Information disclosure and full surplus extraction in mechanism design

Daniel Krähmer*

October 11, 2019

Abstract

I study mechanism design settings with quasi-linear utility where the principal can provide agents with additional private information about their valuations beyond the private information they hold at the outset. I demonstrate that the principal can design information and a mechanism so as to implement the same outcome as if the additional information was publicly known. The key idea is that the principal secretely randomizes over information structures which allows her to cross-check the consistency of agents' reports. As an implication, the principal can fully extract the first-best surplus in a large class of cases.

Keywords: information design, mechanism design, quasi-linear utility, rent extraction
JEL codes: D82, H57

*University Bonn, Department of Economics, Institute for Microeconomics, Adenauer Allee 24-42, D-53113 Bonn (Germany), kraehmer@hcm.uni-bonn.de. I thank two anonymous referees and an associate editor for useful comments. I also thank Paul Heidhues, Stephan Lauermann, Benny Moldovanu, Stephen Morris, Martin Pollrich, Ludovic Renou, David Rodina, Alex Smolin, Nicolas Schutz, Roland Strausz, Balazs Szentes, and Tymon Tatur for inspiring discussions. Thanks also go to seminar audiences at Berlin, Bern, Bonn, LSE, UCL. I am grateful to the UCL Economics department for their hospitality. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 and Germany's Excellence Strategy EXC 2126/1âĂŞ 390838866 is gratefully acknowledged.

1 Introduction

In many mechanism design settings, a principal can disclose additional information to agents prior to implementing an allocation. For example, sellers often offer interested buyers the possibility to test or inspect their products prior to purchase, or grant withdrawal periods to try out an order for a while. Likewise, an auctioneer of an oil well may allow bidders to conduct test drills, or procurement agencies give contractors additional information about the costs of a project.

In this paper, I study mechanism design problems with quasi-linear utility where the principal can, next to a mechanism, design and disclose additional information that affects agents' valuations. Similar to Esö and Szentes (2007a,b) and Li and Shi (2017), I focus on situations in which the information the designer provides becomes an agent's private information, and, moreover, agents already at the outset have some (imperfect) private information about their valuations.

In this context, the literature has shown that agents can secure information rents leading to distortions and welfare losses.[1] The main result of this paper shows that, to the contrary, in a large class of cases the principal can design information and a mechanism so as to fully extract the complete information first-best surplus. Specifically, full surplus extraction is possible when the agents' initial information is payoff-irrelevant, or is appropriately correlated with the information the principal can provide. These cases cover the "static" mechanism design setting when agents perfectly know their valuations at the outset, and the principal can design and privately disclose signals correlated with the agents' private information. Hence, information design when coupled with mechanism design renders the privacy of both ex ante and ex post information irrelevant.

The intuitive idea behind these results is that the principal can reduce information rents by concealing the information structure she uses to inform agents. This allows her to elicit agents' private information by cross-checking whether the information they report is consistent with the true information structure. Thus, my analysis highlights that information design may not only serve the purpose to inform but also to monitor agents.

While my results apply to general mechanism design settings with multiple agents, the logic becomes clearest with a single agent. Thus, my analysis focusses on this case. Specifically, suppose a principal (seller) can design a sales mechanism and any information structure (such as a product

[1]Most notably, this is implied by Li and Shi (2017), Esö and Szentes (2007a,b), Krähmer and Strausz (2015). The literature is reviewed in more detail below.

sample) that provides the agent (buyer) with signals (taste experiences) which are informative about an unknown state that affects his valuation (such as an unknown product feature).

At the core of my analysis is an "equivalence" result which says that the principal can implement the same outcome as when the state is verifiable and becomes publicly known ex post. The equivalence results rests on the *combination* of two features: (i) I allow the principal to design various information structures and (to commit) to secretly randomize between them.[2] (ii) I allow the parties to use rich contracting protocols and to condition the terms of trade on (reports about) the outcome of the principal's randomization, that is, the actual information structure.[3]

This has two implications. First, the private signal supplied to the agent can be elicited at no cost. The idea is to have the principal randomize over a set of information structures with the property that any signal can be generated only by a subset of, yet not by all, possible information structures. Hence, if the agent reports a signal that cannot be generated by the realized information structure, his lie is detected. In fact, in my construction, the agent believes any lie to be detected with positive probability. Truth-telling can then be induced by penalizing detected lies.[4] Second, in my construction, the principal will randomize over information structures which are fully informative: knowing the signal and the information structure reveals the true state. Hence, once the agent's signal is elicited and the information structure is verified, the state is revealed.

Thus, when the agent's valuation depends only on the state, the good can be allocated efficiently and the agent can be charged his valuation, leaving him with no rents. When the agent's valuation depends also on his initial private information, efficiency requires to elicit this information, too. But since the state is revealed ex post, the mechanism can de facto condition on it, as if it was a verifiable ex post signal. Thus, if the correlation between the initial private information and the state satisfies a well-known spanning (or full rank) condition, it can be extracted at no cost as in Riordan and Sappington (1988).

The logic from the single agent case carries over to the setting with multiple agents. Impor-

[2]One way to think about this is that the principal can "frame" the selling environment in which information disclosure takes place. For example, a car dealer may offer test drives, secretly employing various types of tires that affect the agent's driving experience (e.g. standard, sporty, comfortable, etc). An agent who cannot distinguish between the various tires is then uncertain whether his driving experience is due to the car as such or the tires.

[3]In Remark 6, I show that with multiple agents, (ii) can be dropped.

[4]In Remark 2, I show that when the agent's valuation only depends on the state, the required off-path penalties can be made arbitrarily small.

tantly however, my results hold under two additional constraints: First, an agent can only be informed about his own preferences, yet not about the preferences of other agents. This constraint respects the notion that, for example, in a private values auction, a bidder's valuation is his private information not only vis-a-vis the auctioneer, but also vis-a-vis the other bidders.[5] Second, I show that with multiple agents, rich contracting protocols are not needed and the principal (or any other player) does not need to observe the realized information structure. Thus, the construction is fully in line with standard notions of information and mechanism design.

The question of my paper is at the heart of a literature on information disclosure in mechanism design (or screening) where the principal controls the private information agents learn beyond their initial private information. My framework covers cases both when agents' initial information is payoff-irrelevant, as in Li and Shi (2017) and Bergemann and Pesendorfer (2007), or is payoff-relevant, as in Esö and Szentes (2007a,b). While my equivalence result resembles the irrelevance results by Esö and Szentes (2007a,b, 2017), the logic is different, as I explain below.[6,7]

In a similar setting (possibly absent transfers), Zhu (2018) also establishes an equivalence result similar to mine. The problem in Zhu (2018) originates from the constraint that agents have to report their initial and the additionally provided information simultaneously, rather than from the constraint that agents cannot be informed about others' preferences, as in my case. Moreover, my analysis allows for additional and initial information to be correlated, covers the single agent case, and shows that also agents' initial private information can be elicited at no cost.

That a designer can benefit from randomizing over information structures is well-known from the literature on mediation and information design (Myerson, 1982, 1986, Bergemann and Morris, 2016, Kamenica and Gentzkow, 2011), where a mediator randomizes over action recommendations for the agent(s). My paper makes clear that when (some) actions are contractible and can condition on reports about the private signals provided to the agents, randomizing over

[5]Moreover, if the principal was allowed to inform agents about other agents' preferences, she could make them common knowledge among all agents and trivially elicit them through a "shoot-the-liar" scheme.

[6]Bergemann and Wambach (2015) show that in the setting of Bergemann and Pesendorfer (2007), the first-best is attainable when the additional information is gradually disclosed and elicited.

[7]The design and sale of additional information for a privately informed agent is also considered in Bergemann, et al. (2017) and Smolin (2019), but these papers focus on simple contracts, and the richer contracting protocols of my paper are ruled out by assumption. Kolotilin et al. (2017) consider the design of information for a privately informed agent, when no actions are contractible, so there is no role for rich contracting protocols.

information structures has the additional benefit to facilitate the elicitation of these signals.[8] Outside the context of information design, a similar point is made in Rahman (2012) and Rahman and Obara (2010) who show that in team problems, conditioning an agent's pay on secret effort recommendations to the other agents, fosters effort incentives and helps to elicit private signals.

The paper is organized as follows. The next section presents the model. Section 3 presents the main results. Section 4 discusses limitations and extensions.

2 The model

There is a principal (she) and an agent (he), and a set X of contractible allocations. The terms of trade consist of an allocation $x \in X$ and a payment $t \in \mathbb{R}$ from the agent to the principal. The parties have quasi-linear utility, where the principal's and the agent's utilities from (x,t) are denoted by $w_{\theta\omega}(x)+t$ and $v_{\theta\omega}(x)-t$ and may depend on two pieces of information, θ and ω. It is common knowledge that θ is drawn from $\Theta = \{1,\ldots,\bar{\theta}\}$ with distribution $r \in \Delta(\Theta)$, and that, conditional on θ, ω is drawn from $\Omega = \{1,\ldots,\bar{\omega}\}$ with distribution $p_\theta \in \Delta(\Omega)$.[9] I refer to θ as the agent's (ex ante) "type", and to ω as the "state". Both parties have an outside option normalized to 0, and I assume that there is $x_0 \in X$ that replicates the outside option: $w_{\theta\omega}(x_0) = v_{\theta\omega}(x_0) = 0$ for all θ, ω.

I assume that there is a well-defined first-best allocation

$$x^*_{\theta\omega} = \arg\max_x [w_{\theta\omega}(x) + v_{\theta\omega}(x)], \tag{1}$$

and that the expected first-best surplus $Z^* = \sum_{\theta,\omega} r(\theta)p_\theta(\omega)[w_{\theta\omega}(x^*_{\theta\omega}) + v_{\theta\omega}(x^*_{\theta\omega})]$ is positive.

At the outset, the agent privately observes his type θ. In contrast, the state ω is not observable, neither by the agent nor the principal. But the principal (and only the principal) can provide the agent with information about ω by designing any information structure that provides the agent with signals about ω. I assume that what the agent learns from this information is his private information. A ("simple") information structure (S,π) consists of a set S of signals and conditional distributions $\pi : \Omega \to \Delta(S)$, where $\pi_\omega \in \Delta(S)$ denotes the signal distribution, conditional on ω.

[8] That a principal can benefit from endogenously creating correlation through randomization has also been observed in the context of auctions. See Obara (2008) or Krähmer (2012).

[9] For continuous Ω or Θ, Remark 7 shows how my approach can be extended to extract approximate full surplus.

The principal's objective is to design an information structure and a mechanism so as to maximize her expected payoff. The novelty of my approach is the *combination* of two features: I allow the principal to randomize among information structures and to employ "rich contracting protocols" which condition the terms of trade on the realized information structure. This means, the information structure is verifiable ex post.[10] Formally, for all k in some set K, let (S,π_k) be an information structure, where $\pi_{\omega k} \in \Delta(S)$ is the signal distribution, conditional on ω and k.[11] The principal may (commit to) select information structures according to any distribution $\mu \in \Delta(K)$. I denote the resulting ("compound") information structure by (K,Π,μ).

The relationship between the principal and the agent proceeds as follows.

1. The agent privately observes θ.
2. The principal commits to an information structure (K,Π,μ) and a mechanism.
3. The agent decides to accept or reject. (Upon rejection, every party gets their outside option.)
4. If the agent accepts, π_k is selected according to μ, unobserved by the agent; and the agent privately observes a signal s generated by the information structure (S,π_k).
5. The terms of trade are enforced according to the mechanism.

For a given information structure, the revelation principle (Myerson, 1986) implies that an optimal mechanism is in the class of direct and incentive compatible mechanisms which require the agent to submit a report $\hat{\theta} \in \Theta$ after stage 3 and a report $\hat{s} \in S$ after stage 4. Consequently, as I allow the terms of trade to condition on the realized information structure, a mechanism (x,t) consists of contingent allocations $x : \Theta \times S \times K \to X$ and transfers $t : \Theta \times S \times K \to \mathbb{R}$.

Given (K,Π,μ) and (x,t), let $u(\theta,s;\hat{\theta},\hat{s})$ be agent type θ's expected utility from reporting $\hat{s}$ ex post when having reported $\hat{\theta}$ ex ante and observed s ex post (if (θ,s) has positive probability). Let $U_{\theta,\hat{\theta}}$ be type θ's expected utility from reporting $\hat{\theta}$ ex ante. Hence, the principal's problem is:

$$P: \quad \max_{(K,\Pi,\mu),(x,t)} \sum_{\theta,\omega} r(\theta) p_\theta(\omega) \int_K \int_S \left[w_{\theta\omega}(x(\theta,s,k)) + t(\theta,s,k) \right] d\pi_{\omega k}(s) d\mu(k) \quad s.t. \tag{2}$$

$$u(\theta,s;\theta,s) \geq u(\theta,s;\theta,\hat{s}) \quad \forall \theta,s,\hat{s}; \tag{3}$$

$$U_{\theta,\theta} \geq U_{\theta,\hat{\theta}} \quad \forall \theta,\hat{\theta}; \tag{4}$$

[10] In Remark 3, I argue that my results go through if the mechanism, more conventionally, can condition only on a report by the principal about k (and employ a budget breaker). In Remark 6, I present a construction with multiple agents in which, as is standard in information design, no party privately observes the realized information structure.

[11] W.l.o.g., the signal space is the same for all k. Otherwise, define $S = \cup_k S_k$.

$$U_{\theta,\theta} \geq 0 \quad \forall\theta. \tag{5}$$

(3) is the ex post incentive compatibility constraint ensuring truthful reporting of s ex post. Note that the revelation principle requires truthful reporting of s only after a truthful report of θ. (4) is the ex ante incentive compatibility constraint ensuring truthful reporting of θ ex ante. (5) is the individual rationality constraint ensuring that all types accept the mechanism.[12]

I say that the principal has *full informational control* if the parties' valuations only depend on the state: $w_{\theta\omega} = w_{\theta'\omega}$ and $v_{\theta\omega} = v_{\theta'\omega}$ for all θ, θ', ω. In this case, the first-best allocation does not depend on the type: $x^*_{\theta\omega} = x^*_{\theta'\omega}$ for all θ, θ', ω. If some party's valuation depends on θ, then I say the principal has *partial informational control*. I say that the agent's beliefs satisfy the *spanning condition* if there is no type $\tilde{\theta}$ whose belief $p_{\tilde{\theta}}$ is a convex combination of the beliefs p_θ of the other types $\theta \neq \tilde{\theta}$.

An important special case is when the type is fully informative about the state, that is, $\Theta = \Omega$, and p_θ places mass 1 on $\omega = \theta$. This case corresponds to the *static screening model* where the state is the agent's private information at the outset, and the principal can generate any signal correlated with the state, to be privately observed by the agent. By convention, I subsume this case under full informational control.

Remark 1. With full informational control, the model corresponds to a discrete type version of Li and Shi (2017). With partial informational control, and if types are orthogonal to states, that is, $p_\theta = p_{\tilde{\theta}}$ for all $\theta, \tilde{\theta}$, the model corresponds to a discrete type version of Esö and Szentes (2007a,b) (after their orthogonalization). Clearly, the spanning condition is violated in this case.

3 Results

The core insight of my analysis is an "equivalence result" which says that the principal can implement the same outcome as in the benchmark in which ω is contractible and becomes publicly known at stage 4 in the time-line above. By the revelation principle, a mechanism (ξ, τ) in the benchmark consists of allocations $\xi : \Theta \times \Omega \to X$ and transfers $\tau : \Theta \times \Omega \to \mathbb{R}$ contingent on a report about θ by the agent after stage 3 and on the true state ω. The benchmark problem is

$$\tilde{P}: \quad \max_{(\xi,\tau)} \sum_{\theta,\omega} r(\theta) p_\theta(\omega) [w_{\theta\omega}(\xi(\theta,\omega)) + \tau(\theta,\omega)] \quad s.t. \tag{6}$$

[12]As the agent's outside option can be replicated by x_0, it is optimal to induce all types to accept the mechanism.

$$\sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(\xi(\theta,\omega)) - \tau(\theta,\omega)] \geq \sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(\xi(\theta,\omega)) - \tau(\hat{\theta},\omega)] \quad \forall \theta, \hat{\theta}; \quad (7)$$
$$\sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(\xi(\theta,\omega)) - \tau(\theta,\omega)] \geq 0 \quad \forall \theta. \quad (8)$$

(7) and (8) are the incentive compatibility and individual rationality constraints. While the value of P is clearly weakly smaller than that of $\tilde{P}$, the equivalence result says that it is, in fact, equal:

Proposition 1 (Equivalence result). *The value of P is equal to the value of $\tilde{P}$.*

The proof of Proposition 1 (in the appendix) employs the following compound information structure. First, $k = (k_1, \ldots, k_{\bar{\omega}}) \in K = [0,1]^{\bar{\omega}}$ is selected according to the uniform (product) distribution on K. Second, conditional on k and ω, the signal $s = k_\omega \in [0,1]$ is disclosed to the agent with probability 1:[13]

$$\pi_{\omega k}(\{s\}) = \begin{cases} 1 & \text{if there is } \omega \in \Omega \text{ with } s = k_\omega, \\ 0 & else. \end{cases} \quad (9)$$

Hence, s and k identify the true state (for almost all k). Moreover, conditional on k, s can occur only if $s = k_\omega$ for some $\omega \in \Omega$. I say that s and k are *consistent* in this case, and they are *inconsistent* otherwise. Thus, if the agent's report $\hat{s}$ deviates from the true s, this is detected as a lie if the true k turns out to be inconsistent with $\hat{s}$.

The key observation is that, conditional on any θ and s, any lie $\hat{s} \neq s$ is inconsistent with the true k with probability 1, because k is uniformly distributed. Thus, any lie is detected for sure, and by penalizing detected lies sufficiently harshly, the agent is induced to report s truthfully.[14] But once s is elicited, s and k identify the true state ω, and the mechanism can de facto condition on it, as if it was publicly known.[15] Next, I show the implications of the equivalence result for full surplus extraction.

Proposition 2. *(i) Under full informational control, the principal can fully extract the first-best surplus Z^*.*

[13]Note that the information structure is well-defined for almost all k, as the event $\{k \in K \mid k_\omega = k_{\tilde{\omega}} \text{ for some } \omega, \tilde{\omega}, \omega \neq \tilde{\omega}\}$ has probability 0, given that k is uniformly distributed.

[14]In Remark 2, I show how under full informational control, penalties can be made arbitrarily small.

[15]Based on different logic, Esö and Szentes (2007a,b, 2017) show an equivalence (or "irrelevance") result when ω is orthogonal to θ and only simple information structures are feasible. Their argument rests on the "first order approach" and regularity conditions, does not extend to discrete types θ (see Krähmer and Strausz, 2015), and does not cover the case with correlation between θ and ω.

(ii) Under partial informational control, and if the spanning condition is satisfied, the principal can fully extract the first-best surplus Z^*.

The argument is immediate from Proposition 1. When the mechanism can condition on ω directly, then under full informational control it is as if the agent has no (payoff relevant) private information at all, and one can simply, without eliciting θ, implement the first-best allocation x^*_ω and charge the agent his valuation $v_\omega(x^*_\omega)$ if state ω is revealed. This attains the first-best and extracts the full surplus. I stress that this logic applies in particular to the static screening model when the agent knows ω at the outset. Indeed, even if the agent knows ω, the optimal mechanism requires the agent only to make a report about s (not about ω). Since for any ω, the agent believes any lie about s to be detected with probability 1 and appropriately punished, he has no incentive to deviate.

In contrast, under partial informational control, the mechanism also needs to elicit the agent's type θ to attain the first-best. Now notice, since ω is de facto verifiable by Proposition 1, ω effectively amounts to an ex post verifiable signal about θ. But if the principal has access to such a signal, and the correlation between θ and ω satisfies the spanning condition, then she can elicit the agent's type at no cost as in Riordan and Sappington (1988). Thus, the value of $\tilde{P}$ is Z^*.

I conclude the paper by discussing limitations and extensions of my findings.

4 Concluding remarks

Remark 2 (Individual rationality and limited liability)**.** Under full informational control, since the agent pays his valuation in any state, he gets zero expected utility conditional on s (resp. on ω). Thus, individual rationality holds even if the agent can still choose his outside option after observing s, and he does not make a loss "on path", after ω is revealed.

I now argue that, under full informational control, ex post losses can essentially be avoided also "off path". The reason is that under the information structure (9), any lie $\hat{s} \neq s$ is inconsistent with the true k with probability 1, conditional on any θ and s. With full informational control, where the agent receives 0 utility from truth-telling for all θ and s, lying is thus dissuaded by enforcing x_0 and an arbitrarily small penalty $T > 0$ after an inconsistent report.

This also implies that in the static screening model, when the principal can generate a signal correlated with the agent's private information, he can extract full surplus with only tiny off-path losses for the agent. While this conclusion may appear at odds with basic economic intuition, it should be viewed as an extreme implication of the assumptions that the principal can design and commit to any signal that is correlated with the true state, and that she can mix between signals.[16] While this is conceptually not different from standard information design where the principal commits to mix over action recommendations, using a continuum of information structures as in (9) might be infeasible in practice. In Remark 4, I show that full surplus can be extracted with only finitely many simple information structures, but the construction requires the agent not to perfectly know ω at the outset and may require large off-path punishments.[17]

Remark 3 (Non-verifiable information structure)**.** Proposition 1 exploits that the information structure is verifiable. This can be dropped when the principal privately observes k and is required to report k. That is, after stage 4 in the time-line above, principal and agent simultaneously report $\hat{k}$ and $\hat{s}$ respectively. For the same reasons as in the case with verifiable k, truth-telling is a mutual best response if both parties are penalized after inconsistent reports ($\hat{s} \neq \hat{k}_\omega$ for all $\omega \in \Omega$). As both parties are penalized after inconsistent reports, a budget-breaker is required off path.[18]

Remark 4 (Finite information structures)**.** Similarly to ideas in Zhu (2018), Proposition 1 can be shown using only finite S and K. Let $S = K = \{0, 1, \ldots, \bar{\omega}\}$. Conditional on ω and k, let the signal that is released with probability 1 be $s = \omega + k$ modulo $\bar{\omega} + 1$.[19] For the two states case, Table 1 depicts the signal s that is released, conditional on ω and k:

[16]This means also that the principal needs to make the agent understand what mixture she has committed to. While this may require a high degree of agent sophistication, especially when the principal uses many information structures such as with (9), the essence of what the agent needs to understand is that the principal (or the court) has an idea about what information the agent may possibly possess and that this allows her to detect a lie.

[17]With partial informational control, in contrast, the Riordan/Sappington mechanism I employ does impose losses for some θ and ω on path as well as after a lie $\hat{\theta} \neq \theta$ off path. Hence, full surplus extraction generally fails if the agent can only sustain limited losses (Demougin and Garvie, 1991). However, Proposition 1 can be shown to hold even if the agent can only make payments up to a bound and if there is a uniformly worst allocation for the agent for all θ and ω. (A proof is available upon request.)

[18]That courts, in principle, have ways to verify the provision of information and impose off-path fines is not entirely unrealistic. Albeit in a different context but in a similar spirit, after the financial crises, banks had to pay significant fines after having been caught to have misinformed investors. These fines did often not go to investors but to state coffers. See https://www.marketwatch.com/story/banks-have-been-fined-a-staggering-243-billion-since-the-financial-crisis-2018-02-20

[19]Recall that for two natural numbers m, n, m modulo n is the remainder of m/n.

s	$k=0$	$k=1$	$k=2$
$\omega=1$	1	2	0
$\omega=2$	2	0	1

Table 1: Information structure with finitely many signals

As can be seen by inspection, s and k identify the true ω. Moreover, signal s is consistent with all $k \neq s$. A lie $\hat{s} \neq s$ is inconsistent with $k = \hat{s}$, and hence detected if $k = \hat{s}$. Because Θ, S, K are finite, it is easy to see that, conditional on θ and s, the probability that a lie is detected is bounded from below, provided p_θ has full support Ω for all θ, and μ has full support K. Hence, a lie can be deterred by punishing inconsistent reports. This implies Proposition 1.

Note that p_θ have full support Ω for all θ is indispensable: in Table 1, if the agent knew that $\omega = 1$ and observed $s = 1$, he could infer that $k = 0$ and safely deviate to $s = 2$. Thus, the finite construction does not cover the static screening model. Note also that the probability that a lie about s is detected is bounded away from 1. Thus, even with full informational control, off-path punishments might need to be large, as the argument in Remark 2 requires that lies be detected with probability 1, conditional on all θ and s, and hence no longer applies here.

Remark 5 (Unit good). An important setting is the "unit good case", especially with full informational control as in Li and Shi (2017a). In this case, $x \in [0, 1]$, and valuations are $w_\omega x$ and $v_\omega x$. As I show next, full surplus can then be extracted with only three information structures and three signals. Indeed, let $\Omega^+ = \{\omega \mid w_\omega + v_\omega \geq 0\}$ (resp. $\Omega^- = \{\omega \mid w_\omega + v_\omega < 0\}$) be the set of states in which trade is (resp. is not) efficient. Since all that matters for efficiency is whether ω is in Ω^+ or in Ω^-, this effectively corresponds to the two states case. Consider now the information structure in Table 1 (with Ω^+ corresponding to "$\omega = 1$" and Ω^- to "$\omega = 2$"). Then s and k identify whether trade is efficient or not (e.g., $s = 1$ and $k = 0$ identify that it is, while $s = 1$ and $k = 2$ identify that it is not). As above, truth-telling of s can be induced by penalizing inconsistent reports. The following mechanism then extracts full surplus: if the agent's report s and k identify a state in Ω^+, then trade occurs, and the agent is charged his expected valuation $E[v_\omega \mid \omega \in \Omega^+]$. Otherwise, there is no trade, and he is charged nothing. Note also that the mechanism is individually rational, conditional on s, and does not elicit θ, that is, unlike in Li/Shi (2017a), the agent is not screened sequentially. Also, the argument does not depend on the discreteness of Ω or Θ.

Remark 6 (Multiple agents)**.** Propositions 1 and 2 extend to multiple-agents settings. Similarly to Remark 3, a signal s_i disclosed to agent i can be elicited without cost by privately informing agent $j \neq i$ about, and have him report, the information structure used to generate s_i, and by penalizing all agents (cashed in by the principal) if their reports are mutually inconsistent.

While the abstract notion of an information structure does not rule out to privately inform an agent about the information structure used for others, this may be hard to implement. To address this, I now sketch a construction where no party gets to know the realized information structure.

Suppose there are two agents $i = 1, 2$ whose valuations $v^{(i)}_{\theta_i,\omega_i}(x)$ each depend on (agent specific) states $\omega_i \in \{1, 2\}$ about which they can be informed by the principal, and types θ_i. While (θ_1, ω_1) and (θ_2, ω_2) are independent, θ_i may be correlated with ω_i. Let $K = \mathbb{Z}$, and let agent 1 and 2 receive signals

$$s_1 = \omega_1 + k, \quad s_2 = \omega_2 + 1/2 \cdot k. \tag{10}$$

Note that s_i is uninformative about ω_j, $j \neq i$. This captures the idea that preferences are an agent's private information (also vis-a-vis other agents) and cannot be disclosed to another agent.[20,21]

Conditional on k, the support of the distribution of (s_1, s_2) consists of four pairs $\mathscr{S}(k) = \{(1 + k, 1 + k/2), (1 + k, 2 + k/2), (2 + k, 1 + k/2), (2 + k, 2 + k/2), \}$. Since these supports are disjoint for all k, any pair (s_1, s_2) identifies the true k and thus ω_1 and ω_2. Moreover, s_1 and s_2 are now inconsistent if (s_1, s_2) is not in the support $\cup_k \mathscr{S}(k)$ of the (unconditional) joint distribution. As is easy to see, given truth-telling by the other agent, each agent attaches a positive probability to a lie being inconsistent with the other agent's report. Hence, lies can be deterred by penalizing inconsistent reports.[22] While this implies Proposition 1, θ_i can now be elicited as in the single agent case, hence Proposition 2 follows.

Again, I stress that under this construction, no party has private information about k, and the mechanism only conditions on reports about signals s_1 and s_2. Thus, the construction is fully in

[20] Note also that if the principal could disclose information to i which is correlated with ω_j, $i \neq j$, she could simply fully disclose the profile (ω_1, ω_2) to both agents and elicit it by some shoot-the-liar scheme.

[21] An interpretation of (10) is that the principal uses "salesmen" to inform agents. Each k corresponds to a salesman, and while every salesman is truthful, the larger is k, the more "inflated" the language the salesman uses.

[22] Similarly to Remark 4, the argument requires that type θ_i's belief has full support over his types ω_i. Moreover, because lies are detected with a probability bounded away from 1, large punishments may be required to dissuade lies. It is possible, however, to extend the construction so that a lie is detected with a probability arbitrarily close to 1. I omit the details.

line with standard notions of information and mechanism design. Moreover, no agent i's signal is correlated with the other agent j's preferences.

Remark 7 (Continuous state and type space)**.** If $\Omega = [\underline{\omega}, \bar{\omega}]$, $\underline{\omega} < \bar{\omega}$, is continuous, my construction can be adapted by partitioning Ω in L segments $\Omega_\ell = [\omega_\ell, \omega_{\ell+1})$, $\ell = 1, \ldots, L$, $\omega_1 = \underline{\omega}$, $\omega_{L+1} = \bar{\omega}$, $\omega_\ell < \omega_{\ell+1}$. Consider now the information structure analogous to (9) where the principal uniformly randomizes over $k \in [0,1]^L$, and the signal $s = k_\ell$ is released to the agent with probability 1 conditional on k and $\omega \in \Omega_\ell$. As in Proposition 1, one can elicit without cost whether $\omega \in \Omega_\ell$. With full informational control, by choosing the partition sufficiently finely, the principal can thus attain the first best surplus almost fully (be Θ discrete or continuous). Almost full surplus extraction can be attained also under partial informational control if Θ is finite. (The spanning condition can now be made to hold by design of the partition.) If Θ is continuous, almost full surplus extraction can be attained under the belief conditions in McAfee and Reny (1992).

Remark 8 (Necessity)**.** Under partial informational control it can be shown that if the spanning condition fails, then there there are valuations $v_{\theta\omega}(\cdot)$ for which full surplus extraction fails. The argument is the same as in the proof of Theorem 2 in Crémer and McLean (1988).

Appendix

Proof of Proposition 1 Let (9) be given, and let W^* (resp. $\tilde{W}^*$) be the value of P (resp. $\tilde{P}$). Clearly, $W^* \le \tilde{W}^*$. To see that $W^* \ge \tilde{W}^*$, let (ξ, τ) be a solution to $\tilde{P}$. For $T > 0$, define (x, t) by

$$(x(\theta,s,k),\ t(\theta,s,k)) = \begin{cases} (\xi(\theta,\omega),\ \tau(\theta,\omega)) & \text{if } s = k_\omega \text{ for some } \omega \in \Omega \\ (x_0,\ T) & \text{else} \end{cases}. \tag{11}$$

As in footnote 13, the mechanism is well-defined for almost all k. To show $W^* \ge \tilde{W}^*$, it suffices to show that (x, t) satisfies (3), (4), (5), and attains $\tilde{W}^*$.

To see (3), suppose type θ reported $\hat{\theta}$ and observed s. If the agent reports truthfully ($\hat{s} = s$), then for all θ, $\hat{s}$ is consistent with k, and ω is identified with probability 1. Thus, by (11), $(\xi(\hat{\theta}, \omega),\ \tau(\hat{\theta}, \omega))$ is implemented in state ω. Hence,

$$u(\theta,s;\hat{\theta},s) = \sum_{\omega\in\Omega} p_\theta(\omega)[v_{\theta\omega}(\xi(\hat{\theta},\omega)) - \tau(\hat{\theta},\omega)]. \tag{12}$$

On the other hand, if the agent reports $\hat{s} \neq s$, then his report is inconsistent with k with probability 1, and by (11), (x_0, T) is implemented. Hence, $u(\theta, s; \hat{\theta}, \hat{s}) = -T$. This is smaller than $u(\theta, s; \hat{\theta}, s)$ in (12) for sufficiently large T. Hence, (3) is met.

To see (4) and (5), I compute $U_{\theta,\hat{\theta}}$. By the previous step, it is optimal for the agent to report s truthfully ex post for any ex ante report $\hat{\theta}$. Therefore, his report is consistent with k, and ω is identified with probability 1. Thus, by (11),

$$U_{\theta,\hat{\theta}} = \sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(\xi(\hat{\theta}, \omega)) - \tau(\hat{\theta}, \omega)]. \tag{13}$$

By inspection, $U_{\theta,\hat{\theta}}$ coincides with the r.h.s of (7), and $U_{\theta,\theta}$ coincides with the l.h.s of (7) and (8). Since (ξ, τ) is a solution to $\tilde{P}$, (4) and (5) follow from (7) and (8).

Last, with analogous steps that lead to (13), the principal's utility from the (x, t) coincides with the objective (6). Hence, the principal obtains $\tilde{W}^*$ from (x, t). qed

Proof of Proposition 2 The argument is given in the main text. qed

References

Bergemann, D., Bonatti, A., Smolin, A., 2018. The design and price of information. Am. Econ. Rev. 108, 1-48.

Bergemann, D., Morris, S., 2016. Bayes correlated equilibrium and the comparison of information structures in games. Theor. Econ. 11, 487-522.

Bergemann, D., Pesendorfer, M., 2007. Information structures in optimal auctions. J. Econ. Theory 137, 580-609.

Bergemann, D., Wambach, A., 2015. Sequential information disclosure in auctions. J. Econ. Theory 159, 1074-1095.

Crémer, J., McLean, R. P., 1988. Full extraction of the surplus in Bayesian and dominant strategy auctions. Econometrica 56, 1247-1257.

Demougin, D. M., Garvie, D. A., 1991. Contractual design with correlated information under limited liability. Rand J. of Econ. 22, 477-489.

Esö, P., Szentes, B., 2007a. Optimal information disclosure in auctions and the handicap auction. Rev. Econ. Stud. 74, 705-731.

Esö, P., Szentes, B., 2007b. The price of advice. Rand J. of Econ. 38, 863-880.

Esö, P., Szentes, B., 2017. Dynamic contracting: an irrelevance result. Theor. Econ. 12, 109-139.

Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. Am. Econ. Rev. 101, 2590-2615.

Kolotilin, A., Li, M., Myolovanov, T., Zapechelnyuk, A., 2017. Persuasion of a privately informed receiver. Econometrica 85, 1949-1964.

Krähmer, D., 2012. Optimal auction design with endogenously correlated buyer types. J. Econ. Theory 147, 118-141.

Krähmer, D., Strausz, R., 2015. Ex post information rents in sequential screening. Game. Econ. Behav. 90, 257-273.

Myerson, R. B., 1982. Optimal coordination mechanisms in generalized principal-agent problems. J. Math. Econ. 10, 67-81.

Myerson, R. B., 1986. Multistage games with communication. Econometrica 54, 323-358.

Li, H., Shi, X., 2017. Discriminatory information disclosure. Am. Econ. Rev.107, 3363-3385.

McAfee, P., Reny, P., 1992. Correlated information and mechanism design. Econometrica 60, 395-421.

Obara, I., 2008. The full surplus extraction theorem with hidden actions. BE J. Theor. Econ. 8 (Adv.), Article 8.

Rahman, D., 2012. But who will monitor the monitor?. Am. Econ. Rev. 102, 2767-2797.

Rahman, D., Obara, I. (2010). Mediated partnerships. Econometrica 78, 285-308.

Riordan, M. H., Sappington, D. E., 1988. Optimal contracts with public ex post information. J. Econ. Theory 45, 189-199.

Smolin, A., 2019. Disclosure and pricing of attributes. Mimeo.

Zhu, S., 2018. Private disclosure with multiple agents. Mimeo.