Information disclosure and monopolistic screening

Daniel Krähmer*

September 13, 2017

Abstract

I consider a monopolistic seller who can provide additional product information such as product samples to allow consumers to ascertain whether the product fits their tastes. The paper shows that, even if the consumer has private ex ante information and the disclosure of product information increases his private information, the seller can design product information and a selling mechanism so as to fully extract the (unconstrained) first-best surplus if the consumer's ex ante private information is payoff-irrelevant. The result also holds if the consumer's ex ante private information is payoff-relevant and satisfies a spanning condition familiar from the literature on mechanism design with correlated valuations.

Keywords: information design, sequential screening, disclosure, correlation, 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 Stephan Lauermann, Martin Pollrich, David Rodina, Alex Smolin, Roland Strausz, and Tymon Tatur for inspiring discussions.

1 Introduction

In many product markets, sellers offer consumers the possibility to try out, test, or inspect their products prior to purchase. Online stores offer free book previews, music samples, or movie trailers, or grant consumers withdrawal periods to try out an order for a while. Car dealers offer test drives, real estate agents invite interested buyers to on-site inspections, etc.

In these examples, sellers provide information about "product fit" which enables consumers primarily to better ascertain whether the product fits their tastes rather than to verify its objective quality (which often becomes apparent only in the long run). A key feature of product fit information is that while a seller may control its informativeness, for example by setting the time a consumer is allowed to inspect the product or the richness of the product description, how exactly the information influences a consumer's valuation is not verifiable and becomes the consumer's private information. From a welfare point of view, this raises the question whether a seller has incentives to provide product information as well as to allocate the product efficiently.

In this paper, I address this question in the context of a monopolistic seller who faces a buyer who has some, yet imperfect initial private information about whether the product matches his tastes. The seller is allowed to design any "information structure" (such as a test product or product sample) that provides the buyer with "signals" (for instance, taste experiences) which are informative about an otherwise unknown "state" that affects his valuation for the product (such as an unknown product feature that is of interest to the buyer). In this context, the literature has shown that while the provision of product fit information can improve the seller's revenue, the seller typically faces a rent-efficiency trade-off and cannot avoid conceding information rents to the buyer, leading to a distorted product allocation and efficiency losses.[1]

The main result in this paper shows that, in contrast, in a large class of cases the seller can design product fit information that allows her to allocate the product efficiently and to fully extract the resulting first-best surplus. More precisely, if the buyer's initial information is payoff-irrelevant in the sense that it only affects his beliefs about, yet does not directly influence, his valuation of the product, then the seller can always extract the full first-best surplus. I also show that the seller can extract the full first-best surplus if the buyer's initial information is payoff-relevant, provided it

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

is correlated with the state in the sense that the buyer's beliefs about the state satisfy a "spanning condition" that features in the literature on mechanism design with correlated types (Cremer and McLean, 1988, Riordan and Sappington, 1988).

The novelty which distinguishes my approach from the existing literature and which allows the seller to improve efficiency is the combination of two key features: (i) I allow the seller to design various information structures and (to commit) to randomize between them, without disclosing the outcome of the randomization to the buyer. One way to think about this is that the seller can "frame" the selling environment in which information disclosure takes place. For example, a car dealer may offer test drives employing various types of tires that affect the buyer's driving experience (e.g. sporty, comfortable etc). A buyer who cannot distinguish between the various tires is then uncertain whether his driving experience is due to the car as such or the tires. (ii) I allow the parties to employ rich contracting protocols and to condition the terms of trade on (reports about) the outcome of the seller's randomization, that is, the actual information structure.

Allowing for these two features has two implications. First, the seller can elicit the private signal she supplies to the buyer at no cost. The reason is that if the seller randomizes, the signal the buyer obtains is correlated with the true information structure. More specifically, the idea is to have the seller randomize over a set of information structures with the property that any signal that the buyer may observe can be generated by a subset of, yet not by all, possible information structures.[2] In particular, if the agent reports a signal that cannot be generated by the actual information structure, it becomes apparent that the agent must have lied. In fact, I construct an information structure with the property that for any signal, the agent believes any deviation from truth-telling to be detected as a lie with positive probability. Truth-telling can then be induced by penalizing the agent if a lie is detected.

Second, the seller can allocate the product efficiently. In my construction, the seller will randomize over information structures which are each fully informative: knowing the signal the buyer observes and knowing the true information structure reveals the true state. Hence, once the buyer's signal is elicited and the information structure is verified, the state is revealed. As a consequence, if the buyer's initial information is payoff-irrelevant so that only the state matters for his valuation, the product can be allocated efficiently and the buyer can be charged his val-

[2]The information structure I use is similar in spirit to an information structure recently presented by Zhu (2017).

uation. If the initial information is payoff-relevant and correlated with the state, the revelation of the true state can be leveraged to extract the buyer's initial information at no cost using the insights from mechanism design with correlated types (Cremer and McLean, 1988, Riordan and Sappington, 1988).

I discuss various extensions of the main result. Most importantly perhaps, my assumption that the terms of trade can condition on the true information structure requires that the true information structure is verifiable by a third party. I show that this assumption can be dropped provided there is a budget breaker. The idea is that if the information structure is the seller's private information, this information is correlated with the buyer's signal and can therefore be elicited from the seller by cross-checking the "consistency" of the seller's and buyer's reports and by penalizing both if their reports are not consistent. In the latter case, this requires a budget breaker to pocket the penalty, but this threat is never enforced on the equilibrium path.

I also investigate to what extent the spanning condition which is sufficient for full surplus extraction in the payoff-relevant case is necessary. I show that if the spanning condition is violated, then there are always buyer valuations so that for no information structure full surplus extraction is feasible.[3] This includes information structures that depend on a report by the buyer about his initial information[4] (termed "discriminatory information disclosure" by Li and Shi, 2017). However, if the information structure can depend not only on a report by the buyer but also on his true initial information (termed "general disclosure" by Li and Shi, 2017), then full surplus extraction is again possible for all beliefs of the buyer.

The question I address in this paper is at the heart of a recent literature that studies information disclosure in sequential screening by assuming that the seller controls the additional information the buyer learns beyond his initial private information. This literature has focussed on situations in which the seller cannot randomize over disclosure and/or cannot use rich contracting protocols.[5] My framework encompasses cases both in which the buyer's initial information is payoff-irrelevant and correlated with the additional information the seller can disclose, as in Li and Shi (2017), as well as in which the initial information is payoff-relevant and orthogonal to the additional

[3]The argument is essentially identical to that of the "necessity" statement in Cremer and McLean, 1988.

[4]The information structure employed for my main result is not report-dependent.

[5]For sequential screening settings where the additional information cannot be controlled by the seller, see, e.g., Baron and Besanko (1985) or Courty and Li (2000). See Krähmer and Strausz (2015b) for an overview.

information, as in Esö and Szentes (2007a,b).[6] While I show that the first-best is attainable in Li and Shi (2017) type settings, I also show that this is generally not the case in Esö and Szentes (2007a,b) type settings, because if the initial and additional information are orthogonal, then the spanning condition mentioned earlier is violated.[7]

My paper is also closely related to work by Zhu (2017) who considers an information plus mechanism design setting where multiple agents have initial private information, and the designer can disclose additional information about an orthogonal "state" that affects all agents' preferences. The main point of Zhu's (2017) work is an "irrelevance result" which says that the signals provided to the agents can be elicited without cost, and that, in fact, the designer can implement the same outcome as if the state was publicly observable. I adapt the spirit of Zhu's (2017) result, which exploits correlation between signals across players, to provide a similar irrelevance result in a single-agent context by exploiting the correlation between signal and information structure.[8] While Zhu's (2017) irrelevance result, in contrast to mine, covers settings without quasi-linear preferences, I go beyond Zhu (2017) by allowing for correlation between the initial information and the state. Moreover, and most importantly, in contrast to my main results, Zhu (2017) does not focus on implications for efficiency and rent extraction.

[6] To be precise, Esö and Szentes (2007a,b) arrive at and work with such a model after applying their "orthogonalization" approach to a model with payoff-irrelevant initial information. As Li and Shi (2017) make clear, when the seller controls the additional information disclosed to the buyer, the orthogonalization is not innocuous, because it matters whether the seller can disclose information about the buyer's valuation or only about an orthogonal component of it. Moreover, my framework does not literally nest Li and Shi (2017) and Esö and Szentes (2007a,b) because to facilitate tractability, I only allow for discrete information. Finally, Esö and Szentes (2007a) allow for multiple agents.

[7] Li and Shi (2017) do provide an example in which the seller is able to extract the full first-best surplus, but next to exploiting a particular distributional specification, the example rests on the unit good assumption, while my results hold also in the non-unit good case. Moreover, Li and Shi (2017) show that the seller may benefit from using (partial) information disclosure as a price discrimination device. In my setting with richer contracting possibilities, the seller does not need to engage in discriminatory information disclosure, and the information disclosed to the buyer is, by itself, only partially informative, but jointly with knowledge about the information structure fully reveals the buyer's valuation.

[8] These irrelevance results are similar in spirit to irrelevance results by Esö and Szentes (2007a,b, 2017), but, as I explain below, rest on a different logic. One difference is that both Zhu's (2017) and my irrelevance result holds in a setting where the buyer's initial information is discrete in which case Esö and Szentes' (2007a,b) irrelevance result generally does not hold, as shown by Krähmer and Strausz (2015a).

The design of additional information for a privately informed agent is also considered in Bergemann, Bonnati, and Smolin (2017), but in their setting only transfers are contractible, and the agent takes a non-contractible action post information-revelation. While Bergemann, Bonnati, and Smolin (2017) restrict attention to simple contracts where the seller offers a menu of information structures and prices, allowing for the richer contracting protocols of my setting has the potential to improve the seller's revenue, but my first-best results do not directly carry over due to the presence of the additional obedience constraints resulting from the agent's non-contractible action.

The idea that a designer can benefit from using random information structures is well-known from Myerson's (1982, 1986) work on mediation, and the more recent literature on information design (Bergemann and Morris, 2016) or Bayesian persuasion (Kamenica and Gentzkow, 2011). For, randomizing over information structures simply corresponds to the standard notion of a mediator randomizing over action recommendations for the agent(s) and is thus implicit in the appropriate notion of correlated equilibrium. What my paper makes clear is that in a framework where (some) actions are contractible and can condition on reports about the private signals provided to the agents, randomizing over information structures has the additional benefit that it facilitates the elicitation of these signals from the agents.

This point is also made in Rahman (2012) and Rahman and Obara (2010) who show that in team problems, making an agent's pay contingent on secret effort recommendations made to the other agents, fosters effort incentives and allows to elicit signals privately observed by an agent.[9] If the distribution of the private signal depends on others efforts, making a secret (and incentive compatible) random effort recommendation to others corresponds to secretly randomizing over information structures in my setting, and making pay contingent on the effort recommendation made to others corresponds to making the terms of trade contingent on the outcome of the randomization in my setting. As in my approach, an agent in Rahman (2012) and Rahman and Obara (2010) reports truthfully since a lie will be, in a probabilistic sense, inconsistent with the true effort recommendation made to others.

A different force is at work in Rodina (2017) who, in the context of a career concerns framework with moral hazard, shows that making the information provided to the market contingent on a secret random effort recommendation to the agent may increase this agent's effort incentive.

[9] See also Strausz (2012).

The agent will then hold different beliefs about the distribution of the market wage, depending on the recommendation, and this relaxes the effort constraint.[10]

The paper is organized as follows. The next section presents an example that illustrates the main logic behind the paper. Section 3 presents the model, and section 4 contains the main result. Section 5 discusses extensions, and section 6 concludes. All proofs are in the appendix.

2 Example

Consider a seller and a buyer who can consume 0, 1, or 2 units of the seller's product. The buyer may have a low ($\omega = 0$) or a high ($\omega = 1$) valuation. Suppose it is efficient that a low valuation buyer consume one unit and a high valuation buyer two units. To be specific, suppose the buyer's valuation from consumption and the respective surplus are as in the following table (the fist entry in a cell ist the agent's valuation, and the second entry is the total surplus, that is, valuation minus production costs):

valuation, surplus	0 units	1 unit	2 units
$\omega = 0$	$0, 0$	$100, 50$	$200, 0$
$\omega = 1$	$0, 0$	$200, 150$	$400, 200$

Table 1: Valuation and surplus

The buyer does not know his valuation, but holds an ex ante belief about it, $p_\theta \in \Delta(\{0, 1\})$, where θ may be the buyer's ex ante private information. Since the buyer's valuation depends only on ω in this example, the buyer's ex ante private information θ is payoff-irrelevant, however. The seller can inform the buyer about his valuation, for example, by giving him a test sample (henceforth, a "test") of the product. A test provides the buyer with a taste experience, a "signal", that is correlated with his valuation. Suppose that the taste experience is not verifiable and is the buyer's private information. As the literature has shown (see Esö and Szentes, 2007a,b, Krähmer and Strausz, 2015, Li and Shi, 2017), even though the seller can design a single test so as to

[10]That a seller can benefit from endogenously creating correlation through randomization has also been observed in other contexts. Krähmer (2012) shows how an auctioneer can create correlation by randomizing over investments that improve bidders' valuations stochastically. In Obara (2008), bidders can take (hidden) actions that influence the joint distribution of their valuations. He demonstrates that almost full surplus extraction can be attained by a mechanism which implements a mixed action profile by bidders.

reduce the information rent to the buyer, in a situation in which the buyer holds private ex ante beliefs, the seller typically faces an efficiency-rent trade-off and has to leave a positive information rent to the buyer if she wants to implement the first-best.

The key observation of this paper is that the seller can overcome the rent-efficiency trade-off and implement the first-best without paying rents to the buyer by designing various tests and randomizing between them. For example, the seller may influence the intensity of the buyer's taste experience by "framing" the environment in which the test takes place so as to amplify or dim the buyer's taste experience. Intuitively, when having a great taste experience, the buyer is then uncertain whether this is due to the product or the "frame". More formally, suppose the seller has at its disposal a series of tests $k \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, and test k yields the signal

$$s = \omega + k \tag{1}$$

with probability 1. Moreover, the seller (commits to) choose a probability $\mu(k) > 0$ with which she picks test k for the buyer. Crucially, however, the buyer is not informed about the identity k of the test.

I shall now argue that the seller can use this test design, together with an appropriate sales mechanism, to fully extract the efficient surplus. The mechanism will ask the buyer after the test to announce a report $\hat{s}$ about his signal, and, in addition, will condition the terms of trade on the identity k of the true test. Hence, I assume that the identity of the test is verifiable by a third party. (Below, I will argue that this assumption can be dropped if one allows for the presence of a budget breaker). Full surplus extraction becomes possible through two key features: first, the signal and the identity of the test jointly reveal the true valuation, and second, the seller can induce the buyer to report his signal truthfully at no cost.

More precisely, consider the following mechanism where the entries in the second (resp. third) column specify the number of units assigned to (resp. the price charged from) the buyer if he reports $\hat{s}$ and the true test is verified to be equal to k:

	number of units	payments
$\hat{s} - k = 0$	1	100
$\hat{s} - k = 1$	2	400
$\hat{s} - k \notin \{0, 1\}$	0	$T >> 0$

Table 2: Mechanism

Suppose the buyer has observed, say, signal $s_0 = 10$. This reveals to the buyer that the true test $k = s_0 - \omega$ is either $k = 10$ or $k = 9$, and all other tests $k \notin \{9, 10\}$ cannot have occurred. Moreover, the buyer infers that if $k = 10$, then his valuation is low (because the signal $s_0 = 10$ implies that $\omega = s_0 - k = 0$), and if $k = 9$, then his valuation is high (because the signal $s_0 = 10$ implies that $\omega = s_0 - k = 1$). Therefore, if the buyer reports his signal s truthfully, then if $k = 10$, he receives 1 unit which he values at 100 and he pays 100, and if $k = 9$, he receives 2 units which he values at 400 and he pays 400. Consequently, if the buyer reports truthfully, the efficient quantity is implemented, and the buyer gets no rent. In particular, the seller obtains the full surplus.

To see that the mechanism does indeed induce the agent to report his signal truthfully, recall that when $s_0 = 10$, then the buyer knows that the true test $k = s_0 - \omega$ is either $k = 10$ or $k = 9$. Now, the test $k = 10$ can only generate the signals $s = 10$ and $s = 11$, and the test $k = 9$ can only generate the signals $s = 9$ and $s = 10$. Therefore, if the buyer deviates and falsely reports a signal $\hat{s} \notin \{9, 10, 11\}$, he knows for sure that his "lie will be detected". In other words, if the buyer reports $\hat{s} \notin \{9, 10, 11\}$, then, since $k - \hat{s} \notin \{0, 1\}$, he will receive 0 units of the good and has to pay T, leaving him with a utility $-T$.

Moreover, if the buyer falsely reports $\hat{s} = 9$, then this lie will be detected if the true test is $k = 10$, because $k = 10$ is only consistent with signals $s = 10$ and $s = 11$. Thus, in this case, the agent gets 0 units and pays T; and if the buyer falsely reports $\hat{s} = 11$, this lie will be detected if the true test is $k = 9$ (because $k = 9$ is only consistent with signals $s = 9$ and $s = 10$), in which case he gets 0 units and pays T. Since $\mu(k) > 0$ for all k, the buyer attaches positive probability to $k = 10$ and $k = 9$, and he therefore expects to receive utility $-T$ with positive probability if he reports $\hat{s} \in \{9, 11\}$.[11] Hence, if T is sufficiently large, the buyer is deterred from lying about s, and this shows that the mechanism yields the seller the first-best surplus.

Notice that the argument just given is entirely independent from the fact that the buyer's ex ante beliefs about his valuation may be his private information. The reason is that in the example, the buyer's ex ante private information is payoff-irrelevant. In the main analysis, I will allow for the case that the buyer holds private information ex ante which affects both his beliefs and his valuation. I will show that in this case full surplus extraction becomes possible if the buyer's beliefs satisfy the so-called "spanning condition" as it features in the literature on mechanism

[11] To be precise, this requires that the buyer's private ex ante beliefs are non-degenerate and assign positive probability to both states.

design with correlated valuations (Cremer and McLean, 1988, Riordan and Sappington, 1988).

Finally, let me indicate how the assumption that the identity of the test k is verifiable can be relaxed. Suppose that the identity of the test is the seller's private information. Adapt the mechanism as follows: require also the seller to make a report $\hat{k}$ about k, and impose a large penalty F on the seller if, given a report $\hat{s}$ by the agent, $\hat{s}-\hat{k} \notin \{0,1\}$. Given that the buyer reports truthfully, the seller attaches positive probability to the event that a lie $\hat{k} \neq k$ will be detected (that is, she believes that $\hat{s}-\hat{k} \notin \{0,1\}$ with positive probability if $\hat{k} \neq k$). Thus, if the penalty is large, she will be deterred from lying. Notice that this mechanism is no longer budget-balanced, because in the event that one party lies, both parties may be fined. However, the penalties are actually never imposed on the equilibrium path.

3 Model

There is a principal (she) and an agent (he). The principal can produce a quantity $x \geq 0$ of some good at costs $c(x)$ where $c : \mathbb{R}_+ \to \mathbb{R}$. The agent's valuation for consuming x is equal to $v(x)$ where $v : \mathbb{R}_+ \to \mathbb{R}$ with $v(0) = 0$ and $v(x) \geq 0$. The terms of trade consist of a quantity and a payment from the agent to the principal. The parties have quasi-linear utility, that is, if the terms of trade are x and t, the principal's utility is $t - c(x)$, and the agent's utility is $v(x) - t$.[12]

The agent's valuation $v = v_{\theta\omega}$ may depend on two pieces of information, θ and ω.[13] It is common knowledge that θ is drawn from the set $\Theta = \{1, \ldots, \bar{\theta}\}$ with distribution $r \in \Delta(\Theta)$, and that, conditional on θ, ω is drawn from the set $\Omega = \{0, \ldots, \bar{\omega}\}$ with distribution $p_\theta \in \Delta(\Omega)$. I impose the (mild) assumption that p_θ has full support for all θ. I refer to θ as the agent's (ex ante) "type", and to ω as the "state".

I say that the type is payoff-irrelevant when the valuation depends only on the state: $v_{\theta\omega} = v_{\theta'\omega}$ for all θ, θ', ω. Otherwise, I say the type is payoff-relevant.

I assume that there is a well-defined first-best quantity given by

$$x^*_{\theta\omega} = \arg\max_x v_{\theta\omega}(x) - c(x), \tag{2}$$

[12]The specification includes the frequently studied "unit good" case for: $v(x) = vx \cdot 1_{[0,1]}(x)$ and $c(x) = cx + \bar{c} \cdot 1_{(1,\infty)}(x)$ with $v, c \in \mathbb{R}$ and $\bar{c} > 0$ large.

[13]All results go through essentially unchanged if also the seller's costs c depend on θ and ω.

and I denote the first-best surplus by

$$Z^* = \sum_{\theta,\omega} r(\theta) p_\theta(\omega)[v_{\theta\omega}(x^*_{\theta\omega}) - c(x^*_{\theta\omega})]. \tag{3}$$

At the outset, the agent privately observes his type θ, whereas he cannot observe the state ω. However, the principal (and only the principal) can provide the agent with information about ω.[14] For example, the principal may offer product samples or give the agent more or less time to inspect and try out the product. I assume that whatever the agent learns and infers from this information is not verifiable and the buyer's private information.

Formally, the principal can design any information structure that releases signals to the agent. An information structure consists of a set S of signals and conditional signal distributions $\Pi = (\pi_\omega(s))_{\omega,s} \in \Delta(S)^{\bar{\omega}+1}$ where $\pi_\omega(s)$ is the conditional probability that signal s occurs, conditional on ω. I focus on the case that the principal can release at most countably many signals, and let $S = \mathbb{Z}$ be the set of all integers.[15] Having fixed S, I refer to Π as a (pure) information structure.

The novelty of my approach is that I allow the principal to randomize among information structures. Again, I focus on the case that the principal can randomize over at most countably many information structures. Formally, let $K = \mathbb{Z}$, and let Π_k be an information structure for $k \in K$, with $\pi_{\omega k}(s)$ denoting the conditional probability that signal s is observed, conditional on ω and k. The principal may (commit to) select information structures according to any distribution $\mu \in \Delta(K)$ where $\mu(k)$ is the probability with which information structure Π_k is selected.[16] I denote the resulting (mixed) information structure by (Π, μ).

In addition to the information structure, the principal designs a mechanism that specifies the terms of trade. More precisely, 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 (Π, μ) and a mechanism.

3. The agent decides to accept or reject.

[14]Notice that the full support assumption rules out that (some type of) the agent knows ω for sure at the outset.

[15]Restricting the set of signals to at most countably many strengthens my efficiency results below. For the impossibility result in Proposition 3 below, I will allow for fully general spaces S and K.

[16]That the seller can commit to a probability distribution over information structure is (a sometimes implicit) standard assumption in the information design or Bayesian persuasion literature.

– If the agent rejects, every party gets their outside option of 0.

4. If the agent accepts, Π_k is selected with $\mu(k)$, unobserved by the agent; and the agent privately observes a signal s generated by the information structure Π_k.

5. The terms of trade are enforced according to the mechanism.

It is useful to relate my setting to the literature. When the ex ante type is payoff-irrelevant so that the valuation depends only on the state, my model corresponds to a discrete type version of Li and Shi (2017). In this case, the fact that the principal controls information about the state means that she can disclose information about the agent's final valuation. In contrast, when the ex ante type is payoff-relevant, the principal controls information only about an aspect of what makes up the agent's final valuation.[17] In particular, if the type is orthogonal to the state, that is, $p_\theta = p_{\theta'}$ for all θ, θ', the principal can only disclose an orthogonal part of the agent's valuation. In this case, my model can be seen as a single agent discrete version of Esö and Szentes (2007a,b).

3.1 Mechanisms and principal's problem

The principal's objective is to design an information structure and a mechanism to maximize her profits. 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 $\hat{\theta}$ about his ex ante type after stage 3 and a report $\hat{s}$ about the signal observed after stage 4. I refer to $\hat{\theta}$ as an ex ante report and $\hat{s}$ as an ex post report.[18]

In addition, I allow for rich contracting protocols and allow the mechanism to condition the terms of trade on the information structure k that has been realized. In other words, the information structure is verifiable ex post. One way to think about this is that the principal delegates the choice of information structure to a trusted, disinterested third party. The third party observes

[17]For example, if the buyer of a house exclusively cares about the features of the house, his valuation can, in principle, be fully disclosed by a real estate agent. But if the buyer also cares about how many of his friends live in the neighbourhood, his valuation depends on more than what the real estate agent discloses about the features of the house.

[18]It is common in the literature (notably Li and Shi 2017), to also allow the information structure to depend on a report by the agent about his type, or even on the true type. I will show in an extension that allowing for "report-contingent" information structures does not improve the principal's profit, but that allowing for "report- and type-contingent" does.

the randomization, and after the agent has made his reports, discloses k truthfully. Below I argue that my results go through if the mechanism can condition only on a report by the principal about k, and the parties can employ a budget breaker.

Consequently, a mechanism consists of contingent quantities $x : \Theta \times S \times K \to \mathbb{R}_+$ and contingent transfers $t : \Theta \times S \times K \to \mathbb{R}$, where $x(\hat{\theta}, \hat{s}, k)$ (resp. $t(\hat{\theta}, \hat{s}, k)$) denotes the quantity produced (resp. transfer paid) if the agent reports $\hat{\theta}$ and $\hat{s}$, and the information structure is Π_k.

To express the principal's problem formally, I denote by $u(\theta, s; \hat{\theta}, \hat{s})$ agent type θ's expected utility from reporting $\hat{s}$ ex post, conditional on having reported $\hat{\theta}$ ex ante and having observed s ex post (provided the probability of (θ, s) is positive). Moreover, let $U_{\theta,\hat{\theta}}$ be the expected utility of agent type θ from reporting $\hat{\theta}$ ex ante, that is,

$$U_{\theta,\hat{\theta}} = \sum_{\omega,k,s} p_\theta(\omega)\mu(k)\pi_{\omega k}(s) \max_{\hat{s}} [v_{\theta\omega}(x(\hat{\theta}, \hat{s}, k)) - t(\hat{\theta}, \hat{s}, k)]. \tag{4}$$

The principal's problem is then to choose an information structure (Π, μ) and a mechanism (x, t) so as to maximize

$$\sum_{\theta,\omega,k,s} r(\theta) p_\theta(\omega)\mu(k)\pi_{\omega k}(s)[t(\theta, s, k) - c(x(\theta, s, k))] \quad s.t. \tag{5}$$

$$u(\theta, s; \theta, s) \geq u(\theta, s; \theta, \hat{s}) \quad \forall \theta, s, \hat{s}. \tag{6}$$

$$U_{\theta,\theta} \geq U_{\theta,\hat{\theta}} \quad \forall \theta, \hat{\theta}. \tag{7}$$

$$U_{\theta,\theta} \geq 0 \quad \forall \theta. \tag{8}$$

The first constraint is referred to as the ex post incentive compatibility constraint which ensures that the agent reports the signal truthfully ex post. Notice that the revelation principle requires truthful reporting of the signal only "on the path", that is, after a truthful ex ante report. The second constraint is the ex ante incentive compatibility constraint which ensures that the agent reports his type truthfully ex ante. The third constraint is the individual rationality constraint which ensures that all types accept the mechanism.[19]

For a given information structure and a mechanism, I also define type θ's utility from reporting $\hat{\theta}$ ex ante and then reporting s truthfully ex post as

$$V_{\theta,\hat{\theta}} = \sum_{\omega,k,s} p_\theta(\omega)\mu(k)\pi_{\omega k}(s)[v_{\theta\omega}(x(\hat{\theta}, s, k)) - t(\hat{\theta}, s, k)]. \tag{9}$$

[19]As usual, since the agent's outside option can be replicated in the mechanism by producing and charging nothing, it is without loss of generality optimal for the principal to induce all types to accept the mechanism.

This corresponds to the agent's utility in a situation where the signal s is publicly verifiable and the mechanism can directly condition on s without the need to elicit it from the agent. I therefore say that a mechanism is incentive compatible with observable signal if $V_{\theta,\theta} \geq V_{\theta,\hat{\theta}}$ for all $\theta, \hat{\theta}$.

4 Main results

4.1 Information structure

I shall now define an information structure that I use below to construct full surplus extracting first-best mechanisms. Let Π_k^0 be defined by

$$\pi_{\omega k}^0(s) = \begin{cases} 1 & if \quad s = \omega + k, \\ 0 & else \quad . \end{cases} \tag{10}$$

For $\mu \in \Delta(K)$ with $\mu(k) > 0$ for all k, I denote the resulting mixed information structure by (Π^0, μ). Under the mixed information structure (Π^0, μ), if the state is ω and the information structure is Π_k^0, then the signal $s = \omega + k$ is released with probability 1. This has two important implications: First, knowing s and k reveals that the true state is $\omega = s - k$. Second, suppose that agent type θ has observed signal s. Then he assigns positive probability to k having occurred if and only if $s - k \in \Omega$, that is, for all θ:[20]

$$Pr(k \mid \theta, s) > 0 \quad \Longleftrightarrow \quad s - k \in \Omega. \tag{12}$$

I say that in this case, s and k are consistent with one another, while they are inconsistent otherwise. Clearly, if the agent's ex post report $\hat{s}$ deviates from truth-telling, this is "detected" as a lie if the true k turns out to be inconsistent with the reported $\hat{s}$, that is, $\hat{s} - k \notin \Omega$. The next lemma, while straightforward, is key to my results.

[20]The "only if"-part follows directly from the definition of the information structure. The "if"-part follows from my assumption that that $p_\theta(\omega) > 0$ for all θ and ω, and that $\mu(k) > 0$ for all $k \in K$. Indeed, we have $Pr(\theta, s) = \sum_{\omega \in \Omega, k \in K} p_\theta(\omega) \pi_{\omega k}^0(s) \mu(k)$, which is positive for all θ and s, as $p_\theta(\omega) > 0$ for all θ and ω, and $\mu(k) > 0$ for all $k \in K$. Hence, by Bayes' rule:

$$Pr(k \mid \theta, s) = \frac{\sum_{\omega \in \Omega} p_\theta(\omega) \pi_{\omega k}^0(s) \mu(k)}{Pr(\theta, s)}. \tag{11}$$

Because $p_\theta(\omega) > 0$ for all θ and ω, and $\mu(k) > 0$ for all $k \in K$, the enumerator is positive if and only if $s - k \in \Omega$.

Lemma 1 *For all signals s and all reports $\hat{s} \neq s$, there is a $\kappa \in K$ so that s is consistent with κ, but $\hat{s}$ is not.*

Together with (12), the lemma implies that the agent expects any deviation from truth-telling to be detected as a lie with positive probability. As a consequence, by penalizing reports sufficiently harshly if they are detected as a lie, the agent is induced to report his ex post signal truthfully. This is made more precise in the next lemma.

Lemma 2 *Let (Π^0, μ) be given. Consider a mechanism (x, t) which is incentive compatible with observable signal. Then there are payments $\tilde{t}$ so that the mechanism $(x, \tilde{t})$*

(i) induces ex post truth-telling for all ex ante reports $\hat{\theta}$,

(ii) is ex ante incentive compatible, and

(iii) induces for any type θ the same expected payments as does the mechanism (x, t), when the signal is observable, that is:

$$\sum_{\omega,k,s} p_\theta(\omega)\mu(k)\pi^0_{\omega k}(s)t(\theta, s, k) = \sum_{\omega,k,s} p_\theta(\omega)\mu(k)\pi^0_{\omega k}(s)\tilde{t}(\theta, s, k). \tag{13}$$

In particular, both parties receive the same expected utility under the two mechanisms.

The intuition behind the construction of the payments $\tilde{t}$ rests on the previous considerations. If the agent's report $\hat{s}$ is consistent with the true k, then the terms of trade from the mechanism (x, t) are implemented. Otherwise, the agent is penalized, and since for all s, he attaches positive probability to be penalized if he lies and reports $\hat{s} \neq s$, it is optimal to tell the truth. Thus, the outcome of the mechanism with observable signal is replicated even if the signal is the agent's private information.

Lemma 2 is inspired by a similar result in Zhu (2017) for the case with multiple agents. The difference is that in Zhu's construction, k corresponds to a signal released to another agent. The mechanism can then elicit the signals at no cost from each agent by cross-checking the consistency of the agents' reports.

The lemma has some similarity with the irrelevance results of Esö and Szentes (2007a,b, 2017) who show that in a setting with continuous types, private ex post information which is orthogonal to the agent's ex ante information can be elicited without cost. The forces behind the results are, however, very different. Esö and Szentes (2007a,b, 2017) exploit the first-order (or local) approach which implies that the mechanism needs only to deter "local" deviations from truth-telling. This implies that the agent's additional gains from an ex post lie are of "second order".[21] In contrast, Lemma 2 exploits verifiable information that is complementary to the agent's signal to elicit the signal at no cost. This is possible even though, in contrast to Esö and Szentes (2007a,b, 2017), the signal is not orthogonal to the ex ante type, and ex ante types are not continuous.

4.2 Full surplus extraction

I now state the main result of the paper.

Proposition 1 *Let* (Π^0, μ) *be given.*

(i) Suppose the ex ante type is payoff-irrelevant. Then there is a mechanism which implements the first-best, and the seller fully extracts the surplus.

(ii) Suppose the ex ante type is payoff-relevant. If the set of the agent's prior beliefs $\{p_\theta \mid \theta \in \Theta\}$ *satisfies the "spanning condition", that is, no* $p_{\tilde{\theta}}$ *is in the convex hull of the other* p_θ*'s,* $\theta \neq \tilde{\theta}$*, then there is a mechanism which implements the first-best, and the seller fully extracts the surplus.*

The intuition behind the proposition is as follows. By Lemma 2, it is sufficient to construct a mechanism that implements the first-best and extracts the full surplus for the case that the signal s is observable. Now, since s and k reveal that the true state is $\omega = s - k$, this means that the mechanism can effectively condition on the true state directly.

If the ex ante type is payoff-irrelevant, it is therefore as if the agent has no (relevant) private information at all, and one can simply implement the first-best quantity x^*_ω and charge the agent his valuation $v_\omega(x^*_\omega)$ if state ω is revealed. This mechanism attains the first-best and extracts the full surplus.

In contrast, if the ex ante type is payoff-relevant, the mechanism needs to elicit the agent's type θ to attain the first-best. The crucial observation is the following. Because the state ω is de facto verifiable, the state may also serve the role of an ex post verifiable signal which—unless the type is orthogonal to the state—is correlated with the agent's type. But if the principal has access

[21] See Krähmer and Strausz (2015) for a discussion why Esö and Szentes' (2007a,b) irrelevance result my fail in a setting with discrete types.

to such a signal, and the correlation obeys the spanning condition, then she can elicit the agent's type at no cost which is well-known from the literature on mechanism design with correlated types (Riordan and Sappington, 1988, Cremer and McLean, 1988).

5 More general information structures

5.1 Non-verifiable information structures

So far, I have assumed that the information structure k is verifiable so that the mechanism can directly condition on it. I now consider the case that only the principal privately observes k. I extend the notion of a mechanism by requiring that also the principal submit a report $\hat{k}$ about k. More precisely, after stage 4 in the time-line above, the principal and the agent now simultaneously report $\hat{k}$ and $\hat{s}$ respectively. Moreover, I assume that the principal observes the agent's report about θ after stage 3.[22]

Hence, a mechanism now induces a game between the principal and the agent, and I assume that the parties play a Perfect Bayesian Equilibrium (PBE). By the revelation principle, I can again restrict attention to mechanisms for which truth-telling by both parties is a PBE (conditional on ex ante truth-telling by the agent). The next proposition shows that Proposition 1 essentially carries over.

Proposition 2 *Let* (Π^0, μ) *be given, and let* k *be the principal's private information.*

(i) Suppose the ex ante type is payoff-irrelevant. Then there is a mechanism which implements the first-best, and the seller fully extracts the surplus.

(ii) Suppose the ex ante type is payoff-relevant. If the set of the agent's prior beliefs $\{p_\theta \mid \theta \in \Theta\}$ *satisfies the "spanning condition", that is, no* $p_{\tilde{\theta}}$ *is in the convex hull of the other* p_θ*'s,* $\theta \neq \tilde{\theta}$*, then there is a mechanism which implements the first-best, and the seller fully extracts the surplus.*

The mechanisms used under (i) and (ii) penalize both the agent and the principal if their reports $\hat{s}$ *and* $\hat{k}$ *are inconsistent with one another, that is,* $\hat{s} - \hat{k} \notin \Omega$ *(which means that one party must have*

[22] This, in general, makes it more difficult to induce truth-telling by the principal, but is not substantial in the present context.

deviated from truth-telling). Hence, the mechanisms are not budget-balanced off the equilibrium path.

The basic idea is to induce truth-telling by cross-checking the parties' reports and penalizing both parties if their reports $\hat{s}$ and $\hat{k}$ are inconsistent with one another, that is, $\hat{s}-\hat{k} \notin \Omega$. Given the principal reports truthfully, truth-telling is a best response for the agent for the same reasons as in the case with verifiable k. A similar logic applies to the principal. If the principal observes k, she revises her beliefs and assigns positive probability to s having occurred if and only if $s-k \in \Omega$. Analogously to Lemma 1, it can be shown that for any k and $\hat{k} \neq k$, there is a signal s which is consistent with k yet not with $\hat{k}$. Therefore, given that the agent reports the signal truthfully, the principal expects any deviation from truth-telling to be inconsistent with the agent's report with positive probability. As a consequence, by penalizing inconsistent reports sufficiently harshly, the principal is induced to report k truthfully.

5.2 Contingent information structures

The information structure considered so far has the feature that it does not depend on an ex ante report by the agent about his type, or on the type itself. In this section, I relax this feature to address the question whether, if the principal is allowed to use a more general information structure, full surplus extraction is possible even if the agent's beliefs violate the spanning condition.

I say, an information structure Π_k is report-contingent[23], if it depends on an ex ante report $\hat{\theta}$ by the agent, and I denote by $\pi_{\omega k}(s;\hat{\theta})$ the probability that a signal s is generated conditional on ω and k when the agent reports $\hat{\theta}$. I say an information structure Π_k is report- and type-contingent[24] if it depends on an ex ante report $\hat{\theta}$ by the agent and also directly on his true type θ. I denote by $\pi_{\omega k}(s;\hat{\theta},\theta)$ the probability that a signal s is generated conditional on ω and k when the agent is of type θ and reports $\hat{\theta}$.[25]

[23]Li and Shi (2017) refer to this case as "discriminatory disclosure".

[24]Li and Shi (2017) refer to this case as "general disclosure".

[25]As an economic example for a report- and type-contingent information structure, consider a good that consists of various attributes $\theta \in \{1,\dots,\bar{\theta}\}$. The agent cares only about exactly one attribute θ which corresponds to his privately known type. The agent's valuation for the good depends in addition on an unknown state $\omega = (\omega_1,\dots,\omega_{\bar{\theta}})$, and is given by $v_{\theta\omega} = \phi(\omega_\theta,\theta)$ for some function ϕ. Now consider the following disclosure policy by the principal. If the agent announces $\hat{\theta}$, then some information about $\omega_{\hat{\theta}}$ is disclosed to the agent, but none about ω_θ, $\theta \neq \hat{\theta}$.

I first show that allowing for report-contingent information structures does, in general, not help to relax the spanning condition. More precisely, I show that for a given set of beliefs that violates the spanning condition, there is a specification of the agent's valuation for which full surplus extraction is not possible. To make this impossibility result (somewhat) stronger, I now allow the sets S and K to be arbitrary measure spaces endowed with σ-algebras and denote by $\pi_{\omega k}(\cdot;\hat{\theta}) \in \Delta(S)$ and $\mu \in \Delta(K)$ the respective probability measures that capture the conditional signal distribution and the principal's randomization strategy.

Proposition 3 *Suppose the ex ante type is payoff-relevant. If the set of the agent's prior beliefs* $\{p_\theta \mid \theta \in \Theta\}$ *violates the spanning condition, then there are valuations* $v_{\theta\omega}$ *so that for any report-contingent information structure, any mechanism that implements the first-best leaves an information rent to some buyer type* θ.

The argument is a straightforward adaptation of the analogous argument in the proof of Theorem 2 in Cremer and McLean (1988). Intuitively, suppose that the agent's ex ante private information is the result of the agent observing some ex ante signal θ which is correlated with the state, and his beliefs p_θ are the posterior beliefs about the state, conditional on the signal. Failure of the spanning condition then intuitively means that there is one signal $\tilde{\theta}$ that can be generated by pooling some other signals $\theta \neq \tilde{\theta}$. In this sense, the belief $p_{\tilde{\theta}}$ contains less information than the other beliefs p_θ, $\theta \neq \tilde{\theta}$. Therefore, for some model specification, type $\tilde{\theta}$ does strictly worse than the better informed types $\theta \neq \tilde{\theta}$, making it impossible that all types receive utility of 0.

An important class of settings where the spanning condition is violated is the Esö and Szentes (2007a,b) type settings where the type is orthogonal to the state so that p_θ does not, in fact, depend on θ. As the previous proposition shows, full surplus extraction is not guaranteed in these cases.

Next, I allow for report- and type-contingent information structures. In contrast to the preceding result, I show that irrespective of the agent's beliefs, this allows the principal to fully extract the first-best surplus.

Proposition 4 *There is a report- and type-contingent information structure and a mechanism which implements the first-best, and the principal fully extracts the surplus.*

Hence, if the agent announces $\hat{\theta}$, he receives information about his valuation if his true type is $\hat{\theta}$ but no information if his true type is $\theta \neq \hat{\theta}$. Hence, the information he receives depends on his true type.

The intuition is straightforward. When the information structure can be conditioned on the true type, then the principal can release, for all states ω and information structures k, the same signal $s_0 \in S$ ex post if the agent reports a type which differs from the true type. This not only gives the agent no ex post information about the state, it also gives him no ex post information about the information structure k. The idea is now to impose a penalty on the agent if he reports a signal which is inconsistent with k. Hence, if the agent lies about θ, he will receive no additional information about ω and k, and so for any report $\hat{s}$, he expects to be penalized with positive probability. For sufficiently large penalty, this will deter the agent from misreporting ex ante. Thus, the agent's type can be elicited without leaving information rents to him.

6 Conclusion

In this paper I show that in a large class of cases, a monopolistic seller can design additional product information for a privately informed buyer in a way to extract the full first-best surplus. The basic idea is to allow the seller to randomize over information structures and to employ a rich contracting protocol which conditions the terms of trade on (reports about) the outcome of the randomization.

The same idea can be extended to mechanism design settings with multiple agents, such as an auction. In this case, the mechanism does not need to condition directly on the information structure, but the agents' signals can be elicited by cross-checking their reports. This is explored in a companion paper (Krähmer, 2017).

7 Proofs

Proof of Lemma 1 Let $s, \hat{s}$ with $\hat{s} \neq s$ be given. If $\hat{s} < s$, then for $\kappa = s$, we have that $s - \kappa = 0 \in \Omega$ but $\hat{s} - \kappa < 0 \notin \Omega$. If $\hat{s} > s$, then for $\kappa = s - \bar{\omega}$, we have that $s - \kappa = \omega \in \Omega$ but $\hat{s} - \kappa > \bar{\omega} \notin \Omega$. qed

Proof of Lemma 2 Define the payments $\tilde{t}$ as

$$\tilde{t}(\theta, s, k) = \begin{cases} t(\theta, s, k) & if \quad s - k \in \Omega \\ T & if \quad s - k \notin \Omega \end{cases} \tag{14}$$

for some $T > 0$.

As to (i). To see that the mechanism $(x, \tilde{t})$ induces ex post truth-telling, suppose agent type θ has reported $\hat{\theta}$ and observed s. By (12),

$$Pr(k \mid \theta, s) > 0 \quad \Leftrightarrow \quad s - k \in \Omega. \tag{15}$$

Hence, the agent's utility from reporting $\hat{s} = s$ is

$$u(\theta, s; \hat{\theta}, s) = \sum_{k: s-k \in \Omega} Pr(k \mid \theta, s)\{E_{\tilde{\omega}}[v_{\theta\tilde{\omega}}(x(\hat{\theta}, s, k)) \mid \theta, s, k] - \tilde{t}(\hat{\theta}, s, k)\}. \tag{16}$$

Note that this expression is independent of T, because the sum is only over indices k for which $s - k \in \Omega$.

On the other hand, by Lemma 1, for any $\hat{s} \neq s$, there is κ so that $s - \kappa \in \Omega$ but $\hat{s} - \kappa \notin \Omega$. Thus, if the agent reports $\hat{s} \neq s$, then with (at least) probability $Pr(\kappa \mid \theta, s) > 0$, he has to make the payment

$$\tilde{t}(\hat{\theta}, \hat{s}, \kappa) = T, \tag{17}$$

and his utility is

$$\begin{aligned} u(\theta, s; \hat{\theta}, \hat{s}) \quad = \quad & \sum_{k: s-k \in \Omega, k \neq \kappa} Pr(k \mid \theta, s)\{E_{\tilde{\omega}}[v_{\theta\tilde{\omega}}(x(\hat{\theta}, \hat{s}, k)) \mid \theta, s, k] - \tilde{t}(\hat{\theta}, \hat{s}, k)\} & (18) \\ & + Pr(\kappa \mid \theta, s)\{E_{\tilde{\omega}}[v_{\theta\tilde{\omega}}(x(\hat{\theta}, \hat{s}, \kappa)) \mid \theta, s, \kappa] - T\}, & (19) \end{aligned}$$

which becomes smaller than $u(\theta, s; \hat{\theta}, s)$ when T gets large. This shows that if T is sufficiently large, then after any ex ante report $\hat{\theta}$, agent type θ reports s truthfully under the mechanism $(x, \tilde{t})$.

(ii) Next, I argue that $(x, \tilde{t})$ induces ex ante truth-telling. Indeed, because $(x, \tilde{t})$ induces ex post truth-telling by the previous step, and because t and $\tilde{t}$ coincide conditional on ex post truth-telling, agent type θ's first period utility $U_{\theta,\hat{\theta}}$ from reporting $\hat{\theta}$ under $(x, \tilde{t})$ coincides with his utility $V_{\theta,\hat{\theta}}$ if the signal s is publicly verifiable. Because $V_{\theta,\theta} \geq V_{\theta,\hat{\theta}}$ for all $\theta, \hat{\theta}$ by assumption, $(x, \tilde{t})$ is ex ante incentive compatible.

(iii) By the same token as in (ii), because $(x, \tilde{t})$ induces ex post truth-telling by step (i), and because t and $\tilde{t}$ coincide conditional on ex post truth-telling, and because $(x, \tilde{t})$ induces ex ante truth-telling, the claim about the expected payments follows. This shows (iii) and completes the proof. qed

Proof of Proposition 1 Define the allocation rule:

$$x(\theta,s,k)=\begin{cases} x^*_{\theta,s-k} & if \quad s-k\in\Omega, \\ 0 & if \quad s-k\notin\Omega. \end{cases} \tag{20}$$

To define transfers, consider the auxiliary problem over the choice variable $z:\Theta\times\Omega\to\mathbb{R}$:

$$P: \qquad \max_{z(\theta,\omega)} \quad \sum_{\theta,\omega} r(\theta)p_\theta(\omega)[z(\theta,\omega)-c(x^*_{\theta\omega})] \quad s.t. \tag{21}$$

$$\sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(x^*_{\theta\omega})-z(\theta,\omega)] \geq \sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(x^*_{\hat{\theta}\omega})-z(\hat{\theta},\omega)] \quad \forall\theta,\hat{\theta} \tag{22}$$

$$\sum_{\omega} p_\theta(\omega)[v_{\theta\omega}(x^*_{\theta\omega})-z(\theta,\omega)]=0 \quad \forall\theta. \tag{23}$$

Problem P corresponds to a (static) principal agent problem, where the type θ is the agent's private information at the contracting stage, and given the allocation rule x^*, the principal chooses payments z which can condition both on a report $\hat{\theta}$ and on an ex post verifiable signal ω. (The signal is correlated with the agent's type unless the type is orthogonal to the state).

Now, if the type is payoff-irrelevant so that $v_{\theta\omega}$ and $x^*_{\theta\omega}$ do not depend on θ, then a solution to problem P is given by $z(\theta,\omega)=z(\omega)=v_\omega x^*_\omega$. If the type is payoff-relevant, then, as shown by Riordan and Sappington (1988), there is a solution $z(\theta,\omega)$ to this problem if the beliefs p_θ satisfy the spanning condition. In either case, the value of the problem is the first-best surplus Z^*.

Now define

$$t(\theta,s,k)=\begin{cases} z(\theta,s-k) & if \quad s-k\in\Omega, \\ T & if \quad s-k\notin\Omega, \end{cases} \tag{24}$$

for some $T>0$.

I now show that the mechanism (x,t) is incentive compatible and individually rational, and yields the principal the first-best surplus Z^*. To show incentive compatibility, it is sufficient by Lemma 2 to show that (x,t) is incentive compatible with observable signal. To see this, we compute $V_{\theta,\hat{\theta}}$. Recall that

$$V_{\theta,\hat{\theta}} \quad = \quad \sum_{\omega,s,k} p_\theta(\omega)\mu(k)\pi^0_{\omega k}(s)[v_{\theta\omega}(x(\hat{\theta},s,k))-t(\hat{\theta},s,k)]. \tag{25}$$

To understand the sum, fix s and ω and consider the summation over k. By definition of (Π^0,μ), $\pi^0_{\omega k}(s)=1$ if $k=s-\omega$ and $\pi^0_{\omega k}(s)=0$ for all $k\neq s-\omega$. Moreover, by definition of the mechanism, we have for $k=s-\omega$:

$$x(\hat{\theta},s,k)=x^*_{\hat{\theta},\omega}, \quad and \quad t(\hat{\theta},s,k)=z(\hat{\theta},\omega). \tag{26}$$

Therefore, we obtain:

$$\sum_k p_\theta(\omega)\mu(k)\pi^0_{\omega k}(s)[v_{\theta\omega}(x(\hat{\theta},s,k)) - t(\hat{\theta},s,k)] = p_\theta(\omega)\mu(s-\omega)\cdot[v_{\theta\omega}(x^*_{\hat{\theta},\omega}) - z(\hat{\theta},\omega)]. \quad (27)$$

Hence, when now summing over s and ω, we obtain that

$$V_{\theta,\hat{\theta}} = \sum_\omega p_\theta(\omega)[v_{\theta\omega}(x^*_{\hat{\theta},\omega}) - z(\hat{\theta},\omega)]\left(\sum_s \mu(s-\omega)\right) = \sum_\omega p_\theta(\omega)[v_{\theta\omega}(x^*_{\hat{\theta},\omega}) - z(\hat{\theta},\omega)]. \quad (28)$$

By inspection, $V_{\theta,\hat{\theta}}$ coincides with the function that appears on the right hand side of (22), and $V_{\theta,\theta}$ coincides with the functions that appear on the left hand side of (22) and (23). Since z is a solution to P, it follows from (22) that $V_{\theta,\theta} \geq V_{\theta,\hat{\theta}}$ for all $\hat{\theta},\theta$, and hence the mechanism (x,t) is incentive compatible with observable signals, as we wanted to show.

To see that the mechanism is individually rational and delivers the first-best surplus, observe that $V_{\theta,\theta} = 0$ by (23), and since the quantities are the first-best quantities, the principal obtains the first-best surplus. qed

Proof of Proposition 2 Consider the same mechanism (x,t) as in the proof of Proposition 1 with the difference that k, rather than the true information structure, is a report by the principal. The only adjustment made to mechanism is that it imposes a fine $F >> 0$ on the principal if $s-k \notin \Omega$.

I now show that truth-telling by both parties is a PBE. That truth-telling (ex ante and ex post) by the agent is a best response to truth-telling by the principal follows from the proof of Proposition 1. Suppose now that the agent tells the truth (ex ante and ex post). Suppose the principal has observed the agent's ex ante report θ, and she has observed k. Let $Pr(s \mid k,\theta)$ be the probability she attaches to the agent having observed s, conditional on k and the (truthful) report θ. Because $p_\theta(\omega) > 0$ for all ω, we have that

$$Pr(s \mid k,\theta) > 0 \quad \Leftrightarrow \quad s-k \in \Omega. \quad (29)$$

Hence, the principal's utility from reporting $\hat{k} = k$ is

$$\sum_{s:s-k\in\Omega} Pr(s \mid k,\theta)\{t(\theta,s,k) - E_{\tilde{\omega}}[c(x(\theta,s,k)) \mid k]\}. \quad (30)$$

Note that this expression is independent of F, because the sum is only over indices s for which $s-k \in \Omega$.

On the other hand, for any $\hat{k} \neq k$, there is $\sigma \in S$ so that $\sigma - k \in \Omega$ but $\sigma - \hat{k} \notin \Omega$.[26] Thus, if the principal reports $\hat{k} \neq k$, then with (at least) probability $Pr(\sigma \mid k, \theta) > 0$, she has to make the payment F, and hence her utility becomes smaller than (30) if F is sufficiently large. This shows that truth-telling is a best response by the principal to truth-telling by the agent.

That the mechanism implements the first-best and delivers the principal the full surplus now follows as in the proof of Proposition 1. qed

Proof of Proposition 3 Because the spanning condition fails, there are $\tilde{\theta}$ and $\alpha_\theta \in [0,1]$, $\theta \neq \tilde{\theta}$, with $\sum_{\theta \neq \tilde{\theta}} \alpha_\theta = 1$ so that

$$p_{\tilde{\theta}}(\omega) = \sum_{\theta \neq \tilde{\theta}} \alpha_\theta p_\theta(\omega) \quad \forall \omega. \tag{31}$$

Moreover, consider a valuation function with the property that for all θ, ω:

$$v_{\tilde{\theta}\omega}(x^*_{\tilde{\theta}\omega}) < v_{\theta\omega}(x^*_{\tilde{\theta}\omega}). \tag{32}$$

Towards a contradiction, suppose that the principal can extract the full first-best surplus, then by the revelation principle, the agent reports truthfully ex ante and ex post on the equilibrium path, and in state ω, the first-best quantity $x^*_{\theta\omega}$ must be implemented. Hence, agent type $\tilde{\theta}$'s utility under the mechanism is

$$\begin{aligned}
U_{\tilde{\theta},\tilde{\theta}} &= \sum_{\omega} \int_K \int_S p_{\tilde{\theta}}(\omega)[v_{\tilde{\theta}\omega}(x^*_{\tilde{\theta},\omega}) - t(\tilde{\theta}, s, k)]\, d\pi_{\omega k}(s; \tilde{\theta}) d\mu(k) && (33)\\
&= \sum_{\theta \neq \tilde{\theta}} \alpha_\theta \left(\sum_{\omega} \int_K \int_S p_\theta(\omega)[v_{\tilde{\theta}\omega}(x^*_{\tilde{\theta},\omega}) - t(\tilde{\theta}, s, k)]\, d\pi_{\omega k}(s; \tilde{\theta}) d\mu(k) \right) && (34)\\
&< \sum_{\theta \neq \tilde{\theta}} \alpha_\theta \left(\sum_{\omega} \int_K \int_S p_\theta(\omega)[v_{\theta\omega}(x^*_{\tilde{\theta},\omega}) - t(\tilde{\theta}, s, k)]\, d\pi_{\omega k}(s; \tilde{\theta}) d\mu(k) \right) && (35)\\
&\leq \sum_{\theta \neq \tilde{\theta}} \alpha_\theta U_{\theta,\tilde{\theta}}, && (36)
\end{aligned}$$

where in the second line I have used (31), in the third line I have used (32). To understand the final inequality, notice that the expression in the brackets in (35) is the utility of agent type θ when he (untruthfully) reports $\tilde{\theta}$ ex ante and reports s truthfully ex post. But, because after an untruthful report ex ante, it is not necessarily optimal to report truthfully ex post, this expression

[26] If $\hat{k} < k$, then for $\sigma = k + \bar{\omega}$, we have that $\sigma - k = \bar{\omega} \in \Omega$ but $\sigma - \hat{k} < \bar{\omega} \notin \Omega$. If $\hat{k} > k$, then for $\sigma = k$, we have that $\sigma - k = 0 \in \Omega$ but $\sigma - \hat{k} < 0 \notin \Omega$.

is (weakly) smaller than $U_{\theta,\tilde{\theta}}$ which, by definition, is agent type θ's utility when he (untruthfully) reports $\tilde{\theta}$ ex ante and chooses an optimal report ex post.

Now, because the principal extracts the full surplus, all agent types θ get $U_{\theta,\theta} = 0$. Together with incentive compatibility, this implies that $U_{\theta,\tilde{\theta}} \leq U_{\theta,\theta} = 0$, and hence the inequality above implies that $U_{\tilde{\theta},\tilde{\theta}} < 0$, contradicting individual rationality for type $\tilde{\theta}$. qed

Proof of Proposition 4 Let $S = K = \mathbb{Z}$, $\mu \in \Delta(K)$ with $\mu(k) > 0$ for all k, and for some $s_0 \in S$, define

$$\pi_{\omega k}(s;\theta,\hat{\theta}) = \begin{cases} 1 & if \quad \theta = \hat{\theta} \text{ and } s - k = \omega \\ 1 & if \quad \theta \neq \hat{\theta} \text{ and } s = s_0 \\ 0 & \quad else. \end{cases} \tag{37}$$

Hence, if the agent reports his type truthfully, the information structure coincides with (10), and if he misrepresents his type, he gets the signal s_0 with probability 1 which is therefore entirely uninformative, both about ω and k.

Define the mechanism as follows:

$$x(\theta,s,k) = \begin{cases} x^*_{\theta,s-k} & if \quad s-k \in \Omega \\ 0 & if \quad s-k \notin \Omega \end{cases}, \quad t(\theta,s,k) = \begin{cases} v_{\theta,s-k}(x^*_{\theta,s-k}) & if \quad s-k \in \Omega \\ T & if \quad s-k \notin \Omega \end{cases} \tag{38}$$

for some $T > 0$.

For sufficiently large T, it follows as in the proof of Lemma 2 that the agent reports s truthfully, if he has reported θ truthfully ex ante. Therefore, the definition of payments, and the fact that k and s reveal the true state, implies that the agent obtains utility 0 when he reports θ truthfully.

Next, consider the case that agent type θ falsely reports $\hat{\theta} \neq \theta$ ex ante. Then the agent observes s_0 for sure and chooses an optimal report $\hat{s}$ ex post. Because

$$\hat{s} - k \in \Omega \quad \Leftrightarrow \quad k \in \{\hat{s}, \ldots, \hat{s} + \bar{\omega}\}, \tag{39}$$

the agent, at the ex ante reporting stage, anticipates that with (at least) probability $\sum_{\ell \notin \{\hat{s},\ldots,\hat{s}+\bar{\omega}\}} \mu(\ell) > 0$, he receives a quantity of 0 and has to make payments T. This implies that for sufficiently large T, the agent's expected utility from lying ex ante becomes negative. Because his utility from truth-telling is 0, the agent is deterred from lying.

Moreover, because the mechanism implements the first-best quantities and the agent receives 0 rent, the principal fully extracts the first-best surplus. qed

References

Baron, D. and D. Besanko (1984). Regulation and information in a continuing relationship. Information Economics and Policy, 1, 267–302.

Bergemann, D., A. Bonatti, and A. Smolin (2017). The design and price of information. American Economic Review, forthcoming.

Bergemann, D. and S. Morris (2016). Bayes correlated equilibrium and the comparison of information structures in games. Theoretical Economics, 11(2), 487-522.

Courty, P. and H. Li (2000). Sequential screening. Review of Economic Studies, 67, 697–717.

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

Esö, P. and B. Szentes (2007a). Optimal information disclosure in auctions and the handicap auction. Review of Economic Studies, 74(3), 705-731.

Esö, P. and B. Szentes (2007b). The price of advice. The Rand Journal of Economics, 38(4), 863-880.

Esö, P. and B. Szentes (2017). Dynamic contracting: an irrelevance result. Theoretical Economics, 12, 109-139.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. American Economic Review, 101(6), 2590-2615.

Krähmer, D. (2012). Optimal auction design with endogenously correlated buyer types. Journal of Economic Theory, 147, 118-141.

Krähmer, D. (2017). Correlated information structures and optimal auctions. Working paper, mimeo.

Krähmer, D. and R. Strausz (2015a). Ex post information rents in sequential screening. Games and Economic Behavior, 90, 257-273.

Krähmer, D. and R. Strausz (2015b). “Dynamic mechanism design." Chapter 11 in ed. T. Börgers, An Introduction to the Theory of Mechanism Design, 204–234.

Myerson, R. B. (1982). Optimal coordination mechanisms in generalized principalâĂŞagent problems. Journal of Mathematical Economics, 10(1), 67-81.

Myerson, R. (1986). Multistage games with communication. Econometrica, 54, 323-58.

Li, H. and Shi, X. (2017). Discriminatory information disclosure. American Economic Review, forthcoming.

Obara, I. (2008). The full surplus extraction theorem with hidden actions. The B.E. Journal of Theoretical Economics, 8 (1), (Advances), Article 8.

Rahman, D. (2012). But who will monitor the monitor?. The American Economic Review, 102(6), 2767-2797.

Rahman, D. and I. Obara (2010). Mediated partnerships. Econometrica, 78(1), 285-308.

Riordan, M. H. and D. E. Sappington (1988). Optimal contracts with public ex post information. Journal of Economic Theory, 45(1), 189-199.

Rodina, D. (2017). Information design and career concerns. Working paper, Northwestern University, mimeo.

Strausz, R. (2012). Mediated contracts and mechanism design. Journal of Economic Theory, 147(3), 1280-1290.

Zhu, S. (2017). Private disclosure with multiple agents. Working paper, Toulouse School of Economics, mimeo.