University of Heidelberg

Department of Economics

Discussion Paper Series | No. 577

Bayesian signaling

Jonas Hedlund

November 2014

Bayesian signaling

Jonas Hedlund*

November 2014

Abstract

This paper introduces private sender information into a sender-receiver game of Bayesian persuasion with monotonic sender preferences. I derive properties of increasing differences related to the precision of signals and use these to fully characterize the set of equilibria robust to the intuitive criterion. In particular, all such equilibria are either separating, i.e., the sender's *choice* of signal reveals his private information to the receiver, or fully disclosing, i.e., the *outcome* of the sender's chosen signal fully reveals the payoff-relevant state to the receiver. Incentive compatibility requires the high sender type to use sub-optimal signals and therefore generates a cost for the high sender type in comparison to a full information benchmark in which the receiver knows the sender's type. The receiver prefers the equilibrium outcome over this benchmark for large classes of monotonic sender preferences.

Keywords: Bayesian Persuasion, Signaling.

JEL Classification: *D82, D83, D86.*

1 Introduction

The literature on strategic transmission of information typically focuses on the extent to which a *sender* transmits exogenous private information to a *receiver* when preferences are imperfectly

*Department of Economics, University of Heidelberg. Address: Bergheimer str. 58, 69115, Heidelberg, Germany. Email: jonas.hedlund@awi.uni-heidelberg.de. Phone: +49 6221 54-3171. Fax: +49 6221 54-2997

aligned (e.g., Spence 1973; Milgrom 1981; Crawford and Sobel 1982). Recent pioneering work on *Bayesian persuasion* by Kamenica and Gentzkow (2011) departs from this tradition by instead asking what information the sender would *generate* if he initially is as uninformed as the receiver and commits to reveal all information generated. Applications range from prosecutors gathering evidence for presentation at court to firms specifying the terms of free trials of recently developed products. In both examples there is some discretion with respect to which specific information to generate and a conceivable commitment device ensuring that the generated information is passed on to the receiver. The assumption, however, that sender and receiver are initially equally uninformed, is not always plausible. For example, a firm specifying terms for free trials upon launching a new software (e.g., length or functionality) is likely better informed than a typical consumer about the user-friendliness of this software. Further, such private information may be difficult to credibly disclose except through the consumer's own trials. The question then arises to what extent the design of the trials may in themselves signal something about the firm's private information to the consumer.[1] The present paper shows that such private information in fact unravels and can be inferred from the nature of the evidence the sender collects, of the tests he conducts, or of the trials he offers, even if the private information is itself unverifiable and not subject to standard unraveling arguments (as in Milgrom 1981).

I investigate a simple model of Bayesian persuasion in which the sender has unverifiable and imperfect private information about a binary payoff-relevant state prior to generating further information about this state. Contingent on his private information the sender chooses a signal which is informative about the payoff-relevant state and the receiver observes an outcome of the signal and updates her beliefs. The sender's payoff is continuous and strictly monotonic in the receiver's updated belief.[2] I show that confining attention to equilibria robust to Cho and Kreps' (1987) intuitive criterion (henceforth, "equilibria") leads to a number of predictions about

[1] Similarly, a prosecutor might have observed some evidence which is not admissible for presentation at court, such as information about the interrogation of a spouse who later refuses to testify. The question is then to what extent the prosecutor's private information can be inferred from the evidence that he does present.

[2] Such preferences of "pure persuasion" play an important role in the literature on transmission of verifiable information, see, e.g., Milgrom (1981), Milgrom and Roberts (1986) and Hedlund (2014), as well as in several signaling models, see, e.g., Spence (1973) and Mailath (1987). See also Cámara and Alonso (2014a).

the sender's behavior. The main result provides a full characterization of the equilibria and reveals that private information leads to a form of unraveling (Theorem 1 and Proposition 1). In particular, in equilibrium the sender's private information is always transmitted to the receiver in one of two ways. Any equilibrium is either fully separating, i.e., the sender's *choice* of signal reveals his private information to the receiver, or fully disclosing, i.e., the *outcome* of the sender's chosen signal fully reveals the payoff-relevant state to the receiver. This is true even in situations in which the sender would choose an uninformative signal in the absence of private information. The result is not driven by the presence of a signal disclosing the senders type, as in Milgrom's (1981) unraveling result, but follows from conditions of increasing differences in expected payoffs, which arise endogenously in the model and roughly state that sender types with more favorable private information have stronger preferences for more precise signals (Lemma 1 and 2). The increasing differences imply that much of the logic of standard signaling games applies. Roughly, the intuitive criterion requires the receiver to attribute deviations to some sufficiently precise signals to sender types with favorable private information, making it possible to find profitable deviations for such types from any pooling strategy.

To obtain a tractable model I assume that the sender's private information is binary, leading to a binary type-space with a low (high) sender type with unfavorable (favorable) private information. The main result reveals that each sender type's equilibrium signal solves a maximization problem related to a full information benchmark in which the receiver knows the sender's type. In this benchmark case each sender type would choose a signal maximizing his payoff given a common prior on the payoff-relevant state, i.e., what Kamenica and Gentzkow (2011) call an *optimal signal*. In equilibrium the low sender type chooses any such optimal signal. The equilibrium signal of the high sender type, however, maximizes this type's payoff under an incentive compatibility constraint ensuring that the low type does not have incentives to mimic the high type's signal. In most interesting cases this constraint forces the high type to choose a signal which is not optimal in the full information benchmark (Proposition 3).[3] Incentive compatibility thus generates a cost for the

[3] Specifically, the high type's equilibrium signal is optimal if and only if it fully reveals the payoff relevant state, which in turn is possible only if such signals are optimal for both types.

high sender type, as in many models of costly signaling (e.g., Spence 1973). This cost is constant across equilibria and in particular, while there are sometimes several equilibria, all equilibria are payoff-equivalent from the point of view of the sender. Finally, there is an equilibrium at which the signals of both sender types fully reveal the payoff-relevant state if and only if such a signal is optimal for both types (Proposition 2).

While the sender prefers the full information benchmark in which the receiver knows his type over the equilibrium, the receiver sometimes strictly prefers the equilibrium. The reason is that incentive compatibility sometimes forces the high sender type to use a signal which is strictly more informative than the one he would use in the full information benchmark. This occurs, e.g., if the sender's payoff function is strictly concave in the receiver's updated belief, in which case an uninformative signal is optimal for both sender types (see Kamenica and Gentzkow 2011). This may lead to odd situations in which the sender would like to disclose his private information to the receiver prior to choosing a signal, and the receiver refuses to listen.

Finally, the model offers some predictions with respect to "good news" and "bad news" in equilibrium (Remark 2). First, from the sender's point of view, the most favorable updated receiver belief induced by the low type in equilibrium is always more favorable than the least favorable updated belief induced by the high type. If the indued beliefs of both types do not overlap in this way the low type has an incentive to deviate to the high type's signal. There is therefore a sense in which "good news" from low types are better than "bad news" from high types. Second, and relatedly, an uninformative signal is neither the worst nor the best piece of equilibrium news. This holds since only the low type's equilibrium signal can be uninformative. For the updated beliefs to overlap the high type must consequently induce some belief less favorable than the one corresponding to the low type's silence. Hence, while silence is always attributed to the low sender type, as in Milgrom's (1981) unraveling result, silence is never the worst piece of equilibrium news, in contrast to the unraveling result.

The presence of private information in a framework of Bayesian persuasion generates a signaling game with two main properties which together drive the results cited above. First, while I assume the sender's payoff function constant in type, the expected payoff given a signal and an interim receiver belief is linear in type, since different types generate signal outcomes with different frequencies. This dependence can be structured by conditions of increasing differences consistent with an ordering of the signals according to their precision. Second, whenever the sender uses a signal which fully reveals the payoff-relevant state, his payoff is independent of the receiver's belief prior to observing the signal's outcome. Such signals establish type-specific lower bounds on equilibrium expected payoffs. The increasing differences imply that much of the logic of standard signaling games can be applied, which together with the lower bounds on equilibrium payoffs leads to the characterizations of equilibria emphasized here.

Related literature. Surprisingly few papers in the literature following the seminal contribution of Kamenica and Gentzkow (2011) investigate the implications of private sender information. An exception is the paper by Perez-Richet (2014), whose approach, however, differs somewhat from the one here. First, Perez-Richet (2014) considers sender preferences which are constant in the receiver's updated beliefs except for a single discontinuity. Second, while Perez-Richet (2014) assumes the sender perfectly informed about the payoff-relevant state, here the sender is imperfectly informed. Interestingly, these differences allow Perez-Richet (2014) to restrict attention to pooling equilibria, in contrast to the prominence of separation here. The concurrent working paper by Alonso and Câmara (2014b) is technically more related. Their framework nests mine by allowing more general sender preferences, a finite set of payoff-relevant states and a finite type space. However, Alonso and Câmara (2014b) only address the question of whether the sender may obtain higher expected payoff under private information than in its absence (providing a negative answer) and do not attempt further characterizations of equilibria. The present paper shows that a more structured model leads to concrete insights in more limited, but still fairly interesting environments. Alonso and Câmara's (2014b) result is related to my result on the sub-optimality of equilibrium signals, but less specific and slightly different since it compares the sender's equilibrium payoff with a benchmark in which the sender obtains no private information, while my result concerns a benchmark in which the receiver observes the sender's type. Kolotilin (2014a) discusses the impact of verifiable private sender information in a framework of Bayesian persuasion and argues that such information unravels a la Milgrom (1981). In contrast, the present paper emphasizes that unverifiable private sender information unravels in equilibrium and through a very different mechanism. Alonso and Câmara (2014a) analyze Bayesian persuasion when sender and receiver have different prior beliefs. By not modeling the process through which disagreement occurs, however, Alonso and Câmara (2014a) abstract from the incentive compatibility issues in focus here.

Some recent papers have in common with the present one that they investigate situations in which the sender's control of the receiver's information is somewhat weakened. Gentzkow and Kamenica (2012) find that competition between senders may increase the amount of information transmitted. Gentzkow and Kamenica (2014) study costly Bayesian persuasion. Kolotilin (2014b) considers a privately informed receiver.

Another strand of literature analyzes the design of public signals of privately informed senders when, in contrast to the present paper, the set of feasible signals is constrained. Gill and Sgroi (2012) analyze binary pre-launch tests of a privately and perfectly informed monopolist. The precision of tests is fixed and the monopolist's choice is constrained to a "toughness" parameter. In contrast to the results here, all equilibria are pooling. Gill and Sgroi (2008), Rayo and Segal (2010) and Li and Li (2012) impose related constraints on the set of feasible signals.

Finally, this paper is related to the literature on information acquisition followed by information transmission, including the work of Austen-Smith (2000), Henry (2009), Che and Kartik (2009), Ivanov (2010) and Argenziano, Severinov and Squintani (2014). In these papers the sender cannot commit to reveal the information acquired and is also typically constrained in the set of signals he can choose, leading to an analysis substantially different from the one here.

2 The model

Payoff-relevant states and types. There are two players, *sender* and *receiver*. The payoff-relevant states of the world are $\{\omega_L, \omega_H\}$. Sender and receiver agree that the prior probability of ω_H equals $\mu_0 \in (0,1)$. Before anything else occurs, the unverifiable outcome of a binary informative signal is privately revealed to the sender, who rationally updates μ_0 to either μ_L or μ_H, with $0 < \mu_L < \mu_0 < \mu_H < 1$. I refer to $\mu_j \in \{\mu_L, \mu_H\}$ as the sender's *type*.

Signals and beliefs. After observing his private signal the sender's objective is to modify the receiver's prior belief μ_0. In order to accomplish this the sender chooses a signal $\pi = (\pi(\cdot|\omega_L), \pi(\cdot|\omega_H))$, consisting of a pair of conditional probability distributions $\pi(\cdot|\omega_L)$ and $\pi(\cdot|\omega_H)$ over a finite set of outcomes $\widetilde{S}$. Let Π be the set of all signals. For a generic signal π let $S = \{s_1, ..., s_k\}$ be the *support* of π, defined as $S := \{s \in \widetilde{S} : \pi(s|\omega_L) + \pi(s|\omega_H) > 0\}$.[4] I abbreviate $\pi(s_i|\omega_j) = \pi_{ij}$ and assume, for concreteness and without loss of generality, that for $1 \leq i \leq i' \leq k$ either $\pi_{iH}/\pi_{iL} \leq \pi_{i'H}/\pi_{i'L}$ or $\pi_{i'L} = 0$, i.e., higher indexed outcomes in S are associated with higher likelihood ratios (and higher posterior probabilities of ω_H).

The receiver observes the signal π chosen by the sender and makes an interim update of her belief that the payoff-relevant state is ω_H to $\widetilde{\beta}(\pi) \in [\mu_L, \mu_H]$.[5] I.e., the receiver is allowed to make inferences about the sender's private information from the sender's choice of signal. The receiver next observes a realization $s_i \in S$ drawn according to π, and updates her belief to $\hat{\beta}(\pi, s_i)$. Let

$$B(\pi, s_i, \mu) := \frac{\pi_{iH}\mu}{\pi_{iH}\mu + \pi_{iL}(1-\mu)}$$

be the mapping from a signal $\pi \in \Pi$, an outcome $s_i \in S$ and a prior $\mu \in (0,1)$ to a posterior probability that the state is ω_H.

[4] I thus assume the cardinality of S bounded above by $|\widetilde{S}|$, which departs slightly from Kamenica and Gentzkow (2011). This assumption plays a role in the existence proof in Theorem 1, by ensuring compactness of the set of feasible signals. The assumption is plausible if, e.g., there is some limit to the number of signal outcomes that players can reasonably be expected to distinguish and reason about.

[5] While convention would require the receiver's interim belief to be with respect to the sender's type, the approach here provides a shortcut and is without loss of generality as long as $\widetilde{\beta}(\pi) \in [\mu_L, \mu_H]$.

Two types of signals deserve special mention and their own name. First, the sender can fully reveal the payoff-relevant state by choosing a signal π such that $\pi_{ij} > 0$ implies $\pi_{ij'} = 0$ for all $i \in \{1, ..., k\}$ and $\{j, j'\} = \{L, H\}$. The simplest example is a binary signal with $\pi_{1L} = \pi_{2H} = 1$. I refer to such signals as *fully disclosing*. Let π^{FD} denote a generic fully disclosing signal and let $\Pi^{FD} \subset \Pi$ denote the set of fully disclosing signals. This definition of full disclosure differs from the standard definition in persuasion games (see, e.g., Milgrom 2008), where full disclosure typically refers to revealing the private information, i.e., the sender's type. Here full disclosure instead refers to the payoff-relevant state and does not imply that the receiver becomes certain about the sender's type.[6] The sender's second benchmark option is an uninformative signal, i.e., a signal π such that $\pi_{iH} = \pi_{iL}$ for all $i \in \{1, ..., k\}$. I refer to such signals as *silent*.

The setup requires the sender to commit to a signal and the receiver to observe a randomly drawn outcome. E.g., the sender is not allowed to secretly choose a signal, observe an outcome and choose a different signal if unsatisfied. This is somewhat natural if the sender is a firm specifying terms of free trials of a recently developed product. Another interpretation is that the signal is an investigative report containing evidence informative about the payoff-relevant state. The sender is not familiar with the evidence prior to collecting it and commits to a protocol of investigation (a signal) and to report all gathered evidence (the signal's outcome). As Kamenica and Gentzkow (2011) argue, there are then several situations in which the commitment assumption is plausible. For example, in the US a prosecutor is required by law to disclose any admissible evidence in favor of the accused and naturally discloses any evidence against the accused. Pharmaceutical companies must register the design of clinical trials prior to their execution and have clear incentives to report the outcome to the FDA truthfully.

Sender and receiver payoff functions. I assume the sender's payoff strictly and continuously increasing in the receiver's updated belief. For concreteness, I model this by assuming the sender's payoff dependent on a receiver action $a \in A := [\underline{a}, \bar{a}]$ taken after her final update

[6] Indeed, no signal disclosing the sender's type is available here and one can therefore not appeal to the unraveling arguments in Milgrom (1981), Seidmann and Winter (1997), or Hagenbach, Koessler and Perez-Richet (2014) to argue existence or uniquess of fully separating equilibria.

of beliefs. The receiver obtains a payoff depending on her action and the payoff-relevant state and given by a strictly concave and twice continuously differentiable function $u : \Omega \times A \to \mathbb{R}$. Let $a^R(\omega_j) := \arg\max_{a \in A} u(\omega_j, a)$ for $j = L, H$. I assume that $\underline{a} < a^R(\omega_L) < a^R(\omega_H) < \bar{a}$. The receiver's expected payoff given a generic belief $\beta \in [0,1]$ and action $a \in A$ is given by $U(\beta, a) = (1-\beta)u(\omega_L, a) + \beta u(\omega_H, a)$. Under the assumptions above $\hat{a}^R(\beta) := \arg\max_{a \in A} U(\beta, a)$ is well-defined, continuous and strictly increasing.

The sender's payoff depends only on the receiver's action and is given by $v : A \to \mathbb{R}$, where v is assumed continuous on A and strictly increasing. The mapping $\hat{v} : [0,1] \to \mathbb{R}$ such that $\hat{v} = v \circ \hat{a}^R$ then gives the sender's payoff as a continuous and strictly increasing function of the belief of an optimally responding receiver.

Strategies and equilibrium. I focus on pure sender strategies consisting of a pair $(\pi^L, \pi^H) \in \Pi^2$, where π^j is the signal chosen by type μ_j for $j = L, H$. A (pure) receiver strategy is a function $\alpha : \{\pi \times S\}_{\pi \in \Pi} \to A$. Given a receiver strategy α the expected payoff of a type μ_j sender using a signal π is given by the function

$$V(\pi, \alpha, \mu_j) := \sum_{i=1}^{k} \left[\mu_j \pi_{iH} + (1-\mu_j)\pi_{iL}\right] v(\alpha(\pi, s_i)).$$

Let $\hat{V} : \Pi \times [0,1]^2 \to \mathbb{R}$ be defined by

$$\hat{V}(\pi, \mu, \mu_j) := \sum_{i=1}^{k} \left[\mu_j \pi_{iH} + (1-\mu_j)\pi_{iL}\right] \hat{v}(B(\pi, s_i, \mu)),$$

i.e., $\hat{V}(\pi, \mu, \mu_j)$ gives the expected payoff of a type μ_j sender using signal π on a receiver who responds rationally to the signal given interim receiver belief μ. Notice that $\hat{V}$ is continuous in its three arguments. Further, given a signal π and interim receiver belief μ the set of payoffs $\{\hat{v}(B(\pi, s_i, \mu))\}_{i=1}^{k}$ induced by π is independent of the sender's type μ_j. The sender's expected payoff given π and μ is consequently linear in type and the type-dependence arises only since different sender types induce different convex combinations of the induced payoffs. Finally, let

$\hat{V}_j^{FD} := \hat{V}(\pi^{FD}, \mu, \mu_j) = (1-\mu_j)\hat{v}(0) + \mu_j \hat{v}(1)$ denote the expected payoff of a type $j = L, H$ sender using a fully disclosing signal.

The solution concept is standard perfect Bayesian equilibrium (PBE) robust to the intuitive criterion (Cho and Kreps, 1987), in what follows referred to simply as *equilibrium* and defined as follows.

Definition 1 *An equilibrium is a sender strategy* (π^L, π^H)*, a receiver strategy* α*, receiver interim beliefs* $\widetilde{\beta}$ *and receiver final beliefs* $\hat{\beta}$*, such that (i)* $\alpha(\pi, s) = \hat{a}^R(\hat{\beta}(\pi, s))$ *for all* $(\pi, s) \in \{\pi \times S\}_{\pi \in \Pi}$*, (ii) for* $j = L, H$ *we have* $\pi^j \in \arg\max_{\pi \in \Pi} V(\pi, \alpha, \mu_i)$*, (iii) the receiver's beliefs are rational, i.e., for any* $(\pi, s) \in \{\pi \times S\}_{\pi \in \Pi}$ *and* $\{j, j'\} = \{L, H\}$ *we have* $\widetilde{\beta}(\pi) \in [\mu_L, \mu_H]$*,*

$$\widetilde{\beta}(\pi^j) = \begin{cases} \mu_j \text{ if } \pi^j \neq \pi^{j'} \\ \mu_0 \text{ if } \pi^j = \pi^{j'} \end{cases}$$

and $\hat{\beta}(\pi, s) = B(\pi, s, \widetilde{\beta}(\pi))$*, and (iv) for any* $\pi \notin \{\pi^L, \pi^H\}$ *and* $\{j, j'\} = \{L, H\}$ *such that* $V(\pi^j, \alpha, \mu_j) < \hat{V}(\pi, \mu_H, \mu_j)$ *and* $V(\pi^{j'}, \alpha, \mu_{j'}) > \hat{V}(\pi, \mu_H, \mu_{j'})$ *we have* $\widetilde{\beta}(\pi) = \mu_j$*.*

Conditions (i)-(iii) define a standard PBE, while (iv) ensures robustness to the intuitive criterion. Notice that the equilibrium expected payoff of a type μ_j sender is $V(\pi^j, \alpha, \mu_j) = \hat{V}(\pi^j, \widetilde{\beta}(\pi), \mu_j)$. An equilibrium is *separating* if $\pi^L \neq \pi^H$. In a separating equilibrium, the receiver's interim belief agrees with that of the sender, implying that the private information revealed to the sender's before choosing a signal is transmitted to the receiver. An equilibrium is *pooling* if $\pi^L = \pi^H$. Finally, an equilibrium is *fully disclosing* if $\pi^L, \pi^H \in \Pi^{FD}$, implying that the receiver becomes perfectly informed of the payoff-relevant state.

3 Analysis

The nature of the set of feasible signals and the associated equilibrium expected payoff functions generate a signaling game with some particular properties. First, the sender's expected payoff given a fully disclosing signal is independent of the receiver's interim belief. The availability of such signals therefore establish type-specific lower bounds on equilibrium expected payoffs. Second, given a signal and a receiver interim belief, the sender's expected payoff is linear in type. The slope of this linear dependence is related to a dimension of *precision*, along which a fully disclosing signal represents one extreme. The third property of the game is a condition of increasing differences related to precision which exploits this linearity and roughly states that higher types have stronger preferences for more precise signals. As will be shown below, the lower bound on equilibrium expected payoffs resulting from the availability of fully disclosing signals together with the property of increasing differences rule out any pooling equilibrium which is not fully disclosing and pins down the equilibrium as the solution to a maximization problem.

In what follows I first identify conditions of increasing differences in expected payoffs in two preliminary lemmata and then proceed to the characterizations of the equilibria.

3.1 Preliminary results

Since higher types have "better news" for the receiver it seems natural to expect that such types in some sense should have stronger preferences for signals which reveal more information. The following partial order on the set of feasible signals is consistent with such an intuition. For any $\pi, \pi' \in \Pi$ such that $S = S'$ I say that π' is *more precise* than π if for all $i \in \{1, ..., k\}$ either $\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}$ or $\pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}$. If additionally $\pi' \neq \pi$, then π' is *strictly more precise* than π. This definition is stronger than Blackwell (1951) informativeness, i.e., if π' is more precise than π, then π' is more informative than π in the sense of Blackwell (1951), but the converse is not true.[7] In particular, precision disperses the posterior probabilities from the prior, i.e., if π' is more precise than π, then either $B(\pi', s_i, \mu) \leq B(\pi, s_i, \mu) \leq \mu$ or $B(\pi', s_i, \mu) \geq B(\pi, s_i, \mu) \geq \mu$.

Notice that $\hat{V}(\pi, \mu, \mu_j)$ can be written as a linear function of μ_j with slope $\sum_{i=1}^{k}(\pi_{iH} - \pi_{iL})\hat{v}(B(\pi, s_i, \mu))$. The slope determines how the expected payoff of the sender given a signal π and a receiver interim μ depends on the sender's type. It is not difficult to see that the slope

[7] While there are more complete orderings of signals, including Blackwell's (1951), the definition of *more precise* is appropriate for the incentive compatibility analysis which is the main purpose here.

is increasing in the precision of π. The following result exploits this fact to identify conditions of increasing differences which provide a sense in which higher types have stronger preferences for more precise signals (proofs are in the appendix).

Lemma 1 *Consider any $\mu, \mu' \in [0,1]$ and any $\pi, \pi' \in \Pi$ such that $S = S'$. (i) If for all $i \in \{1, .., k\}$ either $\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}$ and $B(\pi', s_i, \mu') \leq B(\pi, s_i, \mu)$ or $\pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}$ and $B(\pi', s_i, \mu') \geq B(\pi, s_i, \mu)$, then*

$$\hat{V}(\pi', \mu', \mu_H) - \hat{V}(\pi, \mu, \mu_H) \geq \hat{V}(\pi', \mu', \mu_L) - \hat{V}(\pi, \mu, \mu_L).$$

(ii) If π' is more precise than π then

$$\hat{V}(\pi', \mu, \mu_H) - \hat{V}(\pi, \mu, \mu_H) \geq \hat{V}(\pi', \mu, \mu_L) - \hat{V}(\pi, \mu, \mu_L),$$

with a strict inequality if π' is strictly more precise than π.

These conditions rely on a payoff maximizing receiver updating beliefs according to Bayes rule and is therefore more of an equilibrium construction than a standard condition of increasing differences, which is usually assumed on the primitives and without requiring particular receiver behavior. The lemma implies, e.g., that if type μ_L is indifferent between two signals which induce the same receiver interim belief and π' is strictly more precise than π, then type μ_H strictly prefers π'. The logic is that if π' is more precise than π, then the difference between the expected payoffs conditional on ω_H and ω_L is larger under π' than under π. Since the state is more likely to be ω_H for type μ_H than for type μ_L it follows that type μ_H is relatively more attracted by π'. The first part of Lemma 1 allows the receiver to respond to two different interim beliefs μ and μ', which is crucial in assessing incentive compatibility below. The second part is essentially a corollary of the first part which obtains a neater condition by setting $\mu = \mu'$ and appealing to the effect of more precise signals on the set of induced receiver beliefs.

The next result relies on Lemma 1 to show that given any signal generating a weakly lower expected payoff than a fully disclosing signal for type μ_H under receiver interim $\mu < \mu_H$, there are (more precise) signals available such that only type μ_H benefits from these under receiver interim μ_H.

Lemma 2 *Consider any $\mu \in [0, \mu_H)$ and any $\pi \in \Pi \backslash \Pi^{FD}$ and suppose $\hat{V}(\pi, \mu, \mu_H) \geq \hat{V}_H^{FD}$. There is some $\pi' \in \Pi$ which is more precise than π and such that*

$$\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0 > \hat{V}(\pi', \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L).$$

Lemma 2 is useful in determining the receiver's out-of-equilibrium beliefs below. The logic of the proof is, roughly, the following. If both types use a signal π then an increase in the receiver's interim belief increases the payoffs of both types. One can then consider a path of increasing precision from π to some $\pi^{FD} \in \Pi^{FD}$ along which increasing differences ensure that type μ_H is favored more than type μ_L. Since $\hat{V}(\pi, \mu, \mu_H) \geq \hat{V}_H^{FD}$ one eventually finds some π' such that type μ_H marginally prefers π' over π and type μ_L strictly prefers π over π'. The main crack in the argument, which leads to the main difficulty in the proof, is that it is not possible to say which type benefits the most from the initial increase in interim belief. Lemma 1 is not helpful here, since *all* posterior probabilities are higher after the interim increase. The trick in the proof is to construct a signal satisfying the hypothesis of Lemma 1 and which is still preferred by type μ_H over π after the interim increase, and then apply the argument of increasing precision outlined above.

3.2 Equilibria

Lemma 2 almost immediately implies that any equilibrium is either separating or fully disclosing. In equilibrium it obviously holds that the payoffs of type μ_L and type μ_H at least equal $\hat{V}_L^{FD}$ and $\hat{V}_H^{FD}$, respectively. If π is a pooling strategy, then the receiver's equilibrium interim beliefs satisfies $\widetilde{\beta}(\pi) = \mu_0 < \mu_H$ and the second part of Lemma 2 then implies the existence of some π'' such that $\widetilde{\beta}(\pi'') = \mu_H$ and which is a profitable deviation for type μ_H.

Proposition 1 *Any equilibrium is either separating or fully disclosing*

In other words, either the sender's choice of signal reveals his type to the receiver, or the outcome of the signal reveals the payoff-relevant state. The result is driven by type μ_H's stronger preferences for precise signals and the possibility of fully disclosing the payoff-relevant state. The argument is roughly the following. A deviation from a pooling strategy π to some signal $\pi' \simeq \pi$ is profitable for both types if the receiver infers $\widetilde{\beta}(\pi') = \mu_H$. As π' becomes more precise and converges to full disclosure, however, the deviation payoffs converge to $\hat{V}_L^{FD}$ and $\hat{V}_H^{FD}$. For π' sufficiently precise both types therefore prefer the equilibrium over π'. Type μ_H's stronger preferences for precise signals, reflected in the increasing differences, implies that there is some sufficiently precise π' such that type μ_L prefers the equilibrium while type μ_H prefers the deviation. The intuitive criterion requires $\widetilde{\beta}(\pi') = \mu_H$ and there is a profitable deviation for type μ_H.

Proposition 1 implies that identifying the set of equilibrium sender strategies is equivalent to identifying the set of incentive compatible fully disclosing and separating strategies. This can be accomplished in terms of solutions to certain maximization problems. Let

$$\Pi_j^* := \arg\max_{\pi\in\Pi} \hat{V}(\pi, \mu_j, \mu_j)$$

for $j \in \{L, H\}$, where it is not difficult to see that $\Pi_j^* \neq \varnothing$. I.e., Π_j^* is what Kamenica and Gentzkow (2011) refer to as type μ_j's *optimal signals*. Under a full information benchmark in which the receiver knows the sender's type, each type would choose one of its respective optimal signals. The set of equilibrium sender strategies is fully characterized by type μ_L's optimal signal and a signal which maximizes type μ_H's payoff under an additional incentive compatibility constraint ensuring that type μ_L does not have incentives to mimic type μ_H's signal.

Theorem 1 *(i) There is an equilibrium at which the sender's strategy is (π^L, π^H) if and only if*

$$\pi^L \in \arg\max_{\pi\in\Pi} \hat{V}(\pi, \mu_L, \mu_L) \tag{1}$$

and

$$\pi^H \in \arg\max_{\pi\in\Pi} \hat{V}(\pi, \mu_H, \mu_H) \ s.t. \ \hat{V}(\pi, \mu_H, \mu_L) \leq \hat{V}(\pi^L, \mu_L, \mu_L). \tag{2}$$

(ii) An equilibrium exists.

While counterexamples (e.g., if $\hat{v}$ is a linear function), reveal that there is not generically a unique equilibrium, an immediate corollary of Theorem 1 is that all equilibria are equivalent from the point of view of the sender.

Corollary 1 *The expected payoffs of the type μ_L and type μ_H sender are constant across equilibria.*

Theorem 1 shows that in spite of the absence of single-crossing assumptions on the primitives, the model's predictions are remarkably similar to those of well behaved signaling models, such as simple versions of Spence's (1973) model of job-market signaling. In particular, to identify a separating PBE it is sufficient to maximize both types' expected payoffs as if the receiver knew their types and subject to an upward incentive compatibility constraint (1 and 2 are sufficient). When adding the intuitive criterion this is the only equilibrium which survives (1 and 2 are necessary).

The result is again driven by increasing differences in precision and the lower bound on equilibrium expected payoffs implied by the availability of a fully disclosing signal, and its proof uses Lemma 1 and Lemma 2 extensively. It is fairly obvious that (1) and the constraint in (2) are necessary for type μ_L not to have a profitable deviation and sufficient if out-of-equilibrium beliefs equal $\widetilde{\beta}(\pi) = \mu_L$. The proof is thus mainly concerned with the necessity and sufficiency of type μ_H's maximization in (2). Sufficiency of (2) follows by setting out-of-equilibrium beliefs equal to $\widetilde{\beta}(\pi) = \mu_L$ and observing that the existence of a profitable deviation would contradict (2). For, one could then invoke the increasing differences underlying Lemma 2 to claim the existence of a sufficiently precise signal which outperforms the deviation under receiver interim μ_H and still satisfies the constraint in (2). The necessity of (2) follows by appealing to the intuitive criterion. If (2) does not hold in some equilibrium then there must be an alternative signal, say π', satisfying the constraint in (2) and generating a higher expected payoff for type μ_H. One can then consider a continuous path of signals of increasing precision from π' to some π^{FD} and invoke Lemma 1 to argue existence of a signal π'' along this path which is marginally preferred by type μ_H over the equilibrium and strictly satisfies the constraint in (2). The intuitive criterion implies a profitable deviation to π'' for type μ_H. Finally, given part (i) of the theorem, the existence of an equilibrium in part (ii) follows by observing that maximization problems (1) and (2) are well defined.

Fully separating versus fully disclosing equilibria. Proposition 1 states that any equilibrium is either separating or fully disclosing. It turns out that fully disclosing equilibria are something of a special case and there is a simple way of checking whether they might occur. Theorem 1 implies that a fully disclosing equilibrium can only exist if a fully disclosing signal is optimal for type μ_L. It follows that if a fully disclosing signal is optimal also for type μ_H, then a fully disclosing equilibrium exists. The following result emphasizes that a fully disclosing equilibrium exists if and only if fully disclosing signals are optimal for both types. In addition, if this is the case, then type μ_H uses a fully disclosing signal in any equilibrium.

Proposition 2 *(i) If a fully disclosing signal is optimal for both types, then any equilibrium signal of type μ_H is fully disclosing. (ii) A fully disclosing equilibrium exists if and only if a fully disclosing signal is optimal for both types.*

The result follows by observing that if a fully disclosing signal is optimal for one type, then such a signal is optimal also for the other type. Kamenica and Gentzkow's (2011) characterization of optimal signals provides a straightforward check of whether fully disclosing signals are optimal. It suffices to draw a line from $(0, \hat{v}(0))$ to $(1, \hat{v}(1))$. A fully disclosing signal is optimal for both types if and only if this line is (weakly) above $\hat{v}$.

Further properties of the equilibrium signals. Theorem 1 establishes that type μ_L's equilibrium signal is optimal, i.e., type μ_L would not behave differently if the receiver could observe his type. The next result shows that the constraint in (2) is active in the sense of forcing type μ_H to choose a signal which is not optimal, with exception only for the special case in which fully disclosing signals are optimal.

Proposition 3 *Suppose (π^L, π^H) is an equilibrium sender strategy. If $\pi^H \in \Pi_H^*$ then $\pi^H \in \Pi^{FD}$.*

Proposition 3 follows by noting that type μ_L strictly prefers posing as type μ_H under any signal $\pi^H \in \Pi_H^* \backslash \Pi^{FD}$ over a signal $\pi^L \in \Pi_L^*$ under a correct receiver interim belief. The proof exploits the characterization of optimal signals provided by Kamenica and Gentzkow (2011). In particular, the choice of a signal is equivalent to the choice of a distribution of updated receiver beliefs. It is easily seen that in equilibrium the lowest updated belief induced by type μ_H must be lower than the highest updated belief induced by type μ_L. The proof then argues that if this is true for some optimal signals of type μ_H and type μ_L, then the optimal distributions of updated receiver beliefs of both types are similar in a precise sense.[8] Intuitively, this similarity implies that if both types use optimal signals in equilibrium and type μ_L deviates to type μ_H's signal, then type μ_L obtains his equilibrium payoff plus a bonus resulting from posing as type μ_H. The deviation is then profitable and type μ_H can therefore not use an optimal signal in equilibrium, except in the special case in which the optimal signal is fully disclosing.

Proposition 3 implies that incentive compatibility generates a cost for type μ_H in comparison to the full information benchmark in which the receiver knows the sender's type. An implication is that if the sender could credibly reveal his private information to the receiver prior to choosing a signal he would prefer doing this.[9] For example, a pharmaceutical company submitting a New Drug Application to the FDA would be keen on passing on research results obtained prior to choosing a protocol for clinical trials.[10] It is not clear, however, that the FDA would be equally keen on receiving this information. In equilibrium the pharmaceutical company reveals its private information anyway through its choice of protocol for clinical trials, and incentive compatibility sometimes leads the pharmaceutical company with favorable private information to reveal more information than what it would if the FDA knew its type. For example, if type μ_H's unique

[8]The idea is most easily visualized by plotting what Kamenica and Gentzkow (2011) refer to as the *concave closure* of $\hat{v}$, i.e., the smallest concave function greater than $\hat{v}$. In particular, here μ_L and μ_H would belong to the same linear segment of the concave closure of $\hat{v}$, implying that the supports of both types' optimal distributions of updated beliefs are similar.

[9]Adding a cheap-talk stage prior to the choice of signal, however, would not affect the set of equilibrium outcomes, since both types would claim to be type μ_H whenever such a message is believed by the receiver.

[10]A similar result is obtained by Kolotilin (2014a), who finds that verifiable private sender information unravels in a related model.

optimal signal is silent, then Proposition 3 implies that type μ_H's equilibrium signal is not silent and therefore strictly Blackwell (1951) more informative than the optimal signal. If additionally type μ_L has a unique optimal signal, the receiver would consequently strictly prefer the equilibrium information structure over the information structure resulting in the full information benchmark in which she knows the sender's type. In other words, there is sometimes an "ignorance-rent" for the receiver. A sufficient condition for such a rent to appear is that $\hat{v}$ is strictly concave, in which case both types' unique optimal signals are silent. The following remark summarizes the observation.

Remark 1 *Fix the signal of type μ_L at some $\pi_L^* \in \Pi_L^*$. If all $\pi_H^* \in \Pi_H^*$ are silent, then the receiver's equilibrium expected payoff is strictly larger than her expected payoff in the full information benchmark in which she knows the sender's type and each sender type chooses an optimal signal. A sufficient condition for this to occur is that $\hat{v}$ is a strictly concave function.*

The next result implies that the equilibrium signal of type μ_H cannot be too imprecise. In particular, there cannot be a signal which is more precise than the equilibrium signal of type μ_H and which would give him a higher expected payoff. Therefore, none of type μ_H's optimal signals can be more precise than the equilibrium signal of type μ_H.

Proposition 4 *Suppose (π^L, π^H) is an equilibrium sender strategy. For any $\pi \in \Pi$ which is more precise than π^H it holds that $\hat{V}(\pi, \mu_H, \mu_H) \leq \hat{V}(\pi^H, \mu_H, \mu_H)$. Consequently, if $\pi_H^* \in \Pi_H^*$ then π_H^* is not more precise than π^H.*

Proposition 4 is established using an increasing differences argument similar to those above. The result implies that the equilibrium signal of type μ_H can never be silent. For, if the equilibrium signal of type μ_H were silent, then Proposition 3 implies that there is an optimal signal which is more precise than type μ_H's equilibrium signal, contradicting Proposition 4. Incentive compatibility hence requires the sender to reveal all private information and some additional information, even when both types' optimal signals are silent.

A remark concerning bad and good news. I will close the discussion with a remark consisting of two parts. The first part has already been briefly mentioned and relates to the distribution of induced receiver beliefs in equilibrium, while the second part relates to the interpretation of silence.

Remark 2 *Suppose* (π^L, π^H) *is an equilibrium sender strategy such that* π^L *and* π^H *have supports* $\{s_1^L, ..., s_{k_L}^L\}$ *and* $\{s_1^H, ..., s_{k_H}^H\}$. *Then (i)* $B(\pi^H, s_1^H, \mu_H) < B(\pi^L, s_{k_L}^L, \mu_L)$ *and (ii) if* π^L *is silent there are* $i, i' \in \{1, ..., k_H\}$ *such that* $B(\pi^H, s_i^H, \mu_H) < \mu_L < B(\pi^H, s_{i'}^H, \mu_H)$.

The first part states that the supports of the distribution of induced receiver beliefs of both types must overlap in equilibrium. This follows trivially by observing that if type μ_L deviates to type μ_H's equilibrium signal he induces the same set of updated beliefs as type μ_H does. An interpretation is that in equilibrium incentive compatibility requires the "worst news" of the "good type" to be worse than the "best news" of the "bad type." In other words, the equilibrium signals are always sufficiently informative to override the private information in some sense.

The second part of the remark is a corollary to the first part emphasizing that if the equilibrium signal of type μ_L is silent, then type μ_H must induce updated beliefs both below and above the updated belief induced by type μ_L's silent signal (which equals μ_L). Since only type μ_L can use a silent signal in equilibrium an interpretation is that an uninformative signal is never the worst nor the best equilibrium news. The idea can be related to Milgrom's (1981) unraveling result, which states that withheld information is always interpreted in the worst possible way in equilibrium, i.e., silence is the worst kind of news. While withholding information is interpreted in the worst possible way here in terms of the receiver's belief regarding the sender's private information, it is neither the worst possible nor the best possible equilibrium news.

4 Concluding remarks

This paper has shown that the introduction of unverifiable private information in a framework of Bayesian persuasion generates a tractable signaling game with concrete predictions. In particular, in perfect Bayesian equilibria robust to a reasonable refinement private information unravels and the set of such equilibria can be fully characterized in terms of maximization problems resembling the ones found in standard signaling models. The results are driven by properties of increasing differences which arise endogenously in the game. An implication is that one can sometimes infer an agent's private information by the design of his experiments, tests, free trials or by the type of evidence he chooses to collect. By studying the equilibrium more closely it turns out that the sender typically would prefer revealing his private information to the receiver prior to choosing a signal. The receiver, however, sometimes obtains an "ignorance rent" and prefers the equilibrium over the full information benchmark in which he knows the sender's type.

There are many ways in which one could generalize the framework considered here. For example, Gentzkow and Kamenica (2014) discuss how costs related to the reduction in entropy (see, e.g., Shannon 1948) can be introduced into a framework of Bayesian persuasion. It would be interesting to see how such costs would affect the unraveling of private information. Another somewhat restrictive feature of the framework here is the binary set of payoff-relevant states. A natural questions is how robust the unraveling of private information is to a more general specification of the set of payoff-relevant states. Finally, this paper investigates particularly simple sender preferences over receiver beliefs. Extensions to other classes preferences, such as single peaked preferences, would be of interest.

Another direction for future research is in applications. For example, in Gill and Sgroi's (2012) analysis of a monopolist's pre-launch tests of a new product signaling through the choice of test can be ruled out if the set of feasible signals is constrained and the monopolist is perfectly informed. The analysis here suggests that if instead the set of feasible signals in unconstrained and the monopolist is imperfectly informed, then such signaling may play an important role.

5 Appendix

Proof. (Lemma 1) Let $\mu, \mu' \in [0,1]$ and let $\pi, \pi' \in \Pi$ such that $S = S'$.

To prove (i) suppose that for all $i \in \{1,..,k\}$ either $\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}$ and $B(\pi', s_i, \mu') \leq B(\pi, s_i, \mu)$ or $\pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}$ and $B(\pi', s_i, \mu') \geq B(\pi, s_i, \mu)$. Let $\bar{k} \in \{1,..,k-1\}$ be the smallest number such that $\pi'_{\bar{k}H} \geq \pi'_{\bar{k}L}$, i.e., such that $B(\pi', s_{\bar{k}}, \mu') \geq B(\pi, s_{\bar{k}}, \mu)$. Abbreviate $\hat{v}_i = \hat{v}(B(\pi, s_i, \mu))$ and $\hat{v}'_i = \hat{v}(B(\pi', s_i, \mu))$. Notice that $\sum_{i<\bar{k}}(\pi_{iH} - \pi_{iL}) = -\sum_{i\geq\bar{k}}(\pi_{iH} - \pi_{iL})$ and likewise for π'. Then

$$
\begin{aligned}
& \hat{V}(\pi', \mu', \mu_H) - \hat{V}(\pi, \mu, \mu_H) - \left[\hat{V}(\pi', \mu', \mu_L) - \hat{V}(\pi, \mu, \mu_L)\right] \\
= \; & (\mu_H - \mu_L)\left[\sum_{i<\bar{k}}(\pi'_{iH} - \pi'_{iL})\hat{v}'_i - (\pi_{iH} - \pi_{iL})\hat{v}_i + \sum_{i\geq\bar{k}}(\pi'_{iH} - \pi'_{iL})\hat{v}'_i - (\pi_{iH} - \pi_{iL})\hat{v}_i\right] \\
\geq \; & (\mu_H - \mu_L)\left[\sum_{i<\bar{k}}(\pi'_{iH} - \pi'_{iL})\hat{v}_i - (\pi_{iH} - \pi_{iL})\hat{v}_i + \sum_{i\geq\bar{k}}(\pi'_{iH} - \pi'_{iL})\hat{v}_i - (\pi_{iH} - \pi_{iL})\hat{v}_i\right] \\
\geq \; & (\mu_H - \mu_L)\left[\sum_{i>\bar{k}}\left[(\pi'_{iH} - \pi'_{iL}) - (\pi_{iH} - \pi_{iL})\right]\hat{v}_i + \sum_{i\geq\bar{k}}\left[(\pi'_{iH} - \pi'_{iL}) - (\pi_{iH} - \pi_{iL})\right]\hat{v}_{\bar{k}}\right] \\
= \; & (\mu_H - \mu_L)\left[\sum_{i<\bar{k}}\left[(\pi'_{iH} - \pi'_{iL}) - (\pi_{iH} - \pi_{iL})\right]\hat{v}_i - \sum_{i<\bar{k}}\left[(\pi'_{iH} - \pi'_{iL}) - (\pi_{iH} - \pi_{iL})\right]\hat{v}_{\bar{k}}\right] \\
= \; & (\mu_H - \mu_L)\sum_{i<\bar{k}}\left[(\pi'_{iH} - \pi'_{iL}) - (\pi_{iH} - \pi_{iL})\right](\hat{v}_i - \hat{v}_{\bar{k}}) \geq 0,
\end{aligned}
$$

which proves (i).

To prove (ii), suppose first that π' is more precise than π. Then for all $i \in \{1,...,k\}$ either $\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}$ and $B(\pi', s_i, \mu) \leq B(\pi, s_i, \mu)$ or $\pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}$ and $B(\pi', s_i, \mu) \geq B(\pi, s_i, \mu)$. By (i) we have $\hat{V}(\pi', \mu, \mu_H) - \hat{V}(\pi, \mu, \mu_H) \geq \hat{V}(\pi', \mu, \mu_L) - \hat{V}(\pi, \mu, \mu_L)$.

Suppose now that π' is strictly more precise than π and that for some $i' \in \{1,...,k\}$ we have $\pi'_{i'H}/\pi'_{i'L} < \pi_{i'H}/\pi_{i'L} \leq 1$, implying $\pi'_{i'H} - \pi'_{i'L} < 0$ and $B(\pi', s_{i'}, \mu) < B(\pi, s_{i'}, \mu)$. The series of inequalities above is then valid and the first inequality is strict so $\hat{V}(\pi', \mu, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > \hat{V}(\pi', \mu, \mu_L) - \hat{V}(\pi, \mu, \mu_L)$. An analogous argument holds if instead $\pi'_{i'H}/\pi'_{i'L} > \pi_{i'H}/\pi_{i'L} \geq 1$, and we have therefore proved (ii) ■

Proof. (Lemma 2) Let $\mu \in [0, \mu_H)$ and suppose $\pi \in \Pi \backslash \Pi^{FD}$ with $\hat{V}(\pi, \mu, \mu_H) \geq \hat{V}_H^{FD}$. I first prove the following preliminary result.

Claim. *There is some $\pi' \in \Pi \backslash \Pi^{FD}$ with support $S' = S$ which is not silent and such that*

$$\begin{aligned}
\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) &\geq \hat{V}(\pi', \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L) \\
\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) &> 0.
\end{aligned}$$

Proof. The proof constructs a signal π' satisfying the hypothesis of part (i) of Lemma 1 and such that $\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0$. Let $\overline{k} \in \{1, ..., k-1\}$ be the smallest number such that $\pi_{\overline{k}H} \geq \pi_{\overline{k}L}$, let $\bar{K} := \{1\} \cup \{1, ..., \overline{k} - 1\}$ and let $K = \{1, ..., k\}$. The construction transfers probability in $\pi(\cdot|\omega_H)$ from outcomes $\{s_i\}_{i \in \bar{K}}$ to outcomes $\{s_i\}_{i \in K \backslash \bar{K}}$ maintaining $B(\pi, s_i, \mu)$ fixed for all $i \in \bar{K}$. Let $\pi'_{iL} = \pi_{iL}$ for all $i \in \{1, ..., k\}$ and let

$$\pi'_{iH} = \frac{\mu(1-\mu_H)\pi_{iH}}{(1-\mu)\mu_H \pi_{iL}} \pi'_{iL} = \frac{\mu(1-\mu_H)\pi_{iH}}{(1-\mu)\mu_H}$$

for all $i \in \bar{K}$. Straightforward algebra shows that for all $i \in \bar{K}$ we have $B(\pi', s_i, \mu_H) = B(\pi, s_i, \mu)$ and $\pi'_{iH} \leq \pi_{iH}$, with equality only if $\pi_{iH} = 0$. For all $i \in K \backslash \bar{K}$ let $\pi'_{iH} = \pi_{iH} + \lambda(1 - \pi_{iH})$ with $\lambda = \sum_{i \in \bar{K}} (\pi_{iH} - \pi'_{iH}) / \sum_{i \in K \backslash \bar{K}} (1 - \pi_{iH})$ if $\sum_{i \in K \backslash \bar{K}} (1 - \pi_{iH}) > 0$ and $\lambda = 0$ otherwise. Notice that π' is not silent, for then $\pi'_{kH} = \pi_{kH}$ but then $\pi'_{1H} = 0$, a contradiction.

For any $i \in \bar{K}$ it holds that $\pi'_{iH} \leq \pi_{iH} \leq \pi_{iL} \leq \pi'_{iL}$ and $B(\pi', s_i, \mu_H) \leq B(\pi, s_i, \mu)$ and for any $i \in K \backslash \bar{K}$ it holds that $\pi'_{iH} \geq \pi_{iH} \geq \pi_{iL} \geq \pi'_{iL}$ and therefore $B(\pi', s_i, \mu_H) \geq B(\pi, s_i, \mu_H) \geq B(\pi, s_i, \mu)$. The hypothesis of part (i) of Lemma 1 is then satisfied which implies

$$\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) \geq \hat{V}(\pi', \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L).$$

Abbreviate $\hat{v}'_i = \hat{v}(B(\pi', s_i, \mu_H))$ and $\hat{v}_i = \hat{v}(B(\pi, s_i, \mu))$. Then

$$\begin{aligned}\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) &= \sum_{i=1}^{k} [\mu_H(\pi'_{iH}\hat{v}'_i - \pi_{iH}\hat{v}_i) + (1-\mu_H)\pi_{iL}(\hat{v}'_i - \hat{v}_i)] \\ &\geq \mu_H \sum_{i=1}^{k} (\pi'_{iH}\hat{v}_i - \pi_{iH}\hat{v}_i) \geq 0,\end{aligned}$$

where the first inequality follows since $\hat{v}'_i \geq \hat{v}_i$ for all $i \in K$ and the second inequality follows since by construction $\pi'(\cdot|\omega_H)$ first order stochastically dominates $\pi(\cdot|\omega_H)$. If the first inequality is an equality then $\hat{v}'_i = \hat{v}_i$ and therefore $B(\pi, s_i, \mu_H) = B(\pi, s_i, \mu)$ for $i \in K\backslash\bar{K}$, which is only possible if $\pi_{iL} = 0$ for all $i \in K\backslash\bar{K}$. Since $\pi \notin \Pi^{FD}$ it must then be that $\pi_{iH} > 0$ for some $i \in \bar{K}$ and then $\pi(\cdot|\omega_H) \neq \pi'(\cdot|\omega_H)$ and the last inequality must be strict. Hence, $\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0$. Finally, since $\hat{V}(\pi', \mu_H, \mu_H) > \hat{V}(\pi, \mu, \mu_H) \geq \hat{V}^{FD}_H$ we have $\pi' \notin \Pi^{FD}$.$\square$

I now use a signal with the properties in the claim to construct a signal satisfying the statement in Lemma 2. Let $\pi' \in \Pi\backslash\Pi^{FD}$ be a signal with support $S' = S$, which is not silent, and such that $\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) \geq \hat{V}(\pi', \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L)$ and $\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0$. Let $\bar{k} \in \{1, ..., k-1\}$ be the smallest number such that $\pi'_{\bar{k}H} \geq \pi'_{\bar{k}L}$, where, since π' is not silent, $\bar{k} > 1$.

Define a signal ξ with support S such that $\xi_{iH} = 0$ for $i < \bar{k}$ and $\xi_{iH} = \pi'_{iH} + \lambda_H(1 - \pi'_{iH})$ for $i \geq \bar{k}$, where $\lambda_H = \sum_{i<\bar{k}} \pi'_{iH} / \sum_{i\geq\bar{k}}(1 - \pi'_{iH})$ if $\sum_{i\geq\bar{k}}(1 - \pi'_{iH}) > 0$ and $\lambda_H = 0$ otherwise. The definition of λ_H ensures that ξ_{iH} distributes the probability mass $\pi'(\cdot|\omega_H)$ puts on outcomes $i < \bar{k}$ and distributes it over outcomes $i \geq \bar{k}$. Define ξ_{iL} analogously by $\xi_{iL} = 0$ for $i \geq \bar{k}$ and $\xi_{iL} = \pi'_{iL} + \lambda_L(1 - \pi'_{iL})$ for $i < \bar{k}$, where $\lambda_L = \sum_{i\geq\bar{k}} \pi'_{iL} / \sum_{i<\bar{k}}(1 - \pi'_{iL})$ if $\sum_{i<\bar{k}}(1 - \pi'_{iL}) > 0$ and $\lambda_L = 0$ otherwise. Then $\xi \in \Pi^{FD}$ with $\xi_{iH} \leq \pi'_{iH} \leq \pi'_{iL} \leq \xi_{iL}$ for $i < \bar{k}$ and $\xi_{iL} \leq \pi'_{iL} \leq \pi'_{iH} \leq \xi_{iH}$ for $i \geq \bar{k}$ and since $\pi' \notin \Pi^{FD}$ we have $\xi \neq \pi'$.

Define a family of signals $g(\gamma)$ with support S by $g_{iH}(\gamma) = (1-\gamma)\pi'_{iH} + \gamma\xi_{iH}$ and $g_{iL}(\gamma) = (1-\gamma)\pi'_{iL} + \gamma\xi_{iL}$ for $\gamma \in [0, 1]$. Then $g(0) = \pi'$ and $g(1) = \xi \in \Pi^{FD}$. For any $\gamma' > \gamma$ we have that

$g(\gamma')$ is strictly more precise than $g(\gamma)$. By Lemma 1 and by hypothesis, for all $\gamma \in (0,1]$

$$\hat{V}(g(\gamma), \mu_H, \mu_H) - \hat{V}(g(\gamma), \mu_H, \mu_L) > \hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi', \mu_H, \mu_L) \geq \hat{V}(\pi, \mu, \mu_H) - \hat{V}(\pi, \mu, \mu_L),$$

and therefore $\hat{V}(g(\gamma), \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > \hat{V}(g(\gamma), \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L)$. Since $\hat{V}(g(\gamma), \mu_H, \mu_H)$ is continuous in γ and since $\hat{V}(g(0), \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0$ and $\hat{V}(g(1), \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) \leq 0$ there is some $\hat{\gamma} \in (0, 1]$ such that

$$0 = \hat{V}(g(\hat{\gamma}), \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > \hat{V}(g(\hat{\gamma}), \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L).$$

Let $\bar{\gamma} = \min\{\gamma \geq 0 : \hat{V}(g(\gamma), \mu_H, \mu_H) = \hat{V}(\pi, \mu, \mu_H)\}$ so $\hat{V}(g(\gamma), \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0$ for all $\gamma \in [0, \bar{\gamma})$ and $\hat{V}(g(\bar{\gamma}), \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L) < 0$. There is then some $\varepsilon \in (0, \bar{\gamma})$ such that

$$\hat{V}(g(\bar{\gamma} - \varepsilon), \mu_H, \mu_H) - \hat{V}(\pi, \mu, \mu_H) > 0 > \hat{V}(g(\bar{\gamma} - \varepsilon), \mu_H, \mu_L) - \hat{V}(\pi, \mu, \mu_L)$$

which concludes the proof. ■

Proof. (Proposition 1) Suppose, to contradiction, that $\pi^L = \pi^H = \pi \in \Pi \backslash \Pi^{FD}$, α and $(\widetilde{\beta}(\pi), \hat{\beta}(\pi, s))$ is an equilibrium. Then $\widetilde{\beta}(\pi) = \mu_0$, $\hat{\beta}(\pi, s_i) = B(\pi, s_i, \mu_0)$ for all $s_i \in S$ and the equilibrium payoff of a type μ_j sender is $\hat{V}(\pi, \mu_0, \mu_i)$. Since we are in equilibrium $\hat{V}(\pi, \mu_0, \mu_H) \leq \hat{V}_H^{FD}$. By Lemma 2 there is some $\pi' \in \Pi$ such that $\hat{V}(\pi', \mu_H, \mu_H) - \hat{V}(\pi, \mu_0, \mu_H) > 0 > \hat{V}(\pi', \mu_H, \mu_L) - \hat{V}(\pi, \mu_0, \mu_L)$, which by (iv) in Definition 1 implies $\widetilde{\beta}(\pi') = \mu_H$. Type μ_H therefore has a profitable deviation to π', a contradiction. ■

Proof. (Theorem 1) First, notice that Π is a compact subset of Euclidean space and that

$$\hat{V}(\pi, \mu, \mu_j) = \sum_{s \in \widetilde{S}} \left[\mu_j \pi(s|\omega_H) + (1 - \mu_j) \pi(s|\omega_H) \right] \hat{v}(B(\pi, s, \mu))$$

is continuous in π. Hence, $\max_{\pi \in \Pi} \hat{V}(\pi, \mu, \mu_j)$ is well defined and therefore $\Pi_j^* \neq \varnothing$ for $j \in \{L, H\}$. Let $\pi_L^* \in \Pi_L^*$ and let $D := \{\pi \in \Pi : \hat{V}(\pi, \mu_H, \mu_L) \leq \hat{V}(\pi_L^*, \mu_L, \mu_L)\}$, where $\Pi^{FD} \subset D \neq \varnothing$.

Step 1. The "if" part of (i). Proof. Suppose $\pi^L = \pi_L^*$, that $\pi^H \in \arg\max_{\pi \in D} \hat{V}(\pi, \mu_H, \mu_H)$, and that the optimally responding receiver's beliefs $(\widetilde{\beta}, \hat{\beta})$ are rational. Notice that if $\pi^L = \pi^H$ then $\pi^L, \pi^H \in \Pi^{FD}$ since otherwise $\hat{V}(\pi^H, \mu_H, \mu_L) > \hat{V}(\pi^L, \mu_L, \mu_L)$. Hence, type μ_j's expected payoff is $\hat{V}(\pi^j, \mu_j, \mu_j)$ for $j = L, H$. Notice that if $\hat{V}(\pi, \mu_H, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H)$ then $\pi \notin D$, so (iv) in Definition 1 imposes no restriction on $\widetilde{\beta}$. Therefore, let $\widetilde{\beta}(\pi) = \mu_L$ for all $\pi \in \Pi \backslash \{\pi^L, \pi^H\}$. Type μ_L then has no profitable deviation by construction.

Suppose, to contradiction, that type μ_H has a profitable deviation π'. Since $\Pi^{FD} \subset D$ the definition of π^H implies $\hat{V}(\pi', \mu_L, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H) \geq \hat{V}_H^{FD}$ and therefore $\pi' \notin \Pi^{FD}$. Lemma 2 then implies the existence of some $\pi'' \in \Pi$ such that

$$\hat{V}(\pi'', \mu_H, \mu_H) - \hat{V}(\pi', \mu_L, \mu_H) > 0 > \hat{V}(\pi'', \mu_H, \mu_L) - \hat{V}(\pi', \mu_L, \mu_L).$$

Then $\hat{V}(\pi'', \mu_H, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H)$ and $\hat{V}(\pi^L, \mu_L, \mu_L) \geq \hat{V}(\pi', \mu_L, \mu_L) > \hat{V}(\pi'', \mu_H, \mu_L)$, so $\pi'' \in D$, contradicting the definition of π^H. Therefore type μ_H has no profitable deviation and we have an equilibrium.$\square$

Step 2. The "only if" part of (i). Proof. Suppose (π^L, π^H), α, $\widetilde{\beta}(\cdot)$ and $\hat{\beta}(\cdot, \cdot)$ and is an equilibrium. Proposition 1 implies that if $\pi^L = \pi^H$ then $\hat{V}(\pi^j, \mu_0, \mu_j) = \hat{V}_j^{FD} = \hat{V}(\pi^j, \mu_j, \mu_j)$. Hence, type μ_j's equilibrium expected payoff is $\hat{V}(\pi^j, \mu_j, \mu_j)$ for $j = L, H$. Notice that $\hat{V}(\pi^L, \mu_L, \mu_L) \leq \hat{V}(\pi_L^*, \mu_L, \mu_L) \leq \hat{V}(\pi_L^*, \widetilde{\beta}(\pi_L^*), \mu_L)$ and hence $\pi^L \in \Pi_L^*$. Further, $\pi^H \in D$, since otherwise type μ_L has a profitable deviation to π^H. Suppose therefore, to contradiction, that there is some $\pi \in \Pi$ such that $\hat{V}(\pi, \mu_H, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H)$ and $\hat{V}(\pi, \mu_H, \mu_L) \leq \hat{V}(\pi^L, \mu_L, \mu_L)$ and notice that $\hat{V}(\pi, \mu_H, \mu_H) > \hat{V}_H^{FD}$ and consequently $\pi \notin \Pi^{FD}$. Let $g(\gamma)$ be a convex combination of π and some $\pi^{FD} \in \Pi^{FD}$ parameterized by γ and such that $g(0) = \pi$, $g(1) = \pi^{FD}$ and $g(\gamma')$ is strictly more precise than $g(\gamma)$ if $\gamma' > \gamma$. For the details on how such a signal can be constructed, see the proof of Lemma 2. Notice that $\hat{V}(g(0), \mu_H, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H)$ and $\hat{V}(g(1), \mu_H, \mu_H) \leq \hat{V}(\pi^H, \mu_H, \mu_H)$.

Let $\bar{\gamma} = \min\{\gamma \geq 0 : \hat{V}(g(\gamma), \mu_H, \mu_H) = \hat{V}(\pi^H, \mu_H, \mu_H)\}$. Then

$$\begin{aligned} \hat{V}(g(\bar{\gamma}), \mu_H, \mu_H) - \hat{V}(\pi^H, \mu_H, \mu_H) &> \\ \hat{V}(g(\bar{\gamma}), \mu_H, \mu_H) - \hat{V}(\pi, \mu_H, \mu_H) &> \hat{V}(g(\bar{\gamma}), \mu_H, \mu_L) - \hat{V}(\pi, \mu_H, \mu_L) \\ &\geq \hat{V}(g(\bar{\gamma}), \mu_H, \mu_L) - \hat{V}(\pi^L, \mu_H, \mu_L), \end{aligned}$$

where the second inequality follows from (ii) in Lemma 1. There is then some $\varepsilon \in (0, \bar{\gamma})$ such that

$$\hat{V}(g(\bar{\gamma} - \varepsilon), \mu_H, \mu_H) - \hat{V}(\pi^H, \mu_H, \mu_H) > 0 > \hat{V}(g(\bar{\gamma} - \varepsilon), \mu_H, \mu_L) - \hat{V}(\pi^L, \mu_H, \mu_L),$$

i.e., $\hat{V}(g(\bar{\gamma} - \varepsilon), \mu_H, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H)$ and $\hat{V}(g(\bar{\gamma} - \varepsilon), \mu_H, \mu_L) < \hat{V}(\pi^L, \mu_L, \mu_L)$. Then (iv) in Definition 1 implies $\widetilde{\beta}(g(\bar{\gamma} - \varepsilon)) = \mu_H$ and a profitable deviation to $g(\bar{\gamma} - \varepsilon)$ for type μ_H, contradicting that we have an equilibrium.$\square$

Step 3. Part (ii): There is an equilibrium. Proof. Since $\hat{V}(\pi, \mu_H, \mu_L)$ is continuous in π and Π is a compact subset of the Euclidean space we have that D is compact and $\max_{\pi \in D} \hat{V}(\pi, \mu_H, \mu_H)$ therefore has a solution. Hence, the problem of finding (π^L, π^H) such that

$\pi^L \in \arg\max_{\pi \in \Pi} \hat{V}(\pi, \mu_L, \mu_L)$ and $\pi^H \in \arg\max_{\pi \in \Pi}\{\hat{V}(\pi, \mu_H, \mu_H) : \hat{V}(\pi, \mu_H, \mu_L) \leq \hat{V}(\pi^L, \mu_L, \mu_L)\}$

has a solution and Step 1 implies that there is an equilibrium. $\blacksquare$

Proof. (Proposition 2) The proof relies on the characterization of optimal signals in terms of the "concavification" of $\hat{v}$ developed by Kamenica and Gentzkow (2011). Let $\varsigma(x)$ be the line drawn from $(0, \hat{v}(0))$ to $(1, \hat{v}(1))$. Corollary 1 and 2 in Kamenica and Gentzkow (2011) imply that a fully disclosing signal is optimal for type μ_j if and only if $\varsigma(x) \geq \hat{v}(x)$ for all $x \in [0,1]$ and in this case $\max_{\pi \in \Pi} \hat{V}(\pi, \mu_j, \mu_j) = \varsigma(\mu_j)$ for $j = L, H$. This immediately implies that if a fully disclosing signal π^{FD} is optimal for type μ_H then π^{FD} is optimal also for type μ_L, and vice versa.

To prove (i), suppose a fully disclosing signal is optimal for both types and that (π^L, π^H) such that $\pi^H \notin \Pi^{FD}$ is an equilibrium strategy. Let $\{s_1^H, ..., s_{k_H}^H\}$ be the support of π^H and let $\{\beta_i^H\}_{i=1}^{k_H}$ be the updated beliefs induced by π^H in equilibrium. Then $\hat{V}(\pi^j, \mu_j, \mu_j) = \varsigma(\mu_j)$ for

$j = L, H$ and it follows from Corollary 2 in Kamenica and Gentzkow (2011) that $\hat{v}(\beta_i^H) = \varsigma(\beta_i^H)$ for any $i \in \{1, ..., k_H\}$. If type μ_L deviates to π^H his payoff therefore equals $\varsigma(\mu^*)$ with $\mu^* = \sum_{i=1}^{k_H}(\mu_L \pi_{iH}^H + (1-\mu_L)\pi_{iL}^H)\beta_i^H$. We have

$$\begin{aligned}\mu_H - \mu^* &= \sum_{i=1}^{k_H}(\mu_H \pi_{iH}^H + (1-\mu_H)\pi_{iL}^H)\beta_i^H - \sum_{i=1}^{k_H}(\mu_L \pi_{iH}^H + (1-\mu_L)\pi_{iL}^H)\beta_i^H \\ &= (\mu_H - \mu_L)\sum_{i=1}^{k_H}(\pi_{iH}^H - \pi_{iL}^H)\beta_i^H < \mu_H - \mu_L,\end{aligned}$$

where $\sum_i(\pi_{iH}^H - \pi_{iL}^H)\beta_i^H < 1$ since π^H is not fully disclosing. Hence, $\mu^* > \mu_L$ and the deviation is profitable, contradicting that (π^L, π^H) is an equilibrium strategy.

Necessity in (ii) follows since a fully disclosing signal is optimal for μ_L if and only if it also optimal for type μ_H and Theorem 1 requires type μ_L's equilibrium signal to be optimal. Sufficiency is an immediate consequence of Theorem 1. ■

Proof. (Proposition 3) Suppose (π^L, π^H) is an equilibrium sender strategy and that the support of π^j is $\{s_1^j, ..., s_{k_j}^j\}$ for $j = L, H$. Suppose, to contradiction, that $\pi^H \in \Pi_H^* \backslash \Pi^{FD}$. The proof again relies on the characterization of optimal signals developed by Kamenica and Gentzkow (2011). For $j = L, H$ let $\{\beta_i^j\}_{i=1}^{k_j}$ be the updated beliefs induced by π^j in equilibrium, where $\beta_1^j \leq \mu_j \leq \beta_{k_j}^j$, so $\beta_1^L < \beta_{k_H}^H$. Let $\varsigma_j(\cdot)$ be the (positively sloped) line going through $(\beta_1^j, \hat{v}_1^j(\beta_1^j))$ and $(\beta_{k_j}^j, \hat{v}_{k_j}^j(\beta_{k_j}^j))$ and defined on $[0,1]$. Corollary 1 and 2 in Kamenica and Gentzkow (2011) imply that since π^L and π^H are optimal signals $\varsigma_j(x) \geq \hat{v}(x)$ for $x \in [0,1]$ and $\hat{v}(\beta_i^j) = \varsigma(\beta_i^j)$ for $j \in \{L, H\}$ and for all $i \in \{1, ..., k_j\}$. Suppose first that $\varsigma_L(1) > \varsigma_H(1)$. It follows that $\varsigma_L(\hat{x}) = \varsigma_H(\hat{x})$ for some $\hat{x} \in (0,1)$ and since $\varsigma_H(x) > \varsigma_L(x) \geq \hat{v}(x)$ for $x \in [0, \hat{x})$ and $\varsigma_L(x) > \varsigma_H(x) \geq \hat{v}(x)$ for $x \in (\hat{x}, 1]$ we have $\beta_1^H \geq \hat{x} \geq \beta_{k_L}^L$. But then μ_L has a profitable deviation to π^H, a contradiction. Suppose instead that $\varsigma_L(1) < \varsigma_H(1)$. Again $\varsigma_L(\hat{x}) = \varsigma_H(\hat{x})$ for some $\hat{x} \in (0,1)$ and now $\mu_L \geq \beta_1^L \geq \hat{x} \geq \beta_{k_H}^H \geq \mu_H$, a contradiction. Hence, $\varsigma_L(1) = \varsigma_H(1)$. But then $\varsigma_L(x) = \varsigma_H(x)$ for all $x \in [0,1]$, since otherwise either $\beta_1^L = 1 > \mu_L$ or $\beta_1^H = 1 > \mu_H$, a contradiction.

Let $\varsigma(x) := \varsigma_L(x) = \varsigma_H(x)$. Since $\sum_{i=1}^{k_j}(\mu_j \pi_{iH}^j + (1-\mu_j)\pi_{iL}^j)\beta_i^j = \mu_j$ we have $\hat{V}(\pi^j, \mu_j, \mu_j) =$

$\varsigma(\mu_j)$ for $j = L, H$. If type μ_L deviates to π^H his deviation payoff equals $\varsigma(\mu^*)$ with $\mu^* = \sum_{i=1}^{k_H}(\mu_L\pi_{iH}^H + (1-\mu_L)\pi_{iL}^H)\beta_i^H$. We have

$$\begin{aligned}\mu_H - \mu^* &= \sum_{i=1}^{k_H}(\mu_H\pi_{iH}^H + (1-\mu_H)\pi_{iL}^H)\beta_i^H - \sum_{i=1}^{k_H}(\mu_L\pi_{iH}^H + (1-\mu_L)\pi_{iL}^H)\beta_i^H \\ &= (\mu_H - \mu_L)\sum_{i=1}^{k_H}(\pi_{iH}^H - \pi_{iL}^H)\beta_i^H < \mu_H - \mu_L,\end{aligned}$$

where $\sum_i(\pi_{iH}^H - \pi_{iL}^H)\beta_i^H < 1$ since π^H is not fully disclosing. Hence, $\mu^* > \mu_L$ and the deviation is profitable, contradicting that (π^L, π^H) is an equilibrium strategy. ■

Proof. (Proposition 4) Suppose (π^L, π^H) is an equilibrium sender strategy. Suppose, to contradiction, that π is more precise than π^H and that $\hat{V}(\pi, \mu_H, \mu_H) > \hat{V}(\pi^H, \mu_H, \mu_H)$. Notice that $\hat{V}(\pi, \mu_H, \mu_H) > \hat{V}_H^{FD}$, so $\pi \notin \Pi^{FD}$. Let $g(\gamma)$ be a convex combination of π and some $\pi^{FD} \in \Pi^{FD}$ parameterized by $\gamma \in [0,1]$ and such that $g(0) = \pi$, $g(1) = \pi^{FD}$ and $g(\gamma')$ is strictly more precise than $g(\gamma)$ if $\gamma' > \gamma$. For the details on how such a signal can be constructed, see the proof of Lemma 2. Let $\bar{\gamma} = \min\{\gamma \geq 0 : \hat{V}(g(\gamma), \mu_H, \mu_H) = \hat{V}(\pi^H, \mu_H, \mu_H)\}$. Since $g(\bar{\gamma})$ is strictly more precise than π^H (ii) in Lemma 1 implies the first of the following inequalities, while type μ_L's incentive compatibility implies the second

$$\begin{aligned}\hat{V}(g(\bar{\gamma}), \mu_H, \mu_H) - \hat{V}(\pi^H, \mu_H, \mu_H) &> \\ \hat{V}(g(\bar{\gamma}), \mu_H, \mu_L) - \hat{V}(\pi^H, \mu_H, \mu_L) &\geq \hat{V}(g(\bar{\gamma}), \mu_H, \mu_L) - \hat{V}(\pi^L, \mu_L, \mu_L),\end{aligned}$$

There is then some $\varepsilon \in (0, \bar{\gamma})$ such that

$$\hat{V}(g(\bar{\gamma}-\varepsilon), \mu_H, \mu_H) - \hat{V}(\pi^H, \mu_H, \mu_H) > 0 > \hat{V}(g(\bar{\gamma}-\varepsilon), \mu_H, \mu_L) - \hat{V}(\pi^L, \mu_L, \mu_L),$$

but this contradicts (2) in Theorem 1, i.e., (π^L, π^H) is not an equilibrium sender strategy. ■

References

[1] Alonso, R. and Câmara, A., 2014a. Bayesian Persuasion with Heterogeneous Priors. *Mimeo, University of Southern California.*

[2] Alonso, R. and Câmara, A., 2014b. On the Value of Persuasion by Experts. *Mimeo, University of Southern California.*

[3] Argenziano, R., Severinov, S. and Squintani, F., 2014. Strategic Information Acquisition and Transmission. *Mimeo.*

[4] Austen-Smith, D., 1994. Strategic Transmission of Costly Information. *Econom. 62*, 955-963.

[5] Blackwell, 1951. Comparison of Experiments. In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93-102.

[6] Che, Y-K. and Kartik, N., 2009. Opinions as Incentives. *J. Polit. Econ. 117*, 815-860.

[7] Cho, I.-K. and Kreps, D., 1987. Signalling and Stable Equilibria. *Q. J. of Econ. 102*, 179-221.

[8] Crawford, V. and Sobel J., 1982. Strategic Information Transmission. *Econom. 50*, 1431-1451.

[9] Gentzkow, M. and Kamenica, E. 2014. Costly Persuasion. *Am. Econ. Rev. P&P 104*, 457-462.

[10] Gentzkow, M. and Kamenica, E. 2012. Competition in Persuasion. *Mimeo, U. Chicago.*

[11] Gill, D. and Sgroi, D., 2008. Sequential decisions with tests. *Games and Econ. Behav. 63*, 663-678.

[12] Gill, D. and Sgroi, D., 2012. The optimal choice of pre-launch reviewer. *J. of Econ. Theory 147*, 1247-1260.

[13] Hagenbach, J., Koessler, F. and Perez-Richet, E. 2014. Certifiable Pre-Play Communication: Full Disclosure. *Econom. 82*, 1093-1131.

[14] Hedlund, J., 2014. Persuasion with communication costs. *Mimeo, University of Heidelberg.*

[15] Henry, E., 2009. Strategic Disclosure of Research Results: The Cost of Proving Your Honesty. *The Econ. J. 119,* 1036-1064.

[16] Ivanov, M., 2010. Informational control and organizational design. *J. of Econ. Theory 145*, 721-751.

[17] Kamenica, E. and Gentzkow, M., 2011. Bayesian Persuasion. *Am. Econ. Rev. 101*, 2590-2615.

[18] Kolotilin, A., 2014a. Experimental Design to Persuade. *Mimeo, University of New South Wales.*

[19] Kolotilin, A., 2014b. Optimal Information Disclosure: Quantity vs. Quality. *Mimeo, University of New South Wales.*

[20] Li, H. and Li, W., 2013. Misinformation. *Int. Econ. Rev. 54*, 253-277.

[21] Mailath, G., 1987. Incentive Compatibility in Signaling Games with a Continuum of Types. *Econom. 55*, 1349-1365.

[22] Milgrom, P., 1981. Good News and Bad News: Representation Theorems and Applications. *The Bell J. of Econ. 12*, 380-391.

[23] Milgrom, P., 2008. What the Seller Won't Tell You: Persuasion and Disclosure in Markets. *J. of Econ. Perspect. 22*, 115-131.

[24] Milgrom, P. and Roberts, J. 1986. Relying on the information of interested parties. *Rand J. of Econ. 17*, 18-32.

[25] Perez-Richet, E. 2014. Interim Bayesian Persuasion: First Steps. *Am. Econ. Rev. P&P 104*, 469-474.

[26] Rayo, L. and Segal, I., 2010. Optimal Information Disclosure. *J. of Polit. Econ. 188*, 949-987.

[27] Seidmann, D. and Winter, E., 1997. Strategic Information Transmission with Verifiable Messages. *Econom. 65*, 163-169.

[28] Shannon, C., 1948. A Mathematical Theory of Communication. *Bell Syst. Tech. J. 27*, 379–423.

[29] Spence, M., 1973. Job Market Signaling. *Q. J. of Econ. 87*, 355-374.