Economic Theory (2021) 72:909–953
https://doi.org/10.1007/s00199-020-01299-5

RESEARCH ARTICLE

Robust persuasion of a privately informed receiver

Ju Hu[1] · Xi Weng[2]

Received: 25 April 2019 / Accepted: 11 July 2020 / Published online: 8 August 2020

Abstract
This paper studies robust Bayesian persuasion of a privately informed receiver in a binary environment, where an ambiguity-averse sender with a maxmin expected utility function has limited knowledge about the receiver's private information source. We develop a novel method to solve the sender's information design problem. Our main result shows that the sender's optimal information structure can be found within the class of linear-contingent-payoff information structures. We also fully characterize the sender's optimal linear-contingent-payoff information structure and analyze the impact of ambiguity on the sender's payoff.

Keywords Bayesian persuasion · Ambiguity aversion · Maxmin utility · Private information · Robustness

JEL Classification D81 · D82 · D83

1 Introduction

Imagine a sender (he) who can provide some information to influence the decision making of a rational Bayesian receiver (she) who has a private source of information. The sender has only limited knowledge about what the receiver privately knows and

Hu and Weng would like to acknowledge the financial support of the National Natural Science Foundation of China (Grant Nos. 71803004 and 71973002).

✉ Ju Hu
juhu@nsd.pku.edu.cn

Xi Weng
wengxi125@gsm.pku.edu.cn

1 National School of Development, Peking University, 5 Yiheyuan Road,, Haidian District Beijing 100871, China

2 Guanghua School of Management, Peking University, 5 Yiheyuan Road,, Haidian District Beijing 100871, China

wants to design a persuasion rule that is robust to this uncertainty. Can the sender gain from persuasion? What is the optimal way to persuade?

The above situation is relevant in many economic settings. For example, when a bond rating agency chooses what information to disclose to investors about bond issuers, the agency knows that investors may also have access to private information from other channels, such as newspapers and the Internet. The agency's knowledge about the investors' private information is limited in the sense that the agency knows the possible channels from which the private information is generated, but does not know from which channel a particular investor obtains her information. In another example, when a school chooses what information to disclose on transcripts to prospective employers about the ability of its students, it knows that employers may also obtain private information from the students' extracurricular activities. The school knows the set of all possible extracurricular activities, but does not know which particular extracurricular activity is observed by an employer.

Our model is built on Kamenica and Gentzkow (2011). There is a sender who designs a disclosure rule to convey information about the state of the world, and a receiver who chooses an action that affects her and the sender's payoffs. We focus on a binary environment where there are two states and two actions for the receiver. The receiver wants to match her action with the state, while the sender strictly prefers one action in all states. The receiver takes the sender's preferred action only if she believes the matching state is more likely to occur. Initially, the sender and receiver hold common prior belief $\pi \in (0, 1)$ about this state.

The receiver receives private information from her private source, about which the sender has limited knowledge. We model the receiver's private information source as a distribution of her private belief. It is common knowledge that the receiver's private belief is bounded between $\alpha < \pi < \beta$ and this is the only knowledge that the sender has about the receiver's private information source. Thus, the sender thinks that every private belief distribution over $[\alpha, \beta]$ whose mean is π is possible. For instance, a special case is $\alpha = 0$ and $\beta = 1$. This is the situation where the sender has no knowledge at all about the receiver's private belief distribution.

A sender's information disclosure rule is an information structure. We investigate how an ambiguity-averse sender with maxmin expected utility optimally designs his robust information structure. The maxmin expected utility criterion, which is perhaps the most commonly adopted model in previous studies involving ambiguity and robustness concerns, simply requires that the sender evaluates each information structure in terms of the worst-case expected payoff across the receiver's possible private belief distributions.[1] The information structure that maximizes this worst-case expected payoff is the sender's optimal information structure.

Each of the sender's information structures defines his expected payoff as a function of the receiver's private belief over $[\alpha, \beta]$. We call this function the sender's *contingent payoff function* from this information structure. Because the sender thinks that the receiver's private belief distribution can be any distribution over $[\alpha, \beta]$ with mean π, the sender's *worst-case payoff function* from each information structure is then the largest convex function below his contingent payoff function over $[\alpha, \beta]$. The

[1] See Gilboa and Schmeidler (1989) for an axiomatic representation of this preference.

sender's worst-case expected payoff from this information structure is just the value of this worst-case payoff function at the prior π. This observation reflects the fact that it is the worst "average performance"of each information structure, rather than its performance at a single private belief, that the sender cares about in the face of ambiguity.

Our main result shows that the sender's optimal information structure is a *linear-contingent-payoff information structure*. The contingent payoff function of such an information structure takes the form of $\max\{0, \ell\}$ over $[\alpha, \beta]$ for some linear function ℓ.[2] In establishing this main result, we develop a novel method for solving the sender's robust information design problem. This method combines convexification and linear approximation. Because the sender's contingent payoff function under a general information structure can be quite arbitrary, its convexification—the worst-case payoff function—is almost impossible to analyze directly. By approximating each worst-case payoff function by linear functions, we effectively transform the comparison of worst-case payoff functions into a much more tractable problem: comparing linear functions. It turns out that the comparison of these linear functions can be accomplished by estimating differential inequalities.

We further characterize the sender's optimal linear-contingent-payoff information structures for different values of α, β and prior $\pi \in [\alpha, \beta]$. For all but one cut-off prior $\hat{\pi} \in (\alpha, \beta)$, the sender has a unique optimal linear-contingent-payoff information structure. More interestingly, the optimal information structures for different priors reflect a fundamental trade-off in the sender's optimization problem. For prior $\pi < \hat{\pi}$, the optimal information structure is the one that maximizes the sender's contingent payoff at private belief α. This is because the receiver's private beliefs are more likely to be close to α for low prior π, and this particular information structure just guarantees himself high contingent payoffs for all private beliefs close to α. In contrast, for prior $\pi > \hat{\pi}$, the optimal information structure delivers low (possibly the lowest) contingent payoffs for private beliefs close to α, but higher contingent payoffs for private beliefs close to β. This is because the receiver's private beliefs are more likely to be close to β for high prior π. Thus, the sender is willing to sacrifice contingent payoffs at low private beliefs in exchange for higher contingent payoffs at high private beliefs.

Even in this simple environment with binary states and actions, we find that some of the optimal information structures involve infinitely many signals. These infinitely many signals are specially distributed to guarantee the linear structure of its contingent payoff function. This is an example where the "revelation principle" (Proposition 1 in Kamenica and Gentzkow 2011) fails. In the current setting, restricting attention to recommendation systems where the signals are interpreted as recommending actions does entail loss of generality. This is because the receiver with different private beliefs will have different posteriors for the same recommendation.

Based on the characterization, we also analyze how the change of ambiguity affects the sender's welfare. Specifically, we ask whether more ambiguity always make the sender strictly worse off. Interestingly, it depends on whether the ambiguity is too biased or not. If α is far away from π compared to β from π, then a decrease in α will

[2] Here, zero is the lowest possible payoff to the sender.

not harm the sender. Similarly, if β is far away from π compared to α from π, then an increase in β will not make the sender strictly worse off either.

1.1 Related literature

Our Bayesian persuasion model is a variation of that of Kamenica and Gentzkow (2011), with the new ingredient that the receiver is privately informed and the sender has only limited knowledge about the receiver's private information source. We study how a sender optimally reveals information that is robust to the receiver's private information. Bayesian persuasion of a privately informed receiver has been studied in Rayo and Segal (2010), Kamenica and Gentzkow (2011), Guo and Shmaya (2019), Kolotilin et al. (2017) and Kolotilin (2018). These papers all assume that the distribution of the receiver's private information is common knowledge, as in the usual mechanism design literature, but we consider the environment in which the sender thinks many distributions are possible. While Rayo and Segal (2010), Guo and Shmaya (2019), Kolotilin et al. (2017) and Kolotilin (2018) model the receiver's private information as her private preference and the latter three papers also consider private persuasion (so named by Kolotilin et al. 2017), our model is closest to Kamenica and Gentzkow (2011), Section VI.A. We model the receiver's private information as her private belief, and focus on public persuasion in which the sender designs a single information disclosure rule for all receiver types. Because the sender in our model is uncertain about the receiver's private information source, he cannot simply form an expectation of the receiver's private beliefs by "integrating over the receiver's private signal," as suggested in Kamenica and Gentzkow (2011), Section VI.A. Consequently, the standard *concavification* approach of Kamenica and Gentzkow (2011) does not apply in our model. Instead, our analysis relies on the opposite of concavification, i.e., *convexification*, to derive our characterizations (see Lemma 1).

Because our model can be equivalently interpreted as a zero-sum game between the sender and nature, our paper is also related to Gentzkow and Kamenica (2017a) who study multiple sender persuasion problems with rich signal spaces. The major difference between their paper and ours is that nature's signal space is restricted by the sender's ambiguity and we do not allow arbitrary correlation between the sender and nature's signals. It is only when the sender has full ambiguity that our model degenerates to that in Gentzkow and Kamenica (2017a) because, in this case, nature's signal space is unrestricted. In other cases, our analysis can be considered as an extension of theirs to one-sided restricted signal space.[3]

More recently, Kosterina (2018) studies a robust persuasion problem that is similar to ours. In her model, the sender and receiver have different priors. The sender does not precisely know the receiver's prior, but thinks that the receiver's prior puts at least some exogenous given mass on states higher than or equal to a certain cut-off. One important difference between her paper and ours is that the receiver and the sender hold a common prior about the states in our model. At the interim stage, the sender

[3] A few recent papers also consider the case of multiple information designers in different environments. For example, Albrecht (2017), Au and Kawai (2020, 2019), Gentzkow and Kamenica (2017b), Li and Norman (2018) and Koessler et al. (2018).

and the receiver may have different beliefs just because the receiver privately received additional information. The sender rationally takes this into account, as Kamenica and Gentzkow (2011) Section VI.A does. As a result, the sender's optimal information structure is independent of the sender's prior in Kosterina (2018), whereas the sender's optimal information structure in our model critically depends on the common prior.

Finally, our paper is also related to the growing literature on robust mechanism design under ambiguity aversion. The literature has studied various contexts, such as auction design, bilateral trade, exchange economy, monopoly pricing, and moral hazard.[4] To the best of our knowledge, our paper is the first to investigate robust Bayesian persuasion of a privately informed receiver.[5] Moreover, in the previous literature, the principal is completely uncertain about the distributions or knows only some moments of the distributions (e.g., Carrasco et al. 2018). In our setup, the principal (the sender) not only knows the mean of the distributions, but also may have further knowledge about the support of the distributions. We believe that the general method developed in this paper can also be applied to study the robust Bayesian persuasion of a privately informed receiver in other frameworks, such as those in Rayo and Segal (2010) and Kolotilin et al. (2017), and other robust mechanism design issues in similar contexts.

2 Model

2.1 Basic setup

Suppose that there are two states of the world, $\omega = 0$ and $\omega = 1$. There are two players: a sender and a receiver. At the beginning of the game, the sender and receiver share a common prior $\pi \in (0, 1)$ on state $\omega = 1$. The sender designs information and the receiver chooses one of the two actions $a = 0$ and $a = 1$. The receiver's ex post payoff function is given by $u(a, \omega) = 1$ if $a = \omega$ and $u(a, \omega) = 0$ if $a \neq \omega$. That is, the receiver earns an ex post payoff 1 if her action matches the underlying state; otherwise she gets 0. Because there are only two states and two actions, assuming that the receiver's ex post payoff is 0 when she chooses the wrong action entails no loss of generality. The assumption that the receiver's payoffs are the same in both states when she chooses the correct action is made purely for ease of exposition. The method

[4] An incomplete list of these studies includes Bergemann and Schlag (2008, 2011) and Carrasco et al. (2018), who study monopoly pricing when the monopolist only has limited knowledge about the distribution of the buyer's valuation; Garrett (2014) analyzes a model of cost-based procurement where the seller is uncertain about the agent's effort cost function; Carroll (2015) considers a principal-agent model in which the principal is uncertain about what the agent can and cannot do; Bose et al. (2006), Bodoh-Creed (2012) and Bose and Renou (2014) investigate auction design problems in which each bidder is uncertain about the other bidders' valuation distributions; Wolitzky (2016) studies efficiency in a bilateral trade model in which the seller and buyer know only the mean of each other's valuations; and de Castro et al. (2011, 2017a, b) and de Castro and Yannelis (2018) study implementation problem in asymmetric information economy where the agents have multiple priors about the states. More recently, Bergemann et al. (2016) and Du (2018) study a robust common value auction design in which the seller is uncertain about the bidders' information structures.

[5] Following Bose and Renou (2014), Beauchêne et al. (2019) consider ambiguous persuasion where the sender can send a signal with multiple likelihood distributions. But our model does not allow for this possibility.

we develop in this paper can be extended to the case where the receiver gets different payoffs in different states when choosing the correct action. The receiver's optimal action when her *posterior belief* about state $\omega = 1$ is q is thus given by $a[q] = 0$ if $q < 1/2$ and $a[q] = 1$ if $q \geq 1/2$. The sender, by contrast, always prefers the receiver taking action $a = 1$, regardless of the underlying states. More specifically, the sender's payoff function is $v(a) = 0$ if $a = 0$ and $v(a) = 1/2$ if $a = 1$. The value $1/2$ is inessential, as our analysis can be applied to any strictly positive $v(1)$. We take $v(1) = 1/2$ rather than other values just to simplify the exposition.

2.2 Modeling the sender's ambiguity

The receiver receives a private signal about the underlying state from her private information source. Upon observing such a signal, the receiver then updates her belief from the common prior π. From an ex ante point of view, this information source leads to a distribution μ of the receiver's *private belief*. If the sender knew the receiver's information source, he would correctly expect μ as the distribution of the receiver's private belief, as discussed in Section VI.A in Kamenica and Gentzkow (2011).

However, we assume that when designing information, the sender neither observes the receiver's private signal nor is aware of her private information source. Instead, the only knowledge that the sender has is that the receiver's private belief is contained in a certain range $[\alpha, \beta]$ where $0 \leq \alpha \leq \pi \leq \beta \leq 1$ no matter what private information source the receiver actually has and what signal is realized. To avoid the trivial case, we assume $\alpha < 1/2$ throughout the paper. The range $[\alpha, \beta]$ then represents the ambiguity faced by the sender. An equivalent interpretation of this ambiguity is that the sender is sure that the receiver's information source has a certain bound on the likelihood ratios, although he has no idea what precisely it is. Moreover, we assume that the sender's knowledge is correct in the sense that all the receiver's possible private beliefs are indeed contained in $[\alpha, \beta]$. This assumption rules out situations where the sender completely misspecifies the receiver's private information structure.[6]

As is a common practice in the Bayesian persuasion literature due to Proposition 1 in Kamenica and Gentzkow (2011), any information source that leads to all the receiver's private beliefs being contained in the range $[\alpha, \beta]$ can be identified as the set of distributions of the receiver's private belief whose mean is the common prior π and whose support is contained in $[\alpha, \beta]$. Formally, let

$$M(\pi; [\alpha, \beta]) \equiv \left\{ \text{probability distribution } \mu \text{ over } [\alpha, \beta] \,\middle|\, \int_{[\alpha, \beta]} p\mu(\mathrm{d}p) = \pi \right\}$$

be the set of all such private belief distributions. The sender believes that the distribution of the receiver's private belief is some $\mu \in M(\pi; [\alpha, \beta])$.

[6] See, for example, Esponda and Pouzo (2016a, b) for a recent discussion of solution concepts under model misspecification.

2.3 Sender's information structure

Aside from the receiver's private information, the sender can design an experiment to supply supplemental information to the receiver. We assume that the receiver's private information and the sender's information are conditionally (on states) independent (as in Kamenica and Gentzkow 2011; Bergemann et al. 2018).

Deviating from the way we model the receiver's private information, we directly model the sender's information design problem as choosing an information structure. Compared to the approach of choosing a posterior belief distribution with mean π, this direct approach has two advantages. First, the receiver, after observing the sender's signal, updates her belief from her *private belief* that may be different from the *common prior* π. If we model the sender's information design problem as choosing a posterior belief distribution, the receiver's true posterior belief distribution will be different from the one designed by the sender. It is not convenient to transform the sender's design into the receiver's true distribution.[7] Second, information structures only involve conditional distributions of signals, and thus are "prior free." This feature makes it easier to evaluate the performance of a given information structure for different priors. As a result, it enables us to solve the sender's information design problem simultaneously for all $\pi \in (\alpha, \beta)$, as we will see in the next section. It also facilitates the comparison of the optimal information structures for different priors.

Formally, an information structure consists of a signal space and two probability measures over the signal space governing the conditional distribution of signals in each state $\omega \in \{0, 1\}$. Theorem 3 in Blackwell (1951) showed that every information structure has a canonical representation.[8] The signal space of such a canonical representation is simply the interval $[0, 1]$. The two conditional distributions of signals are identified by a c.d.f. F over $[0, 1]$ with mean $1/2$. More specifically, the signal distribution in state $\omega = 0$ is given by

$$F_0(s) = \int_{[0,s]} 2\tilde{s} \mathrm{d}F(\tilde{s}), \ \ \forall s \in [0, 1], \tag{1}$$

and that in state $\omega = 1$ is

$$F_1(s) = \int_{[0,s]} 2(1 - \tilde{s}) \mathrm{d}F(\tilde{s}), \ \ \forall s \in [0, 1]. \tag{2}$$

Hence, from the sender's point of view, choosing an information structure, i.e., a pair of signal distributions, is equivalent to choosing such a single c.d.f. F. With slight abuse of terminology, we refer to such an F as the sender's information structure. Let $\mathcal{F} \equiv \left\{\text{c.d.f. } F \text{ over } [0, 1] \,\middle|\, \int_{[0,1]} s \mathrm{d}F(s) = 1/2\right\}$ be the set of all the sender's information structures.

[7] Proposition 1 in Alonso and Câmara (2016) derives a formula for this transformation.

[8] It was called a *standard experiment* in Blackwell (1951).

Given any information structure $F \in \mathcal{F}$, if the receiver's private belief about state $\omega = 1$ is p, her posterior belief after observing signal $s \in [0, 1]$ is[9,10]

$$q(p, s) = \frac{p(1-s)}{p(1-s) + (1-p)s}. \tag{3}$$

From (3), the receiver's posterior belief is decreasing in signal. In particular, signal $s = 1$ perfectly reveals state $\omega = 0$, while $s = 0$ reveals state $\omega = 1$. Because $q(p, s) \geq 1/2$ if and only if $s \leq p$, the receiver will choose the sender's preferred action if and only if she receives a signal below her private belief.

For one example, the c.d.f. F that places all the mass at the atom $1/2$ is the completely uninformative information structure. In both states, it releases signal $1/2$ for sure. For another example, the c.d.f. F that places half of the probability at the atom 0 and the other half of the probability at 1 is the completely informative information structure. It releases signal 1 for sure in state $\omega = 0$ and releases signal 0 for sure in state $\omega = 1$.

For a more relevant example, suppose the common prior is $\pi \in (0, 1/2)$. Consider the following c.d.f.

$$F(s) = \begin{cases} 0, & \text{if } s \in [0, \pi), \\ \frac{1}{2(1-\pi)}, & \text{if } s \in [\pi, 1), \\ 1, & \text{if } s = 1. \end{cases} \tag{4}$$

Because its mean is $1/2$, it is an information structure. From (1) and (2), the associated signal distributions are

$$F_0(s) = \begin{cases} 0, & \text{if } s \in [0, \pi), \\ \frac{\pi}{1-\pi}, & \text{if } s \in [\pi, 1), \\ 1, & \text{if } s = 1, \end{cases} \quad \text{and } F_1(s) = \begin{cases} 0, & \text{if } s \in [0, \pi), \\ 1, & \text{if } s \in [\pi, 1]. \end{cases}$$

Clearly, this information structure contains two signals, π and 1. In state $\omega = 0$, it releases signal $s = \pi$ with probability $\pi/(1-\pi)$ and signal $s = 1$ with probability $1 - \pi/(1-\pi)$. In state $\omega = 1$, it releases signal $s = \pi$ for sure. If the receiver's private belief is π, her posterior belief about state $\omega = 1$ after signal π is $q(\pi, \pi) = 1/2$ and that after signal 1 is $q(\pi, 1) = 0$. Kamenica and Gentzkow (2011) show that this information structure would be the sender's unique optimal information structure if the receiver had no private information (hence the sender faces no ambiguity). This kind of information structure plays an important role in the following analysis. We thus refer to F as the *KG solution for belief* π.

[9] This is because the "probability density" of F_0, i.e., its Radon–Nikodym derivative with respect to F, is $2s$ from (1), and that of F_1 is $2(1-s)$ from (2). Strictly speaking, (3) holds F-almost surely.

[10] From (3), we see that $q(1/2, s) = 1 - s$ for all $s \in [0, 1]$. Equivalently, the posterior belief about state $\omega = 0$ is $1 - q(1/2, s) = s$. Because s is distributed according to F, an information structure F can be equivalently interpreted as a distribution of the posterior belief for state $\omega = 0$, given prior $1/2$. The requirement that $\int s \mathrm{d}F(s) = 1/2$ is then the Bayes plausibility condition in Kamenica and Gentzkow (2011).

2.4 Sender's payoff and information design problem

Suppose the sender designs information $F \in \mathcal{F}$ and the receiver's private belief is p. If the receiver receives a signal $s > p$ from F, then her posterior belief is $q(p, s) < 1/2$. In this case, the receiver will choose $a = 0$ and the sender's payoff is $v(0) = 0$. If, instead, the receiver receives a signal $s \leq p$, her posterior belief is $q(p, s) \geq 1/2$. In this case, the receiver will choose $a = 1$ and the sender's payoff is $v(1) = 1/2$. Thus, the sender's expected payoff from F is $\Pr(s \leq p)/2$. Because the sender and receiver hold a common prior, the sender also has belief p conditional on the event that the receiver's private belief is p. Hence, the sender believes that the signals are distributed according to $(1-p)F_0 + pF_1$, which implies $\Pr(s \leq p) = (1-p)F_0(p) + pF_1(p)$.[11] Therefore, we can write the sender's expected payoff as

$$\begin{aligned} \phi^F(p) &\equiv \frac{1}{2}[(1-p)F_0(p) + pF_1(p)] \\ &= \int_{[0,p]} [(1-p)s + p(1-s)]\mathrm{d}F(s), \end{aligned} \tag{5}$$

where the second equality comes from (1) and (2). We call the function $\phi^F : [0, 1] \to \mathbb{R}$ the sender's *contingent payoff function* from information structure F. It will be the central focus of our analysis.[12]

If the sender knew the receiver's private information structure μ, his ex ante expected payoff from F would be

$$\int_{[\alpha,\beta]} \phi^F(p)\mu(\mathrm{d}p).$$

However, the receiver's information structure is private and the sender is uncertain about it. The sender knows only that the receiver's private information structure is one of those in $M([\alpha, \beta], \pi)$. Following the standard *maxmin expected utility function* assumption in the ambiguity aversion literature (e.g., Gilboa and Schmeidler 1989; Garrett 2014; Carroll 2015), we assume that the sender evaluates an information structure $F \in \mathcal{F}$ in terms of its worst-case expected payoff. When the sender designs F, it is the worst-case expected payoff that he seeks to maximize. Formally, for each $F \in \mathcal{F}$, let

$$V^F(\pi; [\alpha, \beta]) \equiv \inf_{\mu \in M(\pi;[\alpha,\beta])} \int_{[\alpha,\beta]} \phi^F(p)\mu(\mathrm{d}p). \tag{6}$$

be the sender's worst-case expected payoff if he designs information structure F. Then the sender's robust information design problem can be succinctly written as

[11] This is the key difference between our setting of private information and one of noncommon priors. At the interim stage, the sender and receiver may have different beliefs. But this occurs only because the receiver has received her private signal about the underlying states. When the sender forms his belief about the signal distribution, given the receiver's private belief, he should take this fact into account. If, instead, the two agents have non-common priors and intrinsically differ in their beliefs, then the sender's belief about the signal distribution would be $\pi F_0 + (1-\pi)F_1$.

[12] The KG solution for belief $\pi \in (0, 1/2)$ in (4) is the information structure that solves $\max_{F \in \mathcal{F}} \phi^F(\pi)$.

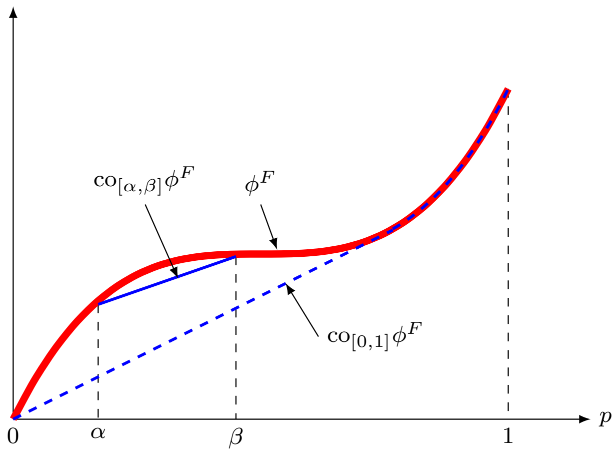


Fig. 1 Convexification on different domains

$$V(\pi; [\alpha, \beta]) \equiv \max_{F \in \mathcal{F}} V^F(\pi; [\alpha, \beta]). \tag{7}$$

2.5 Convexification

For $\alpha < \beta$ and the sender's information structure $F \in \mathcal{F}$, let $\mathrm{co}_{[\alpha,\beta]}\phi^F : [\alpha, \beta] \to \mathbb{R}$ be the *convexification of* ϕ^F *over* $[\alpha, \beta]$. It is the largest convex function below ϕ^F over the interval $[\alpha, \beta]$. Formally,

$$\mathrm{co}_{[\alpha,\beta]}\phi^F(p) \equiv \sup_{\substack{\text{convex } f:[\alpha,\beta]\to\mathbb{R} \\ f \le \phi^F \text{ over } [\alpha,\beta]}} f(p), \ \ \forall p \in [\alpha, \beta].$$

Figure 1 provides an illustration of convexification. It is worth emphasizing that the value of the convexification in general depends on its domain $[\alpha, \beta]$. Different domains $[\alpha, \beta]$ will lead to different values of convexification even for the same sender's information structure F. For instance, the dashed blue curve in Fig. 1 is the convexification over the whole domain $[0, 1]$, while the solid blue line segment is the convexification over the interval $[\alpha, \beta]$. Obviously, these two are quite different.

Applying the concavification result in Kamenica and Gentzkow (2011) (Corollary 2) to the minimization problem in (6), the following lemma provides a characterization of the sender's worst-case expected payoff from an information structure $F \in \mathcal{F}$ using the notion of convexification.

Lemma 1 *For any* $\alpha < \beta$ *and information structure* $F \in \mathcal{F}$*,*

$$V^F(\pi; [\alpha, \beta]) = \mathrm{co}_{[\alpha,\beta]}\phi^F(\pi), \ \ \forall \pi \in (\alpha, \beta).$$

Therefore,

$$V(\pi; [\alpha, \beta]) = \max_{F \in \mathcal{F}} \mathrm{co}_{[\alpha,\beta]}\phi^F(\pi), \ \ \forall \pi \in (\alpha, \beta).$$

Intuitively speaking, the sender's worst-case payoff from an information structure F is the worst average payoff from F. Thus, to maximize his worst-case payoff, the sender should care about the average performance of an information structure, instead of its performance at a particular private belief.

3 Sender's optimal information structure

In this section, we analyze the sender's optimal information structure. We divide the whole analysis into two cases: $\beta = 1$ and $\beta < 1$. Although $\beta = 1$ is the limiting case of $\beta < 1$, we treat it separately because it is simple.

3.1 $\beta = 1$

Consider the KG solution for belief α[13]:

$$F^{\alpha,\alpha}(s) = \begin{cases} 0, & \text{if } s \in [0, \alpha), \\ \frac{1}{2(1-\alpha)}, & \text{if } s \in [\alpha, 1), \\ 1, & \text{if } s = 1. \end{cases} \tag{8}$$

Recall that $F^{\alpha,\alpha}$ uniquely maximizes $\phi^F(\alpha)$ over all $F \in \mathcal{F}$. The sender's contingent payoff function under $F^{\alpha,\alpha}$ can be calculated from (5),

$$\phi^{F^{\alpha,\alpha}}(p) = \frac{1-2\alpha}{2(1-\alpha)}(p-\alpha) + \alpha, \ \ \forall p \in [\alpha, 1].$$

Most importantly, this contingent payoff function is linear over $[\alpha, 1]$. Hence, by Lemma 1, the sender's worst-case payoff for any prior $\pi \in (\alpha, 1)$ is just $\phi^{F^{\alpha,\alpha}}(\pi)$.

Consider a prior $\pi \in (\alpha, 1)$ and any information structure $F \in \mathcal{F}$ other than $F^{\alpha,\alpha}$. If the receiver's private belief happens to be distributed according to $\lambda \circ \alpha + (1-\lambda) \circ 1$ for $\lambda \in (0, 1)$ that satisfies $\lambda\alpha + (1-\lambda) = \pi$, then the sender's payoff from F is $\lambda\phi^F(\alpha) + (1-\lambda)\phi^F(1)$. Therefore, the sender's *worst-case* payoff $\mathrm{co}_{[\alpha,1]}\phi^F(\pi)$ is no higher than $\lambda\phi^F(\alpha) + (1-\lambda)\phi^F(1)$. At private belief α, we have $\phi^F(\alpha) < \phi^{F^{\alpha,\alpha}}(\alpha)$. At private belief 1, $\phi^F(1) = \phi^{F^{\alpha,\alpha}}(1)$ because the sender's information structure becomes irrelevant once the receiver has known the state $\omega = 1$. Therefore,

$$\mathrm{co}_{[\alpha,1]}\phi^F(\pi) < \lambda\phi^{F^{\alpha,\alpha}}(\alpha) + (1-\lambda)\phi^{F^{\alpha,\alpha}}(1) = \phi^{F^{\alpha,\alpha}}(\pi) = \mathrm{co}_{[a,1]}\phi^{F^{\alpha,\alpha}}(\pi),$$

[13] We write double superscripts (α, α) in order to be consistent with our notation in the next subsection. See, for example, Lemma 2.

where both equalities come from the fact that the contingent payoff function $\phi^{F^{\alpha,\alpha}}$ is linear over $[\alpha, 1]$. This observation leads to the following proposition, which characterizes the sender's optimal information structure when $\beta = 1$.

Proposition 1 *Suppose $\beta = 1$. For every $\pi \in (\alpha, \beta)$, $F^{\alpha,\alpha}$ is the unique optimal information structure for the sender.*

A special case is $\alpha = 0$. This represents the sender's full uncertainty. He has no knowledge at all about the receiver's private information structure and simply thinks every distribution of private beliefs with mean π is possible. In this case, the optimal information $F^{0,0}$ corresponds to full information disclosure because only $s = 0$ and $s = 1$ will be generated:

$$F^{0,0}(s) = \begin{cases} \frac{1}{2}, & \text{if } s \in [0, 1), \\ 1, & \text{if } s = 1. \end{cases}$$

This is the special case of the full revelation result in Gentzkow and Kamenica (2017a) for exactly two players, the sender and nature, with zero sum payoffs. According to the same logic as in Gentzkow and Kamenica (2017a), the optimality of full information disclosure holds in general environments with multiple states and actions, and arbitrary sender and receiver payoff specifications.

3.2 $\beta < 1$

3.2.1 Main result

The analysis of the optimal information structure for $\beta < 1$ becomes much more difficult than that for $\beta = 1$. In this subsection, we present our main result. The next two subsections are devoted to explaining the ideas behind it in greater details.

As we have seen, one of the reasons that $F^{\alpha,\alpha}$ is sender optimal is that its contingent payoff function is linear. The following definition generalizes this property to a larger class of information structures.

Definition 1 An information structure $F \in \mathcal{F}$ is a *linear-contingent-payoff information structure (LCPIS)* over $[\alpha, \beta]$ if $\phi^F(p) = \max\{0, \ell(p)\}$ over $[\alpha, \beta]$ for some strictly increasing linear function ℓ.

For example, the KG solution $F^{\alpha,\alpha}$ is an LCPIS over $[\alpha, \beta]$ for any $\beta > \alpha$. Because the contingent payoff function of an LCPIS is convex by definition, the worst-case payoff function $\text{co}_{[\alpha,\beta]}\phi^F$ coincides with ϕ^F for every LCPIS F by Lemma 1.

Our main result can be succinctly stated as follows. Recall that $V(\pi; [\alpha, \beta])$ is the sender's optimal value when the prior is $\pi \in (\alpha, \beta)$.

Proposition 2 (Main result) *For any $\pi \in (\alpha, \beta)$,*

$$V(\pi; [\alpha, \beta]) = \max_{F \in \mathcal{F}} \phi^F(\pi)$$
$$\textit{s.t. } F \textit{ is an LCPIS over } [\alpha, \beta].$$

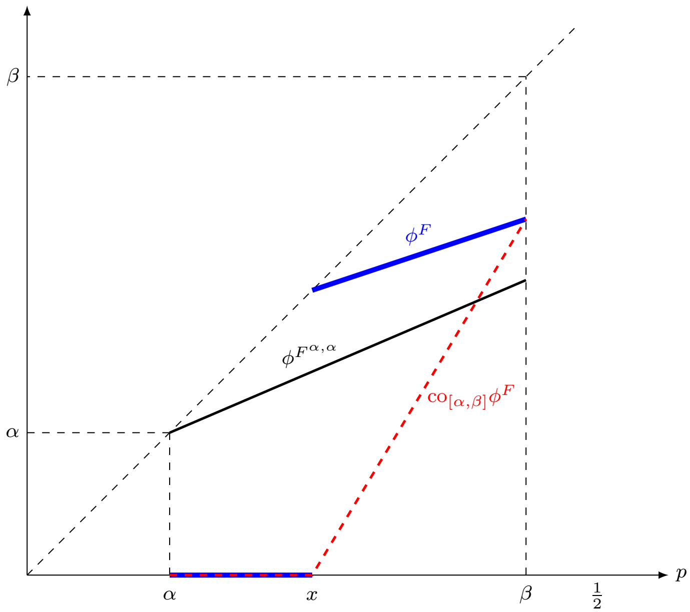


Fig. 2 When $\beta < 1$, $F^{\alpha,\alpha}$ is not optimal for prior π close to β

In words, Proposition 2 states that for any prior $\pi \in (\alpha, \beta)$, we can always find the sender's optimal information structure by looking at only LCPIS's. The proof of this result is very involved. In a nutshell, the whole proof can be divided into two major steps, which correspond to two important properties of the class of LCPIS's: (1) the richness of the class of LCPIS's and (2) the optimality of the LCPIS's. The next two subsections explain these two properties respectively.

3.2.2 Richness

As mentioned above, the KG solution for belief α, $F^{\alpha,\alpha}$, is an LCPIS. But if it were the only LCPIS, then Proposition 2 would not hold when $\beta < 1$. To see this, consider the KG solution for some belief $x \in (\alpha, \min\{\beta, 1/2\})$, say F. The blue curve in Fig. 2 is its contingent payoff function and the dashed red curve is the corresponding worst-case payoff function. Because $\phi^F(\beta) > \phi^{F^{\alpha,\alpha}}(\beta)$ when $\beta < 1$, we see in Fig. 2 that $\mathrm{co}_{[\alpha,\beta]}\phi^F$ exceeds $\phi^{F^{\alpha,\alpha}}$ for priors sufficiently close to β. Thus, $F^{\alpha,\alpha}$ cannot be optimal for those priors.

This example points out a trade-off faced by the sender when $\beta < 1$. Some information structures lead to higher contingent payoffs for private beliefs close to α, e.g., $F^{\alpha,\alpha}$, while others lead to higher contingent payoffs for private beliefs close to β, e.g., F in Fig. 2. When the prior π is close to α, the sender should care more about his contingent payoffs at private beliefs close to α, because this is the area where the receiver's private belief is most likely to occur. In contrast, when the prior π is close to β, the sender should care more about his contingent payoffs at high private beliefs. Note that in the previous analysis of the case $\beta = 1$, this trade-off is absent because

$\phi^F(1)$ is a constant across all $F \in \mathcal{F}$. This is the major reason why the analysis for $\beta = 1$ is much easier than that for $\beta < 1$.

To deal with this trade-off, we need more LCPIS's than just $F^{\alpha,\alpha}$. In general, if F is an LCPIS, then ϕ^F over $[\alpha, \beta]$ takes one of the following two forms, depending on the value of $\phi^F(\alpha)$. If $\phi^F(\alpha) > 0$, then ϕ^F is similar to $\phi^{F^{\alpha,\alpha}}$ and it can be written as

$$\phi^F(p) = a(p - x) + b, \ \forall p \in [\alpha, \beta],$$

where $x = \alpha$, $b \in (0, \alpha]$, and some $a > 0$.[14] If $\phi^F(\alpha) = 0$, then it takes the form

$$\phi^F(p) = \begin{cases} 0 & \text{if } p \in [\alpha, x], \\ a(p - x) + b & \text{if } p \in (x, \beta], \end{cases}$$

for some $x \in [\alpha, \beta)$, $b = 0$, and some $a > 0$. Hence, if F is an LCPIS, its contingent payoff function is characterized by an *initial point* $(x, b) \in \{(\alpha, b)|0 < b \le \alpha\} \cup \{(x, 0)|\alpha \le x < \beta\}$ and a *slope* $a > 0$. In this case, with a slight abuse of terminology, we say that this F has initial point (x, b) and slope a.

For the same initial point, there are potentially many LCPIS's. These LCPIS's differ only in their slopes. Among these LCIPS's, if an LCPIS is the sender's optimal information structure for some prior, it must have the largest slope. This observation leads to the following definition.

Definition 2 An LCPIS F with initial point (x, b) and slope a is *dominant* if there is no LCPIS F' that has the same initial point (x, b) but larger slope $a' > a$.

The KG solution for belief α, $F^{\alpha,\alpha}$, is trivially a dominant LCPIS for any β, since it is the only LCPIS with initial point (α, α). The following lemma shows that we can find the dominant LCPIS for not only initial point (α, α), but a much richer set of initial points.

Lemma 2 (Richness)

(i) *Suppose $\beta \le 1/2$. For every (x, b) in*

$$A(\alpha, \beta) \equiv \{(\alpha, b) \,|\, 0 < b \le \alpha\} \bigcup \{(x, 0) \,|\, \alpha \le x < \beta\},$$

there exists a unique dominant LCPIS $F^{x,b}$ with initial point (x, b).

(ii) *Suppose $\beta > 1/2$ and $\alpha < 1 - \beta$. For every (x, b) in*

$$A(\alpha, \beta) \equiv \{(\alpha, b) \,|\, 0 < b \le \alpha\} \bigcup \{(x, 0) \,|\, \alpha \le x \le 1 - \beta\},$$

there exists a unique dominant LCPIS $F^{x,b}$ with initial point (x, b). Moreover, $\phi^{F^{1-\beta,0}}(\beta) = 1/2$.

[14] Because $\phi^{F^{\alpha,\alpha}}(\alpha) = \alpha$, $b = \phi^F(\alpha) \le \phi^{F^{\alpha,\alpha}}(\alpha) = \alpha$.

(iii) *Suppose* $\beta > 1/2$ *and* $\alpha \geq 1 - \beta$. *There exists* $\hat{b} \in [0, \alpha)$ *such that for every* (x, b) *in*

$$A(\alpha, \beta) \equiv \{(\alpha, b) \mid \hat{b} \leq b \leq \alpha\},$$

there is a unique dominant LCPIS $F^{\alpha,b}$ *with initial point* (α, b). *Moreover,* $\phi^{F^{\alpha,\hat{b}}}(\beta) = 1/2$.

Some remarks are in order. First, in short, Lemma 2 identifies a set $A(\alpha, \beta)$ of initial points at each of which the dominant LCPIS exists. The statement is divided into three parts that correspond to three different cases of the value of β. The three graphs in Fig. 3 illustrate these three cases respectively. In each graph, the thick gray curve represents the corresponding set $A(\alpha, \beta)$. Second, the statement that $\phi^{F^{1-\beta,0}}(\beta) = 1/2$ in part (ii) simply claims that $F^{1-\beta,0}$ is *a* KG solution for belief β: it maximizes the sender's expected payoff when the common prior is β and the receiver has no private information.[15] See $\phi^{F^{1-\beta,0}}$ in Fig. 3b. Similarly, the statement that $\phi^{F^{\alpha,\hat{b}}}(\beta) = 1/2$ in part (iii) also means that $F^{\alpha,\hat{b}}$ is a KG solution for belief β. See $\phi^{F^{\alpha,\hat{b}}}$ in Fig. 3c. Third, the dominant LCPIS's in this lemma in general depend on the value of β. We suppress β from the notation $F^{x,b}$ just for expositional ease.[16]

Figure 4 illustrates a typical dominant LCPIS with initial point $(x, 0)$. Figure 4a depicts $F^{x,0}$ itself and Fig. 4b draws the graphs of its corresponding signal distributions, $F_0^{x,0}$ and $F_1^{x,0}$.[17] There are two important features. First, unlike the KG solution for a belief less than $1/2$, which contains only two signals, $F^{x,0}$ involves infinitely many signals. In particular, the continuum of signals in $[x, \beta]$ are specially distributed to guarantee that $\phi^{F^{x,0}}$ is an LCPIS: $\phi^{F^{x,0}}$ is linear over $[x, \beta]$. Second, the only signal bigger than β is the atom $s = 1$, which perfectly reveals state $\omega = 0$. In other words, the receiver does not choose the sender's preferred action at private belief β only if she observes $s = 1$. This feature is to guarantee that $F^{x,0}$ is dominant. It minimizes the probability of the receiver not choosing the sender's preferred action.

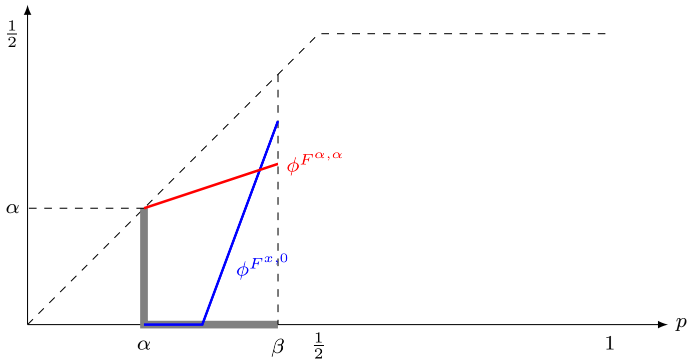
When $\beta > 1/2$ and $\alpha < 1 - \beta$ (part (ii) of Lemma 2), signal $s = 1$ is no longer an atom signal under $F^{1-\beta,0}$. In fact, the support of $F^{1-\beta,0}$ is just $[1 - \beta, \beta]$ and both $F_0^{1-\beta,0}$ and $F_1^{1-\beta,0}$ share this common support. Hence, every signal is below β. This guarantees that, when the receiver's private belief is β, she always chooses the sender's preferred action. This is why $\phi^{F^{1-\beta,0}}(\beta) = 1/2$. Likewise, when $\beta > 1/2$ and $\alpha \geq 1 - \beta$ (part (iii) of Lemma 2), both $F_0^{\alpha,\hat{b}}$ and $F_1^{\alpha,\hat{b}}$ share the common support $[\alpha, \beta]$, which again guarantees that $\phi^{F^{\alpha,\hat{b}}}(\beta) = 1/2$.[18]

[15] Unlike the case where the prior is less than $1/2$, there are many information structures that maximize the sender's expected payoff when the prior greater than $1/2$ and the receiver has no private information. For example, the completely uninformative information structure does the job.

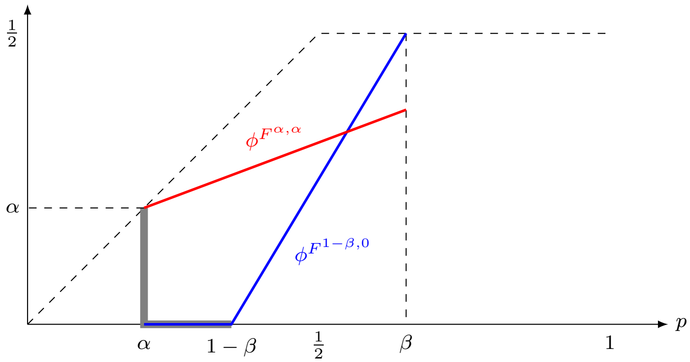
[16] The only exception is $F^{\alpha,\alpha}$. Recall (8).

[17] The functional form of $F^{x,0}$ is given by A.5 in Appendix A.

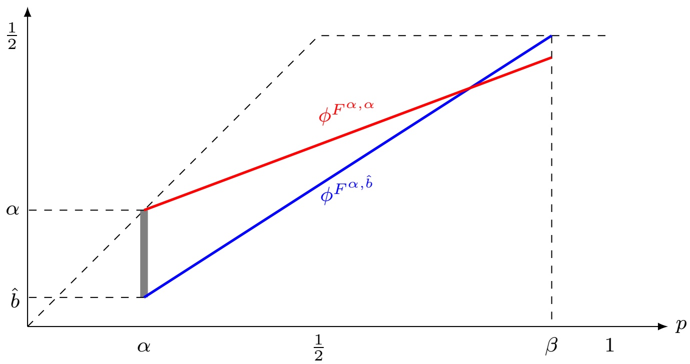
[18] What makes the value $1/2$ special relative to β is the fact that the receiver's cut-off belief at which she is indifferent between actions $a = 0$ and $a = 1$ is $1/2$. Changing the receiver's belief cut-off to an arbitrary number in $(0, 1)$ will not change our results qualitatively. But it will definitely make the analysis much more cumbersome given that the current analysis is already very complex.



(a) $\beta \leq 1/2$

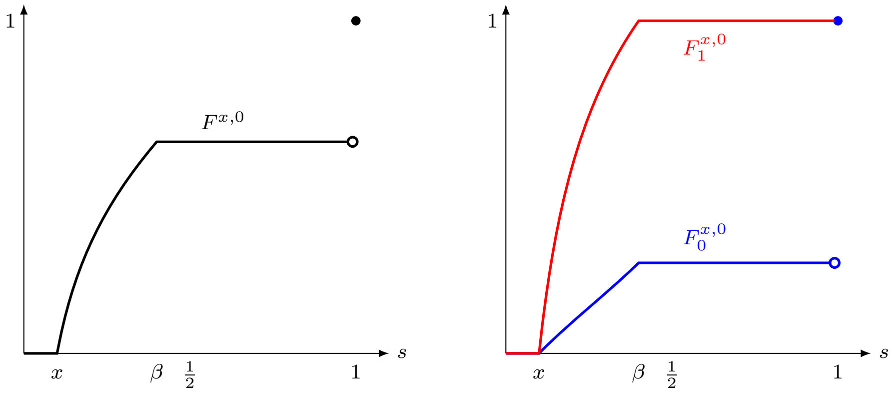


(b) $\alpha < 1-\beta < \frac{1}{2} < \beta$



(c) $1-\beta \leq \alpha < \frac{1}{2} < \beta$

Fig. 3 The set $A(\alpha, \beta)$ and the corresponding LCPIS's



(a) $F^{x,0}$ **(b)** Signal distributions of $F^{x,0}$

Fig. 4 The dominant LCPIS $F^{x,0}$

For any initial point $(x, b) \in A(\alpha, \beta)$, we prove the result by explicitly constructing the dominant LCPIS. For instance, consider an initial point of type $(x, 0) \in A(\alpha, \beta)$. If F is an LCPIS with initial point $(x, 0)$, then $\phi^F(p) = a(p - x)$ for some $a > 0$ over the interval $[x, \beta]$. Using the expression of ϕ^F in (5) and applying integration by parts, we can rewrite this relationship as

$$\phi^F(p) = 2p(1-p)F(p) - (1-2p)\int_x^p F(s)\mathrm{d}s = a(p-x), \ \ \forall p \in [x, \beta]. \quad (9)$$

This continuum of equations defines a differential equation

$$2p(1-p)\frac{\mathrm{d}y}{\mathrm{d}p} - (1-2p)y = a(p-x) \quad (10)$$

over interval $[x, \beta]$ with initial condition $y(x) = 0$. The unique solution of this differential equation together with the condition $F(s) = 0$ for all $s \le x$ (since $\phi^F(x) = 0$) give us the functional form of F over $[0, \beta]$ that satisfies (9). The slope a is a parameter in F. We then find the maximal value of a with which this F function over $[0, \beta]$ can be extended to an information structure. This results in the dominant LCPIS $F^{x,0}$. The construction of the dominant LCPIS for the initial point of type $(\alpha, b) \in A(\alpha, \beta)$ is similar. The only difference is that $F^{\alpha,b}$ has an atom at α in order to guarantee $\phi^{F^{\alpha,b}}(\alpha) = b > 0$.

3.2.3 Optimality

With a rich set of dominant LCPIS's, we can proceed to show our main result, Proposition 2. For this, it suffices to show that the worst-case expected payoff function $\mathrm{co}_{[\alpha,\beta]}\phi^F$ of any arbitrary information structure F is always below the function $\widetilde{V}(\,\cdot\,; [\alpha, \beta]) : [\alpha, \beta] \to \mathbb{R}$ defined as

$$\widetilde{V}(\pi;[\alpha,\beta]) = \max_{(x,b)\in A(\alpha,\beta)} \phi^{F^{x,b}}(\pi), \quad \forall \pi \in [\alpha,\beta]. \tag{11}$$

The value $\widetilde{V}(\pi;[\alpha,\beta])$ is the highest worst-case payoff the sender can obtain by using *only* dominant LCPIS's. However, a direct comparison of $\text{co}_{[\alpha,\beta]}\phi^F$ and $\widetilde{V}(\cdot;[\alpha,\beta])$ is impossible, because ϕ^F can be quite complicated for an arbitrary F, let alone its worst-case payoff function. The trick here is to approximate $\text{co}_{[\alpha,\beta]}\phi^F$ by linear functions below ϕ^F. The comparison between these linear functions and the function $\widetilde{V}(\cdot;[\alpha,\beta])$ turns out to be much more tractable, since, after all, $\widetilde{V}(\cdot;[\alpha,\beta])$ is just the upper envelope of some "linear functions." The following lemma, which is the key to Proposition 2, gives a formal statement of this comparison.

Lemma 3 (Optimality) *For any information structure $F \in \mathcal{F}$ and linear function $\ell : [\alpha,\beta] \to \mathbb{R}$ below ϕ^F, there exists an initial point $(x,b) \in A(\alpha,\beta)$ such that*

$$\ell(p) \le \phi^{F^{x,b}}(p), \;\; \forall p \in [\alpha,\beta].$$

In words, Lemma 3 states that the dominant LCPIS's in Lemma 2 dominate not only all the LCPIS's, but all linear functions below any contingent payoff function. Another way to understand Lemma 3 is to think about how fast on average a contingent payoff function can grow over the interval $[\alpha,\beta]$. Suppose $\phi^F(\alpha) = b$ for some information structure F. Lemma 3 implies that the minimal average growth rate of ϕ^F over $[\alpha,\beta]$,

$$\inf_{p\in[\alpha,\beta]} \frac{\phi^F(p) - b}{p-\alpha},$$

has an upper bound. The dominant LCPIS $F^{\alpha,b}$ achieves this upper bound by ensuring that $\phi^{F^{\alpha,b}}$ grows at a constant rate. Similarly, if $\phi^F(x) = 0$ for some $x \in (\alpha,\beta)$, then its minimal average growth rate over (x,β) is bounded above by that of $F^{x,0}$, whose contingent payoff function grows at a constant rate.

We use the case $\beta \le 1/2$ to sketch the Proof of Lemma 3. Suppose the linear function $\ell(p) = a(p-\alpha) + b$ is below ϕ^F for some $a > 0$ and $b \le \alpha$.[19] Assume $b > 0$ for the moment. Because ℓ is below ϕ^F, we have

$$2p(1-p)F(p) - (1-2p)\int_0^p F(s)\mathrm{d}s \ge a(p-\alpha) + b, \quad \forall p \in [\alpha,\beta]. \tag{12}$$

This inequality is similar to (9). But the critical difference is that we now have a "differential inequality," which cannot be explicitly "solved." What we show is that F over the interval $[\alpha,\beta]$ is bounded below by the solution G to the differential equation

$$2p(1-p)G(p) - (1-2p)\int_\alpha^p G(s)\mathrm{d}s = a(p-\alpha) + b, \quad \forall p \in [\alpha,\beta], \tag{13}$$

[19] Since $l \le \phi^F$, $b = \ell(\alpha) \le \phi^F(\alpha) \le \phi^{F^{\alpha,\alpha}}(\alpha) = \alpha$. If $a \le 0$, then $\ell \le \phi^{F^{\alpha,\alpha}}$.

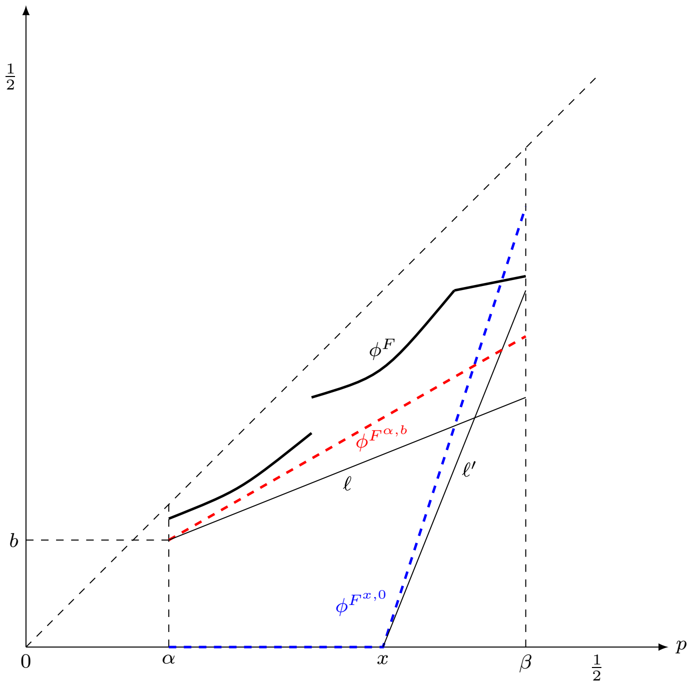


Fig. 5 Illustration of the proof for Lemma 3

with the initial condition $\int_\alpha^\alpha G(s)\mathrm{d}s = 0$. Since $(\alpha, b) \in A(\alpha, \beta)$ by Lemma 2, we can compare the solution G and the dominant LCPIS $F^{\alpha,b}$. If a is greater than the slope of $\phi^{F^{\alpha,b}}$, we know that such solution G can not be part of an information structure; otherwise, it contradicts the dominance of $\phi^{F^{\alpha,b}}$. Using the fact that F is above G over $[\alpha, \beta]$, we further show that F cannot be an information structure either if a is greater than the slope of $\phi^{F^{\alpha,b}}$. This in turn implies that $\ell \le \phi^{F^{\alpha,b}}$, since ℓ and $\phi^{F^{\alpha,b}}$ share the same initial point.

Figure 5 illustrates the above idea. The thick black curve represents an arbitrary contingent payoff function ϕ^F. The straight line ℓ, which takes the form $\ell(p) = a(p-\alpha)+b$ for some $b \in [0, \alpha]$ and $a > 0$, is below ϕ^F over $[\alpha, \beta]$. What we show is that the slope a is no greater than that of $\phi^{F^{\alpha,b}}$. This immediately implies that ℓ can not exceed $\phi^{F^{\alpha,b}}$ over $[\alpha, \beta]$, as shown in the figure.

If $b \le 0$, then there exists $x \in [\alpha, \beta)$ such that $\ell(x) = 0$.[20] In this case, since $(x, 0) \in A(\alpha, \beta)$ by Lemma 2, we can use a similar argument as above to show that $\ell \le \phi^{F^{x,0}}$. The straight line ℓ' in Fig. 5 illustrates this case. It takes the form $\ell'(p) = a(p - x)$ for some $a > 0$. What we show is that its slope a is no greater than that of $\phi^{F^{x,0}}$. This, again, immediately implies that ℓ' can not exceed $\phi^{F^{x,0}}$ over $[\alpha, \beta]$.

[20] If such x does not exist, then ℓ is everywhere below 0. But every LCPIS is bounded below by 0.

When $\beta > 1/2$, the initial point (α, b) or $(x, 0)$ in the above arguments may not be contained in $A(\alpha, \beta)$, as is defined in Lemma 2. However, in this case, $\ell \le \phi^{F^{1-\beta,0}}$ if $\alpha < 1 - \beta$ and $\ell \le \phi^{F^{\alpha,\hat{b}}}$ if $\alpha \ge 1 - \beta$.

3.2.4 Optimal information structure

With Proposition 2, we can now characterize the sender's optimal information structure. Specifically, to find a sender's optimal information structure for prior $\pi \in (\alpha, \beta)$, we only need to solve the optimization problem (11).

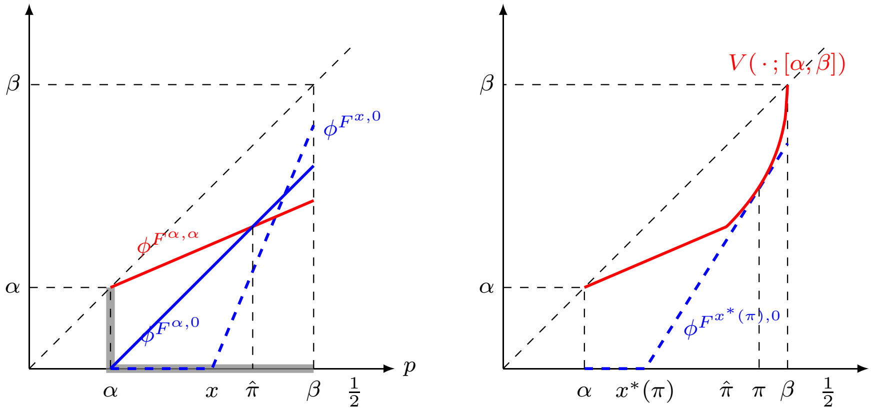
The following proposition fully characterizes the solution to (11) for the first two cases in Lemma 2: $\alpha < \beta \le 1/2$ or $\alpha < 1 - \beta < 1/2 < \beta$. Except for a special cut-off prior, there is a unique sender optimal information structure within the class of LCPIS's.

Proposition 3 *Suppose $\alpha < \beta \le 1/2$ or $\alpha < 1 - \beta < 1/2 < \beta$. Consider the optimization problem (11). There exists a cut-off prior $\hat{\pi} \in (\alpha, \beta)$, such that*

- (i) *if $\pi \in (\alpha, \hat{\pi})$, $F^{\alpha,\alpha}$ is the unique solution;*
- (ii) *if $\pi = \hat{\pi}$, any $F^{\alpha,b}$ for $b \in [0, \alpha]$ is a solution; and*
- (iii) *if $\pi \in (\hat{\pi}, \beta)$, the unique solution takes the form of $F^{x^*(\pi),0}$ for some $x^*(\pi) \in (\alpha, \min\{\pi, 1 - \beta\}]$. As a function of π, x^* is continuous, increasing with range $(\alpha, \min\{\beta, 1 - \beta\})$.*

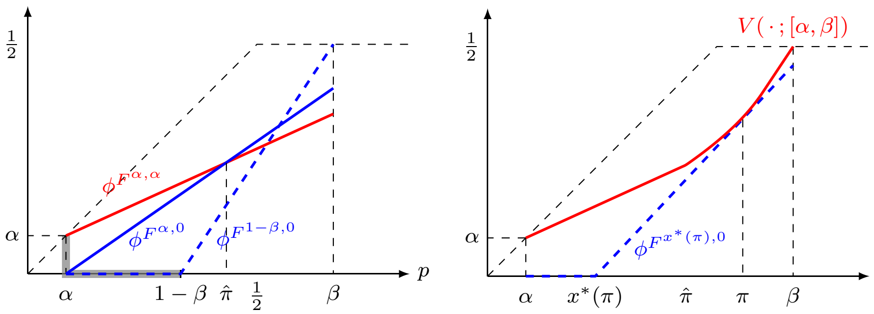
Proposition 3 is best understood by looking at Figs. 6 and 7, which illustrate the dominant linear-contingent-payoff functions and the value functions for the cases $\alpha < \beta \le 1/2$ and $\alpha < 1 - \beta < 1/2 < \beta$, respectively. Take Fig. 6a as an example. Figure 6a depicts some dominant linear-contingent-payoff functions. Starting from (α, α), as the initial point goes down along the line segment between (α, α) and $(\alpha, 0)$, the sender's contingent payoff strictly decreases at private belief α and strictly increases at private belief β. This change can be seen by comparing the red curve, $\phi^{F^{\alpha,\alpha}}$, and the solid blue curve, $\phi^{F^{\alpha,0}}$ in Fig. 6a. Interestingly, all these contingent payoff functions $\{\phi^{F^{\alpha,b}}\}_{b\in[0,\alpha]}$ happen to rotate around one single point, $\hat{\pi}$. Starting from $(\alpha, 0)$, as the initial point moves to the right along the line segment between $(\alpha, 0)$ and $(\beta, 0)$, the sender's contingent payoff at private belief β continues to increase. This change can be seen by comparing the solid blue curve, $\phi^{F^{\alpha,0}}$, and the dashed blue curve, $\phi^{F^{x,0}}$ in Fig. 6a. As x approaches β, $\phi^{F^{x,0}}(\beta)$ also increases to β. These contingent payoff functions intersect $\phi^{F^{\alpha,0}}$ at points to the right of the cut-off prior $\hat{\pi}$.

To solve problem (11) for prior $\pi \in (\alpha, \beta)$, we only need to check which contingent payoff function is the highest at π in Fig. 6a. From the above explanation, at prior π below the cut-off $\hat{\pi}$, only $F^{\alpha,\alpha}$ is optimal. When $\pi = \hat{\pi}$, any $F^{\alpha,b}$ for $b \in [\alpha, \beta]$ is optimal. When π is above the cut-off prior $\hat{\pi}$, LCPIS's of the form $F^{\alpha,b}$ are no longer optimal. Instead, the optimal LCPIS appears in $\{F^{x,0}\}_{x\in(\alpha,\beta)}$. However, no single LCPIS is optimal for all priors above $\hat{\pi}$. Instead, there is a one-to-one correspondence between priors above $\hat{\pi}$ and the optimal $F^{x,0}$. Figure 6b illustrates the sender's value function V. For priors below the cut-off $\hat{\pi}$, V just coincides with $\phi^{F^{\alpha,\alpha}}$ and thus is



(a) Some dominant LCP functions **(b)** Sender's value function

Fig. 6 Dominant LCP functions and the value function for $\alpha < \beta \leq \frac{1}{2}$



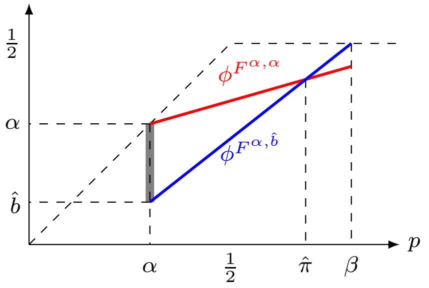
(a) Some dominant LCP functions **(b)** Sender's value function

Fig. 7 Dominant LCP functions and the value function for $\alpha < 1 - \beta < 1/2 < \beta$

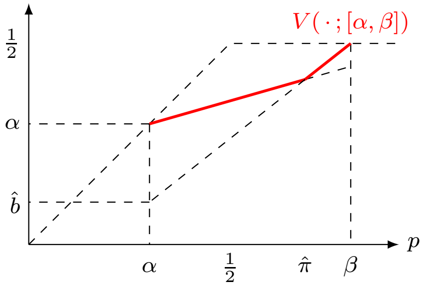
linear. For priors above the cut-off, V is the upper envelope of $\{\phi^{F^{x,0}}\}_{x\in(\alpha,\beta)}$ and is strictly convex.

Holding α and β fixed, the comparative statics of the optimal information structure with respect to prior π also coincide with our intuition. When the prior is below the cut-off $\hat{\pi}$, $F^{\alpha,\alpha}$ is optimal because the sender knows that the receiver's private beliefs are more likely to be close to α and $F^{\alpha,\alpha}$ yields high contingent payoffs over this range. In contrast, when π is above the cut-off $\hat{\pi}$, the sender knows that the receiver's private beliefs are more likely to be close to β. Hence, the optimal information structure guarantees high contingent payoffs over this range, although it gives the sender his lowest possible payoff when the receiver's private belief is low. As the prior approaches β, the sender is willing to sacrifice his contingent payoff for a larger range of low private beliefs in exchange for higher contingent payoffs at high private beliefs.

The next proposition reports the characterization of the solution to (11) for the last case in Lemma 2: $1 - \beta \leq \alpha < 1/2 < \beta$. The idea and result are similar to those of Proposition 3.



(a) Some dominant LCP functions



(b) Sender's value function

Fig. 8 Dominant LCP functions and the value function for $1-\beta \leq \alpha < 1/2 < \beta$

Proposition 4 *Suppose $1-\beta \leq \alpha < 1/2 < \beta$. There exists a cut-off prior $\hat{\pi} \in (\alpha, \beta)$, such that*

(i) *if $\pi \in (\alpha, \hat{\pi})$, $F^{\alpha,\alpha}$ is optimal;*
(ii) *if $\pi = \hat{\pi}$, any $F^{\alpha,b}$ for $b \in [\hat{b}, \alpha]$ is optimal; and*
(iii) *if $\pi \in (\hat{\pi}, \beta)$, $F^{\alpha,\hat{b}}$ is optimal.*

The major qualitative difference between Propositions 3 and 4 is that $F^{\alpha,\hat{b}}$ alone is optimal for all priors above the cut-off $\hat{\pi}$ when $\alpha \geq 1-\beta$. Thus, the optimal information structure is either the KG solution for belief α or a KG solution for belief β. Figure 8 illustrates the dominant linear-contingent-payoff functions and the sender's value function for this case.

We can also show that $\hat{b}$ increases to α as β increases to 1.[21] Thus, as β increases to 1, $\hat{\pi}$ increases to 1 as well and the optimal information structure in Proposition 4 degenerates to $F^{\alpha,\alpha}$ for all π. Therefore, the optimal information structure for the case $\beta = 1$ is indeed the limiting case as β goes to 1.

Given $\alpha < \beta$, both Propositions 3 and 4 show that the optimal information structure is unique within the class of LCPIS's for all $\pi \in (\alpha, \beta)\backslash\{\hat{\pi}\}$. In fact, we can further show that $F^{\alpha,\alpha}$ is the unique optimal information structure for prior $\pi < \hat{\pi}$. When $1-\beta \leq \alpha < 1/2 < \beta$, we can also show that $F^{\alpha,\hat{b}}$ is the unique optimal information structure for prior $\pi > \hat{\pi}$. For other cases, although we conjecture the unique optimal dominant LCPIS is also the unique optimal information structure, unfortunately we are unable to prove it.

3.3 Impacts of limited knowledge

From the analysis so far, we see that the ambiguity faced by the sender $[\alpha, \beta]$ greatly affects the sender's highest worst-case payoff and optimal information design. But when the ambiguity vanishes, the sender can always guarantee himself the payoff he would obtain if there were no ambiguity at all.

[21] See (A.4) for the closed form of $\hat{b}$.

Corollary 1 *Suppose $\alpha < \pi < \beta$. We have*

$$\lim_{\tilde{\alpha}\uparrow\pi} V(\pi; [\tilde{\alpha}, \beta]) = \lim_{\tilde{\beta}\downarrow\pi} V(\pi; [\alpha, \tilde{\beta}]) = \max_{F\in\mathcal{F}} \phi^F(\pi).$$

According to our model specification, when either $\alpha = \pi$ or $\beta = \pi$, the ambiguity completely disappears, because in this case the only private belief that the sender could have is just the prior π. As a result, Corollary 1 implies that ambiguity has little impact on the sender's expected payoff when there is only one-sided ambiguity, e.g., when the receiver's good news or bad news private signal is always sufficiently uninformative. It is important to note that Corollary 1 *does not* claim that the KG solution for belief π, i.e., the one that solves $\max_{F\in\mathcal{F}} \phi^F(\pi)$, is robust when the ambiguity vanishes. For example, consider $\pi < 1/2$. As long as there is ambiguity, i.e., $\alpha < \pi < \beta$, the KG solution for belief π always yields 0 worst-case payoff to the sender, regardless of how little ambiguity the sender is facing. Rather, Corollary 1 only states that the sender can find a way to guarantee himself a worst-case payoff close to $\max_{F\in\mathcal{F}} \phi^F$, e.g., the optimal information structure we found above, as the ambiguity vanishes.

When the ambiguity becomes larger, the sender is never better off. Our previous analysis can further tell us whether the sender becomes *strictly* worse off when facing more ambiguity. We focus on the case $\alpha < 1/2$, as we always do. Otherwise, $V(\pi; [\alpha, \beta]) = 1/2$ is a constant for all $\pi \in [\alpha, \beta] \subset [1/2, 1]$.

Corollary 2 (i) *Suppose $\pi < \beta$. There exists $\hat{\alpha} \in (0, \min\{\pi, 1/2\})$ such that $V(\pi; [\,\cdot\,, \beta])$ is constant over $(0, \hat{\alpha})$ and is strictly increasing over $(\hat{\alpha}, \min\{\pi, 1/2\})$.*
(ii) *Suppose $\alpha < \pi$ and $\alpha < 1/2$. There exists $\hat{\beta} \in (\pi, 1)$ such that $V(\pi; [\alpha, \,\cdot\,])$ is strictly decreasing over $(\pi, \hat{\beta})$ and is a constant over $(\hat{\beta}, 1)$.*

Thus, there is a limit on the effect of increased one-sided ambiguity on lowering the sender's welfare. Intuitively speaking, if the ambiguity is already very biased towards one side, then a further increase in the ambiguity of that side will not harm the sender. To understand the idea, take the case $\pi < \beta < 1/2$ as an example. When α is close to π, the sender's optimal information structure is $F^{\alpha,\alpha}$ as shown in Proposition 3. As we have discussed, this is because the sender believes that the receiver's private beliefs are more likely to be close to α. By choosing $F^{\alpha,\alpha}$, the sender can guarantee himself high payoffs if the receiver's private belief appears in this region. If α decreases a little to α', the sender still wants to make sure that he can obtain high payoffs when the receiver's private belief is close to α'. For this purpose, the sender has to switch to $F^{\alpha',\alpha'}$ to take care of those private beliefs in $[\alpha', \alpha)$. Hence, the sender gets strictly worse off. If, instead, α is small, then the sender thinks that the receiver's private beliefs are more likely to be close to β. Thus, the optimal information structure for the sender is of the form $F^{x,0}$. By choosing this information structure, the sender guarantees high payoff when the receiver's private belief is close to β and essentially ignores the possibility that the receiver's private belief can be below x. If α moves further away from π to α', the sender should even put more emphasis on the receiver's private beliefs close to β. Thus, the sender should continue ignoring the possibility that the receiver's private belief can be below x. That is, it is optimal for the sender to stick to $F^{x,0}$. By doing so, the sender can obtain exactly the same worst-case payoff he would obtain under

α. Hence, the sender is not worse off. The logic behind increasing β is similar. The difference is that when β is sufficiently large, the sender will stick to $F^{\alpha,\alpha}$ to guarantee high payoffs at private beliefs close to α.

4 Conclusion

In this paper, we studied a robust Bayesian persuasion problem where the receiver has private information source about which the sender has limited knowledge. In this two-by-two environment, we showed that the sender's optimal information structure is an LCPIS. In our model, we assumed that the sender uses only public persuasion. That is, the sender designs a public information disclosure rule independent of the receiver's private information. Yet, like Kolotilin et al. (2017) and Bergemann et al. (2018), we can also think of environments in which the sender uses private persuasion. That is, the sender conditions the disclosure rule on the receiver's reported type. In such environments, the sender must design a mechanism of private persuasion that is incentive-compatible and robust to his knowledge about the distribution of the receiver's private beliefs. This is left for future research.

The new technique we developed in characterizing the optimal information structure makes the most of the convexity of the sender's worst-case payoff function. This method can also be applied to the robust pricing problem studied in Carrasco et al. (2018), in which the seller has only first moment information. We believe that this technique can also be applied to many other contexts in which there are mean-restricted ambiguity and robustness concerns.

Acknowledgements We thank Kim-Sau Chung, Fei Li, Qingmin Liu, George Mailath, and Xianwen Shi. The paper has also benefited from the comments of participants at the 2018 International Conference on Economic Theory and Applications (Chengdu), AMES 2018, the Dynamic Models in Economics workshop at NUS, and seminar participants at Fudan University and Renmin University of China. The helpful comments of five referees and Editor are also gratefully acknowledged. Any remaining errors are our own.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A Proof of Lemma 2

A.1 Mathematical preliminaries

For any real numbers $x < \beta$, let $C[x, \beta]$ be the space of all continuous functions over $[x, \beta]$ endowed with the uniform norm $\|\cdot\|$.

Lemma A.1 *Suppose $0 < x < \beta \leq 1/2$. Let $h \in C[x, \beta]$ be an arbitrary function and let $T : C[x, \beta] \to C[x, \beta]$ be the operator defined as follows: for $f \in C[x, \beta]$,*

$$(Tf)(p) \equiv \frac{h(p) + (1-2p)\int_x^p f(s)\mathrm{d}s}{2p(1-p)}, \quad \forall p \in [x, \beta].$$

Then T is a contraction mapping. As a result, T has a unique fixed point $f^* \in C[x, \beta]$ and therefore $\lim_n T^n g = f^*$ for any $g \in C[x, \beta]$.

Proof Clearly, for every $f \in C[x, \beta]$, the function $Tf : [x, \beta] \to \mathbb{R}$ is well defined and continuous. Thus, $T : C[x, \beta] \to C[x, \beta]$ is well defined. For any $f, g \in C[x, \beta]$, we have

$$\begin{aligned}\|Tf - Tg\| &= \max_{x \le p \le \beta} \left| \frac{1-2p}{2p(1-p)} \int_x^p \big(f(s) - g(s)\big)\mathrm{d}s \right| \\ &\le \|f - g\| \max_{x \le p \le \beta} \frac{(1-2p)(p-x)}{2p(1-p)}.\end{aligned}$$

Because

$$\frac{(1-2p)(p-x)}{2p(1-p)} = 1 - \frac{p(1-x) + (1-p)x}{2p(1-p)} \le 1 - \frac{x(1-x)}{\beta(1-\beta)}, \quad \forall p \in [x, \beta],$$

we know that T is a contraction mapping. □

Lemma A.2 *Suppose $1/2 < \beta < 1$. Let $h \in C[1/2, \beta]$ be an arbitrary function and let $T : C[1/2, \beta] \to C[1/2, \beta]$ be the operator defined as follows: for $f \in C[1/2, \beta]$,*

$$(Tf)(p) \equiv \frac{h(p) - (1-2p)\int_p^\beta f(s)\mathrm{d}s}{2p(1-p)}, \ \forall p \in [1/2, \beta].$$

Then T is a contraction mapping. As a result, T has a unique fixed point $f^* \in C[1/2, \beta]$ and $\lim_n T^n g = f^*$ for any $g \in C[1/2, \beta]$.

Proof The proof is analogical to that of Lemma A.1 and thus is omitted. □

Lemma A.3 *Suppose $0 < x < \beta < 1$. Let $F : [0, \beta] \to \mathbb{R}$ be a function. Suppose F satisfies*

$$2p(1-p)F(p) - (1-2p)\int_0^p F(s)\mathrm{d}s = a(p-x) + b, \ \forall p \in [x, \beta], \tag{A.1}$$

for some constants a and b. Then,

$$F(s) = (1-2x)a + 2b + \frac{(1-2x)b - 2x(1-x)a + \int_0^x F(t)\mathrm{d}t}{2\sqrt{x(1-x)}} \frac{1-2s}{\sqrt{s(1-s)}}, \ \forall s \in [x, \beta]. \tag{A.2}$$

Proof It is straightforward (but tedious) to verify that F in (A.2) satisfies (A.1). It remains to show that it is the unique solution.

When $\beta \le 1/2$, uniqueness directly comes from Lemma A.1 since condition (A.1) can be rewritten as

$$F(p) = \frac{h(p) + (1-2p)\int_x^p F(s)\mathrm{d}s}{2p(1-p)}, \quad \forall p \in [x, \beta],$$

where $h(p) = a(p-x) + b + (1-2p)\int_0^x F(s)\mathrm{d}s$ for $p \in [x, \beta]$.

When $\beta > 1/2$, we can divide the interval $[x, \beta]$ into two sub-intervals $[x, 1/2]$ and $[1/2, \beta]$. Then we can apply Lemma A.1 to $[x, 1/2]$ and Lemma A.2 to $[1/2, \beta]$. □

A.2 Construction of LCPIS's

For ease of exposition in the following analysis, define a function

$$H^{x,b}(s; a) \equiv (1-2x)a + 2b + \frac{(1-2x)b - 2x(1-x)a}{2\sqrt{x(1-x)}} \frac{1-2s}{\sqrt{s(1-s)}}. \tag{A.3}$$

When $1-\beta \leq \alpha < 1/2 < \beta$, let

$$\hat{b} \equiv \frac{\frac{1-\alpha}{1-\beta} - \sqrt{\frac{(1-\alpha)\beta}{\alpha(1-\beta)}}}{\left[\sqrt{\frac{(1-\alpha)\beta}{\alpha(1-\beta)}} - 1\right]^2} \in [0, \alpha). \tag{A.4}$$

We will see that this is just the $\hat{b}$ claimed in part (iii) of Lemma 2. Let $A(\alpha, \beta)$ be the set of initial points defined in Lemma 2:

$$A(\alpha, \beta) = \begin{cases} \{(\alpha, b) \mid 0 < b \leq \alpha\} \bigcup \{(x, 0) \mid \alpha \leq x < \beta\}, & \text{if } \beta < 1/2, \\ \{(\alpha, b) \mid 0 < b \leq \alpha\} \bigcup \{(x, 0) \mid \alpha \leq x \leq 1-\beta\}, & \text{if } \alpha < 1-\beta < 1/2 < \beta, \\ \{(\alpha, b) \mid \hat{b} \leq b \leq \alpha\}, & \text{if } 1-\beta \leq \alpha < 1/2 < \beta. \end{cases}$$

We are now ready to construct the desired dominant LCPIS's. For every $(x, b) \in A(\alpha, \beta)$, let $F^{x,b} : [0, 1] \to \mathbb{R}$ be the function defined as follows:

$$F^{x,b}(s) \equiv \begin{cases} 0, & \text{if } s \in [0, x), \\ H^{x,b}\big(s; a^{x,b}\big), & \text{if } s \in [x, \beta), \\ H^{x,b}\big(\beta; a^{x,b}\big), & \text{if } s \in [\beta, 1), \\ 1 & \text{if } s = 1, \end{cases} \tag{A.5}$$

where $a^{x,b} \in \mathbb{R}$ is the unique solution to the following linear equation:

$$2(1-x)\left[1 - \sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right] a^{x,b} + \left[2 + \frac{(1-2x)\sqrt{1-\beta}}{\sqrt{x(1-x)\beta}}\right] b = 1. \tag{A.6}$$

It is easy to verify that, when $(x, b) = (\alpha, \alpha)$, $F^{x,b}$ in (A.5) becomes $F^{\alpha,\alpha}$ as defined in (8), which we have already known is a dominant LCPIS. In what follows, we proceed to verify that $F^{x,b}$ is a dominant LCPIS for every $(x, b) \in A(\alpha, \beta)$.

A.3 Proof of Lemma 2

We first show that $F^{x,b}$ is an LCPIS for every $(x, b) \in A(\alpha, \beta)$ through Lemmas A.4–A.6.

Lemma A.4 *For every* $(x, b) \in A(\alpha, \beta)$, $F^{x,b} \in \mathcal{F}$.

Proof We must verify that $F^{x,b}$ is a c.d.f. over $[0, 1]$ and its mean is $1/2$. By construction, it is obvious that $F^{x,b}(0) = 0$, $F^{x,b}(1) = 1$, and $F^{x,b}$ is right continuous. To show that $F^{x,b}$ is a c.d.f, it remains to verify that it is nondecreasing. The verification is divided into the following three steps. We will frequently use the fact that $b \le x < \beta$ for all $(x, b) \in A(\alpha, \beta)$.

Step 1: $F^{x,b}(x) \ge 0$.

This directly follows from (A.5) and (A.6):

$$F^{x,b}(x) = H^{x,b}\big(x; a^{x,b}\big) = \frac{b}{2x(1-x)} \ge 0.$$

Step 2: $F^{x,b}$ is nondecreasing over $[x, \beta]$.

Equivalently, we have to show that $H^{x,b}(\,\cdot\,; a^{x,b})$ is nondecreasing over $[x, \beta]$. By (A.6),

$$(1-2x)b - 2x(1-x)a^{x,b} = \frac{b-x}{1-\sqrt{\frac{x(1-\beta)}{(1-x)\beta}}} \le 0.$$

Then, because the mapping $s \mapsto (1-2s)/\sqrt{s(1-s)}$ is strictly decreasing over $(0, 1)$, we know that $H^{x,b}(\,\cdot\,; a^{x,b})$ is nondecreasing over $[x, \beta]$.

Step 3: $F^{x,b}(\beta) \le 1$.

Using (A.5) and (A.6), we can calculate

$$F^{x,b}(\beta) = \frac{1}{2} + \frac{x\sqrt{(1-x)\beta}}{2(1-x)\sqrt{x(1-\beta)}} + \frac{\big(\sqrt{x(1-\beta)} - \sqrt{(1-x)\beta}\big)}{2(1-x)\sqrt{x(1-\beta)}}b.$$

Because $x < \beta$, we know the right-hand side is decreasing in b. When $\alpha < \beta \le 1/2$,

$$F^{x,b}(\beta) \le F^{x,0}(\beta) = \frac{1}{2} + \frac{\sqrt{x\beta}}{2\sqrt{(1-x)(1-\beta)}} \le 1,$$

where the second inequality comes from the fact that both $x \le 1-x$ and $\beta \le 1-\beta$ hold. When $\alpha < 1-\beta < 1/2 < \beta$, we have $x \le 1-\beta$ and

$$\begin{aligned} F^{x,b}(\beta) \le F^{x,0}(\beta) &= \frac{1}{2} + \frac{\sqrt{x\beta}}{2\sqrt{(1-x)(1-\beta)}} \\ \le F^{1-\beta,0}(\beta) &= \frac{1}{2} + \frac{\sqrt{(1-\beta)\beta}}{2\sqrt{\beta(1-\beta)}} = 1. \end{aligned}$$

When $1-\beta < \alpha < 1/2 < \beta$, we have $x = \alpha$ and thus

$$F^{\alpha,b}(\beta) \le F^{\alpha,\hat{b}}(\beta) = 1,$$

where the equality comes from the definition of $\hat{b}$ in (A.4).[22] Steps 1–3 together imply that $F^{x,b}$ is nondecreasing. Hence $F^{x,b}$ is a c.d.f.

Finally,

$$\begin{aligned}\int_{[0,1]} s\mathrm{d}F^{x,b}(s) &= \int_{[x,\beta]} s\mathrm{d}F^{x,b}(s) + \big(1 - F^{x,b}(\beta)\big)\\ &= \beta F^{x,b}(\beta) - \int_x^\beta F^{x,b}(s)\mathrm{d}s + \big(1 - F^{x,b}(\beta)\big)\\ &= 1 - (1-x)\left[1 - \sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right]a^{x,b} - \left[1 + \frac{(1-2x)\sqrt{1-\beta}}{2\sqrt{x(1-x)\beta}}\right]b\\ &= \frac{1}{2},\end{aligned}$$

proving that $F^{x,b} \in \mathcal{F}$. □

Lemma A.5 *For any information structure $F \in \mathcal{F}$, we have*

$$\phi^F(p) = 2p(1-p)F(p) - (1-2p)\int_x^p F(s)\mathrm{d}s, \quad \forall p \in [0,1].$$

Proof This directly comes from the expression for the contingent payoff in (5) in the main text and integration by parts. See, for example, Theorem 21.67 in Hewitt and Stromberg (1965) for integration by parts for Lebesgue–Stieltjes integrals. □

Lemma A.6 *For every $(x, b) \in A(\alpha, \beta)$, $F^{x,b}$ is an LCPIS over $[\alpha, \beta]$ with initial point (x, b) and slope $a^{x,b} > 0$.*

Proof Consider $(x, b) \in A(\alpha, \beta)$. By construction, $F^{x,b}$ satisfies condition (A.2) with, a in (A.2) being replaced by $a^{x,b}$. Thus, $F^{x,b}$ satisfies (A.1) by Lemma A.3. By Lemma A.5, we have

$$\phi^{F^{x,b}}(p) = a^{x,b}(p-x) + b, \quad \forall p \in [x, \beta].$$

By (A.6) and $b \le x < \min\{\beta, 1-\beta\} \le 1/2$, we have

$$a^{x,b} = \frac{1 - \left[2 + \frac{(1-2x)\sqrt{1-\beta}}{\sqrt{x(1-x)\beta}}\right]b}{2(1-x)\left[1 - \sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right]} \ge \frac{1 - \left[2 + \frac{(1-2x)\sqrt{1-\beta}}{\sqrt{x(1-x)\beta}}\right]x}{2(1-x)\left[1 - \sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right]} = \frac{1-2x}{2(1-x)} > 0.$$

[22] In fact, $\hat{b}$ is so constructed that it is the unique solution to $H^{\alpha,\hat{b}}(\beta; a^{\alpha,\hat{b}}) = 1$ for $1-\beta \le \alpha < 1/2 < \beta$.

If $x = \alpha$, then $\phi^{F^{x,b}}$ is a strictly increasing linear function over $[\alpha, \beta]$. It is an LCPIS with initial point (x, b) and slope $a^{x,b}$.

If $x > \alpha$, then $b = 0$ and thus $0 \le \phi^{F^{x,b}}(p) \le \phi^{F^{x,b}}(x) = b = 0$ for all $p < x$, where the second inequality comes from the fact that the contingent payoff function is nondecreasing in private belief [see (5)]. Therefore,

$$\phi^{F^{x,b}}(p) = \max\{0, a^{x,b}(p - x)\}, \quad \forall p \in [\alpha, \beta].$$

This is an LCPIS with initial point (x, b) and slope $a^{x,b}$ too. □

Next, we show that $F^{x,b}$ is the unique dominant LCPIS with initial point $(x, b) \in A(\alpha, \beta)$ through Lemmas A.7–A.9.

Lemma A.7 *Suppose $\beta \in (0, 1)$. If $F \in \mathcal{F}$, then*

$$(1 - \beta)F(\beta) + \int_0^\beta F(s)\mathrm{d}s \le \frac{1}{2}, \tag{A.7}$$

with equality if and only if F does not contain signals in $(\beta, 1)$, i.e., $\lim_{s\uparrow 1} F(s) = F(\beta)$.

Proof We have

$$\begin{aligned}
\frac{1}{2} &= \int_{[0,\beta]} s\mathrm{d}F(s) + \int_{(\beta,1]} s\mathrm{d}F(s) \\
&\le \int_{[0,\beta]} s\mathrm{d}F(s) + 1 - F(\beta) \\
&= \beta F(\beta) - \int_0^\beta F(s)\mathrm{d}s + 1 - F(\beta),
\end{aligned}$$

where the last equality comes from integration by parts. Rearranging yields (A.7). Clearly, this inequality becomes an equality if and only if $\int_{(\beta,1]} s\mathrm{d}F(s) = 1 - F(\beta)$ if and only if F does not contain signals in $(\beta, 1)$. □

Lemma A.8 *For every $(x, b) \in A(\alpha, \beta)$, we have*

$$(1-\beta)H^{x,b}(\beta; a^{x,b}) + \int_x^\beta H^{x,b}(s; a^{x,b})\mathrm{d}s = (1-\beta)F^{x,b}(\beta) + \int_0^\beta F^{x,b}(s)\mathrm{d}s = \frac{1}{2}. \tag{A.8}$$

Proof The first equality comes from our construction of $F^{x,b}$. The second equality comes from $\lim_{s\uparrow 1} F^{x,b}(s) = F^{x,b}(\beta)$ and Lemma A.7. □

Lemma A.9 *Consider an initial point $(x, b) \in A(\alpha, \beta)$. Suppose F is an LCPIS with initial point (x, b) and slope $a > 0$. If $F \ne F^{x,b}$, then $a < a^{x,b}$.*

Proof By Lemmas A.3 and A.5, the fact that F is an LCPIS with initial point (x, b) and slope a implies that F over $[x, \beta]$ satisfies (A.2). Using our construction of the function $H^{x,b}$ in (A.3), we can write

$$F(s) = H^{x,b}(s; a) + \frac{\int_0^x F(t)\mathrm{d}t}{2\sqrt{x(1-x)}} \frac{1-2s}{\sqrt{s(1-s)}}, \quad \forall s \in [x, \beta].$$

Thus, we can calculate

$$\begin{aligned}\int_0^\beta F(s)\mathrm{d}s &= \int_0^x F(s)\mathrm{d}s + \int_x^\beta \left(H^{x,b}(s; a) + \frac{\int_0^x F(t)\mathrm{d}t}{2\sqrt{x(1-x)}} \frac{1-2s}{\sqrt{s(1-s)}} \right) \mathrm{d}s \\ &= \int_x^\beta H^{x,b}(s; a)\mathrm{d}s + \frac{\sqrt{\beta(1-\beta)} \int_0^x F(t)\mathrm{d}t}{\sqrt{x(1-x)}}.\end{aligned}$$

Hence, by Lemma A.5, we have

$$\begin{aligned}\frac{1}{2} &\geq (1-\beta)F(\beta) + \int_0^\beta F(s)\mathrm{d}s \\ &= (1-\beta)H^{x,b}(\beta; a) + \int_x^\beta H^{x,b}(s; a)\mathrm{d}s + \frac{(1-\beta)\int_0^x F(t)\mathrm{d}t}{2\sqrt{x(1-x)}\sqrt{\beta(1-\beta)}},\end{aligned} \tag{A.9}$$

with equality if and only if $\lim_{s\uparrow 1} F(s) = F(\beta)$. By Lemma A.8, we then have

$$\begin{aligned}&(1-\beta)H^{x,b}(\beta; a) + \int_x^\beta H^{x,b}(s; a)\mathrm{d}s + \frac{(1-\beta)\int_0^x F(t)\mathrm{d}t}{2\sqrt{x(1-x)}\sqrt{\beta(1-\beta)}} \\ &\quad \leq (1-\beta)H^{x,b}(\beta; a^{x,b}) + \int_x^\beta H^{x,b}(s; a^{x,b})\mathrm{d}s,\end{aligned} \tag{A.10}$$

with equality if and only if $\lim_{s\uparrow 1} F(s) = F(\beta)$. It is easy to check from (A.3) that $H^{x,b}(s; \cdot)$ is strictly increasing for every $s > x$. Thus, (A.10) directly implies $a \leq a^{x,b}$, since the last term in the first line is nonnegative.

If $a = a^{x,b}$, then (A.10) implies that $\int_0^x F(t)\mathrm{d}t = 0$, which in turn implies that (A.10) holds with equality. These observations together imply

$$F(s) = \begin{cases} 0, & \text{if } s \in [0, x), \\ H^{x,b}(s; a^{x,b}), & \text{if } s \in [x, \beta), \\ H^{x,b}(\beta; a^{x,b}), & \text{if } s \in [\beta, 1), \\ 1, & \text{if } s = 1, \end{cases}$$

which exactly coincides with $F^{x,b}$. Equivalently, if $F \neq F^{x,b}$, we must have $a < a^{x,b}$, completing the proof. □

Proof of Lemma 2 Lemma A.9 has already shown that $F^{x,b}$ is the unique dominant LCPIS with initial point $(x, b) \in A(\alpha, \beta)$, as claimed by Lemma 2.

When $\alpha < 1-\beta < 1/2 < \beta$, we have shown $F^{1-\beta,0}(\beta) = 1$ in the proof of Lemma A.4. Therefore,

$$\begin{aligned}\phi^{F^{1-\beta,0}}(\beta) &= \int_{[0,\beta]} [(1-\beta)s + \beta(1-s)]\mathrm{d}F^{1-\beta,0}(s)\\ &= \int_{[0,1]} [(1-\beta)s + \beta(1-s)]\mathrm{d}F^{1-\beta,0}(s)\\ &= \frac{1-\beta}{2} + \frac{\beta}{2}\\ &= \frac{1}{2},\end{aligned}$$

where the third equality comes from $\int_{[0,1]} s\mathrm{d}F^{1-\beta,0}(s) = 1/2$.

Similarly, when $1-\beta \le \alpha < 1/2 < \beta$, we have shown that $F^{\alpha,\hat{b}}(\beta) = 1$ in the proof of Lemma A.4. As above, we can show that $\phi^{F^{\alpha,\hat{b}}}(\beta) = 1/2$. This completes the proof. □

Appendix B Proof of Lemma 3

B.1 Mathematical preliminaries

Lemma B.1 *Assume $\beta \le 1/2$. Let $F \in \mathcal{F}$ be an arbitrary information structure. Assume that $\ell(p) \equiv a(p-x)+b \le \phi^F(p)$ for all $p \in [\alpha, \beta]$, where $(x, b) \in A(\alpha, \beta)$ and $a > 0$. Then $F(s) \ge H^{x,b}(s; a)$ for $s \in [x, \beta]$.*

Proof Define a sequence of continuous functions $\{G_n\}_{n\ge 0}$ over $[x, \beta]$ as follows:

$$G_0(p) \equiv 0, \ \ \forall p \in [x, \beta],$$

and for $n \ge 1$,

$$G_n(p) \equiv \frac{\ell(p) + (1-2p)\int_x^p G_{n-1}(s)\mathrm{d}s}{2p(1-p)}, \quad \forall p \in [x, \beta].$$

By Lemma A.1, $\{G_n\}_{n\ge 0}$ uniformly converges to a function G that satisfies

$$G(p) = \frac{\ell(p) + (1-2p)\int_x^p G(s)\mathrm{d}s}{2p(1-p)}, \quad \forall p \in [x, \beta].$$

By Lemma A.3, $G(p) = H^{x,b}(p; a)$ for $p \in [x, \beta]$.

It remains to show that $F \ge G$ over $[x, \beta]$. Clearly, $F \ge G_0$ over $[x, \beta]$. Suppose $F \ge G_{n-1}$ over $[x, \beta]$ for some $n \ge 1$. Because $\phi^F \ge \ell$ over $[\alpha, \beta]$, we know that

$$\phi^F(p) = 2p(1-p)F(p) - (1-2p)\int_0^p F(s)\mathrm{d}s \ge \ell(p), \quad \forall p \in [x, \beta].$$

This implies, for all $p \in [x, \beta]$,

$$\begin{aligned} 2p(1-p)F(p) &\geq \ell(p) + (1-2p)\int_0^p F(s)\mathrm{d}s \\ &\geq \ell(p) + (1-2p)\int_x^p F(s)\mathrm{d}s \\ &\geq \ell(p) + (1-2p)\int_x^p G_{n-1}(s)\mathrm{d}s, \end{aligned}$$

where the second inequality comes from $p \leq \beta \leq 1/2$ and $F \geq 0$. The last inequality comes from the induction hypothesis. Rearranging yields

$$F(p) \geq \frac{\ell(p) + (1-2p)\int_x^p G_{n-1}(s)\mathrm{d}s}{2p(1-p)} = G_n(p), \quad \forall p \in [x, \beta].$$

Therefore, $F \geq G_n$ over $[x, \beta]$ for all n, implying $F \geq G = H^{x,b}(\,\cdot\,; a)$ over $[x, \beta]$. This completes the proof. □

Lemma B.2 *Assume $1/2 < \beta < 1$. Let $F \in \mathcal{F}$ be an arbitrary information structure. Assume that $\ell(p) \equiv a(p-x) + b \leq \phi^F(p)$ for all $p \in [\alpha, \beta]$, where $(x, b) \in A(\alpha, \beta)$ and $a > 0$. Then $F(s) \geq H^{x,b}(s; a)$ for $s \in [x, 1/2]$ and $F(s) \geq G(s)$ for $s \in [1/2, \beta]$, where*

$$G(s) \equiv (1-2x)a + 2b - \frac{(2(1-x)a + 2b - 1)\beta}{2\sqrt{\beta(1-\beta)}} \frac{1-2s}{\sqrt{s(1-s)}}, \quad \forall s \in [1/2, \beta]. \tag{B.1}$$

Proof The first half of the claim, $F \geq H^{x,b}(\,\cdot\,; a)$ over $[x, 1/2]$, can be obtained by letting $\beta = 1/2$ in Lemma B.1. We only prove the second half of the claim, $F \geq G$ over $[1/2, \beta]$. The proof is similar to that of Lemma B.1. Define $\ell' : [1/2, \beta] \to \mathbb{R}$ as

$$\ell'(p) = \big(a - 2\beta + 2a(\beta - \alpha) + 2b\big)p - a\beta + \beta, \ \forall p \in [1/2, \beta].$$

Define a sequence of continuous functions $\{G_n\}_{n\geq 0}$ over $[1/2, \beta]$ as follows:

$$G_0(p) \equiv 0, \ \forall p \in [1/2, \beta],$$

and

$$G_n(p) \equiv \frac{\ell'(p) - (1-2p)\int_p^\beta G_{n-1}(s)\mathrm{d}s}{2p(1-p)}, \quad \forall p \in [1/2, \beta].$$

By Lemmas A.2 and A.3, $\{G_n\}_{n\geq 0}$ uniformly converges to G in (B.1).

It remains to show that $F \geq G$ over $[1/2, \beta]$. As in the proof of Lemma B.1, it suffices to show that $F \geq G_n$ over $[1/2, \beta]$ for all $n \geq 1$. Clearly $F \geq G_0$ over

$[1/2, \beta]$. Assume $F \geq G_{n-1}$ over $[1/2, \beta]$ for some $n \geq 1$. From (5), we can rewrite ϕ^F as

$$\begin{aligned}\phi^F(p) &= pF(p) + (1-2p)\left(1/2 - \int_{(p,1]} s\mathrm{d}F(s)\right) \\ &= 2p(1-p)F(p) + (1-2p)\int_p^1 F(s)\mathrm{d}s - \frac{1}{2}(1-2p),\end{aligned}$$

where the second equality follows from integration by parts again. Because $\ell \leq \phi^F$, it follows that for all $p \in [1/2, \beta]$,

$$\begin{aligned}&2p(1-p)F(p) \\ &\quad\geq \ell(p) + \frac{1}{2}(1-2p) - (1-2p)\int_p^1 F(s)\mathrm{d}s \\ &\quad\geq \ell(p) + \frac{1}{2}(1-2p) - (1-2p)\int_p^\beta F(s)\mathrm{d}s - (1-2p)(1-\beta)F(\beta), \qquad \text{(B.2)}\end{aligned}$$

where the second inequality follows from $1 - 2p \leq 0$ for $p \in [1/2, \beta]$ and $\int_p^1 F(s)\mathrm{d}s \geq \int_p^\beta F(s)\mathrm{d}s + (1-\beta)F(\beta)$. Letting $p = \beta$ in the above inequality yields

$$(1-\beta)F(\beta) \geq \ell(\beta) + \frac{1}{2}(1-2\beta).$$

Plugging this inequality back into (B.2), it follows that for $p \in [1/2, \beta]$,

$$\begin{aligned}2p(1-p)F(p) &\geq \ell(p) - (1-2p)\int_p^\beta F(s)\mathrm{d}s - (1-2p)\Big(\ell(\beta) - \beta\Big) \\ &= \ell'(p) - (1-2p)\int_p^\beta F(s)\mathrm{d}s \\ &\geq \ell'(p) - (1-2p)\int_p^\beta G_{n-1}(s)\mathrm{d}s,\end{aligned}$$

where the last inequality comes from the induction hypothesis and $1 - 2p \leq 0$ for $p \in [1/2, \beta]$. Equivalently, we have

$$F(p) \geq \frac{\ell'(p) - (1-2p)\int_p^\beta G_{n-1}(s)\mathrm{d}s}{2p(1-p)} = G_n(p), \quad \forall p \in [1/2, \beta].$$

Therefore, $F \geq G_n$ for all n over $[1/2, \beta]$, implying $F \geq G$ over $[1/2, \beta]$. This completes the proof. □

B.2 Proof of Lemma 3

We are now ready to prove Lemma 3. We divide the whole proof into two proofs. The first deals with the case $\beta \le 1/2$ and the second deals with the case $\beta > 1/2$.

Proof of Lemma 3 for $\beta \le 1/2$. Assume $\beta \le 1/2$. Suppose $F \in \mathcal{F}$ is an information structure. Assume linear function $\ell \le \phi^F$ over $[\alpha, \beta]$. If the slope of ℓ is non-positive, then $\ell \le \phi^{F^{\alpha,\alpha}}$ because $\ell(\alpha) \le \phi^F(\alpha) \le \phi^{F^{\alpha,\alpha}}(\alpha)$ and the slope of $\phi^{F^{\alpha,\alpha}}$ is positive. Assume the slope is positive. If $\ell(\beta) \le 0$, then $\ell \le 0 \le \phi^{F^{\alpha,\alpha}}$. Thus, in what follows, we assume positive slope and $\ell(\beta) > 0$.

Then, there must exist $(x, b) \in A(\alpha, \beta)$ such that ℓ can be written as $\ell(p) = a(p - x) + b$ for $p \in [\alpha, \beta]$, where $a > 0$ is its slope. By Lemma B.1, $F(s) \ge H^{x,b}(s; a)$ for $s \in [x, \beta]$. By Lemmas A.7 and A.8, we have

$$
\begin{aligned}
&(1-\beta)H^{x,b}\big(\beta; a^{x,b}\big) + \int_x^\beta H^{x,b}\big(s; a^{x,b}\big)\mathrm{d}s \\
&\ge (1-\beta)F(\beta) + \int_0^\beta F(s)\mathrm{d}s \\
&\ge (1-\beta)F(\beta) + \int_x^\beta F(s)\mathrm{d}s \\
&\ge (1-\beta)H^{x,b}(\beta; a) + \int_x^\beta H^{x,b}(s; a)\mathrm{d}s.
\end{aligned}
$$

Because $H^{x,b}(s; \cdot)$ is strictly increasing when $s > x$, we immediately know that $a \le a^{x,b}$. Therefore, $\ell \le \phi^{F^{x,b}}$. $\square$

Proof of Lemma 3 for $1/2 < \beta < 1$. The idea is similar to the previous proof. We can focus on linear ℓ with positive slope $a > 0$ and $\ell(\beta) > 0$. Then ℓ must intersect the set $\{(\alpha, b)|0 < b \le a\} \cup \{(x, 0)|\alpha \le x < \beta\}$ at some point (x, b) and ℓ can be written as $\ell(p) = a(p - x) + b$ for some $a > 0$.

Suppose $(x, b) \notin A(\alpha, \beta)$. If $\alpha < 1 - \beta$, this can occur only if $x > 1 - \beta$ and $b = 0$. Then

$$
\ell(x) = 0 < \phi^{F^{1-\beta,0}}(x)
$$

and

$$
\ell(\beta) \le \phi^F(\beta) \le \frac{1}{2} = \phi^{F^{1-\beta,0}}(\beta),
$$

together implies $\ell \le \phi^{F^{1-\beta,0}}$ over $[x, \beta]$. Because $\ell < 0 \le \phi^{F^{1-\beta,0}}$ over $[\alpha, x)$, we know $\ell \le \phi^{F^{1-\beta,0}}$ over $[\alpha, \beta]$. Similarly, if $1-\beta \le \alpha$, $(x, b) \notin A(\alpha, \beta)$ implies either $(x, b) = (\alpha, b)$ for some $b < \hat{b}$ or $(x, b) = (x, 0)$ for some $x > \alpha$. In both cases, we can show $\ell \le \phi^{F^{\alpha,\hat{b}}}$ over $[\alpha, \beta]$.

Suppose $(x, b) \in A(\alpha, \beta)$. By Lemmas A.7 and B.2, we know

$$\begin{aligned}\frac{1}{2} \geq &(1-\beta)F(\beta) + \int_0^\beta F(s)\mathrm{d}s \\ \geq &(1-\beta)G(\beta) + \int_x^{\frac{1}{2}} H^{x,b}(s;a)\mathrm{d}s + \int_{\frac{1}{2}}^\beta G(s)\mathrm{d}s \\ = &(1-x)\left[1-\sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right]a + \frac{1}{2}\left[2+\frac{(1-2x)\sqrt{1-\beta}}{\sqrt{x(1-x)\beta}}\right]b, \end{aligned} \tag{B.3}$$

where G is defined in (B.1). Comparing (B.3) and (A.6), we immediately know $a \leq a^{x,b}$, implying $\ell \leq \phi^{F^{x,b}}$. This completes the proof. □

Appendix C Proof of Proposition 2

The next lemma proves that we can approximate the worst-case payoff function from an information structure by the linear functions below its contingent payoff function. This lemma, when combined with Lemma 3, will prove our main result, Proposition 2.

Lemma C.1 *Suppose $F \in \mathcal{F}$. For every $p \in (\alpha, \beta)$, we have*[23]

$$\mathrm{co}_{[\alpha,\beta]}\phi^F(p) = \sup_{\substack{\textit{linear } \ell:[\alpha,\beta]\to\mathbb{R}, \\ \ell\leq\phi^F \textit{ over } [\alpha,\beta]}} \ell(p).$$

Proof Because $\mathrm{co}_{[\alpha,\beta]}\phi^F$ is convex by definition and because $p \in (\alpha, \beta)$ is interior, the left derivative of $\mathrm{co}_{[\alpha,\beta]}\phi^F$ at p exists. Denote this derivative by $a \in \mathbb{R}$. Define $\ell^p(x) \equiv a(x-p) + \mathrm{co}_{[\alpha,\beta]}\phi^F(p)$ for $x \in [\alpha, \beta]$. For any $x \in [\alpha, \beta]$, we have $\phi^F(x) \geq \mathrm{co}_{[\alpha,\beta]}\phi^F(x) \geq \ell^p(x)$. In other words, $\ell \leq \phi^F$ over $[\alpha, \beta]$. Then, we have

$$\mathrm{co}_{[\alpha,\beta]}\phi^F(p) = \ell^p(p) \leq \sup_{\substack{\text{linear } \ell:[\alpha,\beta]\to\mathbb{R}, \\ \ell\leq\phi^F \text{ over } [\alpha,\beta]}} \ell(p).$$

The other direction of the above inequality is obvious. □

We are now ready to prove our main result, Proposition 2.

Proof of Proposition 2 For any $F \in \mathcal{F}$, Lemmas 3 and C.1 together imply

$$\mathrm{co}_{[\alpha,\beta]}\phi^F(\pi) = \sup_{\substack{\text{linear } \ell:[\alpha,\beta]\to\mathbb{R} \\ \ell\leq\phi^F \text{ over } [\alpha,\beta]}} \ell(\pi) \leq \max_{(x,b)\in A} \phi^{F^{x,b}}(\pi), \ \forall \pi \in (\alpha, \beta).$$

[23] If ϕ^F itself is convex, this result is standard. Moreover, this result holds more generally if ϕ^F is replaced by any function f over $[\alpha, \beta]$ that is bounded from below.

Because

$$V(\pi;[\alpha,\beta]) = \max_{F\in\mathcal{F}} \mathrm{co}_{[\alpha,\beta]}\phi^F(\pi)$$

by definition and

$$\max_{(x,b)\in A} \phi^{F^{x,b}}(\pi) \le \max_{F\in\mathcal{F}} \phi^F(\pi)$$
$$\text{s.t. } F \text{ is an LCPIS over } [\alpha,\beta],$$

we know

$$V(\pi;[\alpha,\beta]) \le \max_{F\in\mathcal{F}} \phi^F(\pi)$$
$$\text{s.t. } F \text{ is an LCPIS over } [\alpha,\beta].$$

Since the other direction of the inequality is obvious, we obtain our main result as desired. □

Appendix D Proofs of Propositions 3 and 4

Proof of Proposition 3 We prove Proposition 3 for the case $\alpha < \beta \le 1/2$. The other case is similar.

For notational simplicity, define $f(x) = 2(1-x)\left[1-\sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right]$ and $g(x) = \left[2+\frac{(1-2x)\sqrt{1-\beta}}{\sqrt{x(1-x)\beta}}\right]$ to be the coefficients of $a^{x,b}$ and b in (A.6), respectively. That is, (A.6) can be written as

$$f(x)a^{x,b} + g(x)b = 1.$$

We suppress β from notation f and g for simplicity too. The whole proof is divided into several steps.

Step 1: $\{\phi^{F^{\alpha,b}}\}_{b\in[0,\alpha]}$ all intersect at

$$\hat{\pi} \equiv \alpha + \frac{f(\alpha)}{g(\alpha)} \in (\alpha,\beta). \tag{D.1}$$

Pick any $b \in [0,\alpha)$. Consider the intersection π of $\phi^{F^{\alpha,\alpha}}$ and $\phi^{F^{\alpha,b}}$:

$$\alpha^{\alpha,\alpha}(\pi-\alpha)+\alpha = \alpha^{\alpha,b}(\pi-\alpha)+b.$$

Because $a^{\alpha,\alpha} = (1 - g(\alpha)\alpha)/f(\alpha)$ and $a^{\alpha,b} = (1 - g(\alpha)b)/f(\alpha)$, it is easy to see

$$\pi = \alpha + \frac{f(\alpha)}{g(\alpha)}, \tag{D.2}$$

which is independent of b.

Step 2: $\phi^{F^{\alpha,0}}(\hat{\pi}) > \phi^{F^{x,0}}(\hat{\pi})$ for $x \in (\alpha, \beta)$.

It suffices to consider $x \in (\alpha, \hat{\pi})$. Let $(\tilde{a}, \tilde{b})$ be the solution to the following system of linear equations:

$$\begin{aligned} f(\alpha)\tilde{a} + g(\alpha)\tilde{b} &= 1, \\ \tilde{a}(x - \alpha) + \tilde{b} &= 0. \end{aligned} \tag{D.3}$$

From the same argument as in Step 1, we know $\ell(p) \equiv \tilde{a}(p-\alpha)+\tilde{b}$ also intersects $\{\phi^{F^{\alpha,b}}\}_{b\in[0,\alpha]}$ at $\hat{\pi}$, which implies $\phi^{F^{\alpha,0}}(\hat{\pi}) = \ell(\hat{\pi})$. Because $x < \hat{\pi}$, we immediately know that $\tilde{b} < 0$.

Because ℓ and $\phi^{F^{x,0}}$ intersect at $(x, 0)$, to show $\ell(\hat{\pi}) > \phi^{F^{x,0}}(\hat{\pi})$, it suffices to show $\ell(\alpha) < a^{x,0}(\alpha - x)$, or equivalently $\tilde{b} < a^{x,0}(\alpha - x)$. From (D.3), $\tilde{b} = (f(\alpha)/(\alpha - x) + g(\alpha))^{-1}$. From (A.6), $a^{x,0} = 1/f(x)$. Therefore, to show $\tilde{b} < a^{x,0}(\alpha - x)$, it is equivalent to showing that

$$\frac{1}{\frac{f(\alpha)}{\alpha - x} + g(\alpha)} < \frac{\alpha - x}{f(x)}.$$

Rearranging, it is equivalent to

$$f(x) > f(\alpha) - g(\alpha)(x - \alpha).$$

For this inequality, it suffices to show $f'(\tilde{x}) > -g(\alpha)$ for $\tilde{x} > \alpha$. This is indeed true because

$$f'(\tilde{x}) = -2 - \frac{(1 - 2\tilde{x})\sqrt{1-\beta}}{\sqrt{\tilde{x}(1-\tilde{x})\beta}} = -g(\tilde{x}) > -g(\alpha).$$

Step 3: $F^{\alpha,\alpha}$ is the unique solution to (11) for $\pi \in (\alpha, \hat{\pi})$, and $\{F^{\alpha,b}\}_{b\in[0,\alpha]}$ are solutions for $\pi = \hat{\pi}$.

This is a direct implication of Steps 1 and 2.

Step 4: for every $\pi \in (\hat{\pi}, \beta)$, there is a unique solution $x^*(\pi) \in (\alpha, \beta)$ to

$$\max_{x\in[\alpha,\beta)} \phi^{F^{x,0}}(\pi). \tag{D.4}$$

We can focus on $x \in [\alpha, \pi)$, because $\phi^{F^{x,0}}(\pi) = 0$ when $x \geq \pi$. Because $a^{x,0} = 1/f(x)$, this maximization problem can be written as

$$\max_{x \in [\alpha, \pi)} \frac{\pi - x}{f(x)}.$$

It can be verified that the objective function is strictly concave. Ignoring the constraint $x \in [\alpha, \pi)$ and using $f'(x) = -g(x)$, the first order condition can be written as

$$x^*(\pi) + \frac{f(x^*(\pi))}{g(x^*(\pi))} = \pi.$$

Clearly $x^*(\pi) < \pi$. Because $\pi > \hat{\pi}$, applying (D.1) yields

$$x^*(\pi) + \frac{f(x^*(\pi))}{g(x^*(\pi))} > \alpha + \frac{f(\alpha)}{g(\alpha)}.$$

Because $x + f(x)/g(x)$ is strictly increasing, we know $x^*(\pi) > \alpha$. Therefore, $x^*(\pi) \in (\alpha, \beta)$ is the unique solution to (D.4). Moreover, $x^*(\pi)$ is strictly increasing.

Step 5: $F^{x^*(\pi),0}$ is the unique solution to (11) for $\pi \in (\hat{\pi}, \beta)$.

By Step 1, for every $b \in (0, \alpha]$, $\phi^{\alpha,b}(\pi) < \phi^{\alpha,0}(\pi)$ for all $\pi \in (\hat{\pi}, \beta)$. So $\{F^{\alpha,b}\}_{b \in (0,\alpha]}$ are not optimal for $\pi \in (\hat{\pi}, \beta)$. Then $F^{x^*(\pi),0}$ is the unique solution to (11) for $\pi \in (\hat{\pi}, \beta)$ by Step 4. □

Proof of Proposition 4 As in the proof of Proposition 3, we can also show that $\{\phi^{F^{\alpha,b}}\}_{\hat{b} \leq b \leq \alpha}$ intersect at some $\hat{\pi} \in (\alpha, \beta)$. The desired result then immediately follows. □

Appendix E Proofs of Corollaries 1 and 2

E.1 Proof of Corollary 1

Proof of Corollary 1 We can directly verify from Propositions 3 and 4. Here we provide a simpler proof.

Fix $\pi < \beta$ first. Consider $\{\tilde{\alpha}_n\}_{n \geq 1}$ such that $\tilde{\alpha}_n \uparrow \pi$. If $\pi \leq 1/2$, let F_n be the KG solution for prior $\tilde{\alpha}_n$, $F^{\tilde{\alpha}_n, \tilde{\alpha}_n}$. Then

$$\tilde{\alpha}_n = \phi^{F^{\tilde{\alpha}_n, \tilde{\alpha}_n}}(\tilde{\alpha}_n) \leq \phi^{F^{\tilde{\alpha}_n, \tilde{\alpha}_n}}(\pi) \leq V(\pi; [\tilde{\alpha}_n, \beta]) \leq \max_{F \in \mathcal{F}} \phi^F(\pi) = \pi, \ \forall n.$$

Letting $n \to \infty$ yields the desired result. If $\pi > 1/2$, let F_n be the information structure that is completely uninformative for all n. There exists N such that $\tilde{\alpha}_n > 1/2$ for all $n \geq N$. Then $V(\pi; [\tilde{\alpha}_n, \beta]) = \phi^{F_n}(\pi) = 1/2 = \max_{F \in \mathcal{F}} \phi^F(\pi)$ for all $n \geq N$.

Fix $\alpha < \pi$. Consider $\{\tilde{\beta}_n\}_{n\geq 1}$ such that $\tilde{\beta} \downarrow \pi$. Suppose $\pi \leq 1/2$ first. There exists N such that $x_n \equiv \pi - \sqrt{\tilde{\beta}_n - \pi} \in (\alpha, \pi)$ for $n \geq N$. Note $x_n \uparrow \pi$ and $(\pi - x_n)/(\tilde{\beta}_n - x_n) \uparrow 1$. Let F_n be the KG solution for belief x_n for $n \geq N$. Then, for all $n \geq N$,

$$\mathrm{co}_{[\alpha,\tilde{\beta}_n]}\phi^{F_n}(\pi) = \frac{\phi^{F_n}(\tilde{\beta}_n)}{\tilde{\beta}_n - x_n}(\pi - x_n) = \left[\frac{1-2x_n}{2(1-x_n)}(\tilde{\beta}_n - x_n) + x_n\right]\frac{\pi - x_n}{\tilde{\beta}_n - x_n}.$$

Since $\mathrm{co}_{[\alpha,\tilde{\beta}_n]}\phi^{F_n}(\pi) \leq V(\pi; [\alpha, \tilde{\beta}_n])$ for all $n \geq N$, we have

$$\pi = \lim_{n\to\infty} \mathrm{co}_{[\alpha,\tilde{\beta}_n]}\phi^{F_n}(\pi) \leq \lim_{n\to\infty} V(\pi; [\alpha, \tilde{\beta}_n]) \leq \max_{F\in\mathcal{F}} \phi^F(\pi) = \pi,$$

as desired. Suppose $\pi > 1/2$. Let F_n be the completely uninformative information structure for all n. Then, for all n,

$$\mathrm{co}_{[\alpha,\tilde{\beta}_n]}\phi^{F_n}(\pi) = \begin{cases} \frac{\frac{1}{2}}{\tilde{\beta}_n - 1/2}\left(\pi - \frac{1}{2}\right), & \text{if } \alpha < \frac{1}{2}, \\ \frac{1}{2}, & \text{if } \alpha \geq \frac{1}{2}. \end{cases}$$

Therefore,

$$\frac{1}{2} = \lim_{n\to\infty} \mathrm{co}_{[\alpha,\tilde{\beta}_n]}\phi^{F_n}(\pi) \leq \lim_{n\to\infty} V(\pi; [\alpha, \tilde{\beta}_n]) \leq \max_{F\in\mathcal{F}} \phi^F(\pi) = \frac{1}{2},$$

as desired. □

E.2 Proof of Corollary 2

For Corollary 2, it is crucial to understand how the optimal information structure changes as α and β change. To indicate the dependence and to avoid confusion, in what follows, we will explicit write out α and β in the notation of some of the key variables we identified in the previous sections.

In particular, let $\hat{\pi}(\alpha, \beta)$ be the prior cut-off identified in Propositions 3 and 4. Its formula is given in (D.2). For the cases $\alpha < \beta \leq 1/2$ and $\alpha < 1 - \beta < 1/2 < \beta$, we will frequently refer to the dominant LCPIS of the form $F^{x,0}$ for $(x, 0) \in A(\alpha, \beta)$. Since such LCPIS depends on the value of β, we will explicitly write $F_\beta^{x,0}$. We also write $a_\beta^{x,0}$ to denote the slope of $\phi^{F_\beta^{x,0}}$. Similarly, let $\hat{b}(\alpha, \beta)$ be the $\hat{b}$ identified in Lemma 2 for the case $1 - \beta \leq \alpha < 1/2 < \beta$. Its formula is given in (A.4). In this case, we also write $F_\beta^{\alpha,\hat{b}(\alpha,\beta)}$ and $a_\beta^{\alpha,\hat{b}(\alpha,\beta)}$ to denote the associated LCPIS and the corresponding slope. An exception is the KG solution for belief α. Because it is independent of β, we will still write it as $F^{\alpha,\alpha}$.

To prove Corollary 2, we need the following two lemmas.

Lemma E.1 *For any β, $\hat{\pi}(\cdot, \beta)$ is strictly increasing over* $(0, \min\{\beta, 1/2\})$. *For any $\alpha < 1/2$, $\hat{\pi}(\alpha, \cdot)$ is strictly increasing over* $(\alpha, 1)$.

Proof This is because

$$\frac{\partial \hat{\pi}(\alpha,\beta)}{\partial \alpha} = \frac{(1-\alpha)\sqrt{\alpha(1-\alpha)}}{\sqrt{\beta(1-\beta)}\left[\sqrt{1-\beta} - 2a\sqrt{1-\beta} + 2\sqrt{\alpha(1-\alpha)\beta}\right]^2} > 0,$$

$$\frac{\partial \hat{\pi}(\alpha,\beta)}{\partial \beta} = \frac{\sqrt{\beta(1-\beta)}\left[\sqrt{\frac{1-\alpha}{\alpha}} - \sqrt{\frac{1-\beta}{\beta}}\right]}{\left[\sqrt{1-\beta} - 2a\sqrt{1-\beta} + 2\sqrt{\alpha(1-\alpha)\beta}\right]^2} > 0.$$

□

Lemma E.2 *For any* $\beta > 1/2$, *the function* $\alpha \mapsto a_\beta^{\alpha,\hat{b}(\alpha,\beta)}$ *is strictly decreasing over* $(1-\beta, 1/2)$. *For any* $\alpha < 1/2$, *the function* $\beta \mapsto a_\beta^{\alpha,\hat{b}(\alpha,\beta)}$ *is strictly decreasing over* $(1-\alpha, 1)$.

Proof This is because

$$\frac{\partial a_\beta^{\alpha,\hat{b}(\alpha,\beta)}}{\partial \alpha} = \frac{\sqrt{\beta(1-\beta)}\left[\sqrt{\frac{1-\alpha}{\alpha}} - \sqrt{\frac{\beta}{1-\beta}}\right]}{2\alpha^2(1-\beta)^2\sqrt{\frac{(1-\alpha)\beta}{\alpha(1-\beta)}}\left[\sqrt{\frac{(1-\alpha)\beta}{\alpha(1-\beta)}} - 1\right]^3} < 0,$$

$$\frac{\partial a_\beta^{\alpha,\hat{b}(\alpha,\beta)}}{\partial \beta} = \frac{-(1-2\alpha)\sqrt{\alpha(1-\alpha)}\left[\sqrt{\frac{1-\alpha}{\alpha}} + \sqrt{\frac{\beta}{1-\beta}}\right]}{2\alpha^2(1-\beta)^2\sqrt{\frac{(1-\alpha)\beta}{\alpha(1-\beta)}}\left[\sqrt{\frac{(1-\alpha)\beta}{\alpha(1-\beta)}} - 1\right]^3} < 0.$$

□

We divide the proof of Corollary 2 into two parts for clarity.

Proof of Part (i) of Corollary 2 Fix $\pi < \beta$. We discuss four cases.

Case 1: $\beta > 1/2$ and $\pi \geq \hat{\pi}(1/2, \beta)$.

Because $\hat{\pi}(1/2, \beta) > 1/2$, we know $\pi > 1/2$ in this case. Because $\hat{\pi}(\cdot, \beta)$ is strictly increasing over $(0, 1/2)$, we know $\pi > \hat{\pi}(\alpha, \beta)$ for all $\alpha \in (0, 1/2)$. From Propositions 3 and 4, we have

$$V(\pi; [\alpha,\beta]) = \begin{cases} \max\limits_{x\in[\alpha,1-\beta]} \phi^{F_\beta^{x,0}}(\pi), & \text{if } \alpha \in (0, 1-\beta), \\ \phi^{F_\beta^{\alpha,\hat{b}(\alpha,\beta)}}(\pi), & \text{if } \alpha \in [1-\beta, 1/2). \end{cases}$$

For $\alpha \in [1-\beta, 1/2)$, we can write

$$\phi^{F_\beta^{\alpha,\hat{b}(\alpha,\beta)}}(\pi) = a_\beta^{\alpha,\hat{b}(\alpha,\beta)}(\pi - \beta) + \frac{1}{2},$$

since $\phi^{F_\beta^{\alpha,\hat{b}(\alpha,\beta)}}(\beta) = 1/2$ by Lemma 2. By Lemma E.2, $a_\beta^{\alpha,\hat{b}(\alpha,\beta)}$ is strictly decreasing in α over $(1-\beta, 1/2)$. Therefore, $\phi^{F_\beta^{\alpha,\hat{b}(\alpha,\beta)}}(\pi)$ is strictly increasing over $(1-\beta, 1/2)$, so is $V(\pi; [\cdot, \beta])$.

At $\alpha = 1-\beta$, the optimal information structure is $F_\beta^{1-\beta,\hat{b}(1-\beta,\beta)} = F_\beta^{1-\beta,0}$ since $\hat{b}(1-\beta, \beta) = 0$. Therefore, for all $\alpha \in (0, 1-\beta)$,

$$V(\pi; [\alpha, \beta]) = \max_{x\in[\alpha,1-\beta]} \phi^{F_\beta^{x,0}}(\pi) \geq \phi^{F_\beta^{1-\beta,0}}(\pi) = V(\pi; [1-\beta, \beta]) \geq V(\pi; [\alpha, \beta]),$$

where the last inequality comes from the fact that $V(\pi; [\cdot, \beta])$ is weakly increasing. Therefore, $V(\pi; [\alpha, \beta]) = V(\pi; [1-\beta, \beta])$ for all $\alpha < 1-\beta$.

Let $\hat{\alpha} = 1-\beta$. We have shown that $V(\pi; [\cdot, \beta])$ is constant over $(0, \hat{\alpha})$ and is strictly increasing over $(\hat{\alpha}, 1/2)$.

Case 2: $\beta > 1/2$ and $\hat{\pi}(1-\beta, \beta) \leq \pi < \hat{\pi}(1/2, \beta)$.

Because $\hat{\pi}(1-\beta, \beta) = (3\beta-1)/(4\beta-1) \geq 1/2$, we know $\pi \geq 1/2$ too in this case. Because $\hat{\pi}(\cdot, \beta)$ is strictly increasing over $(0, 1/2)$, there exists a unique $\tilde{\alpha} \in [1-\beta, 1/2)$ such that $\hat{\pi}(\tilde{\alpha}, \beta) = \pi$. Therefore, by Propositions 3 and 4, we have

$$V(\pi; [\alpha, \beta]) = \begin{cases} \max_{x\in[\alpha,1-\beta]} \phi^{F_\beta^{x,0}}(\pi), & \text{if } \alpha \in (0, 1-\beta), \\ \phi^{F_\beta^{\alpha,\hat{b}(\alpha,\beta)}}(\pi), & \text{if } \alpha \in [1-\beta, \tilde{\alpha}), \\ \phi^{F^{\alpha,\alpha}}(\pi), & \text{if } \alpha \in [\tilde{\alpha}, 1/2). \end{cases}$$

Because $\phi^{F^{\alpha,\alpha}}$ can be written as

$$\phi^{F^{\alpha,\alpha}}(\pi) = \frac{\frac{1}{2}-\alpha}{1-\alpha}(\pi - 1) + \frac{1}{2},$$

we know $\phi^{F^{\alpha,\alpha}}(\pi)$ is strictly increasing over $[\tilde{\alpha}, 1/2)$, so is $V(\pi; [\cdot, \beta])$. Moreover, we can use a similar argument as the previous case to show that $V(\pi; [\cdot, \beta])$ is constant over $(0, 1-\beta)$ and is strictly increasing over $[1-\beta, \tilde{\alpha})$.

Again, let $\hat{\alpha} = 1-\beta$. We have shown that $V(\pi; [\cdot, \beta])$ is constant over $(0, \hat{\alpha})$ and is strictly increasing over $(\hat{\alpha}, 1/2)$.

Case 3: $\beta > 1/2$ and $\pi < \hat{\pi}(1-\beta, \beta)$.

Similarly as Case 2, there exists a unique $\hat{\alpha} < \min\{\pi, 1-\beta\}$ such that $\hat{\pi}(\hat{\alpha}, \beta) = \pi$. By Propositions 3 and 4, we know

$$V(\pi; [\alpha, \beta]) = \begin{cases} \max_{x\in[\alpha,1-\beta]} \phi^{F_\beta^{x,0}}(\pi), & \text{if } \alpha \in (0, \hat{\alpha}), \\ \phi^{F^{\alpha,\alpha}}(\pi), & \text{if } \alpha \in [\hat{\alpha}, \min\{\pi, 1/2\}). \end{cases}$$

Similarly as above, we know $V(\pi; [\cdot, \beta])$ is strictly increasing over $[\hat{\alpha}, \min\{\pi, 1/2\})$.

At $\alpha = \hat{\alpha}$, we know $\phi^{F^{\hat{\alpha},\hat{\alpha}}}(\pi) = \phi^{F_\beta^{\hat{\alpha},0}}(\pi)$ because $\pi = \hat{\pi}(\hat{\alpha}, \beta)$. Therefore, we can again show that $V(\pi; [\cdot, \beta]) = V(\pi; [\hat{\alpha}, \beta])$ for all $\alpha < \hat{\alpha}$. Hence, $V(\pi; [\cdot, \beta])$ is constant over $(0, \hat{\alpha})$ and is strictly increasing over $(\hat{\alpha}, 1/2)$.

Case 4: $\beta \leq 1/2$.

Similarly as Case 3, there exists a unique $\hat{\alpha} < \pi$ such that $\hat{\pi}(\hat{\alpha}, \beta) = \pi$. By Propositions 3 and 4, we know

$$V(\pi; [\alpha, \beta]) = \begin{cases} \max\limits_{x \in [\alpha, 1-\beta]} \phi^{F_\beta^{x,0}}(\pi), & \text{if } \alpha \in (0, \hat{\alpha}), \\ \phi^{F^{\alpha,\alpha}}(\pi), & \text{if } \alpha \in [\hat{\alpha}, \pi). \end{cases}$$

Using the same arguments as in Case 3, we can show that $V(\pi; [\cdot, \beta])$ is constant over $(0, \hat{\alpha})$ and is strictly increasing over $(\hat{\alpha}, 1/2)$, completing the proof. □

Proof of Part (ii) of Corollary 2 Fix $\alpha < 1/2$ and $\alpha < \pi$. Because $\hat{\pi}(\alpha, \pi) < \pi$, $\lim_{\beta \uparrow 1} \hat{\pi}(\alpha, \beta) = 1$ and $\hat{\pi}(\alpha, \cdot)$ is strictly increasing by Lemma E.1, there exists a unique $\hat{\beta} \in (\pi, 1)$ such that $\hat{\pi}(\alpha, \hat{\beta}) = \pi$. When $\beta > \hat{\beta}$, we have $\pi < \hat{\pi}(\alpha, \beta)$. By Propositions 3 and 4, we know the optimal information structure is $F^{\alpha,\alpha}$ in this case and thus $V(\pi; [\alpha, \beta]) = \phi^{F^{\alpha,\alpha}}(\pi)$ for all $\beta > \hat{\beta}$. Because $F^{\alpha,\alpha}$ is independent of β, we know $V(\pi; [\alpha, \cdot])$ is a constant over $(\hat{\beta}, 1)$.

It remains to show that $V(\pi; [\alpha, \cdot])$ is strictly decreasing over $(\pi, \hat{\beta})$. We consider three cases.

Case 1: $\pi \leq \hat{\pi}(\alpha, 1/2)$.

In this case, $\hat{\beta} \leq 1/2$. Because $\hat{\pi}(\alpha, \beta) < \pi$ for all $\beta \in (\pi, \hat{\beta})$, we know

$$V(\pi; [\alpha, \beta]) = \max_{x \in [\alpha, \beta)} \phi^{F_\beta^{x,0}}(\pi), \ \forall \beta \in (\alpha, \hat{\beta}),$$

from Proposition 3. As in the proof of Proposition 3, we can also write

$$V(\pi; [\alpha, \beta]) = \max_{x \in [\alpha, \pi]} a_\beta^{x,0}(\pi - x), \ \forall \beta \in (\alpha, \hat{\beta}).$$

Because

$$a_\beta^{x,0} = \frac{1}{2(1-x)\left[1 - \sqrt{\frac{x(1-\beta)}{(1-x)\beta}}\right]}$$

from (A.6), it is easy to see that $a_\beta^{x,0}$ is strictly decreasing in β. Therefore, $V(\pi; [\alpha, \cdot])$ is strictly decreasing over $(\pi, \hat{\beta})$.

Case 2: $\hat{\pi}(\alpha, 1/2) < \pi < \hat{\pi}(\alpha, 1 - \alpha)$.

In this case, $1/2 < \hat{\beta} < 1 - \alpha$, or equivalently $\alpha < 1 - \hat{\beta} < 1/2 < \hat{\beta}$. If $\pi < 1/2$, we have

$$V(\pi; [\alpha, \beta]) = \begin{cases} \max\limits_{x \in [\alpha, \beta)} a_\beta^{x,0}(\pi - x), & \text{if } \beta \in (\pi, 1/2], \\ \max\limits_{x \in [\alpha, 1-\beta]} a_\beta^{x,0}(\pi - x), & \text{if } \beta \in (1/2, \hat{\beta}). \end{cases}$$

If $\pi \geq 1/2$, we have

$$V(\pi; [\alpha, \beta]) = \max_{x \in [\alpha, 1-\beta]} a_\beta^{x,0}(\pi - x), \ \forall \beta \in (\pi, \hat{\beta}).$$

Similarly as Case 1, we can show $V(\pi; [\alpha, \cdot])$ is strictly decreasing over $(\pi, \hat{\beta})$.

Case 3: $\hat{\pi}(\alpha, 1-\alpha) \leq \pi$.

In this case, $1 - \alpha \leq \hat{\beta}$, or equivalently $1 - \hat{\beta} \leq \alpha < 1/2 < \hat{\beta}$. Moreover, because $\hat{\pi}(\alpha, 1-\alpha) = (2-3\alpha)/(3-4\alpha) \geq 1/2$, we know $\pi \geq 1/2$. If $\pi < 1 - \alpha$, we have

$$V(\pi; [\alpha, \beta]) = \begin{cases} \max\limits_{x \in [\alpha, 1-\beta]} a_\beta^{x,0}(\pi - x), & \text{if } \beta \in (\pi, 1-\alpha), \\ \phi^{F_\beta^{\alpha, \hat{b}(\alpha,\beta)}}(\pi), & \text{if } \beta \in [1-\alpha, \hat{\beta}). \end{cases}$$

If $\pi \geq 1 - \alpha$, we have

$$V(\pi; [\alpha, \beta]) = \phi^{F_\beta^{\alpha, \hat{b}(\alpha,\beta)}}(\pi), \ \forall \beta \in (\pi, \hat{\beta}).$$

Similarly as the previous cases, we can show that $V(\pi; [\alpha, \cdot])$ is strictly decreasing over $(\pi, 1-\alpha)$ if $\pi < 1 - \alpha$. What is left is to show that it is strictly decreasing over $(\max\{\pi, 1-\alpha\}, \hat{\beta})$. Pick $\beta_1, \beta_2 \in (\max\{\pi, 1-\alpha\}, \hat{\beta})$ and $\beta_1 < \beta_2$. We can write

$$V(\pi; [\alpha, \beta_1]) = \phi^{F_\beta^{\alpha, \hat{b}(\alpha,\beta_1)}}(\hat{\pi}(\alpha, \beta_1)) + a_\beta^{\alpha, \hat{b}(\alpha,\beta_1)}(\pi - \hat{\pi}(\alpha, \beta_1)),$$
$$V(\pi; [\alpha, \beta_2]) = \phi^{F_\beta^{\alpha, \hat{b}(\alpha,\beta_2)}}(\hat{\pi}(\alpha, \beta_1)) + a_\beta^{\alpha, \hat{b}(\alpha,\beta_2)}(\pi - \hat{\pi}(\alpha, \beta_1)).$$

Because (1) $\phi^{F_{\beta_k}^{\alpha, \hat{b}(\alpha,\beta_k)}}(\hat{\pi}(\alpha, \beta_k)) = \phi^{F^{\alpha,\alpha}}(\hat{\pi}(\alpha, \beta_k))$ for $k = 1, 2$ by the definition of $\hat{\pi}$, (2) $\phi^{F_{\beta_2}^{\alpha, \hat{b}(\alpha,\beta_2)}}$ crosses $\phi^{F^{\alpha,\alpha}}$ from below, and (3) $\hat{\pi}(\alpha, \beta_1) < \hat{\pi}(\alpha, \beta_2)$ by Lemma E.1, we have

$$\phi^{F_{\beta_1}^{\alpha, \hat{b}(\alpha,\beta_1)}}(\hat{\pi}(\alpha, \beta_1)) = \phi^{F^{\alpha,\alpha}}(\hat{\pi}(\alpha, \beta_1)) > \phi^{F_{\beta_2}^{\alpha, \hat{b}(\alpha,\beta_2)}}(\hat{\pi}(\alpha, \beta_1)). \tag{E.5}$$

Because $\pi = \hat{\pi}(\alpha, \hat{\beta}) > \hat{\pi}(\alpha, \beta_1)$ and $a_{\beta_1}^{\alpha, \hat{b}(\alpha,\beta_1)} > a_{\beta_2}^{\alpha, \hat{b}(\alpha,\beta_2)}$, we also have

$$a_{\beta_1}^{\alpha, \hat{b}(\alpha,\beta_1)}(\pi - \hat{\pi}(\alpha, \beta_1)) > a_{\beta_2}^{\alpha, \hat{b}(\alpha,\beta_2)}(\pi - \hat{\pi}(\alpha, \beta_1)). \tag{E.6}$$

Combining (E.5) and (E.6), we obtain

$$V(\pi; [\alpha, \beta_1]) > V(\pi; [\alpha, \beta_2]),$$

completing the proof. □

References

Albrecht, B.C.: Political Persuasion. working paper, University of Minnesota (2017)
Alonso, R., Câmara, O.: Bayesian persuasion with heterogeneous priors. J. Econ. Theory **165**, 672–706 (2016)
Au, P.H., Kawai, K.: Competitive disclosure of correlated information. Econ. Theory (2019). https://doi.org/10.1007/s00199-018-01171-7
Au, P.H., Kawai, K.: Competitive information disclosure by multiple senders. Games Econ. Behav. **119**, 56–78 (2020)
Beauch?ne, D., Li, J., Li, M.: Ambiguous persuasion. J. Econ. Theory **179**, 312–365 (2019)
Bergemann, D., Schlag, K.: Pricing without priors. J. Eur. Econ. Assoc. **6**, 560–569 (2008)
Bergemann, D., Schlag, K.: Robust monopoly pricing. J. Econ. Theory **146**(6), 2527–2543 (2011)
Bergemann, D., Brooks, B., Morris, S.: Informationally Robust Optimal Auction Design. Working paper (2016)
Bergemann, D., Bonatti, A., Smolin, A.: The design and price of information. Am. Econ. Rev. **108**(1), 1–48 (2018)
Blackwell, D.: Comparison of experiments. In: Neyman, J. (ed.) Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, pp. 93–102. University of California Press, Berkeley, California (1951)
Bodoh-Creed, A.L.: Ambiguous beliefs and mechanism design. Games Econ. Behav. **75**(2), 518–537 (2012)
Bose, S., Renou, L.: Mechanism design with ambiguous communication devices. Econometrica **82**, 1853–1872 (2014)
Bose, S., Ozdenoren, E., Pape, A.: Optimal auctions with ambiguity. Theor. Econ. **1**, 411–438 (2006)
Carrasco, V., Luz, V.F., Kos, N., Messner, M., Monteiro, P., Moreira, H.: Optimal selling mechanism under moment conditions. J. Econ. Theory **177**, 245–279 (2018)
Carroll, G.: Robustness and linear contracts. Am. Econ. Rev. **105**(2), 536–563 (2015)
de Castro, L.I., Yannelis, N.C.: Uncertainty, efficiency and incentive compatibility: ambiguity solves the conflict between efficiency and incentive compatibility. J. Econ. Theory **177**, 678–707 (2018)
de Castro, L.I., Pesce, M., Yannelis, N.C.: Core and equilibria under ambiguity. Econ. Theory **48**(2–3), 519 (2011)
de Castro, L.I., Liu, Z., Yannelis, N.C.: Ambiguous implementation: the partition model. Econ. Theory **63**(1), 233–261 (2017a)
de Castro, L.I., Liu, Z., Yannelis, N.C.: Implementation under ambiguity. Games Econ. Behav. **101**, 20–33 (2017b)
Du, S.: Robust mechanisms under common valuation. Econometrica **86**, 1569–1588 (2018)
Esponda, I., Pouzo, D.: Berk–Nash equilibrium: a framework for modeling agents with misspecified models. Econometrica **84**(2), 1093–1130 (2016a)
Esponda, I., Pouzo, D.: Equilibrium in Misspecified Markov Decision Processes. Working paper (2016b)
Garrett, D.F.: Robustness of simple menus of contracts in cost-based procurement. Games Econ. Behav. **87**, 631–641 (2014)
Gentzkow, M., Kamenica, E.: Bayesian persuasion with multiple senders and rich signal space. Games Econ. Behav. **104**, 411–429 (2017a)
Gentzkow, M., Kamenica, E.: Competition in persuasion. Rev. Econ. Stud. **84**(1), 300–322 (2017b)
Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J. Math. Econ. **18**(2), 141–153 (1989)
Guo, Y., Shmaya, E.: The interval structure of optimal disclosure. Econometrica **87**, 653–675 (2019)
Hewitt, E., Stromberg, K.: Real and Abstract Analysis: A Modern Treatment of the Theory of Functions of a Real Variable. Springer, New York (1965)
Kamenica, E., Gentzkow, M.: Bayesian persuasion. Am. Econ. Rev. **101**(6), 2590–2615 (2011)
Koessler, F., Laclau, M., Tomala, T.: Interactive Information Design. Tech. rep (2018)
Kolotilin, A.: Optimal information disclosure: a linear programming approach. Theor. Econ. **13**(2), 607–636 (2018)
Kolotilin, A., Li, M., Mylovanov, T., Zapechelnyuk, A.: Persuasion of a privately informed receiver. Econometrica **85**(6), 1949–1964 (2017)
Kosterina, S.: Persuasion with Unknown Beliefs. Working paper, University of Princeton (2018)
Li, F., Norman, P.: On Bayesian persuasion with multiple senders. Econ. Lett. **170**, 66–70 (2018)
Rayo, L., Segal, I.: Optimal information disclosure. J. Polit. Econ. **118**(5), 949–987 (2010)

Wolitzky, A.: Mechanism design with maxmin agents: theory and an application to bilateral trade. Theor. Econ. **11**(3), 971–1004 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.