A Theorem on Bayesian Updating and Applications to Communication Games *

Navin Kartik† Frances Xu Lee‡ Wing Suen§

June 13, 2017

Abstract

We develop a result on Bayesian updating. Roughly, when two senders have different priors, each believes that a (Blackwell) more informative experiment will, on average, bring the other's posterior closer to his own prior. We apply the result to two models of strategic communication: multi-sender voluntary disclosure and costly falsification/signaling. Under multi-sender voluntary disclosure, senders' information revelation are strategic complements when there is a cost to concealing or falsifying information, but strategic substitutes when there is a cost to disclosing information. Competition between senders benefits a receiver in the former cases but not necessarily in the latter. Under costly signaling, when receiver's external information is better, the sender engages in less wasteful signaling and reveals more information.

Preliminary and incomplete. Please do not circulate without permission.

*Together with its companion paper (Kartik, Lee, and Suen, 2017), this manuscript supersedes an earlier working paper of Xu and Suen, "Investment in Concealable Information by Biased Experts." We thank Nageeb Ali, Ilan Guttman, Alessandro Lizzeri, Steve Matthews, Ariel Pakes, Andrea Prat, Mike Riordan, Satoru Takahashi, Jidong Zhou, and various seminar and conference audiences for comments. We are grateful to Enrico Zanardo for alerting us to a gap in a proof in an earlier draft. Sarba Raj Bartaula and Teck Yong Tan provided excellent research assistance.

†Department of Economics, Columbia University. E-mail: nkartik@columbia.edu.

‡Quinlan School of Business, Loyola University Chicago. E-mail: francesxu312@gmail.com.

§School of Economics and Finance, The University of Hong Kong. E-mail: wsuen@econ.hku.hk.

1. Introduction

Rational senders revise their beliefs upon receiving new information. From an ex-ante point of view, however, information should not systematically push one's beliefs up or down. Indeed, a fundamental property of Bayesian updating is that beliefs are a martingale: a sender's expectation of his own posterior belief is equal to his prior belief. But what about a sender's expectation of the posterior belief of *another sender* when their current beliefs are different? Besides being of intrinsic interest, such predictions can be important in situations of asymmetric information, even when senders begin with a common prior.

This paper makes two contributions. First, we provide a result concerning Bayesian updating by senders who disagree on the distribution of a fundamental. Second, we use this result to derive new economic insights about multi-sender communication games; in particular, we identify some circumstances in which competition promotes the disclosure of information and other circumstances when it does not.

A theorem on Bayesian updating. In Section 2, we develop the following result on Bayesian updating. (Throughout this introduction, technical details are suppressed.) Let $\Omega = \{0, 1\}$ be the possible states of the world. Anne (A) and Bob (B) have mutually-known but possibly-different priors on Ω, $\overline{\beta}_A$ and $\overline{\beta}_B$ respectively, where $\overline{\beta}_i$ is the probability i ascribes to state $\omega = 1$. A signal, s, will be drawn from an information structure or experiment, $\mathcal{E}$, given by a family of probability distributions, $\{p_\omega(\cdot)\}_{\omega\in\Omega}$, where $p_\omega(s)$ is the probability of observing signal s when the true state is ω. Anne and Bob agree on the experiment. Let $\beta_i(s)$ denote i's posterior (on $\omega = 1$) after observing signal s, derived by Bayesian updating. Let $\mathbb{E}_i^{\mathcal{E}}[\beta_j(\cdot)]$ be the ex-ante expectation of i of the posterior of j given the experiment $\mathcal{E}$, where the expectation is taken over signals from i's point of view.

Consider two experiments $\mathcal{E}$ and $\tilde{\mathcal{E}}$ that are comparable in the sense of Blackwell (1951, 1953); specifically, let $\tilde{\mathcal{E}}$ be a *garbling* of $\mathcal{E}$.[1] We prove (Theorem 1) that in this case,

$$\operatorname{sign}\left\{\mathbb{E}_i^{\mathcal{E}}[\beta_j(\cdot)] - \mathbb{E}_i^{\tilde{\mathcal{E}}}[\beta_j(\cdot)]\right\} = \operatorname{sign}\left\{\overline{\beta}_i - \overline{\beta}_j\right\}. \tag{1}$$

In words: if Anne is initially more (resp., less) optimistic than Bob, then Anne expects a more informative experiment to, on average, raise (resp., lower) Bob's posterior to a larger extent. Put differently, Anne expects more information to further validate her prior in the sense of bringing, on average, Bob's posterior closer to her prior; of course, Bob expects just the re-

[1]Throughout, binary comparisons should be understood in the weak sense unless indicated otherwise, e.g. "greater" means "at least as large as".

verse. For brevity, we refer to the result as *information validates the prior*, IVP hereafter.

IVP subsumes the familiar martingale property of Bayesian updating. To see this, simply consider $j = i$ in Equation 1. Then take $\mathbb{E}^{\tilde{\mathcal{E}}}$ to be an uninformative experiment, which leaves a sender's posterior after any signal equal to his prior, and take $\mathbb{E}^{\mathcal{E}}$ to be an informative experiment. IVP also implies that any sender expects another sender's posterior to fall in between their two priors.[2]

Theorem 2 provides a generalization of IVP to non-binary state spaces; the key is to work with a suitable statistic of the senders' beliefs and likelihood-ratio ordering conditions on the senders' priors and the experiments.

Naturally IVP is applicable whenever parties have different prior beliefs. It is also applicable in models that start with a common prior but has pooling in equilibrium. Because the uninformed party often has equilibrium beliefs different from an informed party, IVP is useful for analyzing the informed party's prediction of how the uninformed party would respond to exogenous information. We study one such application by analyzing a multi-sender disclosure game in Section 3. IVP is also applicable in models even if the equilibrium is fully separating because it is relevant for studying the incentive compatibility constraints of the informed party. Incentive compatibility requires that a type prefers to reveal his type than to mimic some other type. By mimicking some other type, this type could induce the uninformed party to have a belief different from his. Thus, IVP is useful at studying the informed party's expectation of how the uninformed party would respond (off the equilibrium path) to his deviation when there is exogenous information. We provide an example of this application in a costly signaling game in Section 4.[3]

Competition and disclosure. Our main economic application is to disclosure of information by multiple senders. Decision makers often seek information from more than one interested sender. For example, judges and arbiters rely on evidence provided by opposing sides, consumers receive product information from multiple sources, and legislative bills are shaped by the information revealed by many interest groups. In all these cases, senders will strategically disclose or conceal their information to influence the decision maker (DM, hereafter). How are a sender's incentives and behavior affected by the presence of other senders? What are

[2]To see this, note that (i) letting $\mathcal{U}$ denote an uninformative experiment, $\mathbb{E}_i^{\mathcal{U}}[\beta_j(\cdot)] = \overline{\beta}_j$, and (ii) letting $\mathcal{F}$ denote a fully informative experiment (one in which any signal reveals the state and hence leaves any two senders with the same posterior no matter their priors), $\mathbb{E}_i^{\mathcal{F}}[\beta_j(\cdot)] = \overline{\beta}_i$. The result follows using Equation 1 and the fact that any experiment is more informative than an uninformative experiment but less informative than a fully informative experiment.

[3]The equilibrium is not fully separating. The pooling is due to a bounded message space which generates interesting implication on how exogenous information affects the information generated by the signaling actions.

the implications for the DM's welfare? In particular, does a DM always benefit from multiple sources of information? Does the answer depend on whether senders' interests conflict or not? These are questions of substantial importance in a host of economic settings.

Section 3 studies a family of *voluntary disclosure* games. In these games (also referred to as persuasion games), biased senders' only instrument of influence is the certifiable or verifiable private information they are endowed with; see Milgrom (2008) and Dranove and Jin (2010) for surveys. Senders cannot explicitly lie but can choose what information to disclose and what to withhold. Much of the existing work on multi-sender disclosure games assumes that informed senders have identical information.[4] By contrast, we posit imperfectly correlated information; specifically, informed senders draw signals that are independently distributed conditional on an underlying state which affects the DM's payoff.[5] We allow for senders to either have opposing or similar biases but assume that a sender's payoff is state-independent and linear (increasing or decreasing) in the DM's belief.

Our focus is on how strategic disclosure interacts with *message costs*: either disclosure or concealment of information can entail direct costs for each sender. For simplicity, we assume this cost is independent of the sender's information. Disclosure costs are natural when the process of certifying or publicizing information demands resources such as time, effort, or hiring an outside expert; there can also be other "proprietary costs". A subset of the literature, starting with Jovanovic (1982) and Verrecchia (1983), has modeled such a cost, although primarily only in single-sender problems. On the flip side, there are contexts in which it is the suppression of information that requires costly resources. Besides direct costs, there can also be a psychic disutility to concealing information, or concealment may be discovered ex-post (by auditors, whistleblowers, or mere happenstance) and result in negative consequences for the sender through explicit punishment or reputation loss.[6] Studies of mandatory disclo-

[4]See, for example, Milgrom and Roberts (1986), Lipman and Seppi (1995), Shin (1998), Bourjade and Jullien (2011), and Bhattacharya and Mukherjee (2013). Two exceptions are Okuno-Fujiwara et al. (1990) and Hagenbach et al. (2014); these papers identify conditions for full disclosure and focus on different issues than we do.

[5]For simplicity, the state is taken to be binary. We assume that with some probability, senders are uninformed; importantly, an uninformed sender cannot certify that he is uninformed. This modeling device was introduced by Dye (1985) and Shin (1994a,b) to prevent "unraveling" (Grossman and Hart, 1980; Milgrom, 1981), and serves the same purpose here.

[6]It is now widely acknowledged that firms in the tobacco industry paid consultants, journalists, and pro-tobacco organizations to debate the evidence on the adverse effects of tobacco, and they engaged in elaborate schemes to destroy or conceal documents that contained damaging information. Ironically, irrefutable evidence of decades of these practices surfaced in the once-secret "tobacco documents" obtained in the 1990s through litigation in the United States. A more recent example of concealment cost is the $70 million dollar fine imposed by the National Highway Traffic Safety Administration on Honda Motor in January 2015 because "it did not report hundreds of death and injury claims. . . for the last 11 years nor did it report certain warranty and other claims in the same period". (The New York Times, 2015)

sure (e.g., Matthews and Postlewaite, 1985; Shavell, 1994; Dahm et al., 2009) can be viewed as dealing with infinitely large concealment costs. Two other recent papers endongenize concealment costs for specific settings. Dye (2017) endogenizes concealment costs in a single-sender game and Daughety and Reinganum (2016) endogenize multiple sources of concealment costs in a two-sender game with perfectly revealing and correlated signals.

In this setting, it is essential for each sender to predict how other senders' messages will affect the DM's posterior belief, and how this depends on his own message (disclosure or nondisclosure). The reason is that a sender must compare the expected benefit from each of his own messages, which depends on other senders' behavior, with the corresponding message cost. Our key insight is to view senders' messages as endogenously-determined experiments and bring the IVP theorem to bear. This approach allows us to provide a unified treatment—regardless of whether senders have similar or opposing biases, and regardless of whether message costs are (net) disclosure or concealment costs—and uncover new insights into the question of whether and when competition promotes disclosure.

We first establish a simple but important benchmark: without message costs, any sender in the multi-sender disclosure game uses the same disclosure threshold as he would in a single-sender game. In other words, there is a *strategic irrelevance*. The intuition is that without message costs, a sender's objective is the same regardless of the presence of other senders: he simply wants to induce the most favorable "interim belief" in the DM based on his own message. Consequently, absent message costs, the DM is strictly better off with more senders, regardless of their biases.

How do message costs alter the irrelevance result? Consider a concealment cost. In a single-sender setting, a sender i's disclosure threshold will be such that the DM's interpretation of nondisclosure is more favorable than i's private belief at the threshold type—this wedge is necessary to compensate i for the concealment cost. Nondisclosure thus generates an *interim disagreement* between the DM's belief and the threshold type's private belief; plainly, disclosure produces no such disagreement. Now add a second sender to the picture, j. Our IVP theorem implies that regardless of j's behavior, the threshold type of i predicts that j's message will, on average, make the DM's posterior less favorable to i as compared to the DM's interim belief following i's nondisclosure. On the other hand, if i discloses his information, then j's message will, on average, leave the DM's posterior equal to the DM's interim belief. Consequently, concealment is now less attractive to sender i: the benefit from making the DM's interim belief more favorable is reduced, while the cost is unchanged. IVP further implies that i's incentives are more strongly affected when j discloses more (in the sense of

j's message being more informative). In sum, senders' disclosures are *strategic complements* under a concealment cost.[7]

The logic reverses under a disclosure cost. In a single-sender setting, the DM's interpretation of nondisclosure is now less favorable than i's private belief at the threshold type—the gain from disclosing information must compensate i for the direct cost. Reasoning analogously to above, IVP now implies that disclosure becomes less attractive in the presence of another sender: the threshold type of i expects the other sender's message to, on average, make the DM's posterior more favorable to i, reducing the gains from disclosure. Consequently, senders' disclosures are *strategic substitutes* under a disclosure cost.

These results have straightforward welfare implications. In the case of concealment cost, a DM always benefits from an additional sender not only because of the information this sender provides, but also because it improves disclosure from other senders. In the case of disclosure cost, however, the strategic substitution result implies that while a DM gains some direct benefit from consulting an additional sender, the indirect effect on other senders' behavior is deleterious to the DM. In general, the net effect is ambiguous; it is not hard to construct examples in which the DM is made strictly worse off by adding a sender, even if this sender has an opposite bias to that of an existing sender. Thus, competition between senders need not increase information revelation nor benefit the DM.[8] The DM can even be made worse off when disclosure costs become lower or a sender is more likely to be informed, although either modification would help the DM in a single-sender setting.

We interpret these "perverse" welfare results as cautionary for some applications: institutional changes that appear to be improvements at first blush may in fact be detrimental. For example, given the importance of disclosure costs in arbitration, litigation, or related judicial settings (e.g., Sobel, 1989), our results qualify arguments made in favor of adversarial procedures that are based on promoting disclosure of information (e.g., Shin, 1998).

Costly signaling games. Our IVP theorem is also useful in models of asymmetric information even when the information asymmetry is entirely eliminated in equilibrium. Consider,

[7]In a different model, Bourjade and Jullien (2011) find an effect related to that we find under concealment cost. Loosely speaking, "reputation loss" in their model plays a similar role to concealment cost in ours.

[8]In quite different settings, Milgrom and Roberts (1986), Dewatripont and Tirole (1999), Krishna and Morgan (2001), and Gentzkow and Kamenica (2015) offer formal analyses supporting the viewpoint that competition between senders helps—or at least cannot hurt—a DM. Carlin et al. (2012) present a result in which increased competition leads to less voluntary disclosure. Their model can be viewed as one in which senders bear a concealment cost that is assumed to decrease in the amount of disclosure by other senders. Elliott et al. (2014) show how a DM can be harmed by "information improvements" in a cheap-talk setting, but the essence of their mechanism is not the strategic interaction between senders.

for example, a canonical sender-receiver signaling model as in Spence (1973). Among the types who separate, the receiver can infer the true type of the sender based on the observed signal, so there is no equilibrium disagreement. Nevertheless, the sender can induce an incorrect belief in the receiver by deviating. If the receiver also obtains information from other sources, the sender needs to predict the receiver's posterior belief after inducing a "wrong" interim belief. For this reason, IVP is useful to analyze the incentive compatibility constraints—and hence the signaling strategy—even for types who do separate.

We develop this point in Section 4 with an application to signaling with lying costs (Kartik, 2009). A sender has imperfect information about a state, and can falsify or manipulate his information by incurring costs. A receiver makes inferences about the state based on both the sender's signal and some other exogenous information.[9] Due to a bounded state and signal space, there is incomplete separation across sender types; we focus on equilibria in which "low" types separate and "high" types pool. IVP is used to show that better exogenous information leads the sender to expect smaller benefits from incurring the cost to generate more favorable signals. This relaxes incentive constraints, and hence better exogenous information reduces wasteful signaling and leads to more information revelation by the sender. We discuss how this result can also be interpreted as showing that competition between multiple senders can promote information revelation when evidence is costly to manipulate.

While this paper applies IVP to two models of strategic communication, we hope that the result will also be useful in other contexts. Indeed, the logic of IVP underlies the mechanisms in a few existing papers that study models with heterogeneous priors under specific information structures. In particular, see the strategic "persuasion motive" that generates bargaining delays in Yildiz (2004), motivational effects of difference of opinion in Che and Kartik (2009) and Van den Steen (2010, Proposition 5), and a rationale for deference in Hirsch (2015, Proposition 8); in a non-strategic setting, see why minorities expect lower levels of intermediate bias in Sethi and Yildiz (2012, Proposition 5).

2. Information Validates the Prior

The backbone of our analysis is a pair of theorems below that relate the informativeness of an experiment to the expectations of individuals with different beliefs. Theorem 1 is a special case of Theorem 2, but for expositional clarity, we introduce them separately.

Throughout, we use the following standard definitions concerning information structures

[9] Feltovich et al. (2002) and Daley and Green (2014) study related settings, but with some important differences both in modeling and substantive focus. Truyts (2015) studies a noisy-signaling model when the receiver receives independent information. We elaborate on the connections with these papers in Section 4.

(Blackwell, 1953). Fix any finite state space Ω with generic element ω. An *experiment* is $\mathcal{E} \equiv (S, \mathcal{S}, \{P_\omega\}_{\omega\in\Omega})$, where S is a measurable space of signals, $\mathcal{S}$ is a σ-algebra on S, and each P_ω is a probability measure over the signals in state ω. An experiment $\tilde{\mathcal{E}} \equiv (\tilde{S}, \tilde{\mathcal{S}}, \{\tilde{P}_\omega\}_{\omega\in\Omega})$ is a *garbling* of experiment $\mathcal{E}$ if there is a Markov kernel from $(S, \mathcal{S})$ to $(\tilde{S}, \tilde{\mathcal{S}})$, denoted $Q(\cdot|s)$,[10] such that for each $\omega \in \Omega$ and every set $\Sigma \in \tilde{\mathcal{S}}$, it holds that

$$\tilde{P}_\omega(\Sigma) = \int_S Q(\Sigma|s)\, dP_\omega(s).$$

This definition captures the statistical notion that, on a state-by-state basis, the distribution of signals in $\tilde{\mathcal{E}}$ can be generated by taking signals from $\mathcal{E}$ and transforming them through the state-independent kernel $Q(\cdot)$. In a sense, $\tilde{\mathcal{E}}$ does not provide any information that is not contained in $\mathcal{E}$. Indeed, $\mathcal{E}$ is also said to be *more informative* than $\tilde{\mathcal{E}}$ because every expected-utility decision maker prefers $\mathcal{E}$ to $\tilde{\mathcal{E}}$.

2.1. Binary states

In this subsection, let $\Omega = \{0, 1\}$, and take all beliefs to refer to the probability of state $\omega = 1$. There are two individuals, m and n, with respective prior beliefs $\overline{\beta}_m, \overline{\beta}_n \in (0, 1)$. Given any experiment, the individuals' respective priors combine with Bayes rule to determine their respective posteriors after observing a signal s, denoted $\beta_i(s)$ for $i \in \{m, n\}$. Let $\mathbb{E}^{\mathcal{E}}_m[\beta_n(\cdot)]$ denote the ex-ante expectation of individual m over the posterior of individual n under experiment $\mathcal{E}$. If $\overline{\beta}_m = \overline{\beta}_n$, then because the individuals' posteriors always agree, it holds that for any $\mathcal{E}$, $\mathbb{E}^{\mathcal{E}}_m[\beta_n(\cdot)] = \overline{\beta}_m$; this is the martingale or iterated expectations property of Bayesian updating. For individuals with different priors, we have:

Theorem 1. *Let $\Omega = \{0, 1\}$ and $\mathcal{E}$ and $\tilde{\mathcal{E}}$ be experiments with $\tilde{\mathcal{E}}$ a garbling of $\mathcal{E}$. Then,*

$$\mathbb{E}^{\mathcal{E}}_m[\beta_n(\cdot)] \leq \mathbb{E}^{\tilde{\mathcal{E}}}_m[\beta_n(\cdot)] \iff \overline{\beta}_m \leq \overline{\beta}_n.$$

Furthermore, $\min\{\overline{\beta}_m, \overline{\beta}_n\} \leq \mathbb{E}^{\mathcal{E}}_m[\beta_n(\cdot)] \leq \max\{\overline{\beta}_m, \overline{\beta}_n\}$.

Suppose that individual m is less optimistic than n, i.e., $\overline{\beta}_m < \overline{\beta}_n$. If an experiment $\tilde{\mathcal{E}}$ is uninformative—no signal realization would change any individual's beliefs—then $\mathbb{E}^{\tilde{\mathcal{E}}}_m[\beta_n(\cdot)] = \overline{\beta}_n > \overline{\beta}_m$. On the other hand, if an experiment $\mathcal{E}$ is fully informative—every signal reveals the state—then for any signal s, $\beta_m(s) = \beta_n(s)$, and hence $\overline{\beta}_m = \mathbb{E}^{\mathcal{E}}_m[\beta_m(\cdot)] = \mathbb{E}^{\mathcal{E}}_m[\beta_n(\cdot)] < \overline{\beta}_n$,

[10] I.e., (i) the map $s \mapsto Q(\Sigma|s)$ is $\mathcal{S}$-measurable for every $\Sigma \in \tilde{\mathcal{S}}$, and (ii) the map $\Sigma \mapsto Q(\Sigma|s)$ is a probability measure on $(\tilde{S}, \tilde{\mathcal{S}})$ for every $s \in S$.

where the first equality is the previously-noted property of Bayesian updating under any experiment. In other words, individual m believes a fully informative experiment will, on average, bring individual n's posterior perfectly in line with m's own prior, whereas an uninformative experiment will obviously entail no such convergence. (Of course, in turn, n expects m to update on average to n's prior under a fully informative experiment.) Theorem 1 generalizes this idea to monotonicity among Blackwell-comparable experiments: m anticipates that a more informative experiment will, on average, bring n's posterior closer to m's prior. For short, we will say the theorem shows that *information validates the prior*, or *IVP*.

Due to the assumption of binary states, there is a simple proof for Theorem 1. As it has some independent interest, we discuss it here. The key is to recognize that, because the individuals agree on the experiment and only disagree in their priors over the state, each individual's posterior can be written as a function of both their priors and the other individual's posterior. This observation is also used by Alonso and Câmara (2015) and Gentzkow and Kamenica (2014) to different ends. For simplicity, consider an experiment with a discrete signal space in which every signal is obtained with positive probability in both states. For any signal realization s, Bayes rule implies that the posterior belief $\beta_i(s)$ for individual $i \in \{m, n\}$ satisfies

$$\frac{\beta_i(s)}{1-\beta_i(s)} = \frac{\overline{\beta}_i}{1-\overline{\beta}_i}\frac{P_1(s)}{P_0(s)},$$

where $P_\omega(s)$ is the probability of observing s in state ω. Eliminating the likelihood ratio $P_1(s)/P_0(s)$ yields $\beta_n(s) = T(\beta_m(s), \overline{\beta}_n, \overline{\beta}_m)$, where

$$T(\beta_m, \overline{\beta}_n, \overline{\beta}_m) := \frac{\beta_m \frac{\overline{\beta}_n}{\overline{\beta}_m}}{\beta_m \frac{\overline{\beta}_n}{\overline{\beta}_m} + (1-\beta_m)\frac{1-\overline{\beta}_n}{1-\overline{\beta}_m}}. \tag{2}$$

It is straightforward to verify that this transformation mapping $T(\cdot, \overline{\beta}_n, \overline{\beta}_m)$ is strictly concave (resp., convex) in m's posterior when $\overline{\beta}_m < \overline{\beta}_n$ (resp., $\overline{\beta}_m > \overline{\beta}_n$). Theorem 1 follows as an application of Blackwell (1953), who showed that a garbling increases (resp., reduces) an individual's expectation of any concave (resp., convex) function of his posterior.[11]

The second part of Theorem 1 is a straightforward consequence of combining the first part with the facts that any experiment is garbling of a fully informative experiment while an uninformative experiment is a garbling of any experiment (and using the properties of these

[11]Since a garbling induces a mean-preserving contraction in m's posterior (by which we mean the opposite of a mean-preserving spread), one can also view the conclusion from the perspective of second-order stochastic dominance (Rothschild and Stiglitz, 1970). We should note that the crux of Blackwell (1951, 1953) and related contributions is in establishing a converse.

extreme experiments noted right after the theorem).

2.2. Many states

Consider an arbitrary finite set of states, $\Omega \equiv \{\omega_1, \ldots, \omega_L\} \subset \mathbb{R}$, and let $\omega_1 < \cdots < \omega_L$, with $L > 1$. We write $\beta(\omega_l)$ as the probability ascribed to state ω_l by a belief $\boldsymbol{\beta}$, with the notation in bold emphasizing that a belief over Ω is now a vector. We say that a belief $\boldsymbol{\beta'}$ *likelihood-ratio dominates* belief $\boldsymbol{\beta}$, written $\boldsymbol{\beta'} \geq_{LR} \boldsymbol{\beta}$ if, for all $\omega' > \omega$,

$$\beta'(\omega')\beta(\omega) \geq \beta(\omega')\beta'(\omega).$$

We denote posterior beliefs given a signal s as $\boldsymbol{\beta}(s) \equiv (\beta(\omega_1|s), \ldots, \beta(\omega_L|s))$.

Definition 1. An experiment $\mathcal{E} \equiv (S, \mathcal{S}, \{P_\omega\}_{\omega\in\Omega})$ is an *MLRP-experiment* if there is a total order on S, denoted $\succeq$ (with asymmetric relation $\succ$), such that the monotone likelihood ratio property holds: $s' \succ s$ and $\omega' > \omega \implies p(s'|\omega')p(s|\omega) \geq p(s'|\omega)p(s|\omega')$.

As is well known, the monotone likelihood ratio property (in the non-strict version above) is without loss of generality when there are only two states: any experiment is an MLRP-experiment when $L = 2$.

For any non-increasing function $h(\omega)$, let $M(\boldsymbol{\beta}) := \sum_{\omega\in\Omega} \omega\beta(\omega)$ represent the expectation of $h(\omega)$ under belief $\boldsymbol{\beta}$. The following result generalizes Theorem 1:

Theorem 2. *Consider priors $\overline{\boldsymbol{\beta}}_m \gg \mathbf{0}$ and $\overline{\boldsymbol{\beta}}_n \gg \mathbf{0}$ that are likelihood-ratio ordered, and any two MLRP-experiments $\mathcal{E}$ and $\tilde{\mathcal{E}}$ with $\tilde{\mathcal{E}}$ a garbling of $\mathcal{E}$. Then,*

$$\mathbb{E}^{\mathcal{E}}_m[M(\boldsymbol{\beta}_n(\cdot))] \leq \mathbb{E}^{\tilde{\mathcal{E}}}_m[M(\boldsymbol{\beta}_n(\cdot))] \iff M(\overline{\boldsymbol{\beta}}_m) \leq M(\overline{\boldsymbol{\beta}}_n).$$

Furthermore, $\min\{M(\overline{\boldsymbol{\beta}}_m), M(\overline{\boldsymbol{\beta}}_n)\} \leq \mathbb{E}^{\mathcal{E}}_m[M(\boldsymbol{\beta}_n(\cdot))] \leq \max\{M(\overline{\boldsymbol{\beta}}_m), M(\overline{\boldsymbol{\beta}}_n)\}$.

The likelihood-ratio ordering assumptions in Theorem 2 are tight in the following sense: (i) there exist priors $\overline{\boldsymbol{\beta}}_m \not\gtrless_{LR} \overline{\boldsymbol{\beta}}_n$ and an MLRP-experiment $\mathcal{E}$ such that $M(\overline{\boldsymbol{\beta}}_m) > M(\overline{\boldsymbol{\beta}}_n) > \mathbb{E}^{\mathcal{E}}_m[M(\boldsymbol{\beta}_{\boldsymbol{n}}(\cdot))]$; and (ii) there exist priors $\overline{\boldsymbol{\beta}}_m >_{LR} \overline{\boldsymbol{\beta}}_n$ and a non-MLRP-experiment $\mathcal{E}$ such that $M(\overline{\boldsymbol{\beta}}_m) > M(\overline{\boldsymbol{\beta}}_n) > \mathbb{E}^{\mathcal{E}}_m[M(\boldsymbol{\beta}_n(\cdot))]$. See the Supplementary Appendix for examples demonstrating these points.

In the following sections, we apply Theorem 1 to strategic communication games in which senders begin with a common prior.

3. Multi-Sender Disclosure Games

3.1. The Model

Players. There is an unknown state of the world, $\omega \in \{0, 1\}$. A decision maker (DM) will form a belief β_{DM} that the state is $\omega = 1$. For much of our analysis, all that matters is the belief that the DM holds; for welfare evaluation, however, it is useful to view the DM as taking an action a with von-Neumann Morgenstern utility function $u_{DM}(a, \omega)$. There is a finite set of N senders indexed by i. Each sender i has state-independent preferences over the DM's belief that are parameterized by a variable $b_i \in \{-1, 1\}$. Sender i's von Neumann-Morgenstern preferences are represented by the function $u(\beta_{DM}, b_i) = b_i \beta_{DM}$. Thus, each sender has linear preferences over the DM's expectation of the state; $b_i = 1$ means that sender i is biased upward (i.e., prefers higher expectations), and conversely for $b_i = -1$. All senders' biases are common knowledge. This is a normalization of the general case where $b_i \neq 0$. We say that two senders have similar biases if their biases have the same sign, and opposing biases if their biases have opposite signs.

Information. The DM relies on the senders for information to form her belief about the state. All players share a common prior π over the state. Each sender may exogenously obtain some private information about the state. Specifically, with independent probability $p_i \in (0, 1)$, a sender i is informed and receives a signal $s_i \in S$; with probability $1 - p_i$, he is uninformed, in which case we denote $s_i = \phi$. If informed, sender i's signal is drawn independently from a distribution that depends upon the true state. Without loss, we equate an informed sender's signal with his *private belief*, i.e., a sender's posterior on state $\omega = 1$ given only his own signal $s \neq \phi$ (as derived by Bayesian updating) is s. For convenience, we assume the cumulative distribution of an informed sender's signals in each state, $F(s|\omega)$ for $\omega \in \{0, 1\}$, have common support $S = [\underline{s}, \overline{s}] \subseteq [0, 1]$ and admit respective densities $f(s|\omega)$.[12]

Communication. Signals are "hard evidence"; a sender with signal $s_i \in S \cup \{\phi\}$ can send a message $m_i \in \{s_i, \phi\}$. In other words, an uninformed sender only has one message available, ϕ, while an informed sender can either report his true signal or feign ignorance by sending the message ϕ.[13] We refer to any message $m_i \neq \phi$ as *disclosure* and the message $m_i = \phi$ as

[12] It is straightforward to allow for heterogeneity across senders in the state-contingent distributions from which they independently draw signals when informed.

[13] Given that each sender's preferences are monotonic in the DM's belief and the direction of bias is known, standard "skeptical posture" arguments imply that our results would not be affected if we were to allow for a richer message space, for example if an informed sender could report any subset of the signal space that contains his true signal. Likewise, allowing for cheap talk would not affect our results as cheap talk would be

nondisclosure. When an informed sender chooses nondisclosure, we say that he is *concealing* his information. The constraint that senders must either tell the truth or conceal their information is standard; a justification is that signals are verifiable and sufficiently large penalties will be imposed on a sender if a reported signal is discovered to be untrue. Note that being uninformed is not verifiable.

Message costs. A sender i who sends message $m_i \neq \phi$ bears a utility cost $c \in \mathbb{R}$.[14] We refer to the case of $c > 0$ as one of *disclosure cost* and $c < 0$ as one of *concealment cost*. A disclosure cost captures the idea that costly resources may be needed to certify or make verifiable the information that one has; a concealment cost captures either a resource-related or psychic dis-utility from concealing available information, or it could represent expectations of possible penalties from ex-post detection of having withheld information. As is well known (Jovanovic, 1982; Verrecchia, 1983), a disclosure cost generally precludes full disclosure. For this reason, our main points under $c > 0$ do not require the assumption that $p_i < 1$. We make the assumption in order to allow for a unified treatment of both $c > 0$ and $c \leq 0$.[15]

Contracts and timing. We assume that transfers cannot be made contingent on messages or decisions, and that no player has any commitment power. The game we study is therefore the following: nature initially determines the state ω and then independently (conditional on the realized state) draws each sender i's private information, $s_i \in S \cup \{\phi\}$; all senders then simultaneously send their respective messages m_i to the DM (whether messages are public or privately observed by the DM is irrelevant); the DM then forms her belief, β_{DM}, according to Bayes rule, whereafter each sender i's payoff is realized as

$$u(\beta_{DM}, b_i) - c \cdot \mathbb{1}_{\{m_i \neq \phi\}}.$$

All aspects of the game except the state and senders' signals (or lack thereof) are common knowledge. Our solution concept is perfect Bayesian equilibrium, which we will refer to simply as "equilibrium." The notion of welfare for any player is ex-ante expected utility.

uninformative in equilibrium.

[14] To accommodate any magnitude of bias $b_i \neq 0$, this cost c can be interpreted as $\frac{c}{b_i}$.

[15] When $c = 0$ our setting is related to Jackson and Tan (2013) and Bhattacharya and Mukherjee (2013). Jackson and Tan (2013) assume a binary decision space, which makes their senders' payoffs non-linear in the DM's posterior and shifts the thrust of their analysis. They are ultimately interested in comparing different voting rules, which is effectively like changing the pivotal DM's preferences; we instead highlight how the presence of disclosure or concealment costs affects the strategic interaction between senders holding fixed the DM's preferences. Bhattacharya and Mukherjee (2013) assume that informed senders' signals are perfectly correlated but allow for senders to have non-monotonic preferences over the DM's posterior.

3.2. A single-sender game

Begin by considering a (hypothetical) game between a single sender i and the DM. In part, our analysis in this subsection generalizes existing results in the literature. For concreteness, suppose the sender is upward biased; straightforward analogs of the discussion below apply if the sender is downward biased.

For any $\beta \in [0,1]$, define $f_\beta(s) := \beta f(s|1) + (1-\beta) f(s|0)$ as the unconditional density of signal s given a belief that puts probability β on state $\omega = 1$. Let F_β be the cumulative distribution of f_β. Since the sender has private belief s upon receiving signal s, disclosure of signal s will lead to the DM also holding belief s. It follows that given any *nondisclosure belief*, i.e., the DM's posterior belief when there is nondisclosure, the optimal strategy for the sender (if informed) is a threshold strategy of disclosing all signals above some *disclosure threshold*, say $\hat{s}$, and concealing all signals below it. If the sender is uninformed, his only available message is ϕ. Suppose the sender is using a disclosure threshold $\hat{s}$. Define the function $\eta : [0,1] \times (0,1) \times [0,1] \to [0,1]$ by

$$\eta(\hat{s}, p, \pi) := \frac{1-p}{1-p+pF_\pi(\hat{s})}\pi + \frac{pF_\pi(\hat{s})}{1-p+pF_\pi(\hat{s})}\mathbb{E}_\pi[s|s<\hat{s}], \tag{3}$$

where $\mathbb{E}_\pi[\cdot]$ refers to an expectation taken with respect to the distribution F_π. This function is simply a posterior derived from Bayes rule in the event of nondisclosure using the prior π, a probability p of the sender being informed, and a conjectured disclosure threshold $\hat{s}$.

An increase in the sender's disclosure threshold has two effects on the DM's nondisclosure belief: first, it increases the likelihood that nondisclosure is due to the sender concealing his signal rather than being uninformed; second, conditional on the sender in fact concealing his signal, it causes the DM to expect a higher signal, i.e., to raise her belief. As the DM's belief conditional on concealment is lower than the prior (since the sender is using a threshold strategy), these two effects work in opposite directions. Moreover, the second effect is stronger than the first if and only if $\hat{s} > \eta(\hat{s}, p, \pi)$. On the other hand, holding the disclosure threshold fixed, an increase in the probability that the sender is informed has an unambiguous effect because it increases the probability that nondisclosure is due to concealed information rather than no information.

Lemma 1. *The nondisclosure belief function, $\eta(\hat{s}, p, \pi)$, has the following properties:*

1. *It is strictly decreasing in $\hat{s}$ when $\hat{s} < \eta(\hat{s}, p, \pi)$ and strictly increasing when $\hat{s} > \eta(\hat{s}, p, \pi)$. Consequently, there is a unique solution to $\hat{s} = \eta(\hat{s}, p, \pi)$, and this solution is interior.*

2. *It is weakly decreasing in p, strictly if $\hat{s} \in (\underline{s}, \overline{s})$.*

All proofs are in the Appendix.

It follows from the above discussion that any equilibrium is fully characterized by the disclosure threshold the sender uses in the equilibrium. If this threshold is interior, the sender must be indifferent between disclosing the threshold signal and concealing it. As sender i's payoff from disclosing any signal s_i is $s_i - c$, we obtain the following equilibrium characterization.

Proposition 1. *Assume there is only one sender i, and this sender is biased upward.*

1. *Any equilibrium has a disclosure threshold $\hat{s}_i^0$ such that: (i) $\hat{s}_i^0$ is interior and $\eta(\hat{s}_i^0, p_i, \pi) = \hat{s}_i^0 - c$; or (ii) $\hat{s}_i^0 = \underline{s}$ and $\pi \leq \underline{s} - c$; or (iii) $\hat{s}_i^0 = \overline{s}$ and $\pi \geq \overline{s} - c$. Conversely, for any $\hat{s}_i^0$ satisfying (i), (ii), or (iii), there is an equilibrium with disclosure threshold $\hat{s}_i^0$.*
2. *If there is either no message cost or there is a concealment cost ($c \leq 0$) then there is a unique equilibrium. Furthermore, if there is no message cost ($c = 0$) the equilibrium disclosure threshold is interior.*
3. *If there is a disclosure cost ($c > 0$) then there can be multiple equilibria.*

Part 1 of the Proposition 1 is straightforward; parts 2 and 3 build on Lemma 1. Multiple equilibria can arise under a disclosure cost because, in the relevant domain (to the right of the fixed point of $\eta(\cdot, p_i, \pi)$), the DM's nondisclosure belief is increasing in the sender's disclosure threshold. In such cases, we will focus on properties of the highest and lowest equilibria in terms of the disclosure threshold. Intuitively, these equilibria respectively correspond to when the sender is least and most informative, and are thus respectively the worst and best equilibria in terms of the DM's welfare. On the other hand, the ranking of those equilibria is reversed for the sender's welfare. To see this, note that because the sender's preferences are linear in the DM's belief and we evaluate welfare at the ex-ante stage, the sender's welfare in an equilibrium with threshold $\hat{s}_i^0$ is $\pi - p_i(1 - F_\pi(\hat{s}_i^0))c$. Thus, when $c > 0$, the sender's welfare is higher when the disclosure threshold is higher: at the ex-ante stage, he cannot affect the DM's belief in expectation and thus would prefer to minimize the probability of incurring the disclosure cost.

When $c = 0$, the sender's belief if he receives the threshold signal is identical to the DM's equilibrium nondisclosure belief. When $c \neq 0$ these two beliefs will differ in any equilibrium: if $c > 0$, the sender's threshold belief, $\hat{s}_i^0$, is higher than the DM's nondisclosure belief, $\eta(\hat{s}_i^0, p_i, \pi)$, and the opposite is true when $c < 0$. This divergence of equilibrium beliefs should the sender withhold information—which, for brevity, we shall refer to as *disagreement*—will prove crucial.[16]

Proposition 1 is stated for an upward biased sender. If the sender is downward biased, he reveals all signals below some threshold, and hence the DM's nondisclosure belief function, $\eta(\hat{s}, p_i, \pi)$, takes the same form as Equation 3 but with $\mathbb{E}_\pi[s|s < \hat{s}]$ replaced by $\mathbb{E}_\pi[s|s > \hat{s}]$. This modified function is single-peaked in $\hat{s}$. The condition for an interior equilibrium is the same as stated in Proposition 1, but the condition for a corner equilibrium becomes: (ii′) $\hat{s}_i^0 = \underline{s}$ and $\pi \geq \underline{s} - c$; or (iii′) $\hat{s}_i^0 = \overline{s}$ and $\pi \leq \overline{s} - c$. Just as with an upward biased sender, equilibrium is unique when $c \leq 0$, but there can be multiple equilibria when $c > 0$. Note that for any $c \neq 0$, the direction of disagreement between the sender and the DM reverses with the direction of sender's bias: for a downward biased sender, $c > 0$ implies that the sender's threshold belief is lower than the DM's nondisclosure belief, and conversely for $c < 0$. Furthermore, when the sender is downward biased, a lower equilibrium threshold corresponds to revealing less information.

The following comparative statics hold with an upward biased sender; the modifications to account for a downward biased sender are straightforward in light of the above discussion.

Proposition 2. *Assume there is only one sender, and this sender is upward biased.*

1. *A higher probability of being informed leads to more disclosure: the highest and lowest equilibrium disclosure thresholds (weakly) decrease.*
2. *An increase in disclosure cost or a reduction in concealment cost leads to less disclosure: the highest and lowest equilibrium disclosure thresholds (weakly) increase.*

The logic for the first part follows from Lemma 1: given any conjectured threshold, a higher p_i leads to a lower nondisclosure belief, which increases the sender's gain from disclosure over nondisclosure of any signal. For the case of $c = 0$, this comparative static has also been noted by other authors, e.g., Jung and Kwon (1988) and Acharya et al. (2011).[17]

The second part of the result is straightforward, as a higher c makes disclosure less attractive. We normalized the bias to be of magnitude of 1, which means our cost c is really cost per unit of bias. A stronger bias is therefore equivalent to a lower message cost. Therefore, an immediate corollary to part 2 of Proposition 2 is that an increase in the magnitude of bias leads to more disclosure when $c > 0$ but less disclosure when $c < 0$.

[16] While the sender knows that there is disagreement, the DM does not. Thus, consistent with Aumann (1976), disagreement is not common knowledge in our setting with a common prior.

[17] A related "prejudicial effect" also arises in Che and Kartik (2009), but in their setting the sender conceals signals that are both sufficiently high and sufficiently low because of non-monotonic preferences.

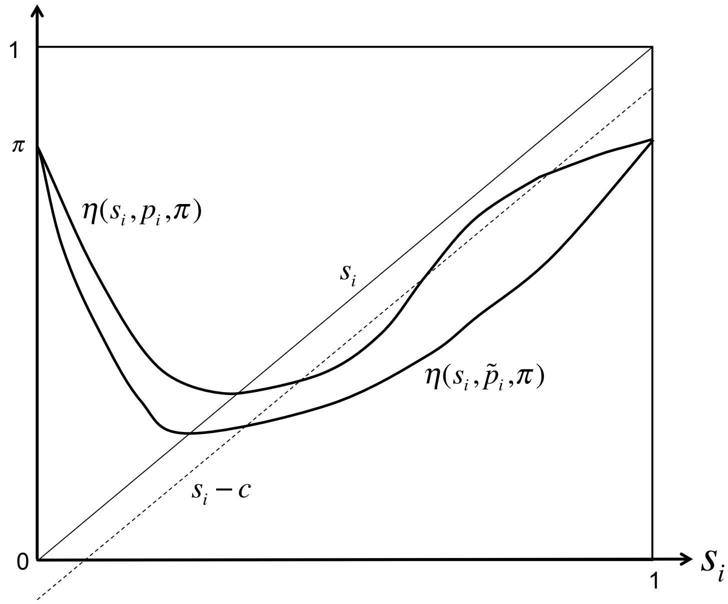

Figure 1. *The single-sender game with an upward biased sender illustrated with parameters $c > 0$ and $\tilde{p}_i > p_i$. Equilibrium disclosure thresholds are given by the intersections of $\eta(s_i, \cdot)$ and $s_i - c$.*

Figure 1 summarizes the results of this section for the case $c > 0$.[18]

Although we postpone a formal argument to Subsection 3.3, it is worth observing now that the comparative statics on disclosure have direct welfare implications. Since the DM always prefers more disclosure, a lower message cost and/or a higher probability of the sender being informed (weakly) increases the DM's welfare in a single-sender setting, subject to an appropriate comparison of equilibria.

3.3. Main results

We are now ready to study the two-sender disclosure game. For concreteness, we will suppose that both senders are upward biased; the modifications needed when one or both senders are downward biased are straightforward.

Lemma 2. *Any equilibrium is a threshold equilibrium, i.e., both senders use threshold strategies.*

Accordingly, we focus our discussion on threshold strategies. A useful simplification afforded by the assumption of conditionally independent signals is that the DM's belief updating is separable in the senders' messages. In other words, we can treat it as though the DM first updates from either sender i's message just as in a single-sender model, and then use this

[18] In the figure, $\eta(\cdot)$ has slope less than 1 when it crosses $s_i - c$ at the highest crossing point. This makes transparent that an increase in p_i leads to a reduction in the highest equilibrium threshold. If the slope of $\eta(\cdot)$ were larger than one at the highest crossing point, then the highest equilibrium threshold would be $\overline{s}$, and a small increase in p_i would not alter this threshold.

updated belief as an interim prior to update again from the other sender j's message without any further attention to i's message. Thus, given any conjectured pair of disclosure thresholds, $(\hat{s}_1, \hat{s}_2)$, there are three relevant nondisclosure beliefs for the DM: if only one sender i discloses his signal s_i while sender j sends message ϕ, the DM's belief is $\eta(\hat{s}_j, p_j, s_i)$; if there is nondisclosure from both senders, the DM's belief is $\eta(\hat{s}_j, p_j, \eta(\hat{s}_i, p_i, \pi))$.[19]

Suppose the DM conjectures that sender i is using a disclosure threshold $\hat{s}_i$. As discussed earlier, if i discloses his signal then his expectation of the DM's belief—viewed as a random variable that depends on j's message—is s_i, no matter what strategy j is using.[20] On the other hand, if i conceals his signal, then he views the DM as updating from j's message based on a prior of $\eta(\hat{s}_i, p_i, \pi)$ that may be different from s_i. Denote j's disclosure threshold by $\hat{s}_j$, and any particular message sent by j as $m_j \in S \cup \{\phi\}$. Let $\beta(\mathcal{I}; q)$ denote the posterior belief given an arbitrary information set $\mathcal{I}$ and prior belief q. Then, sender i's posterior belief about the state given j's message and his own signal can be written as $\beta(m_j; s_i)$. The transformation mapping from Equation 2 implies that i's expected payoff—equivalently, his expectation of the DM's posterior belief—should he conceal his signal is given by

$$\hat{U}(s_i, \eta(\hat{s}_i, p_i, \pi), \hat{s}_j, p_j) := \mathbb{E}_{\hat{s}_j, p_j}[T(\beta(m_j; s_i), \eta(\hat{s}_i, p_i, \pi), s_i)],$$

where $\mathbb{E}_{\hat{s}_j, p_j}$ denotes that the expectation is taken over m_j using the distribution of beliefs that $\hat{s}_j$ and p_j jointly induce in i about m_j (given s_i).

It is useful to study the "best response" of sender i to any disclosure strategy of sender j. More precisely, let $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$ represent the equilibrium disclosure threshold in a (hypothetical) game between sender i and the DM when sender j is conjectured to mechanically adopt disclosure threshold $\hat{s}_j$; we call this sender i's best response. The threshold $\hat{s}_i$ is a best response if and only if it satisfies any one of the following:

$$\begin{cases} U(\hat{s}_i, p_i, \hat{s}_j, p_j) = \hat{s}_i - c & \text{and } \hat{s}_i \in (\underline{s}, \overline{s}); \\ U(\underline{s}, p_i, \hat{s}_j, p_j) \leq \underline{s} - c & \text{and } \hat{s}_i = \underline{s}; \\ U(\overline{s}, p_i, \hat{s}_j, p_j) \geq \overline{s} - c & \text{and } \hat{s}_i = \overline{s}; \end{cases} \tag{4}$$

where $U(s_i, p_i, \hat{s}_j, p_j) := \hat{U}(s_i, \eta(s_i, p_i, \pi), \hat{s}_j, p_j)$.[21] In any equilibrium of the overall game,

[19]If both senders disclose signals, the DM's belief is $(1-\pi)s_1 s_2/((1-\pi)s_1 s_2 + \pi(1-s_1)(1-s_2))$.

[20]The distribution of the DM's beliefs as a function of j's message depends both on j's strategy and the DM's conjecture about j's strategy. As the two must coincide in equilibrium, we bundle them to ease exposition.

[21]Necessity is clear; sufficiency follows from the argument given in the proof of Lemma 2.

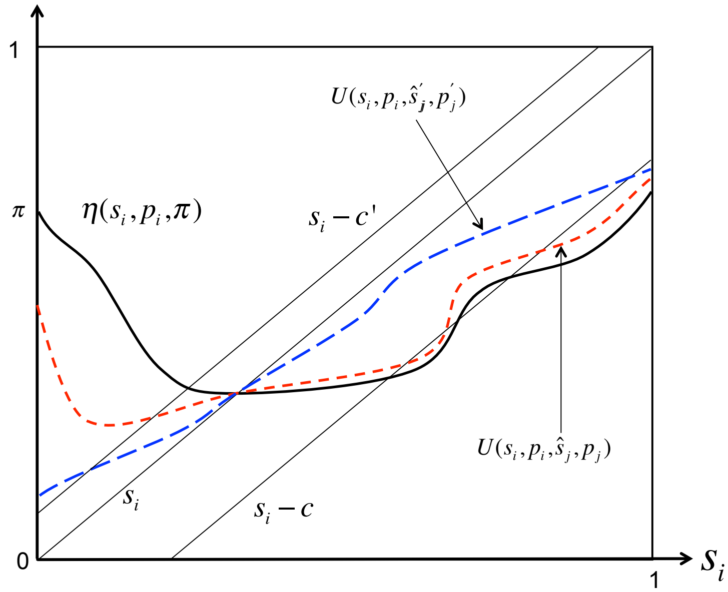

Figure 2. *The "best response" of upward biased sender i to sender j, illustrated with parameters $c > 0 > c'$, $p_j' > p_j$, and $\hat{s}_j' < \hat{s}_j < \overline{s}$, with $\underline{s} = 0$, $\overline{s} = 1$.*

(s_1^*, s_2^*), Equation 4 must hold for each sender i when his opponent uses $\hat{s}_j = s_j^*$.

Lemma 3. *For any $\hat{s}_j$ and p_j,*

$$\begin{aligned} s_i = \eta(s_i, p_i, \pi) &\implies U(s_i, p_i, \hat{s}_j, p_j) = s_i, \\ s_i > \eta(s_i, p_i, \pi) &\implies \eta(s_i, p_i, \pi) \leq U(s_i, p_i, \hat{s}_j, p_j) < s_i, \\ s_i < \eta(s_i, p_i, \pi) &\implies \eta(s_i, p_i, \pi) \geq U(s_i, p_i, \hat{s}_j, p_j) > s_i. \end{aligned}$$

Moreover, both weak inequalities above are strict if and only if $\hat{s}_j < \overline{s}$.

Lemma 3 is a direct consequence of Theorem 1, and it simply reflects that i expects j's information disclosure to, on expectation, bring the DM's posterior belief closer to s_i. Graphically, this is seen in Figure 2 by comparing the red (short dashed) curve that depicts $U(\cdot)$ with the black (solid) curve depicting $\eta(\cdot)$. The former is a rotation of the latter curve around its fixed point toward the diagonal.

It is evident from Figure 2 that j's information disclosure has very different consequences for the best response of i depending on whether there is a cost of disclosure or a cost of concealment. If $c > 0$ (disclosure cost) then the smallest and largest solutions to Equation 4 will be respectively larger than the smallest and largest single-sender disclosure thresholds. If $c < 0$ (concealment cost, depicted as c' in Figure 2) the largest solution will be smaller than the unique single-sender threshold.[22] If $c = 0$, the unique solution is the same as the single-

[22] Although not seen in the figure, there can be multiple solutions to Equation 4 even when $c < 0$.

sender threshold. These contrasting effects are due to the different nature of disagreement induced by the sender. When there is a disclosure cost, the threshold type in any single-sender equilibrium has a *higher* belief than the DM upon nondisclosure, and hence an expected shift of the DM's posterior toward the threshold belief makes concealment more attractive. By contrast, when there is a concealment cost, the threshold type has a *lower* belief than the DM upon nondisclosure, and hence an expected shift of the DM's posterior toward the threshold belief makes concealment less attractive.

Theorem 1 implies that these insights are not restricted to comparisons with the single-sender setting. More generally, the same points hold whenever we compare any $(\hat{s}_j, p_j)$ with $(\hat{s}'_j, p'_j)$ such that $\hat{s}'_j \leq \hat{s}_j$ and $p'_j \geq p_j$. The latter experiment is more informative than the former because sender j is more likely to be informed and discloses more conditional on being informed. More precisely, notice that for any message m'_j under $(\hat{s}'_j, p'_j)$, one can garble it to produce message m_j as follows:

$$m_j = \begin{cases} \phi & \text{if } m'_j = \phi \text{ or } m'_j \in [\hat{s}'_j, \hat{s}_j), \\ m'_j \text{ with prob. } \frac{p_j}{p'_j} \text{ or } \phi \text{ with prob. } 1 - \frac{p_j}{p'_j} & \text{if } m'_j \geq \hat{s}_j. \end{cases}$$

In each state of the world, the distribution of m_j as just constructed is the same as the distribution of sender j's message under $(\hat{s}_j, p_j)$. Thus, the message produced under $(\hat{s}_j, p_j)$ is a garbling of that produced under $(\hat{s}'_j, p'_j)$.

The effect of sender j's message becoming more informative is depicted in Figure 2 as a shift from the red (short dashed) curve to the blue (long dashed) curve. Whether sender i's best response is to disclose less or more of his own information turns on whether $c > 0$ or $c < 0$; the logic is the same as discussed earlier. For any given $\hat{s}_j$, there can be multiple solutions to Equation 4, hence $\hat{s}_i^{BR}(\cdot)$ is a best-response correspondence. We say that sender i's best response increases if the largest and the smallest element of $\hat{s}_i^{BR}(\cdot)$ both increase (weakly); we say that the best response is strictly greater than some $\hat{s}$, written $\hat{s}_i{}^{BR}(\cdot) > \hat{s}$, if the smallest element of $\hat{s}_i^{BR}(\cdot)$ is strictly greater than $\hat{s}$.

Proposition 3. *Assume both senders are upward biased. Any sender i's best-response disclosure threshold $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$ is decreasing in p_i. Furthermore, let $\hat{s}_i^0$ denote the unique (resp. smallest) equilibrium threshold in the single-sender game with i when $c \leq 0$ (resp. $c > 0$).*

1. *(Independence). If $c = 0$, then $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j) = \hat{s}_i^0$ independent of $\hat{s}_j$ and p_j.*
2. *(Strategic complements). If $c < 0$, then (i) $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j) \leq \hat{s}_i^0$, with equality if and only if $\hat{s}_i^0 = \underline{s}$ or $\hat{s}_j = \overline{s}$, and (ii) $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$ increases in $\hat{s}_j$ and decreases in p_j.*

3. *(Strategic substitutes). If $c > 0$, then (i) $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j) \geq \hat{s}_i^0$, with equality if and only if $\hat{s}_i^0 = \overline{s}$ or $\hat{s}_j = \overline{s}$, and (ii) $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$ decreases in $\hat{s}_j$ and increases in p_j.*

Since each sender's best response is monotone, existence of an equilibrium in the two-sender game follows from Tarski's fixed point theorem. When $c < 0$ (concealment cost), the strategic complementarity in disclosure thresholds implies that there is a largest equilibrium, which corresponds to the highest equilibrium thresholds for both senders. Each sender's message in the largest equilibrium is a garbling of his message in any other equilibrium. It follows that the largest equilibrium is the least informative and the *worst* in terms of the DM's welfare. Similarly, the smallest equilibrium is the most informative and the *best* in terms of the DM's welfare. On the other hand, when $c > 0$ (disclosure cost), the two disclosure thresholds are strategic substitutes. There is an *i-maximal* equilibrium that maximizes sender i's threshold and also minimizes sender j's threshold across all equilibria. Likewise, there is a *j-maximal* equilibrium that minimizes sender i's threshold and also maximizes sender j's threshold across all equilibria. These two equilibria are not ranked in terms of (Blackwell) informativeness and in general cannot be welfare ranked for the DM; moreover, neither of these equilibria may correspond to either the best or the worst equilibrium for the DM.[23]

The following result is derived using standard monotone comparative statics arguments.

Proposition 4. *Assume both senders are upward biased. For any $i \in \{1, 2\}$:*

1. *If $c \leq 0$, then an increase in p_i or a decrease in c (a higher concealment cost) weakly lowers the disclosure thresholds of both senders in both the worst and the best equilibria.*
2. *If $c > 0$, then an increase in p_i weakly lowers sender i's disclosure threshold and weakly raises sender j's disclosure threshold in both the i-maximal and the j-maximal equilibria. A decrease in c (a lower disclosure cost) has ambiguous effects on the two senders' equilibrium disclosure thresholds in both the i- and j-maximal equilibria.*

One can view the single-sender game with i as a two-sender game where sender j is never informed, i.e., $p_j = 0$. With this in mind, a comparison of the single-sender game and the two-sender game can be obtained as an immediate corollary to Proposition 3 and Proposition 4.

Corollary 1. *Assume both senders are upward biased and let $\hat{s}_i^0$ denote the unique (resp. smallest) equilibrium threshold in the single-sender game with i when $c \leq 0$ (resp. $c > 0$).*

[23] As explained after Proposition 1, the senders' welfare ranking across equilibria just depends on the probability of disclosure. When $c < 0$, both senders' welfare is lowest in the largest equilibrium and highest in the smallest equilibrium. When $c > 0$, sender i's welfare is highest in the i-maximal equilibrium and lowest in the j-maximal equilibrium.

1. *If $c = 0$, equilibrium in the two-sender game is unique and is equal to $(\hat{s}_1^0, \hat{s}_2^0)$. The DM's welfare is strictly higher in the two-sender game than in a single-sender game with either sender.*
2. *If $c < 0$, every equilibrium in the two-sender game is weakly smaller than $(\hat{s}_1^0, \hat{s}_2^0)$, with equality if and only if $\hat{s}_1^0 = \hat{s}_2^0 = \underline{s}$. The DM's welfare is strictly higher in any equilibrium of the two-sender game than in a single-sender game with either sender.*
3. *If $c > 0$, every equilibrium in the two-sender game is weakly larger than $(\hat{s}_1^0, \hat{s}_2^0)$, with equality if and only if $\hat{s}_1^0 = \hat{s}_2^0 = \overline{s}$. The DM's welfare in the best equilibrium of the two-sender game may be higher or lower than in the best equilibrium of the single-sender game with sender i or sender j alone.*

Part 1 of the corollary follows from part 1 of Proposition 3. When $c = 0$, the best response of each sender is to use the same disclosure threshold as in the single-sender setting, regardless of the other sender's strategy. Since the DM receives two messages instead of just one, and the probability distribution of these messages remain the same as in the single-sender game, she is better off when facing both senders than when facing either sender alone.

Part 2 of Corollary 1 can be obtained by considering the worst equilibrium of the two-sender game. For the case of concealment cost ($c < 0$), let $\mathrm{s}^*(p_i, p_j)$ represent the vector of disclosure thresholds in the worst equilibrium. Proposition 4 implies that $s_i^*(p_i, p_j) \leq s_i^*(p_i, 0) = \hat{s}_i^0$ and $s_j^*(p_i, p_j) \leq s_j^*(0, p_j) = \hat{s}_j^0$. Thus $\mathrm{s}^*(p_i, p_j)$ is weakly smaller than $(\hat{s}_i^0, \hat{s}_j^0)$, and hence every equilibrium is weakly smaller than $(\hat{s}_i^0, \hat{s}_j^0)$. It follows that the DM's welfare is strictly higher than in the unique equilibrium of the single-sender game with either sender. This higher welfare is due to both a direct effect of receiving information from an additional sender, and an indirect effect wherein each sender is now disclosing more than in the single-sender setting.

Finally, for the case of disclosure cost ($c > 0$), let $\mathrm{s}^{i*}(p_i, p_j)$ represent the i-maximal equilibrium and $\mathrm{s}^{j*}(p_i, p_j)$ represent the j-maximal equilibrium. In any equilibrium, i's threshold is at least as large as $s_i^{j*}(p_i, p_j) \geq s_i^{j*}(p_i, 0) = \hat{s}_i^0$, where the inequality is by part 2 of Proposition 4. Analogously, sender j's threshold in any equilibrium is at least as large as $s_j^{i*}(p_i, p_j) \geq s_j^{i*}(0, p_j) = \hat{s}_j^0$. Thus, both senders are (weakly) less informative than in the DM's best equilibrium of the single-sender game. The overall welfare comparison between the two-sender game and the single-sender game is generally ambiguous. While adding a second sender has a direct effect of increasing the DM's information, there is an adverse indirect effect due to strategic substitution that makes the senders disclose less. In the Supplementary Appendix, we provide an explicit example which shows that the net effect can be strictly negative for the DM's welfare, even when the two senders have opposite biases, which is a scenario that is often thought to particularly promote information disclosure. Specifically, a DM with quadratic loss function is better off in the best equilibrium with a single biased sender than with two upward biased senders or two opposite biased senders. Furthermore, given two senders with opposite biases, we show in this example that there is an open and dense set of parameters such that an increase in the disclosure cost $c > 0$ strictly raises the DM's welfare in her best equilibrium.

It is appropriate to compare our welfare results with Bhattacharya and Mukherjee (2013). They study a related model to ours but maintain $c = 0$ and assume perfectly correlated signals. They show that an increase in the probability of a sender being informed can reduce the DM's welfare. However, a necessary condition for this to happen in their model is that at least one sender must have non-monotonic preferences over the DM's posterior, which in turn implies (because senders have single-peaked preferences) that the senders share the same ranking over decisions on a subset of the decision space. In this sense, their result requires that senders are not in "pure conflict," whereas our setup allows senders to have diametrically opposing biases. More broadly, our results on equilibrium behavior and welfare for $c \neq 0$ are orthogonal and complementary to their treatment of non-monotonic preferences.

The assumption that senders' signals are conditionally independent is clearly important for our analytical methodology, as without it we cannot apply Theorem 1. If the signals are conditionally correlated, upon nondisclosure, sender i and the DM disagree not only on the probability assessment of the states, but also on the signal structure of the experiment given by sender j's message. Relaxing the conditional independence assumption to obtain a general analysis is intractable. We illustrate how some of our substantive conclusions would change under a significantly different information structure: the perfectly correlated signal case in the Supplementary Appendix. Another assumption that is important in applying Theorem 1 is that each sender has linear preferences. The Supplementary Appendix discuses how our result under $c = 0$ extends to non-linear preferences and how our results under $c \neq 0$ may or may not hold under non-linear preferences.

3.4. Many senders

Our results readily generalize to any finite number of senders. Suppose in addition to senders i and j, there are K other senders, all of whom simultaneously send messages to the DM. Let $\mathbf{m}$ represent the collection of these K messages. Then, sender i's posterior belief given his own signal s_i, sender j's message m_j, and the K other senders' messages $\mathbf{m}$ is $\beta(m_j, \mathbf{m}; s_i) = \beta(m_j; \beta(\mathbf{m}; s_i))$. The DM's belief given the K senders' messages $\mathbf{m}$ and given nondisclosure by sender i is $\eta(\hat{s}_i, p_i, \beta(\mathbf{m}; \pi))$. Thus, the transformation mapping from Equation 2 and the law of iterated expectations imply that the expected payoff for sender i from concealing his signal is

$$\mathbb{E}\left[\mathbb{E}_{\hat{s}_j,p_j}[T(\beta(m_j;\beta(\mathbf{m};s_i)),\eta(\hat{s}_i,p_i,\beta(\mathbf{m};\pi)),\beta(\mathbf{m};s_i)) \mid \mathbf{m}]\right].$$

The inside conditional expectation (given $\mathbf{m}$) is taken over the distribution of m_j, while the outside expectation is taken over the distribution of $\mathbf{m}$ generated from the equilibrium strategies of the K senders. Given any $\mathbf{m}$, the transformation $T(\cdot)$ in the multi-sender case is the same as that in the two-sender case, with the common prior π replaced by $\beta(\mathbf{m};\pi)$. Since our results hold for any π, the logic of strategic substitution or strategic complementarity continues to apply in the multi-sender case. In particular, when $c < 0$, $\mathbb{E}_{\hat{s}_j,p_j}[T(\cdot) \mid \mathbf{m}]$ increases in $\hat{s}_j$ and decreases in p_j for any $\mathbf{m}$. Consequently, sender i's expected payoff from nondisclosure, $\mathbb{E}\left[\mathbb{E}_{\hat{s}_j,p_j}[T(\cdot) \mid \mathbf{m}]\right]$ also increases in $\hat{s}_j$ and decreases in p_j. Thus disclosure by any two senders are strategic complements. Similarly, in the case of disclosure cost (i.e., $c > 0$), disclosure by any two senders are strategic substitutes.

It follows from these observations that when there is either no message cost or a concealment cost ($c \leq 0$), the DM always benefits from having more senders to supply her with information. When there is disclosure cost ($c > 0$), on the other hand, an increase in the number of senders has ambiguous effects on each sender's disclosure threshold, and can lead to either an increase or decrease in the DM's welfare.

3.5. Sequential reporting

The key insight from our analysis of the multi-sender game with simultaneous disclosure extends to a setting where senders disclose sequentially. For concreteness, consider a two-sender game in which both senders are upward biased but disclosure is sequential: sender 1 reports first and his message m_1 is made public to both the DM and sender 2 before sender 2 submits his report. Sender 2 now effectively faces a single-sender problem where he and the DM share a common prior, say $\beta(m_1;\pi)$, which is a function to be determined in equilibrium. Proposition 2 implies that sender 2 will adopt a disclosure threshold $\hat{s}_2^0$ which depends negatively on p_2.

Consider now the disclosure decision of sender 1 when the DM conjectures that he is using a disclosure threshold $\hat{s}_1$, with corresponding nondisclosure belief $\eta(\hat{s}_1,p_1,\pi)$. If sender 1 discloses his signal s_1, his expectation of the DM's posterior belief is simply s_1. If he chooses nondisclosure, his expectation is $\mathbb{E}_{\hat{s}_2^0,p_2}[T(\beta(m_2;s_1),\eta(\hat{s}_1,p_1,\pi),s_1)]$. Since sender 2 discloses more when he is better informed, a higher p_2 makes the message m_2 more informative, both directly through a higher probability of sender 2 getting a signal and indirectly through a lower disclosure threshold $\hat{s}_2^0$. IVP implies that the DM's belief is expected to move away from $\eta(\hat{s}_1, p_1, \pi)$ toward s_1. The same logic that establishes Proposition 4 therefore gives the following result, whose proof is omitted.

Proposition 5. *Consider sequential disclosure and assume sender 1, the first mover, is upward biased. If $c > 0$ (resp. $c < 0$), a higher p_2 weakly increases (resp. weakly lowers) the equilibrium disclosure threshold of sender 1 in the 1-maximal and 1-minimal equilibria.*

An immediate corollary to Proposition 5 is that, in the case of concealment cost, the first sender discloses more than he does in a single-sender setting. As a result, the DM is always better off in a sequential game than with sender 1 alone. On the other hand, a welfare comparison between the sequential game and the simultaneous move game is generally ambiguous.

We also note that if $c = 0$ the irrelevance result still holds for sender 1: the first sender adopts the same disclosure threshold under sequential reporting as the disclosure threshold in the single-sender problem. The disclosure threshold chosen by the second sender, however, depends on the message sent by sender 1; it may be higher or lower tha sender 2's disclosure threshold were he the only sender.

4. Costly Signaling

An important theme of the multi-sender disclosure game is that message costs induce equilibrium disagreement between sender and receiver. Because a sender needs to predict how the receiver reacts to the message from another sender, whose informativeness is itself an equilibrium object, Theorem 1 is particularly pertinent because it provides a general way of handling expectations about beliefs when people are known to disagree.

Disagreement is a ubiquitous feature in situations involving private information, so we expect that IVP should have wide applicability. In various settings, incentive compatibility constraints require that the informed prefers to reveal his true type than to mimic some other type. If a perfectly separating equilibrium exists, individuals can infer the private information of the informed by observing his behavior, so that disagreement vanishes *in equilibrium*. Nevertheless the fact that an informed *could* induce disagreement by mimicking some other types means that IVP is still relevant for analyzing the incentive constraints in such models. Moreover, IVP is equally applicable regardless of whether the informativeness of an experiment is endogenous or exogenous. It can be used to study the comparative statics of how an exogenous improvement in information affects equilibrium outcomes. In this section, we illustrate such an application in the context of some canonical signaling games.

4.1. The model

To illustrate the applicability of IVP to a wide class of signaling models, we study a setting in which both pooling and separation occur in equilibrium. Specifically, consider a communication game with lying cost, which is a variation of Kartik (2009). A sender and a receiver shares common prior belief about the state $\omega \in \Omega = \{0, 1\}$. A sender has type $t \in [0, 1]$, which is drawn from a distribution $F(\cdot|\omega)$, with corresponding density $f(\cdot|\omega)$. Without loss of generality, we assume that t is the sender's private belief that $\omega = 1$. The sender sends a message $m \in [0, 1]$, from which he suffers a lying cost $c(m, t)$. The receiver forms a belief based on both m and an additional imperfect private signal $s \in [0, 1]$ that is drawn independently conditional on the state from $G(\cdot|\omega)$, with corresponding density $g(\cdot|\omega)$. Alternatively, the signal s can be public but is observed after m is chosen. We assume that the signal s is not perfectly revealing, so that the receiver will not rely solely on it to form her expectations. Denote the receiver's posterior expectation of the state by $\mathbb{E}[\omega|m, s]$. The sender's payoff is linear in the receiver's expectation. His net payoff is given by

$$\mathbb{E}[\omega|m, s] - c(m, t),$$

That is, the sender is upward biased and prefers a higher posterior belief in the receiver. We can allow the sender to have a varying magnitude of bias by interpreting the lying cost as the cost per unit of magnitude of bias. The analysis allows for a downward biased sender as well, a case we omitted to avoid repetition.

We assume that the signals t and s both satisfy the monotone likelihood ratio property. Thus, a higher type of sender has a higher expectation of the state ω. The receiver's belief updating process can be analyzed in two steps: based on the message m from the sender, she forms an interim belief $\pi(m)$ about the state; then she further incorporates the additional signal s into her interim belief to form her posterior belief $\beta(\pi(m), s)$, with

$$\beta(\pi, s) = \frac{\pi g(s|1)}{\pi g(s|1) + (1 - \pi) g(s|0)}.$$

We let $\alpha(m, s)$ denote the sender's payoff from the expectation of the receiver's posterior:

$$\alpha(\pi(m), t) := \mathbb{E}_{s|t}[\beta(\pi(m), s)].$$

The expected payoff of a type-t sender from sending signal m is

$$\alpha(\pi(m), t) - c(m, t)$$

As in Kartik (2009), assume that the lying cost function $c(\cdot)$ is smooth with $\partial c(t, t)/\partial m = 0$ for all t, i.e., the marginal cost of lying when one is telling the truth is zero. The marginal cost of sending a higher message is increasing in m and decreasing in t, i.e., $\partial^2 c(m, t)/\partial m^2 > 0$, and $\partial^2 c(m, t)/\partial m \partial t < 0$. However, the latter assumption is not sufficient to guarantee single-crossing in this model. We impose a stronger assumption.

Assumption 1. *For $m > t$,*

$$\frac{\partial^2 c(m, t)/\partial m \partial t}{\partial c(m, t)/\partial m} \leq -\frac{1}{1 - t}.$$

This assumption can be interpreted as saying that higher type has a sufficiently large marginal cost advantage relative to the marginal cost of message itself.[24] An example of a lying cost function that satisfies these properties is $c(m, t) = (m - t)^2$. The indifference curves of interest are those in the space of actions (messages m) and the receiver's interim beliefs (π). These indifference curves are upward-sloping for $m > t$: a sender is willing to trade off the cost of sending a higher message for a higher interim belief of the receiver. Assumption 1 ensures that the indifference curves are flatter for higher types, meaning that higher types are relatively more willing to send higher messages in order to induce the same increase in interim belief.

Lemma 4. *Assumption 1 implies that $\frac{\partial c(m,t)/\partial m}{\partial \alpha(\pi,t)/\partial \pi}$ decreases in t for $t < m$.*

Although the model is posed as a communication game, we may also interpret it as a variation of the standard signaling model of Spence (1973). Under such an interpretation, a worker possesses some private information t about whether his productivity is high or low, but a potential employer may also observe an imperfect signal s about his productivity (for example, obtained through the job interview process). The wage of a worker depends on the employer's posterior expectation of the worker's productivity, given the schooling level m chosen by the worker and the employer's own signal s. The marginal cost of schooling is decreasing in ability ($\partial^2 c(m, t)/\partial m \partial t < 0$). The constraint that $m \in [0, 1]$ introduces an upper

[24] Similarly, in Daley and Green (2014), a marginal cost advantage of the higher type itself cannot guarantee single crossing of the "belief indifference curves" in the action-interim belief space, which requires the marginal cost advantage to be sufficiently large.

bound on the signaling level m, but this assumption can be relaxed. Further, while we suppose that higher-ability workers intrinsically prefer acquiring more education ($\partial c(t,t)/\partial m = 0$ for all t), the analysis also applies under the more traditional assumption that $\partial c(m,t)/\partial m > 0$ for all m, t. See the discussion after Proposition 6.[25]

4.2. Equilibrium and comparative statics

We focus on *low types separate and high types pool* (hereafter LSHP) equilibria, i.e., we will look for equilibria in which there is a cutoff $\underline{t} \in [0,1]$ such that the sender's (pure) strategy, $\mu(\cdot)$, satisfies:

1. Riley condition: $\mu(0) = 0$.
2. Separation at bottom: $\mu(\cdot)$ is strictly increasing on $[0, \underline{t})$.
3. Pooling at top: $\mu(t) = 1$ for all $t > \underline{t}$.

Because of the assumed upper bound on the sender's strategy, the standard least-costly perfectly separating outcome is not supportable as an equilibrium. See Kartik (2009) for more detailed discussion of this class of equilibria.

In an LSHP equilibrium, the receiver can infer the sender's type if his message m belongs to the interval $[0, \mu(\underline{t})]$. For such m, her interim belief upon observing m is simply $\pi(m) = \mu^{-1}(m)$. If $m = 1$, the receiver infers that $t \geq \underline{t}$; so $\pi(1) = \pi^P(\underline{t}) := \mathbb{E}[\omega | t \geq \underline{t}]$. We note that $\pi^P(\underline{t}) \in (\underline{t}, 1)$.

The equilibrium strategy $\mu(\cdot)$ for $t < \underline{t}$ is pinned down by the differential equation:

$$\frac{\partial c(\mu(t), t)}{\partial m} \mu'(t) = \mathbb{E}_{s|t} \left[\frac{\partial \beta(t,s)}{\partial \pi} \right], \tag{5}$$

with boundary condition $\mu(0) = 0$. Equation 5 is obtained from the upward binding incentive compatibility constraints. The left-hand-side is the marginal cost for type t of mimicking a slightly higher type; the right-hand-side is the marginal benefit, which comes from inducing a higher belief in the receiver. Since $\mu'(t) > 0$ and the right-hand-side is strictly positive, we must have $\mu(t) > t$ for all $t > 0$. Thus, the function $\mu(\cdot)$ will hit 1 at some interior $\bar{t}$, i.e., $\mu(\bar{t}) = 1$ for some $0 < \bar{t} < 1$. The cutoff $\underline{t} \in [0, \bar{t})$ must satisfy:

$$\underline{t} > 0 \implies \mathbb{E}_{s|\underline{t}}[\beta(\pi^P(\underline{t}), s)] - c(1, \underline{t}) = \mathbb{E}_{s|\underline{t}}[\beta(\underline{t}, s)] - c(\mu(\underline{t}), \underline{t}). \tag{6}$$

$$\underline{t} = 0 \implies \mathbb{E}_{s|0}[\beta(\pi^P(0), s)] - c(1, 0) \geq \mathbb{E}_{s|0}[\beta(0, s)] - c(0, 0). \tag{7}$$

[25] The assumption that the type space is continuous is not central either. The model can be reformulated with discrete types with very little change except for the possible adoption of mixed strategy by some type.

Equation 6 requires that, when $\underline{t}$ is interior, type $\underline{t}$ of sender must be indifferent between sending message $m = 1$ (pooling with types above him) and sending message $\mu(\underline{t})$ (revealing his true type). In this case, $\mu(\cdot)$ is discontinuous at $t = \underline{t}$. If Equation 7 holds, then every type prefers to pool and there is a complete pooling equilibrium.

The differential equation (5) with boundary condition $\mu(0) = 0$, and equations (6) and (7) that determine the cutoff type $\underline{t}$, are obtained from local incentive compatibility constraints. They are necessary for an LSHP equilibrium. When a sender's preference satisfies the single-crossing property, these equations are also sufficient for global incentive compatibility.

Lemma 5. *Given Assumption 1, there is a unique LSHP equilibrium.*

In the proof of Lemma 5, we assign the off-equilibrium belief $\pi(m) = \underline{t}$ for $m \in (\mu(\underline{t}), 1)$, and show that no type t has incentive to deviate from $\mu(t)$. It can be shown that this off-equilibrium belief is consistent with the D1 refinement (Cho and Sobel, 1990).

Lemma 5 identifies a unique cutoff type $\underline{t}$ above which pooling occurs. The information the receiver gets about the sender's type is fully characterized by the cutoff: a higher cutoff corresponds to a more informative equilibrium. To study the effect of an improvement in the receiver's exogenous information on the equilibrium cutoff and on equilibrium signaling strategy, we use the following result.

Lemma 6. *If $\tilde{s}$ is drawn from a more informative experiment than s, then for any t,*

1. $\mathbb{E}_{\tilde{s}|t}\left[\frac{\partial \beta(t,\tilde{s})}{\partial t}\right] \leq \mathbb{E}_{s|t}\left[\frac{\partial \beta(t,s)}{\partial t}\right]$.
2. $\mathbb{E}_{\tilde{s}|t}\left[\beta\left(\pi^P(t), \tilde{s}\right)\right] \leq \mathbb{E}_{s|t}\left[\beta\left(\pi^P(t), s\right)\right]$.

For any t and any $\varepsilon > 0$, the belief induced by mimicking type $t + \varepsilon$ likelihood-ratio dominates the belief induced by revealing the true type t. By Theorem 2, it holds that

$$\mathbb{E}_{\tilde{s}|t}[\beta(t+\varepsilon, \tilde{s})] - \mathbb{E}_{\tilde{s}|t}[\beta(t, \tilde{s})] \leq \mathbb{E}_{s|t}[\beta(t+\varepsilon, s)] - \mathbb{E}_{s|t}[\beta(t, s)],$$

where we have used $\mathbb{E}_{\tilde{s}|t}[\beta(t, \tilde{s})] = \mathbb{E}_{s|t}[\beta(t, s)] = t$. Dividing both sides of the inequality by ε and taking the limit as ε goes to 0 proves part 1 of Lemma 6. For part 2, we need to show that the belief induced by learning that the sender's type falls in the interval $[t, 1]$ likelihood-ratio dominates the belief induced by learning that the sender's type is exactly t, which is true because the monotone likelihood ratio property implies

$$\frac{1 - F(t|1)}{1 - F(t|0)} \geq \frac{f(t|1)}{f(t|0)}.$$

Part 2 immediately follows by applying IVP.

Part 1 of Lemma 6 implies that a more informative experiment available to the receiver reduces the right-hand-side of Equation 5. Because each type $t < \underline{t}$ expects a smaller marginal benefit from inducing a higher belief in the receiver, the solution to the differential equation, $\mu(t)$, is point-wise lower. Furthermore, part 2 of Lemma 6 implies that a more informative experiment reduces the left-hand-side of Equation 6, while the right-hand-side is increased because $\mu(t)$ is lower. Thus, type $\underline{t}$ of the sender now strictly prefers to reveal his true type than to pool with higher types. The effect is that the equilibrium cutoff type will increase when the receiver has access to a more informative experiment.

Proposition 6. *Suppose $\tilde{s}$ is drawn from a more informative experiment than s. Then, the equilibrium is (weakly) more informative under $\tilde{s}$ than s, i.e., $\underline{t}_{\tilde{s}} \geq \underline{t}_s$. Furthermore, in the equilibrium under $\tilde{s}$, every type bears a (weakly) lower signaling cost.*

It should be clear that Proposition 6 applies regardless of whether the experiment available to the receiver is endogenous or exogenous. Thus, we can readily adapt our analysis to study a multi-sender signaling game in which each sender gets a conditionally independent signal about the state. Suppose there are two upward biased senders, and each adopts an LSHP strategy.[26] Then, from one sender's perspective, the other sender's message is an MLRP-experiment available to the receiver. Furthermore, the informativeness of this MLRP-experiment is increasing in the other sender's cutoff. Using the same type of argument presented in the analysis of the multi-sender disclosure game, we can establish that the cutoffs adopted by the two senders are strategic complements. If any sender's cost increases in a suitable sense (for example, if each sender i's cost is $k_i c(m,t)$ and $k_i > 0$ increases), or if any sender cares about the receiver's posterior less, then all senders become more informative in equilibrium. Likewise, it is straightforward to show that the smallest or the largest equilibrium cutoffs increase when there are more senders.

The second part of Proposition 6 holds even when there is no pooling at the top. Consider, for example, a canonical Spence (1973) model with unbounded signal space, $m \in [0, \infty)$, and cost function $kC(m,t)$ where $k > 0$, $\partial C/\partial m > 0$, $\partial^2 C/\partial m^2 > 0$, and $\partial^2 C/\partial m \partial t < 0$. In this setting, the (analog of) the familiar least costly separating equilibrium is given by the solution to Equation 5 with the boundary condition $\mu(0) = 0$.[27] Under Assumption 1,

[26] Any combination of biases can be accommodated, so long as we focus on the "right" equilibria, i.e., a downward biased sender deflates rather than inflates his messages.

[27] When there is no exogenous information, standard refinements (e.g., the D1 criterion) select this equilibrium; Cho and Sobel (1990) provide a proof for a discrete type space.

Lemma 6 establishes that the benefit from mimicking a marginally higher type is lower when the receiver has access to better information. As a result, the right-hand-side of Equation 5 is reduced for each type t, and the solution $\mu(t)$ will be pointwise lower. In other words, there will still be full separation, but every type bears a lower signaling cost in equilibrium.

5. Conclusion

We have developed a general theorem of Bayesian updating, the IVP theorem. It says that Anne expects more informative experiments to bring, on average, Bob's posterior belief closer to Anne's prior and away from Bob's prior when the two priors are different. Because a privately informed individual may (either by deviation, or in equilibrium) induce disagreement between his own belief and the uninformed individual's belief, the IVP theorem proves to be useful by providing a general way of analyzing how an individual predicts another individual (who disagrees) to react to informative experiments.

We demonstrate the theorem's applicability by applying it to two games of communication. First is a persuasion game where a sender can choose to either reveal the private information or conceal it. The IVP theorem applies here because the marginal type who is indifferent between revealing and concealing knows that concealing will induce a interim belief in the receiver that is different from his own belief and one can interpret the information provided by competing senders as an experiment available to the receiver. The strategic interaction among competing senders turns out to depend on the direction of disagreement between this sender and the receiver. In this communication game, we have shown that when there is neither a cost of disclosure nor of concealing information ($c = 0$), the presence of other senders is strategically irrelevant for any sender. On the other hand, senders' disclosure are strategic complements under concealment costs ($c < 0$), whereas they are strategic substitutes under disclosure costs ($c > 0$). This implies that under a cost of disclosure, the equilibrium may not be more informative when there are more competing senders.

The second application is a signaling game where a sender can mimic the costly messages sent by other types, knowing that the receiver will have access to exogenous information (possibly from another sender). The IVP theorem applies here because a type can induce a belief in the receiver that is different from its private information by deviating to mimic another type and one can interpret the exogenous information as an experiment available to the receiver. In this communication game, we have shown that better exogenous information leads the sender to expect smaller benefit from mimicking higher types and hence to reduce wasteful signaling.

Appendix

Proof of Theorem 1. This is a special case of Theorem 2 below, because with two states, any pair of priors are likelihood-ratio ordered and any experiment is an MLRP-experiment. $\square$

Proof of Theorem 2. Suppose that $\overline{\boldsymbol{\beta}}_n$ dominates $\overline{\boldsymbol{\beta}}_m$ in the likelihood-ratio order, which is equivalent to $M(\overline{\boldsymbol{\beta}}_n) \geq M(\overline{\boldsymbol{\beta}}_m)$ given that the priors are likelihood-ratio ordered. Let $p(s|\omega)$ represent the density function of s in state ω under experiment $\mathcal{E}$, and let $\tilde{p}(\tilde{s}|\omega)$ represent the density function under experiment $\tilde{\mathcal{E}}$. By the definition of garbling, there exists a non-negative kernal $q(\tilde{s}|s)$ with $\int_{\tilde{s}} q(\tilde{s}|s)\,\mathrm{d}\tilde{s} = 1$ for all s, such that for any state ω,

$$\tilde{p}(\tilde{s}|\omega) = \int_s q(\tilde{s}|s)p(s|\omega)\,\mathrm{d}s.$$

In the following, for $k = m, n$, we let $p_k(s) := \sum_\omega p(s|\omega)\overline{\beta}_k(\omega)$ and $\tilde{p}_k(\tilde{s}) := \sum_\omega \tilde{p}(\tilde{s}|\omega)\overline{\beta}_k(\omega)$. The posterior density function over S conditional on $\tilde{s}$ and with prior belief $\overline{\boldsymbol{\beta}}_k$ is given by

$$\hat{q}_k(s|\tilde{s}) = \frac{q(\tilde{s}|s)p_k(s)}{\tilde{p}_k(\tilde{s})}.$$

Therefore,

$$\frac{\hat{q}_n(s|\tilde{s})}{\hat{q}_m(s|\tilde{s})} = \frac{\tilde{p}_m(\tilde{s})}{\tilde{p}_n(\tilde{s})}\frac{p_n(s)}{p_m(s)}.$$

Because $\mathcal{E}$ is an MLRP-experiment and $\overline{\boldsymbol{\beta}}_n$ dominates $\overline{\boldsymbol{\beta}}_m$, the ratio $p_n(s)/p_m(s)$ increases in s. Therefore, for any $\tilde{s}$, $\hat{q}_n(\cdot|\tilde{s})$ likelihood ratio dominates $\hat{q}_m(\cdot|\tilde{s})$.

Since $M(\boldsymbol{\beta}_n(s))$ increases in s, $\hat{q}_n(\cdot|\tilde{s})$ likelihood-ratio dominates $\hat{q}_m(\cdot|\tilde{s})$ implies

$$\mathbb{E}_n^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(s)\right) \mid \tilde{s}\right] \geq \mathbb{E}_m^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(s)\right) \mid \tilde{s}\right]. \tag{8}$$

For any realization $\tilde{s}$ from experiment $\tilde{\mathcal{E}}$, the left-hand-side of Equation 8 is

$$\mathbb{E}_n^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(s)\right) \mid \tilde{s}\right] = \mathbb{E}_n^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(\tilde{s}, s)\right) \mid \tilde{s}\right] = M\left(\boldsymbol{\beta}_n(\tilde{s})\right), \tag{9}$$

where the first equality follows from the fact that s is sufficient for $\tilde{s}$, and the second follows from the law of iterated expectations. Combining Equation 8 and Equation 9 and take expectation over $\tilde{s}$ (using the prior $\overline{\boldsymbol{\beta}}_m$) gives

$$\begin{aligned}
\mathbb{E}_m^{\tilde{\mathcal{E}}}\left[M\left(\boldsymbol{\beta}_n(\tilde{s})\right)\right] &\geq \mathbb{E}_m^{\tilde{\mathcal{E}}}\left[\mathbb{E}_m^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(s)\right) \mid \tilde{s}\right]\right] \\
&= \int_{\tilde{s}}\left(\int_s M\left(\boldsymbol{\beta}_n(s)\right)\hat{q}_m(s|\tilde{s})\,\mathrm{d}s\right)\tilde{p}_m(\tilde{s})\,\mathrm{d}\tilde{s} \\
&= \int_s M\left(\boldsymbol{\beta}_n(s)\right)\left(\int_{\tilde{s}}\frac{q(\tilde{s}|s)}{\tilde{p}_m(\tilde{s})}\tilde{p}_m(\tilde{s})\,\mathrm{d}\tilde{s}\right)p_m(s)\,\mathrm{d}s \\
&= \mathbb{E}_m^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(s)\right)\right].
\end{aligned}$$

Given that the priors are likelihood-ratio ordered, the above inequality would be reversed if and only if $M(\overline{\boldsymbol{\beta}}_m) \geq M(\overline{\boldsymbol{\beta}}_n)$.

Finally, any experiment $\mathcal{E}$ is a garbling of the perfectly informative experiment $\mathcal{E}^*$ that reveals the true state ω. For experiment $\mathcal{E}^*$, $\mathbb{E}_m^{\mathcal{E}^*}\left[M\left(\boldsymbol{\beta}_n(s)\right)\right] = M\left(\overline{\boldsymbol{\beta}}_m\right)$. Likewise, the perfectly uninformative experiment $\mathcal{E}_*$ that reveals nothing is a garbling of any experiment $\mathcal{E}$. For experiment $\mathcal{E}_*$, $\mathbb{E}_m^{\mathcal{E}_*}\left[M\left(\boldsymbol{\beta}_n(s)\right)\right] = M\left(\overline{\boldsymbol{\beta}}_n\right)$. For $M(\overline{\boldsymbol{\beta}}_n) \geq M(\overline{\boldsymbol{\beta}}_m)$, the experiments $\mathcal{E}^*$ and $\mathcal{E}_*$ respectively provide the lower bound and the upper bound in the second part of the theorem. For $M(\overline{\boldsymbol{\beta}}_n) \leq M(\overline{\boldsymbol{\beta}}_m)$, the experiment $\mathcal{E}^*$ gives the upper bound and $\mathcal{E}_*$ gives the lower bound of $\mathbb{E}_m^{\mathcal{E}}\left[M\left(\boldsymbol{\beta}_n(\cdot)\right)\right]$. □

Proof of Lemma 1. Partially differentiating Equation 3 with respect to the first argument yields

$$\begin{aligned}
\frac{\partial\eta(\hat{s},p,\pi)}{\partial\hat{s}} &= \frac{-p(1-p)f_\pi(\hat{s})}{(1-p+pF_\pi(\hat{s}))^2}(\pi - \mathbb{E}_\pi[s|s<\hat{s}]) + \frac{pF_\pi(\hat{s})}{1-p+pF_\pi(\hat{s})}\frac{f_\pi(\hat{s})}{F_\pi(\hat{s})}(\hat{s} - \mathbb{E}_\pi[s|s<\hat{s}]) \\
&= \frac{pf_\pi(\hat{s})}{1-p+pF_\pi(\hat{s})}\left(\frac{-(1-p)}{1-p+pF_\pi(\hat{s})}(\pi - \mathbb{E}_\pi[s|s<\hat{s}]) + (\hat{s} - \mathbb{E}_\pi[s|s<\hat{s}])\right) \\
&= \frac{pf_\pi(\hat{s})}{1-p+pF_\pi(\hat{s})}(\hat{s} - \eta(\hat{s},p,\pi)).
\end{aligned}$$

Hence, $\operatorname{sign}\left[\partial\eta(\hat{s},p,\pi)/\partial\hat{s}\right] = \operatorname{sign}\left[\hat{s} - \eta(\hat{s},p,\pi)\right]$. Part 1 of the lemma follows from the observation that for any p and π, $\eta(\underline{s},p,\pi) = \eta(\overline{s},p,\pi) = \pi$.

Partially differentiating with respect to the second argument and simplifying yields

$$\frac{\partial\eta(\hat{s},p,\pi)}{\partial p} = \frac{F_\pi(\hat{s})\left(\mathbb{E}_\pi[s|s<\hat{s}] - \pi\right)}{(1-p+pF_\pi(\hat{s}))^2},$$

which proves the second part of the lemma because $\mathbb{E}_\pi[s|s<\hat{s}] < \pi$ if and only if $\hat{s} < \overline{s}$, and $F_\pi(\hat{s}) > 0$ if and only if $\hat{s} > \underline{s}$. □

Proof of Proposition 1. The first part is straightforward and omitted. The second part follows from the fact that for any $p \in (0,1)$ and π, $\eta(\cdot, p, \pi)$ is strictly decreasing on the domain $[0, \hat{s}]$, where $\hat{s}$ is the fixed point of $\eta(\cdot, p, \pi)$, which is interior (Lemma 1). The third part follows because parameters can be chosen such that $c > 0$ and there are multiple solutions in s to $s - c = \eta(s, p_i, \pi)$, as depicted in Figure 1. This can be seen by fixing all parameters except c and p_i and then considering $p_i \to 1$ with a suitable choice of c; details are available on request. $\square$

Proof of Proposition 2. Fix any $p_i > \tilde{p}_i$ and let $\hat{s}_i^0$ and $\tilde{s}_i^0$ denote the corresponding highest equilibrium disclosure thresholds. Suppose by way of contradiction that $\hat{s}_i^0 > \tilde{s}_i^0$. Since $\tilde{s}_i^0$ is the highest equilibrium threshold at $\tilde{p}_i$, it follows that for any $\hat{s} > \tilde{s}_i^0$, $\eta(\hat{s}, \tilde{p}_i, \pi) < \hat{s} - c$. But the fact that $\eta(\hat{s}, p, \pi)$ is weakly decreasing in p (Lemma 1) implies that for any $\hat{s} > \tilde{s}_i^0$, we have $\eta(\hat{s}, p_i, \pi) < \hat{s} - c$, which implies that $\hat{s}_i^0 \leq \tilde{s}_i^0$, a contradiction. A similar argument can be used to establish the result for the lowest equilibrium threshold. We omit the proof of the second part as it follows the same logic as the first part. $\square$

Proof of Lemma 2. Fix any equilibrium and any sender i and sender $j \neq i$. It suffices to show that the difference in the expected payoff for i from disclosing versus concealing is strictly increasing in s_i. Denote the expected payoff from concealing as $\mathbb{E}[\beta_{DM}(m_j, m_i = \phi)]$, where $\beta_{DM}(m_j, m_i = \phi)$ denotes the DM's equilibrium belief following any message m_j and nondisclosure by i, and the expectation is taken over m_j given i's beliefs under s_i. Because m_j is uncorrelated with s_i conditional on the state, and i's belief about the state given s_i is just s_i,

$$\mathbb{E}[\beta_{DM}(m_j, m_i = \phi)] = s_i \mathbb{E}[\beta_{DM}(m_j, m_i = \phi)|\omega = 1] + (1 - s_i)\mathbb{E}[\beta_{DM}(m_j, m_i = \phi)|\omega = 0].$$

The derivative of the right-hand-side of the above equation with respect to s_i is strictly less than one because $\mathbb{E}[\beta_{DM}(m_j, m_i = \phi)|\omega = 1] < 1$, as beliefs lie in $[0,1]$ and m_j cannot perfectly reveal the state. Therefore, the payoff difference, $\mathbb{E}[\beta_{DM}(m_j, m_i = \phi)] - (s_i - c)$ is single-crossing in s_i. $\square$

Proof of Lemma 3. For any p_j, p_i, $\hat{s}_i$, and s_i. $U(s_i, p_i, \hat{s}_j, p_j)$ is i's expectation of the DM's belief (viewed as random variable whose realization depends on j's message) under a prior s_i for i and $\eta(s_i, p_i, \pi)$ for the DM. It follows immediately from Theorem 1 that:

1. $s_i = \eta(s_i, p_i, \pi) \implies U(s_i, p_i, \hat{s}_j, p_j) = s_i$,
2. $s_i > \eta(s_i, p_i, \pi) \implies \eta(s_i, p_i, \pi) \leq U(s_i, p_i, \hat{s}_j, p_j) \leq s_i$,
3. $s_i < \eta(s_i, p_i, \pi) \implies \eta(s_i, p_i, \pi) \geq U(s_i, p_i, \hat{s}_j, p_j) \geq s_i$.

Because $p_j < 1$, the last inequalities in items 2 and 3 above are in fact strict, as an equality in either case requires j's message to be perfectly informative of the state. Finally, the other inequalities in items 2 and 3 are also strict if and only if $\hat{s}_j < \overline{s}$, as j's message is perfectly uninformative if and only if $\hat{s}_j = \overline{s}$. $\square$

Proof of Proposition 3. The transformation $T(\beta_m, \overline{\beta}_n, \overline{\beta}_m)$ defined in Equation 2 is increasing in $\overline{\beta}_n$. Lemma 1 shows that $\eta(\hat{s}_i, p_i, \pi)$ is decreasing in p_i. Hence, sender i's payoff from concealing signal s_i given a correct conjecture by the DM,

$$U(s_i, p_i, \hat{s}_j, p_j) = \mathbb{E}_{\hat{s}_j, p_j}[T(\beta(m_j; s_i), \eta(s_i, p_i, \pi), s_i)],$$

decreases in p_i for any signal s_i, while his payoff from disclosure does not depend on p_i. Following the same argument as in the proof of Proposition 2, the largest and smallest best-response disclosure thresholds must decrease in p_i.

Let $\hat{s}_i^0$ be the smallest equilibrium threshold in the single-sender with i; recall that this is the unique equilibrium threshold if $c \leq 0$.

Consider first $c = 0$. It follows from Part 1 of Lemma 3 that sender i is indifferent between nondisclosure and disclosure when his signal is $\hat{s}_i^0$; hence, $\hat{s}_i^0 \in \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$. Next, we claim there exists no other best-response disclosure threshold. Suppose, to contradiction, that $\hat{s}' > \hat{s}_i^0$ and $\hat{s}' \in \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$. By Lemma 1, $\eta(\hat{s}', p_i, \pi) < \hat{s}'$. Lemma 3 then implies that $U(\hat{s}', p_i, \hat{s}_j, p_j) < \hat{s}'$. Therefore, sender i strictly prefers disclosure to nondisclosure when his signal is $\hat{s}'$, contradicting $\hat{s}' \in \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$. A similar argument establishes that $\hat{s}' < \hat{s}_i^0$ implies $\hat{s}' \notin \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$.

Now consider $c < 0$. If $s_i > \hat{s}_i^0$, then $s_i - c > \eta(s_i, p_i, \pi)$. Since Lemma 3 implies that $U(s_i, p_i, \hat{s}_j, p_j) \leq \max\{s_i, \eta(s_i, p_i, \pi)\}$, it follows that $U(s_i, p_i, \hat{s}_j, p_j) < s_i - c$, and hence $s_i \notin \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$. Furthermore, if $\hat{s}_j < \overline{s}$, then $\hat{s}_i^0 < \eta(\hat{s}_i^0, p_i, \pi)$ and Lemma 3 together imply $U(\hat{s}_i^0, p_i, \hat{s}_j, p_j) > \eta(\hat{s}_i^0, p_i, \pi)$, and hence $\hat{s}_i^0 \in \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$ if and only if $\hat{s}_i^0 = \underline{s}$. Conversely, it is obvious that $\hat{s}_i^0 = \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$ if $\hat{s}_j = \overline{s}$, because sender j's message is completely uninformative. This proves part (i) of the result for $c < 0$. To prove part (ii), we first note from part (i) that if $s_i \in \hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$, $s_i \leq \hat{s}_i^0$ and hence $\eta(s_i, p_i, \pi) > s_i$. Theorem 1 then implies that any garbling of sender j's message decreases $U(s_i, p_i, \hat{s}_j, p_j)$. Thus, an increase in $\hat{s}_j$ or a decrease in p_j—both of which represent a garbling of j's message—lowers sender i's nondisclosure payoff without affecting his disclosure payoff at signal s_i. Following the same argument as in the proof of Proposition 2, the largest and smallest best-response disclosure thresholds must increase.

We omit the proof for $c > 0$ as it follows a symmetric argument to that for $c < 0$; the only point to note is that here the definition of $\hat{s}_i^0$ as the smallest equilibrium threshold in the single-sender game is used to ensure that $s_i < \hat{s}_i^0$ implies $s_i - c < \eta(s_i, p_i, \pi)$. $\square$

Proof of Proposition 4. Consider first the case $c \leq 0$. For each sender i, define the function

$$g_i(\hat{s}_i, \hat{s}_j; p_i, p_j) := \inf\{\hat{s}_i \mid U(\hat{s}_i, p_i, \hat{s}_j, p_j) \leq \hat{s}_i - c\}.$$

That is, $g_i(\cdot)$ gives the smallest element of $\hat{s}_i^{BR}(\hat{s}_j; p_i, p_j)$. Let $g = (g_i, g_j)$, and define

$$\mathbf{s}_*(p_i, p_j) := \inf\{(\hat{s}_i, \hat{s}_j) \mid g(\hat{s}_i, \hat{s}_j; p_i, p_j) \leq (\hat{s}_i, \hat{s}_j)\}.$$

By Proposition 3, $g(\cdot; p_i, p_j)$ is monotone increasing. Hence, $\mathbf{s}_*(p_i, p_j)$ is its smallest fixed point. It remains to be shown that $\mathbf{s}_*(p_i, p_j)$ is the smallest fixed point of the best-response correspondence $(\hat{s}_i^{BR}, \hat{s}_j^{BR})$. Let $\mathbf{s}$ be any other fixed point of the correspondence. Since g is monotone, we have $g(\mathbf{s}_* \wedge \mathbf{s}) \leq g(\mathbf{s}) \leq \mathbf{s}$ and $g(\mathbf{s}_* \wedge \mathbf{s}) \leq g(\mathbf{s}_*) \leq \mathbf{s}_*$. (We follow the notation that $\mathbf{s} \wedge \mathbf{s}' \equiv (\min\{s_1, s_1'\}, \min\{s_2, s_2'\})$.) Thus, $g(\mathbf{s}_* \wedge \mathbf{s}) \leq \mathbf{s}_* \wedge \mathbf{s}$. By the definition of $\mathbf{s}_*$, this in turn implies that $\mathbf{s}_* \leq \mathbf{s}_* \wedge \mathbf{s}$, which is possible only if $\mathbf{s}_* \leq \mathbf{s}$, as required. Thus, this argument establishes that the smallest fixed point of the minimal best response is also the smallest fixed point among all best responses. In other words, $\mathbf{s}_*(p_i, p_j)$ is the *smallest* equilibrium. Proposition 3 establishes that g is decreasing in p_i. It follows from standard monotone comparative statics that the smallest fixed point of g decreases in p_i. It is also straightforward to see from the definition of g_i that g is increasing in c. Hence the best equilibrium increases in c. A parallel argument shows that the worst equilibrium also decreases in p_i and increases in c.

For the case $c > 0$, we keep the sign of $\hat{s}_i$ but flip the sign of $\hat{s}_j$ in the definition of g so that it is monotone in $(\hat{s}_i, -\hat{s}_j)$. The smallest fixed point of g then corresponds to the j-maximal equilibrium. By Proposition 3, a higher p_i decreases sender i's best response but increases sender j's best response. Hence, in a j-maximal equilibrium, $(\hat{s}_i, -\hat{s}_j)$ is decreasing in p_i. The same conclusion holds for an i-maximal equilibrium. $\square$

Proof of Proposition 8. Denote $r \equiv \frac{1-\overline{\beta}_{DM}}{\overline{\beta}_{DM}} \frac{\overline{\beta}_i}{1-\overline{\beta}_i}$, and define $G(\beta, r) := V_i(T(\beta, r)) - V_i(\beta)$, where $T(\beta, r) = \frac{\beta}{\beta + (1-\beta)r}$ is a shorthand for the $T(\beta, \overline{\beta}_{DM}, \overline{\beta}_i)$ transformation defined in Equation 2. When $V_i(\cdot)$ is strictly monotone and $c = 0$, i's best response threshold must be such that $\mathbb{E}[G(\cdot, r)] = 0$ when r is determined by i's threshold type and the DM's nondisclosure belief.

Observe that when $r = 1$, then for any β, $T(\beta, 1) = \beta$ and hence $G(\beta, 1) = 0$. Furthermore, because $T(\beta, r)$ is strictly decreasing in r for all interior β, it follows that for any non-perfectly-informative experiment, $\mathbb{E}[G(\cdot, r)] = 0$ if and only if $r = 1$. Thus, no matter j's disclosure strategy (so long as it is not perfectly informative of the state, which it cannot be since $p_j < 1$), i's best response threshold is such that $r = 1$, i.e., the DM's nondisclosure belief is the same as i's threshold type. But this is the same condition as in the single-sender game. $\square$

Proof of Lemma 4. A sender's indifference curve in the (m, π)-space has a slope equal to

$$\frac{\partial c(m,t)/\partial m}{\partial \alpha(\pi,t)/\partial \pi} = \pi(1-\pi)\frac{\partial c(m,t)/\partial m}{\mathbb{E}_{s|t}[\beta(\pi,s)(1-\beta(\pi,s))|t]}.$$

For $t < m$, the slope is decreasing in sender's type t if

$$\frac{\partial^2 c(m,t)/\partial m \partial t}{\partial c(m,t)/\partial m} < \frac{\mathbb{E}_{s|1}[\beta(\pi,s)(1-\beta(\pi,s))] - \mathbb{E}_{s|0}[\beta(\pi,s)(1-\beta(\pi,s))]}{t\mathbb{E}_{s|1}[\beta(\pi,s)(1-\beta(\pi,s))] + (1-t)\mathbb{E}_{s|0}[\beta(\pi,s)(1-\beta(\pi,s))]}.$$

Since $\mathbb{E}_{s|1}[\beta(\pi,s)(1-\beta(\pi,s))] > 0$,

$$\begin{aligned}
&\frac{\mathbb{E}_{s|1}[\beta(\pi,s)(1-\beta(\pi,s))] - \mathbb{E}_{s|0}[\beta(\pi,s)(1-\beta(\pi,s))]}{t\mathbb{E}_{s|1}[\beta(\pi,s)(1-\beta(\pi,s))] + (1-t)\mathbb{E}_{s|0}[\beta(\pi,s)(1-\beta(\pi,s))]} \\
&> \frac{0 - \mathbb{E}_{s|0}[\beta(\pi,s)(1-\beta(\pi,s))]}{0 + (1-t)\mathbb{E}_{s|0}[\beta(\pi,s)(1-\beta(\pi,s))]} \\
&= -\frac{1}{1-t}.
\end{aligned}$$

So, if Assumption 1 holds, the slope of indifference curve is decreasing in t for $t < m$. $\square$

Proof of Lemma 5. We first establish uniqueness of the cutoff $\underline{t}$. Suppose, by way of contradiction, that $\underline{t} < \underline{t}'$ are both equilibrium cutoffs. From Equation 6 and Equation 7,

$$\mathbb{E}_{s|\underline{t}}[\beta(\pi^P(\underline{t}), s)] - \mathbb{E}_{s|\underline{t}}[\beta(\underline{t}, s)] \geq c(1, \underline{t}) - c(\mu(\underline{t}), \underline{t}).$$

Since $\pi^P(\underline{t}') > \pi^P(\underline{t})$, we obtain

$$\mathbb{E}_{s|\underline{t}}[\beta(\pi^P(\underline{t}'), s)] - \mathbb{E}_{s|\underline{t}}[\beta(\underline{t}, s)] > c(1, \underline{t}) - c(\mu(\underline{t}), \underline{t}).$$

This implies that type $\underline{t}$ would deviate to $m = 1$ from the $\underline{t}'$ equilibrium, a contradiction.

Next, we show that no type has incentive to deviate from $\mu(\cdot)$.

Case (i). $t \leq \underline{t}$. For any $t' \in (t, \underline{t}]$, Equation 5 and Lemma 4 imply

$$\mu'(t') = \frac{\partial \alpha(t', t')/\partial \pi}{\partial c(\mu(t'), t')/\partial m} > \frac{\partial \alpha(t', t)/\partial \pi}{\partial c(\mu(t'), t)/\partial m}.$$

The inequality is due to Lemma 4, which applies because $\mu(t') > t'$, $\mu(t') > t$. The above inequality can be written as

$$\frac{\partial \alpha(t', t)}{\partial \pi} - \frac{\partial c(\mu(t'), t)}{\partial m}\mu'(t') < 0.$$

This inequality is true for any $t' \in (t, \underline{t}]$. For any $\hat{t}$ in the same interval, integrating the inequality over t' from t to $\hat{t}$ gives

$$\alpha(\hat{t}, t) - c(\mu(\hat{t}), t) < \alpha(t, t) - c(\mu(t), t).$$

Thus, type t has no incentive to deviate upward to mimic type $\hat{t} \in (t, \underline{t}]$.

Further, by Equation 6 and Lemma 4, for any $t < \underline{t}$,

$$\alpha(\pi^P(\underline{t}), \underline{t}) - c(1, \underline{t}) = \alpha(\underline{t}, \underline{t}) - c(\mu(\underline{t}), \underline{t}) \implies \alpha(\pi^P(\underline{t}), t) - c(1, t) < \alpha(\underline{t}, t) - c(\mu(\underline{t}), t).$$

Since we have already shown that type $t < \underline{t}$ has no incentive to mimic type $\underline{t}$, type t has no incentive to mimic types higher than $\underline{t}$ by deviating to $m = 1$ either.

Now, let $t' < t$ be such that $t' \geq \mu^{-1}(t)$. Then, an analogous argument establishes that

$$\frac{\partial \alpha(t', t)}{\partial \pi} - \frac{\partial c(\mu(t'), t)}{\partial m}\mu'(t') > 0$$

for any $t' \in [\mu^{-1}(t), t)$. We can apply Lemma 4 for $t' > \mu^{-1}(t)$ because $\mu(t') > t$. The above also holds for $t' = \mu^{-1}(t)$ because $\frac{\partial c(t,t)}{\partial m} = 0$. For any $\hat{t}$ in the same interval, integrating the inequality over t' from $\hat{t}$ to t gives

$$\alpha(t, t) - c(\mu(t), t) > \alpha(\hat{t}, t) - c(\mu(\hat{t}), t).$$

Thus, type $t \leq \underline{t}$ has no incentive to deviate downward to mimic type $\hat{t} \in [\mu^{-1}(t), t)$.

For any $\hat{t} < \mu^{-1}(t)$, we have $\alpha(\hat{t}, t) < \alpha(\mu^{-1}(t), t)$ and $c(\mu(\hat{t}), t) > c(\mu(\mu^{-1}(t)), t) = 0$. Since type t has no incentive to mimic type $\mu^{-1}(t)$, type t has no incentive to mimic type $\hat{t} < \mu^{-1}(t)$ either.

Case (ii). $t \in (\underline{t}, \mu(\underline{t})]$. By Equation 6 and Lemma 4, for any $t > \underline{t}$,

$$\alpha(\pi^P(\underline{t}), \underline{t}) - c(1, \underline{t}) = \alpha(\underline{t}, \underline{t}) - c(\mu(\underline{t}), \underline{t}) \implies \alpha(\pi^P(\underline{t}), t) - c(1, t) > \alpha(\underline{t}, t) - c(\mu(\underline{t}), t).$$

Thus, type t strictly prefers $m = 1$ to $m = \mu(\underline{t})$.

Take any $\hat{t} \in [\mu^{-1}(t), \underline{t})$. In case (i), we have shown that type $\underline{t}$ has no incentive to deviate downward to mimic $\hat{t}$:

$$\alpha(\underline{t}, \underline{t}) - c(\mu(\underline{t}), \underline{t}) > \alpha(\hat{t}, \underline{t}) - c(\mu(\hat{t}), \underline{t}).$$

By Assumption 1, this inequality implies

$$\alpha(\underline{t}, t) - c(\mu(\underline{t}), t) > \alpha(\hat{t}, t) - c(\mu(\hat{t}), t)$$

for $t \in (\underline{t}, \mu(\hat{t})]$. Since t has no incentive to deviate to mimic type $\underline{t}$, type t has no incentive to mimic type $\hat{t} \in [\mu^{-1}(t), \underline{t})$ either.

The same argument as in case (i) shows that type $t \in (\underline{t}, \mu(\underline{t})]$ has no incentive to deviate to any type lower than $\mu^{-1}(t)$.

Case (iii). $t > \mu(\underline{t})$. Let $\hat{\pi} > \underline{t}$ be the interim belief such that type $\underline{t}$ is indifferent between $m = t$ and $m = 1$:

$$\alpha(\pi^P(\underline{t}), \underline{t}) - c(1, \underline{t}) = \alpha(\hat{\pi}, \underline{t}) - c(t, \underline{t}).$$

Since $t > \mu(\underline{t}) > \underline{t}$,

$$\alpha(\pi^P(\underline{t}), t) - c(1, t) > \alpha(\pi^P(\underline{t}), \underline{t}) - c(1, \underline{t})$$

By Assumption 1, this implies

$$\alpha(\pi^P(\underline{t}), t) - c(1, t) > \alpha(\hat{\pi}, t) - c(t, t) > \alpha(\underline{t}, t) - c(\mu(\underline{t}), t).$$

Thus, type t strictly prefers $m = 1$ to $m = \mu(\underline{t})$. The same argument as in case (i) shows that type t has no incentive to deviate to types lower than $\underline{t}$ either.

Finally, it remains to be shown that no type has an incentive to deviate to some $m \in (\mu(\underline{t}), 1)$ which is not used in equilibrium. We assign the off-equilibrium belief $\pi(m) = \underline{t}$ for such m. In cases (i) and (ii), we show that type $t \leq \mu(\underline{t})$ has no incentive to deviate to mimic $\underline{t}$. Since $c(\mu(\underline{t}), t) < c(m, t)$ for off-equilibrium m, this type has no incentive to deviate to m either. In case (iii), we show that type $t > \mu(\underline{t})$ prefers belief-message pair $(\pi^P(\underline{t}), 1)$ to $(\hat{\pi}, t)$, where $\hat{\pi} > \underline{t}$. This type must prefer $(\hat{\pi}, t)$ to $(\underline{t}, m)$ because $c(t, t) \leq c(m, t)$. □

Proof of Proposition 6. We first show that the function $\mu(t)$, defined by Equation 5, is pointwise weakly decreasing in the informativeness of the experiment.

If we show that $\mu(t)$ decreases pointwise when the right-hand-side of Equation 5 decreases for all t, the result then follows Lemma 6. Accordingly, let $\tilde{\mu}(t)$ and $\mu(t)$ be two solutions to Equation 5, with $\tilde{\mu}(0) = \mu(0) = 0$, where $\tilde{\mu}$ solves Equation 5 with a pointwise lower right-hand-side. (These are defined over some respective domains $[0, \tilde{\bar{t}}]$ and $[0, \bar{t}]$. The argument establishes that $\tilde{\bar{t}} \geq \bar{t}$.) For any $t > 0$, if $\tilde{\mu}(t) = \mu(t)$ then $\mu'(t) \geq \tilde{\mu}'(t) > 0$. This implies that at any touching point, μ must touch $\tilde{\mu}$ from below. Consequently, by continuity,

$$\mu(t') \geq \tilde{\mu}(t') \text{ for } t' > 0 \implies \mu(t) \geq \tilde{\mu}(t) \text{ for all } t \geq t'.$$

Now suppose, by way of contradiction, that for some $\hat{t} > 0$, $\tilde{\mu}(\hat{t}) > \mu(\hat{t})$. Then, it must be the case that $\tilde{\mu}(t) > \mu(t)$ for all $t \in (0, \hat{t})$. Since $\partial^2 c/\partial m^2 > 0$ and $\tilde{\mu}$ corresponds to a lower right-hand-side of Equation 5, it follows from Equation 5 that $\tilde{\mu}'(t) < \mu'(t)$ for all $t \in (0, \hat{t})$. But then

$$\tilde{\mu}(\hat{t}) - \mu(\hat{t}) = \int_0^{\hat{t}} [\tilde{\mu}'(t) - \mu'(t)] \, \mathrm{d}t < 0,$$

a contradiction.

Next, we prove that $\underline{t}_{\tilde{s}} \geq \underline{t}_s$. The result is trivial if $\underline{t}_s = 0$, so assume $\underline{t}_s > 0$. Under the more informative experiment $\tilde{s}$ and the corresponding function $\tilde{\mu}$, type $\underline{t}_s$ will (weakly) prefer to separate than pool with the top. There are two reasons, both working in the same direction: (i) separation cost is lower with $\tilde{\mu}$ than with μ (as shown above), i.e., the right-hand-side of Equation 6 increases; and (ii) the benefit of pooling is lower (part 2 of Lemma 6), i.e., the left-hand-side Equation 6 decreases. Since $\bar{t}_{\tilde{s}} \geq \bar{t}_s \geq \underline{t}_s$, and $\bar{t}_{\tilde{s}}$ strictly prefers to pool with $[\bar{t}_{\tilde{s}}, 1]$ rather than separate (same message cost, better inference), by continuity there will be some cutoff $\underline{t}_{\tilde{s}} \geq \underline{t}_s$, and this cutoff is unique by Lemma 5.

The second statement of the proposition follows because all types are sending weakly lower signals in the new equilibrium (and each type's signal is above its true type in both equilibria). □

References

ACHARYA, V. V., P. DEMARZO, AND I. KREMER (2011): "Endogenous Information Flows and the Clustering of Announcements," *American Economic Review*, 101, 2955–79.

ALONSO, R. AND O. CÂMARA (2015): "Bayesian Persuasion with Heterogeneous Priors," Unpublished.

AUMANN, R. J. (1976): "Agreeing to Disagree," *Annals of Statistics*, 4, 1236–1239.

BHATTACHARYA, S. AND A. MUKHERJEE (2013): "Strategic Information Revelation when Experts Compete to Influence," *RAND Journal of Economics*, 522–544.

BLACKWELL, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, vol. 1, 93–102.

——— (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

BOURJADE, S. AND B. JULLIEN (2011): "The Roles of Reputation and Transparency on the Behavior of Biased Experts," *RAND Journal of Economics*, 42, 575–594.

CARLIN, B. I., S. W. DAVIES, AND A. IANNACCONE (2012): "Competition, Comparative Performance, and Market Transparency," *American Economic Journal: Microeconomics*, 4, 202–37.

CHE, Y.-K. AND N. KARTIK (2009): "Opinions as Incentives," *Journal of Political Economy*, 117, 815–860.

CHO, I.-K. AND J. SOBEL (1990): "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50, 381–413.

DAHM, M., P. GONZALES, AND N. PORTEIRO (2009): "Trials, Tricks and Transparency: How Disclosure Rules Affect Clinical Knowledge," *Journal of Health Economics*, 28, 1141–1153.

DALEY, B. AND B. GREEN (2014): "Market Signaling with Grades," *Journal of Economic Theory*, 151, 114–145.

DAUGHETY, A. AND J. REINGANUM (2016): "Information Suppression by Teams and Violations of the Brady Rule," Unpublished.

DEWATRIPONT, M. AND J. TIROLE (1999): "Advocates," *Journal of Political Economy*, 107, 1–39.

DRANOVE, D. AND G. Z. JIN (2010): "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48, 935–63.

DYE, R. A. (1985): "Disclosure of Nonproprietary Information," *Journal of Accounting research*, 23, 123–145.

——— (2017): "Optimal Disclosure When There are Penalties for Nondisclosure," Forthcoming in the *Rand Journal of Economics*.

ELLIOTT, M., B. GOLUB, AND A. KIRILENKO (2014): "How Better Information Can Garble Experts' Advice," *American Economic Review*, 104, 463–468.

FELTOVICH, N., R. HARBAUGH, AND T. TO (2002): "Too Cool for School? Signaling and Countersignaling," *RAND Journal of Economics*, 33, 630–649.

GENTZKOW, M. AND E. KAMENICA (2014): "Costly Persuasion," *American Economic Review*, 104, 457–462.

——— (2015): "Competition in Persuasion," Unpublished.

GROSSMAN, S. J. AND O. D. HART (1980): "Disclosure Laws and Takeover Bids," *Journal of Finance*, 35, 323–334.

HAGENBACH, J., F. KOESSLER, AND E. PEREZ-RICHET (2014): "Certifiable Pre-Play Communication: Full Disclosure," *Econometrica*, 82, 1093–1131.

HIRSCH, A. V. (2015): "Experimentation and Persuasion in Political Organizations," Forthcoming in the *American Political Science Review*.

JACKSON, M. O. AND X. TAN (2013): "Deliberation, Disclosure of Information, and Voting," *Journal of Economic Theory*, 148, 2–30.

JOVANOVIC, B. (1982): "Truthful Disclosure of Information," *Bell Journal of Economics*, 13, 36–44.

JUNG, W.-O. AND Y. K. KWON (1988): "Disclosure When the Market is Unsure of Information Endowment of Managers," *Journal of Accounting Research*, 26, 146–153.

KARTIK, N. (2009): "Strategic Communication with Lying Costs," *Review of Economic Studies*, 76, 1359–1395.

KARTIK, N., F. X. LEE, AND W. SUEN (2017): "Investment in concealable information by biased experts," *The RAND Journal of Economics*, 48, 24–43.

KRISHNA, V. AND J. MORGAN (2001): "A Model Of Expertise," *Quarterly Journal of Economics*, 116, 747–775.

LIPMAN, B. L. AND D. J. SEPPI (1995): "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66, 370–405.

MATTHEWS, S. AND A. POSTLEWAITE (1985): "Quality Testing and Disclosure," *RAND Journal of Economics*, 16, 328–340.

MILGROM, P. (2008): "What the Seller Won't Tell You: Persuasion and Disclosure in Markets," *Journal of Economic Perspectives*, 22, 115–131.

MILGROM, P. AND J. ROBERTS (1986): "Relying on the Information of Interested Parties," *RAND Journal of Economics*, 17, 18–32.

MILGROM, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

OKUNO-FUJIWARA, M., A. POSTLEWAITE, AND K. SUZUMURA (1990): "Strategic Information Revelation," *Review of Economic Studies*, 57, 25–47.

ROTHSCHILD, M. AND J. STIGLITZ (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2, 225–243.

SETHI, R. AND M. YILDIZ (2012): "Public Disagreement," *American Economic Journal: Microeconomics*, 4, 57–95.

SHAVELL, S. (1994): "Acquisition and Disclosure of Information Prior to Sale," *RAND Journal of Economics*, 25, 20–36.

SHIN, H. S. (1994a): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64, 253–264.

——— (1994b): "News Management and the Value of Firms," *RAND Journal of Economics*, 25, 58–71.

——— (1998): "Adversarial and Inquisitorial Procedures in Arbitration," *RAND Journal of Economics*, 29, 378–405.

SOBEL, J. (1989): "An Analysis of Discovery Rules," *Law and Contemporary Problems*, 52, 133–159.

SPENCE, M. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355–374.

THE NEW YORK TIMES (2015): "Honda Fined for Violations of Safety Law," http://nyti.ms/1yFNVUa.

TRUYTS, T. (2015): "Noisy Signaling with Endogenous Information," Unpublished.

VAN DEN STEEN, E. (2010): "Culture Clash: The Costs and Benefits of Homogeneity," *Management Science*, 56, 1718–1738.

VERRECCHIA, R. E. (1983): "Discretionary Disclosure," *Journal of Accounting and Economics*, 5, 179–194.

YILDIZ, M. (2004): "Waiting to Persuade," *The Quarterly Journal of Economics*, 119, 223–248.

A. Supplementary Appendix (Not for Publication)

A.1. Discussion of Theorem 2

This section shows that the conclusion of Theorem 2 can fail with a non-MLRP-experiment (Example 1) or if the priors are not likelihood-ratio ordered (Example 2).

Example 1. Let $\Theta = \{0, 1, 2\}$. Consider two individuals, m and n, with priors $\overline{\boldsymbol{\beta}}_{\boldsymbol{m}}$ and $\overline{\boldsymbol{\beta}}_{\boldsymbol{n}}$ satisfying:

$$\frac{\overline{\beta}_m(0)}{\overline{\beta}_n(0)} > \frac{\overline{\beta}_m(1)}{\overline{\beta}_n(1)} = \frac{1}{2} > \frac{\overline{\beta}_m(2)}{\overline{\beta}_n(2)}.$$

Plainly, $\overline{\boldsymbol{\beta}}_{\boldsymbol{m}} <_{LR} \overline{\boldsymbol{\beta}}_{\boldsymbol{n}}$ and hence $M(\overline{\boldsymbol{\beta}}_{\boldsymbol{m}}) < M(\overline{\boldsymbol{\beta}}_{\boldsymbol{n}})$, where $M(\cdot)$ denotes the expectation operator.

Consider an experiment $\mathcal{E}$ with a binary signal space $\{l, h\}$ and

$$\Pr(l|1) = 1 > \Pr(l|0) = \Pr(l|2) = 0,$$
$$\Pr(h|1) = 0 < \Pr(h|0) = \Pr(h|2) = 1.$$

Thus, signal l reveals state 1 whereas signal h just reveals that the state is not 1, but no relative information about states 0 and 2. Plainly, this not an MLRP-experiment. A direct calculation gives

$$\mathbb{E}_m^{\mathcal{E}}[M(\boldsymbol{\beta}_{\boldsymbol{n}}(\cdot))] = \frac{1}{2}\overline{\beta}_n(1) + 2\left(1 - \frac{1}{2}\overline{\beta}_n(1)\right)\frac{\overline{\beta}_n(2)}{\overline{\beta}_n(0) + \overline{\beta}_n(2)},$$
$$M(\overline{\boldsymbol{\beta}}_{\boldsymbol{n}}) = \overline{\beta}_n(1) + 2\overline{\beta}_n(2).$$

It is possible that $M(\overline{\boldsymbol{\beta}}_{\boldsymbol{m}}) < M(\overline{\boldsymbol{\beta}}_{\boldsymbol{n}}) < \mathbb{E}_m^{\mathcal{E}}[M(\boldsymbol{\beta}_{\boldsymbol{n}}(\cdot))]$; for example, this is the case when $\overline{\beta}_n(0) = 0.3$, $\overline{\beta}_n(1) = 0.2$, and $\overline{\beta}_n(2) = 0.5$ and $\overline{\beta}_m(0) = 0.8$, $\overline{\beta}_n(1) = 0.1$, and $\overline{\beta}_m(2) = 0.1$.

Example 2. Let $\Theta = \{0, 1, 2\}$. Now consider an experiment $\mathcal{E}$ with binary signals $\{l, h\}$ that satisfies MLRP:

$$\frac{\Pr(h|0)}{\Pr(l|0)} < \frac{\Pr(h|1)}{\Pr(l|1)} < \frac{\Pr(h|2)}{\Pr(l|2)}.$$

Let m and n be individuals with priors $\overline{\boldsymbol{\beta}}_{\boldsymbol{m}}$ and $\overline{\boldsymbol{\beta}}_{\boldsymbol{n}}$ such that $\overline{\beta}_m(1) = 1$, $\overline{\beta}_n(0) = \overline{\beta}_n(2) = 1/2$. Plainly, these priors are not likelihood-ratio ordered and $M(\overline{\boldsymbol{\beta}}_{\boldsymbol{m}}) = M(\overline{\boldsymbol{\beta}}_{\boldsymbol{n}}) = 1$. Observe that

$$\mathbb{E}_m^{\mathcal{E}}[M(\boldsymbol{\beta}_{\boldsymbol{n}}(\cdot))] = \Pr(l|1)M(\boldsymbol{\beta}_{\boldsymbol{n}}(l)) + \Pr(h|1)M(\boldsymbol{\beta}_{\boldsymbol{n}}(h)),$$

where by construction $\beta_n(2|l) < 1/2 < \beta_n(2|h)$ and $\beta_n(1|l) = \beta_n(1|h) = 0$. It follows that if $\Pr(h|1) < \Pr(l|1)$, then $\mathbb{E}_m^{\mathcal{E}}[M(\boldsymbol{\beta_n}(\cdot))] < 1 = M(\overline{\boldsymbol{\beta}}_{\boldsymbol{n}}) = M(\overline{\boldsymbol{\beta}}_{\boldsymbol{m}})$.

A.2. Welfare examples under disclosure cost

Example 3. The prior is $\pi = 1/2$. The information structure is parametrized by $\gamma \in (1/2, 1)$ and $\delta \in (0, 1)$. There are four possible signal realizations, with conditional probabilities $\Pr(s|\omega)$ given as follows:

	$\underline{s} = 0$	$s^l = 1 - \gamma$	$s^h = \gamma$	$\overline{s} = 1$
$\omega = 0$	$1 - \delta$	$\gamma\delta$	$(1 - \gamma)\delta$	0
$\omega = 1$	0	$(1 - \gamma)\delta$	$\gamma\delta$	$1 - \delta$

The DM must choose an action $a \in \mathbb{R}$ and her von Neumann-Morgenstern preferences are represented by a quadratic loss function: $u_{DM}(a, \omega) = -(a - \omega)^2$.

In this setting, there is an open and dense set of parameters $(c, p_1, p_2, \gamma, \delta)$ with $c > 0$, such that:

- With a single upward biased sender, the DM's best equilibrium has the sender only disclosing signals s^h and $\overline{s}$ (and symmetrically if the sender is downward biased).
- With two upward biased senders, the DM's best equilibrium has each sender only disclosing signal $\overline{s}$. The DM's welfare in this equilibrium is strictly lower than in the above single-sender equilibrium.
- With opposite biased senders, the DM's best equilibrium has the upward biased sender only disclosing signal $\overline{s}$ and the downward biased sender only disclosing signal $\underline{s}$. The DM's welfare in this equilibrium is strictly lower than in the best equilibrium with either sender alone.

Furthermore, given senders with opposite biases, there is an open and dense set of parameters such that an increase in the disclosure cost $c > 0$ strictly raises the DM's welfare in her best equilibrium.

Note that this example formally violates our assumption of continuous signals. However, one can also perturb the example to make it continuous without affecting the conclusion. We now provides the calculations that verify the claims made in Example 3.

By Lemma 2 of the paper, any equilibrium is a threshold equilibrium, i.e., both senders use threshold strategies. For an upward biased sender i, if $s_i \in \{0, 1 - \gamma, \gamma, 1\}$, then it means respectively that the sender plays a pure strategy of concealing one signal $\underline{s}$ only, concealing two signals up to signal s^l, concealing three signals up to signal s^h, and concealing four signals up to signal $\overline{s}$. We allow mixed strategies. For example, the sender can choose to conceal signals $\underline{s}$ and s^l, and randomize on concealing and disclosing signal s^h with some probability.

A sender's payoffs of disclosing his first, second, third, and fourth signal are, respectively, $0 - c$, $1 - \gamma - c$, $\gamma - c$, and $1 - c$. When $1 - \gamma < c$, a sender at least conceals the first two signals. That is $\hat{s}_i \geq 1 - \gamma$ for an upward biased sender i. For this discrete signal case, let $\eta(\hat{s}_i, p, \pi)$ denote the nondisclosure belief when $\hat{s}_i$ is concealed fully and all signals above $\hat{s}_i$ is fully disclosed.

Parameter case A. $\gamma = 0.7$, $\delta = 0.7$, $p_1 = p_2 = 0.8 \equiv p$, $c = 0.36$

We will use this parameter case to show an example where the best equilibrium under two opposite biased senders is worse for the DM than the best equilibrium under a single sender. In this parameter case, disclosing the third signal gives a sender $\gamma - c = 0.34$ and disclosing the fourth signal gives $1 - c = 0.64$.

Claim A1. With a single upward biased sender, the best equilibrium has the sender only disclosing signals s^h and $\overline{s}$ (and symmetrically if the sender is downward biased).

Proof. Without loss of generality, we prove for a single sender who is upward biased. Since a sender at least conceals two signals, it suffices to show that there is an equilibrium with two signals hidden. That is, given that the sender is believed to conceal two signals, $\hat{s} = 1 - \gamma$, the payoff of concealing any signal is:

$$\eta(1 - \gamma, p, \frac{1}{2}) = \frac{1 - p + p(1 - \gamma)\delta}{2 - p} = 0.306667 < \gamma - c.$$

Therefore, the sender chooses to disclose the third signal. $\square$

Claim A2. With opposite biased senders, the best equilibrium has the upward biased sender only disclosing signal $\overline{s}$ and the downward biased sender only disclosing signal $\underline{s}$. The DM's welfare in this equilibrium is strictly lower than in the best equilibrium with either sender alone.

Proof. Suppose sender 1 is upward biased and sender 2 is downward biased.

Step 1. There is no equilibrium where one sender conceals all four signals and the other sender conceals two or more signals.

First, there does not exist an equilibrium where sender 1 conceals four signals and sender 2 conceals two signals. Suppose there is. Then given that sender 2 conceals two signals and sender 1 is believed to conceal four signals, the payoff to sender 1 of concealing signal $\overline{s}$ is:

$$\delta(1-\gamma)p(1-\gamma) + (1-\delta(1-\gamma)p)\frac{1-p+p(\delta\gamma+1-\delta)}{2-p} = 0.627253 < 1-c.$$

With probability $\delta(1-\gamma)p$, the DM receives s^l from sender 2 and reaches a posterior of $1-\gamma$. With the complementary probability, the DM receives no disclosure from any sender and reaches a posterior of $(1-p+p(\delta\gamma+1-\delta))/(2-p)$. Therefore, sender 1 wants to disclose $\overline{s}$, which forms a contradiction.

By strategic substitution, if sender 2 conceals more than two signals and sender 1 is believed to conceal four signals, sender 1 wants to disclose $\overline{s}$ as well. That is, there is no equilibrium where one sender conceals all four signals and the other sender conceals two or more signals.

Step 2. There is an equilibrium where both senders conceal three signals.

Given that sender 2 conceals three signals and sender 1 is believed to conceal three signals, sender 1's payoff of concealing s^h is:

$$(1-p(1-\gamma)(1-\delta))\frac{1}{2} = 0.464 > \gamma - c$$

If sender 2 discloses a signal, then it must be $\underline{s}$, which gives sender 1 a payoff of 0. When sender 2 conceals a signal, given the symmetric setup, the DM's posterior is $1/2$. Since the payoff of concealing s^h is greater than the payoff of disclosing it, sender 1 conceals s^h.

Step 3. There is no equilibrium where one sender conceals three signals and the other conceals less than three signals.

First we claim that there is no equilibrium where sender 1 conceals two signals and sender 2 conceals three signals. Given that sender 2 conceals three signals and sender 1 is believed to conceal two signals, sender 1's payoff of concealing s^h is:

$$(1-(1-\gamma)(1-\delta)p)\frac{\eta(1-\gamma,p,\frac{1}{2})}{\eta(1-\gamma,p,\frac{1}{2})+(1-\eta(1-\gamma,p,\frac{1}{2}))(1-p+p\delta)} = 0.341395 > \gamma - c.$$

Therefore, sender 1 conceals s^h, which rules out the existence of such an equilibrium. This further implies that there is no equilibrium where both senders conceal two signals.

Now suppose that sender 1 conceals two signals and randomizes on concealing s^h with probability $\lambda \in (0,1)$. Note that the nondisclosure belief is $(1-p+p(1-\gamma)\delta+p\gamma\delta\lambda)/(2-p+p\delta\lambda)$ is increasing in λ for the parameter case we consider. The payoff of concealing s^h is better than when he is believed to conceal only two ($\lambda = 0$). This implies that sender 1 also strictly prefers to conceal s^h.

Step 4. There is no equilibrium where one sender conceals two signals and mixes on concealing the third while the other sender conceals three signals and mixes on concealing the fourth.

Without loss of generality, suppose sender 2 conceals two signals and mixes on the third and sender 1 conceals three signals and mixes on the fourth in equilibrium. Recall that, from Step 2, given sender 2 conceals two signals and sender 1 is believed to conceal four signals, sender 1 wants to disclose the fourth signal. Then, given that sender 2 conceals more than two signals and sender 1 is still believed to conceal four signals, strategic substitution implies that sender 1 wants to disclose the fourth signal as well. Now suppose that sender 1 is believed to conceal three signals and randomize on the fourth signal. Since $\eta(\hat{s}_1, p, 1/2)$ is strictly increasing over $\hat{s}_1 \geq \gamma$ and reaches its maximum at $\eta(1, p, 1/2) = 1/2$, sender 1's payoff of concealing the fourth signal is even worse than if sender 1 is believed to conceal four signals, so sender 1 wants to disclose the fourth signal. Therefore, the above proposed mixed strategy equilibrium does not exist.

Step 5. From Step 3 and Step 4, one learns that the best equilibrium for the DM must be that both senders conceal three signals. In this equilibrium, the expected payoff of the DM is:

$$-(1-p(1-\delta))\frac{1}{4} = -0.19$$

In the best equilibrium under a single sender, i.e., the sender conceals two signals up to s^l, the expected loss of the DM is:

$$-(1-\frac{1}{2}p)\eta(1-\gamma,p,\frac{1}{2})(1-\eta(1-\gamma,p,\frac{1}{2})) + \frac{1}{2}p\delta\gamma(1-\gamma) = -0.186373$$

Therefore, comparing the best equilibrium, the DM is worse off with two opposing senders than with only a single sender. □

Parameter case B. $\gamma = 0.7$, $\delta = 0.7$, $p_1 = p_2 = 0.8 \equiv p$, $c = 0.38$

We will use this parameter case to contrast with parameter case A to show the possibility of an increase in c leading to a higher payoff for the DM under two opposite biased senders. In this parameter case, disclosing the third signal gives a sender $\gamma - c = 0.32$ and disclosing the fourth signal gives $1 - c = 0.62$. Since the difference from Parameter Case A is only in the disclosure cost, the payoff of concealing a particular signal in a particular equilibrium is the same as the corresponding payoff under Parameter Case A.

Claim B. With opposite biased senders, the best equilibrium gives the DM a payoff equal to or higher than an equilibrium where the upward biased sender discloses signal s^h and $\overline{s}$ and the downward biased sender discloses no signal. Contrasting Parameter Case B with Parameter Case A, a higher cost c gives DM a better payoff.

Proof. Step 1. There exists an equilibrium where sender 1 conceals two signals and sender 2 conceals four signals. Given sender 2 conceals all signals, sender 1's payoff of concealing a signal is $\eta(1-\gamma, p, \frac{1}{2}) = 0.306667 < \gamma - c$, so sender 1 conceals only two signals. Given sender 1 conceals two signals, sender 2's payoff of concealing the fourth signal ($\underline{s}$) is $0.627253 > 1-c$, so sender 2 will conceal all of his signals.

There may also exist an equilibrium where sender 2 conceals four signals and sender 1 conceals more than two signals, but this equilibrium is dominated by the above one in terms of the DM's welfare.

Step 2. There is also an equilibrium where both conceal three signals because the payoff of concealing the third signal $0.464 > \gamma - c$.

Step 3. There is no equilibrium where one conceals three signals and the other conceals less than three signals because $0.341395 > \gamma - c$. This further implies that there is no equilibrium where both senders conceal two signals.

Step 4. There may exist a mixed strategy equilibrium where one sender conceals two signals and mixes on concealing the third while the other sender conceals three signals and mixes on concealing the fourth. This equilibrium's payoff may dominate the equilibrium in Step 1.

Step 5. Step 5 of Claim A2 already shows that the equilibrium in Step 1 here is better for the DM than that in Step 2. That is, the best equilibrium for parameter case B is the one in Step 1. Therefore, comparing the best equilibrium of the parameter cases A and B, the DM is better off when c is higher (i.e., under B). □

Parameter case C. $\gamma = 0.8$, $\delta = 0.7$, $p_1 = p_2 = 0.4 \equiv p$, $c = 0.385$

We will use this parameter case to show the possibility that two senders (similar biased or opposite biased) can be worse than a single sender for the DM. In this case, disclosing the third signal gives a sender $\gamma - c = 0.415$ and disclosing the fourth signal gives $1 - c = 0.615$.

Claim C1. With a single upward biased sender, the best equilibrium has the sender only disclosing signals s^h and $\overline{s}$ (and symmetrically if the sender is downward biased).

The proof is the same as that of Claim A1, adapted to the current parameter case.

Claim C2. With opposite biased senders, the best equilibrium has the upward biased sender only disclosing signal $\overline{s}$ and the downward biased sender only disclosing signal $\underline{s}$. The DM's welfare in this equilibrium is strictly lower than in the best equilibrium with either sender alone.

The proof is the same as that of Claim A2, adapted to the current parameter case.

Claim C3. With two upward biased senders, the best equilibrium has each sender only disclosing signal $\overline{s}$. The DM's welfare in this equilibrium is strictly lower than in the above single-sender equilibrium.

Proof. Step 1. There is no equilibrium where one sender conceals all four signals and the other sender conceals two or more signals.

First, there does not exist an equilibrium where sender 1 conceals four signals and sender 2 conceals two signals. Suppose there is. Then given that sender 2 conceals two signals and sender 1 is believed to conceal four signals, the payoff to sender 1 of concealing signal $\overline{s}$ is:

$$\delta\gamma p\gamma + (1-\delta)p + (1-\delta\gamma p - (1-\delta)p)\eta(1-\gamma, p, 1/2) = 0.56816 < 1-c.$$

With probability $\delta\gamma p$, the DM receives s^h from sender 2 and reaches a posterior of γ. With probability $(1-\delta)p$, the DM receives $\overline{s}$ from sender 2 and reaches a posterior of 1. With the remaining probability, the DM receives no disclosure from any sender and reaches a posterior of $\eta(1-\gamma, p, 1/2)$. Therefore, sender 1 wants to disclose $\overline{s}$, which forms a contradiction.

By strategic substitution, if sender 2 conceals more than two signals and sender 1 is believed to conceal four signals, sender 1 wants to disclose $\overline{s}$ as well. That is, there is no equilibrium where one sender conceals all four signals and the other sender conceals two or more signals.

Step 2. There is an equilibrium where both senders conceal three signals.

Given that sender 2 conceals three signals and sender 1 is believed to conceal three signals, sender 1's payoff of concealing s^h is:

$$\gamma(1-\delta)p + (1-\gamma(1-\delta)p)\frac{\eta(\gamma, p, \frac{1}{2})(1-p+p\delta)}{\eta(\gamma, p, \frac{1}{2})(1-p+p\delta) + (1-\eta(\gamma, p, \frac{1}{2}))} = 0.490532 > \gamma - c$$

where $\eta(\gamma, p, 1/2) = (1 - p + p\delta)/(2 - p + p\delta)$. If sender 2 discloses a signal (with probability $\gamma(1-\delta)p$), then it must be $\overline{s}$, which gives sender 1 a payoff of 1. Since the payoff of concealing s^h is greater than the payoff of disclosing it, sender 1 conceals s^h.

Step 3. There is no equilibrium where one sender conceals three signals and the other conceals less than three signals.

First we claim that there is no equilibrium where sender 1 conceals two signals and sender 2 conceals three signals. Given that sender 2 conceals three signals and sender 1 is believed to conceal two signals, sender 1's payoff of concealing s^h is:

$$\gamma(1-\delta)p + (1-\gamma(1-\delta)p)\frac{\eta(1-\gamma, p, \frac{1}{2})(1-p+p\delta)}{\eta(1-\gamma, p, \frac{1}{2})(1-p+p\delta) + (1-\eta(1-\gamma, p, \frac{1}{2}))} = 0.485819 > \gamma - c.$$

Therefore, sender 1 conceals s^h, which rules out the existence of such an equilibrium. This further implies that there is no equilibrium where both senders conceal two signals.

Now suppose that sender 1 conceals two signals and randomizes on concealing s^h with probability $\lambda \in (0,1)$. Note that the nondisclosure belief is $(1-p+p(1-\gamma)\delta+p\gamma\delta\lambda)/(2-p+p\delta\lambda)$ is increasing in λ for the parameter case we consider. The payoff of concealing s^h is better than when he is believed to conceal only two signals ($\lambda = 0$). This implies that sender 1 also strictly prefers to conceal s^h.

Step 4. There are no equilibria where one sender conceals two signals and mixes on concealing the third while the other sender conceals three signals and mixes on concealing the fourth.

Without loss of generality, suppose sender 2 conceals two signals and mixes on the third in equilibrium and sender 1 conceals three signals and mixes on the fourth in equilibrium. Recall that, from Step 2, given sender 2 conceals two signals and sender 1 is believed to conceal four signals, sender 1 wants to disclose the fourth signal. Then, given that sender 2 conceals more than two signals and sender 1 is still believed to conceal four signals, strategic substitution implies that sender 1 wants to disclose the fourth signal as well. Now suppose that sender 1 is believed to conceal three signals and randomize on the fourth signal. Since $\eta(\hat{s}_1, p, 1/2)$ is strictly increasing over $\hat{s}_1 \geq \gamma$ and reaches its maximum at $\eta(1, p, 1/2) = 1/2$, sender 1's payoff of concealing the fourth signal is even worse than if sender 1 is believed to conceal four signals, so sender 1 wants to disclose the fourth signal. Therefore, the above proposed mixed strategy equilibrium does not exist.

Step 5. From Step 3 and Step 4, one learns that the best equilibrium for DM must be that both senders conceal three signals. In this equilibrium, the expected payoff of the DM is:

$$-\left(\frac{1}{2}(1-p(1-\delta))^2+\frac{1}{2}\right)\frac{\eta(\gamma,p,1/2)(1-p+p\delta)}{\eta(\gamma,p,1/2)(1-p+p\delta)+(1-\eta(\gamma,p,1/2))} \times\left(1-\frac{\eta(\gamma,p,1/2)(1-p+p\delta)}{\eta(\gamma,p,1/2)(1-p+p\delta)+(1-\eta(\gamma,p,1/2))}\right)=-0.218215$$

In the best equilibrium under a single sender, i.e., the sender conceals two signals up to s^l, the expected loss of the DM is:

$$-\left(1-\frac{1}{2}p\right)\eta(1-\gamma,p,1/2)(1-\eta(1-\gamma,p,1/2))+\frac{1}{2}p\delta\gamma(1-\gamma)=-0.21592$$

Therefore, comparing the best equilibrium, the DM is worse off with two upward biased senders than with only a single sender. □

A.3. Perfectly correlated signals

Consider the extreme case where informed senders' signals are perfectly correlated and for simplicity, $c=0$. In other words, there is a single signal s drawn from a distribution $F(s|\omega)$, and each sender i is independently either informed of s with probability p_i or remains uninformed. This setting is effectively identical to the "extreme agenda" case of Bhattacharya and Mukherjee (2013).[28] If both senders are biased in the same direction then this model can be mapped to a single-sender problem where the sender is informed with probability $p_i+p_j-p_ip_j$, which is larger than $\max\{p_i,p_j\}$.[29] Proposition 2 then implies that each sender discloses more when there is an additional sender; hence, the DM is always better off with two senders than one.

It is instructive to consider why the irrelevance result no longer holds. For simplicity, suppose both senders are upward biased and symmetric ($p_i=p_j=p$). Let $\hat{s}^0$ denote the common single-sender threshold, so that the nondisclosure belief satisfies $\eta(\hat{s}^0,p,\pi)=\hat{s}^0$. The essential observation is that when sender j is added to the picture, say with the hypothesis that he too discloses all signals weakly above $\hat{s}^0$, type $\hat{s}^0$ of sender i no longer expects the DM's belief to be $\hat{s}^0$ should he conceal his signal, in contrast to the case of conditionally independent signals.

[28] Note that they allow for the senders' utility functions to be non-linear; the ensuing discussion does not depend on linearity either because our single-sender analysis does not require linearity.

[29] Perfect correlation implies that there is only one relevant nondisclosure belief, viz., when both senders don't disclose. So senders who are biased in the same direction must use the same equilibrium disclosure threshold. Given any such threshold, the nondisclosure belief is then computed just as in Equation 3 (assuming the bias is upward), but with $1-p$ replaced by the probability that both senders are uninformed, i.e., $(1-p_i)(1-p_j)$.

Rather, he expects the DM's belief to be strictly lower: if j is informed the DM's belief will be $\hat{s}^0$, and if j is uninformed the DM's belief will be strictly lower because of nondisclosure from two senders rather than just one. This makes type $\hat{s}^0$ of sender i strictly prefer disclosure. From the perspective of Theorem 1, the point is that under conditionally correlated signals, when an informed sender i does not disclose his signal, i and the DM do not agree on the experiment generated by j's message; thus, even if the DM's nondisclosure belief agrees with i's belief (over the state), i's expectation of the DM's posterior belief can be different.

Interestingly, welfare conclusions under perfectly correlated signals are very different when the senders are opposite biased. For simplicity continue to consider $c = 0$. The following proposition shows that when senders are opposite biased, each sender discloses strictly less than he does in his single-sender game.

Proposition 7. *Assume perfectly correlated signals, $c = 0$, and that the two senders are opposite biased. Then, each sender discloses strictly less than what he does in his single-sender game.*

Proof of Proposition 7. We prove it for the upward biased sender; the argument is symmetric for the other sender. Let sender 1 be upward biased and sender 2 be downward biased. Let $\hat{s}_1^0$ denote the single-sender threshold, i.e., $\eta(\hat{s}_1^0) = \hat{s}_1^0$. Write $\eta(\hat{s}_1, \hat{s}_2)$ as the nondisclosure belief (in the event both senders do not disclose) in the two-sender game when the respective thresholds are $\hat{s}_1$ and $\hat{s}_2$. Even though the DM's updating is not separable as in our baseline model, it is clear that $\eta(\hat{s}_1, \hat{s}_2) \geq \eta(\hat{s}_1)$, with equality if and only if $\hat{s}_2 = \underline{s}$. This follows from the simple observation that the nondisclosure event can be viewed as the union of two events: (i) $m_1 = m_2 = \phi$ and $s_2 = \phi$; and (ii) $m_1 = m_2 = \phi$ and $s_2 > \hat{s}_2$. Conditional on the first event, the DM's posterior is just $\eta(\hat{s}_1)$, whereas conditional on the second event, the posterior is larger than $\eta(\hat{s}_1)$ (strictly if and only if $\hat{s}_2 > \underline{s}$). It follows that for any $\hat{s}_2 > \underline{s}$, because $\eta(\hat{s}_1) \geq \hat{s}_1$ for all $\hat{s}_1 \leq s_1^0$, if the DM conjectures thresholds $(\hat{s}_1, \hat{s}_2)$, sender 1 with signal $s_1 = \hat{s}_1$ will strictly prefer nondisclosure to disclosure, where we are using the fact that if sender 2 discloses, he necessarily discloses s_1. Therefore, since sender 2 will use a threshold strictly larger than $\underline{s}$ in any equilibrium, any equilibrium involves sender 1's threshold being strictly larger than $\hat{s}_1^0$. $\square$

Thus, despite the increased availability of information, the overall disclosure of information in the two-sender setting is not more informative than under either single-sender problem. Consequently, for either sender i, there exist preferences of the DM such that she would strictly prefer to face sender i alone rather than the two senders simultaneously. An implication is that the welfare conclusion in Corollary 2 of Bhattacharya and Mukherjee (2013) can be reversed under alternative DM preferences.

In general, for an arbitrary c, an interior equilibrium $(\hat{s}_1, \hat{s}_2)$ requires

$$\Pr[m_2 \neq \phi | s_1 = \hat{s}_1, \hat{s}_2]\hat{s}_1 + \Pr[m_2 = \phi | s_1 = \hat{s}_1, \hat{s}_2]\eta(\hat{s}_1, \hat{s}_2) = \hat{s}_1 - c,$$

or

$$(\hat{s}_1 - \eta(\hat{s}_1, \hat{s}_2)) = \frac{c}{\Pr[m_2 = \phi | s_1 = \hat{s}_1, \hat{s}_2]}.$$

Thus, when $c \geq 0$, using $\eta(\hat{s}_1, \hat{s}_2) > \eta(\hat{s}_1)$ (strictness by interiority of $\hat{s}_2$), it follows that $\hat{s}_1 > \eta(\hat{s}_1)$. Furthermore, for any $c \geq 0$, there is an equilibrium in which $\hat{s}_1$ is weakly larger than the largest single-sender equilibrium. When $c < 0$, the comparison is ambiguous.

A.4. Non-linear utility functions

Another assumption that is important in applying Theorem 1 is that each sender has linear preferences. Suppose, more generally, that a sender i's utility is given by some function, $V_i(\beta_{DM})$. The comparative statics of sender i's disclosure depends on the comparative statics of

$$\mathbb{E}[V_i(T(\beta_i, \overline{\beta}_{DM}, \overline{\beta}_i))] - \mathbb{E}[V_i(\beta_i)] \tag{10}$$

across Blackwell-comparable experiments, where the expectation is taken over the posterior β_i using i's beliefs and $T(\cdot)$ is the transformation of Equation 2. When $V_i(\cdot)$ is linear, one can ignore the second term in Equation 10 as it is constant across experiments and Theorem 1 tells us that the sign of the change in the first term is determined by the sign of $\overline{\beta}_i - \overline{\beta}_{DM}$. Unambiguous comparative statics of Equation 10 cannot be obtained for arbitrary specifications of $V_i(\cdot)$. However, because $T(\beta_i, x, x) = \beta_i$ for any β_i and x, the logic behind our irrelevance result extends generally. In particular:

Proposition 8. *If $c = 0$ and $V_i(\cdot)$ is strictly monotone, then no matter j's disclosure strategy, the best response disclosure threshold for i is the same as when he is a single sender.*

When $c \neq 0$, there are non-linear specifications for $V_i(\cdot)$ under which our themes about strategic complementarity under concealment cost or substitutability under disclosure cost do extend, and there are other specifications which make conclusions ambiguous or even reversed. Below, we show through a family of exponential utility functions how our conclusions are affected by departures from linearity of $V_i(\cdot)$. To this end, consider the more succinct representation: $\mathbb{E}[G(\beta,r)]$, where $G:[0,1]\times\mathbb{R}_+\to\mathbb{R}$ is defined as

$$G(\beta,r):=V_i(T(\beta,r))-V_i(\beta),$$

where $T(\beta,r)=\frac{\beta}{\beta+(1-\beta)r}$. Here, r is a shorthand for $\frac{1-\overline{\beta}_{DM}}{\overline{\beta}_{DM}}\frac{\overline{\beta}_i}{1-\overline{\beta}_i}$; note that $r>1$ if and only if $\overline{\beta}_i>\overline{\beta}_{DM}$ and $r<1$ if and only if $\overline{\beta}_i<\overline{\beta}_{DM}$. Thus, under a disclosure cost $(c>0)$ the relevant case is $r>1$ if the sender is upward biased and $r<1$ if the sender is downward biased; under a concealment cost $(c<0)$ the relevant case is $r<1$ if the sender is upward biased and $r>1$ if the sender is downward biased.

In the following proposition, we say that an upward biased sender i's disclosure is a strategic substitute (resp. complement) to j's if whenever j's message is more Blackwell-informative, i's largest and smallest best response disclosure thresholds increase (resp. decrease).

Proposition 9. *Assume $V_i(\beta)=\gamma\beta^\alpha$, where either $\gamma,\alpha>0$ or $\gamma,\alpha<0$, so that sender i is upward biased. Then, i's disclosure is:*

1. *a strategic substitute to j's under disclosure cost if $0<\alpha\le 1$ and $\gamma>0$.*
2. *a strategic substitute to j's under concealment cost if $\alpha<0$ and $\gamma<0$.*
3. *a strategic complement to j's under disclosure cost if $\alpha\le -1$ and $\gamma<0$.*

Part 1 of Proposition 9 is a generalization of Part 2 of Proposition 3 to some non-linear preferences; on the other hand, Parts 2 and 3 of Proposition 9 show how our main findings of strategic complementary under concealment cost and strategic substitutability under disclosure cost can actually be reversed for other non-linear preferences. Note that one has to be careful with the analog of Proposition 9 for the case of a downward biased sender, because the direction of disagreement between i and the DM reverses. Thus, if $V_i(\beta)=-\gamma\beta^\alpha$, then in each part of Proposition 9 one should replace "disclosure cost" with "concealment cost" and vice-versa.

Given the discussion preceding Proposition 9, and invoking Blackwell's results as in the proof of Theorem 1, Proposition 9 is a straightforward consequence of the following lemma.

Lemma 7. *If $V_i(\beta)=\beta^\alpha$ then $G(\beta,r)$ is:*

1. *convex in β if $0<\alpha\le 1$ and $r>1$;*
2. *concave in β if $\alpha<0$ and $r<1$;*
3. *convex in β if $\alpha<-1$ and $r>1$;*

Proof of Lemma 7. Denoting partial derivatives with subscripts as usual, we obtain that $G_{\beta\beta}(\cdot)$ is equal to

$$V_i''\left(\frac{\beta}{\beta+(1-\beta)r}\right)\left(\frac{r^2}{(\beta+(1-\beta)r)^4}\right)+V_i'\left(\frac{\beta}{\beta+(1-\beta)r}\right)\left(\frac{2r(r-1)}{(\beta+(1-\beta)r)^3}\right)-V_i''(\beta).$$

Plugging in $V_i(\beta)=\beta^\alpha$ and doing some algebra shows that $G_{\beta\beta}(\cdot)$ has the same sign as:

$$\alpha\left[(1-\alpha)+\frac{r(2\beta(r-1)-r(1-\alpha))}{(\beta+(1-\beta)r)^{\alpha+2}}\right]=:H(\beta,\alpha,r).$$

Observe that $H(0,\alpha,r)=\alpha(1-\alpha)(1-r^{-\alpha})$, and hence if $\alpha<1$ and $\alpha\neq 0$ then $\text{sign}[H(0,\alpha,r)]=\text{sign}[r-1]$. Differentiating yields

$$H_\beta(\cdot)=\frac{\alpha(\alpha+1)(r-1)r(\alpha r+2\beta(r-1))}{(\beta+(1-\beta)r)^{\alpha+3}}.$$

We now consider four cases:

1. Suppose $0<\alpha\leq 1$ and $r>1$. Then $H(0,\alpha,r)\geq 0$ and $H_\beta(\cdot)>0$, and hence $H(\beta,\alpha,r)>0$ for all $\beta\in(0,1)$.
2. Suppose $-1\leq\alpha<0$ and $0\leq r<1$. Then $H(0,\alpha,r)<0$ and $H_\beta(\cdot)\leq 0$, and hence $H(\beta,\alpha,r)<0$ for all $\beta\in(0,1)$.
3. Suppose $\alpha<-1$ and $r>1$. Then $H(0,\alpha,r)>0$ and $H(1,\alpha,r)=\alpha(r-1)(\alpha-1+r(1+\alpha))>0$. We will show that $H_\beta(\beta,\alpha,r)=0$ implies $H(\beta,\alpha,r)>0$, which combines with the previous two inequalities to imply that $H(\cdot)>0$. Accordingly, assume $H_\beta(\beta,\alpha,r)=0$, which occurs when $\beta=\frac{\alpha r}{2(1-r)}$, which implies $\alpha\in(-2,-1)$ and $r\geq\frac{2}{2+\alpha}$ (because $\beta\leq 1$ and $\alpha<-1$) . Furthermore,

$$H\left(\frac{\alpha r}{2(1-r)},\alpha,r\right)=\alpha\left[1-\alpha-r^2\left(\frac{2}{r(\alpha+2)}\right)^{\alpha+2}\right].$$

   The derivative of the above expression with respect to r is $\alpha\left(\frac{2}{\alpha+2}\right)^{\alpha+2}r^{-\alpha-1}$, which is strictly positive given $\alpha\in(-2,-1)$ and $r>\frac{2}{\alpha+2}>2$. Moreover, when evaluated with $r=2$, the expression reduces to $1-\alpha-\frac{4}{(\alpha+2)^2}$, which is strictly positive given $\alpha\in(-2,-1)$. Therefore, $H\left(\frac{\alpha r}{2(1-r)},\alpha,r\right)>0$, as was to be shown.
4. Suppose $\alpha<-1$ and $0\leq r<1$. Then $H(0,\alpha,r)<0$ and $H(1,\alpha,r)=\alpha(r-1)(\alpha-1+r(1+\alpha))<0$. As argued in the previous case, $H_\beta(\beta,\alpha,r)=0$ requires $\alpha\in(-2,-1)$ and $r\geq\frac{2}{2+\alpha}>2$, which is not possible given that we have assumed $r<1$. Thus, $H_\beta(\cdot)$

has a constant sign in the relevant domain, which implies that $H(\cdot) < 0$ in the relevant domain. $\square$