could deviate and offer a_j both after being instructed to do so and with a small probability also after being instructed to offer a_i, and the consumer would accept the offer.

To formalize this idea denote by $D(a_i)$ the set of offers that the firm (strictly) prefers to a_i under belief μ_{a_i}:

$$D(a_i) \equiv \{a_j \in A : \mu_{a_i} \cdot v(a_j, \overline{\theta}) > \mu_{a_i} \cdot v(a_i, \overline{\theta})\}.$$

To rule out the deviation described above, the following condition must hold:

$$\text{For every } a_i \in S, \text{ if } a_j \in D(i) \text{ then } \mu_{a_j} \in \partial\Xi_{a_j} \text{ and } \mu_{a_i} \notin \Xi_{a_j}, \tag{4}$$

where ∂X denotes the boundary of the set X.

The other type of incentive-compatibility constraint for the firm deals with more complex deviations.

Suppose that the firm would like to offer a_j after being instructed to offer a_i but knows that were it to change its strategy in that manner the buyer would reject a_j. One possible way in which the firm may still be able to get away with offering a_j instead of a_i in this case is if it can "sweeten" offer a_j for the consumer as follows. The firm could offer a_j not only after being instructed to offer a_i, where the consumer would reject a_j under beliefs μ_{a_i}, but

[16] Recall that the consumer is aware of the firm's deviation.

also after being instructed to offer a_k, that is selected such that the consumer would accept a_i under beliefs μ_{a_k}.

Formally, for a given direct privacy regime, denote by $T(a_j)$ the set of signals that induce beliefs under which the consumer strictly prefers a_j to a_0:

$$T(a_j) \equiv \{a_k \in S : \mu_{a_k} \in int(\Xi_{a_j})\},$$

where $int(X)$ denotes the interior of a set X. In other words, $T(a_j)$ is the set of signals that the firm can use as a sweetener to induce the consumer to accept the firm's offer of a_j instead of a_i.

In order to incentivize the consumer to accept an offer of a_j not only when the designer instructs the firm to offer a_j, but also when, with a small probability when the designer instructs the firm to offer a_i and a_k, the ratio between $p_{a_i}q(a_j|s = a_i)$ and $p_{a_k}q(a_j|s = a_k)$ must be sufficiently small.

Formally, for $a_j \in D(a_i)$ and $a_k \in T(a_j)$ define $\alpha(i, j, k)$ as the solution to the equation:

$$\alpha(p_{a_i}\mu_{a_i} \cdot \Delta u(a_j, \overline{\theta}) + (1 - \alpha)p_{a_k}\mu_{a_k} \cdot \Delta u(a_j, \overline{\theta}) = 0.$$

That is, $\frac{\alpha(i;j,k)}{1-\alpha(i;j,k)}$ is the minimal ratio between $q(a_j|s = a_i)$ and $q(a_j|s = a_k)$ that induces the consumer to accept offer a_j after such a modification.

Incentive compatibility for the firm requires that there be no pair of offers $a_i \in S$, $a_j \in D(a_i)$, and a sweetener $a_k \in T(a_j)$, for which the firm prefers to offer a_j (at the appropriate ratio) instead of following the designer's instructions, i.e.,

$$\begin{aligned}
&\alpha(i, j, k)p_{a_i}\mu_{a_i} \cdot v(a_i, \overline{\theta}) + (1 - \alpha(i, j, k))p_{a_k}\mu_{a_k} \cdot v(a_k, \overline{\theta}) \\
&\qquad \geq \Big(\alpha(i; j, k)p_{a_i}\mu_{a_i} + (1 - \alpha(i; j, k))p_{a_k}\mu_{a_k}\Big) \cdot v(a_j, \overline{\theta}) \\
&\qquad \forall a_i \in S, a_j \in D(a_i), a_k \in T(a_j).
\end{aligned} \tag{5}$$

Solving for $\alpha(i; j, k)$ and simplifying shows that this constraint is equivalent to

$$\begin{aligned}
&\mu_{a_k} \cdot (v(a_k, \overline{\theta}) - v(a_j, \overline{\theta})) \geq \frac{\mu_{a_k} \cdot \Delta u(a_j, \overline{\theta})}{-\mu_{a_i} \cdot \Delta u(a_j, \overline{\theta})}\mu_{a_i} \cdot (v(a_j, \overline{\theta}) - v(a_i, \overline{\theta})) \\
&\qquad \forall a_i \in S, a_j \in D(a_i), a_k \in T(a_j).
\end{aligned} \tag{6}$$

This version of the constraint has the following intuitive interpretation as a cost-benefit analysis. The left-hand side represents the firm's cost from using signal a_k as a sweetener to induce offer a_j. The right-hand side represents the firm's benefit from inducing the buyer to accept a_j when the firm is instructed to offer a_i, scaled by a term that measures the effectiveness of signal a_k as a sweetener for offer a_j relative to the buyer's reluctance to accept a_j under beliefs μ_{a_i}.

Finally, note that under any privacy regime the firm's problem is a linear optimization problem in $q(\cdot,\cdot)$. Thus, if the firm does not have a small improving deviation over obeying the designer's instructions, then it does not have any improving deviation, and obeying the designer's instructions is a best response. This implies there is no need to consider even more complex deviations that the firm may wish to pursue.

Theorem 2. *There exists an optimal solution to the designer's problem in which the designer employs a direct privacy regimes that satisfies the incentive-compatibility constraint for the consumer* (3)*, the incentive-compatibility constraints for the firm* (4) *and* (6)*, and the Bayes-plausibility constraint* (1)*.*

4.2 Examples of Incentive-Compatible Direct Privacy Regimes

To illustrate the nature of the firm's incentive-compatibility constraints in the belief simplex, we describe the restrictions they impose on induced beliefs in some special cases. In all of the examples below we assume for simplicity that $v(a_i, \theta) = i$ for all θ. That is, the firm prefers higher-indexed offers in all states of the world. Recall that Bayes-plausibility implies that the prior must belong to the interior of the convex hull of the induced beliefs, and that the buyer's incentive-compatibility constraints require that $\mu_{a_i} \in X_{a_i}$. In the following examples we take these constraints for granted.

Example 1: Two States, Two Offers.

Suppose that $\Theta = \{\theta_1, \theta_2\}$ and $A = \{a_0, a_1\}$. If the buyer accepts a_1 under the prior beliefs ($\pi(\overline{\theta}) \in \Xi_{a_1}$), then if the firm always offers a_1 regardless of the designer's instruction, this offer would be accepted by the consumer. In this case, there is only one possible incentive-compatible direct privacy regime, where the designer instructs the firm to offer a_1 in all states of the world.

If the buyer rejects a_1 under the prior beliefs ($\pi(\overline{\theta}) \notin \Xi_{a_1}$), then in a direct privacy regime in which a_1 is offered with a positive probability, it must be the case that μ_{a_1} lies on the boundary of Ξ_{a_1}. This is because if μ_{a_1} is in the interior of Ξ_{a_1}, then the firm has the following profitable deviation: when instructed to offer a_0, the firm offers a_1 with some small probability. The fact that μ_{a_1} is in the interior of Ξ_{a_1} implies that upon being offered a_1 the consumer still prefers it to a_0, and would accept the offer.

Figure 3 depicts incentive-compatible and non-incentive-compatible direct privacy regimes (in the left and right panels, respectively). In both panels, with a slight abuse of notation, we let μ_{a_i} denote the probability of state θ_2 after recommendation a_i.

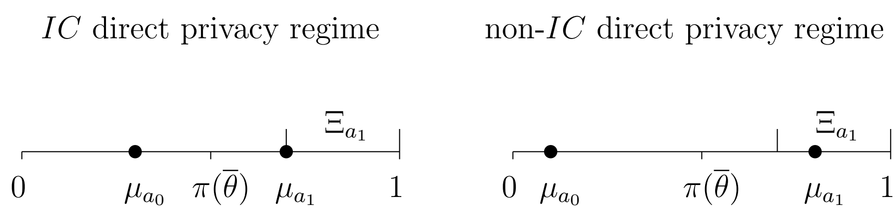

Figure 3: Incentive compatibility with two states and two offers

Example 2: Three States, Three Offers.

Suppose that $\Theta = \{\theta_1, \theta_2, \theta_3\}$ and $A = \{a_0, a_1, a_2\}$. We present two examples in which the designer induces the firm to make all three offers with a positive probability.

First, we consider the case where the sets Ξ_{a_1} and Ξ_{a_2} are disjoint and the prior $\pi(\bar{\theta}) \notin \Xi_{a_1} \cup \Xi_{a_2}$. That is, under the prior beliefs, the consumer prefers a_0 to a_1 and a_2. In this case, the firm's incentive-compatibility constraint (4) has two implications: (i) the consumer must be indifferent between accepting and rejecting each of the offers a_1 and a_2 when they are made, and (ii) the consumer must reject both a_1 and a_2 under belief μ_{a_0}. Since the sets Ξ_{a_1} and Ξ_{a_2} are disjoint, in an incentive-compatible direct privacy regime signal a_1 cannot be used as a sweetener for offer a_2 (and vice versa), and so there are no profitable "complex deviations."[17]

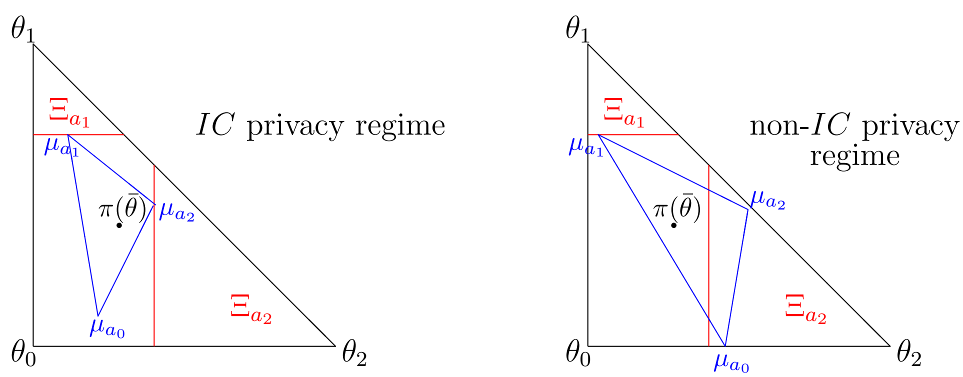

Figure 4: Incentive compatibility with three states and three offers

Note that in the right-hand panel there are two violations of the firm's simple incentive-compatibility constraint. First, the fact that μ_{a_2} is in the interior of Ξ_{a_2} implies that the firm can offer a_2 with a small probability when it is instructed to offer a_1. Second, the fact that μ_{a_0} belongs to Ξ_{a_2} implies that the firm can offer a_2 when it is instructed to offer a_0.

Next, consider the case where $\Xi_{a_2} \subset \Xi_{a_1}$ and assume that the firm is instructed to make all offers with equal probabilities. To satisfy the firm's simple incentive-compatibility

[17] To see this more formally, note that in this case $T(a_1) = T(a_2) = \varnothing$.

constraint (4), the buyer's beliefs upon receiving offer a_i, $i \neq 0$, must be on the boundary of Ξ_{a_i}. In this example, we take those restrictions for granted, and focus on whether it is profitable for the firm to offer a_1 upon being instructed to offer both a_0 and a_2. That is, can the firm benefit from using signal a_2 as a sweetener to incentivize the buyer to accept offer offer a_1 under beliefs μ_{a_0}?

In this example, constraint (5) requires that for any deviation for which $\frac{q(a_1,s=a_0)}{q(a_1,s=a_2)} = \frac{\alpha}{1-\alpha}$ that is incentive compatible for the buyer,

$$\alpha p_{a_0}\mu_{a_0} v(a_0,\overline{\theta}) + (1-\alpha)\mu_{a_2} p_{a_2} v(a_2,\overline{\theta}) \geq (\alpha p_{a_0}\mu_{a_0} + (1-\alpha) p_{a_2}\mu_{a_2}) v(a_1,\overline{\theta}).$$

Given our assumptions on v and p this becomes

$$(1-\alpha)2 \geq (\alpha + (1-\alpha)) \text{ or } \alpha \leq \frac{1}{2}.$$

That is, any complex deviation that would be acceptable to the consumer must be such that the firm offers a_1 when instructed to offer a_0 with a lower probability than the firm offers a_1 when instructed to offer a_2. Geometrically, this implies that mixing μ_{a_2} and μ_{a_0} with equal weights must lead to a belief that is not in the interior of Ξ_{a_1}. Note that the two examples in Figure 5 below depict different utility functions for the consumer.

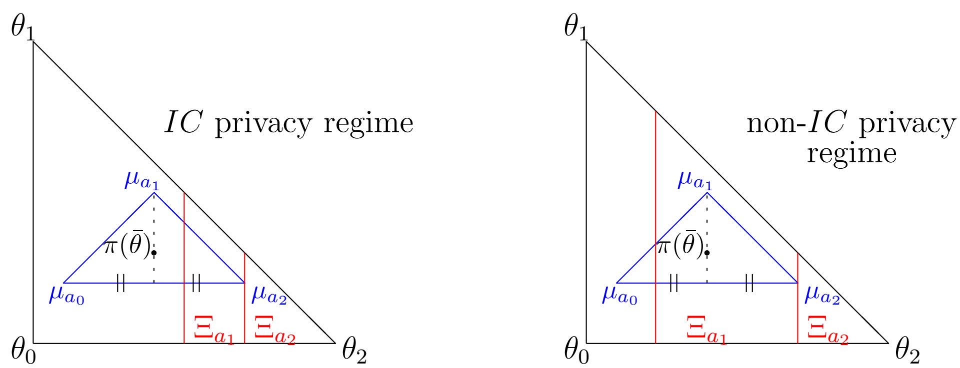

Figure 5: "Complex" incentive compatibility with three states and three offers

4.3 A Geometric Characterization of the Optimal Privacy Regime

Recall that the difference between the consumer's payoff from offer a_i and the consumer's payoff from the default offer a_0 in state θ is denoted by $\Delta u(a_i,\theta) \equiv u(a_i,\theta) - u(a_0,\theta)$. Similarly, denote the difference between the firm's profit from offer a_i and the firm's profit from the default offer a_0 in state θ by $\Delta v(a_i,\theta) \equiv v(a_i,\theta) - v(a_0,\theta)$. Denote the vector of the consumer's and the firm's payoff differences in state θ by

$$W(\theta) \equiv (\Delta u(a_1,\theta),\ldots,\Delta u(a_N,\theta),\Delta v(a_1,\theta),\ldots,\Delta v(a_N,\theta)).$$

With a slight abuse of notation, we refer to the $2N$-dimensional Euclidean space that contains the set of vectors $\{W(\theta)\}_{\theta\in\Theta}$ as the payoff space.

Theorem 1 provides a geometric characterization of the partition of the payoff space into the acceptance and rejection regions. In particular, the two regions form contiguous half-spaces that are separated by a hyperplane that passes through the origin. Clearly, with more than two actions it is impossible to separate the acceptance regions associated with different offers with a single hyperplane.

To generalize Theorem 1 we introduce the following definition:

Definition. *A set $C \in \mathbb{R}^n$ is a polyhedral cone, if there exists a matrix $A \in \mathbb{R}^{m\times n}$ such that $C = \{x \in \mathbb{R}^n : A \cdot x \geq 0\}$.*

Note that if x belongs to a polyhedral cone C, then so does αx for every $\alpha \geq 0$ (that, is a polyhedral cone is a cone), and that a polyhedral cone is convex.

An alternative presentation of Theorem 1 is that the acceptance and rejection regions are both polyhedral cones with disjoint interiors. This property holds for any number of offers.

Theorem 3. *Assume that there exists an optimal privacy regime in which offers $A^* \subseteq A$ are made. There exists a set of polyhedral cones $\{\mathcal{A}_{a_i}\}_{a_i\in A^*}$ with pairwise disjoint interiors that partition the payoff space such that $p(a_i, \theta) > 0$ only if $W(\theta) \in \mathcal{A}_{a_i}$.*

To understand the intuition behind Theorem 3 consider the problem faced by the designer. Ideally, the designer would like to assign each state to the acceptance region of the offer that provides the buyer with the highest payoff in that state. However, were the designer to try and do so the firm could use its information (i.e., the signals it receives) to induce the buyer to accept different offers in some states of nature. Hence, in addition to considering the buyer's direct benefit from adding state θ to the acceptance region of a_i, the designer must also consider how sending signal $s = a_i$ in state θ impacts the firm's ability to use signal $s = a_i$ to induce the buyer to accept offers other than a_i. That is, for each signal $s = a_i$ and offer $a_j \neq a_i$ there is a shadow cost that measures the impact of changing the expected payoffs after signal $s = a_i$ on the firm's incentive to follow instructions. In fact, the designer faces two additional shadow costs that are associated with the buyer's incentive-compatibility constraint and the feasibility constraint $\Sigma_{s\in S} p(s, \theta) = 1$, that affect him in a similar way.

Now, consider the designer's problem in deciding whether to assign state θ to the acceptance region of a_i or the acceptance region of a_j. The value (incorporating the shadow costs) the designer obtains from assigning state θ to the acceptance region of any offer is linear in $W(\theta)$. Moreover, the shadow cost associated with the feasibility constraint is an additive term that does not depend on $W(\theta)$. Hence, there exists a hyperplane that passes through the origin that separates the payoff vectors that the designer would rather assign to the acceptance region a_i from the payoff vectors that the designer would rather assign to the acceptance region[18] a_j. Similarly, for every $a_k \neq a_i$ there exists a hyperplane that separates the payoff vectors in which a_i is better than a_k from the payoff vectors in which the opposite holds. It follows that the designer instructs the firm to offer a_i in state θ only if $W(\theta)$ lies above N different hyperplanes that pass through the origin. That is, the states in which a_i is offered are contained in a polyhedral cone.

Theorem 3 may inform privacy regulation in the following way. Recall that a state represents the consumer's type, or the information that is specific to the consumer.[19] In the problem we are considering, the designer determines the information that the firm is allowed to possess about consumers. Each possible realization of this information can be interpreted as a "cell" in the information partition that the designer permits the firm to have, and is associated with the specific offer that the firm makes to the consumer upon learning this information. Theorem 3 provides a guideline for evaluating privacy policies through the emphasis it places on the dimensions of the consumers' type that should be considered. It shows that states with $W(\theta)$ that lie on nearby rays should be assigned to the same cell of the firm's information partition. That is, under optimal privacy regulation, firms should not be allowed to distinguish between consumer types with similar *ratios* between the different possible payoffs. This simple insight suggests that there may exist two consumer types for which offer a_i is best for both the consumer and the firm, but the designer may nevertheless still require the firm to distinguish between these two types because the impact of the shadow costs of a_i (as captured by the slope of $W(\theta)$) may differ for these two types.

A secondary interpretation of our result is related to the recent paper by Dworczak and Martini (2019). They suggest that Bayesian persuasion can be thought of as a general equilibrium problem where the sender uses the signals he observes about the state of the world as an endowment with which he can purchase beliefs that induce the receiver to choose actions. From this perspective, the designer's problem can be thought of as one where he selects the firm's persuasion endowment with the following objective: maximize the consumer's utility when the firm consumes its endowment (i.e., reveals its signal to the consumer) while making sure that this is indeed the best consumption bundle the firm can afford. In other words, the designer constructs the firm's "persuasion endowment" in such a way that in the resulting persuasion economy the firm will not able to afford to persuade the buyer to take actions other than those intended by the designer. According to this interpretation, it is intuitive that using signal $s = a_i$ to persuade the buyer to choose action

[18]This hyperplane passes through the origin because for the zero payoff vector the designer is indifferent between the two options.

[19]Recall that in our setting, the consumer does not know her type.

a_j must provide the firm with a low utility relative to the cost it imposes on the buyer. Theorem 3 establishes that the optimal way to do so is to construct an endowment that constitutes a collection of polyhedral cones in the payoff space.

5 The Effectiveness of Regulation

Regulation is generally effective. The expected payoff of the consumer under the optimal solution to the designer's problem is larger than her expected payoff under the two extreme benchmarks where the firm is unregulated and fully informed about the state of the world, or where regulation keeps the firm entirely uninformed. However, when stronger assumptions are imposed on the profits of the firm, it is possible to obtain sharper results. In particular, we consider the case where the firm's preferences over the offers are independent of the state of the world. This special case has received a lot of attention in the literature that studies Bayesian persuasion (e.g., the leading example of Kamenica and Gentzkow 2011).

For the rest of this section we assume that the firm's preferences over offers are independent of the state of the world. With no additional loss of generality we assume that $v(a_j, \cdot) > v(a_i, \cdot)$ if[20] $i < j$. We show that under this assumption when there are only two offers, regulation cannot help the consumer; however, this is not true when there are more than two offers.

If the firm can make two offers, then it does not get its preferred offer in the rejection region (if this region is nonempty). As explained above, the firm's incentive-compatibility constraint implies that the consumer must be indifferent between accepting and rejecting offer a_1 in the acceptance region. But if the consumer is indifferent between accepting and rejecting a_1 in the acceptance region then it follows that the consumer's expected payoff is equal to zero both in the acceptance and rejection regions. Hence, under any incentive-compatible direct privacy regime in which both offers are made, regulation fails to increase the consumer's expected payoff above her utility from the default offer.[21] Thus, under any incentive-compatible direct privacy regime, if the consumer weakly prefers a_1 to a_0 under the prior, then the firm is instructed to offer a_1. Otherwise, the consumer's utility is identical to what it would be if the firm were instructed to offer[22] a_0.

[20] As the firm can force the default offer, there is no loss in assuming that this offer is the worse one for the firm.

[21] A similar result is derived by Ichihashi (2019) in a model in which the consumer can choose a_1 even if it is not offered by the firm.

[22] Proposition 4 is a corollary of Proposition 5 below and is presented without proof.

Proposition 4. *Suppose that there are only two offers and the firm prefers a_1 to a_0 in every state of the world. Then, all privacy regimes provide the consumer with the same expected payoff.*

Although the designer cannot affect the consumer's surplus, it can certainly reduce the firm's expected profit compared to the maximum expected profit that the firm could get if it knew the state of the world. Intuitively, unless the designer allows the firm to structure the acceptance and rejection regions such that the consumer is willing to accept a_1 conditional on being offered a_1, the firm may be unable to realize its full potential expected profit.

The next example shows that with more than two offers, regulation can improve the expected payoff of the consumer compared to the situation of an unregulated and fully informed firm, albeit in a limited extent.

Example: Beneficial Regulation

Suppose that there are three equally likely states $\{\theta_1, \theta_2, \theta_3\}$ and the buyer's payoffs are

$$u(a_0, \overline{\theta}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \qquad u(a_1, \overline{\theta}) = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}; \qquad u(a_2, \overline{\theta}) = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}.$$

If the firm knows the state, then it will offer the consumer a_2 in states θ_2 and θ_3, and a_1 in state θ_1. This provides the consumer with an expected payoff of $\frac{1}{3}$.

However, if the firm's induced information partition is given by $\{\theta_1, \theta_2\}, \{\theta_3\}$, then in state θ_3 the firm would offer a_2 and in states $\{\theta_1, \theta_2\}$ the firm would offer the consumer a_1. This offer strategy gives the consumer an expected payoff of $\frac{2}{3}$.

Observe that although this example shows that regulation can help the consumer, it also shows that it can do so in a very particular way. Namely, the regulator can only provide the consumer with a payoff above the default when the lowest-indexed offer is made. The following proposition implies that this observation is in fact true in general and is not a feature of this example.

Proposition 5. *Suppose that the firm prefers higher-indexed to lower-indexed offers in all states of the world. Under any incentive-compatible direct privacy regime, the buyer is indifferent between the default offer and any offer $a_i \in A$ that is not the lowest-indexed offer that is made under that regime.*

Note that Proposition 4 is a corollary of Proposition 5. If there is only one non-default offer, then under any incentive-compatible privacy regime either a_1 is made with certainty, or the lowest-indexed offer is the default offer. In this case, as explained above, regulation cannot help the buyer. By contrast, with multiple non-default offers, there may exist (non-degenerate) incentive-compatible direct privacy regimes in which the lowest- indexed offer is different from the default offer. In this case, regulation can help the buyer, but the effect of regulation is limited to its effect on the buyer's payoff from the lowest-indexed offer and the probability with which this odder is made.

In other words, Proposition 4 shows that when the firm's preferences are independent of the state of nature, the designer must select exactly one offer from which the buyer will benefit and the states in which that offer is made, and then allocate the remaining states to higher-indexed offers in a way that satisfies the incentive-compatibility constraints. This insight not only simplifies the computation of optimal privacy regimes, but also justifies thinking about the designer's problem in the following way. In order to generate surplus for the buyer after offer a_i, the designer is willing to treat all signals $s = a_j$ for $j > i$ as "informational loss leaders." That is, the designer forgoes the opportunity to generate surplus for the buyer after signals $s = a_j$ for every $j > i$ in order to generate surplus for the buyer after signal $s = a_i$.

6 Concluding Remarks

We assume that the set of potential offers is fixed. In practice, a regulator may be able to determine which offers the firm can make (e.g., insurance regulators forbid contracts that do not meet minimal coverage conditions). Disallowing non-default offers may be a useful tool for a regulator because it reduces the set of possible deviations that the firm can consider. However, it is not clear which offers the regulator should disallow because disallowing even a dominated offer might backfire. To see this in the most transparent way, suppose that there is only one state, the buyer's preferences are $a_0 \prec a_2 \prec a_1$, and the firm's preferences are $a_1 \prec a_0 \prec a_2$. In this example, if all offers are allowed then the firm will offer a_2. But, if a_2 if forbidden in order to induce the firm to offer the buyer-dominant offer a_1, then the firm will offer a_0 instead, which is the worst offer the buyer can receive. To determine which offers should be disallowed the regulator should solve the model for every subset of offers, and determine the best subset of permissible offers.

Throughout the analysis we impose the natural assumption that the objective of the designer is to maximize the consumer's surplus. However, our geometric characterization of the optimal privacy regime does not depend on this assumption. If instead the designer's objective is to maximize a weighted average of the consumer's and firm's payoffs, then the acceptance regions under the optimal solution can still be described by polyhedral cones, that depend on the weights given to the consumer's and firm's payoffs. Intuitively, this is due to the fact that assigning states to acceptance regions according to the slope of $W(\theta)$ is the most efficient incentive-compatible way to do so, for any weights given to the consumer's and firm's payoffs.

Our analysis was predicated on the assumption that the buyer can learn about the state of the world only by making inferences from the offer she receives. This is a reasonable assumption in some contexts (e.g., genetic testing that relies on proprietary technology) but it is clearly inadequate in other contexts. However, incorporating a privately informed consumer into the analysis implies that the firm may benefit from screening the consumer's type. This considerably complicates the analysis and is beyond the scope of the present paper. We leave this subject for future research. Furthermore, notice that if the firm were compelled by law to reveal its information about the consumer's type to the consumer, this would not affect our analysis, because the consumer can already infer all the firm's information from the offer she receives.

Finally, we wish to reiterate our observation that our approach suggests that laws against discrimination based on genetic and other private information may be too sweeping to be effective. Regulators may be able to promote social welfare by following a more nuanced approach that takes into account both benefits to consumers and firms' profits and pays close attention to incentive-compatibility considerations.

Appendix: Proofs

Proof of Lemma 1. We prove this lemma by establishing two claims. First, we show that if quadrant three of the payoff space is nonempty, then a_0 is offered (with some probability) in any optimal privacy regime. Second, we show that if quadrant one of the payoff space is nonempty, then there exists an optimal privacy regime in which a_1 is offered.

Suppose by way of contradiction that $\tilde{\theta}$ belongs to the third quadrant and a_1 is offered with probability one under an optimal privacy regime. Observe that moving $\tilde{\theta}$ to the rejection region increases the buyer's expected utility. Moreover, this modification relaxes the buyer's incentive-compatibility constraint after the buyer receives offer a_1. Now, consider the impact of this change on the firm's incentive-compatibility constraints. Since $\tilde{\theta}$ belongs to the third quadrant, in the modified rejection region the firm strictly prefers a_0 to a_1. By assumption, under the prior the firm prefers a_1 to a_0 (were this not the case, offering a_1 with probability one would not be an equilibrium), and so conditional on $\theta \in \Theta \backslash \{\tilde{\theta}\}$ the firm also prefers a_1 to a_0. Hence, moving $\tilde{\theta}$ to the rejection region increases the payoffs of both players and is incentive compatible for both players.

Next, assume that there exists a state $\tilde{\theta}$ in the first quadrant and that the designer selects $S = \{s_0, s_1\}$, where s_1 is realized if and only if the state is $\tilde{\theta}$. After s_1 the buyer strictly prefers a_1 to a_0. Hence, offering a_i after s_i is incentive compatible for the buyer. Moreover, since after s_1 the firm also strictly prefers a_1 to a_0 the strategy of offering a_0 with probability one is suboptimal for the firm under this privacy regime. Because the firm faces a linear optimization problem, it has a best response to this privacy regime. Furthermore, under this best response it must make both offers. Since the buyer's expected utility against this strategy is at least that of receiving the default offer with certainty, it follows that having the firm make the default offer is not the unique optimal privacy regime.

□

Proof of Theorem 1. By Lemma 1 under any optimal privacy regime the set of offers that are made is $A^* = \{a_0, a_1\}$. The result then follows as a special case of the more general Theorem 3 that we derive in Section 4. □

Proof of Theorem 2. First, we show that that outcome obtained under an arbitrary privacy regime can be replicated by a direct privacy regime.

Consider an arbitrary information structure $\hat{S}, \hat{p}(\cdot|\cdot)$ and an offer strategy, $\hat{q}(\cdot|\cdot)$ that is incentive compatible for the buyer and an optimal choice for the firm. Define a direct privacy regime that generates the same joint distribution of actions and states as follows:

$$\tilde{S} = \{a_i : \hat{q}(a_i|s) > 0 \text{ for some } s \in \hat{S}\}$$
$$\tilde{p}(a|\theta) = \sum_{s \in \hat{S}} \hat{q}(a|s)\hat{p}(s|\theta).$$

Note that the buyer's interim belief conditional on any offer made by the firm is the same under both systems. Hence, this direct privacy regime is incentive compatible for the buyer.

Next, we show that selecting $q(a_i, s = a_i) = 1$ for all $a_i \in \tilde{S}$ is an optimal choice for the firm. Assume by way of contradiction that there exists an offer strategy $\tilde{q}(\cdot|\cdot)$ that the firm strictly prefers to following the designer's instructions. This implies that under the information structure $\hat{S}, \hat{p}(\cdot|\cdot)$ the firm prefers the recommendation function

$$q(\cdot|s) = \tilde{q}(\cdot|a)\hat{q}(a|s) \quad \forall s \in \hat{S}$$

to $\hat{q}(\cdot|\cdot)$, a contradiction.

Next, observe that for any privacy regime, the firm's problem is a linear optimization problem in $q(\cdot, \cdot)$. Thus, if the firm does not have a small improving deviation over obeying the designer's instructions, then it does not have any improving deviation, and obeying the designer's instructions is a best response. This implies that there is no need to consider more complex deviations that the firm may wish to pursue. □

Proof of Theorem 3. Assume that there exists an optimal direct privacy regime that induces the offers $A^* \subset A$ via beliefs $\mu^*_{a_i}$ that are induced with probabilities $p^*_{a_i}$, respectively. Denote the offers that the buyer is indifferent to taking in this solution by

$$\overline{A} = \{a_i \in A^* : \mu_{a_i} \in \partial\Xi_{a_i}\}.$$

Then define the constraint

$$\mu_{a_i} \in \partial\Xi_{a_i} \quad \forall a_i \in \overline{A}. \tag{7}$$

Since $A^*, \{p^*_{a_i}, \mu^*_{a_i}\}$ is an optimal direct privacy regime, it follows that the solution to the designer's problem subject to the buyer's incentive-compatibility constraint (3), the firm's incentive-compatibility constraints (4) and (6), and the Bayes-plausibility constraint (1) is also a local maximum to the designer's problem under constraints (7), (6), and (1). Similarly, let $\overline{B} \subset A^* \times A^* \times A^*$ denote the set of triplets for which (6) is binding under $\{p^*_{a_i}, \mu^*_{a_i}\}$, i.e.,

$$\overline{B} = \{(a_i, a_j, a_k) \in (A^*)^3 : \mu_{a_k} \cdot (\overline{v}(a_k) - \overline{v}(a_i)) = \frac{\mu_{a_k} \cdot \Delta u(a_i, \overline{\theta})}{\mu_{a_j} \cdot \Delta u(a_i, \overline{\theta})} \mu_{a_j} \cdot (v(a_j, \overline{\theta}) - v(a_i, \overline{\theta}))\}. \tag{8}$$

Since $A^*, \{p^*_{a_i}, \mu^*_{a_i}\}$ is an optimal information partition, it follows that any local maximum to the designer's problem with constraints (7), (6), and (1) is also a local maximum under constraints (7), (8), and (1). Because (7), (8), and (1) are equations (rather than inequalities), the local maxima of this problem are given by the solution to the following

Lagrangian:

$$L_1 = \sum_{a_i \in A^*} p_{a_i}\Big(\mu_{a_i} \cdot \Delta u(a_i, \overline{\theta})\Big) + u(a_0, \overline{\theta}) \cdot \pi(\overline{\theta}) + \sum_{i \in \overline{A}} \gamma_i \Big(\mu_{a_i} \cdot \Delta u(a_i, \overline{\theta})\Big) +$$

$$\sum_{(a_i, a_j, a_k) \in \overline{B}} \gamma_{i,j,k} \Big(\mu_{a_k} \cdot (v(a_k, \overline{\theta}) - v(a_i, \overline{\theta}))(\mu_{a_j} \cdot \Delta u(a_i, \overline{\theta})) - \mu_{a_k} \cdot \Delta u(a_i, \overline{\theta}) \mu_{a_j} \cdot v(a_j, \overline{\theta}) - v(a_i, \overline{\theta}))\Big)$$

$$+ \overline{\gamma} \cdot \left(\sum_{a_i \in A^*} p_{a_i} \mu_{a_i} - \pi(\overline{\theta}) \right).$$

Recall that $\{p^*_{a_i}, \mu^*_{a_i}\}$ is generated by some privacy regime and recommendation function, and so we can rewrite the Lagrangian in terms of the signals $p(a_i|\theta_l), S$, where $S = A^*$ and the firm uses the offer function given by $R(s, s) = 1$; that is, makes the offer that corresponds to the signal that the firm receives. Doing so and subtracting the constant $\overline{u}(a_0) \cdot \overline{\pi}$ from the objective gives the following Lagrangian:

$$\begin{aligned} L = & \sum_{a_i \in A^*} \sum_{\theta \in \Theta} \pi(\theta) p(a_i, \theta) \Delta u(a_i, \theta) \\ & + \sum_{a_i \in \overline{A}} \lambda_i \Big\{ \sum_{\theta} \pi(\theta) p(a_i, \theta) \Delta u(a_i, \theta) \Big\} \\ & + \sum_{(a_i, a_j, a_k) \in \overline{B}} \lambda_{i,j,k} \Bigg\{ \left(\sum_{\theta \in \Theta} \pi(\theta) p(a_k, \theta)(\Delta v(a_k, \theta) - \Delta v(a_i, \theta)) \right) \left(\sum_{\theta} \pi(\theta) p(a_j, \theta) \Delta u(a_i, \theta) \right) \\ & - \left(\sum_{\theta \in \Theta} \pi(\theta) p(a_k, \theta) \Delta u(a_i, \theta) \right) \left(\sum_{\theta \in \Theta} \pi(\theta) p(a_j, \theta)(\Delta v(a_j, \theta) - \Delta v(a_i, \theta) \right) \Bigg\} \\ & + \sum_{\theta \in \Theta} \lambda_\theta \left\{ \sum_{a_i \in A^*} p(a_i, \theta) - 1 \right\}. \end{aligned}$$

The derivative of L with respect to $p(a_l, \theta)$ is

$$\begin{aligned} \frac{\partial L}{\partial p(a_l, \theta)} = & \pi(\theta) \Delta u(a_l, \theta) + \lambda_i \pi(\theta) \Delta u(a_l, \theta) + \lambda_\theta \\ & + \sum_{a_i, a_k : (a_i, a_l, a_k) \in \bar{B}} \lambda_{i,l,k} \Bigg\{ \left(\sum_{\theta' \in \Theta} \pi(\theta') p(a_k, \theta')(\Delta v(a_k, \theta') - \Delta v(a_i, \theta')) \right) (\pi(\theta) \Delta u(a_i, \theta)) \\ & - \left(\sum_{\theta' \in \Theta} \pi(\theta') p(a_k, \theta') \Delta u(a_i, \theta') \right) (\pi(\theta)(\Delta v(a_l, \theta) - \Delta v(a_i, \theta)) \Bigg\} \\ & + \sum_{a_i, a_j : (a_i, a_j, a_l) \in \bar{B}} \lambda_{i,j,l} \Bigg\{ \pi(\theta) \left(\Delta v(a_l, \theta) - \Delta v(a_i, \theta)\right) \left(\sum_{\theta' \in \Theta} \pi(\theta') p(a_j, \theta') \Delta u(a_i, \theta') \right) \\ & - \pi(\theta) \Delta u(a_i, \theta) \left(\sum_{\theta' \in \Theta} \pi(\theta') p(a_j, \theta')(\Delta v(a_j, \theta') - \Delta v(a_i, \theta') \right) \Bigg\}. \end{aligned}$$

Denote by λ^* the value of the Lagrange multipliers in the solution. Evaluating this derivative for the optimal privacy regime gives

$$\begin{aligned}
&\Delta u(a_l,\theta) + \qquad \lambda_i^* \quad \Delta u(a_l,\theta) + \lambda_\theta^* \\
&+ \sum_{a_i,a_k:(a_i,a_l,a_k)\in\bar{B}} \lambda_{i,l,k}^* \Big\{ p_{a_k}^* \mu_{a_k}^* \cdot (\Delta v(a_k,\overline{\theta}) - \Delta v(a_i,\overline{\theta}))\Delta u(a_i,\theta) \\
&\qquad\qquad - \left(p_{a_k}^* \mu_{a_k}^* \cdot \Delta u(a_i,\overline{\theta})\right) ((\Delta v(a_l,\theta) - \Delta v(a_i,\theta)) \Big\} \\
&+ \sum_{a_i,a_j:(a_i,a_j,a_l)\in\bar{B}} \lambda_{i,j,l}^* \Big\{ ((\Delta v(a_l,\theta) - \Delta v(a_i,\theta))) \left(p_{a_j}^* \mu_{a_j}^* \cdot \Delta u(a_i,\overline{\theta})\right) \\
&\qquad\qquad - (\Delta u(a_i,\theta)) \left(p_{a_j}^* \mu_{a_j}^* \cdot (\Delta v(a_j,\overline{\theta}) - \Delta v(a_i,\overline{\theta})\right) \Big\}.
\end{aligned}$$

Observe that all the dot products do not depend on $W(\theta)$, and so this derivative can be represented as $z_l \cdot W(\theta) + \lambda_\theta^*$ for some $z_{a_l} \in \mathbb{R}^{2N}$. Recall that in the solution to the designer's problem $p(a_l,\theta) > 0$ only if in optimum $\frac{\partial L}{\partial p(a_l,\theta)} \geq 0$. Hence, if there exists a θ for which offers a_l and a_k are made with (strictly) positive probability, then $\frac{\partial L}{\partial p(a_l,\theta)} = \frac{\partial L}{\partial p(a_k,\theta)} = 0$ and so $z_{a_l} \cdot W(\theta) = z_{a_k} \cdot W(\theta) = -\lambda_\theta^*$. Note that the set $\{W \in \mathbb{R}^{2N} : z_{a_l} \cdot W = z_{a_k} \cdot W\}$ is given by a hyperplane $Z_{a_l a_k}$ that contains the origin.

Moreover, for every state θ such that $W(\theta)$ is above $Z_{a_l a_k}$, we have that $\frac{\partial L}{\partial p(a_l,\theta)} > \frac{\partial L}{\partial p(a_k,\theta)}$. Therefore, it follows that if θ is below $Z_{a_l a_k}$ (for some $a_k \in A^*$) it must be that $p(a_l,\theta) = 0$. Hence, the set of states in which offer a_l may be made is contained in the set $\mathcal{A}_l = \cup_{a_k \neq a_l} Y^+_{a_l a_k}$ where $Y^+_{a_l a_k}$ is the half-space above the hyperplane $Z_{a_l a_k}$. Moreover, $p(a_l,\theta) = 1$ if $W(\theta) \in int(\mathcal{A}_{a_l})$. Thus, the set of states in which offer a_l is made is contained in a polyhedral cone. Note that the union $\cup_{a_i:a_i\in A^*}\mathcal{A}_{a_i}$ must cover the entire payoff space. If there exists a payoff vector $W(\theta)$ that is not included in any $\mathcal{A}_{a_i}$, then an optimal offer does not exist for for that payoff vector under the utility function given by the Lagrangian L. Since the set of offers is finite, this leads to a contradiction.

□

Proof of Proposition 5. The proof of the proposition follows immediately from the assumption that the firm prefers higher-indexed offers to lower-indexed ones and the firm's incentive-compatibility constraint (4). To see this, note that under the assumed preferences, $D(a_i) = \{a_j : j > i\}$. Consequently, condition (4) requires that for every $a_i, a_j \in S$ such that $j > i$ it holds that $\mu_{a_j} \in \partial\Xi_{a_j}$. □

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature*, 54(2): 442–492.

Alonso, Ricardo, and Odilon Camara. 2016. "Persuading Voters." *American Economic Review*, 106(11): 3590–3605.

Bergemann, Dirk, and Stephen Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature*, 57(1): 44–95.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The Limits of Price Discrimination." *American Economic Review*, 105(3): 921–957.

Darby, Michael R., and Edi Karni. 1973. "Free Competition and the Optimal Amount of Fraud." *The Journal of Law and Economics*, 16(1): 67–88.

Dulleck, Uwe, and Rudolf Kerschbamer. 2006. "On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods." *Journal of Economic Literature*, 44(1): 5–42.

Dworczak, Piotr, and Giorgio Martini. 2019. "The Simple Economics of Optimal Persuasion." *Journal of Political Economy*, 127(5): 1993–2048.

Galperti, Simone, and Jacopo Perego. 2019. "Belief Meddling in Social Networks: An Information-Design Approach."

Gentzkow, Matthew, and Emir Kamenica. 2017. "Competition in Persuasion." *The Review of Economic Studies*, 84(1): 300–322.

Ichihashi, Shota. 2019. "Limiting Sender's Information in Bayesian Persuasion." *Games and Economic Behavior*, 117: 276–288.

Ichihashi, Shota. 2020. "Online Privacy and Information Disclosure by Consumers." *American Economic Review*, 110(2): 569–595.

Kamenica, Emir. 2019. "Bayesian Persuasion and Information Design." *Annual Review of Economics*, 11(1): 249–272.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. 2017. "Persuasion of a Privately Informed Receiver." *Econometrica*, 85(6): 1949–1964.

Li, Fei, and Peter Norman. 2019. "Sequential Persuasion." *UNC Working Paper.*

Matyskova, Ludmila. 2018. "Bayesian Persuasion with Costly Information Acquisition." *CERGE-EI Working Paper.*

Roesler, Anne-Katrin, and Balázs Szentes. 2017. "Buyer-Optimal Learning and Monopoly Pricing." *American Economic Review*, 107(7): 2072–2080.