

# Information Disclosure in Contests: a Bayesian Persuasion Approach

March 22, 2014

## Abstract

This paper utilizes the Bayesian persuasion approach developed in Kamenica and Gentzkow [10] to examine the optimal information disclosure in a two-player contest. One contestant's valuation is commonly known and the other's is his private information. The contest designer can pre-commit to a signal to influence the uninformed contestant's belief about the informed contestant. We show that, when the state is binary, it is without loss of generality to focus on the no disclosure and full disclosure, which is commonly assumed in this literature. However, when the state goes beyond binary, such a restriction may be of loss of generality. We propose a simple method to compute the optimal signal, which yields explicit solutions in some situations.

**JEL Classification.:** C72, D72, D82

**Keywords:** Bayesian Persuasion, Contests, Information Disclosure

# 1 Introduction

Contests are widely adopted to model R&D, rent seeking, political campaign, patent races, science competitions, job promotions and lobbying. It is now well documented in the literature that contestants often have private information about their own abilities, valuations, competence, etc.<sup>1</sup> In real life contests, the organizer is able to influence contestants' beliefs about each other by information disclosure. For example, in the U.S. lobbies, the government can decide the level of transparency, which requires the lobbying groups to provide information about their business. Such transparency requirement could potentially leak information about their private interests. In job promotions, companies can decide whether to announce the list of candidates and furthermore whether to reveal workers' past experience. Such information conveys signals correlated to workers' private information, and could lead to updates in beliefs once disclosed. In research tournaments, research proposals serve as good signals of firms' research abilities. How to reveal such information back to the firms can influence the competition. In the U.S. political campaigning, candidates are demanded by the Federal Election Campaign Act to reveal the sources of campaign contributions and campaign expenditure, which conveys information about the depth of financial support of a candidate.

In this paper, we will illustrate how information disclosure could be designed optimally. Early works on this issue usually assume zero or one choice by comparing no and full disclosure.<sup>2</sup> With no disclosure, beliefs remain the prior; and with full disclosure, contestants' exact types become common knowledge before contests take place. However, in reality, something in between could often arise. It is quite a debate whether restricting to the no and full disclosure is with loss of generality or not. However, due to lack of proper technical tools, it is quite difficult to go beyond the zero or one choice. Fortunately, the recent developed Bayesian persuasion approach, pioneered by Kamenica and Gentzkow [10], provides us with the possibility to tackle such a question. As to be shown, one general message from the paper is that restricting to the no and full disclosure is indeed with loss of generality, which suggests the need for a more general treatment along this literature.

In our model, there are two contestants: contestant A with commonly known valuation and contestant B with privately known valuation. The contest designer and contestant A share the same prior about contestant B's valuation. The most important assumption is that the contest designer can pre-commit to a signal before the contest takes place. Since the signal is correlated to contestant B's private valuation, contestant A will update his belief about contestant B after observes a signal realization. Finally, contestant A and B engage in the contest by simultaneously

---

<sup>1</sup>See Zhang and Wang [17], Moldovanue and Sela [14], Fey [3], Hurely and Shogren [8].

<sup>2</sup>Some literatures will be discussed later.

choosing effort levels in the competition. The contest designer aims to maximize the total effort from the two contestants by choosing the signal.

We show that when contestant B's valuation follows a binary distribution, it is without of generality to focus on no and full disclosure, and one of them is optimal among all feasible disclosure policies. The necessary and sufficient condition for each of them to be optimal is provided. This condition does not depend on the prior distribution of contestant B's valuation. However, when contestant B's valuation takes more than (including) three different values, we illustrate that the simple zero or one choice fails in maximizing the contest designer's objective in general. The novelty of the paper is to show that it is without loss of generality to focus on posterior beliefs on the edge of the simplex. The edge of the simplex has the same properties as the binary case, which is fully solved in our model. Such observation enables us to propose a simple method to compute the optimal signal. This method allows us to explicit characterize the optimal signal in some situations. First, we show that when contestant A is strong enough, full disclosure is optimal. Second, when contestant A is a bit weaker, pooling the highest two valuations together and fully separates the others is optimal.

How to reveal information to influence the outcome of a game has been extensively studied in the literature.<sup>3</sup> Kamenica and Gentzkow [10] are the first to investigate how to disclose information through Bayesian persuasion. In their paper, there are a single sender and a single receiver. At the beginning, there is a state of nature unknown to everyone. Sender pre-commits to an informative signal about the state of the world. After the Receiver observes a signal realization and updates his belief about the state of nature, he takes an action. What make the information disclosure a Bayesian persuasion is the assumption that the Sender cannot distort or conceal information once the signal is realized. This key assumption is quite likely to hold in contests and makes our paper a natural application. For instance, in political campaigns and lobbies, governments' commitment is legally mandated. Furthermore, contest organizers usually need to hold the same contests over and over again, and reputation concerns often enforce the organizer to commit.

The novelty of Kamenica and Gentzkow [10] is showing that finding the optimal signal is equivalent to solving the concavification of a value function defined on the set of all posteriors. This observation is particularly powerful if the state follows a binary distribution since the concavification has a graphical representation. However, going beyond binary distribution is usually hard. Our paper is based on general distribution, and unfortunately, none of their results have direct implication on our model. In addition, our paper differs from their original paper in two aspects. First, there are multiple receivers. Second, each receiver knows their own valuations when making

---

<sup>3</sup>Information can be revealed through many different ways. For example, in Vincent Crawford and Sobel [1], the sender can disclose information through cheap talk. In Grossman [7] and Milgrom [13], the sender cannot lie about the truth although he does not need to tell the whole truth. In Kartik [11] and Spence [15], there is a cost of lying.

decisions.

The theory in Kamenica and Gentzkow [10] is then extended in several directions. Kamenica and Gentzkow [6] allow multiple senders and investigate whether competitions among senders will lead to more information to be revealed. Yun [16] applies the theory of Bayesian persuasion to voting games. Her model is more general than Kamenica and Gentzkow [10] in the aspect that it allows multiple receivers. As shown in Kamenica and Gentzkow [10], allowing multiple receivers will not result in more complications if the persuasion is public in the sense that the signal realization is publicly observed. Unfortunately, when persuasion could be private, serious problem arises since “the key simplifying step in the analysis-reducing the problem of finding an optimal signal to one of maximizing over distributions of posterior beliefs-does not apply”. The contribution of Yun [16] is to be the first one to investigate private persuasion, which is restricted to independent private persuasion, and compares it with public persuasion. To accommodate this complication, the author focuses on binary state of nature. Our paper is the first paper to apply the Bayesian persuasion theory in Kamenica and Gentzkow [10] to contest theory, and we deal with general distributions.

The literature on information disclosure in contests motivates our paper. Fu et. al [4], Fu et. al [5], and Lim and Matros[12] consider how to reveal the information about the entry result when entries are stochastic. Denter et al. [2] analyze the incentive for a privately informed contestant to disclose his information to his rival, the incentive for the uninformed contestant to acquire information, and the incentive for the designer to mandate transparency. All of these papers focus on comparing the no disclosure and full disclosure. As a result, a natural question is whether restricting to the zero or one choice is with loss of generality or not. The approach of Bayesian persuasion covers both no and full disclosure as special cases, and allows much more instruments for the organizer. Our paper analysis demonstrates that when the state is binary, it is plausible to restricting to the no and full disclosure since either of them will be optimal among all feasible disclosure policies. However, when the state takes more than three values, such a simplification could be with loss of generality.

The rest of the paper is organized as follows. In Section 2, we describe the model. In Section 3, we characterize the equilibrium in the posterior contest game. In Section 4, we solve the optimal signal. In Section 5, we conclude. All the technical proofs are relegated to the appendix.

## 2 The model

Consider the following static contest under one-sided incomplete information. The basic framework is borrowed from Hurley and Shogren [8, 9] and Denter et al. [2]. There are two risk neutral contes-

tants, A and B, competing with each other for a prize by exerting irreversible efforts simultaneously. The success function of contestant  $i \in \{A, B\}$  under the effort portfolio  $(x_A, x_B)$  is given by

$$p_i(x_A, x_B) = \frac{x_i}{x_A + x_B} \quad (1)$$

If both exert zero effort, then each wins with half probability. The payoff of a contestant is simply his valuation of winning multiplied by the winning probability, and minus the cost of effort, which is assume to be linear.

Contestant A's valuation of winning is commonly known as  $v_A$ . Contestant B's valuation  $V_B$  of winning is his private information; the contest designer and contestant A share a common prior about it. More specifically,  $V_B$  is a random variable on  $\Omega$  with  $N \geq 2$  values,  $v_{B1} < \dots < v_{BN}$ .<sup>4</sup> Let  $\Delta^{N-1} = \{\mu \in \mathbf{R}^N | \mu^n \geq 0, \sum_{n=1}^N \mu^n = 1\}$  denote the standard  $(N-1)$ -simplex in  $\mathbf{R}^N$ , and  $\text{int}(\Delta^{N-1})$  denote the interior of  $\Delta^{N-1}$ . Each point  $\mu \in \Delta^{N-1}$  is also identified as a probability distribution on  $\{v_{B1}, \dots, v_{BN}\}$ . Denote the prior distribution of  $V_B$  as  $\mu_0 = \{\mu_0^1, \dots, \mu_0^N\}$  and assume  $\mu_0^n > 0, \forall n$ .<sup>5</sup>

The difference between our work and the previous literature is that the contest designer can pre-commit to a signal before the contest starts in order to maximize her objective, the total expected effort from the two contestants. A signal  $\pi$  consists of a realization space  $S$  with  $N$  elements and a family of likelihood distributions  $\pi = \{\pi(\cdot | v_{Bi})\}_{i=1}^N$  over  $S$ .<sup>6</sup> Potentially, the instruments for the contest organizer is quite rich. For example, both no and full disclosure policies are special cases of the signal. It also includes many policies such as partitions, and no lie about the truth. As noted in Kamenica and Gentzkow [10], the optimal signal also provides an upper bound when the contest designer's commitment power is absent, which means that any information disclosure through cheap talk, signaling, etc, cannot generate more total effort.

Note that the signal is a conditional distribution on contestant B's valuation. Thus, when a signal  $s \in S$  is realized, contestant A needs to update his belief about contestant B using Bayes's rule. Denote the posterior belief as  $\mu_s \in \Delta^{N-1}$ . Although we assume that the prior  $\mu_0$  belongs to  $\text{int}(\Delta^{N-1})$ , the posterior belief  $\mu_s$  could lie on the boundary of  $\Delta^{N-1}$ .

The timing of the game is as follows.

---

<sup>4</sup>When  $N = 1$ , there is no private information, and there is no role for information disclosure. Furthermore, all results hold for continuous distribution with  $N$  goes to infinity.

<sup>5</sup>It is without loss of generality to assume  $\mu_0 \in \text{int}(\Delta^{N-1})$ , since we can simply reduces the dimension of  $N$  when some prior probabilities are zeros.

<sup>6</sup>As shown in Kamenica and Gentzkow [10], it is without loss of generality to assume that the size of the signal is less than the minimum of the size of action space and the type space. In our model, the action space is continuous and the type space has  $N$  values.

1. The contest organizer chooses and pre-commits to a signal  $\pi$ .
2. Nature moves and draws a valuation for contestant B, say  $v_{Bn}$ .
3. The contest organizer carries out his commitment and a signal realization  $s \in S$  is generated according to  $\pi(s|v_{Bn})$ .
4. The signal realization  $s$  is observable by the public and leads to a posterior belief of contestant B, denoted as  $\mu_s$ .
5. The contest takes place and both contestants choose effort levels simultaneously.

Decisions are made only in stage 1 and 5. We call stage 1 the Bayesian persuasion stage and stage 5 the posterior contest game. The posterior game is a one-side incomplete information contest between two contestants who simultaneously choose their efforts. In the Bayesian persuasion stage, the contest designer's problem is to choose the optimal signal  $\pi$  to maximize the total effort. The equilibrium concept we employ is perfect Bayesian Nash equilibrium. We work from backward and first examine the posterior contest game, i.e., stage 5.

### 3 The posterior contest game

In the posterior contest game, contestant A's valuation is commonly known as  $v_A$  and contestant B's valuation is commonly believed as drawn from the distribution  $\mu_s$ . The equilibrium of such a game is summarized in the following proposition.

**Proposition 1 (Equilibrium in one-sided incomplete information contest)** *In a two-contestant, A and B, one-sided incomplete information contest, where A's valuation is commonly known as  $v_A$  and B's valuation is distributed according to  $\mu_s \in \Delta^{N-1}$  on  $(v_{B1}, \dots, v_{BN})$ , there exists a unique pure strategy equilibrium in which contestant A chooses effort*

$$x_A^* = \left( \frac{\mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_B}}]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_B}}]}} \right)^2,$$

*and contestant B chooses effort according to*

$$x_B^*(v_{Bn}) = \sqrt{v_{Bn}} \left( \frac{\mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_B}}]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_B}}]}} \right) - \left( \frac{\mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_B}}]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_B}}]}} \right)^2, \quad n = 1, 2, \dots, N.$$

The expected total effort in this equilibrium is

$$F(\mu_s) = \frac{\mathbf{E}_{\mu_s}[\sqrt{V_B}] \mathbf{E}_{\mu_s}[\frac{1}{\sqrt{V_B}}]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s}[\frac{1}{V_B}]}. \quad (2)$$

The notation  $\mathbf{E}_{\mu_s}\{\cdot\}$  is the expectation under belief  $\mu_s$ .

Note that we assume interior solutions here.<sup>7</sup> The formula works for any distribution  $\mu_s$ , even when  $\mu_s$  is a continuous probability distribution. Now we can examine the contest organizer's optimal signal, i.e., the optimal Bayesian persuasion in stage 1.

## 4 Bayesian Persuasion

In stage 1, the contest designer chooses the signal  $\pi$  to maximize the total effort in the contest. Given a signal realization  $s$ , it leads to a posterior belief  $\mu_s$  and total effort  $F(\mu_s)$  defined in equation (2) in Proposition 1. Due to the complexity in the choice of  $\pi$ , the contest designer's problem is not tractable in general.

Denote a distribution of posteriors as  $\tau \in \Delta(\Delta^{N-1})$ .  $\tau$  is called Bayes-plausible if the expected posterior probability equals the prior, i.e.,  $\sum_{Supp(\tau)} \mu d\tau(\mu) = \mu_0$ . Kamenica and Gentzkow [10] shows that finding the optimal signal  $\pi$  is equivalent to searching over the Bayes-plausible distribution of posteriors  $\tau$  to maximize the expected value of the posterior total effort:

$$\begin{aligned} & \max_{\{\alpha_k, \mu_k\}_{k=1}^N} \sum_{k=1}^N \alpha_k F(\mu_k) \\ & s.t. \quad \sum_{k=1}^N \alpha_k \mu_k = \mu_0, \\ & \quad \sum_{k=1}^N \alpha_k = 1, \alpha_k \geq 0, \text{ and } \mu_k \in \Delta^{N-1}, k = 1, 2, \dots, N. \end{aligned} \quad (3)$$

Please refer to their original paper for details. If we treat  $\mu_k$  as a lottery and  $F(\mu_k)$  as the value of the lottery, then we can treat the distribution of posterior beliefs  $\tau$  as the compound lottery  $(\mu_1, \dots, \mu_N; \alpha_1, \dots, \alpha_N)$ . As a result, the above problem is to maximize the expected value of  $F$

---

<sup>7</sup>A sufficient condition to guarantee this is to assume  $1/v_{B1} \leq 1/v_A + 1/v_{BN}$ . Note that contestant B's effort function is increasing in his valuation. Therefore, in the case where some valuations's optimal efforts hit zero, we can transform the model by assuming that those valuations are replaced by zero valuation.

among all possible compound lotteries  $\tau$  whose reduced lottery remains  $\mu_0$ . The above formula is a bit different from the original one in Kamenica and Gentzkow [10]. In general, the support of  $\tau$  should include a continuum of posterior beliefs. However, as shown in their online appendix, it is without loss of generality to assume that the size of signal as well as the number of posteriors to be less than the minimum of the size of action space and the type space. In our model, the action space is continuous and the type space has  $N$  values. Thus, we can assume that there are at most  $N$  posterior beliefs in  $\tau$ .

Mathematically, the indirect value function from the above maximization program (3) is exactly the value of the concavification of  $F$  evaluated at the prior, denote as  $\mathbf{cav}F(\mu_0)$ . The following result is established in Kamenica and Gentzkow [10].

**Proposition 2** *The optimal signal always exists and achieves an expected total effort equal to  $\mathbf{cav}F(\mu_0)$ .*

As a result, we need to construct the concavification of  $F$  on the simplex  $\Delta^{N-1}$ . In our model, given any posterior belief, the expected total effort in the contest is  $F(\cdot)$  defined in (2).

Let  $\mathbf{e}^i \in \Delta^{N-1}$  denote the vector with 1 on the  $i$ -th slot, and 0s everywhere else. We also call  $\mathbf{e}^i$  the vertex of the simplex. Denote the set of the vertexes as  $Vertex(\Delta^{N-1})$ . Let  $\mathbf{e}^{ij} = \{t\mathbf{e}^i + (1-t)\mathbf{e}^j, t \in [0, 1]\}$  denote the line segment connecting  $\mathbf{e}^i$  and  $\mathbf{e}^j$ . We also call  $\mathbf{e}^{ij}$  the edge of the complex connecting the vertexes  $\mathbf{e}^i$  and  $\mathbf{e}^j$ . Denote the set of edges as  $Edge(\Delta^{N-1})$ . Let  $\mathbf{e}^{ijk} = \{\alpha\mathbf{e}^i + \beta\mathbf{e}^j + (1-\alpha-\beta)\mathbf{e}^k, \alpha, \beta \geq 0 \text{ and } \alpha + \beta \leq 1\}$  denote the plane connecting the vertexes  $\mathbf{e}^i$ ,  $\mathbf{e}^j$  and  $\mathbf{e}^k$ . We also call  $\mathbf{e}^{ijk}$  the face of the complex connecting the vertexes  $\mathbf{e}^i$ ,  $\mathbf{e}^j$  and  $\mathbf{e}^k$ . Denote the set of faces as  $Face(\Delta^{N-1})$ . Note that  $\mathbf{e}^{ii}$  and  $\mathbf{e}^{iii}$  degenerates to the vertex  $\mathbf{e}^i$ . These terminologies will be very convenient later on.

Under no disclosure, the expected total effort is  $F(\mu_0)$ , and under full disclosure, the expected total effort is

$$\zeta(\mu_0) := \sum_{n=1}^N \mu_0^n F(\mathbf{e}^n) = \sum_{n=1}^N \mu_0^n \frac{1}{\frac{1}{v_A} + \frac{1}{v_{Bn}}}. \quad (4)$$

Finding the concavification is relatively straightforward if the valuation space is binary since the posterior belief can be represented by a single variable and the concavification has a graphical representation.<sup>8</sup> In the following, we will first examine the binary case, and then the general case. The binary case is an important building block for solving the general case.

---

<sup>8</sup>If the objective function only depends on a single measure of the posterior belief such as the mean, it also simplifies the problem. Unfortunately, this is not the case in our model.



#### 4.1 Binary case: N=2

When contestant B's valuation follows a binary distribution, let  $v_{B1} = v_L$  and  $v_{B2} = v_H$ . In the posterior contest game, let  $\mu_s = (q, 1 - q)$ . In this case, the posterior belief is characterized by a single variable  $q$ , the probability of low valuation. Note that, although the prior belongs to  $\text{int}(\Delta^{N-1})$ , the posterior belief could lie on the boundary, i.e.,  $q \in [0, 1]$ . We can rewrite the expected total effort in the posterior contest game in (2) as:

$$\phi(q) := F(q, 1 - q) = \frac{(\frac{q}{\sqrt{v_L}} + \frac{1-q}{\sqrt{v_H}})(q\sqrt{v_L} + (1-q)\sqrt{v_H})}{\frac{1}{v_A} + \frac{q}{v_L} + \frac{1-q}{v_H}}, \quad q \in [0, 1]. \quad (5)$$

Direct calculation shows that

$$\phi''(q) = \frac{2v_A v_H v_L (\sqrt{v_H} - \sqrt{v_L})^2 (v_A(v_H + \sqrt{v_H v_L} + v_L) + v_H v_L)}{(v_H v_L + v_A(qv_H + (1-q)v_L))^3} \times (v_A - \sqrt{v_L v_H}) \quad (6)$$

The first term is positive for any  $q \in [0, 1]$ , hence  $\text{sign}\{\phi''(q)\} = \text{sign}\{v_A - \sqrt{v_L v_H}\}$ . Therefore

$$\phi(q) \text{ is } \begin{cases} \text{convex,} & \text{if } v_A \geq \sqrt{v_H v_L}; \\ \text{concave,} & \text{if } v_A \leq \sqrt{v_H v_L}. \end{cases} \quad (7)$$

When  $\phi$  is concave, the concavification of  $\phi$  is, by definition, just  $\phi$  itself. When  $\phi$  is convex, by Jensen's inequality, the concavification of  $\phi$  is  $\zeta$  defined in (4). In summary

$$\mathbf{cav}\phi = \begin{cases} \zeta & \text{if } v_A \geq \sqrt{v_H v_L}; \\ \phi & \text{if } v_A \leq \sqrt{v_H v_L}. \end{cases} \quad (8)$$

Note that when  $v_A = \sqrt{v_H v_L}$ ,  $\phi(q)$  is actually linear, and  $\zeta$  coincides with  $\phi$ . Hence, we have the following characterization for the binary case.

**Proposition 3 (Optimal Disclosure Policy: binary case)** *When  $N = 2$ , either full disclosure or no disclosure is optimal. More specifically, full disclosure is optimal if  $v_A \geq \sqrt{v_H v_L}$ , and no disclosure is optimal if  $v_A \leq \sqrt{v_H v_L}$ .*

Note that when  $v_A = \sqrt{v_H v_L}$ , all signals yield the same expected total effort. The theorem shows that the full (no) disclosure is optimal if contestant A is strong (weak). Here is the intuition. Note that a more balance contest will induce more total effort as it can induce more competition. Let's first consider the case  $v_A < v_L$ . The other cases are similar. When contestant B is of low valuation,

it is better to reveal this information as  $v_L$  is the closest valuation to that of contestant A. When contestant B is of high valuation, it is better to conceal this information since in this case contestant A will think contestant B of average valuation and the contest will be more balanced. As a result, there is a tradeoff between high and low valuations if more information is revealed. When contestant is very the benefit from concealing the high valuation dominates the cost from separating the low valuation.

In the binary case, the above proposition fully characterizes the optimal signal. Denter et al. [2] yields the same result when restricting to the no and full disclosure. Our analysis demonstrates that their result is robust if more sophisticated disclosure policies are allowed in the contest designer's choice. However, when we go beyond the binary case, this message could fail in general.

## 4.2 More than two states: $N \geq 3$

From now on, we assume that there are more than two states, i.e.,  $N \geq 3$ . It can be shown that in this case the expected total effort in the posterior contest  $F(\mu_s)$  is neither concave nor convex.<sup>9</sup> Furthermore, since  $F(\mu_s)$  cannot be characterized by a single variable, the concavification cannot be solved by examining the graph. As a result, it becomes much more complicated than the binary case, and we need to find the concavification of  $F$  directly. The following two lemmas are important for our main results.

**Lemma 1**  $\forall \mu \notin \text{Face}(\Delta^{N-1})$ , there exists weights  $\lambda_k$  and vector  $\mu_k \in \text{Face}\{\Delta^{N-1}\}$ ,  $k = 1, \dots, K$ , such that:

$$F(\mu) < \sum_{k=1}^K \lambda_k F(\mu_k), \quad (9)$$

$$\text{with } \sum_{k=1}^K \lambda_k = 1, \lambda_k > 0, k = 1, \dots, K;$$

$$\sum_{k=1}^K \lambda_k \mu_k = \mu,$$

**Lemma 2**  $\forall \mu \notin \text{Edge}(\Delta^{N-1})$ , there exists weights  $\lambda_k$  and vector  $\mu_k \in \text{Edge}\{\Delta^{N-1}\}$ ,  $k = 1, \dots, K$ ,

---

<sup>9</sup>See the appendix for the proof.

such that:

$$\begin{aligned}
F(\mu) &\leq \sum_{k=1}^K \lambda_k F(\mu_k), \\
\text{with } \sum_{k=1}^K \lambda_k &= 1, \lambda_k > 0, k = 1, \dots, K; \\
\sum_{k=1}^K \lambda_k \mu_k &= \mu,
\end{aligned} \tag{10}$$

Lemma 1 shows that for any lottery  $\mu$  not in the face of the complex, we can find a compound lottery  $(\mu_1, \dots, \mu_K; \lambda_1, \dots, \lambda_K)$ , whose elements are lotteries on the face and whose reduced lottery remains  $\mu$ , such that it yields *strictly* higher expected value than the lottery  $\mu$ . Lemma 2 shows that for any lottery  $\mu$  not on the edge of the complex, we can find a compound lottery  $(\mu_1, \dots, \mu_K; \lambda_1, \dots, \lambda_K)$ , whose elements are lotteries on the edge and whose reduced lottery remains  $\mu$ , such that it yields *weakly* higher expected value than the lottery  $\mu$ . The equality in Lemma 2 holds only when  $v_A$  is equal to some very specific values, which are solely determined by the parameters  $v_{B1}, \dots, v_{Bn}$ <sup>10</sup>. For example, if  $v_A = \sqrt{v_{Bi}v_{Bj}}$ , then any probability mixture over  $v_{Bi}$  and  $v_{Bj}$  will yield the same expected value, similar to the binary case. Mathematically, we can show that the strict relationship in Lemma 2 arises generically. The above lemmas yield the following proposition.

**Proposition 4** *In the posterior distribution induced by an optimal signal, all the posteriors must lie on the face of the simplex; furthermore, generically, all the posteriors must lie on the edge of the simplex.*

The above proposition demonstrates that it is never optimal to pool more than four (including four) valuations together. Furthermore, it is never optimal to pool more than three (including three) valuations together unless  $v_A$  is equal to some specific values, which are solely determined by the parameters  $v_{B1}, \dots, v_{Bn}$ .

Although Proposition 4 characterizes some necessary properties of the optimal signal, the solution remain unknown. The maximization problem in (3) provides a general way to compute the optimal signal, but the programming could be too complicated to solve. In the follows, we will propose a simplified method for computing an optimal signal<sup>11</sup>. The second part of Lemma 2 implies

<sup>10</sup>See the proofs of these two lemmas for more details.

<sup>11</sup>The optimal signal may not be unique as we have observed in binary case in section 2.

the following result.

**Lemma 3** *An optimal signal can be achieved by using posteriors on the edge.*

Lemma 3 on its own does not simplify the whole problem too much. Although we can restrict to posteriors on the edges, their exact locations are unknown and one edge could have more than one posteriors. However, restricting on a edge of the complex, the maximization program is simply equivalent to our fully solved binary case , and we can utilize our results in the previous section 2.

We know the posterior total effort function  $F(\cdot)$ , restricted on an edge of the simplex, is either concave or convex . If an edge is concave, we can use at most one point on this edge; if an edge is convex, we can use at most two points on this edge and the points must be the vertexes; if an edge is linear, we treat it as convex. Those observations together with Lemma 3 greatly simplify the whole problem. We have the following proposition.

**Proposition 5** *The following simplified programming can be adopted to solve the optimal signal.*

*Step 1: Determine the shapes of the edges. The edge  $e^{ij}$  is convex (strictly concave) if and only if  $v_A \geq (<) \sqrt{v_{Bi}v_{Bj}}$ .*

*Step 2: Fully separate all valuations without any associated strictly concave edge, i.e., valuations less than  $v_A^2/v_N$ .*

*Step 3: Now the remaining valuations should have at least one associate strictly concave edge. For convex edges, it has no weight and no parameter is needed; for strictly concave edges, assign one parameter for the weight of the edge and one parameter to identify the position of the point.*

This proposition simplifies the computation of the optimal signal to a much easier problem. The number of edges could be potentially very small by excluding convex edges; each edge of the simplex can have at most one posterior; each posterior is identified by a single parameter. In some cases, Proposition 5 actually pins down the optimal signal.

**Corollary 1** *Suppose  $N \geq 3$ . If  $v_A \geq \sqrt{v_{B(N-1)}v_{BN}}$ , full disclosure is optimal.*

The intuition can be drawn from the binary case. When contestant A is very strong, it is never optimal to pool any two valuations together. This is because the loss from the higher valuation is less than the gain from the lower valuation. Another fully solvable case is the following.

**Corollary 2** *Suppose  $N \geq 3$ . If  $v_A \in (\sqrt{v_{B(N-2)}v_{BN}}, \sqrt{v_{B(N-1)}v_{BN}})$ , the following signal is optimal. Whenever the valuation is less or equal to  $v_{B(N-2)}$ , reveal it truthfully; when the valuation is either  $v_{B(N-1)}$  or  $v_{BN}$ , reveal that it lies in the set  $\{v_{B(N-1)}, v_{BN}\}$ .*

For the case  $v_A < \sqrt{v_{B(N-2)}v_{BN}}$ , a closed form characterization of the optimal signal is usually not available. However, the following example illustrates the power of Proposition 5.

**Example 1** Suppose  $v_A = 5/2$ ,  $v_{B1} = 1$ ,  $v_{B2} = 4$ ,  $v_{B3} = 9$ , and  $\pi = (\pi_1, \pi_2, \pi_3)$ . Note that the edge  $e^{12}$  is convex and the edges  $e^{13}$  and  $e^{23}$  are strictly concave. According to Proposition 5, we can restrict attention to the following posteriors:  $(s, 0, 1 - s)$ ,  $s \in [0, 1]$  or  $(0, \mu, 1 - \mu)$ ,  $\mu \in [0, 1]$ . Let  $\lambda$  and  $1 - \lambda$  be the weights. Then the program reduces to

$$\begin{aligned} \max \quad & \lambda * F(s, 0, 1 - s) + (1 - \lambda) * F(0, \mu, 1 - \mu) \\ \text{st.} \quad & \lambda * (s, 0, 1 - s) + (1 - \lambda) * (0, \mu, 1 - \mu) = (\pi_1, \pi_2, \pi_3) \\ & 0 \leq s, \mu, \lambda \leq 1 \end{aligned}$$

One can solve  $s, \mu$  in terms of  $\lambda$ :  $s = \frac{\pi_1}{\lambda}, \mu = \frac{\pi_2}{1-\lambda}$ . Plugging these conditions into the objective function yields

$$\max \lambda * F\left(\frac{\pi_1}{\lambda}, 0, 1 - \frac{\pi_1}{\lambda}\right) + (1 - \lambda) * F\left(0, \frac{\pi_2}{1-\lambda}, 1 - \frac{\pi_2}{1-\lambda}\right), \text{ s.t. } \lambda \in [\pi_1, 1 - \pi_2], \quad (11)$$

which is a single-variable maximization problem, hence easy to solve.

For example, when  $\pi = (1/3, 1/3, 1/3)$ , the optimal

$$\lambda^* = \frac{2(223244 - 461\sqrt{97027})}{285867} \approx 0.557227.$$

The maximal value is about 1.44705, which can be implemented by the following likelihood matrix:

$$L = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.671681 & 0.328319 \end{bmatrix}$$

where the row corresponds to states, and the column corresponds to signals. Only two signals  $\{s1, s2\}$  are used in the optimal disclosure rule: when the state is the lowest one, i.e.,  $v_{B1}$ , send the signal  $s1$ ; when the state is the medium one, i.e.,  $v_{B2}$ , send the signal  $s2$ ; when the state is the highest, i.e.,  $v_{B3}$ , send signal either  $s1$  or  $s2$  with probabilities 0.671681 and 0.328319 respectively. The posterior belief after observing  $s1$  is  $(\frac{1}{3\lambda^*}, 0, 1 - \frac{1}{3\lambda^*}) \approx (0.5982, 0, 0.4018)$ , and the posterior belief after observing  $s2$  is  $(0, \frac{1}{3(1-\lambda^*)}, 1 - \frac{1}{3(1-\lambda^*)}) \approx (0, 0.752831, 0.247169)$ .

Note that the example also applies when the state takes more than three values with all other valuations less than 1. This is because any such valuations should be fully separated according to

Proposition 5. Also note that if we want to solve this example by using the original program (3) directly, it will involve in much more computations.

## 5 Conclusion and discussion

In this paper, we investigate how information disclosure through Bayesian persuasion can be utilized to enhance the total effort in contests. In our model, one contestant's valuation is commonly known and the other has private valuation. We show that in the binary case, it is without loss of generality to focus on no and full disclosure as either will be optimal even if more sophisticated disclosure policies can be adopted. However, going beyond the binary case not only brings up technical challenges, but also could fail this message. In any optimal signal, the induced posteriors cannot pool more than three (including) three valuations together. We show that if the commonly known contestant is strong enough, the full disclosure is optimal; if he is a bit weaker, pooling the highest two valuations together and fully separating the others is optimal. We also illustrate an efficient method to compute the optimal signal in the general case.

In Kamenica and Gentzkow[10], their first question is when the sender could benefit from persuasion. The same question can be asked in our framework as well. Proposition 4 actually implies that with  $N \geq 4$ , the contest designer always benefit from persuasion, and with  $N = 3$ , the contest designer always benefit from persuasion generically. Denter et al. [2] compare full disclosure and on disclosure with binary distributions. In our framework with more than two valuations, it can be shown that the full disclosure dominates the no disclosure if and only if the commonly known contestant's valuation is below a cutoff.<sup>12</sup>

## A Proofs

### Proof of Proposition 1:

Fixing  $x_A$ ,  $x_B^*(v_{Bn})$  solves

$$\max_{x_B \geq 0} \frac{x_B}{x_B + x_A} v_{Bn} - x_B.$$

Assuming interior solutions, we get  $x_{Bn}^* = \sqrt{v_{Bn}}\sqrt{x_A} - x_A$ . Player  $A$  choose  $x_A^*$  to maximize

$$\max_{x_A \geq 0} \mathbf{E}_{\mu_s} \left[ \frac{x_A}{x_{V_B} + x_A} v_A - x_A \right].$$

---

<sup>12</sup>Proofs are available in the online appendix.

The FOC is

$$\mathbf{E}^{\mu_s} \left[ \frac{x_{Bi}}{(x_{Bi} + x_A^*)^2} v_A \right] = 1.$$

Plugging in  $x_{Bi} = \sqrt{v_{Bi}}\sqrt{x_A^*} - x_A^*$  and solving for  $x_A^*$  yields:

$$x_A^* = \left( \frac{\mathbf{E}_{\mu_s} \left[ \frac{1}{\sqrt{v_{Bi}}} \right]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s} \left[ \frac{1}{v_{Bi}} \right]} \right)^2,$$

Therefore,

$$x_{Bi}^* = \sqrt{v_{Bi}}\sqrt{x_A^*} - x_A^* = \sqrt{v_{Bi}} \left( \frac{\mathbf{E}_{\mu_s} \left[ \frac{1}{\sqrt{v_{Bi}}} \right]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s} \left[ \frac{1}{v_{Bi}} \right]} \right) - \left( \frac{\mathbf{E}_{\mu_s} \left[ \frac{1}{\sqrt{v_{Bi}}} \right]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s} \left[ \frac{1}{v_{Bi}} \right]} \right)^2.$$

While the expected total effort is

$$F(\mu_s) = \mathbf{E}_{\mu_s}[x_{Bi}^*] + x_A^* = \mathbf{E}_{\mu_s}[\sqrt{v_{Bi}}\sqrt{x_A^*} - x_A^*] + x_A^* = \mathbf{E}_{\mu_s}[\sqrt{v_{Bi}}]\sqrt{x_A^*} = \frac{\mathbf{E}_{\mu_s}[\sqrt{v_{Bi}}]\mathbf{E}_{\mu_s}[\frac{1}{\sqrt{v_{Bi}}}]}{\frac{1}{v_A} + \mathbf{E}_{\mu_s}[\frac{1}{v_{Bi}}]}.$$

□

#### Proof of Lemma 1 and Lemma 2:

Take a fixed prior  $\mu \in \text{int}(\Delta^{N-1})$ , and a fixed vector  $\mathbf{u}$  with  $\sum_{i=1}^n u_i = 0$ , defined the following function

$$\eta_{\mathbf{u}}(\epsilon) := F(\mu + \epsilon \mathbf{u}) \tag{12}$$

Clearly  $\eta_{\mathbf{u}}$  is well defined on the interval  $[\delta_1^{\mathbf{u}}, \delta_2^{\mathbf{u}}]$  where

$$\delta_1^{\mathbf{u}} = \min(\epsilon | \mu_s + \epsilon \mathbf{u} \in \Delta^{N-1})$$

and

$$\delta_2^{\mathbf{u}} = \max(\epsilon | \mu_s + \epsilon \mathbf{u} \in \Delta^{N-1}).$$

Since  $\mu_s$  is interior,  $\delta_1^{\mathbf{u}} < 0 < \delta_2^{\mathbf{u}}$ . We need the following key Lemma 4 to proceed.

The following lemma summaries the properties of the function  $\eta$ .

**Lemma 4** Fix  $N \geq 3$  and  $\{v_{Bi}, i = 1, \dots, n\}$

- (1) for any  $\mathbf{u}$  with  $\sum u_i = 0$  and any positive  $v_A$ ,  $\eta''$  doesn't change sign on the interval  $[\delta_1^{\mathbf{u}}, \delta_2^{\mathbf{u}}]$ .
- (2) for any positive  $v_A$ , there exists a vector  $\mathbf{u}$  with  $\sum u_i = 0$  such that  $\eta''_{\mathbf{u}} = 0$  on the interval  $[\delta_1^{\mathbf{u}}, \delta_2^{\mathbf{u}}]$ .

(3) for any positive  $v_A$  (with exception at at most one point when  $N = 3$ ), there exists a vector  $\mathbf{u}'$  with  $\sum u'_i = 0$  such that  $\eta''_{\mathbf{u}'} > 0$  on the interval  $[\delta_1^{\mathbf{u}'}, \delta_2^{\mathbf{u}'}]$ .

**Proof of Lemma 4:**

Part (1):

Define  $\mathbf{w}^1 = (1, 1, \dots, 1)$ ,  $\mathbf{w}^2 = (\sqrt{v_{B1}}, \sqrt{v_{B2}} \dots, \sqrt{v_{Bn}})$ ,  $\mathbf{w}^3 = (\frac{1}{\sqrt{v_{B1}}}, \frac{1}{\sqrt{v_{B2}}} \dots, \frac{1}{\sqrt{v_{Bn}}})$ ,  $\mathbf{w}^4 = (\frac{1}{v_{B1}}, \frac{1}{v_{B2}} \dots, \frac{1}{v_{Bn}})$ , then we have

$$\eta_{\mathbf{u}}(\epsilon) = F(\mu + \epsilon \mathbf{u}) = \frac{\mathbf{E}^{\mu + \epsilon \mathbf{u}}[\sqrt{v_{Bi}}] \mathbf{E}^{\mu + \epsilon \mathbf{u}}[\frac{1}{\sqrt{v_{Bi}}}]}{\frac{1}{v_A} + \mathbf{E}^{\mu + \epsilon \mathbf{u}}[\frac{1}{v_{Bi}}]} = \frac{\langle \mu + \epsilon \mathbf{u}, \mathbf{w}^2 \rangle \times \langle \mu + \epsilon \mathbf{u}, \mathbf{w}^3 \rangle}{\frac{1}{v_A} + \langle \mu + \epsilon \mathbf{u}, \mathbf{w}^4 \rangle}$$

Define  $f(\epsilon) := \langle \mu + \epsilon \mathbf{u}, \mathbf{w}^2 \rangle$ ,  $g(\epsilon) = \langle \mu + \epsilon \mathbf{u}, \mathbf{w}^3 \rangle$ , and  $h(\epsilon) := \frac{1}{v_A} + \langle \mu + \epsilon \mathbf{u}, \mathbf{w}^4 \rangle$ . Clearly  $f, g, h$  are linear in  $\epsilon$ , therefore,  $g'' = f'' = h'' = 0$ . Moreover,

$$\eta_{\mathbf{u}} = \frac{fg}{h}, \text{ with } f > 0, g > 0, h > 0.$$

Using product rule, we have

$$\begin{aligned} \eta'_{\mathbf{u}} &= \left(\frac{fg}{h}\right)' = \frac{(f'g + fh')h - fgh'}{h^2} \\ \eta''_{\mathbf{u}} &= \left(\frac{fg}{h}\right)'' = \left(\frac{(f'g + fh')h - fgh'}{h^2}\right)' = \frac{2f'g'h^2 - 2hh'(f'g + fh') + 2fg(h')^2}{h^3} \end{aligned} \quad (13)$$

Here we have used the fact that  $f'' = h'' = g'' = 0$ . Note that the denominator  $h^3$  is positive, while the numerator is actually a constant, independent of  $\epsilon$  as

$$(2f'g'h^2 - 2hh'(f'g + fh') + 2fg(h')^2)' = 2f'g'2hh' - 2h'h'(f'g + fh') - 2hh'2f'g' + 2(f'g + fh')(h')^2 = 0$$

again the linearity of  $f, g, h$  is used. Therefore  $\eta''$  doesn't change sign on the interval  $[\delta_1^{\mathbf{u}}, \delta_2^{\mathbf{u}}]$ , so part (1) is proved.

Part (2):



Pick  $\mathbf{u}$  such that

$$\begin{aligned}\langle \mathbf{u}, \mathbf{w}^1 \rangle &= 0, \\ \langle \mathbf{u}, \mathbf{w}^2 \rangle &= \langle \mu, \mathbf{w}^2 \rangle, \\ \langle \mathbf{u}, \mathbf{w}^4 \rangle &= \frac{1}{v_A} + \langle \mu, \mathbf{w}^4 \rangle.\end{aligned}$$

This is possible as the vectors  $\mathbf{w}^1$ ,  $\mathbf{w}^2$  and  $\mathbf{w}^4$  are linearly independent as  $N \geq 3$  and  $v_{Bi}$  are pairwise different. Then  $f(\epsilon) = (1 + \epsilon)\langle \mu, \mathbf{w}^2 \rangle$  and  $h(\epsilon) = (1 + \epsilon) \left( \frac{1}{v_A} + \langle \mu, \mathbf{w}^4 \rangle \right)$ , therefore,

$$\eta_{\mathbf{u}}(\epsilon) = \frac{f(\epsilon)g(\epsilon)}{h(\epsilon)} = \frac{\langle \mu, \mathbf{w}^2 \rangle}{\left( \frac{1}{v_A} + \langle \mu, \mathbf{w}^4 \rangle \right)} g(\epsilon).$$

is linear in  $\epsilon$ , therefore  $\eta''_{\mathbf{u}} = 0$  on the interval  $[\delta_1^{\mathbf{u}}, \delta_2^{\mathbf{u}}]$ .

Part (3)

When  $N \geq 4$ , the vectors  $\mathbf{w}^1$ ,  $\mathbf{w}^2$ ,  $\mathbf{w}^3$  and  $\mathbf{w}^4$  are linearly independent, therefore we can pick  $\mathbf{u}'$  such that

$$\begin{aligned}\langle \mathbf{u}', \mathbf{w}^1 \rangle &= 0, \\ \langle \mathbf{u}', \mathbf{w}^2 \rangle &= 1, \\ \langle \mathbf{u}', \mathbf{w}^3 \rangle &= 1, \\ \langle \mathbf{u}', \mathbf{w}^4 \rangle &= 0.\end{aligned}$$

In this case  $h$  is a constant function, or  $h' = 0$ , so by Equation 13,

$$\text{sign}(\eta''_{\mathbf{u}'} ) = \text{sign}(f'g') = \text{sign}(\langle \mathbf{u}', \mathbf{w}^2 \rangle \langle \mathbf{u}', \mathbf{w}^3 \rangle) = \text{sign}(1)$$

Therefore  $\eta''_{\mathbf{u}'} > 0$ .

When  $N = 3$ , we need another construction of  $\mathbf{u}'$ . Pick  $\mathbf{u}'$  such that

$$\begin{aligned}\langle \mathbf{u}', \mathbf{w}^1 \rangle &= 0, \\ \langle \mathbf{u}', \mathbf{w}^2 \rangle &= \langle \mu, \mathbf{w}^2 \rangle, \\ \langle \mathbf{u}', \mathbf{w}^3 \rangle &= \langle \mu, \mathbf{w}^3 \rangle,\end{aligned}$$

again such  $\mathbf{u}'$  exists by the independence of  $\mathbf{w}^1$ ,  $\mathbf{w}^2$  and  $\mathbf{w}^3$ . In this case  $f(\epsilon) = (1 + \epsilon)\langle \mu, \mathbf{w}^2 \rangle$ ,  $g(\epsilon) = (1 + \epsilon)\langle \mu, \mathbf{w}^3 \rangle$ , by Equation 13,

$$\eta''_{\mathbf{u}'} = \langle \mu, \mathbf{w}^2 \rangle \langle \mu, \mathbf{w}^3 \rangle \frac{2(h(\epsilon) - (1 + \epsilon)h'(\epsilon))^2}{h^3} = \langle \mu, \mathbf{w}^2 \rangle \langle \mu, \mathbf{w}^3 \rangle \frac{2\left(\frac{1}{v_A} + \langle \mu, \mathbf{w}^4 \rangle - \langle \mathbf{u}, \mathbf{w}^4 \rangle\right)^2}{h^3} \quad (14)$$

Therefore we have  $\eta''_{\mathbf{u}'} > 0$  as long as  $\frac{1}{v_A} + \langle \mu_s, \mathbf{w}^4 \rangle - \langle \mathbf{u}, \mathbf{w}^4 \rangle \neq 0$ , which rules at most one  $v_A$ . Direct calculations show that the critical  $v_A$  satisfies the following condition

$$\frac{1}{v_A} = \left( \frac{1}{\sqrt{v_{B1}v_{B2}}} + \frac{1}{\sqrt{v_{B1}v_{B3}}} + \frac{1}{\sqrt{v_{B2}v_{B3}}} \right), \quad (15)$$

If  $N = 3$ ,  $v_{B1} < v_{B2} < v_{B3}$ . Note that this critical  $v_A$  actually depends only on the values of player B, but not on the belief  $\mu$ . Therefore Lemma 4 is proved.  $\square$

Now we start to prove Lemma 1 and Lemma 2.

Prove of Lemma 1: For  $N \geq 4$ , first let us assume that  $\mu$  have full support. According to part (3) of Lemma 4, for any positive  $v_A$  there exists a vector  $\mathbf{u}'$  with  $\sum u'_i = 0$  such that  $\eta''_{\mathbf{u}'} > 0$  on the interval  $[\delta_1^{\mathbf{u}'}, \delta_2^{\mathbf{u}'}]$ . Let  $\lambda = \frac{\delta_2^{\mathbf{u}'}}{\delta_2^{\mathbf{u}'} - \delta_1^{\mathbf{u}'}} \in [0, 1]$ , then  $0 = \lambda\delta_1^{\mathbf{u}'} + (1 - \lambda)\delta_2^{\mathbf{u}'}$ , therefore, by Jensen's inequality

$$\eta_{\mathbf{u}}(0) = \eta_{\mathbf{u}}(\lambda\delta_1^{\mathbf{u}'} + (1 - \lambda)\delta_2^{\mathbf{u}'}) < \lambda\eta_{\mathbf{u}}(\delta_1^{\mathbf{u}'}) + (1 - \lambda)\eta_{\mathbf{u}}(\delta_2^{\mathbf{u}'})$$

or equivalently

$$F(\mu) < \lambda F(\mathbf{q}') + (1 - \lambda)F(\mathbf{q}'') \quad (16)$$

with  $\mathbf{q}' = \mu_s + \delta_1^{\mathbf{u}'}\mathbf{u}$  and  $\mathbf{q}'' = \mu_s + \delta_2^{\mathbf{u}'}\mathbf{u}$ . Clearly  $\mathbf{q}' \neq \mathbf{q}''$  moreover  $\mu = \lambda\mathbf{q}' + (1 - \lambda)\mathbf{q}''$ . By definition of  $\delta_1^{\mathbf{u}'}$  and  $\delta_2^{\mathbf{u}'}$ , vector  $\mathbf{q}'$  and  $\mathbf{q}''$  are not in the interior of  $\Delta^{N-1}$ , therefore they lie on the boundary of  $\Delta^{N-1}$ . Suppose, for example that  $\mathbf{q}'$  doesn't lie on the face of the simplex, then its supports contains at least four  $v_{Bi}$ 's, we can continue this decomposition process and applying the above construction to  $\mathbf{q}'$  Iteratively until the all the posteriors found lie on some face of the simplex.

Prove of Lemma 2: The proof is similar to the proof of Lemma 2. We may part (2) of the Lemma 4 at the exceptional  $v_A$  if necessary. However, we now only have weak inequality, not strict less inequality.  $\square$

**Proof of proposition 4 :** Directly follows from Lemma 1, 2, Equation 3 and Proposition 2.  $\square$

**Proof of Lemma 3:** Directly follows from Lemma 2, Equation 3 and Proposition 2.  $\square$

**Proof of proposition 5 :** By Lemma 3, we can restrict attention to posteriors that lie on the

edges. The concavity/convexity of the function  $F()$  on each edge  $e^{ij}$  is analyzed in section 2. The Proposition just follows.  $\square$

**Proof of corollary 1** : Directly follows from proposition 5.  $\square$

**Proof of corollary 2** : Directly follows from proposition 5.  $\square$

## References

- [1] CRAWFORD, V. P., AND SOBEL, J. Strategic information transmission. *Econometrica: Journal of the Econometric Society* (1982), 1431–1451.
- [2] DENTER, P., AND SISAK, D. Where ignorance is bliss, tis folly to be wise: Transparency in rent-seeking contests.
- [3] FEY, M. Rent-seeking contests with incomplete information. *Public Choice* 135, 3 (2008), 225–236.
- [4] FU, Q., JIAO, Q., AND LU, J. On disclosure policy in contests with stochastic entry. *Public Choice* 148, 3-4 (2011), 419–434.
- [5] FU, Q., LU, J., AND ZHANG, J. Enhancing competitions by disclosure policy in contests with stochastic entry. *Public Choice* 148, 3-4 (2011), 419–434.
- [6] GENTZKOW, M., AND KAMENICA, E. Competition in persuasion. Tech. rep., National Bureau of Economic Research, 2011.
- [7] GROSSMAN, S. J. The informational role of warranties and private disclosure about product quality. *Journal of Law & Economics* 24, 3 (December 1981), 461–83.
- [8] HURELY, T., AND SHOGREN, J. Asymmetric information contests. *European Journal of Political Economy* 14 (1998), 645–665.
- [9] HURELY, T., AND SHOGREN, J. Effort levels in a cournot nash contest with asymmetric information. *Journal of Public Economics* 69 (1998), 195–210.
- [10] KAMENICA, E., AND GENTZKOW, M. Bayesian persuasion. *The American Economic Review* 101, 6 (2011), pp. 2590–2615.
- [11] KARTIK, N. Strategic communication with lying costs. *The Review of Economic Studies* 76, 4 (2009), 1359–1395.

- [12] LIM, W., AND MATROS, A. Contests with a stochastic number of players. *Games and Economic Behavior* 67, 2 (2009), 584–597.
- [13] MILGRON, P. Good news and bad news: Representation theorems and applications. *Bell journal of Economics* 12, 2 (1981), 380–391.
- [14] MOLDOVANU, B., AND SELA, A. Contest architecture. *Journal of Economic Theory* 126, 1 (2006), 70–96.
- [15] SPENCE, A. M. Job market signaling. *The Quarterly Journal of Economics* 87, 3 (August 1973), 355–74.
- [16] WANG, Y. Bayesian persuasion with multiple receivers. *Job Market Paper, Department of Economics, University of Pittsburgh* (2012), 1–45.
- [17] ZHANG, J., AND WANG, R. The role of information revelation in elimination contests. *Economic Journal* 119 (2009), 613–641.