

WORKING PAPERS

N° TSE - 936

July 2018

“Optimal Public Information Disclosure by Mechanism Designer”

Takuro Yamashita

Optimal Public Information Disclosure by Mechanism Designer*

Takuro Yamashita[†]

July 18, 2018

Abstract

We consider a mechanism design environment where a principal can partially control agents' information before they play a mechanism (e.g., a seller disclosing quality information). Assuming that the principal can ex ante commit to his disclosure policy, this is a Bayesian persuasion problem with an endogenous payoff function as a consequence of optimal mechanism design. We first show that, if the principal's and agents' information are independent or affiliated, and if

*I am grateful for Thomas Mariotti, Balazs Szentes, Ricardo Alonso, Ronny Razin, Alessandro Bonatti, Andrew Rhodes, Andy Zapechelnuk, Marco Ottaviani, Tymofiy Mylovanov, Jean-Charles Rochet, Frederic Koessler, Philippe Jehiel, Emir Kamenica, seminar participants at LSE, Hitotsubashi, Toulouse, Contract Theory Workshop (Osaka), SUFE (Shanghai), LMU (Munich), Cambridge, Bocconi, SMU (Singapore), PSE (Paris). I am indebted to Tong Su, a former PhD student in Toulouse School of Economics, for his contribution in early stages of the paper (see also our companion paper, Su and Yamashita (2014), "Mechanism design and information disclosure"). This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No 714693)

[†]Toulouse School of Economics, University of Toulouse, Capitole.
takuro.yamashita@tse-fr.eu

the implementable set of (non-monetary) allocation rules is invariant to disclosure policies, then it is optimal for the principal to disclose all the relevant information. In case of negative correlation or in case the set of implementable allocation rules varies with disclosure policies, then full disclosure is not necessarily optimal. We then characterize the optimal (non-full) disclosure policy under additional assumptions, which, viewed as a Bayesian persuasion problem, provides a solution to a novel class of environments.

1 Introduction

In this paper, we consider a mechanism design environment where a principal can partially control agents' information before they play a mechanism.

For example, a seller of a second-hand car, who designs a monopoly-pricing mechanism, may disclose a quality certificate of the car. As another example, a regulation authority, who designs a regulation mechanism, may announce some statistics about the future economic situation. In both cases, the principal (seller / authority) may not fully disclose the true information: he may have an incentive to strategically control the quality of the information in order to manipulate the agents' (buyers / industries) behavior. This paper aims to develop some tools to analyze the optimal information disclosure by the principal and to study the interaction between the optimal mechanism design and information disclosure.

The crucial assumption in this paper is that the principal can commit (at the ex ante stage) to any information disclosure policy, as well as to any mechanism. Methodologically, this commitment assumption makes our problem a

Bayesian persuasion problem.¹ A simple interpretation of this commitment assumption is that the principal himself does not observe his information, but can ask a third-party certifier to generate (possibly noisy) hard evidence about it. As is often the case in practice, this hard evidence is assumed to become public information.² Because of this commitment assumption to public disclosure, our approach differs from the (non-committed) signaling literature such as in the informed-principal literature,³ and, in environments with multiple agents, differs from the private-disclosure literature.⁴ These assumptions are admittedly restrictive: a principal in practice may enjoy some commitment power, though not as strongly as assumed in this paper. Nevertheless, we believe that the analysis in this paper could be a useful benchmark by providing some basic economic tradeoffs for a principal who can affect both agents' information and their allocations. Moreover, we argue that our main messages would be robust with respect to some timing and commitment assumptions (Remark ??).

Hence, in our setting, information disclosure means that *both* the principal and the agents become more informed, and therefore, its basic tradeoff is as follows. For the principal's side, more disclosure means that his mechanism can be contingent on more precise information. This *flexibility effect*

¹See, for example, Kamenica and Gentzkow (2011), Rayo and Segal (2010), Gentzkow and Kamenica (2015), and Kolotilin et al. (2015)

²See Dranove and Jin (2010) for public certification and other information disclosure exercises in practice. This assumption excludes a possibility that the principal observes a certificate *without* disclosing it to the agent (or equivalently, the agent must make all the decisions before he observes the certificate), as studied by Skreta (2011).

³See, for example, Myerson (1983), Maskin and Tirole (1990, 1992), Mylovanov and Tröger (2012), and Koessler and Skreta (2016).

⁴See, for example, Bergemann and Morris (2013), Kolotilin et al. (2015), and Bergemann et al. (2017).

in the mechanism choice makes, given everything else equal, the principal favor more disclosure. For the agent’s side, more disclosure means that his incentive compatibility and participation constraint become more stringent (e.g., under full disclosure, these constraints must be satisfied “ex post”, that is, for each realization of the principal’s information; while under no disclosure, they must be satisfied only “on average”). More generally, the set of implementable allocation rules may become smaller by more disclosure. This *implementability effect* makes, given everything else equal, the principal favor less disclosure. The optimal disclosure policy is determined by appropriately balancing these two effects.⁵

Our first main result builds on the simple idea that, if the second implementability effect is null, then full disclosure is optimal for the principal. This simple observation unifies some known results in the literature: a class of single-agent, monotone and quasilinear environments with affiliated information, such as the monopoly-pricing environment of Ottaviani and Prat (2001); and a class of multi-agent, monotone and linear environments with independent information, such as the auction environment of (a benchmark case of) Eső and Szentes (2007). Those results are obtained in quite different environments (and with different motivations), based on different economic intuitions. We uncover a hidden connection between those results based on the nullity of the second implementability effect above, which is appli-

⁵Note that, in our paper, any certification policy is assumed to be costless for the principal, which, in particular, excludes the (third-party) certifier’s pricing decision. This simplifying assumption is made in order for us to focus on the study of basic tradeoffs brought by information disclosure in the mechanism design stage, that is, the effect of information on the flexibility of the mechanism choice and on the agents’ incentives. See, for example, Lizzeri (1999) and Stahl and Strausz (2017), as the study of certifier’s incentive in his pricing decision.

cable more generally in environments other than revenue maximization in monopoly pricing and auction.

This general insight is also useful in understanding the limit of the full-disclosure result. We show that, in certain other environments (e.g., negatively correlated information, non-monotone or nonlinear payoffs, and restriction on monetary transfers such as due to a budget-balance restriction) the implementability effect can be significant, and hence full disclosure may not be optimal. For example, in the bilateral-trade environment of Myerson and Satterthwaite (1983), full disclosure is strictly suboptimal under very mild assumption.

Given that full disclosure is not optimal in certain class of environments, a natural next question is characterization of the (not-necessarily-full) optimal disclosure policy. As is known in the Bayesian persuasion literature, characterization of the optimal policy is a hard problem, especially when the principal's information is a continuous random variable.⁶ As in Kolotilin et al. (2015) and Gentzkow and Kamenica (2015), we assume that the agents only care about the posterior mean (rather than the entire posterior distribution) so that the principal's choice variable is a one-dimensional function, a distribution over the posterior means. Although the problem is still infinite-dimensional, under a weak regularity assumption on the payoff functions, we characterize the optimal information disclosure policy, and show that it can always be implemented by a combination of full-disclosure regions (i.e., if the realization of the principal's information lies in this region, it is (fully) disclosed) and "binary lower-truncation" regions (i.e., if the realization lies

⁶On the other hand, the shape of the sender's value function is characterized as a concavification in a quite general environment by Kamenica and Gentzkow (2011).

in this region, only its (randomly-chosen) lower bound is disclosed; moreover, only two realizations are possible in this region). This means that the principal's result is now finite-dimensional (under additional conditions, it can be as simple as even one- or two-dimensional), and hence many standard techniques can straightforwardly be applied.

1.1 Some related literature

In terms of the motivation for studying the effect of the principal's information on implementability, our work is related to the informed-principal literature, where the principal fully knows his information but cannot generate any hard evidence by himself. There are two main differences. First, in our paper, the principal can commit, at the ex ante stage, to any disclosure policy, while in the informed-principal literature, he does not have such a commitment power. Second, in our paper, the principal cannot make a mechanism contingent on any non-disclosed part of the principal's information, while in the the informed-principal literature, a mechanism can be contingent on the principal's information (without disclosing it to the agent), although such contingency must be consistent with the principal's own incentive constraints. In this sense, the comparison of these two approaches are not trivial, and one might wonder if our results crucially depend on the assumption that a mechanism cannot be contingent on the non-disclosed part of the principal's information. In Remark ??, we conjecture that similar methodology as in this paper would be useful in an extended model where a mechanism can be contingent *both* on the non-disclosed part of the principal's information (as in the informed-principal literature, i.e., subject to

the principal’s incentive constraints) and also on the disclosed part of the principal’s information as in our main model.

In the literature of committed information disclosure in mechanism design, Esó and Szentes (2007) (except for their “benchmark case”), Bergemann and Pesendorfer (2007), Li and Shi (2015), and Zhu (2017) consider situations where the disclosed information by the principal is not observed by the principal himself (but only by the agents).⁷ Their assumption fits well, for example, to a situation where a seller of an experimental good (i.e., a buyer does not know his true value before consuming it) can allow for its “trial” to buyers. The seller can perhaps control the quality of the signal delivered by the trial, but cannot usually observe the signal realization (and hence the price cannot be directly contingent on it). On the other hand, our assumption that everyone (including the principal) can publicly observe the signal realization fits better to the context of third-party certification. In this sense, these two approaches are complementary.

Methodologically, our paper is most related to the Bayesian-persuasion literature. Our contribution to the Bayesian-persuasion literature is two-fold. First, we characterize the shape of the “sender’s” (principal’s) payoff function in mechanism design contexts as a consequence of optimal mechanism design, while in the standard Bayesian persuasion model, the sender’s payoff function is usually exogeneously given. Indeed, it has been considered as an important question in the literature under what conditions the principal’s expected payoff given the optimal mechanism is maximized by the full-disclosure policy. Second, we characterize the optimal information dis-

⁷See also Bergemann and Wambach (2015) and Esó and Szentes (2017) for similar approaches in dynamic environments.

closure policy in a class of environments where the sender’s information is a *continuous* random variable, as a combination of full-disclosure regions and “lower-truncation” regions. Gentzkow and Kamenica (2015), Kolotilin et al. (2015), Kolotilin (2016), and Dworzak and Martini (2017) study such continuous cases under various assumptions. Gentzkow and Kamenica (2015) consider cases where a receiver’s action is either binary or trinary, implying that a sender’s payoff is a two- or three-step function of the posterior mean. Similarly, a receiver in Kolotilin et al. (2015) has a binary action space, but he also has (independent) private information, and hence a sender’s payoff is in general more complicated than simple step functions. Under the assumption that the sender’s payoff function is linear in the receiver’s payoff, they characterize the optimal disclosure policy, and show that it can be interpreted as a “censorship” policy. Our characterization is obtained under less structures and hence could be useful in more general environments. Kolotilin (2016) and Dworzak and Martini (2017) apply duality theory to establish that, under fairly general conditions, the (primal) problem of finding an optimal disclosure policy is equivalent to the (dual) problem of choosing a function which satisfies some lower bound conditions. Their dual-based approaches are particularly useful when the researcher has some candidate optimal policies to confirm its optimality. Our primal-based approach is useful when the researcher does not yet have an idea for an optimal policy. Also, its simple implementation as combination of full disclosure and truncation policies is novel (see also Footnote 26).

The paper is structured as follows. We consider a single-agent model in Section 2, and a multi-agent model in Section 3. The main results in these

sections are: (i) in linear environments with independent or affiliated information, full disclosure is optimal (Theorem 1, 2, and 3); and (ii) in other environments, full disclosure may not be optimal (Example 1 and Theorem 4). Given that full disclosure is not optimal in certain class of environments, Section 4 characterizes the (not-necessarily-full) optimal information disclosure policy under additional assumptions. Section 5 concludes. The proofs are in Appendix and Supplementary Materials.

2 Single agent

2.1 Environment

In this section, we consider a situation where there is one principal and one agent.

The agent has payoff-relevant private information (“type”) $t \in T = [0, 1]$. Another payoff-relevant information, whose disclosure is controlled by the principal, is denoted by $\theta \in \Theta = [\underline{\theta}, \bar{\theta}] \subseteq \mathbb{R}$.⁸ At the ex ante stage, the principal and agent share a common prior for $(\theta, t) \in \Theta \times T$. The marginal prior for θ is denoted by F_0 with density f_0 .⁹ The conditional for $t|\theta$ is denoted by $F_1(\cdot|\theta)$ for each θ with density $f_1(\cdot|\theta)$. In Section 2.1, θ, t are assumed to be independently drawn. In Section 2.2, θ, t are allowed to be correlated, and in particular, are affiliated.

⁸Some results can be generalized for the case where Θ is a more general space. For example, Theorem 1 and 3 hold true for any Θ that is a separable, complete metric space. Theorem 2 holds true for the same domain if endowed with a partial order (monotonicity conditions in Section 2.4 are defined based on that partial order).

⁹In this paper, by abuse of notation, I treat a probability measure for a real random variable and its representation as a cumulative distribution function interchangeably.

The principal assigns an allocation $\chi \in X$ through a mechanism. The principal's payoff is given by $u_0(\chi, \theta, t)$, and the agent's payoff is given by $u_1(\chi, \theta, t)$. Later, we impose more assumptions on X , u_0 and u_1 .

The timing of the game is as follows. First, the principal chooses an *information disclosure* strategy (M, G) , where M is a measurable space and $G : \Theta \rightarrow \Delta(M)$. We interpret M as the principal's message space, and G as the rule that generates a randomized message in M , whose distribution depends on the realized state $\theta \in \Theta$. Second, after $m \in M$ is publicly observed, the principal designs a direct mechanism $\chi_m : T \rightarrow X$. Finally, the agent sends a message to the mechanism, and the allocation is realized.

There are several implicit assumptions in this formulation. First, the principal has a commitment power both in terms of the mechanism and information disclosure policy. Second, the allocation can depend on any disclosed information (m), but not on any non-disclosed part of θ . Also, m is observed equally by any agent of any type ("public" disclosure). Although we believe that this formulation is sensible in certain applications and at least as a benchmark, admittedly there are many other possible alternative formulations. We briefly discuss them in the concluding remarks.

2.2 Principal's problem

2.2.1 Agent's posterior

The principal's message m is important because it affects the agent's posterior for θ and hence affects his incentive. Thus, our first step to understand the problem is to compute the agent's posterior given m . However, unless θ, t

are independently distributed, the agent's posterior for θ given m depends on his type t , which makes the problem nontrivial.

A useful observation is that the principal's message m can correlate with t only through θ . In other words, regardless of the principal's information disclosure, the *conditional* distribution $t|\theta$ cannot be affected, and hence is always given by $F_1(\cdot|\theta)$. Therefore, once a distribution *solely* of θ is obtained (given m), say $\Psi_0(\cdot|m)$, then the distribution *jointly* for (θ, t) is given by the product of $\Psi_0(\cdot|m)$ and $F_1(\cdot|\theta)$. Then, the conditional of $\theta|(t, m)$ is obtained in the standard way.

Existence of a distribution of $\theta|m$ is given by the following well-known result in probability theory.

Proposition 1. (Product-regular-conditional-probability property¹⁰) Each (M, G) induces a pair $(\mu, (\Psi^m)_{m \in M})$ such that (i) $\mu \in \Delta(M)$, (ii) $\Psi^m \in \Delta(\Theta)$ for each m , and (iii) $\int_{m \in B} \Psi^m(A) d\mu = \int_{\theta \in A} G(B|\theta) dF_0(\theta)$ for each measurable $A \subseteq \Theta$ and $B \subseteq M$.

μ is the marginal distribution over the message space M , and for each realization $m \in M$, Ψ^m is the posterior over Θ . The last condition implies (by taking $B = M$) $\int_{m \in M} \Psi^m(A) d\mu = F_0(A)$ for each A , i.e., the system of posterior distributions $(\Psi^m)_{m \in M}$ must satisfy a martingale property.¹¹

As discussed above, given $\Psi_0 \in \Delta(\Theta)$ induced by an arbitrary message, the agent's posterior for θ given his type t is given by $\Psi_1(\cdot|t, \Psi_0)$, where for

¹⁰The crucial assumption for this property to hold is that Θ is a separable, complete metric space. On the other hand, no assumption is necessary on M (except that it is a measurable space). For the proof, see Faden (1985), for example.

¹¹This is called *Bayesian plausibility* by Kamenica and Gentzkow (2011).

each measurable subset $\Theta' \subseteq \Theta$,

$$\Psi_1(\Theta'|t, \Psi_0) = \frac{\int_{\theta \in \Theta'} f_1(t|\theta) d\Psi_0(\theta)}{\int_{\theta \in \Theta} f_1(t|\theta) d\Psi_0(\theta)}.$$

2.2.2 Mechanism design

Because the principal essentially makes a sequential decision of information disclosure and then mechanism design, we first consider the mechanism-design problem.

Given $\Psi_0 \in \Delta(\Theta)$ induced by an arbitrary message, let $\Pi(\Psi_0)$ denote the principal's expected payoff under the optimal mechanism, i.e.,

$$\begin{aligned} \Pi(\Psi_0) &= \sup_{\chi: T \rightarrow X} \int_{\theta \in \Theta} \int_{t \in T} u_0(\chi(t), t, \theta) dF_1(t|\theta) d\Psi_0(\theta) \\ &\text{sub. to} \quad \int_{\theta \in \Theta} u_1(\chi(t), t, \theta) d\Psi_1(\theta|t, \Psi_0) \\ &\quad \geq \int_{\theta \in \Theta} u_1(\chi(t'), t, \theta) d\Psi_1(\theta|t, \Psi_0), \quad \forall t, t', \\ &\quad \int_{\theta \in \Theta} u_1(\chi(t), t, \theta) d\Psi_1(\theta|t, \Psi_0) \geq 0, \quad \forall t. \end{aligned}$$

The first constraint corresponds to the agent's Bayesian incentive compatibility condition, and the second constraint is the agent's interim individual rationality or participation condition. Because the agent observes a public message by the principal, the agent's expected utility is computed using his posterior $\Psi_1(\cdot|t, \Psi_0)$ over θ .

2.2.3 Information disclosure

Next, we consider the information disclosure problem. Recall that, by Proposition 1, each disclosure policy (M, G) induces a distribution (μ) over the set of posteriors $((\Psi^m)_{m \in M})$ that satisfies the martingale property. The converse is also true: for each distribution over the set of posteriors, denoted by $\lambda \in \Delta(\Delta(\Theta))$, that satisfies the martingale property, there exists a disclosure policy (M, G) that induces it. Thus, in what follows, we consider the principal's problem of choosing $\lambda \in \Delta(\Delta(\Theta))$, instead of choosing (M, G) .

Proposition 2. For each (M, G) (and its induced $(\mu, (\Psi^m)_{m \in M})$ as in Proposition 1), there exists $\lambda \in \Delta(\Delta(\Theta))$ such that, for each measurable¹² $C \subseteq \Delta(\Theta)$, we have $\lambda(C) = \mu(\{m \in M | \Psi^m \in C\})$ with the martingale property that $\int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(\cdot) d\lambda(\Psi_0) = F_0(\cdot)$.

Conversely, for each $\lambda \in \Delta(\Delta(\Theta))$ with the martingale property that $\int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(\cdot) d\lambda(\Psi_0) = F_0(\cdot)$, there exists (M, G) (and its induced $(\mu, (\Psi^m)_{m \in M})$ as in Proposition 1) such that, for each measurable $C \subseteq \Delta(\Theta)$, we have $\lambda(C) = \mu(\{m \in M | \Psi^m \in C\})$.

By this proposition, the principal's optimal information disclosure problem can be written as follows.

$$\begin{aligned} \Pi^* &= \sup_{\lambda \in \Delta(\Delta(\Theta))} \int_{\Psi_0 \in \Delta(\Theta)} \Pi(\Psi_0) d\lambda \\ &\text{sub. to } \int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(\cdot) d\lambda = F_0(\cdot). \end{aligned}$$

¹²In what follows, we endow $\Delta(\Theta)$ with a Prokhorov metric (which metrizes its weak-* topology) and Borel σ -algebra. Then, $\Delta(\Theta)$ is a complete, separable metric space.

In the standard Bayesian persuasion problem, applying Jensen’s inequality, optimality of full disclosure is obtained when Π is convex.¹³ In the next subsection, we obtain sufficient conditions on the mechanism design environment which imply the convexity of Π .

In our model, the principal’s information disclosure affects both the principal and the agent in the following ways. For the principal’s side, more disclosure means that his mechanism can be contingent on more precise information. This *flexibility effect* in the mechanism choice makes, given everything else equal, the principal favor more disclosure. For the agent’s side, more disclosure means that his incentive compatibility and participation constraint become more stringent (e.g., under full disclosure, these constraints must be satisfied “ex post”, that is, for each realization of the principal’s information; while under no disclosure, they must be satisfied only “on average”). More generally, the set of implementable allocation rules becomes smaller by more disclosure. This *implementability effect* makes, given everything else equal, the principal favor less disclosure. The optimal disclosure policy is determined by appropriately balancing these two effects.

2.3 Linear and independent environment

For the rest of this section, we mainly consider the following linear environment. An allocation χ comprises two elements, $(q, p) \in [0, 1] \times \mathbb{R}$, where q represents a non-monetary allocation (e.g., the probability that the princi-

¹³In fact, convexity (concavity) is not sufficient for Π to admit Jensen’s inequality unless $\Delta(\Theta)$ is finite-dimensional (see, for example, Perlmán (1974)). Therefore, in the subsequent proofs, we additionally show an appropriate version of continuity of Π to apply Jensen’s inequality.

pal sells the good to the agent), and p represents a monetary transfer from the agent to the principal. This class of environments includes, for example, monopoly pricing (Mussa and Rosen (1978)) and regulation (Laffont and Tirole (1993)).

The agent's payoff is $qv_1(\theta, t) - p$, where the marginal valuation function v_1 is bounded, nondecreasing in t , and differentiable in t . The principal's payoff is $p - c(q, \theta, t)$ where c is bounded. Linearity in p means the players' risk neutrality in money, which is a standard assumption in the literature. Linearity in q for the agent's payoff can be weakened to some extent, though not fully. We discuss it in Supplementary Materials (Section K).

Given any posterior $\Psi_0 \in \Delta(\Theta)$ after the principal's information disclosure, let

$$V(t|\Psi_0) = \int_{\theta \in \Theta} v(\theta, t)\Psi_1(\theta|t, \Psi_0)$$

denote the agent's expected marginal valuation (recall that $\Psi_1(\cdot|t, \Psi_0)$ is the agent's posterior for θ given t and Ψ_0 , taking into account potential correlation between θ and t).

Suppose, for the moment, that V is non-decreasing in t (clearly, independence of (θ, t) is sufficient for this). Then, applying the standard technique in mechanism design based on an envelope theorem, the principal's mechanism

design problem can be written as follows.

$$\begin{aligned}
\Pi(\Psi_0) &= \sup_{(q,p)} \int_{\theta \in \Theta} \int_{t \in T} p(t) - c(q(t), \theta, t) dF_1(t|\theta) d\Psi_0(\theta) \\
&\text{sub.to } q(t)V(t|\Psi_0) - p(t) \geq 0, \\
&\quad q(t)V(t|\Psi_0) - p(t) \geq q(t')V(t|\Psi_0) - p(t'), \quad \forall t, t' \\
&= \sup_q \pi(q, \Psi_0) \\
&\text{sub.to } q \text{ nondecreasing,}
\end{aligned}$$

where

$$\pi(q, \Psi_0) \equiv \int_{\theta \in \Theta} \int_{t \in T} q(t)v(\theta, t) - c(q(t), \theta, t) - \left(\int_0^t q(\tilde{t})V'(\tilde{t}|\Psi_0)d\tilde{t} \right) dF_1(t|\theta) d\Psi_0(\theta).$$

The expected value of the first two terms, $q(t)v(\theta, t) - c(q(t), \theta, t)$, represents the expected total surplus generated in a mechanism that allocates $q(\cdot)$. Note that this expected value is linear in Ψ_0 . The expected value of the last term, $\int_0^t q(\tilde{t})V'(\tilde{t}|\Psi_0)d\tilde{t}$, represents the expected information rent paid to the agent. This expected value is not necessarily linear in Ψ_0 . However, as we see now, the expected information rent is linear in Ψ_0 (and hence $\pi(q, \Psi_0)$ is linear in Ψ_0 too) if (θ, t) are independent. This linearity of $\pi(q, \Psi_0)$ plays an important role for the optimality of full disclosure.

Lemma 1. If (θ, t) are independent, then $\pi(q, \Psi_0)$ is linear in Ψ_0 .

Now we state the first main result of this paper.

Theorem 1. If (θ, t) are independent, then full disclosure is optimal.

The proof basically exploits a simple property that, if an objective $(\Pi(q, \Psi_0))$

is linear in a parameter (Ψ_0) and the feasible set (the set of all monotonic q) does not vary with this parameter, then the value function ($\Pi(\Psi_0)$) is convex in this parameter.

To provide some economic intuition for the result, recall that there are two channels where information disclosure affects implementable allocation rules. First, more disclosure implies more flexibility in the mechanism. In its extreme, q can vary with θ under full disclosure, while q must be constant in θ under no disclosure. This first effect makes the principal favor more disclosure. Second, more disclosure implies (weakly) tighter incentive constraints. Again in its extreme, truth-telling must be optimal given every possible realization θ under full disclosure, while truth-telling needs to be optimal only “in expectation” with respect to Θ under no disclosure. Because of this, either expected information rent may be higher or implementable allocation rules may be smaller under full disclosure (than under less disclosure). Hence, this second effect makes the principal favor less disclosure.

However, in the current environment, the second effect is null. Because every player has a linear payoff, the expected information rent is linear in Ψ_0 , and hence, more disclosure *does not* imply strictly tighter incentive constraints. Therefore, only the first effect exists, leading to optimality of full disclosure.

2.4 Correlated case

If θ and t are correlated, then the expected information rent is not in general linear, because the change in θ affects $F_1(t|\theta)$ in a nontrivial manner, and hence the entire information rent expression. Indeed, it is possible that even

the no-disclosure policy is optimal.

Example 1. Assume $v(\theta, t) = 2 + \theta t$ and $c \equiv 0$, and the joint distribution over (θ, t) is given by

Pr	$t = 3$	$t = 6$
$\theta = 1$	1/12	5/12
$\theta = 3$	5/12	1/12

Under no disclosure, the agent's expected marginal valuation is $E[v|t] = 2 + E[\theta|t]t = 10$ for any t . Therefore, the optimal mechanism sets $(q(t), p(t)) = (1, 10)$ for any t , yielding the revenue 10. This is clearly the first-best outcome.

Under more disclosure, the agent earns some information rent and/or the second-best allocation is inefficient. Therefore, disclosing no information is optimal.

Notice that, in this example where v is non-decreasing in θ and (θ, t) are negatively correlated,¹⁴ less information disclosure implies less information rent, and hence preferred to the principal.

On the other hand, if (v is non-decreasing in θ and) (θ, t) are positively correlated, or more specifically, if F_1 , the conditional distribution of $t|\theta$, exhibits monotonicity in the sense that $\frac{1-F_1(t|\theta)}{f_1(t|\theta)}$ is increasing in θ ,¹⁵ then the expected information rent is concave in Ψ_0 for any feasible (i.e., nondecreasing) $q(\cdot)$, which implies convexity of $\pi(q, \Psi_0)$ and $\Pi(\Psi_0)$. Therefore, full disclosure is again optimal.

¹⁴Equivalently (up to redefinition of θ), v is non-increasing in θ and (θ, t) are positively correlated.

¹⁵A sufficient condition for this property is that F_1 has monotone likelihood ratio property, i.e., (θ, t) are affiliated.

Intuitively, concavity of the information rent means that information disclosure implies smaller ex ante information rent. This is related to the *linkage principle* of Milgrom and Weber (1982) in a correlated-value auction environment, and more closely, Ottaviani and Prat (2001) in a monopoly-pricing environment.

For each monotone $q : [0, 1] \rightarrow [0, 1]$ and $\Psi_0 \in \Delta(\Theta)$, let

$$R(q, \Psi_0) = \int_{\theta \in \Theta} \int_{t \in T} \int_0^t q(\tilde{t}) V'(\tilde{t} | \Psi_0) d\tilde{t} dF_1(t | \theta) d\Psi_0(\theta)$$

denote the expected information rent.

Lemma 2. For each t , assume that $v(\theta, t)$ and $\frac{1-F_1(t|\theta)}{f_1(t|\theta)}$ are non-decreasing in θ . Then, for each monotone $q : [0, 1] \rightarrow [0, 1]$ and $\Psi_0 \in \Delta(\Theta)$,

$$R(q, \Psi_0) \geq \int_{\theta \in \Theta} R(q, \delta_\theta) d\Psi_0(\theta),$$

where δ_θ denotes a Dirac measure on θ .

Theorem 2. For each t , assume that $v(\theta, t)$ and $\frac{1-F_1(t|\theta)}{f_1(t|\theta)}$ are non-decreasing in θ . Then, full disclosure is optimal.

3 Multiple agents

Similar arguments as in the single-agent case are applicable even with multiple agents.¹⁶ In Section 3.1, we show that, in the linear and independent

¹⁶Obviously, a crucial assumption is that we only consider *public* information disclosure. Although we believe that it is a reasonable assumption in the context of third-party certification, in other contexts, *private* disclosure may be more relevant. Zhu (2017) shows

environment as in Section 2 but with multiple agents, full disclosure is still optimal, as long as no restriction is imposed on feasible monetary transfers. This class of environments include applications such as auctions (Myerson (1981), Eső and Szentes (2007)) and two-sided markets (Gomes and Pavan (2016), Jeon et al. (2016)).

The assumption of no restriction on feasible monetary transfers is crucial. As we see in Section 3.2, In the same environment except that monetary transfers are restricted by a budget-balance (or no-deficit) requirement, full disclosure is not necessarily optimal. Indeed, in a general class of bilateral-trading environments as in Myerson and Satterthwaite (1983), we show that full disclosure is strictly suboptimal.

3.1 Linear and independent environment

We consider the same linear model as before, except that there are now N agents. An allocation comprises $(q_i, p_i)_{i=1}^N \in \mathbb{R}^{2N}$ where q_i represents a non-monetary allocation to agent i , and p_i represents monetary transfer from agent i to the principal. Let $Q \subseteq \mathbb{R}^N$ denote the feasible set of the non-monetary allocations to the agents (e.g., $q_i \geq 0$ for all i and $\sum_i q_i \leq 1$ in a single-good auction), while no restriction is imposed on each p_i .

As in Section 2, the principal's information is denoted by $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$, and each agent i 's type is denoted by $t_i \in T_i = [0, 1]$. We assume that $(\theta, t_1, \dots, t_N)$ are mutually independently distributed, where F_0 denotes the

that, under a version of monotone virtual value conditions in a linear environment, it is without loss of generality to focus on the public disclosure. That is, no private disclosure policy yields a strictly higher expected payoff of the principal than the optimal public disclosure policy.

distribution for θ , and each F_i denotes the distribution for t_i . Also, let F_T denote the joint distribution for t .¹⁷

Each agent's payoff is given by $q_i v_i(\theta, t_i) - p_i$, where the marginal valuation v_i is bounded, nondecreasing in t_i , and differentiable in t_i . As we discuss later, the result can be extended to the case of interdependent values (i.e., v_i is a function of (θ, t_i, t_{-i}) instead of only (θ, t_i)). The principal's payoff is $\sum_{i=1}^N p_i - c(q, \theta, t)$.

Again, applying the standard technique in mechanism design based on an envelope theorem and integration by parts, the principal's mechanism design problem can be written as follows:

$$\begin{aligned} \Pi(\Psi_0) = & \sup_{q: [0,1]^N \rightarrow Q} \pi(q, \Psi_0) \\ & \text{sub.to } q_i(\cdot, t_{-i}) \text{ nondecreasing, } \forall i, t_{-i}, \end{aligned}$$

where

$$\begin{aligned} \pi(q, \Psi_0) \equiv & \int_{\theta \in \Theta} \int_{t \in T} \sum_i \left(q_i(t) v_i(\theta, t_i) \right) - c(q(t), \theta, t) - \left(\int_0^{t_i} \int_{\theta \in \Theta} \frac{\partial v_i}{\partial t_i}(\theta, t_i) d\Psi_0(\theta) d\tilde{t}_i \right) \\ & dF_T(t) d\Psi_0(\theta). \end{aligned}$$

With independence, we show that the expected value of the last term, the sum of the agents' information rent, is linear in Ψ_0 . Therefore, full disclosure is optimal. Because it is a straightforward extension of Theorem 1, we omit the proof.

¹⁷Again, Θ can be generalized to any separable, complete metric space. For example, Θ could be multidimensional, such as $\Theta \subseteq \mathbb{R}^N$ (where, in some applications, each agent i might only care about the i -th argument of θ).

Theorem 3. Full disclosure is optimal.

The assumptions of linearity and private values can be weakened to some extent (though not completely dispensable). See the Supplementary Materials (Section K) for alternative assumptions and solution concepts that imply similar full-disclosure results.

3.2 Budget balance and suboptimality of full disclosure

As discussed in the last section, the property that the set of implementable q does not vary with Ψ_0 is crucial for the convexity of Π (and hence for optimality of full disclosure).

In this section, we observe that the feasible set varies with Ψ_0 in balanced-budget bilateral trading environments. This is because the budget balance condition constrains the agents' total expected information rents, an expression that varies with Ψ_0 . In this environment, we show that full disclosure is suboptimal. More specifically, there exists a subset of Θ where the principal strictly prefers not to reveal its realization.

Following Myerson and Satterthwaite (1983), we assume that there are two agents, a seller ($i = 1$) and a buyer ($i = 2$). (θ, t_1, t_2) are mutually independent. For each $i = 1, 2$, let F_i denote the cdf for t_i with a full-support density f_i . For θ , we assume that F_0 has a full-support density f_0 on $\Theta = [\underline{\theta}, \bar{\theta}]$ with $\underline{\theta} < 1 < \bar{\theta}$. The seller's payoff is $-v_1(\theta, t_1)q_1 - p_1$ and the buyer's payoff is $v_2(\theta, t_2)q_2 - p_2$, where $q_1 = q_2 \in [0, 1]$ is the trade probability and $p_i \in \mathbb{R}$ is the monetary transfer from i . The budget balance condition

requires that $p_1 + p_2 \geq 0$.¹⁸ Therefore, the feasible allocation set is given by

$$X = \{(q_i, p_i)_{i=1}^2 \in [0, 1]^2 \times \mathbb{R}^2 \mid q_1 = q_2, p_1 + p_2 \geq 0\}.$$

The principal's objective is the trade surplus, $v_2(\theta, t_2)q_2 - v_1(\theta, t_1)q_1$, as in Myerson and Satterthwaite (1983). Then, the value function for the principal given any posterior Ψ_0 is $\Pi(\Psi_0)$, given by

$$\begin{aligned} & \sup_{(q, p_1, p_2): [0, 1]^2 \rightarrow X} \int_{\theta \in \Theta} \int_{t \in T} (v_2(\theta, t_2) - v_1(\theta, t_1))q(t) dF_T(t) d\Psi_0(\theta) \\ & \text{sub. to} \int_{\theta \in \Theta} \int_{t_2 \in T_2} (-v_1(\theta, t_1)q(t) - p_1(t)) dF_2(t_2) d\Psi_0(\theta) \\ & \quad \geq \max \left\{ 0, \int_{\theta \in \Theta} \int_{t_2 \in T_2} (-v_1(\theta, t_1)q(t'_1, t_2) - p_1(t'_1, t_2)) dF_2(t_2) d\Psi_0(\theta) \right\}, \forall t_1, t'_1, \\ & \quad \int_{\theta \in \Theta} \int_{t_1 \in T_1} (v_2(\theta, t_2)q(t) - p_2(t)) dF_1(t_1) d\Psi_0(\theta) \\ & \quad \geq \max \left\{ 0, \int_{\theta \in \Theta} \int_{t_1 \in T_1} (v_2(\theta, t_2)q(t_1, t'_2) - p_2(t_1, t'_2)) dF_1(t_1) d\Psi_0(\theta) \right\}, \forall t_2, t'_2, \\ & \quad p_1(t) + p_2(t) \geq 0, \forall t. \end{aligned}$$

We assume the following “regularity” conditions.

Assumption 1.

1. $\frac{\partial v_2}{\partial t_2}(\theta, t_2) \frac{1-F_2(t_2)}{f_2(t_2)}$ is strictly decreasing in t_2 , and $\frac{\partial v_1}{\partial t_1}(\theta, t_1) \frac{F_1(t_1)}{f_1(t_1)}$ is strictly increasing in t_1 .¹⁹

¹⁸This is also called a “no-deficit” condition. In the optimal mechanism, the constraint binds with equality anyway.

¹⁹A stronger set of conditions is the following: v_2 is concave in t_2 , v_1 is convex in t_1 , $\frac{1-F_2(t_2)}{f_2(t_2)}$ is decreasing in t_2 , and $\frac{F_1(t_1)}{f_1(t_1)}$ is increasing in t_1 . The last two conditions are known as monotone hazard rate conditions.

2. There exist $0 < b_1 \leq b_2 < \infty$ such that (i) for each $i, t_i, f_i(t_i) \in [b_1, b_2]$,
(ii) for each $i, t_i, \theta, \frac{\partial v_i}{\partial t_i}(\theta, t_i) \in [b_1, b_2]$, and (iii) for each $\theta, \rho(\theta) = v_2(\theta, 0) - v_1(\theta, 1) \in [(\theta - 1)b_1, (\theta - 1)b_2]$.²⁰

The first assumption is standard in order to make sure that no bunching occurs in the optimal mechanism. The second set of assumptions is to regulate the shape of some key functions in the following analysis. The first two of them simply say that F_i and v_i are Lipschitz with (uniformly) positive slopes. The third one implies the following: if $\theta > 1$ is common knowledge, then the *full-trade* outcome (i.e., $q(t) = 1$ for all t) is efficient, which is implementable by a posted-price mechanism (“gap case”), yielding the full-trade surplus

$$\Pi^1(\theta) = \int_{t \in T} (v_2(\theta, t_2) - v_1(\theta, t_1)) dF_T(t);$$

while if $\theta < 1$ is common knowledge (“no-gap case”), then the full-trade outcome is not efficient, and any mechanism can implement neither the full-trade outcome nor the first-best efficient outcome (Myerson and Satterthwaite (1983)).

Under Assumption 1, $\Pi(\Psi_0)$ is equivalently given by

$$\begin{aligned} & \sup_{q: [0,1]^2 \rightarrow [0,1]} \int_{\theta \in \Theta} \int_{t \in T} (v_2(\theta, t_2) - v_1(\theta, t_1)) q(t) dF_T(t) d\Psi_0(\theta) \\ & \text{sub. to} \quad \int_{\theta \in \Theta} \int_{t \in T} \left(v_2(\theta, t_2) - v_1(\theta, t_1) - \frac{\partial v_2}{\partial t_2}(\theta, t_2) \frac{1 - F_2(t_2)}{f_2(t_2)} - \frac{\partial v_1}{\partial t_1}(\theta, t_1) \frac{F_1(t_1)}{f_1(t_1)} \right) \\ & \quad \quad \quad q(t) dF_T(t) d\Psi_0(\theta) \geq 0. \end{aligned}$$

²⁰A seemingly weaker assumption assigns different bounds for each of these three functions. However, it is equivalent to our assumption here, by setting b_1 as the minimum of those three lower bounds, and b_2 as the maximum of those three upper bounds.

Theorem 4. Full disclosure is suboptimal.

The key intuition of this suboptimality is the following. As $\theta \uparrow 1$, the Lagrange multiplier for the constraint, say $\lambda(\theta)$, is vanishing, as is known in the literature. However, its rate of convergence is slower than a linear rate. Indeed, $\lambda'(\theta) \uparrow \infty$ as $\theta \uparrow 1$.²¹ This means that, under full disclosure, as θ decreases from 1, $\Pi(\delta_\theta)$ decreases drastically. On the other hand, it is easy to see that the full-trade surplus $\Pi^1(\theta)$ decreases only in a linear rate. Thus, for θ close to 1, *if the full-trade outcome is ever implementable*, then it is strictly better than the best outcome under full disclosure. In the proof, we identify θ^*, θ^{**} with $\theta^* < 1 < \theta^{**}$ such that, if the agents only know that $\theta \in (\theta^*, \theta_2^{**})$, then the full-trade outcome is indeed implementable. Then, based on this logic, the principal strictly prefers not to disclose θ for $\theta \in (\theta^*, \theta_2^{**})$. Notice that this argument does not depend much on the specific functional forms (except for the regularity conditions), which allows us to obtain this suboptimality result in a fairly general environment.

4 Optimal information disclosure

When full disclosure is suboptimal, a natural next step is characterization of the optimal disclosure policy. However, this problem in the general environment is difficult, because the principal's choice variable $\lambda \in \Delta(\Delta(\Theta))$,

²¹More precisely, $\lambda(\theta)$ is vanishing at rate $\sqrt{1-\theta}$, as implied in the inequality in the proof:

$$\frac{\lambda(\theta)}{1+\lambda(\theta)} \geq \frac{b_1}{2b_2} \left(\rho(\theta) + \frac{b_1}{b_2} \sqrt{\frac{-\rho(\theta)b_1}{2b_2 - \rho(\theta)b_1}} \right).$$

a distribution over posteriors, is a probability distribution over an infinite-dimensional space. We make the following assumption to simplify the analysis.

Assumption 2. For each $x \in [\underline{\theta}, \bar{\theta}]$ and $\Psi_0, \Psi'_0 \in \Delta(\Theta)$ with $x = \int_{\theta \in \Theta} \theta d\Psi_0(\theta) = \int_{\theta \in \Theta} \theta d\Psi'_0(\theta)$, we have $\Pi(\Psi_0) = \Pi(\Psi'_0)$, which we denote by $\Pi(x)$.

Under this assumption, the principal's choice variable is now a distribution over posterior *means*, denoted by $F \in \Delta(\Theta)$. Since F is identified by a cdf over $\Theta \subseteq \mathbb{R}$, it is a one-dimensional function. In this sense, the space of feasible F is still infinite-dimensional, but simpler than the fully general case.²²

Although admittedly restrictive, the following lemma shows that this mean-only assumption is satisfied in some environments.

Lemma 3. Assumption 2 is satisfied either (i) if $(\theta, t_1, \dots, t_N)$ are mutually independent, and $u_0(\theta, t), u_1(\theta, t), \dots, u_N(\theta, t)$ are affine in θ ; or (ii) if there exist a random variable $s \in \{0, 1\}$ (“unobservable fundamental”) and functions of (s, t) , $\hat{u}_0, \hat{u}_1, \dots, \hat{u}_N$, such that $u_j(\theta, t) = E[\hat{u}_j(s, t)|\theta, t]$ for $j = 0, 1, \dots, N$, and $(\theta, t_1, \dots, t_N)|s$ are conditionally (on each s) mutually independent with density $f_0^s(\theta)$ for $\theta|s$ and $f_i^s(t)$ for $t_i|s$.

²²Assumption 2 can be replaced by the following (seemingly weaker) assumption: there exists a measurable function $\xi : \Theta \rightarrow \mathbb{R}$ such that, for any x and Ψ_0, Ψ'_0 with $x = \int_{\theta \in \Theta} \xi(\theta) d\Psi_0(\theta) = \int_{\theta \in \Theta} \xi(\theta) d\Psi'_0(\theta)$, we have $\Pi(\Psi_0) = \Pi(\Psi'_0)$ (which we denote by $\Pi(x)$). This is just a matter of notation in the sense that, redefining $\xi(\theta)$ as θ , this alternative assumption becomes equivalent to the original one. However, this alternative version may be useful in applications where θ has some intrinsic economic meaning. For example, in Lemma 3 (ii), $\xi(\theta)$ represents the posterior for some random variable. Also, in case Θ (endowed with some intrinsic economic meaning) is multidimensional, this alternative assumption would be important to admit a one-dimensional representation of the problem, so that the results of this section become applicable.

We further assume the following property on Π .

Assumption 3. (i) Π is upper-semi-continuous. (ii) There exist $K \in \mathbb{N}$ and $\{x_k\}_{k=1}^K$ with $0 = x_0 < \dots < x_K = 1$ such that, on each open interval (x_{k-1}, x_k) , $k = 1, \dots, K$,

- (i) Π is differentiable with derivative Π' , and
- (ii) Π' is monotone (i.e., either nondecreasing or nonincreasing), and absolutely continuous with derivative Π'' (which exists almost everywhere).

For example, any polynomial function satisfies this assumption. Although Π must be well-behaved on each open interval (x_{k-1}, x_k) , Π can be more “irregular” at each x_k : Π may be discontinuous at x_k (although we need upper-semi-continuity), or even if not, Π' may not exist or be discontinuous at x_k .

The main result of this section is characterization of the optimal information disclosure policy.

Definition 1. (i) For each $A \subseteq [\underline{\theta}, \bar{\theta}]$, we say that (M, G) exhibits *full disclosure* on A if $M \supseteq A$ and $G(\{\theta\}|\theta) = 1$ for each $\theta \in A$.

(ii) For each $a, b \in [\underline{\theta}, \bar{\theta}]$ with $a < b$, we say that (M, G) exhibits *binary lower truncation* on sub-interval (a, b) if there exist $y \in (a, b)$ and $w \in (0, 1]$ such that

- (ii-i) $G(\{z_1\}|\theta) = 1$ for $\theta \in (a, y)$; and
- (ii-ii) $G(\{z_1\}|\theta) = w$ and $G(\{z_2\}|\theta) = 1 - w$ for $\theta \in (y, b)$;

where

$$z_j = \left\{ \begin{array}{ll} \frac{\int_a^y \theta dF_0(\theta) + \int_y^b w \theta dF_0(\theta)}{\int_a^y 1 dF_0(\theta) + \int_y^b w dF_0(\theta)} & \text{if } j = 1 \\ \frac{\int_y^b \theta dF_0(\theta)}{\int_y^b 1 dF_0(\theta)} & \text{if } j = 2 \end{array} \right\} \left(= E[\theta | z_j \text{ is announced}] \right).$$

To explain these properties, we first provide a few examples.

Example 2. For simplicity, assume that F_0 corresponds to $U(0, 1)$ in the following examples.

- Consider (M, G) which exhibits binary lower truncation on the entire $\Theta = [0, 1]$ with $w = 1$. It is the no-disclosure policy, because for any θ , $G(\{z_1\}|\theta) = 1$ where $z_1 = \frac{1}{2}$.
- If there exists $\theta^* \in (0, 1)$ such that (M, G) exhibits binary lower truncation on $\theta > (<) \theta^*$ with $w = 1$ and full-disclosure otherwise, it corresponds to upper (lower) censorship of Kolotilin et al. (2015).
- If there exists $\theta^* \in (0, 1)$ such that (M, G) exhibits binary lower truncation on $(0, \theta^*)$ with $w = 1$ and binary lower truncation on $(\theta^*, 1)$ again with $w = 1$, then it corresponds to a (two-message) “monotone partition” structure, as in the cheap-talk literature (Crawford and Sobel (1982)).
- Consider (M, G) which exhibits binary lower truncation on the entire $\Theta = [0, 1]$ with $y = \frac{1}{3}$ and $w = \frac{1}{4}$. Then we have $z_1 = \frac{1}{3}$ and $z_2 = \frac{2}{3}$, and for each $\theta < \frac{1}{3}$,

$$G\left(\left\{\frac{1}{3}\right\}|\theta\right) = 1,$$

while for each $\theta > \frac{1}{3}$,

$$\begin{aligned} G\left(\left\{\frac{1}{3}\right\}|\theta\right) &= \frac{1}{4}, \\ G\left(\left\{\frac{2}{3}\right\}|\theta\right) &= \frac{3}{4}. \end{aligned}$$

□

The disclosure policy in the last example can be interpreted in the following alternative way: instead of z_1 or z_2 , imagine that the principal truthfully announces a *lower bound* of the realized θ , say $\underline{\theta}$, either $\underline{\theta} = 0$ or $\underline{\theta} = \frac{1}{3}$. More specifically, if $\theta < \frac{1}{3}$, then $\underline{\theta} = 0$ is announced for sure; if $\theta \geq \frac{1}{3}$, then $\underline{\theta} = 0$ is announced with probability $\frac{1}{4}$, while $\underline{\theta} = \frac{1}{3}$ is announced with probability $\frac{3}{4}$.²³ By Bayesian updating, the posterior mean of θ given announcement of $\underline{\theta} = 0$ is $\frac{1}{3}$, while given announcement of $\underline{\theta} = \frac{1}{3}$, it is $\frac{2}{3}$. Announcing a lower bound in this way essentially makes the posterior distribution a lower (or “left”) truncation of the prior; hence we call it a (binary) lower truncation disclosure policy.

Theorem 5. There exist $L \leq K$ and disjoint open sub-intervals $\{I_l\}_{l=1}^L$ of $[0, 1]$ such that an optimal disclosure policy exhibits binary lower truncation on each I_l and full-disclosure otherwise.

There are several important properties of the optimal policy in the statement. First, the message space can always be taken as a subset of $[\underline{\theta}, \bar{\theta}]$, because in both binary-lower-truncation regions and full-disclosure region,

²³We mean by truthfulness that the principal never announces $\underline{\theta} > \theta$ in case θ is realized.

the principal's message is a number in $[\underline{\theta}, \bar{\theta}]$. In fact, conditional on $x \in [\underline{\theta}, \bar{\theta}]$ being announced, the (principal's) posterior mean of θ is always x .

Second, although infinitely many messages may be necessary, it is only because of the full-disclosure region, i.e., outside the full-disclosure region, only finitely many messages are necessary.²⁴ Furthermore, *conditional on* θ , the support of $G(\cdot|\theta)$ is always an (at most) binary set. In these senses, the theorem implies a significant bound on the size of necessary messages, even though Θ is an infinite space.

Third, the optimal policy in the statement exhibits the following monotone-likelihood-ratio property: for any two states $\theta < \theta'$ and any two messages $x < x'$, we have

$$G(\{x\}|\theta)G(\{x'\}|\theta') \geq G(\{x'\}|\theta)G(\{x\}|\theta').$$

In various applications, such monotonic disclosure policies may look natural. For example, imagine that a seller of goods with uncertain quality (e.g., used cars, drugs, or agricultural products) can generate some hard evidence about the quality through a third-party certifier. The certifier examines the quality of the goods through some physical experiments, generating noisy signals. In such a case, a feasible signal distribution would be constrained by the nature of the physical experiments. The monotone likelihood ratio property may be one such natural constraint. Moreover, recall that a lower truncation policy can be interpreted as a random announcement of a lower bound of realized θ . This would fit well into a situation where a certifier runs

²⁴In fact, as is made clear in the proof, the total number of messages outside the full-disclosure region is bounded by $2K$.

multiple pass-fail tests (so that each test, if passed, provides a lower bound of realized θ), and randomly announces which tests have been passed.²⁵

Another situation where monotonicity is important is when the sender has a weaker commitment power. Recall that, in the Bayesian persuasion literature, the sender is assumed to have a strong commitment power in that he (i) sets up an experiment (M, G) at the ex ante stage, and then (ii) inputs the realized θ “truthfully” to this G . The first assumption seems reasonable, for example, in environments where the sender can delegate signal generation to a third-party certifier (e.g., FDA in the context of a drug industry) at the ex ante stage. However, the second assumption may be more controversial. Even if the sender delegates the experiments to the third-party certifier, often times, it is the sender himself who provides a sample of the goods to be examined to the certifier. If the sender can observe the quality of the goods and manipulate it *before* providing it to the certifier, the sender may find it profitable to do so. This means that, in such “limited commitment” situations, feasible (M, G) would be restricted by the sender’s incentive compatibility in “inputting” θ to G . Relevant incentive constraints would vary depending on the applications, and one possible situation may be that the sender can *reduce* the quality of the goods for free, while cannot *increase* it at all. In such a case (and with an increasing Π , such as in pure revenue maximization), non-monotonic G is not incentive compatible, because the

²⁵By a similar proof, the optimal information disclosure policy can also be described as a combination of full-disclosure regions and (binary) upper-truncation regions. Such an upper-truncation policy can be interpreted as a random announcement of failures instead of passes. Also, any convex combination (in an appropriate sense) of lower- and upper-truncation is also optimal, which can be interpreted as a random announcement of passes and failures. Note that the linearity assumption plays an important role for the multiplicity of optimal policies.

sender may find it profitable to decrease θ to enjoy higher posterior means. The theorem, which says that an optimal policy can always be found in the class of monotone disclosure policies, has an important implication that this principal’s incentive compatibility constraint is satisfied in the current environment.²⁶

4.1 Example 1: Bilateral trade

We consider two applications where full disclosure is suboptimal. Here, we consider a balanced-budget bilateral trade environment, and characterize the optimal disclosure policy (and the optimal mechanisms) in expected surplus. In Supplementary Materials (Section M), we consider another example which is in a single-agent correlated environment, and characterize the optimal disclosure policy (and the optimal mechanisms) in expected profit.

Consider a balanced-budget bilateral trade example where each agent’s type t_i follows an independent uniform distribution over $[0, 1]$. As in Section 3.2, an allocation is represented by (q, p) , where $q \in [0, 1]$ denotes the probability of trading and $p \in \mathbb{R}$ denotes the payment from the buyer (agent 2) to the seller (agent 1). Agent 1’s payoff is $p - qt_1$, and agent 2’s payoff is $q(\theta + t_2) - p$, where $\theta \in U(a, b)$ with $a < 1 < b$.

²⁶ As claimed in the next lemma, our problem can be interpreted as a choice problem of F (a cdf) that is less riskier than F_0 in the sense of Rothschild and Stiglitz (1970). In this sense, once we could characterize the optimal F in that problem, we can directly apply the method by Rothschild and Stiglitz (1970) (or more precisely, the method by Machina and Pratt (1997) for continuous random variables) to construct a policy (M, G) that implements F . However, their construction is different from ours in that it does not satisfy monotone likelihood ratio property in the above sense, and does not admit its interpretation as combination of full disclosure and “announcement of lower bounds of realization”. For the reasons discussed above, these properties could be advantageous in some economic applications.

Because the support of θ is around 1, as shown in Section 3.2, full disclosure is suboptimal. We first examine the shape of $\Pi(\cdot)$ more fully.

Lemma 4. Π' exists and continuous for all x . Π'' exists and continuous for all $x \neq 1$. More specifically, there exists $\hat{x} (\simeq 0.87) \in (\frac{1}{3}, 1)$ such that

$$\Pi''(x) \begin{cases} \geq 0 & \text{if } x < \hat{x}, \\ < 0 & \text{if } x \in (\hat{x}, 1), \\ = 0 & \text{if } x > 1. \end{cases}$$

With this simple threshold structure, the optimal disclosure policy is proved to be an (upper) *censorship* policy of Kolotilin et al. (2015).

Proposition 3. There exists $x^* \in [a, \hat{x}]$ such that the following disclosure strategy is optimal: fully disclose the realized θ if $\theta \leq x^*$, and not (at all) disclose it if $\theta > x^*$. For example, if $a \leq -1$ and $b \geq 3$, then $x^* = \frac{1}{3}$, yielding ex ante expected surplus $\frac{9b^2+1}{18(b-a)}$.

The main tradeoff for information disclosure is between the budget-balance requirement and flexibility in the mechanism choice. Disclosing less information can be beneficial for the principal because it mitigates the budget-balance requirement, especially around $\theta = 1$. On the other hand, disclosing more information can be beneficial, because it brings more flexibility in the mechanism, especially for smaller θ .

4.2 Example 2: Type-dependent outside options

5 Concluding remarks

In certain mechanism design environments, the principal finds it optimal to fully disclose the information. Although those environments include many “standard” mechanism design environments, in other cases, full disclosure may be suboptimal. With additional assumptions, we characterize the optimal disclosure policy as a combination of full-disclosure regions and binary lower truncation regions.

We conclude the paper with a discussion of some alternative models. Two important assumptions to obtain the main results of the paper are (i) public disclosure and (ii) commitment assumptions.

Regarding the first assumption, possible alternative models may allow the principal to (i-a) communicate with each of multiple agents *privately*, i.e., send different messages to different agents, or (i-b) even for a single agent, send a *type-contingent* message, i.e., different messages to different types.²⁷ As discussed in the introduction (see Dranove and Jin (2010) cited in Footnote 2), we believe that the assumption of public disclosure is reasonable in many applied contexts, but at least theoretically, such alternative models should be considered. Note that public disclosure is always a special case of those two alternatives, and hence, the principal can only be better off. In this sense, the result about suboptimality of full disclosure (Theorem 4) continues

²⁷This (i-b) can only be possible by reversing the timing of the game so that the principal first asks each agent’s type, and then chooses a (possibly private) disclosure policy depending on their type reports. In principle, the problem can be much more complicated because which disclosure policy to adopt may be used as another source of incentive provision (see Bergemann et al. (2017)).

to hold. On the other hand, the results about optimality of full disclosure may not be robust. Nevertheless, under additional assumptions, the results can be shown to be robust to those alternative models. For example, in the model of Section 2.3, assume that (θ, t) is independent, $\frac{1-F_1(t)}{f_1(t)}$ is non-increasing, i.e., a standard monotone hazard rate condition, and that v_1 is concave in t . Then, full disclosure continues to be optimal as follows. First, recall that, under those assumptions, Skreta (2011) considers a much better situation (in view of the principal) where the principal can make the allocation contingent on θ in a fully committed manner without disclosing to the agents. The principal's value in her setting is obviously an upper bound of what the principal achieve in our problem (and under alternative models such as (i-a) or (i-b)), but this value can be indeed achieved under full disclosure.

The second crucial assumption of the current paper that the principal can commit both to his mechanism and disclosure policy is arguably too strong to be true in reality, while these are often assumed in the literature. Perhaps the current results with a strong commitment power may be seen as a benchmark analysis which could be useful to discuss more realistic “weaker commitment” cases. As such, we do not have a strong argument for this strong commitment assumption, but in what follows, we discuss a few points which suggest that our results might possess some robust feature with respect to the commitment assumption regarding the disclosure policy. First, even if the principal in practice cannot fully commit to his desired disclosure policy, he may be able to *delegate* information disclosure to some third-party certifier, and this delegation (especially if it can be done before the principal observes θ) can provide some commitment power to the principal, as often discussed in the

industrial organization literature.

Second, even in environments where the principal has less commitment power, our optimal disclosure policy may still show up. For example, in the linear environment studied in Section 2 and 3, if the principal cannot commit to his disclosure policy, then a natural guess would be that some sort of *unraveling* argument (?) implies full disclosure in any equilibrium. Then, our result shows that this full-disclosure outcome without commitment is actually the best the principal attain even if he can commit. As a related point, as discussed in Theorem 5, some (non-full) optimal policy also exhibits certain monotonicity property, which suggests that, even if the principal has weaker commitment policy, the same policy may still be implementable.

Other than those cases, in general, the optimal disclosure policy with different timing and / or under limited commitment can be different from our fully-committed case. Nevertheless, we believe that the methodology developed in this paper may still be useful in deriving the optimal disclosure policy in some of those alternative environments. As one such instance, imagine a situation where the principal not only can engage in the public, committed disclosure as in this paper, but also can reveal other information as in the standard informed-principal manner, that is, the principal can make the allocation rule contingent on even the part of θ that is not publicly disclosed, though subject to the *principal's own incentive compatibility*. This alternative assumption may correspond to the situation where the principal can both use their-party public certification and engage in his own information acquisition. In the current paper, we assume that the allocation rule cannot depend on the part of θ that is not publicly disclosed, and in this sense, this

alternative assumption makes the principal weakly better off. Whether it makes the principal strictly better off would depend on the situation, but in either case, the methodology developed in this paper could be useful to consider this problem. To explain the main idea, let Ψ_0 denote the posterior induced by public (committed) disclosure. Then, the principal essentially faces a standard informed-principal problem where he can make the allocation contingent on the non-disclosed part of θ but subject to his incentive compatibility. Let $\tilde{\Pi}(\Psi_0)$ denote the principal's value of this second-stage problem (which is in general higher than $\Pi(\Psi_0)$). Once $\tilde{\Pi}(\cdot)$ is given, the information-disclosure stage stays the same. Therefore, the qualitative aspect of our results would not change: (i) if this “modified $\Pi(\cdot)$ ” is convex, then full disclosure is optimal; and (ii) if it satisfies Assumption 2, then the analysis in Section 4 would be applicable.²⁸

²⁸Of course, even if we consider the same economic environment, the shapes of our $\Pi(\cdot)$ and $\tilde{\Pi}(\cdot)$ can be different. Indeed, Example 1 in Koessler and Skreta (2016) illustrates a single-agent environment with independent information (where our $\Pi(\cdot)$ is convex) where the $\tilde{\Pi}(\cdot)$ is not convex.

A Proof of Proposition 2

For the first statement, by Proposition 1, (M, G) induces $(\mu, (\Psi^m)_{m \in M})$ such that (i) $\mu \in \Delta(M)$, (ii) $\Psi^m \in \Delta(\Theta)$ for each $m \in M$, and (iii) $\int_{m \in B} \Psi^m(A) d\mu = \int_{\theta \in A} G(B|\theta) dF_0$ for each measurable $A \subseteq \Theta$ and $B \subseteq M$.

Because a mapping $m \mapsto \Psi^m$ is measurable, for each measurable $C \subseteq \Delta(\Theta)$, $\{m \in M | \Psi^m \in C\}$ is measurable. Thus, we define $\lambda \in \Delta(\Delta(\Theta))$ by $\lambda(C) = \mu(\{m \in M | \Psi^m \in C\})$ for each measurable $C \subseteq \Delta(\Theta)$. Taking $B = M$, Property (iii) above implies

$$\begin{aligned} F_0(A) &= \int_{m \in M} \Psi^m(A) d\mu \\ &= \int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(A) d\lambda. \end{aligned}$$

For the second statement, because $\Delta(\Theta)$ is a complete, separable metric space, the product-regular-conditional-probability property implies that there exists $(\tilde{\mu}, (\tilde{\Psi}^\theta)_{\theta \in \Theta})$ such that (i) $\tilde{\mu} \in \Delta(\Theta)$, (ii) $\tilde{\Psi}^\theta \in \Delta(\Delta(\Theta))$ for each $\theta \in \Theta$, and (iii) for each measurable $A \subseteq \Theta$ and $C \subseteq \Delta(\Theta)$,

$$\int_{\theta \in A} \tilde{\Psi}^\theta(C) d\tilde{\mu} = \int_{\Psi_0 \in C} \Psi_0(A) d\lambda.$$

First, taking $C = \Delta(\Theta)$, we obtain

$$\begin{aligned} \tilde{\mu}(A) &= \int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(A) d\lambda \\ &= F_0(A), \end{aligned}$$

where the second equality is because of the martingale property of λ . Hence, $\tilde{\mu} = F_0$.

Consider (M, G) such that $M = \Delta(\Theta)$ and $G(\cdot|\theta) = \tilde{\Psi}^\theta(\cdot)$. As in Proposition 1, it induces $(\mu, (\Psi^m)_{m \in M})$ such that, for each measurable $A \subseteq \Theta$ and $C \subseteq \Delta(\Theta)$,

$$\begin{aligned} \int_{m \in C} \Psi^m(A) d\mu &= \int_{\theta \in A} G(C|\theta) dF_0 \\ &= \int_{\theta \in A} \tilde{\Psi}^\theta(C) d\tilde{\mu} \\ &= \int_{\Psi_0 \in C} \Psi_0(A) d\lambda. \end{aligned}$$

Therefore, taking $A = \Theta$, we obtain $\mu(C) = \lambda(C)$. Hence, $\mu = \lambda$. This implies that, for each measurable $A \subseteq \Theta$ and $C \subseteq \Delta(\Theta)$,

$$\int_{m \in C} \Psi^m(A) d\mu = \int_{m \in C} m(A) d\mu,$$

and thus, a mapping $m \mapsto \Psi^m$ must be an identity map for μ -a.e. m . That is, for each measurable $C \subseteq \Delta(\Theta)$, $\mu(C) = \mu(\{m \in M | \Psi^m \in C\})$, establishing $\lambda(C) = \mu(\{m \in M | \Psi^m \in C\})$.

B Proof of Lemma 1

Because of independence, let F_1 denote the distribution of t given any θ . Also, $\Psi_1(\cdot|t, \Psi_0)$ reduces to $\Psi_0(\cdot)$ for any t . Thus $V'(\tilde{t}|\Psi_0) = \int_{\theta \in \Theta} \frac{\partial v}{\partial t}(\theta, \tilde{t}) d\Psi_0(\theta)$.

Therefore, the expected information rent is

$$\begin{aligned} & \int_{\theta \in \Theta} \int_{t \in T} \left(\int_0^t q(\tilde{t}) V'(\tilde{t} | \Psi_0) d\tilde{t} \right) dF_1(t|\theta) d\Psi_0(\theta). \\ &= \int_{\theta \in \Theta} \int_{t \in T} \left(\int_0^t q(\tilde{t}) \frac{\partial v}{\partial t}(\theta, \tilde{t}) d\tilde{t} \right) dF_1(t) d\Psi_0(\theta), \end{aligned}$$

which is linear in Ψ_0 .

C Proof of Theorem 1

Let Σ denote the set of all finite signed measures on Θ . Endowed with a total variation norm (denoted by $\|\cdot\|$), Σ is a normed vector space, which includes $\Delta(\Theta)$ as a (norm-)closed and convex subset.

By Perlman (1974), Jensen's inequality applies if Π is convex and norm-continuous on $\Delta(\Theta)$. Thus, we show each of those properties in the following lemmas.

Lemma 5. Π is convex on $\Delta(\Theta)$.

Proof. Fix arbitrary $\Psi_0, \Psi'_0 \in \Delta(\Theta)$ and $\alpha \in (0, 1)$, and let $\Psi''_0 = \alpha\Psi_0 + (1 - \alpha)\Psi'_0$. Fix $\varepsilon > 0$ and let $q : T \rightarrow [0, 1]$ represent an ε -optimal allocation rule given Ψ_0 , i.e.,

$$\pi(q, \Psi_0) \geq \Pi(\Psi_0) - \varepsilon.$$

Similarly, let $q' : T \rightarrow [0, 1]$ ($q'' : T \rightarrow [0, 1]$) represent an ε -optimal allocation rule given Ψ'_0 (Ψ''_0).

Because q'' is monotonic, it is implementable given Ψ_0 and Ψ'_0 . Thus, we have

$$\Pi(\Psi_0) \geq \pi(q, \Psi_0) \geq \Pi(\Psi_0) - \varepsilon \geq \pi(q'', \Psi_0) - \varepsilon,$$

and

$$\Pi(\Psi'_0) \geq \pi(q, \Psi'_0) \geq \Pi(\Psi'_0) - \varepsilon \geq \pi(q'', \Psi'_0),$$

The weighted sum of the left-most and right-most expressions with weight $(\alpha, 1 - \alpha)$ is

$$\begin{aligned} \alpha\Pi(\Psi_0) + (1 - \alpha)\Pi(\Psi'_0) &\geq \alpha\pi(q'', \Psi_0) + (1 - \alpha)\pi(q'', \Psi'_0) - \varepsilon \\ &= \pi(q'', \Psi''_0) - \varepsilon \\ &= \Pi(\Psi''_0) - 2\varepsilon, \end{aligned}$$

where the first equality is because of linearity of π in the second argument. Because the inequality holds for any $\varepsilon > 0$, we have

$$\alpha\Pi(\Psi_0) + (1 - \alpha)\Pi(\Psi'_0) \geq \Pi(\Psi''_0),$$

implying that Π is convex. □

Lemma 6. Π is norm-continuous on $\Delta(\Theta)$, i.e., for each $\varepsilon > 0$, there exists $\delta > 0$ such that, if $\Psi_0, \Psi'_0 \in \Delta(\Theta)$ satisfies $\|\Psi_0 - \Psi'_0\| < \delta$,²⁹ then $|\Pi(\Psi_0) - \Pi(\Psi'_0)| < \varepsilon$.

²⁹ $\|\cdot\|$ represents a total variation norm on Σ , i.e., for each $\tilde{\Psi}_0 \in \Sigma$, $\|\tilde{\Psi}_0\|$ is the supremum of $\tilde{\Psi}_0(A)$ among all measurable $A \subseteq \Theta$.

Proof. We first show that, for any q , $\Pi(q, \cdot)$ is norm-continuous on Σ . Applying the same argument as in the proof of the last lemma, $\pi(q, \cdot) : \Psi_0 \mapsto \pi(q, \Psi_0)$ is a linear functional on a normed vector space Σ (not only on $\Delta(\Theta)$). Moreover, it is a bounded linear functional, because for $\Psi_0 \in \Sigma$,

$$|\pi(q, \Psi_0)| \leq \int_{\theta \in \Theta} \sup_{(q, \theta, t)} |qv(\theta, t) - c(q, \theta, t)| d\Psi_0(\theta) \leq \overline{M} \|\Psi_0\|,$$

where $\overline{M} = \sup_{(q, \theta, t)} |qv(\theta, t) - c(q, \theta, t)|$ is assumed to be bounded.

For an arbitrary $\varepsilon > 0$, let $\delta_\varepsilon = \frac{\varepsilon}{\overline{M}}$. Then, for any $\Psi_0, \Psi'_0 \in \Sigma$, if $\|\Psi_0 - \Psi'_0\| < \delta_\varepsilon$, then

$$\begin{aligned} |\pi(q, \Psi_0) - \pi(q, \Psi'_0)| &= |\pi(q, \Psi_0 - \Psi'_0)| \\ &\leq \overline{M} \|\Psi_0 - \Psi'_0\| < \varepsilon, \end{aligned}$$

that is, $\Pi(q, \cdot)$ is norm-continuous on Σ .

Now, contrarily to the statement, suppose that there exists $\varepsilon^* > 0$ such that, for all $\delta > 0$, there exist $\Psi_0, \Psi'_0 \in \Delta(\Theta)$ with $\|\Psi_0 - \Psi'_0\| < \delta$ and $\Pi(\Psi_0) \geq \Pi(\Psi'_0) + \varepsilon^*$. In particular, let $\delta = \delta_{\varepsilon^*/2} = \frac{\varepsilon^*}{2\overline{M}}$, and let $\Psi_0, \Psi'_0 \in \Delta(\Theta)$ be such that $\|\Psi_0 - \Psi'_0\| < \frac{\varepsilon^*}{2\overline{M}}$ and $\Pi(\Psi_0) \geq \Pi(\Psi'_0) + \varepsilon^*$. Note that $|\pi(q, \Psi_0) - \pi(q, \Psi'_0)| \leq \frac{\varepsilon^*}{2}$ for all nondecreasing q .

By definition of $\Pi(\Psi_0)$, there is nondecreasing q^* such that $\pi(q^*, \Psi_0) + \frac{\varepsilon^*}{2} >$

$\Pi(\Psi_0)$. Thus,

$$\begin{aligned}\pi(q^*, \Psi'_0) &\geq \pi(q^*, \Psi_0) - \frac{\varepsilon^*}{2} \\ &> \Pi(\Psi_0) - \varepsilon^* \\ &\geq \Pi(\Psi'_0),\end{aligned}$$

which contradicts that, for any nondecreasing q , we must have $\Pi(\Psi'_0) \geq \pi(q, \Psi'_0)$. \square

With these lemmas, Π^* admits Jensen's inequality in the sense that, for each $\Psi_0 \in \Delta(\Theta)$,

$$L(\Psi_0) \equiv \int_{\theta \in \Theta} \Pi(\delta_\theta) d\Psi_0(\theta) \geq \Pi(\Psi_0).$$

For each $\lambda \in \Delta(\Delta(\Theta))$ such that $\int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(\cdot) d\lambda = F_0(\cdot)$, because $L(\cdot)$ is linear,

$$\begin{aligned}\int_{\Psi_0 \in \Delta(\Theta)} \Pi(\Psi_0) d\lambda &\leq \int_{\Psi_0 \in \Delta(\Theta)} L(\Psi_0) d\lambda \\ &= L\left(\int_{\Psi_0 \in \Delta(\Theta)} \Psi_0 d\lambda\right) \\ &= L(F_0) \\ &= \int_{\theta \in \Theta} \Pi(\delta_\theta) dF_0(\theta),\end{aligned}$$

implying optimality of full disclosure.

D Proof of Lemma 2

First, because q is monotone, letting $q^{-1}(x) = \inf\{t|q(t) \geq x\}$,

$$R(q, \Psi_0) = \lim_{K \rightarrow \infty} \sum_{k=1}^K \int_{t=q^{-1}(\frac{k}{K})}^1 V'(t|\Psi_0) \int_{\theta \in \Theta} (1 - F_1(t|\theta)) d\Psi_0(\theta) dt.$$

Let $R_\tau(\Psi_0) = \int_{t=\tau}^1 V'(t|\Psi_0) \int_{\theta \in \Theta} (1 - F_1(t|\theta)) d\Psi_0(\theta) dt$. Then,

$$\begin{aligned} R_\tau(\Psi_0) &= -V(\tau|\Psi_0) \int_{\theta \in \Theta} (1 - F_1(\tau|\theta)) d\Psi_0(\theta) + \int_{\tau}^1 V(t|\Psi_0) \int_{\theta \in \Theta} f_1(t|\theta) d\Psi_0(\theta) dt \\ &= -V(\tau|\Psi_0) \int_{\theta \in \Theta} (1 - F_1(\tau|\theta)) d\Psi_0(\theta) + \int_{\tau}^1 \int_{\theta \in \Theta} v(\theta, t) f_1(t|\theta) d\Psi_0(\theta) dt, \end{aligned}$$

where the second equality is by definition of $V(t|\Psi_0)$.

Similarly,

$$\int_{\theta \in \Theta} R_\tau(\delta_\theta) d\Psi_0(\theta) = \int_{\theta \in \Theta} -v(\theta, \tau)(1 - F_1(\tau|\theta)) d\Psi_0(\theta) + \int_{\theta \in \Theta} \int_{\tau}^1 v(\theta, t) f_1(t|\theta) dt d\Psi_0(\theta).$$

Therefore,

$$\begin{aligned} &R_\tau(\Psi_0) - \int_{\theta \in \Theta} R_\tau(\delta_\theta) d\Psi_0(\theta) \\ &= \int_{\theta \in \Theta} v(\theta, \tau)(1 - F_1(\tau|\theta)) d\Psi_0(\theta) - V(\tau|\Psi_0) \int_{\theta \in \Theta} (1 - F_1(\tau|\theta)) d\Psi_0(\theta) \\ &= \int_{\theta \in \Theta} v(\theta, \tau)(1 - F_1(\tau|\theta)) d\Psi_0(\theta) - \left(\frac{\int_{\theta \in \Theta} v(\theta, \tau) f_1(\tau|\theta) d\Psi_0(\theta)}{\int_{\theta \in \Theta} f_1(\tau|\theta) d\Psi_0(\theta)} \right) \int_{\theta \in \Theta} (1 - F_1(\tau|\theta)) d\Psi_0(\theta) \\ &= \left[E_{\theta|\tau} \left[v(\theta, \tau) \frac{1 - F_1(\tau|\theta)}{f_1(\tau|\theta)} \right] - E_{\theta|\tau} [v(\theta, \tau)] E_{\theta|\tau} \left[\frac{1 - F_1(\tau|\theta)}{f_1(\tau|\theta)} \right] \right] \left(\int_{\theta \in \Theta} f_1(\tau|\theta) d\Psi_0(\theta) \right) \\ &\geq 0, \end{aligned}$$

where in the second last line,

$$E_{\theta|\tau}[\cdot] = \frac{\int_{\theta \in \Theta} (\cdot) f_1(\tau|\theta) d\Psi_0(\theta)}{\int_{\theta \in \Theta} f_1(\tau|\theta) d\Psi_0(\theta)}$$

represents expectation with respect to $\Psi_1(\cdot|t, \Psi_0)$. The last inequality is because $v(\theta, \tau)$ and $\frac{1-F_1(\tau|\theta)}{f_1(\tau|\theta)}$ are both increasing in θ .³⁰

E Proof of Theorem 2

We only provide a sketch of the proof because it is mostly analogous to that of Theorem 1.

For each nondecreasing q , define $\tilde{\pi}(q, \Psi_0)$ by

$$\tilde{\pi}(q, \Psi_0) = \int_{\theta \in \Theta} \int_t q(t) v(\theta, t) - c(q(t), \theta, t) dF_1(t|\theta) d\Psi_0(\theta) - \int_{\theta \in \Theta} R(q, \delta_\theta) d\Psi_0(\theta),$$

and $\tilde{\Pi}(\Psi_0)$ as the supremum of $\tilde{\pi}(q, \Psi_0)$ among all feasible q . Then, $\tilde{\pi}(q, \Psi_0) \geq \pi(q, \Psi_0)$ for all feasible q , and $\tilde{\Pi}(\Psi_0) \geq \Pi(\Psi_0)$, where the equality is obtained (for both of these inequalities) if $\Psi_0 = \delta_\theta$ for some θ .

Because $\tilde{\pi}(q, \Psi_0)$ is linear in Ψ_0 , as in the proof of Theorem 1, we can show that $\tilde{\Pi}(\Psi_0)$ is convex and norm-continuous on $\Delta(\Theta)$. Thus, applying Jensen's

³⁰See Schmidt (2014).

inequality, for each $\lambda \in \Delta(\Delta(\Theta))$ such that $\int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(\cdot) d\lambda = F_0(\cdot)$, we have

$$\begin{aligned} \int_{\Psi_0 \in \Delta(\Theta)} \Pi(\Psi_0) d\lambda &\leq \int_{\Psi_0 \in \Delta(\Theta)} \tilde{\Pi}(\Psi_0) d\lambda \\ &\leq \int_{\theta \in \Theta} \tilde{\Pi}(\delta_\theta) dF_0(\theta) \\ &= \int_{\theta \in \Theta} \Pi(\delta_\theta) dF_0(\theta), \end{aligned}$$

implying optimality of full disclosure.

F Proof of Theorem 4

To prove the theorem, we show that there exist θ^*, θ^{**} with $\theta^* < 1 < \theta^{**}$ such that a strictly higher expected surplus can be achieved than under full disclosure if the principal (i) fully discloses when $\theta \notin (\theta^*, \theta^{**})$ and (ii) discloses nothing when $\theta \in (\theta^*, \theta^{**})$. More specifically, for this second case, the principal offers a posted-price mechanism so that the full-trade outcome is implemented, where the implementability is implied by the following lemma.

Lemma 7. There exist $\theta_1 < 1 < \theta_2$ and $\tau : [\theta_1, 1] \rightarrow [1, \theta_2]$ such that (i) $\tau(\theta_1) = \theta_2$ and $\tau(1) = 1$, and (ii) for each $\theta \in (\theta_1, 1)$, $\int_\theta^{\tau(\theta)} \rho(\theta) dF_0(\theta) = 0$.³¹

Proof. Take $\theta_1 < 1$ so that $\int_{\theta_1}^{\bar{\theta}} \rho(\theta) dF_0(\theta) > 0$. Because $\int_1^{\bar{\theta}} \rho(\theta) dF_0(\theta) > 0$, by continuity, such θ_1 exists.

Then, again by continuity, for each $\theta \in [\theta_1, 1)$, we can take $\tau(\theta) > 1$ so that $\int_\theta^{\tau(\theta)} \rho(\theta) dF_0(\theta) = 0$. Finally, we let $\theta_2 = \tau(\theta_1)$. \square

³¹Recall $\rho(\theta) \equiv v_2(\theta, 0) - v_1(\theta, 1)$ defined in Assumption 1.

Let $\theta^* \in (\theta_1, 1)$ and $\theta^{**} = \tau(\theta^*)$. Suppose that the principal does not disclose any information when $\theta \in (\theta^*, \theta^{**})$. Then the full-trade outcome, $q(t) = 1$ for all t , is implementable by setting $-p_1(t) = p_2(t) = E[v_1(\theta, 1) | \theta \in (\theta^*, \theta^{**})]$ for all t .

Recall that $\Pi^1(\theta) = \int_{t \in T} (v_2(\theta, t_2) - v_1(\theta, t_1)) dF_T(t)$ represents the expected surplus in state θ under the full-trade outcome. It is not the first-best efficient outcome when $\theta < 1$, but we show that, if θ is sufficiently close to 1, then such a full-trade outcome is better than the best outcome under full disclosure. More specifically, we show the following lemma.

Lemma 8. There exist $\theta^* \in (\theta_1, 1)$ and a continuous function $\zeta : (\theta^*, 1) \rightarrow \mathbb{R}_{++}$ such that, for each $\theta \in (\theta^*, 1)$,

$$\Pi(\delta_\theta) \leq \Pi^1(\theta) - \zeta(\theta),$$

where δ_θ represents a Dirac measure on θ .

Proof. Fix $\theta \in (\theta_1, 1)$ arbitrarily. The Lagrangian of the problem of $\Pi(\delta_\theta)$ is

$$\int_t (1 + \lambda)(v_2(\theta, t_2) - v_1(\theta, t_1) - \lambda \left(\frac{\partial v_2}{\partial t_2}(\theta, t_2) \frac{1 - F_2(t_2)}{f_2(t_2)} + \frac{\partial v_1}{\partial t_1}(\theta, t_1) \frac{F_1(t_1)}{f_1(t_1)} \right)) q(t) dF_T(t).$$

Therefore, given $\lambda \geq 0$, the pointwise maximization of the Lagrangian yields $q(t) = 1$ if

$$v_2(\theta, t_2) - v_1(\theta, t_1) \geq \frac{\lambda}{1 + \lambda} \left(\frac{\partial v_2}{\partial t_2}(\theta, t_2) \frac{1 - F_2(t_2)}{f_2(t_2)} + \frac{\partial v_1}{\partial t_1}(\theta, t_1) \frac{F_1(t_1)}{f_1(t_1)} \right),$$

and $q(t) = 0$ otherwise.

Note that such q satisfies the monotonicity (i.e., nonincreasing in t_1 and

nondecreasing in t_2) because of Assumption 1. Therefore, the optimal q is determined by identifying the smallest λ such that the constraint is satisfied, which we denote by $\lambda(\theta)$. As shown by Myerson and Satterthwaite (1983), the Lagrangian multiplier for the budget constraint, $\lambda(\theta)$, is strictly positive and approaches 0 as $\theta \uparrow 1$.

Let $\beta(\theta)$ be the supremum value of t_1 such that $q(t_1, 0) = 1$ given $\lambda(\theta)$. Assumption 1 implies that there exists $\theta_3 > 0$ such that $\beta(\theta) > 0$ for any $\theta \in (\theta_3, 1)$. In the following, θ is always taken in this range.

For each t_1 , let $\alpha(\theta, t_1)$ be the infimum value of t_2 such that $q(t_1, t_2) = 1$. Then,

$$\begin{aligned} v_2(\theta, \alpha(t_1, \theta)) &= v_1(\theta, t_1) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \left(\frac{\partial v_2}{\partial t_2}(\theta, t_2) \frac{1 - F_2(t_2)}{f_2(t_2)} + \frac{\partial v_1}{\partial t_1}(\theta, t_1) \frac{F_1(t_1)}{f_1(t_1)} \right) \\ &\in \left[v_1(\theta, 1_1) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{b_1}{b_2} F_1(\beta(\theta_3)), v_1(\theta, t_1) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{2b_2}{b_1} \right], \end{aligned}$$

and similarly,

$$v_1(\theta, \beta(\theta)) \in \left[v_2(\theta, 0) - \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{2b_2}{b_1}, v_2(\theta, 0) - \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{b_1}{b_2} F_1(\beta(\theta_3)) \right],$$

Because $v_2(\theta, \alpha(t_1, \theta)) \in [v_2(\theta, 0) + b_1 \alpha(t_1, \theta), v_2(\theta, 0) + b_2 \alpha(t_1, \theta)]$, we have $\alpha(t_1, \theta) \in [\underline{\alpha}(t_1, \theta), \bar{\alpha}(t_1, \theta)]$, where

$$\begin{aligned} \bar{\alpha}(t_1, \theta) &\leq \left(v_1(\theta, t_1) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{2b_2}{b_1} - v_2(\theta, 0) \right) \frac{1}{b_1}, \\ \underline{\alpha}(t_1, \theta) &\geq \left(v_1(\theta, t_1) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{b_1}{b_2} F_1(\beta(\theta_3)) - v_2(\theta, 0) \right) \frac{1}{b_2}. \end{aligned}$$

Similarly, because $v_1(\theta, \beta(\theta)) \in [v_1(\theta, 1) - b_2(1 - \beta(\theta)), v_1(\theta, 1) - b_1(1 -$

$\beta(\theta))$], we have $\beta(\theta) \in [\underline{\delta}(\theta), \bar{\delta}(\theta)]$, where

$$\begin{aligned} 1 - \underline{\delta}(\theta) &\leq \left(-v_2(\theta, 0) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{2b_2}{b_1} + v_1(\theta, 1) \right) \frac{1}{b_1}, \\ 1 - \bar{\delta}(\theta) &\geq \left(-v_2(\theta, 0) + \frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{b_1}{b_2} F_1(\beta(\theta_3)) + v_1(\theta, 1) \right) \frac{1}{b_2}. \end{aligned}$$

Let $B(\theta, t) = v_2(\theta, t_2) - v_1(\theta, t_1) - \left(\frac{\partial v_2}{\partial t_2}(\theta, t_2) \frac{1 - F_2(t_2)}{f_2(t_2)} + \frac{\partial v_1}{\partial t_1}(\theta, t_1) \frac{F_1(t_1)}{f_1(t_1)} \right)$. Let θ be sufficiently close to 1 so that $B(\theta, t) < 0$ at $t = (\underline{\delta}(\theta), \bar{\alpha}(1, \theta))$. This is possible because B is continuous in (θ, t) and $B(1, (0, 1)) < 0$. Then,

$$\begin{aligned} 0 &\leq \int_{\beta(x)}^1 \int_{\alpha(t_1, x)}^1 B(x, t) dF_2(t_2) dF_1(t_1) + \int_0^{\beta(x)} \int_0^1 B(x, t) dF_2(t_2) dF_1(t_1) \\ &= \int_0^1 \int_0^1 B(\theta, t) dF_T(t) + \int_{\beta(x)}^1 \int_0^{\alpha(t_1, x)} (-B(x, t)) dF_2(t_2) dF_1(t_1) \\ &\leq \rho(\theta) + \int_{\underline{\delta}(x)}^1 \int_0^{\bar{\alpha}(x, t_1)} \left(\frac{2b_2}{b_1} - \rho(\theta) \right) b_2^2 dt_2 dt_1 \\ &\leq \rho(\theta) + \frac{b_2^2}{b_1^2} \left(\frac{2b_2}{b_1} - \rho(\theta) \right) \left(\frac{\lambda(\theta)}{1 + \lambda(\theta)} \frac{2b_2}{b_1} - \rho(\theta) \right)^2, \end{aligned}$$

where the equality is because $\int_0^1 \int_0^1 B(1, t) dF_T(t) = 0$, the second inequality is because $-B$ is positive and increasing in $(t_1, -t_2)$, for all $t \geq (\underline{\delta}(x), \bar{\alpha}(1, x))$. Therefore,

$$\frac{\lambda(\theta)}{1 + \lambda(\theta)} \geq \frac{b_1}{2b_2} \left(\rho(\theta) + \frac{b_1}{b_2} \sqrt{\frac{-\rho(\theta)b_1}{2b_2 - \rho(\theta)b_1}} \right).$$

Because $\rho(\theta) \in [(\theta - 1)b_1, (\theta - 1)b_2]$, the inequality above implies the following. We omit the proof because it is straightforward.

Lemma 9. There exists $\theta_4 \in (\theta_3, 1)$ such that, for all $\theta \in (\theta_4, 1)$,

$$\begin{aligned} 1 - \bar{\delta}(\theta) &\geq \phi\sqrt{1 - \theta}, \\ \underline{\alpha}(t_1, \theta) &\geq \phi\sqrt{1 - \theta} - 1 + t_1, \end{aligned}$$

where $\phi = \frac{b_1^2 F_1(\beta(\theta_3))}{2b_2^4} \sqrt{\frac{b_1^2}{2b_2 + b_1 b_2(1 - \theta_3)}} (> 0)$.

Finally, letting $y = \sqrt{1 - \theta}$, we have

$$\begin{aligned} \Pi^1(\theta) - \Pi(\delta_\theta) &\geq \int_{\bar{\delta}(\theta)}^1 \int_0^{\underline{\alpha}(t_1, \theta)} (v_2(\theta, t_2) - v_1(\theta, t_1)) dF_2(t_2) dF_1(t_1) \\ &\geq \int_{1-\phi y}^1 \int_0^{t_1-1+\phi y} (-b_2 y^2 + b_1(t_2 - t_1 + 1)) dF_2(t_2) dF_1(t_1) \\ &\geq -\frac{1}{2} b_2^3 y^4 \phi^2 + b_1^3 \int_{1-\phi y}^1 \int_0^{t_1-1+\phi y} (t_2 - t_1 + 1) dt_2 dt_1 \\ &= \frac{y^3 \phi^2}{6} (2b_1^3 \phi - 3b_2^3 y), \end{aligned}$$

which is strictly positive if y is sufficiently close to 0. Therefore, there exists $\theta^* \in (\theta_4, 1)$ such that, for all $\theta \in (\theta^*, 1)$, we have

$$\zeta(\theta) \equiv \frac{y^3 \phi^2}{6} (2b_1^3 \phi - 3b_2^3 y) > 0,$$

where ζ is obviously continuous for $\theta \in (\theta^*, 1)$. □

With this lemma, we complete the proof of the theorem as follows. As described above, consider a disclosure policy such that the principal (i) fully discloses when $\theta \notin (\theta^*, \theta^{**})$ and (ii) discloses nothing when $\theta \in (\theta^*, \theta^{**})$. For case (i), the principal offers the same mechanism as under full disclosure. For

case (ii), the principal implements the full-trade outcome by a posted-price mechanism so that $\Pi^1(\theta)$ is achieved. Then, the ex ante expected surplus under this disclosure policy is higher than under full disclosure by

$$\int_{\theta^*}^1 (\Pi^1(\theta) - \Pi(\delta_\theta)) dF_0(\theta) \geq \int_{\theta^*}^1 \zeta(\theta) dF_0(\theta) > 0.$$

G Proof of Lemma 3

(i) Fix $x \in [\underline{\theta}, \bar{\theta}]$ arbitrarily, and let $\Psi_0 \in \Delta(\Theta)$ be such that $x = \int_{\theta \in \Theta} \theta d\Psi_0(\theta)$. The principal's mechanism design problem is given by

$$\begin{aligned} \Pi(\Psi_0) = & \sup_{\chi: T \rightarrow X} \int_{t \in T} u_0(\chi(t), x, t) dF_T(t) \\ \text{sub. to} & \quad \forall i, t_i, t'_i, \\ & \int_{t_{-i} \in T_{-i}} u_i(\chi(t), x, t) dF_{-i}(t_{-i}) \\ & \geq \max\{0, \int_{t_{-i} \in T_{-i}} u_i(\chi(t'_i, t_{-i}), x, t) dF_{-i}(t_{-i})\}. \end{aligned}$$

Because the right-hand side depends on Ψ_0 only through x , we have $\Pi(\Psi_0) = \Pi(\Psi'_0)$ if $\int_{\theta \in \Theta} \theta d\Psi_0(\theta) = \int_{\theta \in \Theta} \theta d\Psi'_0(\theta)$.

(ii) Let $\sigma = \Pr(s = 1)$ be the (marginal) probability for $s = 1$. For each θ , let

$$\xi(\theta) = \frac{\sigma f_0^1(\theta)}{\sigma f_0^1(\theta) + (1 - \sigma) f_0^0(\theta)}$$

be the conditional probability for $s = 1$ given θ .

Let $\underline{\theta} = 0$, $\bar{\theta} = 1$, and fix $x \in [0, 1]$ arbitrarily. Let $\Psi_0 \in \Delta(\Theta)$ be such that $x = \int_{\theta \in \Theta} \xi(\theta) d\Psi_0(\theta)$. Given Ψ_0 , agent i 's posterior for $s = 1$ with type t_i is

$$\frac{x f_i^1(t_i)}{x f_i^1(t_i) + (1-x) f_i^0(t_i)},$$

which we denote by $\sigma_i(x, t_i)$.

Then, the principal's mechanism design problem is given by

$$\begin{aligned} \Pi(\Psi_0) &= \sup_{\chi: T \rightarrow X} \int_{t \in T} [x u_0(\chi(t), 1, t) f_T^1(t) + (1-x) u_0(\chi(t), 0, t) f_T^0(t)] dt \\ &\text{sub. to } \forall i, t_i, t'_i, \\ &\int_{t_{-i} \in T_{-i}} [\sigma_i(x, t_i) u_i(\chi(t), 1, t) f_{-i}^1(t_{-i}) + (1 - \sigma_i(x, t_i)) u_i(\chi(t), 0, t) f_{-i}^0(t_{-i})] dt_{-i} \\ &\geq \max\{0, \int_{t_{-i} \in T_{-i}} [\sigma_i(x, t_i) u_i(\chi(t'_i, t_{-i}), 1, t) f_{-i}^1(t_{-i}) \\ &\quad + (1 - \sigma_i(x, t_i)) u_i(\chi(t'_i, t_{-i}), 0, t) f_{-i}^0(t_{-i})] dt_{-i}\}. \end{aligned}$$

Because the right-hand side depends on Ψ_0 only through x , we have $\Pi(\Psi_0) = \Pi(\Psi'_0)$ if $\int_{\theta \in \Theta} \xi(\theta) d\Psi_0(\theta) = \int_{\theta \in \Theta} \xi(\theta) d\Psi'_0(\theta)$.

H Proof of Theorem 5

The first key step of the proof of the theorem is to observe that our problem is to choose F that is less riskier than F_0 in the sense of Rothschild and Stiglitz (1970) (i.e., F has the same mean as F_0 , and F second-order stochastically dominates F_0).

Lemma 10.

$$\begin{aligned} \Pi^* &= \sup_{F \in \Delta(\Theta)} \int_{x \in [\underline{\theta}, \bar{\theta}]} \Pi(x) dF(x) \\ \text{sub. to} & \int_{\underline{\theta}}^x F(y) dy \leq \int_{\underline{\theta}}^x F_0(y) dy, \quad \forall x \\ & \int_{\underline{\theta}}^{\bar{\theta}} F(y) dy = \int_{\underline{\theta}}^{\bar{\theta}} F_0(y) dy. \end{aligned}$$

A version of this Lemma (and its explanation based on the riskiness of Rothschild and Stiglitz (1970)) appears in Gentzkow and Kamenica (2015).³² The key intuition is the following. Imagine that the principal recommends the posterior mean $x \in [\underline{\theta}, \bar{\theta}]$ according to a distribution F in an “honest” way, that is, after Bayesian updating, the (principal’s) posterior mean conditional on recommendation x is indeed x . This means that $E[\theta|x] = x$, which implies: (i) by taking expectation with respect to $x \sim F$ on both sides, we obtain $E[\theta] = E[x]$; and (ii) the distribution of θ , F_0 , is second-order stochastically dominated by the distribution of x , F . Property (i) is equivalent to $\int_{\underline{\theta}}^{\bar{\theta}} F(y) dy = \int_{\underline{\theta}}^{\bar{\theta}} F_0(y) dy$, and Property (ii) is equivalent to $\int_{\underline{\theta}}^x F(y) dy \leq \int_{\underline{\theta}}^x F_0(y) dy$ for all x .

Proof. First, we show that Π^* is not lower than the right-hand side value of the statement. Take any F that is feasible in the right-hand side problem. As in Rothschild and Stiglitz (1970), if $x \sim F$, there exists a random variable ε such that $E[\varepsilon|x] = 0$ and $(x + \varepsilon) \sim F_0$. Now, let $\Psi_0^x \in \Delta(\Theta)$ denote the conditional distribution for $(x + \varepsilon)|x$. Define $\lambda \in \Delta(\Delta(\Theta))$ so that, for each

³²See also Kolotilin et al. (2015).

measurable $A \subseteq \Delta(\Theta)$,

$$\lambda(A) = \int_{x \in [\underline{\theta}, \bar{\theta}]} 1\{\Psi_0^x \in A\} dF(x).$$

Then, for each measurable $B \subseteq \Theta$,

$$\int_{\Psi_0 \in \Delta(\Theta)} \Psi_0(B) d\lambda = \int_{x \in [\underline{\theta}, \bar{\theta}]} \Psi_0^x(B) dF(x) = F_0(B),$$

and hence, this λ is feasible in the original problem (presented right after Proposition 2 in Section 2.3).

Moreover,

$$\int_{\Psi_0 \in \Delta(\Theta)} \Pi(\Psi_0) d\lambda = \int_{x \in [\underline{\theta}, \bar{\theta}]} \Pi(\Psi_0^x) dF(x) = \int_{x \in [\underline{\theta}, \bar{\theta}]} \Pi(x) dF(x),$$

and therefore, the value of the original problem (i.e., Π^*) is not lower than the right-hand side value above.

Next, we show that the right-hand side value is not lower than Π^* . Take any λ that is feasible in the original problem. Define F so that, for $y \in [\underline{\theta}, \bar{\theta}]$,

$$F(y) = \int_{\Psi_0 \in \Delta(\Theta)} 1\{e_{\Psi_0} \leq y\} d\lambda,$$

where $e_{\Psi_0} = \int_{z \in [\underline{\theta}, \bar{\theta}]} z d\Psi_0(z)$.

Thus, for any $x \in [\underline{\theta}, \bar{\theta}]$, we have

$$\begin{aligned} \int_{\underline{\theta}}^x (F_0(y) - F(y))dy &= \int_{\underline{\theta}}^x \int_{\Psi_0 \in \Delta(\Theta)} (\Psi_0([\underline{\theta}, y]) - 1\{e_{\Psi_0} \leq y\}) d\lambda dy \\ &= \int_{\Psi_0 \in \Delta(\Theta)} \left(\int_{\underline{\theta}}^x \Psi_0([\underline{\theta}, y]) dy - \int_{\min\{x, e_{\Psi_0}\}}^x 1 dy \right) d\lambda. \end{aligned}$$

For each $x \leq e_{\Psi_0}$, it is nonnegative because $\min\{x, e_{\Psi_0}\} = x$. For $x = \bar{\theta}$, it is zero because

$$\begin{aligned} \int_{\underline{\theta}}^{\bar{\theta}} \Psi_0([\underline{\theta}, y]) dy &= 1 - \int_{y \in [\underline{\theta}, \bar{\theta}]} y d\Psi_0(y) = 1 - e_{\Psi_0}, \\ \int_{\min\{\bar{\theta}, e_{\Psi_0}\}}^1 1 dy &= 1 - e_{\Psi_0}. \end{aligned}$$

For $x \in (e_{\Psi_0}, \bar{\theta})$, it is nonnegative because

$$\int_{\underline{\theta}}^x \Psi_0([\underline{\theta}, y]) dy - \int_{\min\{x, e_{\Psi_0}\}}^x 1 dy = 0 - \int_x^{\bar{\theta}} (\Psi_0([\underline{\theta}, y]) - 1) dy \geq 0.$$

Therefore, this F is feasible in the new problem.

Moreover,

$$\int_{\Psi_0 \in \Delta(\Theta)} \Pi(\Psi_0) d\lambda = \int_{\Psi_0 \in \Delta(\Theta)} \Pi(e_{\Psi_0}) d\lambda = \int_{x \in [\underline{\theta}, \bar{\theta}]} \Pi(x) dF(x),$$

and therefore, the value of the new problem (i.e., the right-hand side value in the statement) is not lower than the value of the original problem (i.e., Π^*). \square

In the rest of this section, we sketch the proof idea of the theorem under

the stronger assumption that Π is twice continuously differentiable on $[\underline{\theta}, \bar{\theta}]$ (so Π' and Π'' exist everywhere and are continuous on $[\underline{\theta}, \bar{\theta}]$).³³ Hence, in what follows, without loss of generality, we assume that $\Pi''(x) \geq 0$ on (x_{k-1}, x_k) for each k odd, and $\Pi''(x) \leq 0$ on (x_{k-1}, x_k) for each k even. For the general case (with Assumption 2), see Supplementary Materials (Section L).

Let $H_0(x) = \int_{\underline{\theta}}^x F_0(y)dy$ for $x \in [\underline{\theta}, \bar{\theta}]$. The following lemma further rewrites the problem, by viewing $H(\cdot) = \int_{\underline{\theta}}^{(\cdot)} F(y)dy$ as a choice variable, instead of $F(\cdot)$ itself. The proof is omitted.

Lemma 11.

$$\begin{aligned} \Pi^* &= \sup_{H: \Theta \rightarrow \mathbb{R}} && \Pi(\bar{\theta}) - \Pi'(\bar{\theta})H_0(\bar{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)H(x)dx \\ &\text{sub. to} && H(x) \leq H_0(x), \forall x, \\ &&& H(\underline{\theta}) = H_0(\underline{\theta}) = 0, \quad H(\bar{\theta}) = H_0(\bar{\theta}), \\ &&& H \text{ is convex, } \frac{H(x) - H(x')}{x - x'} \in [0, 1], \forall x > x'. \end{aligned}$$

The first two lines of the constraints are because F is less riskier than F_0 . The last line of the constraints is because H is obtained by integrating a cdf F : because F is nondecreasing, H must be convex; and because F takes a value between 0 and 1, the slope of H must be between 0 and 1. Finally, the objective is obtained by applying integration-by-parts (twice) on $\int_{\underline{\theta}}^{\bar{\theta}} \Pi(x)dF(x)$.³⁴ Because the first two terms in the new objective are constants, we consider maximization of $\int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)H(x)dx$.

We can solve this problem in two steps. First, fix an arbitrary vector

³³As a convention, we define a function's derivative at $x = 0$ ($x = 1$) by its right (left) derivative.

³⁴For the validity of integration-by-parts, see the proof in the appendix.

$\{h_k\}_{k=0}^K$ such that there exists a feasible \hat{H} with $\hat{H}(x_k) = h_k$ for each k . As the first step, we consider a constrained problem where the principal chooses H that is feasible and satisfies $H(x_k) = h_k$ for $k = 0, \dots, K$. Then, as the second step, we optimize the first-step value function with respect to $\{h_k\}_{k=0}^K$.

We define the following $\tilde{H} : [\underline{\theta}, \bar{\theta}] \rightarrow \mathbb{R}$. First, $\tilde{H}(x_k) = h_k$ for all k . Second, for each k odd, \tilde{H} coincides with the largest convex function that is below both H_0 and affine function:

$$\bar{\xi}^k : x \mapsto \frac{h_k - h_{k-1}}{x_k - x_{k-1}}(x - x_{k-1}) + h_{k-1}.$$

Let $f_0 = 0$, $f_K = 1$, and for each $k = 3, \dots, K - 1$ odd, let $f_{k-1} = \tilde{H}'(x_{k-1} + 0)$ and $f_k = \tilde{H}'(x_k - 0)$. Finally, for each k even and each $x \in (x_{k-1}, x_k)$, $\tilde{H}(x) = \underline{\xi}^k(x)$, where:

$$\underline{\xi}^k : x \mapsto \max\{h_{k+1} - f_{k+1}(x_{k+1} - x), h_k + f_k(x - x_k)\}.$$

Lemma 12. \tilde{H} is feasible. Moreover, for any feasible H such that $H(x_k) = h_k$ for $k = 0, \dots, K$, we have

$$\int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)H(x)dx \leq \int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)\tilde{H}(x)dx,$$

that is, \tilde{H} solves the first-step problem.

Proof. First, we show that \tilde{H} is feasible.

Claim 1: $\tilde{H} \leq H_0$.

It is obvious by construction for $x \in [x_{k-1}, x_k]$ on which Π is convex. For

$x \in [x_{k-1}, x_k]$ on which Π is concave, we have

$$H_0(x) \geq \hat{H}(x) \geq \underline{\xi}^k(x) = \tilde{H}(x).$$

Claim 2: $\tilde{H}(\underline{\theta}) = H_0(\underline{\theta})$ and $\tilde{H}(\bar{\theta}) = H_0(\bar{\theta})$.

We have $\tilde{H}(\underline{\theta}) = h_0 = H_0(\underline{\theta})$ and $\tilde{H}(\bar{\theta}) = h_K = H_0(\bar{\theta})$.

Claim 3: $\frac{\tilde{H}(x) - \tilde{H}(x')}{x - x'} \in [0, 1]$ for $x > x'$.

Fix x, x' with $\underline{\theta} \leq x' < x \leq \bar{\theta}$. By construction, clearly we have $\frac{\tilde{H}(x) - \tilde{H}(x')}{x - x'} \geq 0$. In order to show the other inequality, let $k \in \{1, \dots, K\}$ be such that $x \leq x_k$. Then:

$$\frac{\tilde{H}(x) - \tilde{H}(x')}{x - x'} \leq f_k \leq f_K = 1.$$

Claim 4: \tilde{H} is convex.

For each open subinterval (x_{k-1}, x_k) , \tilde{H} is clearly convex. Thus, we complete the proof of the claim by showing that \tilde{H} is convex around each boundary point x_k , $k = 1, \dots, K - 1$. This is indeed true, because $H'(x_{k+}) \geq f_k \geq H'(x_{k-})$.

Next, we show that \tilde{H} is optimal. If there exists H that is feasible in the first-step problem and strictly better than \tilde{H} , either (i) there exist k odd and $x \in (x_{k-1}, x_k)$ such that $H(x) > \tilde{H}(x)$, or (ii) there exist k even and $x \in (x_{k-1}, x_k)$ such that $H(x) < \tilde{H}(x)$.

If (i) holds with convex H with $H \leq H_0$, then $H \leq \bar{\xi}^k$ is violated, because \tilde{H} is the largest convex function that is below both H_0 and $\bar{\xi}^k$. Then, we have either $H(x_{k-1}) \neq h_{k-1}$ or $H(x_k) \neq h_k$, which contradicts that H is feasible in the first-step problem.

If (ii) holds with convex H , then $H'(x_{k-1}+) < f_{k-1}$ or $H'(x_k-) > f_k$. Again, it contradicts that H is feasible in the first-step problem.

In conclusion, \tilde{H} is optimal in the first-step problem. \square

The lemma implies that, for any $\{h_k\}_{k=0}^K$ fixed in a feasible way, the solution to the first-step problem satisfies the following property: there exist $L \leq K$ and disjoint open sub-intervals $\{I_l\}_{l=1}^L$ such that (i) \tilde{H} is piecewise-linear on each I_l and $\tilde{H}(x) < H_0(x)$ for $x \in I_l$, and (ii) $\tilde{H}(x) = H_0(x)$ otherwise. Thus, for the optimal F^* in the original problem (obtained by the right derivative of \tilde{H} given the optimal choice of $\{h_k\}_{k=0}^K$) there exist $L \leq K$ and disjoint open sub-intervals $\{I_l\}_{l=1}^L$ such that (i) $F^*(x)$ is piecewise constant and $\int_{\underline{\theta}}^x F^*(y)dy < \int_{\underline{\theta}}^x F_0(y)dy$ for each l and each $x \in I_l$, and (ii) $F^*(x) = F_0(x)$ and $\int_{\underline{\theta}}^x F^*(y)dy = \int_{\underline{\theta}}^x F_0(y)dy$ otherwise.

So far, the optimal F^* is partially characterized by examination of the first step of the problem, i.e., with arbitrarily fixed $\{h_k\}_{k=0}^K$. Next, we show that optimal choice of $\{h_k\}_{k=0}^K$ implies that F^* can have at most two jumps on each interval I_l (on which F^* is piecewise constant).

Contrarily, suppose that F^* has three or more (but finitely many) jumps on some interval I_l .³⁵ Consider any three consecutive jump points, $c_1, c_2, c_3 \in I_l$ with $c_1 < c_2 < c_3$ (i.e., $F^*(c_\kappa-) < F^*(c_\kappa) = F^*(c_{\kappa+1}-)$ for each κ). Define

³⁵The previous argument implies that the first-step optimal \tilde{H} can have at most one kink on each interval (x_k, x_{k-1}) , and thus, \tilde{H} can have at most finitely many kinks overall. That is, the optimal F^* can have at most finitely many jumps.

\tilde{F} as follows:

$$\tilde{F}(x) = \begin{cases} F^*(c_1) + \varepsilon & \text{if } x \in [c_1, c_2), \\ F^*(c_2) - \frac{c_2 - c_1}{c_3 - c_2} \varepsilon & \text{if } x \in [c_2, c_3), \\ F^*(x) & \text{otherwise.} \end{cases}$$

Notice that $\int_{c_1}^{c_3} \tilde{F}(x) dx = \int_{c_1}^{c_3} F^*(x) dx$. Thus, for sufficiently small $\varepsilon > 0$, \tilde{F} is feasible.³⁶ \tilde{F} attains the principal's expected payoff higher than F^* does by the following amount:

$$\left(\Pi(c_1) - \frac{c_2 - c_1}{c_3 - c_2} \Pi(c_2) \right) \varepsilon.$$

Thus, if this is strictly positive, then it contradicts that F^* is optimal. So assume that it is non-positive. If this is strictly negative, however, we again obtain a contradiction, by considering an alternative \hat{F} as follows:

$$\hat{F}(x) = \begin{cases} F^*(c_1) - \varepsilon & \text{if } x \in [c_1, c_2), \\ F^*(c_2) + \frac{c_2 - c_1}{c_3 - c_2} \varepsilon & \text{if } x \in [c_2, c_3), \\ F^*(x) & \text{otherwise.} \end{cases}$$

Therefore, we must have $\frac{c_2 - c_1}{c_3 - c_2} \Pi(c_2) = \Pi(c_1)$. In this case, F^* and \hat{F} attain exactly the same expected payoff for the principal, and this is true for any $\varepsilon > 0$ as long as the corresponding \hat{F} is feasible. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\} (> 0)$, where ε_1 solves $\hat{F}(c_1) = \hat{F}(c_1 -)$ and ε_2 solves $\hat{F}(c_3) = \hat{F}(c_3 -)$. In either

³⁶Because F^* has finitely many jumps, for sufficiently small $\varepsilon > 0$, \tilde{F} is a cdf; in particular, non-decreasing. Also, because $\int_{\underline{\theta}}^x F^*(x) dx < \int_{\underline{\theta}}^x F_0(x) dx$ for all $x \in I_l$, for sufficiently small $\varepsilon > 0$, $\int_{\underline{\theta}}^x \tilde{F}(x) dx < \int_{\underline{\theta}}^x F_0(x) dx$. Finally, we clearly have $\int_{\underline{\theta}}^{\bar{\theta}} F^*(x) dx = \int_{\underline{\theta}}^{\bar{\theta}} F_0(x) dx$.

case with $\varepsilon = \varepsilon_1$ or $\varepsilon = \varepsilon_2$, the corresponding \hat{F} is feasible, and has a strictly smaller number of jumps on I_l than the original F^* . We can continue this process of reducing jumps in I_l without changing the value of the objective, until at most two jumps exist in I_l .

Finally, we construct (M, G) that implements F^* , exhibits binary lower truncation on each I_l , and exhibits full-disclosure otherwise. Let $M = \Theta$, and let $G(\{\theta\}|\theta) = 1$ if $\theta \notin I_l$ for any l .

For each I_l where F^* has only one jump (say at z^*), we have $z^* = \frac{\int_a^b \theta dF_0(\theta)}{F_0(b) - F_0(a)}$ due to the requirement that F^* and F_0 have the same mean. Thus, on this I_l , we consider a binary lower truncation policy with (arbitrary $y \in I_l$ and) $w = 1$: $G(\{z_1\}|\theta) = 1$ for all $\theta \in I_l$. This is essentially a full-pooling policy on I_l , implying that (M, G) implements F^* on I_l .

Finally consider each $I_l(= (a, b))$ where F^* has two (non-trivial) jumps, say at z_1 and z_2 . Note that $F_0(a) = F^*(a) < F^*(z_2) = F^*(b) = F_0(b)$, and that $z_1 < \frac{\int_a^b \theta dF_0(\theta)}{F_0(b) - F_0(a)} < z_2$ due to the requirement that F^* and F_0 have the same mean. First, we define $y \in (a, F_0^{-1}(F^*(z_1)))$ ³⁷ by

$$\int_a^y \theta dF_0(\theta) + z_2(F^*(z_1) - F_0(y)) - z_1(F^*(z_1) - F_0(a)) = 0.$$

To see that y is well-defined, notice that, as a function of \tilde{y} ,

$$\int_{y_0}^{\tilde{y}} x dF_0(x) + z_2(F^*(z_1) - F_0(\tilde{y})) - z_1(F^*(z_1) - F_0(a))$$

is continuous, strictly positive at $\tilde{y} = a$,³⁸ and strictly negative at $\tilde{y} =$

³⁷Note that $z_1 > a$ and $F^*(a) = F_0(a)$ imply $F_0^{-1}(F^*(z_1)) > F_0^{-1}(F_0(a)) = a$; and that $z_1 < b$ and $F^*(b) = F_0(b)$ imply $F_0^{-1}(F^*(z_1)) < F_0^{-1}(F_0(b)) = b$.

³⁸If $\tilde{y} = a$, we have $(z_2 - z_1)(F^*(z_1) - F_0(a)) > 0$.

$F_0^{-1}(F^*(z_1))$,³⁹ which means that a root of this function exists in $(y_0, F_0^{-1}(F^*(z_1)))$.

Next, we define w by

$$w = \frac{F^*(z_1) - F_0(y)}{F_0(b) - F_0(y)}.$$

Consider a binary lower truncation policy on $I_l = (a, b)$ identified by y and w such that

$$G(\{z_1\}|\theta) = \begin{cases} 1 & \text{if } \theta \in (a, y), \\ w & \text{if } \theta \in (y, b), \end{cases}$$

and $G(\{z_2\}|\theta) = 1 - w$ for $\theta \in (y, b)$, where we can verify that

$$z_j = \begin{cases} \frac{\int_a^y \theta dF_0(\theta) + \int_y^b w \theta dF_0(\theta)}{\int_a^y 1 dF_0(\theta) + \int_y^b w dF_0(\theta)} & \text{if } j = 1, \\ \frac{\int_y^b \theta dF_0(\theta)}{\int_y^b 1 dF_0(\theta)} & \text{if } j = 2. \end{cases} \left(= E[\theta | z_j \text{ is announced}] \right).$$

Observe that this (M, G) implements F^* on $I_l = (a, b)$, because in the ex

³⁹If $\tilde{y} = F_0^{-1}(F^*(z_1))$, then $\tilde{y} < b$ because $F_0(\tilde{y}) = F^*(z_1) < F^*(b)$. Thus, $\int_a^{F_0^{-1}(F^*(z_1))} F_0(x) dx > \int_a^{F_0^{-1}(F^*(z_1))} F^*(x) dx = (z_1 - a)F_0(a) + \int_{z_1}^{F_0^{-1}(F^*(z_1))} F^*(x) dx$, which implies

$$\begin{aligned} & \int_a^{F_0^{-1}(F^*(z_1))} x dF_0(x) - z_1(F^*(z_1) - F_0(a)) \\ &= F^*(z_1)(F_0^{-1}(F^*(z_1)) - z_1) + (z_1 - a)F_0(a) - \int_a^{F_0^{-1}(F^*(z_1))} F_0(x) dx \\ &< \int_{z_1}^{F_0^{-1}(F^*(z_1))} F^*(z_1) - F^*(x) dx \leq 0. \end{aligned}$$

ante perspective, z_1 is announced with probability

$$\int_a^y 1dF_0(\theta) + \int_y^b w dF_0(\theta) = F^*(z_1) - F^*(a),$$

and z_2 is announced with probability

$$\int_y^b (1 - w)dF_0(\theta) = F_0(b) - F^*(z_1) = F^*(z_2) - F^*(z_1).$$

I Proof of Lemma 4

Given that t follows an independent uniform distribution on $[0, 1]$, the second-best trading rule satisfies $q(t) = 1$ if

$$\begin{aligned} (t_2 + x - t_1)(1 + \lambda(x)) &\geq (1 - t_2 + t_1)\lambda(x) \\ \Leftrightarrow t_2 + x - t_1 &\geq \frac{(1 + x)\lambda(x)}{1 + 2\lambda(x)} \equiv \eta(x), \end{aligned}$$

where $\lambda(x)$ is the Lagrange multiplier of the problem of $\Pi(x)$, and $q(t) = 0$ otherwise. The first-best efficiency is achieved if and only if $\eta(x) = 0$, and the size of $\eta(x)$ represents inefficiency of the second-best trading.

We first characterize $\eta(x)$ for each $x \in \mathbb{R}$. By Myerson and Satterthwaite (1983), $\eta(x) = 0$ if and only if $x \leq -1$ or $x \geq 1$ (called the ‘‘gap’’ cases), where the first-best efficiency is achieved by a simple posted-price mechanism, implying $\Pi(x) = 0$ for $x \leq -1$ and $\Pi(x) = E[v_2 + x - v_1] = x$ for $x \geq 1$. Thus, in the following, we consider the other case with $x \in (-1, 1)$, where $\eta(x)$ must be strictly positive.

Recall that $\lambda(x)$ (and hence $\eta(x)$) is determined to satisfy the budget-

balance constraint with equality. Given arbitrarily given $\eta > 0$, consider a trading rule such that $q(t) = 1$ if $t_2 + x - t_1 \geq \eta$, and $q(t) = 0$ otherwise. The budget surplus is then

$$\begin{aligned}
B(x, \eta) &= \int_{t \in T} (2t_2 - 2t_1 + x - 1)q(t)dt \\
&= \begin{cases} 0 & \text{if } \eta \geq 1 + x, \\ (1 + x - \eta)^2 \left(\frac{1+x-\eta}{3} - \frac{1+x-2\eta}{2} \right) & \text{if } \eta \in [x, 1 + x), \\ x - 1 + \frac{(1-x+\eta)^2(5+x-4\eta)}{6} & \text{if } \eta \in (0, x). \end{cases}
\end{aligned}$$

With respect to η , B is continuously differentiable and has single-crossing. We have

$$\eta(x) \geq x \Leftrightarrow B(x, x) \leq 0 \Leftrightarrow x \leq \frac{1}{3}.$$

Case (I): $x \in (-1, \frac{1}{3})$.

In this case, we have

$$(1 + x - \eta(x))^2 \left(\frac{1 + x - \eta(x)}{3} - \frac{1 + x - 2\eta(x)}{2} \right) = 0,$$

and hence, $\eta(x) = \frac{1+x}{4}$.

We obtain $\Pi(x) = \frac{9(1+x)^3}{64}$, and therefore, $\Pi(x)$ is convex in this region. Π' and Π'' exist and continuous in this region, and moreover, $\lim_{x \downarrow -1} \Pi'(x) = \lim_{x \downarrow -1} \Pi'(x) = 0$, and hence, Π' and Π'' are continuous at $x = 0$ too. At the other extreme point, $\lim_{x \uparrow \frac{1}{3}} \Pi' = \frac{3}{4}$ and $\lim_{x \uparrow \frac{1}{3}} \Pi'' = \frac{9}{8}$.

Case (II): $x \in (\frac{1}{3}, 1)$.

In this case, we have

$$x - 1 + \frac{(1 - x + \eta(x))^2(5 + x - 4\eta(x))}{6} = 0,$$

and hence, we do not have a closed-form expression for $\eta(x)$.

To simplify the expression, let $z(x) = 1 - x + \eta(x) \in (0, 1)$. Then the budget-balance condition becomes

$$x = \frac{9z(x)^2 - 4z(x)^3 - 6}{3z(x)^2 - 6}.$$

By the implicit function theorem, we have

$$\begin{aligned} z'(x) &= \frac{-3(2 - z(x)^2)^2}{4(z(x)^4 - 6z(x)^2 + 6z(x))} < 0, \\ z''(x) &= \frac{-9(2 - z(x)^2)^3}{8(z(x)^4 - 6z(x)^2 + 6z(x))^3}(-2z(x)^3 + 9z(x)^2 - 12z(x) + 6) < 0, \\ z'''(x) &= \frac{-27(2 - z(x)^2)^4}{32(z(x)^4 - 6z(x)^2 + 6z(x))^5}6(1 - z(x))^2(z(x)^4(2 - z(x))^2 + 24(1 - z(x))^2 + 12) < 0. \end{aligned}$$

The expected social surplus is

$$\begin{aligned} \Pi(x) &= \int_{t_1=0}^1 \int_{t_2=0}^1 t_2 + x - t_1 dt - \int_{t_1=1-z(x)}^1 \int_{t_2=0}^{t_1-(1-z(x))} t_2 + x - t_1 dt \\ &= x - \frac{z(x)^2(z(x) + x - 1)}{2} + \frac{z(x)^3}{6}, \end{aligned}$$

and thus, under the budget-balance condition,

$$\begin{aligned}\Pi(x) &= \frac{9z(x)^2 - 4z(x)^3 - 6}{3z(x)^2 - 6} - \frac{z(x)^2(z(x) + x - 1)}{2} + \frac{z(x)^3}{6} \\ &= \frac{z(x)^3}{6} - z(x)^2 + 1.\end{aligned}$$

To examine the shape of Π , let $T(z) = \frac{z^3}{6} - z^2 + 1$ for $z \in (0, 1)$. Then, $T'(z) = z^2 - 2z$, $T''(z) = 2z - 2$, $T'''(z) = 2$, and

$$\begin{aligned}\Pi'(x) &= T'(z(x))z'(x), \\ \Pi''(x) &= T''(z(x))(z'(x))^2 + T'(z(x))z''(x), \\ \Pi'''(x) &= T'''(z(x))(z'(x))^3 + 3T''(z(x))z'(x)z''(x) + T'(z(x))z'''(x).\end{aligned}$$

Hence, $\lim_{x \downarrow \frac{1}{3}} \Pi'(x) = \frac{3}{4}$, $\lim_{x \uparrow 1} \Pi'(x) = 1$, $\lim_{x \downarrow \frac{1}{3}} \Pi''(x) = \frac{9}{8}$, and $\lim_{x \uparrow 1} \Pi''(x) = -\infty$. Therefore, Π' exists and continuous for $x \in [\frac{1}{3}, 1]$. Π'' exists and continuous for $x \in [\frac{1}{3}, 1)$, but $\Pi''(1-) = -\infty \neq 0 = \Pi''(1+)$.

Finally, observe that

$$\begin{aligned}\Pi'''(x) &= T'''(z(x))(z'(x))^3 + 3T''(z(x))z'(x)z''(x) + T'(z(x))z'''(x) \\ &= \frac{-27(2 - z(x)^2)^4 z(x)^2}{32(z(x)^4 - 6z(x)^2 + 6z(x))^5} [1 + \\ &\quad (1 - z(x))(71(1 - 2z(x))^2 + 67z(x) + 579z(x)^2(1 - z(x))^2 \\ &\quad + z(x)^3(401 - 217z(x)^2 - 9z(x)^3 - z(x)^5 - z(x)^6) + 44z(x)^4 + 21z(x)^7)] \\ &< 0,\end{aligned}$$

which means that there exists a unique $\hat{x} \in (\frac{1}{3}, 1)$ such that

$$\Pi''(x) \begin{matrix} \leq \\ \geq \end{matrix} 0 \Leftrightarrow x \begin{matrix} \geq \\ \leq \end{matrix} \hat{x}.$$

A numerical search suggests $\hat{x} \simeq 0.87$.

J Proof of Proposition 3

Because Π' exists everywhere and is absolutely continuous, we can rewrite the objective function as $\int_0^1 \Pi''(x)H(x)dx$ (plus a constant). Because $\Pi(x)$ is convex for $x < \hat{x}$ and concave for $x > \hat{x}$ (hence, we have $K = 2$), our “first-step” problem is to find the optimal H such that $H(\hat{x}) = h_1$, where h_1 is arbitrarily fixed (up to feasibility, which implies $h_1 \leq H_0(\hat{x}) (= \frac{\hat{x}^2}{2})$ and $h_1 \geq \hat{x} - \frac{1}{2}$). The “second-step” problem is optimization with respect to h_1 .

Applying Theorem 5, the solution of the first-step problem, denoted by \tilde{H} , is given as follows. Letting $\tilde{x} = \hat{x} - \sqrt{\hat{x}^2 - 2h_1}$,

$$\tilde{H}(x) = \begin{cases} H_0(x) (= \frac{x^2}{2}) & \text{if } x < \tilde{x}, \\ \max\{\tilde{x}(x - \hat{x}) + h_1, x - \frac{1}{2}\} & \text{if } x > \tilde{x}. \end{cases}$$

Geometrically, \tilde{x} is the point at which H_0 is supported by the affine function that (i) supports H_0 on the left of $x = \hat{x}$ and (ii) takes value h_1 at $x = \hat{x}$.

The second step would be then to maximize the first-step value function with respect to h_1 , and the corresponding \tilde{x} would be x^* in the statement.

As an example, assume that $a \leq -1$ and $b \geq 3$. In what follows, however,

instead of working on the first-step value function, we rather characterize x^* directly in the original problem. Let $\hat{\Pi}(x)$ be the expected surplus if the principal adopts the upper-censorship policy that (i) truthfully discloses the realized θ if $\theta < x$, and (ii) discloses nothing if $\theta > x$. By the analysis above, we know that the optimal value of x , x^* , is not greater than \hat{x} . Thus,

$$\hat{\Pi}(x) = \begin{cases} \frac{x+b}{2} \frac{b-x}{b-a} & \text{if } x \leq -1, \\ \frac{1}{b-a} \left[\int_{-1}^x \frac{9}{64} (1+y)^3 dy + \frac{x+b}{2} (b-x) \right] & \text{if } x \in (-1, \hat{x}], \end{cases}$$

which implies that x^* is not lower than -1 .

Therefore, our problem reduces to the following:

$$\max_{x \in [-1, \hat{x}]} \frac{1}{b-a} \left[\int_{-1}^x \frac{9}{64} (1+y)^3 dy + \frac{x+b}{2} (b-x) \right],$$

and we obtain $x^* = \frac{1}{3}$.

References

- BERGEMANN, D., A. BONATTI, AND A. SMOLIN (2017): “The Design and Price of Information,” Cowles Foundation Discussion Paper No. 2049R.
- BERGEMANN, D. AND S. MORRIS (2013): “Robust Predictions in Games With Incomplete Information,” *Econometrica*, 81, 1251–1308.
- BERGEMANN, D. AND M. PESENDORFER (2007): “Information structures in optimal auctions,” *Journal of Economic Theory*, 137, 580–609.

- BERGEMANN, D. AND A. WAMBACH (2015): “Sequential Information Disclosure in Auctions,” *Journal of Economic Theory*, 159, 1074–1095.
- BORDER, K. C. (2016): “Notes on Integration by Parts,” Retrieved from <http://people.hss.caltech.edu/kcb/Notes/IntegrationbyParts.pdf>.
- CRAWFORD, V. P. AND J. SOBEL (1982): “Strategic Information Transmission,” *Econometrica*, 50, 1431–1451.
- DRANOVE, D. AND G. Z. JIN (2010): “Quality Disclosure and Certification: Theory and Practice,” *Journal of Economic Literature*, 48, 935–963.
- DWORCZAK, P. AND G. MARTINI (2017): “The Simple Economics of Optimal Persuasion,” Mimeo, Stanford University.
- ESÓ, P. AND B. SZENTES (2007): “Optimal Information Disclosure in Auctions and the Handicap Auction,” *Review of Economic Studies*, 74, 705–731.
- (2017): “Dynamic Contracting: An Irrelevance Result,” *Theoretical Economics*, 12, 109–139.
- FADEN, A. M. (1985): “The Existence of Regular Conditional Probabilities: Necessary and Sufficient Conditions,” *The Annals of Probability*, 13, 288–298.
- GENTZKOW, M. AND E. KAMENICA (2015): “A Rothschild-Stiglitz Approach to Bayesian Persuasion,” *American Economic Review, Papers and Proceedings*, 106 (5).

- GOMES, R. AND A. PAVAN (2016): “Many-to-many matching and price discrimination,” *Theoretical Economics*, 11, 10051052.
- JEON, D., B. KIM, AND D. MENICUCCI (2016): “A Seconddegree Price Discrimination by a Twosided Monopoly Platform,” TSE Working Paper.
- KAMENICA, E. AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 2590–2615.
- KOESSLER, F. AND V. SKRETA (2016): “Informed Seller with Taste Heterogeneity,” *Journal of Economic Theory*, 165, 456–471.
- KOLOTILIN, A. (2016): “Optimal Information Disclosure: A Linear Programming Approach,” Mimeo, University of New South Wales.
- KOLOTILIN, A., M. LI, T. MYLOVANOV, AND A. ZAPECHELNYUK (2015): “Persuasion of a Privately Informed Receiver,” Working paper.
- LAFFONT, J. J. AND J. TIROLE (1993): *A Theory of Incentives in Procurement and Regulation*, Cambridge, MA: MIT Press.
- LI, H. AND X. SHI (2015): “Discriminatory Information Disclosure,” Working Paper.
- LIZZERI, A. (1999): “Information Revelation and Certification Intermediaries,” *The RAND Journal of Economics*, 30, 214–231.
- MACHINA, M. AND J. PRATT (1997): “Increasing Risk: Some Direct Constructions,” *Journal of Risk and Uncertainty*, 14, 103–127.

- MASKIN, E. AND J. TIROLE (1990): “The Principal-Agent Relationship with an Informed Principal: The Case of Private Values,” *Econometrica*, 58, 379–409.
- (1992): “The Principal-Agent Relationship with an Informed Principal, II: Common Values,” *Econometrica*, 60, 1–42.
- MILGROM, P. AND R. J. WEBER (1982): “The Value of Information in A Sealed-Bid Auction,” *Journal of Mathematical Economics*, 10, 105–114.
- MUSSA, M. AND S. ROSEN (1978): “Monopoly and product quality,” *Journal of Economic Theory*, 18, 301 – 317.
- MYERSON, R. B. (1981): “Optimal Auction Design,” *Mathematics of Operation Research*, 6, 58–73.
- (1983): “Mechanism Design by An Informed Principal,” *Econometrica*, 51, 1767–1797.
- MYERSON, R. B. AND M. A. SATTERTHWAITTE (1983): “Efficient Mechanisms for Bilateral Trading,” *Journal of Economic Theory*, 29, 265–281.
- MYLOVANOV, T. AND T. TRÖGER (2012): “Informed-principal problems in environments with generalized private values,” *Theoretical Economics*, 7, 465–488.
- OTTAVIANI, M. AND A. PRAT (2001): “The Value of Public Information in Monopoly,” *Econometrica*, 69, 1673–1683.

- PERLMAN, M. D. (1974): “Jensen’s inequality for a convex vector-valued function on an infinite-dimensional space,” *Journal of Multivariate Analysis*, 4, 52–65.
- RAYO, L. AND I. SEGAL (2010): “Optimal Information Disclosure,” *Journal of Political Economy*, 118, 949–987.
- ROTHSCHILD, M. AND J. STIGLITZ (1970): “Increasing Risk: I. A definition,” *Journal of Economic Theory*, 2, 225–243.
- SCHMIDT, K. D. (2014): “On inequalities for moments and the covariance of monotone functions,” *Insurance: Mathematics and Economics*, 55, 91 – 95.
- SKRETA, V. (2011): “On the informed seller problem: optimal information disclosure,” *Review of Economic Design*, 15, 1–36.
- STAHL, K. AND R. STRAUZ (2017): “Certification and Market Transparency,” *forthcoming, Review of Economic Studies*.
- YAMASHITA, T. (2016): “Revenue Guarantees in Auctions with a (Correlated) Common Prior and Information Acquisition,” TSE Working Paper.
- ZHU, S. (2017): “Private Disclosure with Multiple Agents,” Working paper.

Supplementary Materials

(not for publication)

K Two remarks on the full-disclosure results

Remark 1. Regarding Theorem 1–3, the assumptions of linearity and private values can be weakened to some extent (though not completely dispensable). We briefly see that, in the multi-agent framework studied in Section 3.1, the same result is obtained if either (i) $u_i(\chi, \theta, t) = z_i(q, t_{-i})y_i(\theta, t_i) - p_i$, (ii) $u_i(\chi, \theta, t) = z_i(q, \theta, t_{-i})y_i(t_i) - p_i$, or (iii) $u_i(\chi, \theta, t) = z_i(q, t)y_i(\theta) - p_i$. Case (i) is a direct generalization of the one used in the main text, allowing for externalities in q and interdependence, as long as those are multiplicably separated from $y_i(\theta, t_i)$. Either t_i or θ (but not both) can be moved into z_i instead of y_i , leading to Case (ii) or (iii).

For (i), agent i 's incentive compatibility is equivalent to combination of (a) (interim) monotonicity:

$$E_{t_{-i}}[z_i(q(t_i, t_{-i}), t_{-i})] \text{ is nondecreasing in } t_i,$$

and (b) envelope formula:

$$\begin{aligned} & E_{t_{-i}, \theta}[z_i(q(t), t_{-i})y_i(\theta, t_i) - p_i(t)] \\ &= E_{t_{-i}, \theta}[z_i(q(0, t_{-i}), t_{-i})y_i(\theta, t_i) - p_i(0, t_{-i})] \\ &+ \int_0^{t_i} E_{t_{-i}}[z_i(q(0, t_{-i}), t_{-i}) \frac{\partial y_i}{\partial t_i}(\theta, \tilde{t}_i)] d\tilde{t}_i, \end{aligned}$$

where expectation with respect to θ is based on a posterior Ψ_0 .

Therefore, the set of implementable q does not vary with Ψ_0 and the information rent is linear in Ψ_0 , which implies optimality of full disclosure. We omit the other cases (ii) and (iii).

Remark 2. Although this paper only considers Bayesian incentive compatibility, the same sort of exercise is possible for other solution concepts, and the optimal disclosure strategy naturally varies with the underlying solution concept. For example, we may consider ex post incentive compatibility with respect to the agents' types, i.e., each agent i finds truth-telling optimal regardless of the other agents' type realization, t_{-i} . On the other hand, we still keep the assumption that the principal has a full control over disclosure of θ and can design the agents' posterior about θ .⁴⁰

To provide a more concrete idea, consider an interdependent-value auction environment where each i 's utility is $q_i v_i(\theta, t) - p_i$ with $\frac{\partial v_i}{\partial t_i} > 0$. Then, ex post incentive compatibility (with respect to t) means, for each i, t_i, t'_i, t_{-i} ,

$$q_i(t_i, t_{-i})E_\theta[v_i(\theta, t_i, t_{-i})] - p_i(t_i, t_{-i}) \geq q_i(t'_i, t_{-i})E_\theta[v_i(\theta, t_i, t_{-i})] - p_i(t'_i, t_{-i}),$$

where expectation with respect to θ is based on a posterior Ψ_0 .

⁴⁰Although this “double standard” in the agents' informational assumption may look strange, there exist some situations where this assumption is relevant. For example, imagine the same Bayesian environment as in the main model (both in terms of θ and t), but assume that each agent i can engage in covert information acquisition about the other agents types t_{-i} . Yamashita (2016) shows that, if the principal does not know the agents' information acquisition technology and hence desires to maximize his worst-case (or “guaranteed”) expected payoff across all possible information structures, then the optimal mechanism must satisfy ex post incentive compatibility with respect to t_{-i} .

Again, this is equivalent to combination of (a) ex post monotonicity:

$$q_i(t_i, t_{-i}) \text{ is nondecreasing in } t_i,$$

and (b) envelope formula:

$$\begin{aligned} & q_i(t) E_\theta[v_i(\theta, t)] - p_i(t) \\ &= q_i(0, t_{-i}) E_\theta[v_i(\theta, 0, t_{-i})] - p_i(0, t_{-i}) \\ &+ \int_0^{t_i} q_i(\tilde{t}_i, t_{-i}) E_\theta\left[\frac{\partial v_i(\theta, \tilde{t}_i, t_{-i})}{\partial t_i}\right] d\tilde{t}_i. \end{aligned}$$

Therefore, the set of implementable q does not vary with Ψ_0 and the information rent is linear in Ψ_0 , which implies optimality of full disclosure.

L Proof of Theorem 5 (general case)

Even without continuity or differentiability of Π or Π' , the problem can be rewritten by changing the choice variable from F (a cdf) to H (an integrate of a cdf), but with a more complicated expression of the objective function.

First, we provide a formal proof for a version of integration by parts as an application of Fubini theorem.⁴¹

Lemma 13. For $a, b \in \mathbb{R}$ with $a < b$, let $h : (a, b) \rightarrow \mathbb{R}$ be absolutely continuous with derivative h' (which exists almost everywhere), and $F :$

⁴¹The proof is based on Border (2016).

$\mathbb{R} \rightarrow [0, 1]$ be a cdf. Then,

$$\int_{x \in (a,b)} h(x) dF(x) = h(b-)F(b-) - h(a+)F(a) - \int_{x \in (a,b)} h'(x)F(x) dx.$$

Proof. Consider two independent random variables $x \sim U(a, b)$ and $y \sim F$. Let μ denote the joint distribution for (x, y) . Let $\phi : (a, b)^2 \rightarrow \mathbb{R}$ be such that $\phi(x, y) = 1\{x \leq y\}h'(x)$. Then,

$$\int_{(x,y) \in (a,b)^2} |\phi(x, y)| d\mu(x, y) \leq \int_{(x,y) \in (a,b)^2} |h'(x)| d\mu(x, y) \leq \int_{x \in (a,b)} |h'(x)| dx < \infty,$$

where the last inequality is because h' is integrable. This validates an application of Fubini theorem for ϕ , which implies

$$\int_{x \in (a,b)} \int_{y \in [x,b]} h'(x) dF(y) dx = \int_{y \in (a,b)} \int_{x \in (a,y]} h'(x) dx dF(y).$$

The left-hand side equals

$$\int_{x \in (a,b)} h'(x)(F(b-) - F(x)) dx = (h(b-) - h(a+))F(b-) - \int_{x \in (a,b)} h'(x)F(x) dx,$$

and the right-hand side equals

$$\int_{y \in (a,b)} (h(y) - h(a+)) dF(y) = \int_{y \in (a,b)} h(y) dF(y) - h(a+)(F(b-) - F(a)).$$

Therefore,

$$\int_{y \in (a,b)} h(y) dF(y) = h(b-)F(b-) - h(a+)F(a) - \int_{x \in (a,b)} h'(x)F(x) dx.$$

□

Let $H_0(x) = \int_{\underline{\theta}}^x F_0(y)dy$ for each x . The following lemma corresponds to Lemma 11 in the main body.

Lemma 14.

$$\begin{aligned} \Pi^* = \sup_{H: [\underline{\theta}, \bar{\theta}] \rightarrow \mathbb{R}} & \quad \Pi(\bar{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)H(x)dx + \sum_{k=1}^K \left[H'(x_{k-1+})(\Pi(x_{k-1}) - \Pi(x_{k-1+})) \right. \\ & \quad \left. + H'(x_k-)(\Pi(x_k-) - \Pi(x_k)) - H(x_k)\Pi'(x_k-) + H(x_{k-1})\Pi'(x_{k-1+}) \right] \\ \text{sub. to} & \quad H(x) \leq H_0(x), \quad \forall x, \\ & \quad H(\underline{\theta}) = H_0(\underline{\theta}) = 0, \quad H(\bar{\theta}) = H_0(\bar{\theta}), \\ & \quad H \text{ is convex, } \frac{H(x) - H(x')}{x - x'} \in [0, 1], \quad \forall x > x'. \end{aligned}$$

Proof. It suffices to show that the new objective is the same as that of the

original problem with $H(\cdot) = \int_{\underline{\theta}}^{(\cdot)} F(x)dx$ and $H_0(\cdot) = \int_{\underline{\theta}}^{(\cdot)} F_0(x)dx$. Indeed:

$$\begin{aligned}
& \int_{x \in [\underline{\theta}, \bar{\theta}]} \Pi(x) dF(x) \\
&= \sum_{k=1}^K \left(\int_{x \in (x_{k-1}, x_k)} \Pi(x) dF(x) + \Pi(x_k)(F(x_k) - F(x_{k-})) \right) \\
&= \sum_{k=1}^K \left(\Pi(x_{k-})F(x_{k-}) - \Pi(x_{k-1}+)F(x_{k-1}) \right. \\
&\quad \left. + \Pi(x_k)(F(x_k) - F(x_{k-})) - \int_{x \in (x_{k-1}, x_k)} \Pi'(x)F(x)dx \right) \\
&= \Pi(\bar{\theta}) + \sum_{k=1}^K \left(F(x_{k-1})(\Pi(x_{k-1}) - \Pi(x_{k-1}+)) + F(x_{k-})(\Pi(x_{k-}) - \Pi(x_k)) \right. \\
&\quad \left. - \Pi'(x_{k-1}+)(H(x_k) - H(x_{k-1})) - (\Pi'(x_{k-}) - \Pi'(x_{k-1}+))H(x_k) \right. \\
&\quad \left. + \int_{x \in (x_{k-1}, x_k)} \Pi''(x)H(x)dx \right) \\
&= \Pi(\bar{\theta}) + \int_0^1 \Pi''(x)H(x)dx + \sum_{k=1}^K \left[H'(x_{k-1}+)(\Pi(x_{k-1}) - \Pi(x_{k-1}+)) \right. \\
&\quad \left. + H'(x_{k-})(\Pi(x_{k-}) - \Pi(x_k)) - H(x_k)\Pi'(x_{k-}) + H(x_{k-1})\Pi'(x_{k-1}+) \right].
\end{aligned}$$

□

Note that upper-semi-continuity of Π implies $\Pi(x_{k-1}) - \Pi(x_{k-1}+) \geq 0$ and $\Pi(x_{k-}) - \Pi(x_k) \leq 0$.

We can solve this problem in two steps. First, fix arbitrarily $\{h_k, f_k^-, f_k^+\}_{k=0}^K$ such that there exists a feasible \hat{H} with $\hat{H}(x_k) = h_k$, $\hat{H}'(x_{k-}) = f_k^-$, and $\hat{H}'(x_{k+}) = f_k^+$, for each k . As the first step, we consider a constrained problem where the principal chooses H that is feasible and satisfies, for $k = 0, \dots, K$, (i) $H(x_k) = h_k$, $H'(x_{k-}) \leq f_k^-$, and $H'(x_{k-1}+) \geq f_{k-1}^+$ if Π' is

nondecreasing on the interval (x_{k-1}, x_k) , and (ii) $H(x_k) = h_k$, $H'(x_k-) = f_k^-$, and $H'(x_{k-1}+) = f_{k-1}^+$ if Π' is nonincreasing on the interval (x_{k-1}, x_k) . Then, as the second step, we optimize the first-step value function with respect to $\{h_k, f_k^-, f_k^+\}_{k=0}^K$. In what follows, we only consider the first step, because the rest of the proof is the same as in the twice-continuously differentiable case demonstrated in the main text.

For the first-step problem, we define the following $\tilde{H} : [\underline{\theta}, \bar{\theta}] \rightarrow \mathbb{R}$. First, $\tilde{H}(x_k) = h_k$ for all k . Second, for each interval (x_{k-1}, x_k) where Π' is nondecreasing, \tilde{H} coincides with the largest convex function that is below both H_0 and affine function:

$$\bar{\xi}^k : x \mapsto \frac{h_k - h_{k-1}}{x_k - x_{k-1}}(x - x_{k-1}) + h(x_{k-1}).$$

Finally, for each interval (x_{k-1}, x_k) where Π' is nonincreasing, for each $x \in (x_k, x_{k+1})$, $\tilde{H}(x) = \underline{\xi}^k(x)$, where:

$$\underline{\xi}^k : x \mapsto \max\{h_{k+1} - f_{k+1}^+(x_{k+1} - x), h_k + f_k^-(x - x_k)\}.$$

Lemma 15. \tilde{H} is feasible. Moreover, for any feasible H such that $H(x_k) = h_k$ for $k = 0, \dots, K$, we have

$$\int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)H(x)dx \leq \int_{\underline{\theta}}^{\bar{\theta}} \Pi''(x)\tilde{H}(x)dx,$$

that is, \tilde{H} solves the first-step problem.

Proof. First, we show that \tilde{H} is feasible.

Claim 1: $\tilde{H} \leq H_0$.

It is obvious by construction for $x \in [x_{k-1}, x_k]$ on which Π is convex. For $x \in [x_{k-1}, x_k]$ on which Π is concave, we have

$$H_0(x) \geq \hat{H}(x) \geq \underline{\xi}^k(x) = \tilde{H}(x).$$

Claim 2: $\tilde{H}(\underline{\theta}) = H_0(\underline{\theta})$ and $\tilde{H}(\bar{\theta}) = H_0(\bar{\theta})$.

We have $\tilde{H}(0) = h_0 = H_0(0)$ and $\tilde{H}(1) = h_K = H_0(1)$.

Claim 3: $\frac{\tilde{H}(x) - \tilde{H}(x')}{x - x'} \in [0, 1]$ for $x > x'$.

Fix x, x' with $\underline{\theta} \leq x' < x \leq \bar{\theta}$. By construction, clearly we have $\frac{\tilde{H}(x) - \tilde{H}(x')}{x - x'} \geq 0$. In order to show the other inequality, let $k \in \{1, \dots, K\}$ be such that $x \leq x_k$. Then:

$$\frac{\tilde{H}(x) - \tilde{H}(x')}{x - x'} \leq f_K^+ \leq 1.$$

Claim 4: \tilde{H} is convex.

For each open subinterval (x_{k-1}, x_k) , \tilde{H} is clearly convex. Thus, we complete the proof of the claim by showing that \tilde{H} is convex around each boundary point x_k , $k = 1, \dots, K-1$. This is indeed true, because $H'(x_k+) \geq f_k^+ \geq f_k^- \geq H'(x_k-)$.

Next, we show that \tilde{H} is optimal. If there exists H that is feasible in the first-step problem and strictly better than \tilde{H} , there exist k and $x \in (x_{k-1}, x_k)$ such that, either (i) Π is convex on (x_{k-1}, x_k) and $H(x) > \tilde{H}(x)$, or (ii) Π is

concave on (x_{k-1}, x_k) and $H(x) < \tilde{H}(x)$.

If (i) holds with convex H with $H \leq H_0$, then $H \leq \bar{\xi}^k$ is violated, because \tilde{H} is the largest convex function that is below both H_0 and $\bar{\xi}^k$. Then, we have either $H(x_{k-1}) \neq h_{k-1}$ or $H(x_k) \neq h_k$, which contradicts that H is feasible in the first-step problem.

If (ii) holds with convex H , then $H'(x_{k-1}+) < f_{k-1}^-$ or $H'(x_k-) > f_k^+$. Again, it contradicts that H is feasible in the first-step problem.

In conclusion, \tilde{H} is optimal in the first-step problem. \square

The lemma implies that, for any $\{h_k, f_k^-, f_k^+\}_{k=0}^K$ fixed in a feasible way, the solution to the first-step problem satisfies the following property: there exist $L \leq K$ and disjoint open sub-intervals $\{I_l\}_{l=1}^L$ such that (i) $\tilde{H}(x)$ is piecewise-linear on each I_l and $\tilde{H}(x) < H_0(x)$, and (ii) $\tilde{H}(x) = H_0(x)$ otherwise. Thus, for the optimal F^* in the original problem, there exist $L \leq K$ and disjoint open sub-intervals $\{I_l\}_{l=1}^L$ such that (i) $F^*(x)$ is piecewise-constant and $\int_{\underline{\theta}}^x F^*(y)dy < \int_{\underline{\theta}}^x F_0(y)dy$ for each l and each $x \in I_l$, and (ii) $F^*(x) = F_0(x)$ and $\int_{\underline{\theta}}^x F^*(y)dy = \int_{\underline{\theta}}^x F_0(y)dy$ otherwise.

Moreover, the total number of kinks in \tilde{H} is not greater than $2K$, because, as is obvious in the construction of \tilde{H} , \tilde{H} can kink at most once on each (x_{k-1}, x_k) , and may kink on each boundary point x_k . Therefore, the total number of jumps in F^* is not greater than $2K$ either (and hence is bounded).

The rest of the proof is the same as in the previous case with twice-continuously differentiable Π , and hence is omitted.

M Example 2: Employment with correlated information

We consider a single-agent environment where the agent's information t and the principal's information θ are correlated. As we show in Section 2.4, if $v(\theta, t)$ is increasing in both of the arguments and (θ, t) are affiliated, then full disclosure is optimal. If they are negatively correlated — or equivalently up to relabeling, if $v(\theta, t)$ is increasing in $(\theta, -t)$ (i.e., decreasing in t) and (θ, t) are affiliated — then full disclosure can be suboptimal. In such a case, Theorem 5 is proved to be useful in identifying the optimal disclosure policy.

Imagine that the agent is a potential employee of a firm owned by the principal. If the agent is hired, the principal earns 1, while the agent incurs cost $1 - s$, where $s \in \{0, 1\}$ may be interpreted as an unobserved match value between the principal and the agent. The (marginal) probability for $s = 1$ is $\frac{3}{4}$. The principal has a noisy signal θ about s , which is distributed according to density $2(1 - \theta)$ conditional on $s = 0$, and according to density $\frac{2}{3}(1 + \theta)$ conditional on $s = 1$. Then we have $E[s|\theta] = \Pr[s = 1|\theta] = \frac{1+\theta}{2}$, and θ is marginally distributed according to a uniform distribution over $[0, 1]$.

If not hired, the agent enjoys the outside option $\frac{1+t}{6}$, where $t \in \{1, 2, 3\}$ is the agent's private information and correlated with s as follows:⁴²

⁴²Rigorously, this example does not fit into the previous analysis because t is discrete. However, the difference is not essential. Similar results hold by replacing the pdf of t by its pmf.

$\Pr(t, s)$	$s = 0$	$s = 1$
$t = 1$	$\frac{1}{15}$	$\frac{1}{10}$
$t = 2$	$\frac{3}{20}$	$\frac{9}{20}$
$t = 3$	$\frac{1}{30}$	$\frac{1}{5}$

Note that (s, t) are affiliated, and in particular, $E[s|t]$ is strictly increasing in t . Imagine that an employee with a higher match value with an employer is more likely to have a high match value with another employer too. The affiliation of s and t captures such positive relationship of the agent's match values with the principal (s) and with the other employers (summarized by the agent's outside option t). Conditional on s , we assume that (θ, t) are independent.

An allocation is given by a pair (q, p) , where $q \in [0, 1]$ represents the probability of hiring, and p represents the payment from the principal to the agent. Then, the principal's payoff is $q - p$, and the agent's payoff (net the reservation payoff t) is $p - q(1 - s + t)$, where $1 - s$ denotes the agent's (material) cost, and t denotes his opportunity cost. Observe that the agent's valuation for employment is increasing in s but decreasing in t . Equivalently, if we interpret $\tau = -t$ as the agent's private information, then the agent's valuation for employment is increasing in (s, τ) , while (s, τ) are negatively correlated. Thus, as observed previously, the optimal disclosure policy may not be full-disclosure.

Given any message from the principal which implies that the (t -unconditional) posterior for $s = 1$ is x , the agent's posterior for $s = 1$ given $(x$ and) his type

t is as follows:

$$E[s|x, t] = \begin{cases} \frac{1+x}{3-x} & \text{if } t = 1 \\ \frac{1+x}{2} & \text{if } t = 2 \\ \frac{2+2x}{3+x} & \text{if } t = 3 \end{cases}$$

and thus, the payment necessary to hire the agent with type t is, conditional on x ,

$$1 - E[s|x, t] + \frac{1+t}{6} = \begin{cases} \frac{4}{3} - \frac{1+x}{3-x} & \text{if } t = 1 \\ \frac{3}{2} - \frac{1+x}{2} & \text{if } t = 2 \\ \frac{5}{3} - \frac{2+2x}{3+x} & \text{if } t = 3. \end{cases}$$

Therefore, the optimal mechanism for each x is given in the following lemma. The proof is omitted.

Lemma 16. Given x , the optimal mechanism $(q_x(\cdot), p_x(\cdot))$ satisfies the following: letting $x^* = \frac{9-\sqrt{65}}{2} (\simeq 0.47)$,

- (i) for $x < x^*$, $q_x(t) = 1$ and $p_x(t) = \frac{5}{3} - \frac{2+2x}{3+x}$ for all t , yielding the expected payoff $\frac{4x}{9+3x}$ for the principal (conditional on x);
- (ii) for $x > x^*$,

$$(q_x(t), p_x(t)) = \begin{cases} (1, \frac{3}{2} - x) & \text{if } t = 1, 2 \\ (0, 0) & \text{if } t = 3, \end{cases}$$

yielding the expected payoff $\frac{x(12-x)}{30}$ for the principal (conditional on x).

Therefore, it is concave for $x < x^*$, and again concave for $x > x^*$. At x^* , it is kinked “upward”, making it not globally concave.

To provide a more economic intuition, notice that the concavity of the principal’s expected payoff in each of the two regions above means that the principal would prefer less disclosure if the mechanism was *fixed*. This is because, as in Section 2.4, in case (θ, t) are negatively correlated (and v_1 is increasing in (θ, t)), or equivalently up to relabelling, in case (θ, t) are positively correlated and v_1 is increasing in $(\theta, -t)$, then more disclosure means higher expected information rent.

On the other hand, because the optimal mechanism changes with x , this flexibility makes the principal favor more disclosure. Indeed, the kink in the above lemma is because the optimal mechanism drastically changes at this point. In this sense, mechanism design plays a crucial role in this problem.

Therefore, in the optimal policy, information is “minimally disclosed” so that such flexibility in the mechanism choice is attained, and in each region where the same mechanism is optimal, no information is to be disclosed.

Proposition 4. There exist $(y^*, w^*; z_1^*, z_2^*, p^*) \simeq (0.34, 0.30, 0.35, 0.67, 0.54)$ such that the following binary lower truncation policy (M, G) is optimal: $M = [0, 1]$, and

- for $\theta \in (0, y^*)$,

$$G(\{z_1^*\}|\theta) = 1,$$

- for $\theta \in (y_1^*, 1)$,

$$\begin{aligned} G(\{z_1^*\}|\theta) &= w^*, \\ G(\{z_2^*\}|\theta) &= 1 - w^*, \end{aligned}$$

which implements F^* such that

$$F^*(x) = \begin{cases} 0 & \text{if } x < z_1^*, \\ p^* & \text{if } x \in [z_1^*, z_2^*), \\ 1 & \text{if } x \geq z_2^*. \end{cases}$$

Proof. First, applying Theorem 5, the principal's objective is rewritten as follows:

$$\frac{11}{30} + \Pi'(x^*)H(x^*) + \int_0^1 \Pi''(x)H(x)dx.$$

Because $\Pi''(x)$ is negative for any $x \in (0, x^*), (x^*, 1)$, there exist $z_1 \in [0, x^*]$, $z_2 \in [x^*, 1]$ such that the optimal H^* is a (continuous and convex) piecewise-linear function with kinks potentially at z_1 and z_2 .⁴³ That is, there

⁴³Another potential kink is at x^* , but we can easily see that it is suboptimal to have a kink at x^* . Indeed, if any feasible H admits a non-trivial kink exists at x^* , then we have $H'(x^+)(x - x^*) + H(x^*) < 0 = H'(x^*)(x - x^*) + H(x^*)$ at $x = z_1$. In this case, consider an alternative \tilde{H} such that

$$\tilde{H}(x) = \begin{cases} \max\{0, H'(x^+)(x - x^*) + H(x^*)\} < H(x) & \text{if } x \in (z_1, x^*), \\ H(x) & \text{if } x \notin (z_1, x^*), \end{cases}$$

which is strictly better than H .

exists $p \in [0, 1]$ such that

$$H^*(x) = \max\{0, p(x - z_1), x - (1 - p)z_2 - pz_1\},$$

where feasibility of H^* implies $p(x - z_1) \leq \frac{x^2}{2}$ for all x , and $1 - (1 - p)z_2 - pz_1 = \frac{1}{2}$. Or equivalently, $p \leq 2z_1$ and $(1 - p)z_2 + pz_1 = \frac{1}{2}$.

The optimal $F^* \in \Delta([0, 1])$ is hence written as

$$F^*(x) = \begin{cases} 0 & \text{if } x < z_1 \\ p & \text{if } x \in [z_1, z_2) \\ 1 & \text{if } x \geq z_2. \end{cases}$$

Therefore, our (originally infinite-dimensional) problem becomes the following two-dimensional problem:

$$\begin{aligned} \Pi^* &= \max_{z_1, z_2, p} \frac{4z_1}{9 + 3z_1}p + \frac{z_2(12 - z_2)}{30}(1 - p) \\ \text{sub. to} & \quad pz_1 + (1 - p)z_2 = \frac{1}{2}, \quad z_1 \leq x^* \leq z_2, \quad p \leq 2z_1. \end{aligned}$$

We first ignore all the inequality constraints and solve that relaxed problem. Then, we confirm that the solution satisfies the ignored inequalities, and hence it is the solution to the problem above.

Let (z_1^*, z_2^*, p^*) denote the solution to the relaxed problem. First, slightly perturbing the solution by (dz_1, dz_2, dp) where $p^*dz_1 + (1 - p^*)dz_2 = dp = 0$ must not improve the objective, and thus:

$$\frac{4}{(3 + z_1)^2} = \frac{6 - z_2^*}{15}.$$

Second, slightly perturbing the solution by (dz_1, dz_2, dp) where $(z_1^* - z_2^*)dp + p^*dz_1 = dz_2 = 0$ must not improve the objective, and thus:

$$\frac{4z_1}{9 + 3z_1} - \frac{z_2(12 - z_2)}{30} + (z_2^* - z_1^*)\frac{4}{(3 + z_1)^2} = 0.$$

Applying these two necessary conditions and the constraint that $p z_1 + (1 - p)z_2 = \frac{1}{2}$, the (relaxed) problem is a one-dimensional optimization, which is solved in a standard manner. The solution of this relaxed problem is $(z_1^*, z_2^*, p^*) \simeq (0.35, 0.67, 0.54)$, and it is easy to see that all the inequality constraints in the original problem are satisfied with strict inequalities. Therefore, this is also the solution to the original problem. The value of the objective is $\Pi^* \simeq 0.19$.

Finally, the optimal binary lower-truncation policy is identified by solving for (y, w) so that it induces H^* . We omit this step which is straightforward.

□