Ambiguous Persuasion in Contests *

Xin Feng†

November 2023

Abstract

We study optimal information disclosure via an ambiguous persuasion approach in a two-player contest. The designer can precommit to an ambiguous device to influence the uninformed contestant's belief about his opponent's private type. We fully characterize the optimal ambiguous information structures when players are maxmin expected utility (MMEU) maximizers. Depending on the prior, it is optimal to either induce ambiguity or fully conceal information. We provide a necessary and sufficient condition under which an effort-maximizing organizer can benefit strictly more from using ambiguous persuasion than from using the optimal Bayesian device.

JEL Classification Numbers: D72, D81, D82.

Keywords: Contests, Bayesian persuasion, Ambiguity Aversion, Information Disclosure.

*We are grateful to the Associate Editor and two anonymous reviewers for their insightful comments and constructive suggestions, which significantly improved the quality of the paper. We thank Satoru Takahashi and Jun Zhang for their insightful comments and suggestions. We thank Huiyi Guo, Xiang Sun, Jie Zheng and the seminar participants at the Wuhan University, and the conference participants of the 8th Annual Contests: Theory and Evidence Conference, Asian Meeting of the Econometric Society in China 2022, 2022 International Conference on Industrial Economics, 2022 Conference on Mechanism and Institution Design, and The 4th Forum For Chinese Microeconomics Theories for their helpful feedbacks. We thank the National Natural Science Foundation of China (Nos. 72273063 and 71803019). Any remaining errors are our own.

†Xin Feng: School of Economics, Nanjing University, 22 Hankou Road, Nanjing, Jiangsu 210093, China. Email: a0078175@u.nus.edu.

1 Introduction

Contest organizers often promote productive effort from contestants through contests. For example, employees within a company compete for promotion, school students make efforts to compete for college admissions, athletes spend years in training to rank better in sports events, and firms engage in research tournaments to win a patent race. In those competitive situations, the resources/effort devoted are usually costly and irreversible. When a contestant decides how much effort to expend strategically, information about his rival's private type becomes crucial. A contest organizer who aims to maximize effort can therefore manipulate a contestant's belief about his opponent by disclosing relevant information. In this paper, we allow a contest organizer to disclose *ambiguous* information in a contest with ambiguity-averse players.[1]

Intentional ambiguity is a ubiquitous phenomenon, as organizers often reveal information without explaining its credibility, which naturally creates ambiguity in real-world situations. For instance, when a new entrant competes against an incumbent employee for a job position, a manager could disclose relevant information about the new candidate in a mild, vague and roundabout way, which leaves the incumbent uncertain about his competitor's capability. Moreover, ambiguity arises when there exist multiple information channels. Consider two firms competing for a procurement contract: A contest organizer usually possesses information about the firms' capacities as well as their proposals, and can inform an uninformed firm that his opponent is strong but does not specify how the conclusion is drawn. In this case, the message of "strong" is ambiguous, as the uninformed firm is uncertain whether the information released is based on the firm's competence or the quality of the proposal. A question naturally arises here: Should an effort-maximizing organizer leave his message ambiguous without specifying the exact information channel that generates the message? More generally, we ask the following research questions: Would an effort-maximizing organizer choose to disclose ambiguous information intentionally? What is the optimal (ambiguous) information structure in general? Could an organizer do strictly better by using an ambiguous information device than by using the optimal probabilistic device?

To seek the answers, we study the optimal design of information disclosure when an organizer is allowed to use ambiguous information structures in a binary-state environ-

[1]See Gilboa and Marinacci (2016) for a recent survey of the literature on ambiguity aversion and its axiomatic foundations. See Mukerji and Tallon (2004) and Epstein and Schneider (2010) for two surveys on the economic and financial applications of ambiguity aversion.

ment. Specifically, an ambiguous information device consists of multiple probabilistic devices/signals. It is shown in Kamenica and Gentzkow (2011) that each probabilistic device/signal induces a Bayes-plausible distribution over posteriors. Meanwhile, by observing a message induced by an ambiguous device, the uninformed player (receiver) forms a set of posteriors. We assume that the uninformed player applies the full Bayesian rule to update his belief (see (Epstein and Schneider, 2007; Pires, 2002)) and that players are maxmin expected utility (MMEU) maximizers (see (Gilboa and Schmeidler, 1989)). In particular, the organizer maximizes his maxmin objective, which is the expected total effort.

In our model, two contestants compete in a one-sided incomplete information contest. The valuation of the uninformed contestant is commonly known, while the informed player's valuation is his own private information, which can be either high or low. The binary distribution of the valuation is common knowledge for all players including the contest organizer. Prior to the contest, the organizer can design and precommit to an (ambiguous) information device, which would be announced to the public once decided.[2] According to the information policy, a message is drawn and observed by all players. Based on the revealed information device and the realized message, the uninformed contestant forms a set of posteriors and then both contestants make their efforts simultaneously.[3]

In the paper, we allow a contest organizer to disclose information via an ambiguous device. As an MMEU maximizer, the uninformed contestant would take the worst-case belief among all the posteriors induced by the realized message. In our contest game, the more likely his opponent is to be a high-type, the lower the payoff that a contestant expects to receive. By introducing proper ambiguity, the organizer can persuade an ambiguity-averse contestant to *always* "overestimate" the strength of his opponent.[4] However, it is infeasible to persuade an uninformed contestant to "underestimate" his opponent *on average*. This is because the worst-case posterior that the uninformed contestant takes first-order stochastically dominates all others within the induced set of posteriors, as it assigns the greatest possible probability to the event that his opponent is a high-type.

We fully characterize the optimal ambiguous information structures in the contest game. As we will show, when the prior probability of a high-type is lower than a cutoff, the orga-

[2] In contrast to Zhang and Zhou (2016) studying probabilistic information disclosure in this setting, we examine the use of ambiguous information disclosure in maximizing the total effort in the same setting.

[3] If the information device is a Bayesian one, the set of posteriors boils down to a single posterior.

[4] The notion of overestimation differs from overconfidence in Deng, Fang, Fu, and Wu (2021), who assume that players hold different priors. In contrast, we compare the uninformed contestest's belief before and after persuasion, i.e., his prior and worst-case posterior belief.

nizer can induce both contestants to always exert the greatest effort by disclosing ambiguous information; otherwise, a deterministic information policy would be optimal. When it is optimal to introduce ambiguity, we construct the effort-maximizing ambiguous signal under which the uninformed player "overestimates" his opponent with probability one, regardless of the realized message. For the remaining case, we apply the concavification approach by Kamenica and Gentzkow (2011) and Beauchêne, Li, and Li (2019) to establish the optimality of full concealment.[5] Our results indicate that depending on the prior, an effort-maximizing organizer should either induce ambiguity or simply stay silent. More precisely, the comparison between the prior probability of a high-type and the identified cutoff determines: (1) which kind of policy is optimal and (2) when an effort-maximizing organizer can benefit strictly more from using the optimal ambiguous device than from using the optimal Bayesian one. Moreover, we show that there exist uncountably many optimal information structures whenever it is optimal to induce ambiguity. Recall that a maxmin contestant's effort choice is determined solely by his worst-case belief. We formally establish that an ambiguous device is optimal if and only if the worst posterior belief of the contestant can attain the desired one.

In this contest setting, we consider ambiguous information structures for two reasons. From an applied perspective, ambiguous information is commonly observed in practice, for example, a contest organizer can disclose relevant information without specifying the exact information generating mechanism. As a result, the uninformed player does not know the credibility of the information.[6] More theoretically, allowing ambiguous devices grants an organizer a great deal of flexibility in designing the information structures. Our result demonstrates that the additional flexibility can benefit an effort-maximizing organizer in many circumstances. This implies that an effort-maximizing organizer would intentionally generate ambiguous information in a competitive situation.

Related Literature. The literature on information disclosure/Bayesian persuasion has grown rapidly in recent years, see, e.g., Rayo and Segal (2010), Kamenica and Gentzkow (2011), and Bergemann and Morris (2016a,b), among others. Economic applications of information design to contests have also attracted much attention recently. Earlier literature concerning information disclosure in contests focuses mainly on deterministic information

[5] Beauchêne, Li, and Li (2019) generalize the concavification approach by Kamenica and Gentzkow (2011) to characterize the value of optimal *ambiguous* persuasion.

[6] If the organizer is restricted to use a Bayesian device, the organizer has to specify the information structure. In that case, the uninformed player knows which information structure that generates the revealed information.

policies. Some of these studies focus on the comparison between full disclosure and full concealment, e.g., Wärneryd (2003), Fu, Jiao, and Lu (2014), Denter, Morgan, and Sisak (2014), and Chen, Jiang, and Knyazev (2017), etc. While others compare the complete- and incomplete-information settings, e.g., Morath and Münster (2008), Wasser (2013), and Kovenock, Morath, and Münster (2015), etc. In addition, Serena (2021) explores type-dependent deterministic information disclosure in Tullock contests and Lu, Ma, and Wang (2018) further investigate such deterministic information policies in all-pay auctions. More recent studies allow a contest organizer to provide *probabilistic* information through Bayesian persuasion approach by Kamenica and Gentzkow (2011), e.g., Zhang and Zhou (2016), Chen, Kuang, and Zheng (2019), Feng (2020), Chen (2021), Deng, Fang, Fu, and Wu (2021), Chen, Kuang, and Zheng (2022), and Kuang, Zhao, and Zheng (2022), among others. The current paper differs from the previous literature in that here, the contest organizer can choose to use *ambiguous* signals. In our contest model, we show that an effort-maximizing organizer can elicit more effort by using ambiguous information structures than by using Bayesian devices in many cases.

This paper is closest to Zhang and Zhou (2016), who apply a Bayesian persuasion approach to explore the optimal probabilistic information disclosure in contests. They find that when the private type is binary, comparing only deterministic information policies does not sacrifice generality.[7] The paper complements Zhang and Zhou (2016) by adopting the same model setup but allowing for the use of *ambiguous* information structures. We show that ambiguous devices can elicit strictly more effort than Bayesian devices in many circumstances; as a result, focusing on deterministic information policies would cause a loss in a binary-state environment when ambiguous information structures are taken into account. As a by-product, we solve for the two-player contest game when one player makes his effort decision under uncertainty.

Technically, our paper benefits from Beauchêne, Li, and Li (2019). They characterize the optimal value of ambiguous persuasion and provide useful constructions to approach this value.[8] Relying on their approach, we identify the optimal effort in the contest game. However, it remains unclear whether there exists an optimal information structure that can induce the exact optimal effort identified. We show that the optimal effort can be always achieved in our two state-infinite actions contest game and two messages would be sufficient.

[7]When the state goes beyond two, they provide a useful procedure to compute the optimal signal, which yields partial characterization of the optimal information structure, e.g., full disclosure and semi-separating.

[8]See Proof of Proposition 1, Examples 3 and 4, and Footnote 23 in Beauchêne, Li, and Li (2019).

Moreover, we provide a full characterization by constructing *all* the optimal information structures that can attain the optimal effort.

The current paper is also related to recent studies on ambiguous communication between strategic players.[9] Blume and Board (2014), Bose and Renou (2014), Kellner and Le Quement (2017), and Kellner and Le Quement (2018) investigate ambiguous communication in a cheap-talk model in which the sender has no commitment power. While Beauchêne, Li, and Li (2019), Cheng (2021), and Tang (2021) assume that the sender can fully commit to the ambiguous communication device.

The remainder of this paper is organized as follows. In Section 2, we introduce the contest model with one-sided incomplete information. In Section 3, we study the belief-updating process, characterize the equilibrium, and formalize the organizer's problem, successively. Combining the results, we derive the optimal ambiguous information disclosure. Section 5 provides discussions and concludes.

2 Model

We consider a static contest game between two players under one-sided incomplete information.[10] The two risk-neutral players, indexed by $i \in \{A, B\}$, compete for a single prize by exerting irreversible efforts simultaneously. We employ a lottery contest success function (CSF) to model the competition: Given any effort profile (x_A, x_B) with $x_A,\ x_B \in [0, +\infty)$, the winning probability of player i equals

$$p_i(x_A, x_B) = \begin{cases} x_i/(x_A + x_B) & \text{if } (x_A, x_B) \neq (0,0), \\ 1/2 & \text{if } (x_A, x_B) = (0,0). \end{cases}$$

Each player incurs a unity marginal cost for exerting effort. Denoting player i's valuation

[9] In addition, there is a growing literature on mechanism design with ambiguity-averse players, e.g., Bose, Ozdenoren, and Pape (2006), Frankel (2014), Di Tillio, Kos, and Messner (2016), Wolitzky (2016), Ayouni and Koessler (2017), Guo (2019), and Lopomo, Rigotti, and Shannon (2021), among others. Di Tillio, Kos, and Messner (2016) show that the seller can increase his profit by using an ambiguous mechanism.

[10] Hurley and Shogren (1998a,b) and Denter, Morgan, and Sisak (2014) consider the contest framework. Zhang and Zhou (2016) and Deng, Fang, Fu, and Wu (2021) adopt a Bayesian persuasion approach to analyze the contest game.

of winning as v_i, the expected payoff of player i equals

$$u_i(x_A, x_B) = p_i(x_A, x_B)v_i - x_i, \forall i \in \{A, B\}.$$

The two players compete in a one-sided incomplete information contest: Player A's valuation, v_A, is commonly known, while player B's valuation, v_B, is his private information. In particular, v_B is distributed over $\Omega := \{v_B^H, v_B^L\}$. Both the contest organizer and player A share a common prior belief about $v_B \in \Omega$, which is captured by $\mu_0 := \Pr(v_B = v_B^H) \in (0, 1)$. To avoid the corner solution, we assume that $v_B^L \geq v_A/4$ throughout the paper.

Unlike in the previous literature, in our model, the contest organizer can choose to use an *ambiguous communication device* to induce effort from both contestants. An ambiguous communication device Π consists of a finite set of signals/probabilistic devices, π_1, ..., π_K.[11] Each probabilistic device π_k consists of probability distributions that are contingent on the state, i.e., $\pi_k = \{\pi_k(\cdot|v_B)\}_{v_B \in \Omega}$. Specifically, for each $v_B \in \Omega$, probabilistic device π_k generates a distribution over a common message space $M = \{h, l\}$, i.e., $\pi_k(\cdot|v_B) \in \Delta M$. We will show that focusing on $M = \{h, l\}$ in Section 4.3 does not entail loss of generality, as the optimal expected total effort remains the same for a finite message space M with $|M| \geq 2$.

As will be discussed in detail in Section 3.1, an ambiguous device Π that consists of $\{\pi_1, ..., \pi_K\}$ generates a message $m \in M$. By observing m, player A updates his own belief probability-by-probability, which leads to a set of posteriors. Each posterior distribution over $\Omega := \{v_B^H, v_B^L\}$ is formed based on π_k and message m using Bayes' rule. As the state is binary, we introduce $\mu_m^k := \Pr(v_B = v_B^H|\pi_k, m)$ to represent the corresponding posterior and $\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}$ to represent the aforementioned set of posteriors for simplicity.

We follow Epstein and Schneider (2007) and Pires (2002) to assume that player A (Receiver) forms posteriors using the full Bayesian updating rule and make decisions based on the worst-case expected payoff by following Gilboa and Schmeidler (1989)'s maxmin expected utility model.[12]

Assumption 1 *Player A is an interim maxmin expected utility maximizer and updates his*

[11] When $K = 1$, an ambiguous device boils down to a Bayesian/probabilistic device, which has been analyzed by Zhang and Zhou (2016) in the contest game. For simplicity of notation, we will also use K to denote the set of probabilistic devices.

[12] Many studies concerning mechanism design with ambiguity adopt the maxmin expected utility model of Gilboa and Schmeidler (1989), e.g., Bose, Ozdenoren, and Pape (2006), Frankel (2014), Di Tillio, Kos, and Messner (2016), Wolitzky (2016), Ayouni and Koessler (2017), Guo (2019), and Lopomo, Rigotti, and Shannon (2021), among others.

own belief probability-by-probability.

Without loss of generality, we can assume that player B is either an MMEU maximizer or simply an expected utility (EU) maximizer, as player B knows his own valuation.[13] The timing of the game is as follows.

Time 1: The contest organizer chooses and precommits to an ambiguous communication device Π that consists of $\{\pi_1, ..., \pi_K\}$.

Time 2: Player B observes his own valuation $v_B \in \Omega = \{v_B^L, v_B^H\}$, which is distributed according to μ_0.

Time 3: A message $m \in M$ is generated according to Π and revealed to the public. By observing the message m, player A forms a posterior set, i.e., $\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}$.

Time 4: Both contestants exert efforts simultaneously. Payoffs are determined accordingly.

We call stage 1 the ambiguous persuasion stage. The contest organizer's problem is to design an ambiguous information structure in stage 1, in order to maximize her ex ante maxmin expected objective, which is the total effort from both contestants. We assume that the organizer fully commits to the ambiguous device. Alternatively, we can assume that the organizer commits to one probabilistic device that belongs to $\{\pi_1, ..., \pi_K\}$ but neither player knows which device has been chosen.

3 Analysis

In this section, we study the optimal ambiguous information structure for the environment in Section 2. In Section 3.1, given an ambiguous device, by considering player A's belief-updating process, we identify the set of posteriors that results from a realized message. In Section 3.2, we characterize the equilibrium and calculate the expected total effort when player A holds multiple posteriors. In Section 3.3, we formulate the organizer's optimization problem. Combining the results, we discuss two cases according to the range of the prior and fully solve for the optimal ambiguous device in Section 3.4.

[13] As Player A would follow Assumption 1 to evaluate his own payoff, there is no further layer of ambiguity even we assume that player B is an MMEU maximizer.

3.1 Belief Updating

Given $\Omega = \{v_B^H, v_B^L\}$ and the message space $M = \{h, l\}$, an ambiguous communication device Π consists of $\pi_1, ..., \pi_K$ with each signal/probabilistic device $\pi_k = \{\pi_k(\cdot|v_B^L), \pi_k(\cdot|v_B^H)\}$, where $\pi_k(\cdot|v_B^L)$, $\pi_k(\cdot|v_B^H) \in \Delta M$. We denote the convex hull of $\{\pi_1, ..., \pi_K\}$ by Π, as in the literature on ambiguous persuasion and MMEU models.[14] Formally,

$$\begin{aligned}\Pi &= co((\pi_k)_K) \\ &= \left\{\pi \in \Delta M \text{ s.t. } \pi = \sum_{k \in K} \lambda(k)\pi_k \text{ for some } \lambda \in \Delta K\right\}. \end{aligned} \tag{1}$$

Each $\pi \in \Pi$ is a linear combination of $\pi_1, ..., \pi_K$. To interpret (1), imagine that the organizer follows a distribution $\lambda(\cdot) \in \Delta\{1, ..., K\}$ to draw a probabilistic device, which will be used to generate a signal, while player A only knows that $\pi_1, ..., \pi_K$ are used but not $\lambda(\cdot)$. Note that $\pi(\cdot|v_B^L)$, $\pi(\cdot|v_B^H) \in \Delta M$, as $\pi_k(\cdot|v_B^L)$, $\pi_k(\cdot|v_B^H) \in \Delta M$, $\forall k$. For an ambiguous communication device consisting of only two devices, π_1 and π_2, the resulting $\Pi = co(\pi_1, \pi_2) = \{\pi \in \Delta M | \pi = \lambda\pi_1 + (1-\lambda)\pi_2 \text{ for a } \lambda \in [0, 1]\}$.

By receiving a message $m \in \{h, l\}$, player A updates his belief probability-by-probability, which leads to the following set of posterior that can be represented by

$$\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}.$$

Each posterior, or μ_m^k equivalently, is induced by π_k and message m. Specifically,

$$\mu_m^k := \Pr(v_B = v_B^H | \pi_k, m) = \frac{\mu_0 \pi_k(m|v_B^H)}{\mu_0 \pi_k(m|v_B^H) + (1-\mu_0)\pi_k(m|v_B^L)},$$

and $\Pr(v_B = v_B^L | \pi_k, m) = \frac{(1-\mu_0)\pi_k(m|v_B^L)}{\mu_0 \pi_k(m|v_B^H) + (1-\mu_0)\pi_k(m|v_B^L)}$.

For an easier exposition, given ambiguous device $\Pi = co((\pi_k)_K)$, we denote by $\{\mu_h^1, \mu_h^2, ..., \mu_h^K\}$ (resp. $\{\mu_l^1, \mu_l^2, ..., \mu_l^K\}$) the set of posteriors induced by the message h (resp. l).

[14] Beauchêne, Li, and Li (2019) point out that the equilibrium is unaffected by the choice between $(\pi_k)_K$ and Π, as only the extreme points of the set will be minimizing probabilities for an MMEU agent. This result remain valid in our context.

3.2 Contest Game with Multiple Posteriors

We first consider the contest game with a single posterior and extend the analysis with a set of posteriors. When player A holds a single posterior belief $(\mu,\, 1-\mu)$ over $\{v_B^H,\, v_B^L\}$, where $\mu = \Pr(v_B = v_B^H)$, player A exerts effort

$$x_A^*(\mu) = \left(\frac{\frac{\mu}{\sqrt{v_B^H}} + \frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}} \right)^2 .$$

Player B makes his effort according to his valuation v_B, and his effort strategy is

$$x_B(v_B, \mu) = \sqrt{v_B x_A^*(\mu)} - x_A^*(\mu),\ v_B \in \{v_B^H, v_B^L\},$$

which is non-negative, as we assume $v_B^L \geq v_A/4$.

Therefore, the expected total effort that results from a single posterior equals

$$\begin{aligned} TE(\mu) &= E_\mu \left[x_A^*(\mu) + x_B(v_B, \mu) \right] \\ &= E_\mu \left[\sqrt{v_B x_A^*(\mu)} \right] \\ &= \frac{\left[\mu\sqrt{v_B^H} + (1-\mu)\sqrt{v_B^L} \right] \left[\frac{\mu}{\sqrt{v_B^H}} + \frac{1-\mu}{\sqrt{v_B^L}} \right]}{\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}}, \end{aligned} \qquad (2)$$

as in Zhang and Zhou (2016).

We next turn to the contest game under ambiguity. To proceed, we first investigate how player A, who holds a set of posteriors, makes an effort choice. Since player A's preference is represented by MMEU, he would evaluate his payoff under the worst-case scenario when choosing a level of effort under ambiguity. In the following lemma, we solve the contest game with a set of posteriors and derive the corresponding total effort function.

Lemma 1 *Under Assumption 1, given the posterior set, i.e.,* $\{\mu^1, \mu^2, ..., \mu^K\}$, *we have*

$$TE(\{\mu^1, \mu^2, ..., \mu^K\}) = TE(\max(\mu^1, \mu^2, ..., \mu^K)),$$

where $TE(\{\mu^1, \mu^2, ..., \mu^K\})$ *is the expected total effort that results from the the posterior set*

$\{\mu^1, \mu^2, ..., \mu^K\}$ and $TE(\cdot)$ is given by (2).

Proof. By Assumption 1, player A is ambiguity averse with maxmin expected utility. Given player B's effort, player A's problem is

$$\max_{x_A \in [0,+\infty)} \min_{\mu \in \{\mu^1, \mu^2, ..., \mu^K\}} E_\mu \left[\frac{x_A}{x_A + x_B(v_B)} v_A - x_A \right]$$
$$= \max_{x_A \in [0,+\infty)} E_{\mu^*} \left[\frac{x_A}{x_A + x_B(v_B)} v_A - x_A \right],$$

where $\mu^* = \max(\mu^1, \mu^2, ..., \mu^K)$. Specifically, μ^* also represents the distribution that assigns $\max(\mu^1, \mu^2, ..., \mu^K)$ to v_B^H over $\{v_B^H, v_B^L\}$.

The reason why the last equality holds is that the more likely player B is to be a high-type, the lower player A's payoff would be. Therefore, player A would maximize his utility according to μ^*, which is his worst-case belief among $\{\mu^1, \mu^2, ..., \mu^K\}$.

The remaining calculation is analogous to the case wherein the posterior set is a singleton. Specifically, player A's equilibrium effort is

$$x_A^*(\mu^*) = \left(\frac{\frac{\mu^*}{\sqrt{v_B^H}} + \frac{1-\mu^*}{\sqrt{v_B^L}}}{\frac{1}{v_A} + \frac{\mu^*}{v_B^H} + \frac{1-\mu^*}{v_B^L}} \right)^2$$

and player B's equilibrium effort is

$$x_B(v_B, \mu^*) = \sqrt{v_B x_A^*(\mu^*)} - x_A^*(\mu^*), \text{ for } v_B \in \{v_B^H, v_B^L\}.$$

The expected total effort therefore equals $TE(\mu^*)$, where $\mu^* = \max(\mu^1, \mu^2, ..., \mu^K)$ and $TE(\cdot)$ is given by (2). ■

Lemma 1 says that when player A holds a set of posteriors $\{\mu^1, \mu^2, ..., \mu^K\}$, the expected total effort is solely determined by the worst one that assigns $\max(\mu^1, \mu^2, ..., \mu^K)$ to v_B^H. This is because player A is an MMEU maximizer and exerts effort according to his worst-case belief μ^*.

3.3 The Organizer's Problem

In this subsection, we will formally describe the organizer's problem. The organizer chooses an ambiguous communication device to maximize his ex ante *maxmin* payoff, which is the expected total effort. For an ambiguous communication device that consists of $\{\pi_1$, ..., $\pi_K\}$ and the message space $M = \{h, l\}$, each probabilistic device π_k induces a posterior beilef μ_m^k upon receiving a message $m \in M$. Let $\tau_k \in \Delta M$ denote the marginal distribution over message space M. The distribution over posteriors, τ_k, is Bayes plausible if it satisfies $\sum_{m\in M} \tau_k(m)\mu_m^k = \mu_0$, i.e., $\tau_k(h)\mu_h^k + \tau_k(l)\mu_l^k = \mu_0$, $\forall k \in \{1, ..., K\}$. Probabilistic devices can be viewed as a special class of ambiguous devices by letting $K = 1$.

The organizer's problem is to choose an ambiguous communication device Π to maximize his ex ante maxmin objective in the following:

$$\begin{aligned} & \sup_{\Pi} \min_{\pi\in\Pi} E_\pi \left[TE(\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}) \right] \\ = & \sup_{\Pi} \min_{\pi\in\Pi} \sum_{v_B\in\{v_B^H, v_B^L\}} \left[\Pr(v_B) \sum_{m\in\{h,l\}} \pi(m|v_B) TE(\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}) \right], \end{aligned} \tag{3}$$

where $\Pr(v_B^H) = \mu_0$ is the prior probability that player B is a v_B^H-type and $\Pr(v_B^L) = 1 - \mu_0$ is the prior probability of being a v_B^L-type.

Recall that an ambiguous device Π is a closed and convex set of multiple probabilistic devices with common support, $\pi_1, ..., \pi_K$. More precisely, $\Pi = co((\pi_k)_K)$ as in (1). The organizer has flexibility to design each probabilistic device, π_k, and to decide how many devices to be used, K, which determines Π uniquely.[15] This ambiguous device is utilized to generate a message. After the message is revealed, player A forms his posterior belief as described in Section 3.1 and players A and B compete as described in Section 3.2. Given Π, the organizer's worst-case payoff is $\min_{\pi\in\Pi} \sum_{v_B\in\{v_B^H, v_B^L\}} \left[\Pr(v_B) \sum_{m\in\{h,l\}} \pi(m|v_B) TE(\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}) \right]$. The organizer wishes to maximize his worst-case payoff by designing an ambiguous device Π.

By Kamenica and Gentzkow (2011) and Beauchêne, Li, and Li (2019), an ambiguous device that consists of $\{\pi_1$, ..., $\pi_K\}$ can be considered as a corresponding set of distributions over posteriors, $(\tau_k)_{k\in K}$. More precisely, we define R as the set of distributions over posteriors

[15] Even the same Π can sometime be generated by different probabilistic devices $(\pi_k)_K$. Nevertheless, it suffices to pin down the optimal Π.

induced by these K probabilistic devices, i.e.,

$$R = \{(\tau_k)_{k \in K} \colon \tau_k \in \Delta M \text{ s.t. } \tau_k(\cdot) = \mu_0 \pi_k(\cdot | v_B^H) + (1 - \mu_0)\, \pi_k(\cdot | v_B^L)\}.$$

The set of distributions above, R, is determined by the choice of $\{\pi_1, \ldots, \pi_K\}$. As the choice of ambiguous device varies, we may obtain different sets of distributions over posteriors. We further define $\mathcal{R}$ as the collection containing all the sets of the distributions over posteriors, where each distribution satisfies the Bayes-plausible condition.[16] The organizer's problem of designing an ambiguous device is equivalent to choosing a set of distributions that belongs to $\mathcal{R}$. Furthermore, given a set of distributions over posteriors $R \in \mathcal{R}$, the organizer's worst-case payoff equals $\min_{\tau_k \in R} E_{\tau_k} \left[TE(\{\mu_m^1, \mu_m^2, \ldots, \mu_m^K\}) \right]$. We therefore rewrite the organizer's problem in (3) as follows:

$$\sup_{R \in \mathcal{R}} \min_{\tau_k \in R} E_{\tau_k} \left[TE(\{\mu_m^1, \mu_m^2, \ldots, \mu_m^K\}) \right],$$

which is equivalent to

$$\sup_{R \in \mathcal{R}} \min_{\tau_k \in R} \left[\tau_k(h) TE(\max\{\mu_h^1, \mu_h^2, \ldots, \mu_h^K\}) + \tau_k(l) TE(\max\{\mu_l^1, \mu_l^2, \ldots, \mu_l^K\}) \right] \tag{4}$$

using Lemma 1.

3.4 Optimal Ambiguous Persuasion

In this subsection, we will study the organizer's problem in (4) whose solution yields the optimal ambiguous device. To accomplish this, we consider two cases according to the range of the prior μ_0 and derive the optimal ambiguous information devices for each case in Propositions 1 and 3, respectively. We summarize the results in Theorem 1, which implies that an effort-maxmizing organizer can strictly benefit from using ambiguous devices if and only if μ_0 is lower than a cutoff.

To begin our analysis, we introduce some useful properties and definitions. We first introduce the following properties of $TE(\cdot)$, which prove useful when we explore the optimal ambiguous device.

[16] Kamenica and Gentzkow (2011) show that a distribution of posteriors τ can be induced by a probabilistic device if τ is Bayes plausible.

Property 1 *$TE(\cdot)$ in (2) satisfies the following properties:*

(i) $TE(\cdot)$ is continuous within $[0,1]$;

(ii) $TE(\cdot)$ is concave (resp. convex) when $v_A < \sqrt{v_B^H v_B^L}$ (resp. $v_A > \sqrt{v_B^H v_B^L}$) and is linear when $v_A = \sqrt{v_B^H v_B^L}$.

Property 1(i) can be verified directly using equation (2). Property 1(ii) has been shown by Zhang and Zhou (2016).[17] Since $TE(\cdot)$ given by (2) is continuous in $[0,1]$, by the Weierstrass extreme value theorem, there exists an optimal μ_{opt} that maximizes $TE(\cdot)$ within $[0,1]$. Formally, we introduce the following definition.

Definition 1 *We define*

$$\mu_{opt} := \arg\max_{\mu \in [0,1]} TE(\mu).$$

In other words, $TE(\mu_{opt}) \geq TE(\mu)$, $\forall \mu \in [0,1]$. We further prove the uniqueness and fully characterize μ_{opt} in the following lemma.

Lemma 2 *μ_{opt} is unique. In particular,*

(i) if $v_A \geq \sqrt{v_B^H v_B^L}$, $\mu_{opt} = 1$;

(ii) if $v_A < \sqrt{v_B^H v_B^L}$, $\mu_{opt} = \min\{\widehat{\mu}, 1\}$, where $\widehat{\mu}$ solves $TE'(\mu) = 0$.

Proof. The proof and the calculation details are relegated into the Appendix. ■

We next define a special class of ambiguous devices parameterized by μ, which proves helpful in our follow-up analysis.

Definition 2 *Define $\Pi^*(\mu)$ as the ambiguous device that consists of $\{\pi_1, \pi_2\}$, where*

$$\pi_1(m = h|v_B^H) = 0,\ \pi_1(m = h|v_B^L) = \frac{\mu - \mu_0}{(1-\mu_0)\mu};$$

$$\pi_2(m = h|v_B^H) = 1,\ \pi_2(m = h|v_B^L) = \frac{\mu_0(1-\mu)}{(1-\mu_0)\mu}.$$

[17] We also provide the calculation details when proving the uniqueness of μ_{opt}.

Given the above ambiguous device $\Pi^*(\mu)$, a message could be interpreted differently using π_1 and π_2, which generates ambiguity. In the following, we identify all the posterior sets resulting from $\Pi^*(\mu)$ when the prior $\mu_0 < \mu$.

Lemma 3 *If $\mu_0 < \mu$, the ambiguous device $\Pi^*(\mu)$ defined in Definition 2 always induces the same set of posteriors $\{0,\, \mu\}$, regardless of the realized message.*

Proof. Given ambiguous device $\Pi^*(\mu)$, by receiving a message, h or l, player A forms a set of posteriors, $P_h(v_B^H) = \{0,\, \mu\}$ or $P_l(v_B^H) = \{0,\, \mu\}$, respectively. The result holds whenever $\mu_0 < \mu$. ■

From Lemma 3, if the prior $\mu_0 < \mu_{opt} := \arg\max_{\mu\in[0,1]} TE(\mu)$, as in Definition 2, we could construct ambiguous device $\Pi^*(\mu_{opt}) = \{\pi_1,\, \pi_2\}$, which induces the posterior sets $\{0,\, \mu_{opt}\}$ and $\{\mu_{opt},\, 0\}$. $\forall k \in \{1, 2\}$, the resulting expected total effort equals

$$\begin{aligned} & \tau_k(h)TE(\{0, \mu_{opt}\}) + \tau_k(l)TE(\{\mu_{opt}, 0\}) \\ = \; & TE(\mu_{opt}), \end{aligned}$$

which is the highest possible expected effort induced by an ambiguous device. We formally establish the result in the following proposition, which is the first main result of the paper.

Proposition 1 *If the prior $\mu_0 < \mu_{opt}$, the greatest expected total effort induced by an information device equals $TE(\mu_{opt})$. In particular, when $\mu_0 < \mu_{opt}$, $TE(\mu_{opt})$ can be induced by the ambiguous device $\Pi^*(\mu_{opt})$.*

Proof. To explain why $TE(\mu_{opt})$ is the optimal effort, we first recall that $TE(\mu_{opt}) \geq TE(\mu)$, $\forall \mu \in [0, 1]$, which follows directly from Definition 1. Therefore, for any distribution $\tau_k \in \Delta M$, we have

$$TE(\mu_{opt}) \geq \tau_k(h)TE(\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\}) + \tau_k(l)TE(\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\}).$$

As a result,

$$TE(\mu_{opt}) \geq \sup_{R\in\mathcal{R}} \min_{\tau_k\in R} \left[\tau_k(h)TE(\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\}) + \tau_k(l)TE(\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\})\right],$$

which is given by (4).

Moreover, $TE(\mu_{opt})$ is achievable whenever $\mu_0 < \mu_{opt}$. In particular, when $\mu_0 < \mu_{opt}$, $TE(\mu_{opt})$ can be induced by the ambiguous device $\Pi^*(\mu_{opt}) = \{\pi_1, \pi_2\}$, where

$$\pi_1(m = h|v_B^H) = 0,\ \pi_1(m = h|v_B^L) = \frac{\mu_{opt} - \mu_0}{(1-\mu_0)\mu_{opt}};$$

$$\pi_2(m = h|v_B^H) = 1,\ \pi_2(m = h|v_B^L) = \frac{\mu_0(1-\mu_{opt})}{(1-\mu_0)\mu_{opt}}.$$

By Lemma 3, $\Pi^*(\mu_{opt})$ induces posterior sets $\{0, \mu_{opt}\}$ and $\{\mu_{opt}, 0\}$. With Lemma 1, the expected total effort equals $TE(\mu_{opt})$. ■

Proposition 1 solves the optimal design of the ambiguous information structure when $\mu_0 < \mu_{opt}$. In this case, the greatest effort attains $TE(\mu_{opt})$, which maximizes $TE(\cdot)$ in (2). It is worth noting that the organizer cannot elicit strictly more effort than $TE(\mu_{opt})$, even when she is allowed to design player A's belief directly. Moreover, the level of effort, $TE(\mu_{opt})$, can never be achieved by a probabilistic/Bayesian device in this case.

More precisely, it is infeasible to persuade player A to always take belief μ_{opt} via a probabilistic/Bayesian device unless $\mu_0 = \mu_{opt}$, as the Bayes plausible condition requires that the expected posterior equals the prior. In contrast, when ambiguous information devices are available, the organizer can introduce ambiguity, under which player A would pick the worst-case posterior that does not necessarily satisfy the Bayes plausible condition. In particular, the worst-case posterior always assigns higher probability to a high-type than the prior belief does. As a result, an ambiguity-averse contestant may rationally "overestimate" the strength of his opponent under ambiguity. To determine the optimal ambiguous device, it remains to investigate "how much" ambiguity to induce and how many devices should be used.

From (4), it would be optimal to persuade player A to always take belief μ_{opt} whenever possible. As shown in Proposition 1, this turns out to be feasible if the prior $\mu_0 < \mu_{opt}$. Moreover, two devices would be sufficient to induce the desired ambiguity. To explain the result, consider the optimal ambiguous device $\Pi^*(\mu_{opt})$, each message $m \in M = \{h, l\}$ has two different interpretations by using π_1 and π_2, that is, player B is either a low-type for sure or a high-type with probability μ_{opt}. As a maxmin expected utility maximizer, player A would take the worst-case belief μ_{opt} when facing the two possibilities, 0 and μ_{opt}. When player A acts as if his belief is μ_{opt}, both players make their own effort accordingly and the total effort therefore equals $TE(\mu_{opt})$.

Although we derive the optimal ambiguous device in Proposition 1, it remains to be investigated whether the identified optimal information structure is unique. We show that there are uncountably many the optimal ambiguous devices. In the following proposition, we first describe sufficient and necessary conditions that the optimal ambiguous devices with $K \geq 2$ must satisfy and then characterize all optimal ambiguous devices with $K = 2$.

Proposition 2 *If $\mu_0 < \mu_{opt}$, there exist uncountably many ambiguous devices that attain the greatest effort $TE(\mu_{opt})$. More precisely,*

(i) for any $K \geq 2$, a device Π^* is optimal if and only if the induced set of posterior $\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}$ satisfies $\max\{\mu_m^1, \mu_m^2, ..., \mu_m^K\} = \mu_{opt}$ for each realized message $m \in \{h, l\}$;

(ii) for $K = 2$, the collection of optimal devices is $\{\Pi^*(\mu_{opt}, \mu', \mu'') | \mu', \mu'' < \mu_0\}$. Given $\mu', \mu'' < \mu_0$, $\Pi^*(\mu_{opt}, \mu', \mu'')$ denotes the ambiguous device that consists of $\{\pi_1, \pi_2\}$, where

$$\pi_1(m|v_B^H) = \frac{\mu'\left(\mu_{opt} - \mu_0\right)}{\mu_0(\mu_{opt} - \mu')}, \ \pi_1(m|v_B^L) = \frac{(1-\mu')\left(\mu_{opt} - \mu_0\right)}{(1-\mu_0)\left(\mu_{opt} - \mu'\right)};$$

$$\pi_2(m|v_B^H) = \frac{\mu_{opt}\left(\mu_0 - \mu''\right)}{\mu_0(\mu_{opt} - \mu'')}, \ \pi_2(m|v_B^L) = \frac{\left(1-\mu_{opt}\right)\left(\mu_0 - \mu''\right)}{(1-\mu_0)\left(\mu_{opt} - \mu''\right)},$$

for $m \in \{h, l\}$.

Proof. (i) $\Rightarrow$: Consider an ambiguous device Π that consists of $\pi_1, ..., \pi_K$. If Π induces the greatest effort $TE(\mu_{opt})$, $\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\} = \mu_{opt}$ and $\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\} = \mu_{opt}$ must hold by (4) and the definition of μ_{opt}.

$\Leftarrow$: If $\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\} = \max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\} = \mu_{opt}$, it follows from (4) that the induced total effort equals $TE(\mu_{opt})$.

(ii) For $K = 2$, consider an ambiguous device Π that consists of two probabilistic devices $\{\pi_1, \pi_2\}$, denote the induced posteriors μ_1^m and μ_2^m when message is $m \in \{h, l\}$. An ambiguous device is optimal if and only if the induced posteriors always satisfy $\max\{\mu_1^h, \mu_2^h\} = \max\{\mu_1^l, \mu_2^l\} = \mu_{opt}$. In addition, recall that for each Bayesian device π_i, the expected posteriors must equal prior μ_0, which implies that if $\mu_i^h > \mu_0$, $\mu_i^l < \mu_0$ must hold. Since $\mu_{opt} > \mu_0$, if $\mu_i^h = \mu_{opt}$ (resp. $\mu_i^l = \mu_{opt}$), we must have $\mu_i^l < \mu_0$ (resp. $\mu_i^h < \mu_0$), $\forall i \in \{1, 2\}$. As a result, there are only two possibilities: either $\mu_1^l = \mu_2^h = \mu_{opt}$ or $\mu_1^h = \mu_2^l = \mu_{opt}$. We can then recover the corresponding π_1 and π_2. One can check that the constructed π_1 induces a posterior set $\{\mu', \mu_{opt}\}$ and π_2 induces a posterior set $\{\mu'', \mu_{opt}\}$, where $\mu', \mu'' < \mu_0$. The details are relegated into the Appendix. ■

With Proposition 2, we fully characterize the optimal ambiguous information devices when $\mu_0 < \mu_{opt}$. The reason why the optimal ambiguous device is not unique if $\mu_0 < \mu_{opt}$ is that a maxmin contestant would make his effort decision according to his worst belief among the posteriors, i.e., $\max\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}$. To induce the maximal effort, the key is to maintain $\max\{\mu_m^1, \mu_m^2, ..., \mu_m^K\} = \mu_{opt}$ and other posteriors have no effect on the contestant's effort choice. When K goes beyond 2, one way to construct the desired information devices is to add more probabilistic devices $\pi_3, ..., \pi_K$, in addition to π_1 and π_2 given by Proposition 2(i), whenever the induced posteriors are (weakly) less than μ_{opt}. Additionally, $\Pi^*(\mu_{opt})$ in Proposition 1 can be viewed as a special case of $\Pi^*(\mu_{opt}, \mu', \mu'')$, since $\Pi^*(\mu_{opt}) = \Pi^*(\mu_{opt}, \mu', \mu'')|_{\mu'=\mu''=0}$ holds by definition. To better explain our optimality result, we provide an example in the following.

Example 1: Consider a contest game in which $v_B^H = 2$, $v_B^L = 0$, and $v_A = 1$. We plot $TE(\cdot)$ given by (2) in Figure 1. In this case, $\mu_{opt} = 1$, any prior μ_0 satisfies the condition that $\mu_0 \leq \mu_{opt}$. From Proposition 1, the optimal expected total effort that results from ambiguous devices equals $TE(\mu_{opt}) = TE(1) \approx 0.667$, regardless of the prior. It can be easily verified that $TE(\cdot)$ is convex; full disclosure is thus the optimal probabilistic device, and the resulting expected total effort equals $\mu_0 TE(0) + (1 - \mu_0)TE(1) = (1 - \mu_0)TE(1)$, which is less than $TE(\mu_{opt})$.

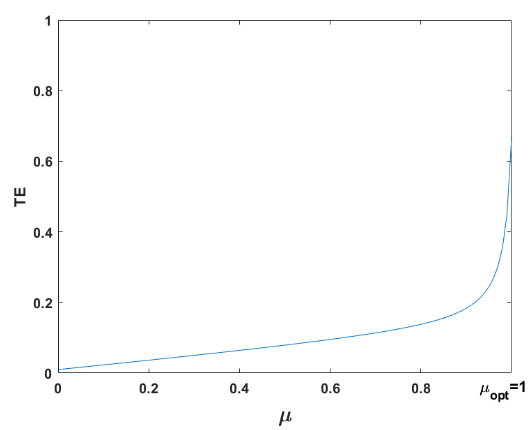


Figure 1

In Example 1, when the two players are more or less evenly matched, it is always beneficial to persuade player A to believe that his opponent is a high-type. However, when the difference between v_A and v_B^H is sufficiently large, this is not the case, i.e., it would discourage player A if he believes that he would compete with a high-type. We provide an example to illustrate the latter case. In the following example, we find that μ_{opt} could be strictly less than 1. Consequently, there exists prior $\mu_0 \in [0, 1]$ that violates the condition $\mu_0 < \mu_{opt}$.

Example 2: In this example, we assume that $v_B^H = 20$, $v_B^L = 1$, and $v_A = 1$. We plot $TE(\cdot)$ given by (2) in Figure 2. From Proposition 1, when the prior $\mu_0 = 0.3$ is less than $\mu_{opt} \approx 0.71$, the optimal expected total effort that results from ambiguous devices equals $TE(\mu_{opt}) \approx 1.173$, which is greater than $TE(\mu_0) \approx 0.915$, the optimal total effort induced by probabilistic devices.

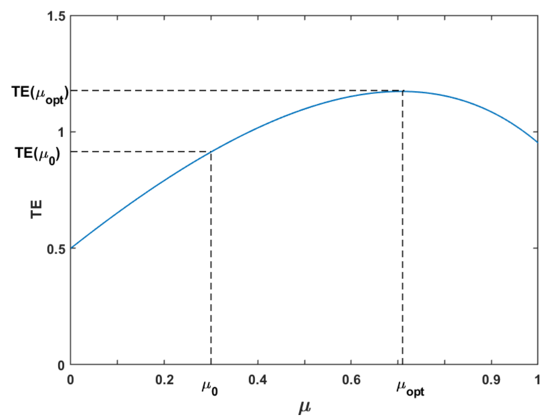


Figure 2: $\mu_0 < \mu_{opt}$

However, when $\mu_0 \geq \mu_{opt}$, the analysis of Proposition 1 no longer applies. For completeness, it remains to tackle the problem in (4) with $\mu_0 \geq \mu_{opt}$. To solve the case of $\mu_0 > \mu_{opt}$, we apply the concavification approach by Kamenica and Gentzkow (2011) and Beauchêne, Li, and Li (2019). The latter characterizes the optimal value of ambiguous persuasion. To apply the concavification approach in our contest game, we first identify the greatest expected total effort, which is given by the maximal projection of the concave closure of the total effort function $TE(\cdot)$ in (2). Given the optimal effort, we construct a Bayesian/probabilistic device to achieve the optimum, which depends on the convexity of $TE(\cdot)$. Recall that $TE(\cdot)$ is either convex or concave by Property 1 (ii).

Proposition 3 *If the prior $\mu_0 \geq \mu_{opt}$, the greatest expected total effort induced by an infor-*

mation device equals $TE(\mu_0)$, which can be induced by no disclosure.

Sketch of proof. In this case, we apply the concavification approach to solve for the *maxmin* effort-maximizing ambiguous device, which takes three steps. In the first step, we derive the maximal projection of the concave closure of the total effort function $TE(\cdot)$, in order to derive the greatest expected total effort that results from the optimal ambiguous persuasion. In the second step, we find that the optimal effort, which is identified in the first step, can be attained by a deterministic device. In the third step, we note that $\mu_0 > \mu_{opt}$ never occurs when $v_A > \sqrt{v_B^H v_B^L}$. For two remaining cases $v_A = \sqrt{v_B^H v_B^L}$ and $v_A < \sqrt{v_B^H v_B^L}$, no disclosure is optimal. The details are relegated into the Appendix. $\square$

Proposition 3 shows that introducing ambiguity does not help in increasing the total effort, if the prior $\mu_0 \geq \mu_{opt}$. In particular, the condition of $\mu_0 > \mu_{opt}$ indicates that player A would be discouraged by facing an overly strong opponent, i.e., player B with a large v_B^H. In this case, the total effort would increase if player A "underestimates" the strength of his opponent, as $TE(\cdot)$ decreases with $\mu \in (\mu_{opt}, 1]$ (see also Figure 2). However, it is impossible to persuade player A to "underestimate" his opponent *on average,* even via an ambiguous information structure.

More precisely, consider an arbitrary ambiguous device that consists of probabilistic devices, $\pi_1, ..., \pi_K$. For each π_k, it follows from the Bayes-plausible condition that $\tau_k(h)\mu_h^k + \tau_k(l)\mu_l^k = \mu_0$. When $m = l$, player A follows the full updating rule to form the posterior set $\{\mu_l^1, \mu_l^2, ..., \mu_l^K\}$ and makes his own effort according to the worst-case belief $\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\}$. A useful observation is that if $\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\} < \mu_0$, $\mu_l^k < \mu_0$ must hold for each k. By the Bayes-plausible condition, $\mu_h^k > \mu_0$ must hold for each k; as a result, $\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\} > \mu_0$. This implies that when $m = h$, player A's worst-case belief is $\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\}$, which must be strictly greater than μ_0. In other words, player A would not systematically "underestimate" his opponent, due to the Bayes-plausible condition. In fact, under ambiguity, player A with maxmin perference always behaves as if he "overestimates" the strength of his opponent, which would lower the total expected effort if $\mu_0 > \mu_{opt}$. As a result, an ambiguous device does no strictly better than any Bayesian device in persuading player A to "underestimate" his opponent. Moreover, no disclosure is the optimal deterministic/probabilistic device. We therefore conclude that if $\mu_0 > \mu_{opt}$, a contest organizer cannot improve the expected total effort by using an ambiguous device over a Bayesian one (or no disclosure).

Combining Propositions 1 and 3, we characterize the optimal effort $TE_{AP}(\mu_0)$ and the corresponding information policy in the following.

Theorem 1 *Depending on the prior μ_0, it is optimal to either induce ambiguity or simply stay silent. More precisely, we have*

(i) If $\mu_0 < \mu_{opt}$, the optimal effort $TE_{AP}(\mu_0) = TE(\mu_{opt})$, which can be induced by the ambiguous device $\Pi^*(\mu_{opt})$; (ii) if $\mu_0 \geq \mu_{opt}$, $TE_{AP}(\mu_0) = TE(\mu_0)$, which can be induced by no disclosure.

To improve the total effort, the organizer wants to persuade player A to "overestimate" (resp. "underestimate") the strength of his opponent when $\mu_0 < \mu_{opt}$ (resp. $\mu_0 > \mu_{opt}$). Generating ambiguous information leads to overestimation while underestimation never occurs, since player A is ambiguity-averse with maxmin preference. Hence, an effort-maximizing organizer can strictly improve the effort by exploiting the ambiguity-aversion of player A if and only if $\mu_0 < \mu_{opt}$. In that case, it is essential for such an organizer to induce ambiguity by using at least two Bayesian devices.

Specifically, it is optimal to persuade player A to act as if his opponent is a v_H-type with probability μ_{opt}. Any other belief would cause a fall in the total effort, since μ_{opt} uniquely maximizes the total effort function. Intuitively, the belief μ_{opt} balances and thus intensifies the competition between the two players. However, it is feasible to induce player to take belief μ_{opt} if and only if $\mu_0 \leq \mu_{opt}$. Recall that by Property 1 and Lemma 2, if $v_A < \sqrt{v_B^H v_B^L}$, $TE(\cdot)$ is concave and $\mu_{opt} \in (0, 1]$, which implies that inducing proper "overestimation" is optimal when player A is not that strong; if $v_A > \sqrt{v_B^H v_B^L}$, $TE(\cdot)$ is convex and increasing, $\mu_{opt} = 1$, in this case, the designer should induce "overestimation" as much as possible as player A is strong.

We next compare ambiguous devices and probabilistic devices in terms of the resulting optimal total efforts, $TE_{AP}(\cdot)$ and $TE_{BP}(\cdot)$ in the following corollary.

Corollary 1 *The organizer strictly benefits from using ambiguous devices than using probabilistic device, i.e., $TE_{AP}(\mu_0) > TE_{BP}(\mu_0)$ if and only if $\mu_0 < \mu_{opt}$.*

Proof. By Theorem 1(ii), if $\mu_0 \geq \mu_{opt}$, $TE_{AP}(\mu_0) = TE(\mu_0)$, which means that no disclosure is the optimal ambiguous mechanism. It then remains to consider the case $\mu_0 < \mu_{opt}$. On one hand, Theorem 1(i) implies that $TE_{AP}(\mu_0) = TE(\mu_{opt})$ if $\mu_0 < \mu_{opt}$. On the other hand, Zhang and Zhou (2016) shows that with a binary distribution, the optimal probabilistic device is either full disclosure or full concealment, which implies that $TE_{BP}(\mu_0)$ equals either $\mu_0 TE(1) + (1 - \mu_0) TE(0)$ or $TE(\mu_0)$. By Lemma 2, μ_{opt} is unique, and therefore for any $\mu \in [0,1]$, $TE(\mu)$ is strictly less than $TE(\mu_{opt})$, unless $\mu = \mu_{opt}$. Corollary 1 thus follows. ■

Corollary 1 provides the sufficient and necessary condition under which the optimal ambiguous device strictly outperforms the optimal probabilistic device. Note that $TE_{AP}(\mu_0) \geq TE_{BP}(\mu_0)$ always holds.

Remark 1 *$TE_{BP}(\cdot)$ is obtained when both the players and the organizer are expected utility (EU) maximizers, while $TE_{AP}(\cdot)$ is obtained when both the players and the organizer are maxmin expected utility (MMEU) maximizers.*

It is worth to noting that if a maxmin expected utility (MMEU) organizer can strictly benefit from using ambiguous persuasion, a expected utility (EU) organizer can also do, but not vice versa. In the following, we provide an example to illustrute the comparison of $TE_{AP}(\cdot)$ and $TE_{BP}(\cdot)$.

Example 2 (Continued): We maintain the assumptions of $v_B^H = 20$, $v_B^L = 1$, and $v_A = 1$. In this example, $TE(\cdot)$ is concave, which can be easily verified using (2) (see also Figure 2). For a better illustration, we plot the optimal effort $TE_{AP}(\cdot)$ induced by an ambiguous device and the optimal effort $TE_{BP}(\cdot)$ induced by a probabilistic device in the following figure. From Theorem 1, when $\mu_0 \leq \mu_{opt}$, the optimal effort induced by the optimal ambiguous device $TE_{AP}(\mu_0) = TE(\mu_{opt})$, which is greater than $TE_{BP}(\mu_0) = TE(\mu_0)$; when $\mu_0 > \mu_{opt}$, $TE_{AP}(\mu_0)$ coincides with $TE_{BP}(\mu_0)$, which equals $TE(\mu_0)$.

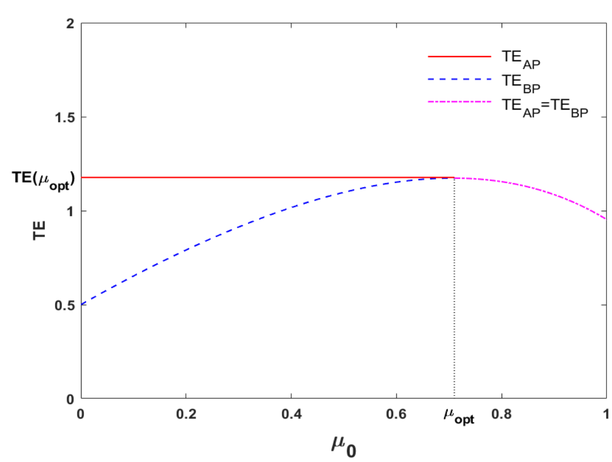


Figure 3: Comparison of $TE_{AP}(\cdot)$ and $TE_{BP}(\cdot)$.

4 Discussions

In this Section, we advance our analysis by examining the following aspects. Specifically, we study the impact of μ_0, v_A, v_B^H and v_B^L on the total effort in Section 4.1. We show that it is without loss of generality to focus on two devices, i.e., $K = 2$ in Section 4.2 and to focus on the message space $M = \{h, l\}$ in Section 4.3, respectively. In Section 4.4, we provide a short discussion on the applicability of our analysis to other settings.

4.1 Comparative Statics

By Theorem 1, $TE_{AP}(\mu_0) = TE(\mu_{opt})$ when $\mu_0 < \mu_{opt}$ and $TE_{AP}(\mu_0) = TE(\mu_0)$ when $\mu_0 \geq \mu_{opt}$. As the cutoff μ_{opt} plays a crucial role, we first consider how the cutoff varies with v_A in the following.

Proposition 4 *μ_{opt} increases with v_A.*

Proof. We rely on the result of Milgrom and Shannon (1994) to show the lemma. To do so, we prove that $TE(\mu, v_A) = K(\mu, v_A)(\mu\sqrt{v_B^H} + (1-\mu)\sqrt{v_B^L})$ obeys single-crossing property, where $K(\mu, v_A) = \frac{\frac{\mu}{\sqrt{v_B^H}} + \frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}}$. The details are relegated into the appendix. ■

The result says that in order to induce the greaterst effort, the optimal persuasion requires to persuade player A to believe player B is more likely a strong one as player A's value grows.

We next investigate how total effort varies with μ_0, v_A, v_B^H and v_B^L, respectively. Since state is binary, the corresponding prior distribution $(\mu_0, 1-\mu_0)$ on $\{v_B^H, v_B^L\}$ exhibits first-order stochastic dominance as the prior μ_0 increases, i.e., $(\mu_0', 1-\mu_0') \overset{FOSD}{\succ} (\mu_0'', 1-\mu_0'')$, when $\mu_0' > \mu_0''$.

Proposition 5 *(i) As μ_0 increases from 0 to 1, the optimal effort either stays constant or first remains constant then falls.*

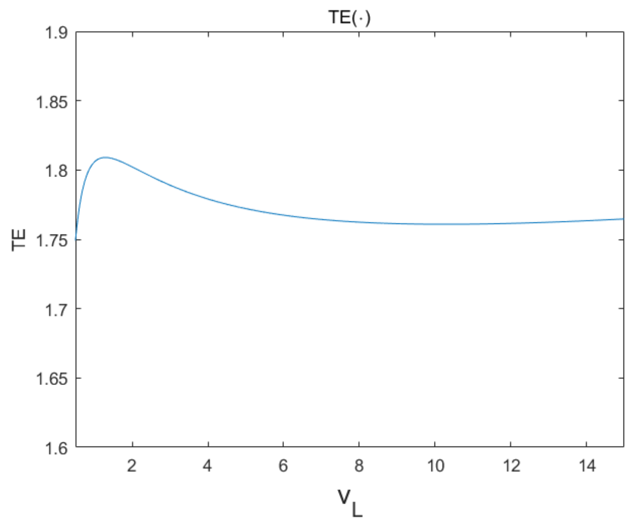
(ii) The total effort increases with v_A.

(iii) The total effort increases with v_H.

(iv) The total effort changes non-monotonically with v_L.

Proof. (i) Consider two cases: $\mu_{opt} = 1$ and $\mu_{opt} < 1$, the proposition follows directly from Theorem 1. For (ii)-(iii), we prove the results by considering the corresponding derivative. As for (iv), we consider an example by letting $v_B^H = 15$, $v_A = 2$, $\mu = 0.9$, and the resulting total effort is non-monotone in v_B^L, since $\frac{dTE}{dv_B^L}|_{v_B^L=0.6} \approx 0.182\,14 > 0$, $\frac{dTE}{dv_B^L}|_{v_B^L=2} \approx -1.368\,5 \times 10^{-2} < 0$, and $\frac{dTE}{dv_B^L}|_{v_B^L=14} \approx 1.195\,2 \times 10^{-3} > 0$. The details are relegated into the appendix. ■

By letting $v_B^H = 15$; $v_A = 2$; $\mu = 0.9$, we plot the total effort as a function of v_B^L as follows.



Example: Proposition 5 (iv)

For Proposition 5(i), when μ_0 is in a low range (i.e., $\mu_0 < \mu_{opt}$), the organizer can always persuade the uninformed player to take belief μ_{opt} by disclosing ambiguous information so that the optimal effort stays constant in that range; when μ_0 is in a high range (i.e., $\mu_0 \geq \mu_{opt}$), the uninformed player initially believes that his opponent is more likely a high-type, which discourges the uninformed and causes a fall in the total effort. Moreover, Theorem 1(ii) says that no information can induce the maximal total effort in that case.

To explain Proposition 5(ii), note that player A with a higher value v_A would spend larger effort, which would further sitimulate his opponent when player B is not overly weak.[18] Regarding Proposition 5(iii) and (iv), it states that a higher v_H would always increase the total effort, while a higher v_L would not. As v_L (or v_H) increases, player B would certainly raise his own effort, which would initially stimulate but later discourage player A as player B becomes overly stronger. That is why the total effort changes non-monotonically with v_L. It is worth noting that a v_H-type player would spend more effort than a v_L-type player, since

[18] Recall that we assume that $v_B^L \geq v_A/4$ throughout the paper.

the marginal effort cost of a v_H-type is less costly. This helps to explain why the resulting total effort always rises as v_H increases.

4.2 The Sufficiency of $K = 2$

When analyzing ambiguous devices, our proofs of Propositions 1 and 3 do not depend on the number of the probabilistic devices used, K. It in fact suffices to focus on $K = 2$, which has been shown in Proposition 1 of Beauchêne, Li, and Li (2019). In the following, we provide a more intuitive proof in our context.

Proposition 6 *Focusing on the ambiguous devices that consist of two probabilistic devices, i.e., $K = 2$, does not involve a loss of generality.*

Suppose that Π_1, which consists of $\{\pi_1, ..., \pi_K\}$, is the maxmin effort-maximizing ambiguous device, which induces a set of distributions over posteriors, $(\tau_k)_{k\in K}$. Each distribution over posteriors, τ_k, must be Bayes plausible, i.e., $\tau_k(h)\mu_h^k + \tau_k(l)\mu_l^k = \mu_0$. The resulting maxmin expected total effort equals

$$\min_{k\in\{1,...,K\}} \left[\tau_k(h)TE(\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\}) + \tau_k(l)TE(\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\})\right]. \tag{5}$$

We now prove that the optimal value above can be induced by an ambiguous device with $K = 2$. First, there must exist $\mu_h^{k(h)}$ and $\mu_l^{k(l)}$ such that $\mu_h^{k(h)} = \max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\})$ and $\mu_l^{k(l)} = \max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\}$. We next consider the ambiguous device Π_2 that consists of $\{\pi_{k(h)}, \pi_{k(l)}\}$, which leads to a maxmin expected total effort equal to

$$\min_{k\in\{k(h),k(l)\}} \left[\tau_k(h)TE(\max\{\mu_h^{k(h)}, \mu_h^{k(l)}\}) + \tau_k(l)TE(\max\{\mu_l^{k(h)}, \mu_l^{k(l)}\})\right]. \tag{6}$$

We now prove that the total effort in (5) must coincide with the total effort in (6). On one hand, since Π_1 is the maxmin effort-maximizing ambiguous device, (5) must be greater than (6). On the other hand, it follows from the construction that $TE(\max\{\mu_h^1, \mu_h^2, ..., \mu_h^K\}) = TE(\max\{\mu_h^{k(h)}, \mu_h^{k(l)}\})$ and $TE(\max\{\mu_l^1, \mu_l^2, ..., \mu_l^K\}) = TE(\max\{\mu_l^{k(h)}, \mu_l^{k(l)}\})$; as a result, (5) must be less than (6), since $\{k(h), k(l)\} \subset \{1, ..., K\}$.

4.3 The Sufficiency of $M = \{h, l\}$

Consider a finite message space M with $|M| \geq 2$, upon receiving a message $m \in M$, player A updates his belief probability-by-probability, which leads to the following set of posteriors:

$$\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}.$$

We solve the contest game with multiple posteriors as in the Section 3.2 and obtain Lemma 1. Analogous to (4), the organizer's problem can be rewritten as

$$\sup_{R \in \mathcal{R}} \min_{\tau_k \in R} \left[\sum_{m \in M} \tau_k(m) TE(\max\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}) \right]. \tag{7}$$

We show that Propositions 1 and 3 remain valid in the following.

Proposition 7 *Focusing on $M = \{h, l\}$ does not involve loss of generality. In particular, Propositions 1 and 3 hold for a finite message space M with $|M| \geq 2$.*

Proof.

(i) Recall that $TE(\mu_{opt}) \geq TE(\mu)$, $\forall \mu \in [0, 1]$, which follows directly from Definition 1. Therefore, for any distribution $\tau_k \in \Delta M$, we have

$$TE(\mu_{opt}) \geq \sum_{m \in M} \tau_k(m) TE(\max\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}).$$

As a result,

$$TE(\mu_{opt}) \geq \sup_{R \in \mathcal{R}} \min_{\tau_k \in R} \left[\sum_{m \in M} \tau_k(m) TE(\max\{\mu_m^1, \mu_m^2, ..., \mu_m^K\}) \right],$$

which is given by (4).

Moreover, when $\mu_0 \leq \mu_{opt}$, $TE(\mu_{opt})$ can be achieved by the information policies described in Proposition 1. Therefore, the expected total effort equals $TE(\mu_{opt})$.

(ii) In the proof of Proposition 3, we derive the maximal projection of the concave closure of the total effort function, $\overline{TE}(\mu_0)$, which yields the greatest expected total effort that results from the optimal ambiguous persuasion. It thus suffices to show that $\overline{TE}(\mu_0)$ remains the same for the messages space M with $|M| \geq 2$. By checking the Step 1 in the proof of Proposition 3, $\overline{TE}(\mu_0)$ remains $\max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} \mathbf{cav}TE(\max\{\mu_0, \mu^2, ..., \mu^K\})$, which gives the maximum total effort that results from the optimal *ambiguous* persuasion. ■

4.4 A Short Discussion

In this paper, we adopt the two-state setting of one-sided incomplete-information lottery contest. In this subsection, we discuss the extent to which we expect our results to apply or not apply in a more general contest setting, including two-state settings but with different contest technologies, settings with three or more states, and contests with two-sided information asymmetry.

In the current setting, our main result consists two parts: first, the optimal effort could reach the maximal value $TE(\mu_{opt})$ when the initial prior is in the low range ($\mu_0 < \mu_{opt}$); second, no disclosure is optimal when the prior is in the high range ($\mu_0 \geq \mu_{opt}$). We conjecture that the main insight of the first part could be extended to other two-player two-state contest settings, provided that the total effort function TE can be characterized. If the total effort function is continuous over the belief within $[0, 1]$, one could define μ_{opt} as in Definition 1 and construct an information device as in Definition 2, which always induces a maxmin player to take the desired belief μ_{opt}.[19] While for the second part, no disclosure could not be the optimal in general, even within Bayesian devices, for example, Chen, Kuang, and Zheng (2019) consider a two-player all-pay auction contest with one-sided asymmetric information. In this case, one could apply the concavification approach by Kamenica and Gentzkow (2011) and Beauchêne, Li, and Li (2019) to characterize the optimal value of ambiguous persuasion and construct a Bayesian/probabilistic device to achieve the optimum. When applying the aforementioned approach, several difficulties may arise, including the explicit characterization of the effort strategies involved, the concavification of the total effort function, and the construction of the optimal device.[20]

The same difficulties carry on to settings with three or more states. More specifically, there is a lack of a closed-form characterization of optimal Bayesian persuasion, although Zhang and Zhou (2016) provide an algorithm to search for optimal Bayesian persuasion with at least three states in this contest game, an explicit formula is still unavailable. Moreover,

[19]If μ_{opt} is not unique, it is without loss to take the maximum.

[20]To identify the greatest expected total effort, the concavification approach requires to compute the maximal projection of the concave closure of the total effort function.

it is not straightforward to apply concavification procedure when going beyond binary cases, since now the expected total function must be multi-dimensional. In addition, with this binary-state assumption, all prior distributions are completely ordered using the first-order stochastic dominance. However, not all prior distributions are ordered for three or more states, which further complicate the problem.

It is also reasonable to consider a contest with two-sided information asymmetry wherein both contestants are uninformed of each other's types. As Zhang and Zhou (2016) point out "It is natural to ask what occurs if both contestants possess private information. Several technical challenges emerge accordingly, [...] a characterization of the equilibrium is usually not obtainable." Due to the lack of a closed-form solution, one has to develop an approach to circumvent this difficulty to derive the optimal ambiguous signals.

5 Concluding Remarks

We reinvestigate the optimal design of an information structure when the organizer is allowed to disclose ambiguous information in a two-player contest. One contestant's valuation is publicly known, while the other's valuation is his private type, which follows a binary distribution. We characterize the optimal information structure through an ambiguous persuasion approach. The effort-maximizing disclosure policy depends on the prior. More precisely, it is optimal to either induce maximal ambiguity or fully conceal information. By introducing ambiguity, a contest organizer can persuade the uninformed contestant to "overestimate" the strength of his opponent. In contrast, such a contestant never "underestimates" his opponent on average under ambiguity. In the current setting, we show that an effort-maximizing organizer can benefit from generating ambiguous information in many circumstances. There still many other settings worth investigating. We leave these directions to future work.

Appendix

This appendix covers the proofs of Lemma 2, Propositions 2, 3, 4, and 5.

Proof of Lemma 2

Recall that $TE(\mu) = K(\mu)\left(\mu\sqrt{v_B^H} + (1-\mu)\sqrt{v_B^L}\right)$, where $K(\mu) = \frac{\frac{\mu}{\sqrt{v_B^H}} + \frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}}$. It follows from direct calculation that

$$K'(\mu) = \left(\frac{1}{\sqrt{v_B^H}} - \frac{1}{\sqrt{v_B^L}}\right) \frac{\frac{1}{v_A} - \frac{1}{\sqrt{v_B^H v_B^L}}}{\left(\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}\right)^2} \begin{cases} < 0, & \text{when } v_A < \sqrt{v_B^H v_B^L}; \\ = 0, & \text{when } v_A = \sqrt{v_B^H v_B^L}; \\ > 0, & \text{when } v_A > \sqrt{v_B^H v_B^L}. \end{cases}$$

and

$$K''(\mu) = -2\left(\frac{1}{\sqrt{v_B^H}} - \frac{1}{\sqrt{v_B^L}}\right)^2 \frac{\frac{1}{v_A} - \frac{1}{\sqrt{v_B^H v_B^L}}}{\left(\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}\right)^3} \begin{cases} < 0, & \text{when } v_A < \sqrt{v_B^H v_B^L}; \\ = 0, & \text{when } v_A = \sqrt{v_B^H v_B^L}; \\ > 0, & \text{when } v_A > \sqrt{v_B^H v_B^L}. \end{cases}$$

Hence, we have

$$\begin{aligned} TE'(\mu) &= K'(\mu)\left(\mu\sqrt{v_B^H} + (1-\mu)\sqrt{v_B^L}\right) + K(\mu)\left(\sqrt{v_B^H} - \sqrt{v_B^L}\right) \\ &> 0, \text{ when } v_A \geq \sqrt{v_B^H v_B^L}. \end{aligned}$$

and

$$\begin{aligned} TE''(\mu) &= K''(\mu)\left(\mu\sqrt{v_B^H} + (1-\mu)\sqrt{v_B^L}\right) + 2K'(\mu)\left(\sqrt{v_B^H} - \sqrt{v_B^L}\right) \\ & \begin{cases} < 0, & \text{when } v_A < \sqrt{v_B^H v_B^L}; \\ = 0, & \text{when } v_A = \sqrt{v_B^H v_B^L}; \\ > 0, & \text{when } v_A > \sqrt{v_B^H v_B^L}. \end{cases} \end{aligned}$$

Therefore, if $v_A \geq \sqrt{v_B^H v_B^L}$, $TE'(\mu) > 0$, which implies that $\mu_{opt} = 1$; if $v_A < \sqrt{v_B^H v_B^L}$, $TE''(\mu) < 0$, which implies the uniqueness of μ_{opt}.

To solve for μ_{opt} explicitly, recall that $\mu_{opt} = \arg\max_{\mu\in[0,1]} TE(\mu)$. We next apply Lagrange method with Karush-Kuhn-Tucker conditions. μ_{opt} maximizes $L(\mu) = TE(\mu) + \lambda\mu + \gamma(1-\mu)$ and thus satisfies

$$\begin{aligned} TE'(\mu) + \lambda - \gamma &= 0; \\ \lambda\mu &= 0; \\ \gamma(1-\mu) &= 0. \end{aligned}$$

In the following, we prove that $\mu_{opt} = 0$ never occurs by contradiction. Suppose that $\mu_{opt} = 0$, which means that $\gamma = 0$ and $TE'(\mu)|_{\mu=0} = -\lambda < 0$. By direct substitution, we have

$$TE'(\mu)|_{\mu=0} = \left(\frac{1}{\sqrt{v_B^H}} - \frac{1}{\sqrt{v_B^L}}\right) \frac{\frac{1}{v_A} - \frac{1}{\sqrt{v_B^H v_B^L}}}{\left(\frac{1}{v_A} + \frac{1}{\sqrt{v_B^L}}\right)^2} \sqrt{v_B^L} + \frac{\frac{1}{\sqrt{v_B^L}}}{\frac{1}{v_A} + \frac{1}{v_B^L}} \left(\sqrt{v_B^H} - \sqrt{v_B^L}\right) < 0,$$

which is equivalent to

$$\frac{\sqrt{v_B^H} - \sqrt{v_B^L}}{\frac{1}{v_A} + \frac{1}{\sqrt{v_B^L}}} \left[\frac{1}{\sqrt{v_B^H}} \frac{\frac{1}{\sqrt{v_B^H v_B^L}} - \frac{1}{v_A}}{\frac{1}{v_A} + \frac{1}{v_B^L}} + \frac{1}{\sqrt{v_B^L}} \right] < 0,$$

i.e.,

$$\left(\sqrt{\frac{v_B^L}{v_B^L}} \frac{1}{v_B^L} + \frac{1}{\sqrt{v_B^H v_B^L}}\right) v_A < 1 - \sqrt{\frac{v_B^H}{v_B^L}}.$$

which contradicts, since $\left(\sqrt{\frac{v_B^L}{v_B^L}} \frac{1}{v_B^L} + \frac{1}{\sqrt{v_B^H v_B^L}}\right) v_A > 0$ and $1 - \sqrt{\frac{v_B^H}{v_B^L}} < 0$.

As a result, $\mu_{opt} \neq 0$, which implies that $\lambda = 0$ and $TE'(\mu)|_{\mu=\mu_{opt}} = \gamma \geq 0$. If $\gamma > 0$, $\mu_{opt} = 1$. If $\gamma = 0$, μ_{opt} must solve $TE'(\mu) = 0$.

We therefore conclude that if $v_A < \sqrt{v_B^H v_B^L}$, $\mu_{opt} = \begin{cases} \widehat{\mu}, & \text{if } \widehat{\mu} < 1 \\ 1, & \text{if } \widehat{\mu} \geq 1 \end{cases}$, where $\widehat{\mu}$ solves $TE'(\mu) = 0$. As a by-product, $TE'(\mu)|_{\mu=\mu_{opt}} \geq 0$. Therefore, if $v_A < \sqrt{v_B^H v_B^L}$, $\mu_{opt} = \min\{\widehat{\mu}, 1\}$, where $TE'(\widehat{\mu}) = 0$.

Proof of Proposition 2

Our goal is to construct all optimal ambiguous devices with $K = 2$. Consider an ambiguous device Π that consists of two probabilistic devices $\{\pi_1, \pi_2\}$, denote the induced posteriors μ_1^m and μ_2^m when message is $m \in \{h, l\}$. An ambiguous device is optimal if and only if the maximal posteriors always coinciding μ_{opt}, which means $\max\{\mu_1^h, \mu_2^h\} = \max\{\mu_1^l, \mu_2^l\} = \mu_{opt}$. As a result, there are only two possibilities: either $\mu_1^l = \mu_2^h = \mu_{opt}$ or $\mu_1^h = \mu_2^l = \mu_{opt}$. In addition, recall that for each Bayesian device π_i, the expected posteriors must equal prior μ_0, which implies that if $\mu_i^h > \mu_0$, $\mu_i^l < \mu_0$ must hold. Since $\mu_{opt} > \mu_0$, if $\mu_i^h = \mu_{opt}$ (resp. $\mu_i^l = \mu_{opt}$), we must have $\mu_i^l < \mu_0$ (resp. $\mu_i^h < \mu_0$), $\forall i \in \{1, 2\}$.

Consider that $(\mu_1^h, \mu_1^l) = (\mu', \mu_{opt})$ and $(\mu_2^h, \mu_2^l) = (\mu_{opt}, \mu'')$, where μ', $\mu'' < \mu_0$. We next show how to recover the corresponding π_1 and π_2. It follows from Bayes' rule that

$$\begin{aligned}
\mu_1^h &= \frac{\mu_0 \pi_1(h|v_B^H)}{\mu_0 \pi_1(h|v_B^H) + (1-\mu_0)\pi_1(h|v_B^L)} = \mu', \\
\mu_1^l &= \frac{\mu_0 \pi_1(l|v_B^H)}{\mu_0 \pi_1(l|v_B^H) + (1-\mu_0)\pi_1(l|v_B^L)} = \frac{\mu_0\left[1-\pi_1(h|v_B^H)\right]}{\mu_0\left[1-\pi_1(h|v_B^H)\right] + (1-\mu_0)\left[1-\pi_1(h|v_B^L)\right]} = \mu_{opt}; \\
\mu_2^h &= \frac{\mu_0 \pi_2(h|v_B^H)}{\mu_0 \pi_2(h|v_B^H) + (1-\mu_0)\pi_2(h|v_B^L)} = \mu_{opt}, \\
\mu_2^l &= \frac{\mu_0 \pi_2(l|v_B^H)}{\mu_0 \pi_2(l|v_B^H) + (1-\mu_0)\pi_2(l|v_B^L)} = \frac{\mu_0\left[1-\pi_2(h|v_B^H)\right]}{\mu_0\left[1-\pi_2(h|v_B^H)\right] + (1-\mu_0)\left[1-\pi_2(h|v_B^L)\right]} = \mu''.
\end{aligned}$$

Solving the above system yields

$$\begin{aligned}
\pi_1(h|v_B^H) &= \frac{\mu'\left(\mu_{opt}-\mu_0\right)}{\mu_0(\mu_{opt}-\mu')},\ \pi_1(h|v_B^L) = \frac{(1-\mu')\left(\mu_{opt}-\mu_0\right)}{(1-\mu_0)\left(\mu_{opt}-\mu'\right)}; \\
\pi_2(h|v_B^H) &= \frac{\mu_{opt}\left(\mu_0-\mu''\right)}{\mu_0(\mu_{opt}-\mu'')},\ \pi_2(h|v_B^L) = \frac{\left(1-\mu_{opt}\right)\left(\mu_0-\mu''\right)}{(1-\mu_0)\left(\mu_{opt}-\mu''\right)}.
\end{aligned}$$

Analoguously, if $(\mu_1^h, \mu_1^l) = (\mu_{opt}, \mu')$ and $(\mu_2^h, \mu_2^l) = (\mu'', \mu_{opt})$, one could apply the same procedure to derive π_1 and π_2, which are given by

$$\pi_1(l|v_B^H) = \frac{\mu'\left(\mu_{opt}-\mu_0\right)}{\mu_0(\mu_{opt}-\mu')},\ \pi_1(l|v_B^L) = \frac{(1-\mu')\left(\mu_{opt}-\mu_0\right)}{(1-\mu_0)\left(\mu_{opt}-\mu'\right)};$$

$$\pi_2(l|v_B^H) = \frac{\mu_{opt}(\mu_0-\mu'')}{\mu_0(\mu_{opt}-\mu'')},\ \pi_2(l|v_B^L) = \frac{(1-\mu_{opt})(\mu_0-\mu'')}{(1-\mu_0)(\mu_{opt}-\mu'')}.$$

One could easily check the optimality of the constructed devices. The posteriors induced by each message satisfy $\max\{\mu_1^h, \mu_2^h\} = \max\{\mu_1^l, \mu_2^l\} = \mu_{opt}$. As a result, player A would take belief μ_{opt}, regardless of the realized message. The optimality of the constructed devices thus follows.

Proof of Proposition 3

Recall that the organizer's objective is to maximize his ex ante *maxmin* payoff, which is the expected total effort. When $\mu_0 = \mu_{opt}$, $TE(\mu_{opt})$ can be induced by no disclosure; When $\mu_0 > \mu_{opt}$, we take three steps to solve the problem in (4) as follows.

Step 1: In order to derive the greatest maxmin expected total effort, we apply the technique developed by Beauchêne, Li, and Li (2019) in the following. Given that player A holds the posterior beliefs $\{\mu^1, \mu^2, ..., \mu^K\}$, we will first solve for the concave closure of the subgraph of TE, i.e., $\mathbf{cav}TE(\mu^1, \mu^2, ..., \mu^K)$, and subsequently derive its maximal projection $\overline{TE}(\mu_0)$, which gives the optimal total effort.

It follows from Lemma 1 that the total effort $TE(\mu^1, \mu^2, ..., \mu^K) = TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$. The subgraph of TE is $\{(\mu^1, \mu^2, ..., \mu^K, z)$ s.t. $TE(\max\{\mu^1, \mu^2, ..., \mu^K\}) \geq z\}$. We denote its concave closure as $\mathbf{cav}TE(\mu^1, \mu^2, ..., \mu^K) := \sup\{z|(\mu^1, \mu^2, ..., \mu^K, z) \in co(Subgrah(TE))\}$. Note that $\mathbf{cav}TE(\mu^1, \mu^2, ..., \mu^K) = \mathbf{cav}TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$.

Therefore, the maximal projection of $\mathbf{cav}TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$ is

$$\overline{TE}(\mu^1) = \max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} \mathbf{cav}TE(\max\{\mu^1, \mu^2, ..., \mu^K\}).$$

Beauchêne, Li, and Li (2019) show that $\overline{TE}(\mu_0)$ is the optimal value that results from the optimal *ambiguous* persuasion. Zhang and Zhou (2016) study the optimal probabilistic device, i.e., $K = 1$. They find that $TE(\cdot)$ is either concave or convex, and therefore $\mathbf{cav}TE(\mu)$ equals either $TE(\mu)$ or $\mu TE(1)+(1-\mu)TE(0)$. To investigate ambiguous devices with $K \geq 2$, we discuss three cases according to the convexity of $TE(\cdot)$.

(i) When $TE(\cdot)$ is convex, the concave closure of $TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$ equals

$$\mathbf{cav}TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$$

$$
\begin{aligned}
&= \mathbf{cav}TE(\max\{\mu^1, \max\{\mu^2, ..., \mu^K\}\}) \\
&= \max\{\mu^1, \max\{\mu^2, ..., \mu^K\}\}TE(1) + (1 - \max\{\mu^1, \max\{\mu^2, ..., \mu^K\}\})TE(0),
\end{aligned}
$$

since $\mathbf{cav}TE(\mu) = \mu TE(1) + (1-\mu)TE(0)$, which follows from the convexity of $TE(\cdot)$.

Beauchêne, Li, and Li (2019) prove that the optimal value equals $\overline{TE}(\mu_0)$. In our model,

$$
\begin{aligned}
&\overline{TE}(\mu_0) \\
= &\max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} \mathbf{cav}TE(\max\{\mu_0, \mu^2, ..., \mu^K\}) \\
= &\max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} \max\{\mu_0, \max\{\mu^2, ..., \mu^K\}\}TE(1) + (1 - \max\{\mu_0, \max\{\mu^2, ..., \mu^K\}\})TE(0) \\
= &\mu_0 TE(1) + (1-\mu_0)TE(0).
\end{aligned}
$$

The last equality holds since $\mu_{opt} = 0$, i.e., $TE(0) > TE(\mu)$, $\forall \mu \in (0,1]$. To see this, when $TE(\cdot)$ is convex, either $\mu_{opt} = 0$ or $\mu_{opt} = 1$. Since $\mu_0 > \mu_{opt}$, we must have $\mu_{opt} = 0$.[21]

(ii) When $TE(\cdot)$ is linear, $TE(\cdot)$ is either (weakly) increasing or (weakly) decreasing, and $\mu_0 > \mu_{opt}$ implies that $TE(\cdot)$ must be (weakly) decreasing.[22] As a result,

$$
\begin{aligned}
&\overline{TE}(\mu_0) \\
= &\max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} \mathbf{cav}TE(\max\{\mu_0, \mu^2, ..., \mu^K\}) \\
= &\max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} TE(\max\{\mu_0, \mu^2, ..., \mu^K\}) \\
= &TE(\mu_0),
\end{aligned}
$$

since $TE(\cdot)$ is a weakly decreasing linear function. In particular, $TE(\mu_0) = \mu_0 TE(1) + (1 - \mu_0)TE(0)$.

(iii) When $TE(\cdot)$ is concave, the concave closure of $TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$ equals

$$
\mathbf{cav}TE(\max\{\mu^1, \mu^2, ..., \mu^K\}) = TE(\max\{\mu^1, \mu^2, ..., \mu^K\}),
$$

[21] When $\mu_{opt} = 1$, $\mu_0 \leq \mu_{opt}$, we go back to Proposition 1.

[22] If $TE(\cdot)$ is a constant, μ_{opt} can be any value between 0 and 1 and $\mu_0 > \mu_{opt}$ is impossible; if $TE(\cdot)$ is an increasing function, $\mu_{opt} = 1$ and $\mu_0 > \mu_{opt}$ never occurs.

using $\mathbf{cav}TE(\mu) = TE(\mu)$. Therefore, the maximal projection of $\mathbf{cav}TE(\max\{\mu^1, \mu^2, ..., \mu^K\})$ is

$$
\begin{aligned}
& \overline{TE}(\mu_0) \\
= & \max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} \mathbf{cav}TE(\max\{\mu_0, \mu^2, ..., \mu^K\}) \\
= & \max_{(\mu^2,...,\mu^K)\in(\Delta\Omega)^{K-1}} TE(\max\{\mu_0, \mu^2, ..., \mu^K\}) \\
= & TE(\mu_0),
\end{aligned}
$$

as the prior $\mu_0 > \mu_{opt}$.

Step 2: We construct probabilistic devices to induce the optimal effort identified in Step 1 as follows.

(i) When $TE(\cdot)$ is convex, full disclosure could induce the optimal expected total effort $\mu_0 TE(1) + (1 - \mu_0)TE(0)$.

(ii) When $TE(\cdot)$ is linear, the optimal effort, $TE(\mu_0) = \mu_0 TE(1) + (1 - \mu_0)TE(0)$, can be induced by any probabilistic device.

(iii) When $TE(\cdot)$ is concave, the optimal effort, $TE(\mu_{opt}) = TE(\mu_0)$, can be induced by no disclosure.

Step 3: By Lemma 2, when $v_A > \sqrt{v_B^H v_B^L}$, $\mu_{opt} = 1$, which implies that $\mu_0 > \mu_{opt}$ never occurs. It thus suffices to consider two remaining cases $v_A = \sqrt{v_B^H v_B^L}$ and $v_A < \sqrt{v_B^H v_B^L}$. Note that $v_A =$ (resp. $<$)$\sqrt{v_B^H v_B^L}$ if and only if $TE(\cdot)$ is linear (resp. strictly concave).

Combining the results, we conclude that if $\mu_0 = \mu_{opt}$, $TE(\mu_{opt})$ can be induced by no disclosure; if $\mu_0 > \mu_{opt}$, the following results hold.

(i) When $v_A = \sqrt{v_B^H v_B^L}$, the optimal total effort equals $TE(\mu_0) = (1 - \mu_0)TE(0) + \mu_0 TE(1)$, which can be induced by full/no disclosure or any other Bayesian device;

(ii) When $v_A < \sqrt{v_B^H v_B^L}$, the optimal total effort equals $TE(\mu_0)$, which can be induced by no disclosure. Proposition 3 thus follows.

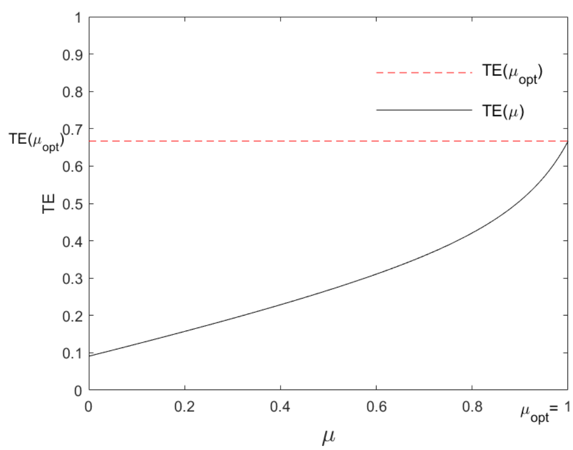
Proof of Proposition 4

By Lemma 2, it suffices to show that $\widehat{\mu}$ increases with v_A, where $\widehat{\mu}$ solves $TE'(\mu) = 0$. By Milgrom and Shannon, it is equivalent to prove that $TE(\mu, v_A) = K(\mu, v_A)(\mu\sqrt{v_B^H} +$

$(1-\mu)\sqrt{v_B^L})$ obeys single-crossing property, where $K(\mu, v_A) = \frac{\frac{\mu}{\sqrt{v_B^H}} + \frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A} + \frac{\mu}{v_B^H} + \frac{1-\mu}{v_B^L}}$. For $\mu'' > \mu'$, let $\delta(v_A) = TE(\mu'', v_A) - TE(\mu', v_A)$. To verify the single-crossing property, we need to show that $\delta(v_A') \geq 0$ implies that $\delta(v_A'') \geq 0$. More detailedly, $\delta(v_A') = K(\mu'', v_A')(\mu''\sqrt{v_B^H} + (1-\mu'')\sqrt{v_B^L}) - K(\mu', v_A')(\mu'\sqrt{v_B^H} + (1-\mu')\sqrt{v_B^L})$ and $\delta(v_A'') = K(\mu'', v_A'')(\mu''\sqrt{v_B^H} + (1-\mu'')\sqrt{v_B^L}) - K(\mu', v_A'')(\mu'\sqrt{v_B^H} + (1-\mu')\sqrt{v_B^L})$. Note that $\delta(v_A'') \geq 0$ is equivalent to show that $\frac{K(\mu'', v_A'')}{K(\mu', v_A'')} \geq \frac{\mu'\sqrt{v_B^H} + (1-\mu')\sqrt{v_B^L}}{\mu''\sqrt{v_B^H} + (1-\mu'')\sqrt{v_B^L}}$. It then suffices to prove that $\frac{K(\mu'', v_A'')}{K(\mu', v_A'')} \geq \frac{K(\mu'', v_A')}{K(\mu', v_A')}$, i.e., $\frac{\frac{1}{v_A''} + \frac{\mu'}{v_B^H} + \frac{1-\mu'}{v_B^L}}{\frac{1}{v_A''} + \frac{\mu''}{v_B^H} + \frac{1-\mu''}{v_B^L}} \geq \frac{\frac{1}{v_A'} + \frac{\mu'}{v_B^H} + \frac{1-\mu'}{v_B^L}}{\frac{1}{v_A'} + \frac{\mu''}{v_B^H} + \frac{1-\mu''}{v_B^L}}$. This holds, since $F(v_A) = \frac{\frac{1}{v_A} + \frac{\mu'}{v_B^H} + \frac{1-\mu'}{v_B^L}}{\frac{1}{v_A} + \frac{\mu''}{v_B^H} + \frac{1-\mu''}{v_B^L}}$ increases with v_A.

Proof of Proposition 5

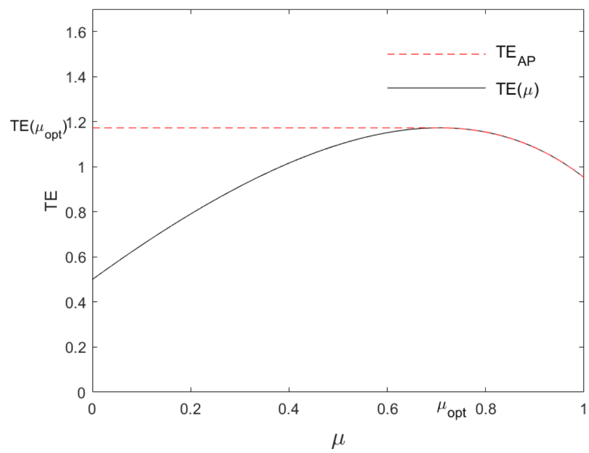
(i) If $\mu_{opt} = 1$, the first case occurs, $TE^*(\mu_0)$ stays at the level of $TE(\mu_{opt})$ as μ_0 moves from 0 to 1, since $\mu_{opt} = 1$ and $\mu_0 \leq 1$ always holds. We plot an example in the following.



Case 1: $\mu_{opt} = 1$

If $\mu_{opt} < 1$, the second case occurs, i.e., $TE^*(\mu_0)$ remains constant at the level of $TE(\mu_{opt})$

until $\mu_0 = \mu_{opt}$ and falls once μ_0 exceeds μ_{opt}. We plot an example in the following figure.



Case 2: $\mu_{opt} < 1$

(ii) The total effort $TE(\mu) = \frac{\left[\mu\sqrt{v_B^H}+(1-\mu)\sqrt{v_B^L}\right]\left[\frac{\mu}{\sqrt{v_B^H}}+\frac{1-\mu}{\sqrt{v_B^L}}\right]}{\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}}$ given by (2) increases with v_A.

(iii) We rewrite $TE(\mu) = K(\mu)\left(\mu\sqrt{v_B^H} + (1-\mu)\sqrt{v_B^L}\right)$, where $K(\mu) = \frac{\frac{\mu}{\sqrt{v_B^H}}+\frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}}$.

$$
\begin{aligned}
&\frac{dTE}{dv_B^H} \\
= \;& \frac{d}{dv_B^H}\left(\frac{\frac{\mu}{\sqrt{v_B^H}}+\frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}}\right)\left(\mu\sqrt{v_B^H}+(1-\mu)\sqrt{v_B^L}\right)+\left(\frac{\frac{\mu}{\sqrt{v_B^H}}+\frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}}\right)\frac{\mu}{2\sqrt{v_B^H}} \\
= \;& \frac{-\frac{1}{2}\frac{\mu}{v_A\sqrt{\left(v_B^H\right)^3}}+\frac{1}{2}\frac{\mu^2}{\left(v_B^H\right)^2\sqrt{v_B^H}}+\frac{\mu(1-\mu)}{\left(v_B^H\right)^2\sqrt{v_B^L}}-\frac{1}{2}\frac{\mu(1-\mu)}{\sqrt{\left(v_B^H\right)^3 v_B^L}}}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)^2}\left(\mu\sqrt{v_B^H}+(1-\mu)\sqrt{v_B^L}\right) \\
& +\frac{\frac{1}{2}\frac{\mu^2}{v_B^H}+\frac{1}{2}\frac{\mu(1-\mu)}{\sqrt{v_B^H v_B^L}}}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)^2}\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)^2}\left[\begin{array}{c} -\frac{1}{2}\frac{\mu^2}{v_Av_B^H}+\frac{1}{2}\frac{\mu^3}{\left(v_B^H\right)^3}+\frac{\mu^2(1-\mu)}{\sqrt{\left(v_B^H\right)^3v_B^L}}-\frac{1}{2}\frac{\mu^2(1-\mu)}{v_B^Hv_B^L} \\ -\frac{1}{2}\frac{\mu(1-\mu)\sqrt{v_B^L}}{v_A\sqrt{\left(v_B^H\right)^3}}+\frac{1}{2}\frac{\mu^2(1-\mu)\sqrt{v_B^L}}{\left(v_B^H\right)^2\sqrt{v_B^H}}+\frac{\mu(1-\mu)^2}{\left(v_B^H\right)^2}-\frac{1}{2}\frac{\mu(1-\mu)^2}{\sqrt{\left(v_B^H\right)^3v_B^L}} \end{array}\right] \\
&\quad +\frac{1}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)^2}\left[\begin{array}{c} \frac{1}{2}\frac{\mu^2}{v_Av_B^H}+\frac{1}{2}\frac{\mu(1-\mu)}{v_A\sqrt{v_B^Hv_B^L}}+\frac{1}{2}\frac{\mu^3}{\left(v_B^H\right)^2}+\frac{1}{2}\frac{\mu^2(1-\mu)}{\sqrt{\left(v_B^H\right)^3v_B^L}} \\ +\frac{1}{2}\frac{\mu^2(1-\mu)}{v_B^Hv_B^L}+\frac{1}{2}\frac{\mu(1-\mu)^2}{\sqrt{v_B^H\left(v_B^L\right)^3}} \end{array}\right] \\
&= \frac{1}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)^2}\left[\begin{array}{c} \frac{1}{2}\frac{\mu^3}{\left(v_B^H\right)^3}+\frac{\mu^2(1-\mu)}{\sqrt{\left(v_B^H\right)^3v_B^L}}+\frac{1}{2}\frac{\mu^2(1-\mu)\sqrt{v_B^L}}{\left(v_B^H\right)^2\sqrt{v_B^H}}+\frac{\mu(1-\mu)^2}{\left(v_B^H\right)^2} \\ +\left(\frac{1}{2}\frac{\mu(1-\mu)}{v_A\sqrt{v_B^Hv_B^L}}-\frac{1}{2}\frac{\mu(1-\mu)\sqrt{v_B^L}}{v_A\sqrt{\left(v_B^H\right)^3}}\right)+\left(\frac{1}{2}\frac{\mu(1-\mu)^2}{\sqrt{v_B^H\left(v_B^L\right)^3}}-\frac{1}{2}\frac{\mu(1-\mu)^2}{\sqrt{\left(v_B^H\right)^3v_B^L}}\right) \\ +\frac{1}{2}\frac{\mu^3}{\left(v_B^H\right)^2}+\frac{1}{2}\frac{\mu^2(1-\mu)}{\sqrt{\left(v_B^H\right)^3v_B^L}} \end{array}\right] \\
&> 0,
\end{aligned}
$$

since $\frac{1}{2}\frac{\mu(1-\mu)}{v_A\sqrt{v_B^Hv_B^L}}-\frac{1}{2}\frac{\mu(1-\mu)\sqrt{v_B^L}}{v_A\sqrt{\left(v_B^H\right)^3}}>0$ and $\frac{1}{2}\frac{\mu(1-\mu)^2}{\sqrt{v_B^H\left(v_B^L\right)^3}}-\frac{1}{2}\frac{\mu(1-\mu)^2}{\sqrt{\left(v_B^H\right)^3v_B^L}}>0$.

As for (iv), we consider an example by letting $v_B^H=15$, $v_A=2$, $\mu=0.9$, and the resulting total effort is non-monotone in v_B^L, since $\frac{dTE}{dv_B^L}|_{v_B^L=0.6}\approx 0.182\,14>0$, $\frac{dTE}{dv_B^L}|_{v_B^L=2}\approx -1.368\,5\times 10^{-2}<0$, and $\frac{dTE}{dv_B^L}|_{v_B^L=14}\approx 1.195\,2\times 10^{-3}>0$, where

$$
\begin{aligned}
&\frac{dTE}{dv_B^L} \\
&= \frac{d}{dv_B^L}\left(\frac{\frac{\mu}{\sqrt{v_B^H}}+\frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}}\right)\left(\mu\sqrt{v_B^H}+(1-\mu)\sqrt{v_B^L}\right)+\left(\frac{\frac{\mu}{\sqrt{v_B^H}}+\frac{1-\mu}{\sqrt{v_B^L}}}{\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}}\right)\frac{1-\mu}{2\sqrt{v_B^L}} \\
&= \frac{-\frac{1}{2}\frac{1-\mu}{v_A\sqrt{\left(v_B^L\right)^3}}+\frac{1}{2}\frac{\mu(1-\mu)}{v_B^H\sqrt{\left(v_B^L\right)^3}}+\frac{\mu(1-\mu)}{\left(v_B^L\right)^2\sqrt{v_B^H}}-\frac{1}{2}\frac{(1-\mu)^2}{\left(v_B^L\right)^2\sqrt{v_B^L}}}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)^2}\left(\mu\sqrt{v_B^H}+(1-\mu)\sqrt{v_B^L}\right) \\
&\quad +\frac{1}{\left(\frac{1}{v_A}+\frac{\mu}{v_B^H}+\frac{1-\mu}{v_B^L}\right)}\left[\frac{1}{2}\frac{\mu(1-\mu)}{\sqrt{v_B^Hv_B^L}}+\frac{1}{2}\frac{(1-\mu)^2}{v_B^L}\right].
\end{aligned}
$$

References

M. Ayouni and F. Koessler. Hard evidence and ambiguity aversion. *Theory and Decision*, 82(3):327–339, 2017.

D. Beauchêne, J. Li, and M. Li. Ambiguous persuasion. *Journal of Economic Theory*, 179: 312–365, 2019.

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

D. Bergemann and S. Morris. Information design, bayesian persuasion, and bayes correlated equilibrium. *American Economic Review*, 106(5):586–91, 2016b.

A. Blume and O. Board. Intentional vagueness. *Erkenntnis*, 79(4):855–899, 2014.

S. Bose and L. Renou. Mechanism design with ambiguous communication devices. *Econometrica*, 82(5):1853–1872, 2014.

S. Bose, E. Ozdenoren, and A. Pape. Optimal auctions with ambiguity. *Theoretical Economics*, 1(4):411–438, 2006.

B. Chen, X. Jiang, and D. Knyazev. On disclosure policies in all-pay auctions with stochastic entry. *Journal of Mathematical Economics*, 70:66–73, 2017.

J. Chen, Z. Kuang, and J. Zheng. Bayesian persuasion in sequential tullock contests. *Working Paper*, 2019.

J. Chen, Z. Kuang, and J. Zheng. Persuasion and timing in asymmetric-information all-pay auction contests. *Working Paper*, 2022.

Z. Chen. All-pay auctions with private signals about opponents' values. *Review of Economic Design*, 25(1-2):33–64, 2021.

X. Cheng. A concavification approach to ambiguous persuasion. *arXiv preprint arXiv:2106.11270*, 2021.

S. Deng, H. Fang, Q. Fu, and Z. Wu. Confidence management in tournaments. *Working Paper*, 2021.

P. Denter, J. Morgan, and D. Sisak. 'where ignorance is bliss,'tis folly to be wise': Transparency in contests. *Available at SSRN 1836905*, 2014.

A. d. Di Tillio, N. Kos, and M. Messner. The design of ambiguous mechanisms. *The Review of Economic Studies*, 84(1):237–276, 2016.

L. G. Epstein and M. Schneider. Learning under ambiguity. *The Review of Economic Studies*, 74(4):1275–1303, 2007.

L. G. Epstein and M. Schneider. Ambiguity and asset markets. *Annual Review of Financial Economics*, 2(1):315–346, 2010.

X. Feng. Information disclosure on the contest mechanism. *Journal of Mathematical Economics*, 91:148–156, 2020.

A. Frankel. Aligned delegation. *American Economic Review*, 104(1):66–83, 2014.

Q. Fu, Q. Jiao, and J. Lu. Disclosure policy in a multi-prize all-pay auction with stochastic abilities. *Economics Letters*, 125(3):376–380, 2014.

I. Gilboa and M. Marinacci. Ambiguity and the bayesian paradigm. In *Readings in formal epistemology*, pages 385–439. Springer, 2016.

I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

H. Guo. Mechanism design with ambiguous transfers: An analysis in finite dimensional naive type spaces. *Journal of Economic Theory*, 183:76–105, 2019.

T. M. Hurley and J. F. Shogren. Asymmetric information contests. *European Journal of Political Economy*, 14(4):645–665, 1998a.

T. M. Hurley and J. F. Shogren. Effort levels in a cournot nash contest with asymmetric information. *Journal of Public Economics*, 69(2):195–210, 1998b.

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

C. Kellner and M. T. Le Quement. Modes of ambiguous communication. *Games and Economic Behavior*, 104:271–292, 2017.

C. Kellner and M. T. Le Quement. Endogenous ambiguity in cheap talk. *Journal of Economic Theory*, 173:1–17, 2018.

D. Kovenock, F. Morath, and J. Münster. Information sharing in contests. *Journal of Economics & Management Strategy*, 24(3):570–596, 2015.

Z. Kuang, H. Zhao, and J. Zheng. Ridge distributions and information design in simultaneous all-pay auction contests. *Working Paper*, 2022.

G. Lopomo, L. Rigotti, and C. Shannon. Uncertainty in mechanism design. *arXiv preprint arXiv:2108.12633*, 2021.

J. Lu, H. Ma, and Z. Wang. Ranking disclosure policies in all-pay auctions. *Economic Inquiry*, 56(3):1464–1485, 2018.

P. Milgrom and C. Shannon. Monotone comparative statics. *Econometrica: Journal of the Econometric Society*, pages 157–180, 1994.

F. Morath and J. Münster. Private versus complete information in auctions. *Economics Letters*, 101(3):214–216, 2008.

S. Mukerji and J.-M. Tallon. An overview of economic applications of david schmeidlerâĂŹs models of decision making under uncertainty. *Uncertainty in economic theory*, pages 299–318, 2004.

C. P. Pires. A rule for updating ambiguous beliefs. *Theory and Decision*, 53(2):137–152, 2002.

L. Rayo and I. Segal. Optimal information disclosure. *Journal of political Economy*, 118(5): 949–987, 2010.

M. Serena. Harnessing beliefs to optimally disclose contestants' types. *Economic Theory*, Forthcoming, 2021.

R. Tang. A theory of updating ambiguous information. *arXiv preprint arXiv:2012.13650*, 2021.

K. Wärneryd. Information in conflicts. *Journal of Economic Theory*, 110(1):121–136, 2003.

C. Wasser. Incomplete information in rent-seeking contests. *Economic Theory*, 53(1):239–268, 2013.

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

J. Zhang and J. Zhou. Information disclosure in contests: A bayesian persuasion approach. *The Economic Journal*, 126:2197–2217, 2016.