Bayesian signaling

Jonas Hedlund*

February 2015

Abstract

This paper introduces private sender information in 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 selected by the D1 criterion. The sender’s equilibrium strategy consists of signals which 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. Whether the equilibrium signals are separating or fully disclosing is completely determined by the optimality properties of fully disclosing signals. Incentive compatibility requires the sender to use suboptimal signals in any equilibrium which is not fully disclosing and then generates a cost for the sender in comparison to a full information benchmark in which the receiver knows the sender’s type.

Keywords: Bayesian Persuasion, Signaling, Information Transmission.

JEL Classification: D82, D83, D86.

Acknowledgement: I thank Carlos Oyarzun, Cristoph Vanberg and seminar participants at the University of Heidelberg for many helpful comments and suggestions.

*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
1 Introduction

The recent literature on Bayesian persuasion pioneered by Kamenica and Gentzkow (2011) studies the strategic design of experiments by a sender whose objective is to influence the beliefs of a receiver. In contrast to most established models on information transmission (e.g., Spence 1973; Milgrom 1981; Crawford and Sobel 1982), sender and receiver share the same prior information and the sender influences the receiver by committing to an informative experiment. The experiment may consist of, e.g., a public signal, a protocol of information acquisition, or independent research sponsored by the sender. Situations fitting the description include firms specifying the terms of free trials of new products, or interest groups funding research for lobbying purposes. In both examples the sender (firm or interest group) controls the type of information generated and is able to commit to reveal the generated information to the receiver. Applications of the framework has lead to novel insights in, e.g., collective decision making (Alonso and Câmara 2014c) and industrial organization (Bergemann, Brooks and Morris 2014).

In some situations, however, it is unlikely that sender and receiver would share the same prior information. For example, a firm specifying terms (e.g., length or functionality) of free trials upon launching a new software is likely better informed than a typical consumer about the user-friendliness of the software. An interest group promoting a product may have private information regarding health concerns prior to funding further independent research on the subject. How would such private sender information impact the equilibrium design of experiments? To which extent is it possible to infer the sender’s private information from the type of experiment he runs?\footnote{I.e., will the terms of the free trials signal the firm’s private information about user-friendliness? Can the interest group’s private information about health risks be inferred from the design of the research project it funds?} The present paper shows that private sender 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).

The framework consists of a simple model of Bayesian persuasion in which the sender has unverifiable and imperfect private information about a binary payoff-relevant state, prior to gen-
erating further information about this state. Contingent on his private information the sender chooses a random signal (i.e., an experiment) which can be informative about the payoff-relevant state. The receiver observes an outcome of the signal, updates her beliefs, and the sender’s collects a payoff which is continuous and strictly increasing in the receiver’s updated belief.\footnote{Such monotonic preferences play an important role in the literature on transmission of verifiable information, see, e.g., Milgrom (1981) and Milgrom and Roberts (1986), as well as in several signaling models, see, e.g., Spence (1973), Mailath (1987) and Cho and Sobel (1990).}

I show that confining attention to equilibria selected by Cho and Kreps’ (1987) D1 criterion (henceforth, "equilibria") leads to a number of predictions about 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, the sender’s equilibrium strategy consists of signals which are either separating or fully reveal the payoff-relevant state. I.e., either the sender’s choice of signal reveals his private information to the receiver, or the outcome of the chosen signal fully reveals the payoff-relevant state. The result follows from conditions of increasing differences in expected payoffs which arise endogenously in equilibrium and roughly state that sender types with more favorable private information have stronger preferences for more precise signals (Lemmata 1 and 2). The increasing differences in precision imply that much of the logic of standard signaling games applies. Roughly, the D1 criterion requires the receiver to attribute deviations to sufficiently precise signals to sender types with favorable private information, making it possible to find profitable deviations for such types from most pooling strategies.

The sender’s private information is assumed to be finite, leading to a finite type-space in which sender types are ordered according to how favorable their private information is.\footnote{A previous version of this paper analyzed an identical model with a binary type-space, obtaining an identical set of results by confining attention to equilibria selected by Cho and Kreps’ (1987) intuitive criterion rather than the D1 criterion. I.e., in the binary case the same results hold also under a weaker solution concept.} 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 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 lowest sender type chooses an optimal signal, while each remaining sender type chooses a signal
maximizing this type’s payoff under a local upward incentive compatibility constraint. Whether the constraint forces the sender to choose suboptimal signals is completely determined by the optimality properties of signals which fully reveal the payoff-relevant state, i.e., of fully disclosing signals. Whenever fully disclosing signals are optimal for all types the equilibrium signals are optimal and fully disclosing, and otherwise they are neither (Proposition 2). I.e., either the sender provides the receiver with perfect information, or the equilibrium is fully separating and incentive compatibility generates a cost for the sender in comparison to the full information benchmark, as in many models of costly signaling (e.g., Spence 1973). The 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.

Private information thus leads to a typical incentive compatibility dilemma in which the sender must incur a cost to signal his prior. The cost is not of the money burning type, but is associated with a loss of control over the receiver’s informational environment. This loss of control sometimes benefits the receiver by providing her with a more informative information structure. This occurs, e.g., if the sender’s payoff function is strictly concave in the receiver’s updated belief, in which case all sender types would choose uninformative signals in the full information benchmark. Paradoxically, in such cases the sender strictly prefers the full information benchmark over the equilibrium, while the receiver strictly prefers the equilibrium, i.e., the receiver collects an "ignorance rent." This may lead to odd situations in which the sender would like to disclose his private information prior to choosing a signal, and the receiver would prefer not to listen.

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 any sender type in equilibrium is always more favorable than the least favorable updated belief induced by any higher type, i.e., the beliefs different types induce must overlap. There is therefore a sense in which "good news" from low types are better than "bad news" from high types. Second, and related, an uninformative signal is neither the worst nor the best piece of equilibrium news. This holds since only the lowest type’s equilibrium signal can be
uninformative. For the induced beliefs to overlap higher types must consequently induce some belief less favorable than the one corresponding to the lowest type’s uninformative signal. Hence, while an uninformative signal is viewed skeptically, as in Milgrom’s (1981) unraveling result, it 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 in which two main properties drive many of the results. 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 the same signal outcome 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, the sender’s expected payoff given a fully disclosing signal is independent of the receiver’s belief regarding his type. Such signals establish type-specific lower bounds on equilibrium expected payoffs. The increasing differences imply that much of the logic of signaling games in which the D1 criterion is known to select clear predictions (e.g., Spence 1973; Cho and Sobel 1990) applies here. Combining with the lower bounds on equilibrium expected payoffs leads to the characterizations of equilibria discussed above.

While most of the analysis is concerned with an imperfectly informed sender, allowing the sender to sometimes become perfectly informed is straightforward and leads to particularly sharp results. If the sender is certain that the payoff-relevant state is unfavorable with positive (even arbitrarily small) probability, then any equilibrium perfectly reveals the payoff-relevant state (Proposition 4). Intuitively, in this case there is a sender type which has everything to win by mimicking other types and which can only be discouraged if these types use fully disclosing signals. The resulting outcome is far from optimal for the sender, while the receiver has perfect information and could not be better off.

Finally, in the absence of private information it would be without loss of generality to confine the analysis to particularly simple signals with only two outcomes, given that the set of payoff-relevant states is binary (Kamenica and Gentzkow 2011; Alonso and Câmara 2014a). A technical side note is that this simplification is no longer valid in the presence of private information. In
particular, by using signals with more than two outcomes the sender can sometimes obtain higher expected payoff while remaining incentive compatible. Private information can thus motivate the design of more complex experiments. This possibility is discussed in more detail in Section 5.4 in the appendix.

**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 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, 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 relationship between the results is discussed in more detail in Section 3.3, which extends the model here to allow a perfectly informed sender.

The concurrent working paper by Alonso and Câmara (2014b) is technically more related. While our models are similar, however, our focuses are different. Alonso and Câmara (2014b) investigate more generally whether the sender can be strictly better off under private information than in its absence, while my objective is to obtain concrete insights concerning equilibrium behavior in more structured environments. Their negative answer is complementary to my result on the suboptimality of equilibrium signals, the main difference being that Alonso and Câmara (2014b) compare 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 circumstances which might weaken the sender’s control of the receiver’s information. 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) and Li and Li (2012) impose related constraints on the set of feasible signals.

This paper also relates to work on optimal information structures preceding Kamenica and Gentzkow (2011). E.g., Rayo and Segal (2010) examine the optimal disclosure of information by a firm choosing the informativeness of its ads. Ostrowsky and Schwartz (2008) study schools’ equilibrium revelation of information about students through the design of grade transcripts. Brocas and Carrillo (2007) characterize the rents that can be obtained by controlling the flow of public information.

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 initially agree that the prior probability of \(\omega_H\) equals \(\mu_0 \in (0, 1)\). The sender then privately observes one of \(n \geq 2\) unverifiable outcomes of an exogenous signal which is informative about \(\omega_j\), and rationally updates \(\mu_0\) to \(\mu_t \in T := \{\mu_1, \mu_2, \ldots, \mu_n\}\), where \(0 < \mu_1 < \ldots < \mu_n < 1\). I refer to \(\mu_t \in T\) as the sender’s type. Sender types are thus ordered according to their beliefs that the payoff-relevant state is \(\omega_H\) and no type is perfectly informed.\(^4\)

Signals and beliefs. After observing his private signal the sender’s objective is to modify the receiver’s prior belief \(\mu_0\). 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 \(\widetilde{S}\), with \(|\widetilde{S}| \geq 2\). Let \(\Pi\) be the set of all signals over \(\widetilde{S}\). The sender is allowed to choose any signal in \(\Pi\). For a generic signal \(\pi\) let \(S = \{s_1, \ldots, s_k\}\) be the support of \(\pi\), defined as \(S := \{s \in \widetilde{S} : \pi(s|\omega_L) + \pi(s|\omega_H) > 0\}\).\(^5\) I abbreviate \(\pi(s_i|\omega_j) = \pi_{ij}\) and assume, for concreteness and without loss of generality, that for \(i < i'\) either \(\pi_{i_H}/\pi_{i_L} \leq \pi_{i'H}/\pi_{i'L}\) or \(\pi_{i_L} = 0\). I.e., signal outcomes are ordered according to likelihood ratios such that higher indexed outcomes correspond to higher posterior probabilities of \(\omega_H\).

The receiver observes the signal \(\pi\) chosen by the sender and makes an interim update of her belief that the payoff-relevant state is \(\omega_H\) to \(\tilde{\beta}(\pi) \in [\mu_1, \mu_n]\).\(^6\) I.e., the receiver is allowed to make inferences about the sender’s private information from the sender’s choice of signal. The receiver

\(^4\)For ease of exposition, the special case in which the sender is allowed to be perfectly informed, i.e., where \(\mu_1 = 0\) or \(\mu_n = 1\), is dealt with separately in Section 3.3.

\(^5\)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. While, as argued in Section 5.4 of the appendix, an upper bound on the number of signal outcomes is not necessarily without loss of generality, 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.

\(^6\)While the interpretation is that the receiver forms a belief regarding the sender’s type, I treat the receiver’s interim belief as a belief regarding the payoff-relevant state, since this provides a shortcut and is without loss of generality as long as \(\tilde{\beta}(\pi) \in [\mu_1, \mu_n]\).
next observes a realization $s_i \in S$ drawn according to $\pi$, and updates to final belief $\widehat{\beta}(\pi, s_i)$. 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 an arbitrary interim belief $\mu \in (0,1)$, to a posterior probability that the state is $\omega_H$. I.e., if the receiver updates to final beliefs according to Bayes rule given interim belief $\tilde{\beta}(\pi)$, then her final belief equals $\widehat{\beta}(\pi, s_i) = B(\pi, s_i, \tilde{\beta}(\pi))$.

Two classes of signals deserve special mention and their own name. First, the sender can fully reveal the payoff-relevant state by choosing a signal $\pi$ such that $\pi_{ij} > 0$ implies $\pi_{ij'} = 0$ for all $s_i \in S$ and $j \neq j'$. For such a signal $B(\pi, s_i, \mu) \in \{0,1\}$ for all $s_i \in S$. The simplest example is a binary signal with $\pi_{1L} = \pi_{2H} = 1$. I refer to such signals as fully disclosing. Let $\pi^{FD}$ denote a generic fully disclosing signal and let $\Pi^{FD} \subset \Pi$ denote the set of fully disclosing signals. This notion of full disclosure differs from the standard notion in persuasion games (see, e.g., Milgrom 2008), where full disclosure typically means using a message which discloses the sender’s type. Since here the sender is not perfectly informed, disclosing the type is not equivalent to disclosing the payoff-relevant state, and while there is a signal disclosing the payoff-relevant state there is no signal disclosing the sender’s type.\(^7\) The sender’s second benchmark option is an uninformative signal, i.e., a signal $\pi$ such that $\pi_{iH} = \pi_{iL}$ for all $s_i \in S$. For such a signal $B(\pi, s_i, \mu) = \mu$ for all $s_i \in S$. 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 new product. It also seems natural if the signal is public, as when a firm provides a public product demonstration, or if the outcome is independently revealed by a third party. Another interpretation is that the signal is an investigative report, prepared by the sender and containing evidence concerning the payoff-relevant state. The sender is not familiar with the

\(^7\) The unraveling arguments in Milgrom (1981), Seidmann and Winter (1997), or Hagenbach, Koessler and Perez-Richet (2014) can consequently not be applied to argue existence or uniqueness of fully separating equilibria.
evidence prior to collecting it and commits to a protocol of investigation (a signal) and to report all collected evidence (the signal’s outcome). As Kamenica and Gentzkow (2011) argue, there are several such situations in which the commitment assumption is plausible. For example, in the US a prosecutor is required by law to disclose any 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.8

**Sender and receiver payoffs.** I assume the sender’s payoff continuous and strictly increasing in the receiver’s final belief. For concreteness, I model this by assuming the sender’s payoff dependent on a receiver action \( a \in A := [a, \pi] \) taken after her final update of beliefs. The receiver collects a payoff depending on her action and the payoff-relevant state and given by a function \( u : \{\omega_L, \omega_H\} \times A \to \mathbb{R} \). I assume that \( u \) is differentiable and strictly concave in its second argument and that \( a < \arg \max_{a \in A} u(\omega_L, a) < \arg \max_{a \in A} u(\omega_H, a) < \pi \). The receiver’s expected payoff given final belief \( \mu \in [0, 1] \) and action \( a \in A \) equals \( U(\mu, a) := (1 - \mu)u(\omega_L, a) + \mu u(\omega_H, a) \). The receiver’s optimal action \( a^R(\mu) := \arg \max_{a \in A} U(\mu, a) \) is then 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 and strictly increasing. The mapping \( \hat{v} : [0, 1] \to \mathbb{R} \) such that \( \hat{v} = v \circ a^R \) then gives the sender’s payoff as a continuous and strictly increasing function of the belief of an optimally responding receiver. E.g., an interpretation is that the sender provides free trials of a product of uncertain quality and extracts a price which is increasing in the receiver’s ultimate belief that the quality is high. Sender preferences which are monotonic in the receiver’s belief play an important role in several models of information transmission (Spence 1973; Milgrom 1981; Cho and Sobel 1990). Combining with the absence of signaling costs implies that no single-crossing condition imposed on the preferences drive the results below.

---

8Another concern is that the sender may have incentives to "renegotiate" and provide additional information once the receiver has updated to final beliefs. As long as the receiver can modify her beliefs with respect to the sender’s type after unexpectedly being offered additional information, however, this possibility would not alter the set of equilibrium outcomes here.
Strategies and expected payoffs. A (pure) sender strategy is a tuple \((\pi^1, \ldots, \pi^n) \in \Pi^n\), where \(\pi^t\) is the signal chosen by type \(\mu_t\). A (pure) receiver strategy is a function \(\alpha : \{\pi \times S\}_{\pi \in \Pi} \rightarrow A\). Given a receiver strategy \(\alpha\) the expected payoff of a type \(\mu_t\) sender using a signal \(\pi\) is given by

\[
V(\pi, \alpha, \mu_t) := \sum_{i=1}^{k} [\mu_t \pi_i H + (1 - \mu_t) \pi_i L] v(\alpha(\pi, s_i)).
\]

It will be useful to define a function which gives the sender’s expected payoff for rational receiver behavior rationally given arbitrary interim belief \(\mu \in (0, 1)\). Let \(\hat{V} : \Pi \times (0,1) \times [0,1] \rightarrow \mathbb{R}\) be defined by

\[
\hat{V}(\pi, \mu, \mu_t) := \sum_{i=1}^{k} [\mu_t \pi_i H + (1 - \mu_t) \pi_i L] \hat{v}(B(\pi, s_i, \mu)).
\]

\(\hat{V}(\pi, \mu, \mu_t)\) gives the expected payoff of a type \(\mu_t\) sender using signal \(\pi\), provided that the receiver updates to final beliefs according to Bayes rule given interim belief \(\mu\) and chooses an action which is optimal given the updated final beliefs. The function \(\hat{V}\) plays a key-role in the equilibrium analysis and it is convenient to make some of its properties explicit. First, \(\hat{V}\) is continuous. Second, given a signal \(\pi\) and interim belief \(\mu\) the set of payoffs \(\{\hat{v}(B(\pi, s_i, \mu))\}_{i=1}^{k}\) induced by \(\pi\) is independent of the sender’s type \(\mu_t\). \(\hat{V}\) is consequently linear in type and the type-dependence arises only since different sender types induce different convex combinations of the induced payoffs. Third, the expected payoff of a type \(\mu_t\) sender using a fully disclosing signal equals \(\hat{V}^{FD}_t := \hat{V}(\pi^{FD}, \mu, \mu_t) = (1 - \mu_t) \hat{v}(0) + \mu_t \hat{v}(1)\), which is independent of the receiver’s interim belief \(\mu\). For any other signal it holds that \(\hat{V}(\pi, \cdot, \mu_t)\) is strictly increasing.

Equilibrium. The solution concept is perfect Bayesian equilibrium (PBE) selected by Cho and Kreps’ (1987) D1 criterion.

Definition 1 A PBE is a sender strategy \((\pi^1, \ldots, \pi^n)\), a receiver strategy \(\alpha\) and receiver interim and final beliefs \((\hat{\beta}, \hat{\beta})\) such that (i) \(\alpha(\pi, s) = a^{R}(\hat{\beta}(\pi, s))\) for all \((\pi, s) \in \{\pi \times S\}_{\pi \in \Pi}\). (ii) \(\pi^t \in \ldots\)

The focus on pure strategies is mostly for notational convenience. E.g., allowing the sender to randomize over finite sets of signals does not significantly alter the results presented below.
arg max_{\pi \in \Pi} V(\pi, \alpha, \mu_t) for all t \in \{1, ..., n\} and (iii) the receiver’s beliefs are rational, i.e.,

$$\tilde{\beta}(\pi) = \frac{\sum_{t: \pi^t = \pi} \mu_t \Pr(\mu_t)}{\sum_{t: \pi^t = \pi} \Pr(\mu_t)}$$

for all \(\pi \in \{\pi^1, ..., \pi^n\}\) and \(\tilde{\beta}(\pi, s) = B(\pi, s, \tilde{\beta}(\pi))\) for all \((\pi, s) \in \{\pi \times S\}_{\pi \in \Pi}\).

The expected payoff of a type \(\mu_t\) sender in any PBE and given any signal \(\pi\) equals \(V(\pi, \alpha, \mu_t) = \tilde{V}(\pi, \tilde{\beta}(\pi), \mu_t)\). It is thus possible to reformulate the requirement regarding the sender’s strategy in Definition 1 to \(\pi^t \in \arg \max_{\pi \in \Pi} \tilde{V}(\pi, \tilde{\beta}(\pi), \mu_t)\) for all \(t \in \{1, ..., n\}\). This offers a considerable simplification and the equilibrium analysis will therefore henceforth be in terms of \(\tilde{V}\).

The D1 criterion adds a restriction on the receiver’s out-of-equilibrium interim beliefs. Given a sender strategy \((\pi^1, ..., \pi^n)\) and receiver interim beliefs \(\tilde{\beta}(\cdot)\), define for any \(\pi \in \Pi\) and \(\mu_t \in T\) the sets \(D^0(\pi, \mu_t) := \{\mu \in [\mu_1, \mu_n] : \tilde{V}(\pi, \mu, \mu_t) \geq \tilde{V}(\pi^t, \tilde{\beta}(\pi^t), \mu_t)\}\) and \(D(\pi, \mu_t) := \{\mu \in [\mu_1, \mu_n] : \tilde{V}(\pi, \mu, \mu_t) > \tilde{V}(\pi^t, \tilde{\beta}(\pi^t), \mu_t)\}\). I.e., given a PBE \(D^0(\pi, \mu_t)\) and \(D(\pi, \mu_t)\) are the receiver interim beliefs such that type \(\mu_t\) weakly and strictly, respectively, prefers deviating to \(\pi\), provided that the receiver behaves optimally given her interim belief.\(^{10}\)

**Definition 2** A PBE is selected by the D1 criterion if for any out-of-equilibrium signal \(\pi\) and sender types \(\mu_t\) and \(\mu_{t'}\), \(D^0(\pi, \mu_t) \subset D(\pi, \mu_{t'})\) implies that \(\tilde{\beta}(\pi)\) is a convex combination of \(T \setminus \{\mu_t\}\).

The D1 criterion requires the receiver not to attribute a deviation to a sender type, if there is another type willing to make the deviation for a strictly larger set of receiver responses. The idea is that the receiver considers it more likely that a deviation would come from types which are more likely to profit from this particular deviation. PBE selected by the D1 criterion constitutes a reference point in a wide range of signaling models (e.g., Bernheim 1994; Bernheim and Severinov 2003; Austen-Smith and Fryer 2005; Kartik 2009).

Henceforth, I refer to a PBE selected by the D1 criterion simply as an *equilibrium*. A sender strategy \((\pi^1, ..., \pi^n)\) consistent with some equilibrium is an *equilibrium sender strategy*. If \(\pi^{t'} \neq \pi^t\)

---

\(^{10}\)These objects are usually defined in terms of receiver actions, but here it is equivalent and more convenient to define them in terms of receiver interim beliefs.
for all $t' \neq t$ I say that $\pi^t$ is a separating. An equilibrium is separating if the sender’s strategy consists of separating signals and an equilibrium is fully disclosing if the sender’s strategy consists of fully disclosing signals.

3 Analysis

The predictions of the model are closely linked to the properties of the function $\hat{V}$. Two main properties together drive many of the results below. First, the fact that $\hat{V}$ is independent of the receiver’s interim belief for any fully disclosing signal implies that any type $\mu_t$ earns $\hat{V}_t^{FD}$ by deviating from any equilibrium to a fully disclosing signal. The availability of fully disclosing signals therefore establish type-specific lower bounds $\hat{V}_t^{FD}$ on equilibrium expected payoffs. Second, the linear dependence of $\hat{V}$ in type can be structured by conditions of increasing differences over an ordering of the signals according to a dimension of precision, along which fully disclosing signals represent one extreme. In particular, it roughly holds that higher sender types have stronger preferences for more precise signals. The result is a signaling game with a structure reminiscent of signaling games in which the D1 criterion is known to select clear predictions, such as the class of signaling games analyzed by Cho and Sobel (1990). By combining the lower bounds on equilibrium expected payoffs and the increasing differences with the D1 criterion it is possible both to rule out pooling on signals which are not fully disclosing, and to pin down the equilibrium as the solution to a set of maximization problems. In what follows I first identify conditions of increasing differences of $\hat{V}$ 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 \( \pi' \) is more precise than \( \pi \) if for all \( i \in \{1, \ldots, k\} \) either
\[
\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}\quad \text{or}\quad \pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}
\]
If additionally \( \pi' \neq \pi \), then \( \pi' \) is strictly more precise than \( \pi \). This definition is stronger than Blackwell (1951) informativeness. I.e., if \( \pi' \) is more precise than \( \pi \), then \( \pi' \) is more informative than \( \pi \) in the sense of Blackwell (1951), but the converse is not true.\(^{11}\) In particular, precision disperses the posterior probabilities from the prior. If \( \pi' \) is more precise than \( \pi \), then either
\[
B(\pi', s_i, \mu) \leq B(\pi, s_i, \mu) \leq \mu \quad \text{or} \quad B(\pi', s_i, \mu) \geq B(\pi, s_i, \mu) \geq \mu.
\]

Notice that \( \hat{V} (\pi, \mu, \mu_t) \) can be written as a linear function of \( \mu_t \) 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 \( \pi \) and a receiver interim \( \mu \) depends on the sender’s type. It is not difficult to see that the slope is increasing in the precision of \( \pi \). 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 types \( \mu_{i'} > \mu_i \), receiver interim beliefs \( 0 < \mu' < 1 \) and \( 0 < \mu < 1 \) and signals \( \pi, \pi' \in \Pi \) with support \( S \) and such that \( \pi' \) is more precise than \( \pi \). (i) If for all \( s_i \in S \) either
\[
B(\pi', s_i, \mu') \leq B(\pi, s_i, \mu) \leq \mu \quad \text{or} \quad B(\pi', s_i, \mu') \geq B(\pi, s_i, \mu) \geq \mu,
\]
then
\[
\hat{V} (\pi', \mu', \mu_{i'}) - \hat{V} (\pi, \mu, \mu_t) \geq \hat{V} (\pi', \mu', \mu_t) - \hat{V} (\pi, \mu, \mu_t).
\]

(ii) We have
\[
\hat{V} (\pi', \mu, \mu_{i'}) - \hat{V} (\pi, \mu, \mu_t) \geq \hat{V} (\pi', \mu, \mu_t) - \hat{V} (\pi, \mu, \mu_t),
\]
with a strict inequality if \( \pi' \) is strictly more precise than \( \pi \).

These conditions rely on the responses of a rational receiver and are therefore rather an equilibrium construction than the usual condition of increasing differences imposed on the primitives. The lemma implies, e.g., that if type \( \mu_t \) is indifferent between two signals which induce the same
\(^{11}\)There are of course different and more complete orderings of signals, including Blackwell’s (1951) and, e.g., those of Ganuza and Penalva (2010). The purpose here, however, is to define a simple ordering of signals consistent with conditions of increasing differences to be used in the equilibrium analysis. The definition of more precise turns out to be appropriate for this purpose.
receiver interim belief and $\pi'$ is strictly more precise than $\pi$, then any type $\mu_{t'} > \mu_t$ strictly prefers $\pi'$. The logic is that if $\pi'$ is more precise than $\pi$, then the difference between the expected payoffs conditional on $\omega_H$ and $\omega_L$ is larger under $\pi'$ than under $\pi$. Since the state is more likely to be $\omega_H$ for type $\mu_{t'}$ than for type $\mu_t$ it follows that type $\mu_{t'}$ is relatively more attracted by $\pi'$. The first part of Lemma 1 allows the receiver to respond to two different interim beliefs $\mu$ and $\mu'$. 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 qualifier in the first part of Lemma 1 implies that the result is not valid for arbitrary $\mu$ and $\mu'$ when keeping the remaining variables fixed. In the absence of this restriction, several of the characterizations below would follow from Lemma 1.\footnote{Without the qualifier the first part of Lemma 1 would be an increasing differences analogue to the single-crossing assumption under which Cho and Sobel (1990) characterize sequential equilibria selected by the D1 criterion (their condition A4).} While the qualifier cannot be dispensed with, the next result derives an alternative property which proves equally useful. The result establishes that for any signal giving type $\mu_{t'}$ a lower expected payoff than a fully disclosing signal under receiver interim $\mu < \mu_{t'}$, there is an alternative signal which type $\mu_{t'}$ but no type $\mu_t < \mu_{t'}$ would prefer provided the receiver’s interim were instead $\mu_{t'}$.

**Lemma 2** Consider any type $\mu_{t'} > \mu_1$, receiver interim belief $0 < \mu < \mu_{t'}$ and signal $\pi \in \Pi \setminus \Pi^{FD}$ and suppose that $\hat{V}(\pi, \mu, \mu_{t'}) \geq \hat{V}_{t'}^{FD}$. There is then some $\pi' \in \Pi$ which is more precise than $\pi$ and such that for all $\mu_t \in \{\mu_1, \ldots, \mu_{t'-1}\}$

$$\hat{V}(\pi', \mu_t, \mu_{t'}) - \hat{V}(\pi, \mu, \mu_{t'}) > 0 > \hat{V}(\pi', \mu_t, \mu_t) - \hat{V}(\pi, \mu, \mu_t).$$

Lemma 2 is invoked to establish several of the results below. The logic of the proof is roughly the following. If types $\mu_{t'}$ and $\mu_t < \mu_{t'}$ use a signal $\pi$, then an increase in the receiver’s interim belief from $\mu$ to $\mu_{t'}$ increases both types’ expected payoffs. One can then consider a path of signals of increasing precision from $\pi$ to some $\pi^{FD}$, along which increasing differences ensure that type $\mu_{t'}$ is favored more than type $\mu_t$. Since $\hat{V}(\pi, \mu, \mu_{t'}) \geq \hat{V}_{t'}^{FD}$ one eventually finds some $\pi'$ such that

$$\hat{V}(\pi', \mu_t, \mu_{t'}) - \hat{V}(\pi, \mu, \mu_{t'}) > 0.$$
type $\mu_{t'}$ marginally prefers $\pi'$ over $\pi$ and type $\mu_t$ strictly prefers $\pi$ over $\pi'^{13}$.

### 3.2 Equilibria

The first main result follows almost immediately by combining Lemma 2 with the observation that the equilibrium expected payoff of any type $\mu_t$ is bounded below by $\hat{V}_{FD}^t$. The result states that whenever two or more types pool on the same signal in equilibrium, this signal must be fully disclosing. To illustrate the logic, suppose $\mu_1, ..., \mu_t$ pool on $\pi \in \Pi \setminus \Pi_{FD}$ in equilibrium. Then $\tilde{\beta}(\pi) < \mu_t$ and $\hat{V}(\pi, \tilde{\beta}(\pi), \mu_t) \geq \hat{V}_{FD}^t$, and Lemma 2 can be combined with D1 to argue the existence of a deviation $\pi'$ such that $\tilde{\beta}(\pi') \geq \mu_t$ and which is profitable for type $\mu_t$.

**Proposition 1** If $(\pi^1, ..., \pi^n)$ is an equilibrium sender strategy, then for each $t \in \{1, ..., n\}$ it holds that $\pi^t$ is either separating or fully disclosing.

Hence, either the sender’s choice of signal reveals his type to the receiver, or the outcome of the signal reveals the payoff-relevant state. I.e., whenever tests or experiments designed by the sender are not perfectly informative, one can infer the sender’s prior by observing the design of the tests or experiments. If the type-space is binary the statement strengthens to "any equilibrium is either separating or fully disclosing." In the binary case this is true also if the solution concept is weakened to PBE selected by the intuitive criterion. The result is driven by higher types’ stronger preferences for precise signals and the possibility of fully disclosing the payoff-relevant state. These properties combined with D1 imply that at any equilibrium at which there is some pooling at signals which are not fully disclosing, the receiver’s beliefs must be non-skeptical for some sufficiently precise out-of-equilibrium signals. The highest type in any set of pooling types can then upset any such equilibrium by deviating to some sufficiently precise signal inducing non-skeptical beliefs.

---

$^{13}$The major complication in the formal argument is that it is generally not possible to say whether $\mu_{t'}$ or $\mu_t$ benefits the most from the initial increase in interim belief. The proof resolves the difficulty by constructing an intermediate signal, which satisfies the qualifier in part (i) of Lemma 1 and which type $\mu_{t'}$ still prefers over $\pi$ after the interim increase, and then defining a path of signals of increasing precision.
Proposition 1 implies that identifying the set of equilibrium sender strategies is equivalent to identifying the set of incentive compatible strategies consisting of fully disclosing and separating signals. This can be accomplished in terms of solutions to certain maximization problems. Let

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

for \(t \in \{1, ..., n\}\), where it is not difficult to see that \(\Pi_t^* \neq \emptyset\). I.e., \(\Pi_t^*\) is what Kamenica and Gentzkow (2011) refer to as type \(\mu_t\)’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 \(\mu_1\)’s optimal signal and signals maximizing the expected payoff of each type \(\mu_t > \mu_1\) under a local incentive compatibility constraints ensuring that type \(\mu_{t-1}\) does not have incentives to mimic type \(\mu_t\)’s signal.

**Theorem 1** (i) There is an equilibrium at which the sender’s strategy is \((\pi^1, ..., \pi^n)\) if and only if

$$\pi^1 \in \arg \max_{\pi \in \Pi} \hat{V}(\pi, \mu_1, \mu_1)$$  \hspace{1cm} (1)

and

$$\pi^t \in \arg \max_{\pi \in \Pi} \hat{V}(\pi, \mu_t, \mu_t) \text{ s.t. } \hat{V}(\pi, \mu_t, \mu_{t-1}) \leq \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_{t-1})$$  \hspace{1cm} (2)

for all \(t \in \{2, ..., n\}\). (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 payoff of each sender type is constant across equilibria.

Theorem 1 reveals that in spite of the absence of single-crossing assumptions on the primitives, the structure of the predictions is remarkably similar to that of standard signaling models, such as simple versions of Spence’s (1973) model of job-market signaling, or the more general class of
signaling games analyzed by Cho and Sobel (1990). In particular, to identify a PBE it is sufficient to maximize each type’s expected payoff as if the receiver knew the type and subject to a local upward incentive compatibility constraint (1 and 2 are sufficient). When adding D1 only such PBE survive (1 and 2 are necessary).\footnote{The corresponding maximization problems in Cho and Sobel’s (1990) characterization of equilibria are more complex than those in Theorem 1, however, since they must account for the possibility that types might prefer pooling on the "highest" signal, rather than separating. Such a qualification is redundant here, since the "highest" signal is fully disclosing, and $\tilde{V}_t^{FD}$ is independent of the receiver’s interim belief.} In a simple model of unproductive job-market signaling this is analogous to a "Riley equilibrium" at which the lowest type obtains zero education and the remaining types obtain just enough education for the type below not to have incentives to locally mimic upwards.

The result is driven by the linearity of $\tilde{V}$ in type, increasing differences in precision and the lower bound on equilibrium expected payoffs implied by the availability of fully disclosing signals. Sufficiency follows by first noting that the maximization in (2) implies that no type $\mu_t$ has a profitable local downward deviation. For, one could then invoke Lemma 2 to claim the existence of a sufficiently precise signal which outperforms the deviation under receiver interim $\mu_t$ and still satisfies the constraint in (2). The linearity of $\tilde{V}$ in type next implies that local incentive compatibility is sufficient for global incentive compatibility. Similar arguments rule out profitable deviations to out-of-equilibrium signals. The necessity of (2) follows by appealing to D1. If (2) does not hold in some equilibrium then there must be a type $\mu_t$ and a deviation $\pi$ which satisfies the constraint in (2) and is profitable for $\mu_t$ provided that $\tilde{\beta}(\pi) \geq \mu_t$. The linearity of $\tilde{V}$ in type implies that all types $\mu_{t'} < \mu_t$ prefer the equilibrium over $\pi$ under interim $\mu_t$. D1 then implies $\tilde{\beta}(\pi) \geq \mu_t$ and a profitable deviation for type $\mu_t$.

**The optimality of equilibrium signals.** Theorem 1 establishes that type $\mu_1$’s equilibrium signal is optimal, i.e., type $\mu_1$ would not behave differently if the receiver could observe his type. The remaining types, however, maximize their expected payoff under a constraint and Theorem 1 does not clarify whether the constraint will be active in the sense of forcing these types to choose suboptimal signals. On the other hand, Proposition 1 states that the sender’s equilibrium strategy consists of separating and fully disclosing signals, but does not clarify under which conditions each
of these possibilities occur. The next result reveals that both questions are related and have a common and exhaustive answer. The equilibrium signals of all types except the lowest type can be optimal only if they are fully disclosing. Conversely, the equilibrium signal of any sender type is fully disclosing only if such signals are optimal.

**Proposition 2** Suppose \((\pi^1, \ldots, \pi^n)\) is an equilibrium sender strategy. For any \(t \in \{2, \ldots, n\}\) the following statements are equivalent: (i) \(\pi^t \in \Pi^*_t\), (ii) \(\pi^t \in \Pi^{FD}\) and (iii) \(\Pi^*_t \cap \Pi^{FD} \neq \emptyset\).

The claim that whenever a type \(\mu_t > \mu_1\) uses an optimal signal it must be fully disclosing (i\(\Rightarrow\)ii) follows by noting that type \(\mu_1\) strictly prefers posing as any type \(\mu_t > \mu_1\) under \(\pi^t \in \Pi^*_t \setminus \Pi^{FD}\) over a signal \(\pi^1 \in \Pi^*_1\) 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 \(\mu_t\) must be lower than the highest updated belief induced by type \(\mu_1\). It can then be argued that if this is true for optimal signals of type \(\mu_t\) and type \(\mu_1\), then the optimal distributions of updated receiver beliefs of both types are similar in a precise sense.\(^{15}\) Intuitively, this similarity implies that if both types use optimal signals in equilibrium and type \(\mu_1\) deviates to type \(\mu_t\)’s signal, then type \(\mu_1\) obtains his equilibrium payoff plus a bonus resulting from posing as type \(\mu_t\). The deviation is then profitable and type \(\mu_t\) cannot use an optimal signal in equilibrium, except in the special case in which the optimal signal is fully disclosing, in which case posing as some other type provides no bonus.

The claim that a fully disclosing signal can only be used if it is optimal (ii\(\Rightarrow\)i) follows by first noting that fully disclosing signals are either optimal for all types, or for no type. If some type \(\mu_t > \mu_1\) uses a suboptimal fully disclosing signal, this type can obtain a higher expected payoff under interim \(\mu_t\) by perturbing the fully disclosing signal in the direction of an optimal signal. Since type \(\mu_1\) uses an optimal signal he strictly prefers the equilibrium over the perturbation.

\(^{15}\)The idea can be 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 \(\mu_1\) and \(\mu_t\) 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.
It then roughly follows that since the perturbation is almost fully disclosing and such signals are suboptimal, other types prefer the equilibrium over the perturbation, which contradicts \( \mu_t \)'s maximization in Theorem 1.\(^{16}\) Intuitively, the use of suboptimal fully disclosing signals implies that the sender has gone too far in the pursuit of incentive compatibility and there are then profitable deviations to out-of-equilibrium signals. Section 5.4 in the appendix discusses an example which illustrates why fully disclosing signals cannot be used in equilibrium unless they are optimal.

Finally, (iii) implies (i) since fully disclosing signals provide lower bounds on equilibrium expected payoffs. Each sender type must therefore choose an optimal signal whenever fully disclosing signals are optimal.

The case in which fully disclosing signals are optimal is somewhat special and Kamenica and Gentzkow’s (2011) characterization of optimal signals provides a straightforward check of when this case occurs. It suffices to draw a line from \((0, \hat{v}(0))\) to \((1, \hat{v}(1))\). A fully disclosing signal is optimal for all types if and only if this line is (weakly) above \( \hat{v} \). Proposition 2 thus roughly pins down the equilibrium in terms of two possible cases. Either fully disclosing signals are optimal and any equilibrium is fully disclosing (except possibly for the signal of type \( \mu_1 \)). Otherwise, fully disclosing signals are not optimal and are then not used in equilibrium, i.e., the sender’s strategy consists of separating signals. In this case all sender types except \( \mu_1 \) incur a cost in comparison to the full information benchmark in which the receiver knows the sender’s type. I summarize the conclusion in the following corollary.

**Corollary 2** (i) If \( \Pi^{FD} \cap (\cup_{t=1}^n \Pi_t^*) = \emptyset \), then any equilibrium is separating and no equilibrium signal is fully disclosing. (ii) If \( \Pi^{FD} \cap (\cup_{t=1}^n \Pi_t^*) \neq \emptyset \), then there is a fully disclosing equilibrium and any equilibrium signal of any type except possibly \( \mu_1 \) is fully disclosing.

**Further properties of the equilibrium signals.** The cost incurred by the sender in separating equilibria is not of the money-burning type, but rather arises from a loss of control over the receiver’s informational environment. This loss of control sometimes benefits the receiver by

---

\(^{16}\)The precise argument in the proof derives such a contradiction for at least some type \( \mu_T \leq \mu_t \).
providing her with a more informative information structure. An implication is that if the sender
could credibly disclose his private information to the receiver prior to choosing a signal, he might
strictly prefer doing this, while the receiver may prefer not to receive this information.\footnote{17} For ex-
ample, 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.\footnote{18} It is not
clear, however, that the FDA should 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 forces the pharmaceutical company to
reveal more information than what it would if the FDA knew its type. Such a situation arises,
e.g., if all optimal signals are silent. In this case Proposition 2 implies that the equilibrium signal
of any type $\mu_t > \mu_1$ is not silent and therefore strictly Blackwell (1951) more informative than
the 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 all
types' unique optimal signals are silent (see Kamenica and Gentzkow 2011). The following remark
summarizes the observation.

**Remark 1** Fix the signal of type $\mu_1$ at some $\pi \in \Pi_1^*$. If all signals in $\bigcup_{t=2}^{n} \Pi_t^*$ 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 any type $\mu_t > \mu_1$ cannot be too imprecise.
In particular, there cannot be a signal which is more precise than the equilibrium signal of type
$\mu_t$ and which would give this type a higher expected payoff. Therefore, none of type $\mu_t$'s optimal
signals can be more precise than the equilibrium signal of type $\mu_t$.

\footnote{17}{Adding a cheap-talk stage prior to the choice of signal, however, would not affect the set of equilibrium outcomes.}
\footnote{18}{A similar result is obtained by Kolotilin (2014a), who finds that verifiable private sender information unravels in a related model.}
Proposition 3 Suppose \((\pi^1, ..., \pi^n)\) is an equilibrium sender strategy. For any \(t \in \{2, ..., n\}\) it holds that if \(\pi \in \Pi\) is more precise than \(\pi^t\), then \(\tilde{V}(\pi, \mu_t, \mu_t) \leq \tilde{V}(\pi^t, \mu_t, \mu_t)\). Consequently, if \(\pi \in \Pi^*_t\), then \(\pi\) is not more precise than \(\pi^t\).

Proposition 3 follows from an increasing differences argument similar to those above. The result implies that the equilibrium signal of any type \(\mu_t > \mu_1\) can never be silent. For, if the equilibrium signal of type \(\mu_t\) were silent, then Proposition 2 implies that there is an optimal signal which is more precise than type \(\mu_t\)'s equilibrium signal, contradicting Proposition 3. Incentive compatibility hence requires the sender to reveal his private information as well as some additional information, even when all types' optimal signals are silent.

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

Remark 2 Suppose \((\pi^1, ..., \pi^n)\) is an equilibrium sender strategy such that \(\pi^t\) has support \(\{s^t_1, ..., s^t_k\}\) for each \(t \in \{1, ..., n\}\). (i) If \(t < t'\), then \(B(\pi^t, s^t_i, \mu_t) < B(\pi^{t'}, s^{t'}_{k_t}, \mu_t)\). (ii) If \(\pi^1\) is silent and \(t \in \{2, ..., n\}\), then \(B(\pi^t, s^t_1, \mu_t) < \mu_1 < B(\pi^t, s^t_{k_t}, \mu_t)\).

The first part states that the supports of the distribution of induced receiver beliefs of different types must overlap in equilibrium. This follows trivially by observing that if type \(\mu_t\) deviates to the equilibrium signal of type \(\mu_{t'} > \mu_t\) he induces the same set of updated beliefs as type \(\mu_{t'}\) does. An interpretation is that in equilibrium incentive compatibility requires the "worst news" of "better types" to be worse than the "best news" of "worse types." In other words, the equilibrium signals are sufficiently informative to override the private information to some extent. If, e.g., there are only two types \(\mu_1 \sim 0\) and \(\mu_2 \sim 1\) and the optimal signal of type \(\mu_1\) is silent, the equilibrium becomes quite polarized, consisting of a silent signal for type \(\mu_1\) and an almost fully disclosing signal for type \(\mu_2\).
The second part of the remark is a corollary of the first part, which emphasizes that if the equilibrium signal of type $\mu_1$ is silent, then type $\mu_i > \mu_1$ must induce updated beliefs both below and above the updated belief induced by type $\mu_1$’s silent signal (which equals $\mu_1$). Since only type $\mu_1$ can use a silent signal in equilibrium an interpretation is that a silent 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 type, it is neither the worst nor the best equilibrium news.

3.3 Perfectly informed sender

The analysis above precludes the possibility that the sender is perfectly informed, i.e., that $\mu_1 = 0$ or $\mu_n = 1$. It is relatively straightforward to extend the analysis to account for this possibility. Here a comment is in place regarding the receiver’s reaction to events which are impossible given her interim beliefs. Suppose, e.g., that type $\mu_n = 1$ separates by using a fully disclosing signal $\pi^n$. If some other type deviates to $\pi^n$, then signal outcomes such that $\pi^n_{iL} > 0$ occur with positive probability, but these outcomes are impossible given $\bar{b}(\pi^n) = 1$. As Perez-Richet (2014), I assume that the receiver abandons the interim beliefs in such cases, so in the example her final beliefs would switch from 1 to 0 (given that $\pi^n$ was fully disclosing). I.e., the information generated by the signal is taken as hard evidence and has some pre-eminence off the equilibrium path. The issue arises only for interim belief $\mu \in \{0, 1\}$, and the analysis can therefore be handled by appropriately extending the function $\tilde{V}$ to interim beliefs $\mu \in \{0, 1\}$.

Allowing $\mu_n = 1$ is straightforward and preserves most of the results derived above. In this case, type $\mu_n$ cannot pool on a signal $\pi$ with some other type unless $\pi$ is fully disclosing. For, if this were the case there would be some outcome $s_i$ such that $\pi_{iL} > 0$ and $\pi_{iH} > 0$, and therefore $B(\pi, s_i, \tilde{b}(\pi)) < 1$. Type $\mu_n$ could then profitably deviate to a fully disclosing signal and earn $\tilde{v}(1)$. The equilibrium signal of type $\mu_n$ is therefore separating or fully disclosing and the equilibrium
expected payoff of type $\mu_n$ equals $\tilde{v}(1)$. Any equilibrium then simply consists of signals satisfying Theorem 1 for types $\mu_1, \ldots, \mu_{n-1}$ and any signal $\pi^n$ such that these types do not have incentives to mimic type $\mu_n$, which holds if $\pi^n \in \Pi_{FD}$. I.e., the equilibrium signals $\pi^1, \ldots, \pi^{n-1}$ preserve the properties derived in Section 3.2 and $\pi^n$ must simply be chosen to preserve upward incentive compatibility.

Allowing $\mu_1 = 0$ has a more profound and somewhat surprising effect on the predictions of the model. Whenever $\pi^1$ is separating or fully disclosing type $\mu_1$ earns expected payoff $\tilde{v}(0)$ and it is then notoriously difficult to discourage type $\mu_1$ from mimicking the signals of other types. In fact, this is only possible if all types except $\mu_1$ choose fully disclosing signals. At the same time, Lemma 1 and 2 and Proposition 1 make no use of the assumption $\mu_1 > 0$, which implies that any equilibrium consists of separating or fully disclosing signals. The implication is that whenever $\mu_1 = 0$, then any type except $\mu_1$ must use a fully disclosing signal in equilibrium. We have the following result.

**Proposition 4** Suppose $\mu_1 = 0$. Then $(\pi^1, \ldots, \pi^n) \in \Pi^n$ is an equilibrium sender strategy if and only if $\pi^t \in \Pi_{FD}$ for all $t \in \{2, \ldots, n\}$.

The "if" statement follows by noting that type $\mu_1$ benefits from any deviation $\pi$ which is not fully disclosing, provided that $\tilde{\beta}(\pi) > \mu_1$. The D1 criterion thus allows skeptical beliefs $\tilde{\beta}(\pi) = 0$, and no type has incentives to deviate given such beliefs.

Proposition 4 implies quite a remarkable unraveling of information. The sender always perfectly informs the receiver of the payoff-relevant state. This is true even if the sender becomes informed that the payoff-relevant state is $\omega_L$ with arbitrarily small probability, and even if the sender would choose uninformative signals in the full information benchmark. Intuitively, the unraveling of information occurs due to the presence of a sender type which always has everything to win and is willing to mimic any signal which may identify him as a higher type. From the sender’s point of view this can be a rather inefficient state of affairs, while the receiver has perfect information and could not be better off.

Proposition 4 provides some insight regarding the contrast between the results in Section 3.2
and the results of Perez-Richet (2014). In his model the type-space is binary, the sender is always perfectly informed, i.e., $\mu_1 = 0$ and $\mu_2 = 1$, and the sender’s payoff is constant in the receiver’s belief, except for a single discontinuity. In such a setting it is without loss of generality to confine attention to pooling PBE. Here, it is without loss of generality to focus on pooling equilibria only if $\mu_1 = 0$. E.g., if $\mu_1 > 0$ and $\mu_n = 1$, then type $\mu_1$ would choose an optimal signal and type $\mu_n$ would fully disclose in any PBE selected by the D1 criterion. In other words, whether all equilibrium outcomes can be obtained as pooling equilibria seems mostly driven by assumptions regarding whether and when the sender is perfectly informed.\footnote{The extent to which it is without loss of generality to focus on pooling equilibria in a model with sender preferences as in Perez-Richet (2014) and a type space like the one here remains, however, an open question.}

4 Concluding remarks

The introduction of unverifiable private information in a framework of Bayesian persuasion generates a tractable signaling game which produces a number of concrete predictions. In perfect Bayesian equilibria selected by a standard refinement private information unravels and such equilibria are fully characterized by maximization problems reminiscent of those found in canonical signaling models. The results are driven to a large extent by properties of increasing differences which arise endogenously in equilibrium. The analysis predicts that either it is possible to infer private information from the sender’s design of experiments, tests or free trials, or the sender reveals as much information as feasible. Whenever the former case occurs it does so at a cost for the sender. In particular, the use of fully disclosing and optimal signals in equilibrium are two sides of the same coin. The sender’s equilibrium signal is either optimal and fully disclosing, or neither. Whenever fully disclosing signals are suboptimal the sender would strictly prefer revealing his private information to the receiver prior to choosing a signal. The receiver, however, sometimes strictly prefers the suboptimal equilibrium signals, since these may generate better information. I.e., the receiver may collect an "ignorance rent."

There are many ways in which one could generalize the framework considered here. For ex-
ample, 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 information. Another restrictive feature of the framework here is the binary set of payoff-relevant states. A natural question 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 of 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 is unconstrained and the monopolist is imperfectly informed, then such signaling may play an important role.

5 Appendix

5.1 Proofs of Lemmata 1 and 2

Proof. (Lemma 1) Consider any types $\mu_t < \mu_{t'}$, receiver interims $0 < \mu < 1$ and $0 < \mu' < 1$ and $\pi, \pi' \in \Pi$ with support $S$ such that $\pi'$ is more precise than $\pi$, i.e., such that for all $i \in \{1, ..., k\}$ either $\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}$ or $\pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}$.

To prove (i) suppose that for all $s_i \in S$ either $B(\pi', s_i, \mu') \leq B(\pi, s_i, \mu) \leq \mu$ or $B(\pi', s_i, \mu') \geq B(\pi, s_i, \mu) \geq \mu$. Let $\bar{k} := \min\{i \in \{1, ..., k\} : \pi'_{iL} \leq \pi'_{iH}\}$. Abbreviate $\hat{\nu}_i = \hat{\nu}(B(\pi, s_i, \mu))$ and $\hat{\nu}'_i = \hat{\nu}(B(\pi', s_i, \mu'))$. Then $\pi'_{iL} \geq \pi_{iL} \geq \pi_{iH} \geq \pi'_{iH}$ and $\hat{\nu}'_i \leq \hat{\nu}_i$ for any $i < \bar{k}$ and $\pi'_{iL} \leq \pi_{iL} \leq \pi_{iH} \leq \pi'_{iH}$.
and $\nu'_i \geq \nu_i$ for any $i \geq k$. Since $\sum_{i<k}(\pi_{iH} - \pi_{iL}) = - \sum_{i\geq k}(\pi_{iH} - \pi_{iL})$ and likewise for $\pi'$

$$\hat{V}(\pi', \mu', \mu_t) - \hat{V}(\pi, \mu, \mu_t) = [\hat{V}(\pi', \mu', \mu_t) - \hat{V}(\pi, \mu, \mu_t)]$$

$$= (\mu - \mu) \sum_{i<k}(\pi_{iH}' - \pi_{iL}')\nu_i' - (\pi_{iH} - \pi_{iL})\nu_i + \sum_{i\geq k}(\pi_{iH}' - \pi_{iL}')\nu_i' - (\pi_{iH} - \pi_{iL})\nu_i$$

$$\geq (\mu - \mu) \sum_{i<k}[(\pi_{iH}' - \pi_{iL}') - (\pi_{iH} - \pi_{iL})]\nu_i' + \sum_{i\geq k}[(\pi_{iH}' - \pi_{iL}') - (\pi_{iH} - \pi_{iL})]\nu_i'$$

$$= (\mu - \mu) \sum_{i<k}[(\pi_{iH}' - \pi_{iL}') - (\pi_{iH} - \pi_{iL})](\nu_i' - \nu_k') \geq 0,$$

which proves (i).

Since $\pi'$ is more precise than $\pi$ we have that for all $s_i \in S$ either $B(\pi', s_i, \mu) \leq B(\pi, s_i, \mu) \leq \mu$ or $B(\pi', s_i, \mu) \geq B(\pi, s_i, \mu) \geq \mu$, which by (i) proves the first part of (ii).

Suppose $\pi'$ is strictly more precise than $\pi$. Then $k > 1$, $\sum_{i<k}[(\pi_{iH}' - \pi_{iL}') - (\pi_{iH} - \pi_{iL})] < 0$ and $\nu_i' < \nu(\mu') \leq \nu_k'$ for all $i < k$ and the last inequality above is strict. We have proved (ii). \blacksquare

Proof. (Lemma 2) Suppose $\mu > \mu_1$, $\mu \in (0, \mu_v)$ and $\pi \in \Pi \setminus \Pi^{FD}$ with $\hat{V}(\pi, \mu, \mu_v) \geq \hat{V}^{FD}_v$. I first prove the following preliminary result.

Claim. There is some $\pi' \in \Pi \setminus \Pi^{FD}$ with support $S' = S$ which is not silent, more precise than $\pi$ and such that $\hat{V}(\pi', \mu', \mu_v) - \hat{V}(\pi, \mu, \mu_v) \geq \hat{V}(\pi', \mu_v, \mu_v) - \hat{V}(\pi', \mu_v, \mu_v)$ for any $\mu \in \{\mu_1, ..., \mu_{v-1}\}$

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

Proof. The proof constructs a signal $\pi'$ satisfying the hypothesis of part (i) of Lemma 1 and such that $\hat{V}(\pi', \mu_v, \mu_v) - \hat{V}(\pi, \mu, \mu_v) > 0$. Let $k = \min\{i \in \{1, ..., k\} : \pi_{iL} \leq \pi_{iH}\}$, let $K = \{1\}$ if $k = 1$ and $K = \{1, ..., k - 1\}$ otherwise, and let $K = \{1, ..., k\}$. I.e., $\pi_{iL} \geq \pi_{iH}$ and $B(\pi, s_i, \mu) \leq \mu$ for $i \in K$ and $\pi_{iL} \leq \pi_{iH}$ and $B(\pi, s_i, \mu) \geq \mu$ for $i \in K \setminus K$. Let $\pi_{iL}' = \pi_{iL}$ for all $i \in K$ and let

$$\pi_{iH}' = \mu(1 - \mu_v) \pi_{iH} - \pi_{iL}' = \mu(1 - \mu_v) \pi_{iH} - (1 - \mu) \mu_v$$

27
for all \( i \in \overline{K} \). Straightforward algebra shows that for all \( i \in \overline{K} \) we have \( B(\pi', s_i, \mu') = B(\pi, s_i, \mu) \leq \mu \) and \( \pi'_i = \pi_i \geq \pi_i \geq \pi'_i, \) where \( \pi_i = \pi'_i \) if and only if \( \pi_i = 0 \). Notice that \( \pi' \) is not silent, for then \( \pi'_i = \pi_i = 0 \), a contradiction.

For all \( i \in K \setminus \overline{K} \) let \( \pi'_i = \pi_i + \lambda(1 - \pi_i) \) with \( \lambda = \sum_{i \in \overline{K}} (\pi_i - \pi'_i) / \sum_{i \in \overline{K}} (1 - \pi_i) \) if \( \sum_{i \in \overline{K}} (1 - \pi_i) > 0 \) and \( \lambda = 0 \) otherwise. For all \( i \in K \setminus \overline{K} \) we then have \( \pi'_i = \pi_i \leq \pi_i \leq \pi'_i \) and \( B(\pi', s_i, \mu') \geq B(\pi, s_i, \mu') \geq B(\pi, s_i, \mu) \geq \mu \). Since \( \pi' \) is more precise than \( \pi \), (i) of Lemma 1 implies \( \widehat{V}(\pi', \mu', \mu') - \widehat{V}(\pi, \mu, \mu) \geq \widehat{V}(\pi', \mu', \mu) - \widehat{V}(\pi, \mu, \mu) \) for any \( \mu \in \{\mu_1, \ldots, \mu_{t-1}\} \).

Let \( \widehat{v}'_i := \widehat{v}(B(\pi', s_i, \mu')) \) and \( \widehat{v}_i := \widehat{v}(B(\pi, s_i, \mu)) \). Since \( \widehat{v}'_i \geq \widehat{v}_i \) for all \( i \in K \)

\[
\widehat{V}(\pi', \mu', \mu') - \widehat{V}(\pi, \mu, \mu') = \sum_{i=1}^{k} [\mu'_i (\widehat{v}'_i - \widehat{v}_i) + (1 - \mu'_i) \pi_i (\widehat{v}'_i - \widehat{v}_i)] \\
\geq \mu'_i \sum_{i=1}^{k} (\pi'_i - \pi_i) \geq 0,
\]

where the second inequality follows since by construction \( \pi'(\omega_H) \) first order stochastically dominates \( \pi(\omega_H) \). If the first inequality is an equality, then \( B(\pi, s_i, \mu') = B(\pi, s_i, \mu) = 1 \) for all \( i \in K \setminus \overline{K} \). Since \( \pi \notin \Pi^{FD} \) it must then be that \( B(\pi, s_i, \mu) \in (0,1) \) and therefore \( \pi_i > 0 \) for some \( i \in \overline{K} \). By the definition of \( \pi' \) we have \( \pi'_i > \pi_i \) and therefore \( \pi'(\omega_H) \neq \pi(\omega_H) \) and the last inequality is strict. Hence, \( \widehat{V}(\pi', \mu', \mu') - \widehat{V}(\pi, \mu, \mu') > 0 \), and since \( \widehat{V}(\pi', \mu', \mu') > \widehat{V}(\pi, \mu, \mu') \) we have \( \pi' \notin \Pi^{FD} \).

The proof concludes by using a signal with the properties in the claim to construct a signal satisfying the statement in Lemma 2. Therefore, let \( \pi' \in \Pi \) have the properties in the claim above and let \( \overline{k} = \min\{i \in \{1, \ldots, k\} : \pi'_i \leq \pi_i \} \) where, since \( \pi' \) is not silent, \( \overline{k} > 1 \).

Let \( \xi \in \Pi^{FD} \) have support \( S \) and be such that \( \xi_i = 0 \) for \( i < \overline{k} \) and \( \xi_i = \pi'_i \) for all \( i \geq \overline{k} \), and \( \xi_i = 0 \) for all \( i \geq \overline{k} \) and \( \xi_i \geq \pi'_i \) for all \( i < \overline{k} \). Then \( \xi_i \geq \pi_i \geq \pi'_i \geq \xi_i \) for \( i < \overline{k} \) and \( \xi_i \leq \pi'_i \leq \pi'_i \leq \xi_i \) for \( i \geq \overline{k} \) and since \( \pi' \notin \Pi^{FD} \) we have \( \xi \neq \pi' \).

Define a family of signals \( g(\gamma) \) with support \( S \) by \( g_i(\gamma) = (1 - \gamma) \pi'_i + \gamma \xi_i \) and \( g_i(\gamma) = (1 - \gamma) \pi'_i + \gamma \xi_i \) for \( \gamma \in [0,1] \). Then \( g(0) = \pi' \) and \( g(1) = \xi \in \Pi^{FD} \), and if \( \gamma > \gamma \), then \( g(\gamma') \) is strictly more precise than \( g(\gamma) \).
Let \( \overline{\gamma} := \min\{\gamma \in [0,1] : \overline{V}(g(\gamma), \mu_0, \mu_t) = \overline{V}(\pi, \mu, \mu_t)\} \), where \( \overline{\gamma} > 0 \) is well defined since \( \overline{V}(g(\gamma), \mu_0, \mu_t) \) is continuous in \( \gamma \) and \( \overline{V}(g(0), \mu_0, \mu_t) > \overline{V}(\pi, \mu, \mu_t) \geq \overline{V}(g(1), \mu_0, \mu_t) \). By Lemma 1 and by hypothesis, for all \( \mu_t \in \{\mu_1, ..., \mu_{t-1}\} \)

\[
\overline{V}(g(\gamma), \mu_0, \mu_t) - \overline{V}(g(\gamma), \mu_0, \mu_t) > \overline{V}(\pi', \mu_0, \mu_t) - \overline{V}(\pi', \mu_0, \mu_t) \geq \overline{V}(\pi, \mu, \mu_t) - \overline{V}(\pi, \mu, \mu_t),
\]

and therefore \( \overline{V}(g(\gamma), \mu_0, \mu_t) - \overline{V}(\pi, \mu, \mu_t) = 0 > \overline{V}(g(\gamma), \mu_0, \mu_t) - \overline{V}(\pi, \mu, \mu_t) \).

Since \( \overline{V}(g(\gamma), \mu_0, \mu_t) > \overline{V}(\pi, \mu, \mu_t) \) for all \( \gamma \in [0, \overline{\gamma}) \) there is some \( \gamma > \overline{\gamma} \) such that \( \overline{V}(g(\gamma), \mu_0, \mu_t) - \overline{V}(\pi, \mu, \mu_t) > 0 > \overline{V}(g(\gamma), \mu_0, \mu_t) - \overline{V}(\pi, \mu, \mu_t) \) for all \( \mu_t \in \{\mu_1, ..., \mu_{t-1}\} \), which concludes the proof. \[ \blacksquare \]

### 5.2 Proofs of Propositions 1-3 and Theorem 1

**Proof.** (Proposition 1) Suppose that \((\pi^1, ..., \pi^n)\), \(\alpha\) and \((\beta, \beta')\) is an equilibrium and, to contradiction, that \(\pi^t = \pi \in \Pi \setminus \Pi^{FD} \) for all \(\mu_t \in T \subseteq T\) with \(|T| \geq 2\) and \(\pi^t \neq \pi\) otherwise. Then \(\mu_T := \max T > \beta(\pi)\) and \(\overline{V}(\pi, \beta(\pi), \mu_T) \geq \overline{V}_T^{FD}\). By Lemma 2 there is some \(\pi' \in \Pi\) such that \(\overline{V}(\pi', \mu_T, \mu_T) - \overline{V}(\pi, \beta(\pi), \mu_T) > 0 > \overline{V}(\pi', \mu_T, \mu_T) - \overline{V}(\pi, \beta(\pi), \mu_T)\) for all \(\mu_t \in \{\mu_1, ..., \mu_{T-1}\}\) and by incentive compatibility \(\overline{V}(\pi', \beta(\pi), \mu_T) \geq \overline{V}(\pi, \beta(\pi), \mu_T) > \overline{V}(\pi', \mu_T, \mu_T)\). Since \(\overline{V}\) is increasing in its second argument, \(D^0(\pi', \mu_T) \not\subset D(\pi', \mu_T)\) for any \(\mu_t \in \{\mu_1, ..., \mu_{T-1}\}\). \(D^1\) requires \(\beta(\pi') \geq \mu_T\) and type \(\mu_T\) has a profitable deviation, contradicting the equilibrium. \[ \blacksquare \]

**Proof.** (Theorem 1) For any \(t \in \{2, ..., n\}\) and \(\pi \in \Pi\) let \(Z_t(\pi) := \{\pi' \in \Pi : \overline{V}(\pi', \mu_t, \mu_{t-1}) \leq \overline{V}(\pi, \mu_{t-1}, \mu_{t-1})\}\).

**Step 1. The "if" part of (i).**

Proof. Suppose \(\pi^1 \in \Pi_1^*\) and \(\pi^t \in \arg\max_{\pi \in Z_t(\pi^{t-1})} \overline{V}(\pi, \mu_t, \mu_t)\) for \(t \in \{2, ..., n\}\). Then \(\Pi^{FD} \subseteq Z_t(\pi^{t-1})\) and therefore \(\overline{V}(\pi^t, \mu_t, \mu_t) \geq \overline{V}_t^{FD}\) for all \(t \in \{2, ..., n\}\). Suppose \(\alpha\) is optimal given rational beliefs \((\beta, \beta')\). For each \(\pi \in \Pi \setminus \{\pi^1, ..., \pi^n\}\) let \(\hat{T}(\pi) := \{\mu_t \in T : D^0(\pi, \mu_t) \not\subset D(\pi, \mu_t)\}\) for some \(\mu_t \in T\) and let \(\beta(\pi) = \min T \setminus \hat{T}(\pi)\). Since \(T \setminus \hat{T}(\pi) \neq \emptyset\) is the set of types not ruled out by \(\text{D}1\) \(\beta\) is consistent with \(\text{D}1\). The proof proceeds through three claims. Claims 1 and 2 rule out
profitable deviations to any \( \pi \in \{ \pi^1, ..., \pi^n \} \), while Claims 2 and 3 rule out profitable deviations to any \( \pi \in \Pi \backslash \{ \pi^1, ..., \pi^n \} \).

**Claim 1.** For each \( t, t' \in \{1, ..., n\} \) we have \( \hat{V}(\pi^t, \mu_t, \mu_t) \geq \hat{V}(\pi^{t'}, \mu_{t'}, \mu_t) \).

Proof. Notice that for \( t \in \{2, ..., n\} \) we have \( \hat{V}(\pi^t, \mu_t, \mu_t) \geq \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_t) \). For, otherwise \( \hat{V}^D_t \leq \hat{V}(\pi^t, \mu_t, \mu_t) < \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_t) \) so \( \pi^{t-1} \not\in \Pi^D \), and by Lemma 2 there then is some \( \pi' \) such that \( \hat{V}(\pi', \mu_t, \mu_t) - \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_t) > 0 > \hat{V}(\pi', \mu_t, \mu_t) - \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_t) \). Then \( \pi' \in Z_t(\pi^{t-1}) \) and \( \hat{V}(\pi', \mu_t, \mu_t) > \hat{V}(\pi^t, \mu_t, \mu_t) \), a contradiction.

Consider any \( t \in \{1, ..., n\} \) and \( \hat{t} \in \{2, ..., n\} \). Then \( \hat{V}(\pi^{\hat{t}}, \mu_{\hat{t}}, \mu_{\hat{t}}) \geq \hat{V}(\pi^{\hat{t}-1}, \mu_{\hat{t}-1}, \mu_{\hat{t}}) \) and \( \hat{V}(\pi^{\hat{t}}, \mu_{\hat{t}}, \mu_{\hat{t}}) \leq \hat{V}(\pi^{\hat{t}-1}, \mu_{\hat{t}-1}, \mu_{\hat{t}}) \). Since \( \hat{V} \) is linear in its third argument \( \hat{t} \in \{1, ..., t\} \) implies \( \hat{V}(\pi^{\hat{t}-1}, \mu_{\hat{t}-1}, \mu_{\hat{t}}) \leq \hat{V}(\pi^{\hat{t}}, \mu_{\hat{t}}, \mu_{\hat{t}}) \) and \( \hat{t} \in \{t + 1, ..., n\} \) implies \( \hat{V}(\pi^{\hat{t}-1}, \mu_{\hat{t}-1}, \mu_{\hat{t}}) \geq \hat{V}(\pi^{\hat{t}}, \mu_{\hat{t}}, \mu_{\hat{t}}) \).

Applying the inequalities for \( \hat{t} = 2, ..., n \) gives \( \hat{V}(\pi^{t}, \mu_t, \mu_t) \geq \hat{V}(\pi^{t'}, \mu_{t'}, \mu_t) \) for any \( t' \in \{1, ..., n\} \).

**Claim 2.** For any \( t, t' \in \{1, ..., n\} \) we have \( \hat{V}(\pi^{t'}, \beta(\pi^{t'}), \mu_t) = \hat{V}(\pi^{t'}, \mu_{t'}, \mu_t) \).

Proof. The claim is obvious if \( \pi^{t'} \) is separating. Suppose instead \( \pi^{t'} = \pi'' \) = \( \pi \) for some \( t' < t'' \). If \( \pi \not\in \Pi^D \) then \( \hat{V}(\pi^{t'}, \mu_{t'}, \mu_{t'}) < \hat{V}(\pi^{t''}, \mu_{t''}, \mu_{t''}) \), contradicting Claim 1. Hence \( \pi \in \Pi^D \), implying \( \hat{V}(\pi^{t'}, \beta(\pi^{t'}), \mu_t) = \hat{V}(\pi^{t'}, \mu_{t'}, \mu_t) \).

**Claim 3.** For any \( t \in \{1, ..., n\} \) and \( \pi \in \Pi \backslash \{ \pi^1, ..., \pi^n \} \) we have \( \hat{V}(\pi^t, \mu_t, \mu_t) \geq \hat{V}(\pi, \beta(\pi), \mu_t) \).

Proof. Suppose, to contradiction, that for some \( t \in \{1, ..., n\} \) and \( \pi \in \Pi \backslash \{ \pi^1, ..., \pi^n \} \) we have \( \hat{V}(\pi^t, \mu_t, \mu_t) < \hat{V}(\pi, \beta(\pi), \mu_t) \), implying \( \pi \not\in \Pi^D \). Let \( \mu_{\hat{t}} := \min T \backslash \hat{T}(\pi) = \beta(\pi) \).

(a) Suppose \( \mu_t \geq \mu_{\hat{t}} \). First notice that \( \hat{V}(\pi^{\hat{t}}, \mu_{\hat{t}}, \mu_{\hat{t}}) < \hat{V}(\pi, \mu_{\hat{t}}, \mu_{\hat{t}}) \). For, if for some \( \hat{t} < t' \leq t \) it holds that \( \hat{V}(\pi^{t'}, \mu_{t'}, \mu_{t'}) < \hat{V}(\pi, \mu_{t'}, \mu_{t'}) \) and \( \hat{V}(\pi^{t'-1}, \mu_{t'-1}, \mu_{t'-1}) \geq \hat{V}(\pi, \mu_{t'}, \mu_{t'}) \), then by Lemma 2 there is some \( \pi' \in \Pi \) such that \( \hat{V}(\pi', \mu_{t'}, \mu_{t'}) = \hat{V}(\pi, \mu_{t'}, \mu_{t'}) \). Then \( \pi' \in Z^D_t(\pi^{t'-1}) \) and \( \hat{V}(\pi', \mu_{t'}, \mu_{t'}) = \hat{V}(\pi^{t'}, \mu_{t'}, \mu_{t'}) \), a contradiction.

Since \( \pi^1 \in \Pi^1 \) it must be that \( \hat{t} > 1 \) and we have \( \mu_{\hat{t}-1} \in \hat{T}(\pi) \) and \( \mu_{\hat{t}} \in T \backslash \hat{T}(\pi) \). Since \( \hat{V}(\pi, \beta(\pi)) \) is strictly increasing in its second argument \( D^0(\pi, \mu_{\hat{t}-1}) \subseteq D(\pi, \mu_{\hat{t}}) \). Combining these observations with \( \mu_{\hat{t}} \in D(\pi, \mu_{\hat{t}}) \) it follows that there is some \( \mu' \in [\mu_1, \mu_{\hat{t}}] \) such that \( \hat{V}(\pi^{\hat{t}}, \mu_{\hat{t}}, \mu_{\hat{t}}) < \hat{V}(\pi, \mu', \mu_{\hat{t}}) \) and \( \hat{V}(\pi^{\hat{t}-1}, \mu_{\hat{t}-1}, \mu_{\hat{t}-1}) \geq \hat{V}(\pi, \mu', \mu_{\hat{t}-1}) \). By Lemma 2 there is some \( \pi' \in \Pi \) such
that $\tilde{V}(\pi', \mu_t, \mu_t^i) - \tilde{V}(\pi, \mu_t, \mu_t^i) > 0 > \tilde{V}(\pi', \mu_t, \mu_{t-1}) - \tilde{V}(\pi, \mu_{t-1}, \mu_{t-1})$. Then $\pi' \in Z_t(\pi_{t-1})$ and $\tilde{V}(\pi', \mu_t, \mu_t^i) > \tilde{V}(\pi_t^i, \mu_t^i, \mu_t^i)$, a contradiction.

(b) Suppose $\mu_t < \mu_t^i$. There is then some $\mu_{t'} \in T \setminus \tilde{T}(\pi)$ such that $D^0(\pi, \mu_t) \not\subseteq D(\pi, \mu_{t'})$. By the previous case $V(\pi', \mu_{t'}, \mu_{t'}) \geq V(\pi, \mu_t, \mu_t^i)$, so $\mu_t^i \not\in D(\pi, \mu_{t'})$. Since $D^0(\pi, \mu_t) \not\subseteq D(\pi, \mu_{t'})$ we have $\mu_t^i \not\in D^0(\pi, \mu_t)$ and therefore $V(\pi^t, \mu_t, \mu_t^i) > V(\pi, \mu_t, \mu_t^i)$, a contradiction. $\Box$

**Step 2.** The "only if" part of (i).

Proof. Suppose $(\pi^1, ..., \pi^n)$, $\alpha$ and $(\tilde{\beta}, \tilde{\beta})$ is an equilibrium. Proposition 1 implies that if $\pi_t^1 = \pi_t^{i'}$ for some $t \neq t'$ then $\pi^t \in \Pi^{FD}$. Hence, for any $t, t' \in \{1, ..., n\}$ we have $\tilde{V}(\pi_t^{i'}, \tilde{\beta}(\pi_t^{i'}), \mu_t) = \tilde{V}(\pi_t^{i'}, \mu_{t'}, \mu_t)$. Obviously, $\pi^1 \in \Pi_t^1$ and $\pi^t \in Z_t(\pi^t-1)$ for all $t \in \{2, ..., n\}$. Suppose therefore, to contradiction, that there is some $t \in \{2, ..., n\}$ and $\pi \in \Pi$ such that $\tilde{V}(\pi, \mu_t, \mu_t) > \tilde{V}(\pi_t^{i'}, \mu_{t'}, \mu_t)$ and $\tilde{V}(\pi, \mu_t, \mu_t-1) \leq \tilde{V}(\pi_t^{i-1}, \mu_{t-1}, \mu_{t-1})$, implying $\pi \not\in \Pi^{FD}$. Since $\tilde{V}(\pi, \mu_t, \mu_t) > \tilde{V}(\pi_t^{i-1}, \mu_{t-1}, \mu_{t-1})$ and $\tilde{V}$ is linear in its third argument, $\tilde{V}(\pi, \mu_t, \mu_{t'}) \leq \tilde{V}(\pi_t^{i-1}, \mu_{t-1}, \mu_{t'})$ for all $t' \in \{1, ..., t-1\}$. By incentive compatibility $\tilde{V}(\pi, \mu_t, \mu_{t'}) \leq \tilde{V}(\pi_t^{i-1}, \mu_{t-1}, \mu_{t'}) \leq \tilde{V}(\pi_t^{i'}, \mu_{t'}, \mu_{t'})$. Since $\pi \not\in \Pi^{FD}$ we have that $\tilde{V}(\pi, \cdot, \mu_{t'})$ is strictly increasing and therefore $D^0(\pi, \mu_{t'}) \not\subseteq D(\pi, \mu_t)$ for all $t' \in \{1, ..., t-1\}$. D1 requires $\tilde{\beta}(\pi) \geq \mu_t$, so $\tilde{V}(\pi, \tilde{\beta}(\pi), \mu_t) \geq \tilde{V}(\pi, \mu_t, \mu_t) > \tilde{V}(\pi_t^{i'}, \mu_{t'}, \mu_{t'})$ and type $\mu_t$ has a profitable deviation to $\pi$, a contradiction. $\Box$

**Step 3.** Part (ii): There is an equilibrium.

Proof. Since $\Pi$ is a compact subset of Euclidean space and $\tilde{V}$ is continuous $\max_{\pi \in \Pi} \tilde{V}(\pi, \mu, \mu_t)$ is well defined and therefore $\Pi_t^1 \neq \emptyset$. Let $\pi^1 \in \Pi_t^1$. Since $\Pi^{FD} \subseteq \Pi$ we have $\Pi^{FD} \subseteq Z_t^1(\pi^1) \neq \emptyset$. Since $\tilde{V}$ is continuous $Z_t^1(\pi^1)$ is compact and $\max_{\pi \in Z_t^1(\pi^1)} \tilde{V}(\pi, \mu_2, \mu_2)$ is well defined. Iterating the argument for $t \in \{3, ..., n\}$ the problem of finding $(\pi^1, ..., \pi^n)$ solving (1) and (2) has a solution and Step 1 implies that there is an equilibrium. $\blacksquare$

**Proof.** (Proposition 2) Suppose $(\pi^1, ..., \pi^n)$ is an equilibrium sender strategy with supports $\{s_t^1, ..., s_t^{k_t}\}$ and final updated equilibrium beliefs $\{\beta_{t_i}^k\}_{i=1}^{k_t}$ for each $t \in \{1, ..., n\}$. The proof relies on the characterization of optimal signals in terms of the "concavification" of $\tilde{v}$ developed by Kamenica and Gentzkow (2011). Corollaries 1 and 2 in Kamenica and Gentzkow (2011) imply that $\pi^t \in \Pi_t^1$ if and only if there is a line $\xi^t(\cdot)$ such that $\xi^t(\beta) = \tilde{v}(\beta)$ for all $\beta \in \{\beta_{t_i}^k\}_{i=1}^{k_t}$ and
\( \zeta(x) \geq \tilde{v}(x) \) for all \( x \in [0, 1] \). In this case \( \tilde{V}(\pi^t, \mu_t, \mu_t) = \zeta(x) \).

**Step 1.** For all \( t \in \{2, ..., n\} \) we have \( \pi^t \in \Pi^{FD}_t \Rightarrow \pi^t \in \Pi^*_t \).

Proof. Suppose, to contradiction, that \( \pi^t \in \Pi^{FD}_t \) and \( \pi^t \notin \Pi^*_t \) for some \( t \in \{2, ..., n\} \). Since \( \pi^t \notin \Pi^*_t \) the line \( \zeta^{FD}(\cdot) \) through \((0, \tilde{v}(0))\) and \((1, \tilde{v}(1))\) satisfies \( \zeta^{FD}(x) < \tilde{v}(x) \) for some \( x \in (0, 1) \), which implies \( \Pi^{FD}_t \cap (\bigcup_{j=1}^{n} \Pi^*_j) = \emptyset \). By Theorem 1 we then have \( \tilde{V}(\pi^1, \mu_1, \mu_1) > \tilde{V}^{FD}_t \). Let \( t := \max\{t' \in \{1, ..., t\} : \tilde{V}(\pi^{t'}, \mu_{t'}, \mu_{t'}) > \tilde{V}^{FD}_t \} + 1 \), so \( \tilde{V}(\pi^t, \mu_t, \mu_t) < \tilde{V}_t \) and therefore \( \pi^t \notin \Pi^*_t \).

Consider some \( \pi^* \in \Pi^*_t \) with support \( S^* = \{s_{1}^*, ..., s_{k^*}^*\} \), where \( \tilde{V}(\pi^*, \mu_t, \mu_t) < \tilde{V}(\pi^*, \mu_t, \mu_t) \).

Let \( g(\gamma) \) denote a family of signals parameterized by \( \gamma \in [0, 1] \), with support \( S^* \cup \{s_0^*, s_{k^*+1}^*\} \) and such that \( g_{ij}(\gamma) = \gamma \pi^*_{ij} \) for any \( s_i^* \in S^* \) and \( j \in \{L, H\} \) and \( g_{0L}(\gamma) = g_{k^*+1H}(\gamma) = 1 - \gamma \). Then \( B(g(\gamma), s_i^*, \mu_t) = B(\pi^*, s_i^*, \mu_t) \) for any \( s_i^* \in S^* \), \( B(g(\gamma), s_0^*, \mu_t) = 0 \) and \( B(g(\gamma), s_{k^*+1}^*, \mu_t) = 1 \).

Further, \( g(0) \in \Pi^{FD}_t \). Then

\[
\tilde{V}(g(\gamma), \mu_t, \mu_t) = \sum_{i=1}^{k^*} (\mu_i \pi^*_{ij} + (1 - \mu_i) \pi^*_{iL}) \tilde{v}(B(\pi^*, s_i^*, \mu_t)) + (1 - \mu_i) \tilde{v}(0) + \mu_i \tilde{v}(1) \]

for any \( \gamma \in (0, 1] \). Since by construction \( \tilde{V}(g(0), \mu_t, \mu_t) = \tilde{V}^{FD}_{t-1} < \tilde{V}(\pi^{t-1}, \mu_{t-1}, \mu_{t-1}) \) there is some \( \gamma > 0 \) such that \( \tilde{V}(\tilde{\gamma}(\gamma), \mu_t, \mu_t) > \tilde{V}(\pi^t, \mu_t, \mu_t) \) and \( \tilde{V}(\tilde{\gamma}(\gamma), \mu_t, \mu_t) < \tilde{V}(\pi^{t-1}, \mu_{t-1}, \mu_{t-1}) \), contradicting (2) in Theorem 1. □

**Step 2.** For all \( t \in \{2, ..., n\} \) we have \( \pi^t \in \Pi^*_t \Rightarrow \pi^t \in \Pi^{FD}_t \).

Proof. Suppose, to contradiction, that \( \pi^t \in \Pi^*_t \) and \( \pi^t \notin \Pi^{FD}_t \) for some \( t \in \{2, ..., n\} \) and recall that \( \pi^1 \in \Pi^*_1 \). For \( t' \in \{1, t\} \) let \( \zeta^{t'}(\cdot) \) be a line such that \( \zeta^{t'}(\beta_i^{t'}) = \tilde{v}(\beta_i^{t'}) \) for all \( i \in \{1, ..., k_{t'}\} \) and \( \zeta^{t'}(x) \geq \tilde{v}(x) \) for all \( x \in [0, 1] \).

First notice that \( \zeta^1(\mu_1) \geq \zeta^1(\mu_1) \). To see this, suppose instead \( \zeta^1(\mu_1) < \zeta^1(\mu_1) \) and notice that \( \beta_1^1 \leq \mu_1 \leq \beta_1^k \). The properties of \( \zeta^1(\cdot) \) and \( \zeta(\cdot) \) imply that either \( \tilde{v}(x) \leq \zeta^1(x) < \zeta^1(x) \) on some \([0, \tilde{x}]\) with \( \tilde{x} > \mu_1 \) and therefore \( \beta_1^1 > \mu_1 \), a contradiction. Otherwise \( \tilde{v}(x) \leq \zeta^1(x) < \zeta^1(x) \) on some \([\tilde{x}, 1]\) with \( \tilde{x} < \mu_1 \) and \( \beta_1^k < \mu_1 \), a contradiction.

We have \( \tilde{V}(\pi^{t'}, \mu_{t'}, \mu_{t'}) = \zeta^{t'}(\mu_{t'}) \) for \( t' \in \{1, t\} \) and by deviating to \( \pi^t \) type \( \mu_1 \) obtains \( \zeta^t(\mu^*) \).
with \( \mu^* = \sum_{i=1}^{k_t} (\mu_t \pi_{iH}^t + (1 - \mu_t) \pi_{iL}^t) \beta_i^t \). We have

\[
\mu_t - \mu^* = \sum_{i=1}^{k_t} (\mu_t \pi_{iH}^t + (1 - \mu_t) \pi_{iL}^t) \beta_i^t - \sum_{i=1}^{k_t} (\mu_1 \pi_{iH}^t + (1 - \mu_1) \pi_{iL}^t) \beta_i^t = (\mu_t - \mu_1) \sum_{i=1}^{k_t} (\pi_{iH}^t - \pi_{iL}^t) \beta_i^t < \mu_t - \mu_1,
\]

where \( \sum_i (\pi_{iH}^t - \pi_{iL}^t) \beta_i^t < 1 \) since \( \pi^t \not\in \Pi^{FD} \). Hence, \( \mu^* > \mu_1 \) and since \( \varsigma^t(\mu^*) > \varsigma^t(\mu_1) \geq \varsigma^t(\mu_1) \) the deviation is profitable, contradicting that \( (\pi^1, \ldots, \pi^n) \) is an equilibrium strategy. \( \square \)

Finally, \( \Pi^*_t \cap \Pi^{FD}_t \neq \emptyset \Rightarrow \pi^t \in \Pi^*_t \) since \( \hat{V}^t \) is a lower bound on type \( \mu_t \)'s equilibrium payoff. The reverse implication follows from Step 2. \( \blacksquare \)

**Proof.** (Proposition 3) Suppose \( (\pi^1, \ldots, \pi^n) \) is an equilibrium sender strategy and, to contradiction, that \( \pi \) is more precise than \( \pi^t \) and \( \hat{V}(\pi, \mu_t, \mu_t) > \hat{V}(\pi^t, \mu_t, \mu_t) \) for some \( t \in \{2, \ldots, n\} \). Since \( \hat{V}(\pi, \mu_t, \mu_t) > \hat{V}^{FD}_t, \pi \not\in \Pi^{FD} \). Let \( g(\gamma) \) be a convex combination of \( \pi \) 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. We have \( \hat{V}(g(0), \mu_t, \mu_t) > \hat{V}(\pi^t, \mu_t, \mu_t) \geq \hat{V}(g(1), \mu_t, \mu_t) \). Let \( \gamma = \min\{\gamma \in [0, 1] : \hat{V}(g(\gamma), \mu_t, \mu_t) = \hat{V}(\pi^t, \mu_t, \mu_t)\} \). Since \( g(\gamma) \) is strictly more precise than \( \pi^t \) Lemma 1 implies the first of the following inequalities, while type \( \mu_{t-1} \)'s incentive compatibility implies the second

\[
\hat{V}(g(\gamma), \mu_t, \mu_t) - \hat{V}(\pi^t, \mu_t, \mu_t) > \hat{V}(g(\gamma), \mu_t, \mu_{t-1}) - \hat{V}(\pi^t, \mu_t, \mu_{t-1}) \geq \hat{V}(g(\gamma), \mu_t, \mu_{t-1}) - \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_{t-1}).
\]

There is then some \( \tilde{\gamma} < \gamma \) such that \( \hat{V}(g(\tilde{\gamma}), \mu_t, \mu_t) - \hat{V}(\pi^t, \mu_t, \mu_t) > 0 > \hat{V}(g(\tilde{\gamma}), \mu_t, \mu_{t-1}) - \hat{V}(\pi^{t-1}, \mu_{t-1}, \mu_{t-1}) \), contradicting (2) in Theorem 1. \( \blacksquare \)

### 5.3 Proof of Proposition 4

**Proof.** (Proposition 4) \( \hat{V} \) is here extended to interim beliefs \( \mu \in \{0, 1\} \) in accordance with the receiver responses to unexpected events assumed in Section 3.3. It suffices to extend \( B(\pi, s, \mu) \)
to $\mu \in \{0,1\}$. In particular, whenever $B(\pi, s_i; \mu)$ is not defined for some $\pi \in \Pi$, $s_i \in S$ and $\mu \in \{0,1\}$, set $B(\pi, s, \mu) = 1$ if $\mu = 0$ and $B(\pi, s, \mu) = 0$ if $\mu = 1$.

**Step 1.** The "only if" statement.

Proof. Suppose $\mu_1 = 0$ and that $(\pi^1, \ldots, \pi^n)$, $\alpha$ and $(\tilde{\beta}, \bar{\beta})$ is an equilibrium. First notice that the proofs of Lemmata 1 and 2 and Proposition 1 do not require the assumption $\mu_1 > 0$. Proposition 1 therefore extends to the case $\mu_1 = 0$, so for each $t \in \{1, \ldots, n\}$, $\pi^t$ is separating or fully disclosing. Then $\tilde{V}(\pi^1, \tilde{\beta}(\pi^1), \mu_1) = \hat{v}(0)$. Suppose, to contradiction, that $\pi^t \not\in \Pi^{FD}$ for some $t \in \{2, \ldots, n\}$. There is then some outcome $s_i$ such that $\pi^t_{iL} > 0$, $\pi^t_{iH} > 0$ and $B(\pi^t, s_i, \mu_i) > 0$. By deviating to $\pi^t$ type $\mu_1$ obtains at least $\pi^t_{iL}\hat{v}(B(\pi, s_i, \mu_1)) + (1 - \pi^t_{iL})\hat{v}(0) > \hat{v}(0)$, i.e., the deviation is profitable, a contradiction.$\square$

**Step 2.** The "if" statement.

Proof. Suppose $\mu_1 = 0$, $\pi^1 \in \Pi$, $\pi^t \in \Pi^{FD}$ for all $t \in \{2, \ldots, n\}$ and that $\alpha$ is optimal given rational beliefs $(\tilde{\beta}, \bar{\beta})$. Then $\tilde{V}(\pi^t, \tilde{\beta}(\pi^t), \mu_t) = \tilde{V}^{FD}_t$ for all $t \in \{1, \ldots, n\}$. For any $\pi \in \Pi \setminus \Pi^{FD}$ and $\mu > \mu_1$ it holds that $\tilde{V}(\pi, \mu, \mu_1) > \hat{v}(0) = \tilde{V}(\pi^1, \tilde{\beta}(\pi^1), \mu_1)$ (see Step 1). For any $t \in \{1, \ldots, n\}$

$$\tilde{V}(\pi, \mu_1, \mu_t) \leq \sum_{i=1}^{k} [\mu_i\pi_{iH}\hat{v}(1) + (1 - \mu_i)\pi_{iL}\hat{v}(B(\pi, s_i, \mu_1))] = \mu_t\hat{v}(1) + (1 - \mu_t)\hat{v}(0) = \tilde{V}^{FD}_t.$$

Therefore $D(\pi, \mu_t) \subseteq D(\pi, \mu_1)$ for all $t \in \{2, \ldots, n\}$ and $D_1$ allows setting $\tilde{\beta}(\pi) = \mu_1$ for all $\pi \in \Pi \setminus \Pi^{FD}$. Since $\tilde{V}(\pi, \mu_1, \mu_t) \leq \tilde{V}^{FD}_t$ for all $t \in \{1, \ldots, n\}$ and $\pi \in \Pi \setminus \Pi^{FD}$ it follows that no type has a profitable deviation and we have an equilibrium.$\square$

**5.4 Binary signals and loss of generality**

In the absence of private information it would be without loss of generality to confine attention to binary signals, i.e., to set $|\tilde{S}| = 2$, as long as the only concern is maximizing the sender’s expected payoff. This is true even if sender and receiver have different prior beliefs (see Alonso and Câmara 2014a). Intuitively, the sender’s expected payoff can be represented as a point in the convex hull of the graph of $\hat{v}$ and one can argue that it suffices to consider convex combinations of two such
Here, however, confining attention to binary signals does imply a loss of generality. While \( \max_{\pi \in \Pi} \tilde{V}(\pi, \mu, \mu_i) \) does not depend on \( |\tilde{S}| \), the same is not true when adding the constraint in (2) in Theorem 1. Here I present an example which proves this point.

Suppose there are only two types, that \( \mu_1 = 1/8 \) and \( \mu_2 = 3/4 \) and let

\[
\tilde{v}(x) = \begin{cases} 
    x & \text{if } 0 \leq x < \frac{1}{4} \\
    2x - \frac{1}{4} & \text{if } \frac{1}{4} \leq x < \frac{1}{2} \\
    \frac{1}{2}x + \frac{1}{2} & \text{if } \frac{1}{2} \leq x < \frac{3}{4} \\
    2x - \frac{5}{8} & \text{if } \frac{3}{4} \leq x \leq 1
\end{cases}
\]

Here any \( \pi^1 \in \Pi^*_1 \) induces posterior beliefs \( (\beta_1^1, \beta_2^1) = (0, 1/2) \) and \( \tilde{V}(\pi^1, \mu_1, \mu_1) = 3/16 \). A binary equilibrium signal \( \pi^2 \) of type \( \mu_2 \) must then induce posterior beliefs \( \beta_1^2 < 1/4 \) and \( \beta_2^2 > 3/4 \). Type \( \mu_2 \)'s expected payoff given such a signal is \( \tilde{V}(\pi, \mu_2, \mu_2) = \frac{1}{32} (-9\pi_{1H} - 5(1 - \pi_{1L}) + 33) \). The unique maximizer of this expression is \( \pi_{1H} = 0 \) and \( \pi_{1L} = 1 \) and the corresponding expected payoff is 33/32. Since fully disclosing signals provide lower bounds on equilibrium expected payoffs, the fully disclosing signal \( \pi_{1H} = 0 \) and \( \pi_{1L} = 1 \) is the only possible binary equilibrium signal of type \( \mu_2 \).

The fully disclosing signal is not optimal, however. Type \( \mu_2 \)'s optimal signal is given by \( \pi_{1H} = 1/3 \) and \( \pi_{1L} = 1 \), induces posterior beliefs \( (\beta_1^2, \beta_2^2) = (1/2, 1) \) and provides type \( \mu_2 \) expected payoff 34/32. Proposition 2 implies that type \( \mu_2 \)'s equilibrium signal cannot be fully disclosing and it can therefore not be binary.

Figure 1 below illustrates the example. In both panels the receiver’s posterior belief is on the horizontal axis. The left panel illustrates the induced posterior beliefs of the optimal signals of both types and the corresponding expected payoffs. The right panel illustrates the fully disclosing signal and corresponding expected payoffs (upper dashed line), as well as a perturbation of the fully disclosing signal which induces posterior beliefs \( \beta_1' \) and \( \beta_2' \) (lower dashed line). The expected payoff of type \( \mu_2 \) given the perturbation (i.e., \( \tilde{V}'_2 \)) is lower than the expected payoff given the fully disclosing signal.

---

The graphical representation of signals in terms of updated receiver beliefs used here is due to Kamenica and Gentzkow (2011).
Figure 1. In the left panel, $\hat{V}_i^*$ and $\beta_i^t$ represent the expected payoff and induced posterior beliefs of type $\mu_i$’s optimal signal. In the right panel, $\hat{V}_2'$ and $\beta_i'$ represent the expected payoff of type $\mu_2$ and induced posterior beliefs given a perturbation of a fully disclosing signal.

In the example $\tilde{v}$ is first convex, then concave and then convex. This implies that while a fully disclosing signal is not optimal, it is locally optimal, and generating a higher expected payoff for type $\mu_2$ with a binary signal requires perturbing the fully disclosing signal to the point of violating incentive compatibility. By instead perturbing the fully disclosing signal by adding a third signal outcome generating a posterior belief $1/2$ with small probability, type $\mu_2$ increases his expected payoff and remains incentive compatible. This is essentially the argument used in the proof of Proposition 2 to explain why fully disclosing signals can only be used in equilibrium if they are optimal.

References


*Econom.* 55, 1349-1365.

*The Bell J. of Econ.* 12, 380-391.


 of Econ.* 17, 18-32.


469-474.


423.