Information Design in Simultaneous All-Pay Auction Contests*

Zhonghong Kuang† Hangcheng Zhao‡ Jie Zheng§

First Version January 2, 2018; Updated December 31, 2019

Abstract

We study the information design problem of the contest organizer in a simultaneous 2-player 2-type all-pay auction contest environment, where players have limited information about own/others valuations of the prize. The contest organizer can send a public message to the contestants about their type distribution, in order to maximize expected total effort. We allow the players' ex-ante symmetric type distributions to be correlated, and the information disclosure policy to take the stochastic approach of Bayesian persuasion. The optimal design, the structure of which depends on the degree of the correlation of players' types, is completely characterized and shown to achieve higher effort than the traditional discrete type-dependent information disclosure policy. Given players' types are private information, if there is a strong positive correlation, an optimal design consists of two posteriors with one representing a perfect positive correlation and the other representing a positive correlation identified by a cutoff condition; if there is a weak positive correlation or negative correlation between types, the optimal design consists of two posteriors with one where both being high types is impossible and the other where a positive correlation is identified by the same cutoff condition. We also consider the case in which types are unknown to both players and the case in which the type information is asymmetric between the two players. Welfare comparisons are also conducted across different informational setups. Our work is the first study on full characterization of information design for games with two-sided asymmetric information and infinite action space.

JEL Classification.: C72, D72, D82.

Keywords: Bayesian Persuasion; Information Design; All-Pay Auction Contest; Information Disclosure

*We are especially grateful to Jian Chen for detailed comments which helped us improve the manuscript. For helpful comments and insights, we thank Zhuoqiong Chen, Fuhai Hong, Ju Hu, Kai Konrad, Dan Kovenock, Jianpei Li, Jaimie Lien, Jingfeng Lu, Marco Serena, Adam Wong, Zaifu Yang, Zhongxiang Zhang, seminar participants at the University of International Business and Economics, Lingnan University and conference participants at the Max Planck Institute for Tax Law and Public Finance Tax Day workshop, 2018 International Conference on Economic Theory and Applications, 3rd Conference on Energy, Environment, Industrial Economics and Policy Applications, 2019 Asian Meeting of Econometric Society, 30th International Conference on Game Theory and 2019 Annual Conference of China Information Economics Society. This research is funded by National Natural Science Foundation of China (Projects No. 61661136002 and No. 71873074), Tsinghua University Initiative Scientific Research Grant (Project No. 20151080397). All errors are our own.

†Department of Management Science and Engineering, School of Economics and Management, Tsinghua University, Beijing, China, 100084 (e-mail: kuangzhh.15@sem.tsinghua.edu.cn).

‡Department of Economics, The University of Chicago, Chicago, IL, USA, 60637 (e-mail: hangcheng.academic@gmail.com).

§Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China, 100084 (e-mail: jie.academic@gmail.com).

1 Introduction

Contests, where agents expend irreversible resources to win a reward, are prevalent and essential in social and economic activities. Examples of contests include sports competitions, patent races among firms, research tournaments, promotion tournaments in organizations, rent-seeking, lobbying and election campaigns among politicians. The common objective of the organizer is to maximize the expected total effort of contestants. In a sports competition, the organizer hopes to encourage athletes to do their best. In a patent race, the industry organizer wants to maximize total investment on some specific technology. The employer who organizes a promotion contest wishes to boost employees' efforts.

In a contest, participants decide the amount of effort invested in the contest based on not only his or her own private information, but also his or her belief about the opponents' information. Hence, a contest organizer is able to influence contestants' beliefs about each other via disclosing information about the information held by their opponents. For example, middle manager can reveal capability of candidates through education and work experience before job promotion. In a research tournament, the principal can assess and announce the quality of proposals publicly. In such competing environment with competitors of private types, information disclosure plays an important role and is often used as a popular policy to help achieve the goal of the designer. A sales manager wants to maximize his employees' total efforts in a sales contest, and needs to decide under what conditions he will disclose the information about employees' abilities.

In this paper, we study the information design problem of the designer in a simultaneous all-pay auction contest environment where players have limited information about own/others valuation of the prize. We allow the information disclosure policy to take the stochastic approach of Bayesian persuasion proposed by Kamenica and Gentzkow (2011), which is a generalization of the traditional discrete information disclosure policy. The Bayesian persuasion framework has requirements for commitment power, and this assumption is generally satisfied in contests because of reputation concerns of the contest designer.

In our model, there are two players and a single prize. Players' valuations about the prize (called "type") are private information and are drawn from a symmetric joint binary distribution. Players' types can be ex-post heterogeneous, which takes the value of either high or low. This prior distribution is common knowledge to the contest designer and both contestants. In addition to their own valuations, before the contest starts, each of them can observe a public signal which contains information about the opponent's valuation, designed and broadcast by the contest designer. After receiving the signal realization, both contestants update their belief simultaneously to reach a common posterior. Finally, both contestants engage in an all-pay auction contest by simultaneously choosing their efforts. Kuang (2019) has shown the ineffectiveness of private persuasion under the Bayes correlated equilibrium framework in a two player all-pay auction contest. Therefore, it is without loss of generality that we focus on public persuasion in this paper.

The Bayesian persuasion framework as well as the concept of information design has received great attention in recent years (Bergemann and Morris, 2016*a*,*b*, Kamenica and Gentzkow, 2011, Mathevet, Perego and Taneva, 2019, Rayo and Segal, 2010). Bergemann and Morris (2019) and Kamenica (2019) provide literature reviews on the general topic of Bayesian persuasion and information design. The theory of Bayesian persuasion is then extended in several directions. Our paper belongs to the stream of literature on public persuasion with multiple receivers, where multiple players receive a common and public signal from the sender (Alonso and Camara, 2016*a*,*b*, Arieli and Babichenko, 2019, Bardhi and Guo, 2018, Chan et al., 2019, Hoshino, 2017, Laclau and Renou, 2016, Michaeli, 2017, Shimoji, 2016, Song and Zhao, 2019, Wang, 2018). In this paper, after receiving the signal generated by the contest designer, two contestants update beliefs and then compete against each other.

For studies related to Bayesian persuasion, there are generally two noteworthy issues, the existence and optimality of Bayesian persuasion signal. Whether and when does contest designer benefit from Bayesian persuasion? What is the optimal Bayesian persuasion signal? In order to answer these two questions, generally three steps are required: (1) characterizing equilibrium strategies, (2) expressing the persuader's payoff as a function of posterior distribution, and (3) concavification of the persuader's payoff function. Our analysis is based on the equilibrium characterized in all-pay auction contests with correlated information (Liu and Chen, 2016, Lu and Parreiras, 2017, Siegel, 2014). For the remaining two steps, we first present some examples of public signals to illustrate that our information design approach could do better than all information disclosure policies introduced by previous studies. We then show that the contest designer benefits from such a public persuasion process compared with the full concealment policy for almost every prior distribution, which confirms the effectiveness of information design.

The main contribution of this paper is threefold. First, we utilize the public Bayesian persuasion approach with existence of two sided information asymmetry. To the best of our knowledge, we make the first attempt to apply the methodology of Bayesian persuasion into a setting with two sided information asymmetry. The novelty of work by Kamenica and Gentzkow (2011) is their illustration that finding the optimal signal is equivalent to solving the concave closure 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 concave closure has a graphical representation. However, going beyond the binary distribution is often technically non-trivial. To solve for optimal Bayesian persuasion with more than one parameter, our work along with Kuang (2019) provides the method of elimination of strictly[1] dominated posteriors to refine posterior distributions.

Second, this paper analyzes the properties of total effort function in two-player all-pay auction contests. The symmetric Bayes Nash equilibrium for two-player all-pay auction contests with correlated prior has been characterized by Liu and Chen (2016). Based on this equilibrium, we pin down the expected payoff for the contest designer with any possible prior distributions. We find out that when types of two players are positively correlated to some extent, the expected payoff for the organizer can achieve the *first best solution*. The first best solution refers to the situation where the prize is allocated efficiently and all contestants receive zero expected payoff. We call the set of posteriors achieving first best solution as ridge and refer to this property as the *ridge phenomenon of positive correlation*. In such a scenario, providing more accurate information may reduce the payoff for the contest designer.

Third, this paper completely characterizes the optimal information disclosure policy via Bayesian

[1]Elimination of weakly dominated posteriors in Kuang (2019).

persuasion in all-pay auction contests with two sided information asymmetry. Previous study has compared the expected payoff among multiple candidates of information disclosure polices. Typical strategies include degenerated policies (full disclosure or full concealment), and more complex type-dependent revelation methods (Serena (2016) for Tullock contests, Lu, Ma and Wang (2018) for all-pay auction contests). In this paper, taking advantage of the *ridge phenomenon of positive correlation*, we find that applying information design can benefit the organizer with almost any possible prior distribution, which is contrary to the optimality of no disclosure when considering type-dependent discrete policies only (Lu, Ma and Wang, 2018). In addition, we derive the optimal signal when prior distribution is sufficiently positively correlated, in which case the signal only generates two posteriors. When prior distribution is sufficiently negatively correlated or mildly correlated, we also prove that at most two posteriors is generated by optimal signal and one of them is located at ridge.

The literature on information disclosure in contests motivates our research. Some of these studies are conducted under the framework of the Tullock contests (Chen, Kuang and Zheng, 2017*b*, Fu, Jiao and Lu, 2011, Hurley and Shogren, 1998*a*,*b*, Serena, 2016, Warneryd, 2003, Wasser, 2013), while the others focus on all-pay auction framework. Morath and Munster (2008) compares the policies of full disclosure and full concealment in a single-prize contest environment, which is later extended to allow for multiple prizes by Fu, Jiao and Lu (2014). Chen (2017) compares the public signal and private signal in an independent private value setting. Cai et al. (2019) investigates information disclosure joint with a reserve price under a continuous-type all-pay auction framework. In addition to analyzing the information disclosure from the perspective of the contest organizer, there are also studies that analyze the information sharing behavior from the perspective of the participants. Wu and Zheng (2017) and Kovenock, Morath and Mnster (2015) studies the incentives to share private information ahead of contests under the Tullock contest framework and the all-pay auction contest framework, respectively. Recently, there are several studies working on Bayesian persuasion and information design in Tullock contests (Chen, Kuang and Zheng, 2017*a*, Zhang and Zhou, 2016) or all-pay auction contests (Chen, Kuang and Zheng, 2018, Kuang, 2019).

We also incorporate our analysis in the framework of surplus triangle (or outcome triangle), proposed by Bergemann, Brooks and Morris (2015) when analyzing price discrimination. This method is further applied to the analysis on first price auctions (Bergemann, Brooks and Morris, 2017), buyer optimal learning and monopoly pricing (Roesler and Szentes, 2017), and adverse selection (Kartik and Zhong, 2019). If we use the horizontal axis and vertical axis to represent the ex ante surplus for contestants and organizer respectively, the potential outcome is bounded by three constraints: the efficient frontier characterized by maximum social welfare and two individual rationality conditions for contestants and contest organizer. According to the equilibrium characterization in Liu and Chen (2016), when prior distribution is independent or mildly correlated, monotonic equilibrium arise and hence prize is always allocated efficiently and the surplus pair is located at efficient frontier. When prior distribution is sufficiently positively correlated, contestant get zero utility and the surplus pair is located at vertical axis. Therefore, for distribution that belongs to both regions we mentioned above, or ridge, the contest organizer reaches first best solution.

In the first extension, we study the situations where players are not informed about their realized winning value. In previous sections, we assume that players know their true valuations before posterior contest game, thus the contest designer cannot manipulate players about their own types. However, if players do not know their realized type, should the designer disclose their own type information to them? We consider three different scenarios on players' information regarding their own types, no information where neither player knows own winning value, private information where both players know their own winning value, and asymmetric information where exactly one player knows own winning value. We show that no information is always superior to asymmetric information. The preference over private information and other two schemes depends on the parameters.

In the second extension, we extend the ridge phenomenon of positive correlation into arbitrary symmetric joint distribution with finite number of types. We prove that given the marginal distribution, we can uniquely pin down the joint distribution such that contest organizer reaches first-best solution. When the prior distribution is sufficiently positively correlated such that it can be expressed as weighted average of posteriors that located on the ridge, contest designer can apply Bayesian persuasion alone to achieve first best solution.

The rest of the paper is organized as follows: Section 2 sets up the model. Section 3 introduces the analytic frameworks for public persuasion. In Section 4 we study the posterior contest game. In Section 5 we study the optimal information design problem for the benchmark model of public persuasion with privately informed players. Section 6 extends the benchmark model to allow for alternative initial informational settings. Section 7 extends the ridge phenomenon of positive correlation into arbitrary symmetric joint distribution. Section 8 concludes.

2 Model

2.1 Settings

Consider the following static all-pay auction contest with two-sided incomplete information. Two risk neutral players ($i = 1, 2$) participate in a single-prize all-pay auction contest by exerting irreversible efforts simultaneously. The winning value of participant i, denoted as v_i is drawn from a finite set Ω_i. Both players share the same finite state space with 2 elements, $\Omega_i = \Omega = \{v_L, v_H\}$ where $v_L < v_H$. Without loss of generality, we define $d = \frac{v_H}{v_L}$ and normalize $v_L = 1$. Contestant i chooses an non-negative action x_i from infinite set $\mathbb{R}_+ = [0, +\infty)$. We denote $\boldsymbol{x} = (x_1, x_2)$ as the effort profile. The prize is assigned to the contestant with higher effort. Ties are broken with equal probabilities. The success function of contestant $i \in \{1, 2\}$ under effort portfolio $\boldsymbol{x}$ is given by

$$p_i(\boldsymbol{x}) = \begin{cases} 1 & \text{if} \quad x_i > x_{-i} \\ \frac{1}{2} & \text{if} \quad x_i = x_{-i} \\ 0 & \text{if} \quad x_i < x_{-i} \end{cases} \tag{1}$$

Player's utility has a linear form, with his valuation v_i of winning multiplied by the winning probability p_i minus the cost of effort. Hence, the ex post utility is denoted by $\Pi_i(\boldsymbol{x}) = p_i(\boldsymbol{x})v_i - x_i$. The linearity is derived by risk neutrality directly. Contestant's surplus is expressed as

$$\Pi_C = \Pi_1 + \Pi_2 \tag{2}$$

Contest organizer's utility is defined by summation of effort exerted by two players,

$$\Pi_O = x_1 + x_2 \tag{3}$$

Both $\mathbb{E}(\Pi_C)$ and $\mathbb{E}(\Pi_O)$ is nonnegative by individual rationality condition.

	High	Low
High	p	$\frac{1-p-q}{2}$
Low	$\frac{1-p-q}{2}$	q

Table 1: General Form of Joint Distribution

It is commonly known that the joint probability distribution of types is $\Pr(v_i, v_i)$, which is symmetric, $\Pr(H, L) = \Pr(L, H)$. Let μ denote the joint distribution, such prior distribution μ_0 can characterized by Table 1. In the subsequent analysis, we use (p, q) to represent the joint distribution because these two parameters are sufficient characterize the full distribution. The distribution is independent distribution if and only if $\sqrt{p} + \sqrt{q} = 1$.

Using p, q, the upper bound of social surplus is given by

$$\mathbb{E}(\Pi_C) + \mathbb{E}(\Pi_O) = \mathbb{E}(\max(v_1, v_2)) = d(1-q) + q \tag{4}$$

which we refer to efficient frontier. If we use the horizontal axis and vertical axis to represent the ex ante surplus for contestants and organizer respectively, the set of potential outcome must be located in gray area of Figure 1.

2.2 Realization Schemes

Typically we assume that players know their true valuations before posterior contest game, thus the contest designer cannot manipulate players about their own types. However, if players do not know their realized type, should the designer disclose their own type information to them? Three potential realization schemes arise,

1. No Information. Denoted as $\mathbb{N}$. (Neither player knows his/her own winning value.)
2. Private Information. Denoted as $\mathbb{P}$. (Both players know their own winning values.)
3. Asymmetric Information. Denoted as $\mathbb{A}$. (Exactly one player know his/her own winning value.)

We can either give the schemes exogenously or let contest designer chooses its optimal schemes. In benchmark, $\mathbb{P}$ is given exogenously. In the first extension, we let contest designer choose the realization schemes.

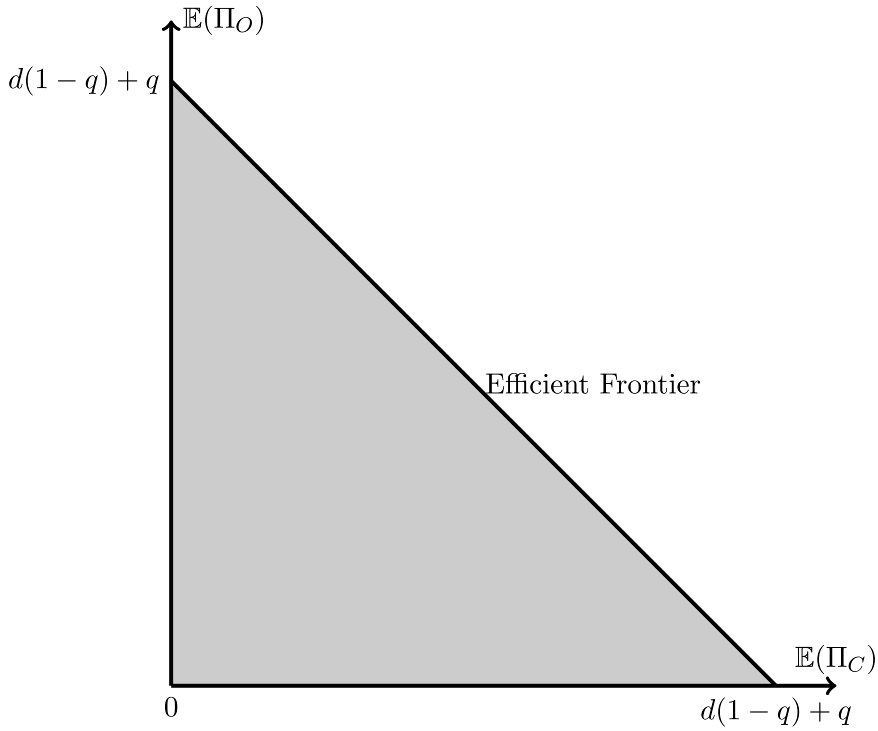

Figure 1: Surplus Triangle

2.3 Signal Decomposition

Mathevet, Perego and Taneva (2019) shows that the information design problem can be decomposed as optimizing over private information in the first step and then adding an optimal public signal in the second step. Since the problem is solved backward, the private signal is received separately after both contestants observing the public signal.

2.3.1 Public Signal

The public signal π consists of a realization space S and a family of likelihood distributions $\pi = \{\pi(\cdot|v_1, v_2)\}_{v_1,v_2 \in \Omega}$ over S. Hence, the public signal is interpreted as a mapping

$$\pi : \Omega^2 \to \Delta(S)$$

For each pair of (v_1, v_2), the signal generates a distribution over the signal space S. In this paper, we focus on anonymous public signal that requiring the symmetric condition $\pi(s|v_H, v_L) = \pi(s|v_L, v_H)$. This constraint guarantee symmetric posterior distributions. For each realization $s \in S$, the posterior distribution is derived as (p_s, q_s) by Bayes rule. Potential instruments for public persuasion are quite rich, including full disclosure and full concealment.

2.3.2 Private Signal

When a public signal $s \in S$ is realized, both contestants need to update their belief about the distribution of winning valuations simultaneously. Denote this posterior belief as (p_s, q_s). Private persuasion comes after this updating procedure.

The private signal σ consists of two realization spaces X_1, X_2, where X_1, X_2 denotes the private signal set for player i, and a family of likelihood distributions $\sigma = \{\sigma(\cdot, \cdot|s, v_1, v_2)\}_{s \in S, v_1, v_2 \in \Omega}$. over $X_1 \times X_2$. Since public signal is generated firstly, the private signal may depend on the realization of public signal. A private signal is interpreted as a mapping

$$\pi : \Omega^2 \times S \to \Delta(X_1 \times X_2)$$

For each signal realization $(x_1, x_2) \in X_1 \times X_2$, x_1 is privately sent to player 1 while x_2 is privately sent to player 2. By the revelation principle, it is without loss of generality focus on direct signaling policies (Arieli and Babichenko, 2019, Bergemann and Morris, 2016*a*). One way to understand the direct private signal is to view σ as the strategy of an omniscient mediator who first observes the realization of (v_1, v_2) chosen according to prior distribution (p_0, q_0) and the realization of s chosen according $\pi(\cdot|v_1, v_2)$, and then picks an action profile (x_1, x_2) and privately announces to each player i the draw of x_i. For players to have an incentive to follow the recommendation in this scenario, it would have to be the case that the recommended action x_i was always preferred to any other. This is reflected as the obedience condition in the Definition 1 of Bergemann and Morris (2016*a*). The private signal σ constructs a Bayes correlated equilibrium if the obedience condition holds for both players and all possible signal realizations.

2.4 Timeline

The timing of the game is as follows. Before started, a realization scheme is chosen from $\{\mathbb{N}, \mathbb{P}, \mathbb{A}\}$ exogenously (or by contest designer).

1. The contest designer chooses and pre-commits to a signal π and a private signal σ.
2. Nature moves and draws valuation profile $\boldsymbol{V} = (v_1, v_2)$ from prior distribution.
3. The contest designer carries out his commitment and a public signal realization $s \in \boldsymbol{S}$ is generated according to $\pi(s|\boldsymbol{V})$.
4. Signal realization s is observable by the public and both players update their common posterior belief as (p_s, q_s).
5. All players realize their valuation v_i according to realization scheme.
6. The contest designer carries out his commitment and a pair of recommended strategies (x_1, x_2) is generated according to $\sigma(x_1, x_2|v_1, v_2)$. x_1 is privately sent to player 1 and x_2 is privately sent to player 2.
7. The contest takes place, and all contestants choose efforts simultaneously.

Note that decisions are made only in stage 1 (contest designer) and stage 7 (contestants). We call stage 1 the *persuasion stage* and stage 6 the *contest game*, following the terminologies in Zhang and Zhou (2016). The posterior game is an all-pay auction with correlated distribution. In the persuasion stage, the contest designer is willing to choose the public signal π and private signal σ to maximize the expected total effort. If we allow contest designer to chooses its preferred realization scheme, then contest designer should make decision at stage 0 as well.

Stage 2 through 6 are implemented automatically by stochastic devices once contest designer made the commitment. Stage 3 is called *public signaling stage* and stage 6 is called *private signaling stage*. In stage 4, both contestants form a common posterior belief about (v_1, v_2). In stage 5, contestants are informed about (v_1, v_2) according to the scheme chosen by contest organizer. The solution concept employed in this paper is Bayesian Nash equilibrium for the posterior contest game and perfect Bayesian equilibrium for the whole game.

3 Information Design

Similar to the definition made by Mathevet, Perego and Taneva (2019), we first define the concept of public persuasion and private persuasion. Signal is public if and only if all players receive the same signal. This is equivalent to that sender announces the signal realization publicly. If only public signal is available, then the persuasion process is called public persuasion. Otherwise, we call it private persuasion.

In this essay, we are mainly focus on public persuasion because the private persuasion is ineffective for incomplete information all-pay auctions with two players (Kuang, 2019).

In stage 1, the contest designer chooses the signal π to maximize the expected total effort in the contest. Given a signal realization s, this leads to a common posterior joint distribution $\mu_s \in \Delta(\Omega^2)$ for all players. If the persuasion signal is private, no common joint distribution exists and we need to use belief-hierarchy to represent such distribution. Under the symmetric condition, the distribution of posterior also has the form Table 1. Let τ be a random variable that takes value in the $\Delta(\Delta(\Omega^2))$. Namely, it assigns a probability measure on the posteriors in the support of $\Delta(\Omega^2)$. We call τ *Bayesian-plausible* if the expected posterior probability equals the prior. Kamenica and Gentzkow (2011) show that finding optimal public signal π is equivalent to searching over Bayesian-plausible distribution of posteriors τ, which maximize the expected value of the posterior expected total effort $\mathbb{E}(\Pi_O)$. Mathematically, the indirect value function is exactly equal to the value of the concave closure of Π_O at the prior, denoted as $\mathbf{cav}\Pi_O(\mu_0)$, as is shown in the following proposition (Kamenica and Gentzkow, 2011).

Proposition 3.1. *The optimal signal always exists and achieves an expected total effort equal to* $\mathbf{cav}\Pi_O(\mu_0)$.

The optimal signals are quite simple in some special cases. When $\Pi_O(\mu)$ is concave, then no disclosure is optimal. When $\Pi_O(\mu)$ is convex, then full disclosure is optimal. In order to find the optimal signal, we need to construct the concave disclosure of Π_O on the $\Delta(\Omega^2)$. Slightly abuse the symbol, we use $\Pi_O(p,q)$ to represent expected total effort as a function of p,q, where p and q are defined in Table 1.

4 Persuasion with Privately Informed Players

4.1 Joint Distribution

Both contestants' valuations of winning are binary variables that chosen from prior joint distribution Table 1. Given the symmetric condition, we have for example $\Pr(v_1 = v_H | v_2 = v_L) = \Pr(v_2 = v_L | v_2 = v_H)$. We hence use the shorthand notation $\Pr(H|H), \Pr(L|H), \Pr(L|L), \Pr(H|L)$ hereafter. Since the prior distribution can uniquely defined by $\Pr(H,H), \Pr(L,L)$ and hence has degree of freedom two. By using parameters p,q, we can define the prior distribution as

$$\begin{aligned}
\Pr(H,H) &= p_0 \\
\Pr(L,L) &= q_0 \\
\Pr(H,L) &= \frac{1-p_0-q_0}{2} \\
\Pr(L,H) &= \frac{1-p_0-q_0}{2}
\end{aligned}$$

Then, we can illustrate the joint prior distribution μ_0 by Table 1 with parameters $p = p_0$ and $q = q_0$.

4.2 The Posterior Contest Game

In the posterior contest game, we assume that the valuation of both players followed the following symmetric joint distribution μ described by Table 1. The types' interim beliefs are

$$\begin{aligned}
\Pr(H|H) &= \frac{\Pr(H,H)}{\Pr(H)} = \frac{2p}{1+p-q} \\
\Pr(L|H) &= 1 - \Pr(H|H) = \frac{1-p-q}{1+p-q} \\
\Pr(L|L) &= \frac{\Pr(L,L)}{\Pr(L)} = \frac{2q}{1-p+q} \\
\Pr(H|L) &= 1 - \Pr(L|L) = \frac{1-p-q}{1-p+q}
\end{aligned}$$

The equilibrium of such a game is summarized in the following proposition according Liu and Chen (2016).

Proposition 4.1. *Consider the contest with the correlated information structure in Table 1. There is a symmetric Bayesian Nash equilibrium where both types play a mixed strategy.*

Case(1) (Sufficiently negatively correlated)*: When $d < \frac{\Pr(H|L)}{\Pr(H|H)}$, the symmetric equilibrium is* ***non-***

monotonic:

$$F_L(x) = \begin{cases} \frac{x}{\Pr(L|L)} & x \in [0, \frac{\Pr(L|L)(d-1)}{\Pr(L|H)d - \Pr(L|L)}] \\ \frac{(\Pr(H|L)-\Pr(H|H)d)x+\Pr(H|L)(d-1)}{(\Pr(H|L)-\Pr(H|H))d} & x \in [\frac{\Pr(L|L)(d-1)}{\Pr(L|H)d-\Pr(L|L)}, 1] \end{cases}$$

$$F_H(x) = \frac{(\Pr(L|H)d - \Pr(L|L))x - \Pr(L|L)(d-1)}{(\Pr(L|H) - \Pr(L|L))d}, x \in [\frac{\Pr(L|L)(d-1)}{\Pr(L|H)d - \Pr(L|L)}, 1]$$

Case(2) (Mildly correlated)*: When* $d > \frac{\Pr(L|L)}{\Pr(L|H)}$ *and* $d > \frac{\Pr(H|L)}{\Pr(H|H)}$*, the symmetric equilibrium is* ***monotonic:***

$$F_L(x) = \frac{x}{\Pr(L|L)}, x \in [0, \Pr(L|L)]$$

$$F_H(x) = \frac{x - \Pr(L|L)}{\Pr(H|H)d}, x \in [\Pr(L|L), \Pr(L|L) + \Pr(H|H)d]$$

Case(3) (Sufficiently positively correlated)*: When* $d < \frac{\Pr(L|L)}{\Pr(L|H)}$*, the symmetric equilibrium is* ***non-monotonic:***

$$F_L(x) = \frac{\Pr(H|H)d - \Pr(H|L)}{(\Pr(L|L) - \Pr(L|H))d}x, x \in [0, \frac{\Pr(L|L) - \Pr(L|H)}{\Pr(H|H)d - \Pr(H|L)}d]$$

$$F_H(x) = \begin{cases} \frac{\Pr(L|L)-\Pr(L|H)d}{(\Pr(L|L)-\Pr(L|H))d}x & x \in [0, \frac{\Pr(L|L)-\Pr(L|H)}{\Pr(H|H)d-\Pr(H|L)}d] \\ \frac{x-\Pr(L|H)d}{\Pr(H|H)d} & x \in [\frac{\Pr(L|L)-\Pr(L|H)}{\Pr(H|H)d-\Pr(H|L)}d, d] \end{cases}$$

Here, $F_L(x)$ and $F_H(x)$ are the distribution functions. And we could see that all the distribution functions are uniform distributions. When equilibrium is **monotonic**, the social surplus reaches maximum level of $d(1-q)+q$. We may also call such equilibrium **efficient**.

Proposition 4.2. *When distribution is mildly correlated,* $d > \frac{\Pr(L|L)}{\Pr(L|H)}$ *and* $d > \frac{\Pr(H|L)}{\Pr(H|H)}$*,* $\mathbb{E}(\Pi_C) + \mathbb{E}(\Pi_O) = d(1-q) + q$*.*

For the binary type symmetric distribution described by benchmark model, there are three possible ex post contest circumstances, high type compete with high type, low type compete with low type, and high type compete with low type. Inefficiency arises only in the last circumstance when it is possible for low type contestant bid higher than high type contestant. In other words, non-monotonicity implies inefficiency. Therefore, when distribution is mildly correlated, social surplus is maximized. When bidding range of both player contain zero, bidding zero is equivalent to bidding any values in the support of mixed strategy and hence both type of contestants get zero utility. We call such equilibria **exploitative** because contestants get zero utility ex ante.

Proposition 4.3. *When distribution is sufficiently positively correlated,* $d < \frac{\Pr(L|L)}{\Pr(L|H)}$*,* $\mathbb{E}(\Pi_C) = 0$*.*

Applying Proposition 4.1, we can then summarize the expected total effort for contest designer, $\Pi_O(p,q)$ by the following theorem.

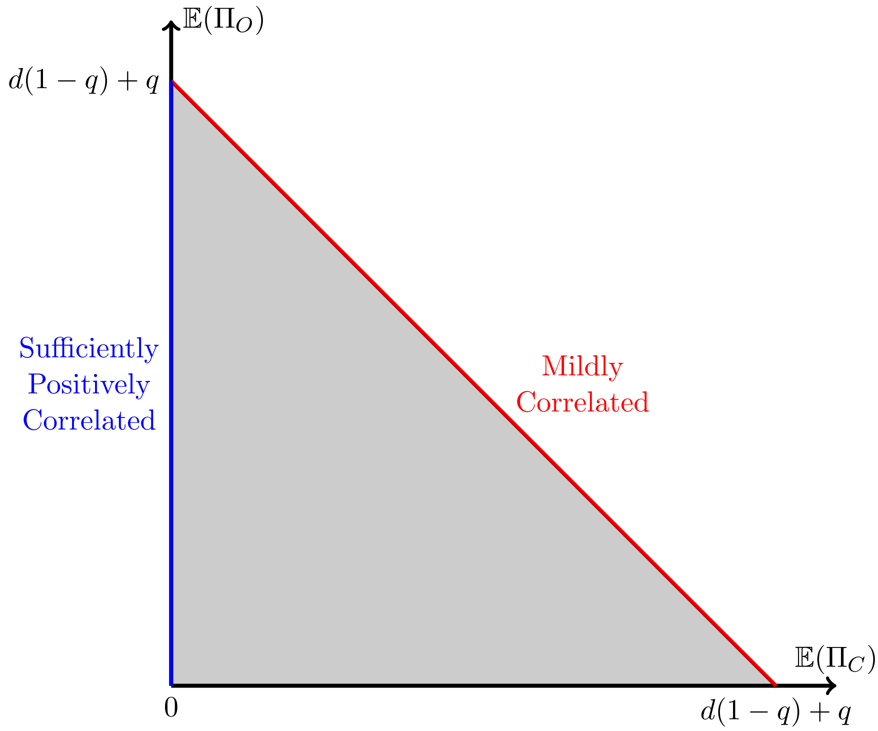

Figure 2: Surplus Triangle

Theorem 4.4. *The expected total effort for contest designer,*

$$\Pi_O(p,q) = (1+p-q)\mathbb{E}(x_H) + (1-p+q)\mathbb{E}(x_L) \tag{5}$$

When the two players' valuation follows symmetric joint distribution μ described by Table 1, the expected number of players that have high valuation is $1+p-q$. Similarly, the the expected number of players that have low valuation is $1-p+q$. And according to Proposition 4.1, the expected effort of a player conditional on his type is easily obtained from the probability density and the average value of the two extreme points. So we characterize the total effort in the three cases.

Case(1) (Strong negative correlation):

$$\Pi_O(p,q) = 1 + \frac{(d-1)(1+p-q)\left[4q-(1-p+q)^2\right]}{2(1-p+q)(1-p-q)d - 4q(1+p-q)} \tag{6}$$

Case(2) (Low correlation):

$$\Pi_O(p,q) = q + pd + \frac{2q(1+p-q)}{1-p+q} \tag{7}$$

Case(3) (Strong positive correlation):

$$\Pi_O(p,q) = 1 + \frac{(d-1)(1+p-q)\left[d(1+p-q)(1-p+q)-2(1-p-q)\right]}{4pd(1-p+q)-2(1-p-q)(1+p-q)} \tag{8}$$

4.2.1 2-Dimensional Properties of $\Pi_O(p,q)$

The equilibrium characterization can be divided into three parts according to the degree of correlation: case (1) corresponds to strong negative correlations between v_1 and v_2, case (2) corresponds to low correlation and case (3) corresponds to strong positive correlations.

In this subsection, by observing the changing of division with d via graphically illustration, we propose conjectures and then prove those conjectures. The following two plots characterized the three regions with two different values of d, which are defined as $d = \frac{v_H}{v_L}$ in benchmark model. We choose the cases where d equals 2 and 3.

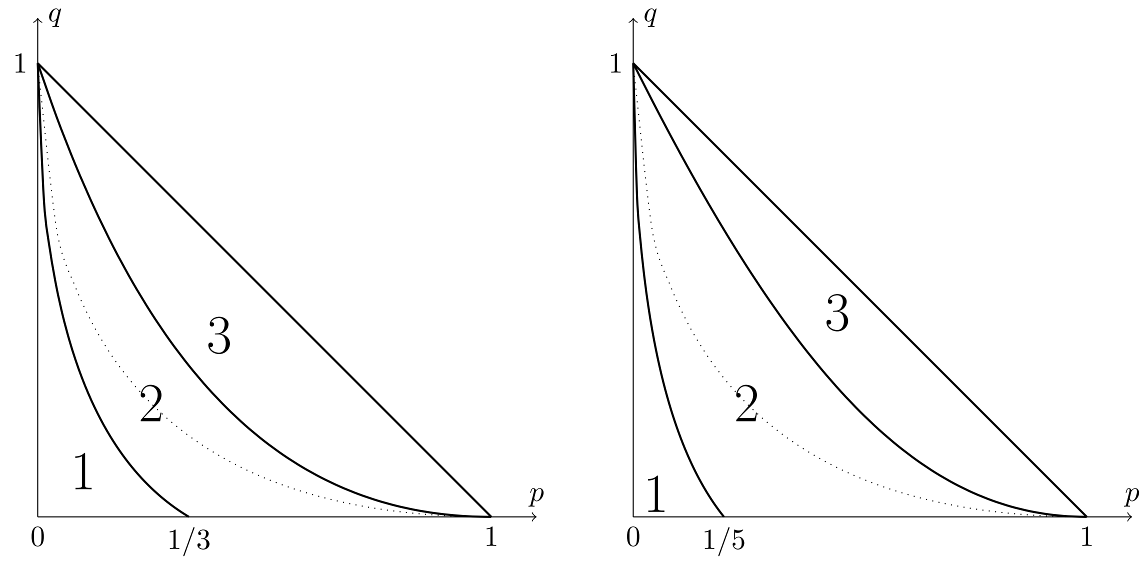

Figure 3: Left ($d = 2$), Right ($d = 3$)

For the sake of simplicity, we call the lowest part region 1 (case (1), strong negative correlation), the middle part region 2 (case (2), low correlation), the highest part region 3 (case (3), strong positive correlation). By Proposition 4.1, the boundary curve between region 1 and region 2 is defined as

$$d = \frac{\Pr(H|L)}{\Pr(H|H)} = \frac{(1-p-q)(1+p-q)}{2p(1-p+q)} \tag{9}$$

or written as quadratic equation of p, q with parameter d

$$(2d-1)p^2 + q^2 - 2dpq - 2dp - 2q + 1 = 0 \tag{10}$$

We call this boundary curve **valley**. For the implicit function described above, we define $p_v(q)$ and $q_v(p)$ respectively that denote the explicit relationship between p, q. By Proposition 4.1, the boundary curve between region 2 and region 3 is defined as

$$d = \frac{\Pr(L|L)}{\Pr(L|H)} = \frac{2q(1+p-q)}{(1-p-q)(1-p+q)} \tag{11}$$

or written as quadratic equation of p, q with parameter d

$$dp^2 + (2-d)q^2 - 2pq - 2dp - 2q + d = 0 \tag{12}$$

We call this boundary curve **ridge**. For the implicit function described above, we define $p_r(q)$ and $q_r(p)$ respectively that denote the explicit relationship between p, q.

From the graphical illustration described above, we can easily conclude that as d increases, region 2 expands while both region 1 and 3 shrink. Mathematically,

Observation 4.5. *For boundary between region 1 and region 2 described by Equation 10, we have the following condition*

$$\frac{\partial p_v(q)}{\partial d} < 0 \tag{13}$$

For boundary between region 2 and region 3 described by Equation 12, we have the following condition

$$\frac{\partial p_r(q)}{\partial d} > 0 \tag{14}$$

The above observation has the following implication. According to the equilibrium characterization in Proposition 4.1, monotone equilibrium exists only in region 2. In monotone equilibrium, the prize is never allocated to the bidder with lower value when valuations are different ex post. That is, only region 2 is efficient on the behalf of social planner. As d increases, region 2 grows, and monotone equilibrium may exist even in highly unbalanced distribution.

What's more, as it is shown in previous observation, when d increases, region 3 shrinks. And region 3 always keeps its convexity.

Observation 4.6. *Region 3 is convex.*

In fact, the convexity of region 3 plays great role in optimal signal characterization in Section 5.

Last but not least, we figure out that valley has unique intersection point with p-axis,

Observation 4.7. *The valley and p-axis has unique intersection point $\frac{1}{2d-1}$.*

By letting $q = 0$, the formula of valley turns to be

$$(2d-1)p^2 - 2dp + 1 = 0 \Longrightarrow [(2d-1)p - 1](p-1) = 0$$

The solution is hence $\frac{1}{2d-1}$.

4.2.2 3-Dimensional Properties of $\Pi_O(p, q)$

Conditional on those three regions, we construct expected total effort function $\Pi_O(p, q)$ in 3-d coordinate system by drawing all possible $(p, q, \Pi_O(p, q))$. In this subsection, we focus on special properties of $\Pi_O(p, q)$. Applying the surplus triangle methodology, we provide the upper bound of this function, which is defined by the upper bound of social surplus $\mathbb{E}[\max(v_1, v_2)]$,

Proposition 4.8. $\Pi_O(p, q) \leq \mathbb{E}[\max(v_1, v_2)] = d(1-q) + q$

We begin with some boundary cases. Since the domain of (p, q) is a 2-D simplex, we first analyze three edges of such simplex, including

1. $p = 0, q \in [0, 1]$ or known as q-axis,
2. $q = 0, p \in [0, 1]$ or known as p-axis,
3. $p + q = 1, p \in [0, 1]$ or known as hypotenuse.

q-axis

Proposition 4.9. *When posterior locates in q-axis, the resulting expected total effort can be expressed as a function of q alone,*

$$\Pi_O(0, q) = \frac{(1 + 4q - q^2)d + (1 - 6q + q^2)}{(2 + 2q)d - 4q} \tag{15}$$

The above equation is derived by using case (1) of Theorem 4.4 and letting $p = 0$.

Lemma 4.10. $\Pi_O(0, q)$ *is concave with respect to q.*

This lemma holds because the second order derivative,

$$\frac{\partial^2}{\partial q^2}\Pi_O(0, q) = \frac{-4(d-1)^3}{(d + dq - 2q)^3} < 0 \tag{16}$$

p-axis

Proposition 4.11. *When posterior locates in p-axis, the resulting expected total effort can be expressed as a function of p alone,*

$$\Pi_O(p, 0) = \begin{cases} \frac{d+1}{2d} - \frac{(d-1)p}{2d} & p \in [0, \frac{1}{2d-1}] \\ pd & p \in (\frac{1}{2d-1}, 1] \end{cases} \tag{17}$$

According to Observation 4.7, point $(p, 0)$ falls in region 1 when $p \leq \frac{1}{2d-1}$, and falls in region 2 otherwise. When $p \leq \frac{1}{2d-1}$, putting $q = 0$ into the expression of Theorem 4.4 case (1), we have $\Pi_0(p, 0) = \frac{d+1}{2d} - \frac{(d-1)p}{2d}$. When $p > \frac{1}{2d-1}$. putting $q = 0$ into the expression of Theorem 4.4 case (2), we have $\Pi_0(p, 0) = pd$.

Hypotenuse

Proposition 4.12. *When posterior locates in the hypotenuse, expected total effort function reaches the upper bound of $\Pi_O(p, q)$,*

$$\Pi_O(1 - q, q) = d(1 - q) + q \tag{18}$$

In this situation, the winning value of prize is common for both contestant. Hence, equilibrium is both efficient and exploitative. Therefore, expected total effort function reaches the maximum social surplus.

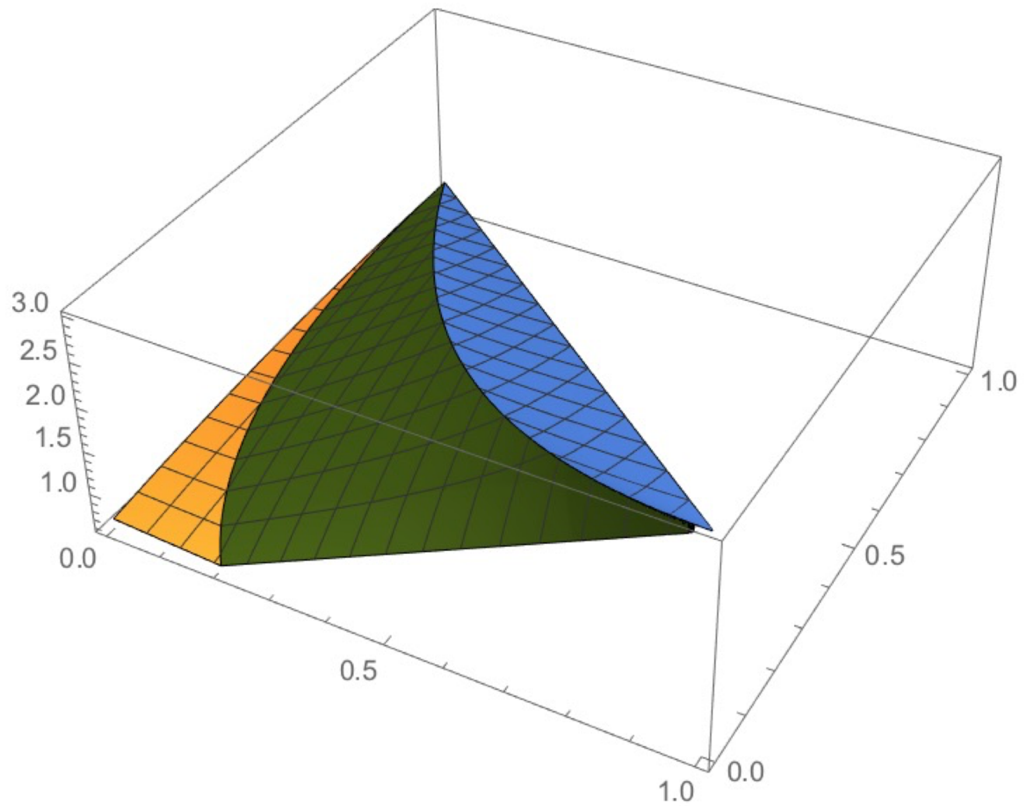

Figure 4: 3D schematic representation of $\Pi_O(p, q)$

Now we move to valley (boundary line of region 1 and region 2) and ridge (boundary line of region 2 and region 3). Figure 4 provides us a graphical illustration of $\Pi_O(p, q)$ where yellow represents the sufficiently negatively correlated case, green represents the mildly correlated case and blue represents the sufficiently positively correlated case.

Ridge

We find out that boundary of region 2 and region 3 with Equation 12 is indeed a ridge, which we refer to *ridge phenomenon of positive correlation.* Mathematically,

Proposition 4.13. *When posterior locates on the ridge, expected total effort function reaches the upper bound of* $\Pi_O(p, q)$*,*

$$\Pi_O(p_r(q), q) = d(1 - q) + q \tag{19}$$

When p, q satisfies Equation 12, according to Theorem 4.4, replacing $2q(1 + p - q)$ by $d(1 - p - q)(1 - p + q)$, we have the expected total effort

$$\Pi_O(p_r(q), q) = q + pd + \frac{2q(1 + p - q)}{1 - p + q} = q + pd + (1 - p - q)d = d(1 - q) + q$$

Counter intuitively, $\Pi_O(p_r(q), q) = \mathbb{E}(\max(v_1, v_2))$. That is, the upper bound of the $\Pi_O(p, q)$ function can be obtained on interior inside the 2-D simplex, see Figure 5.

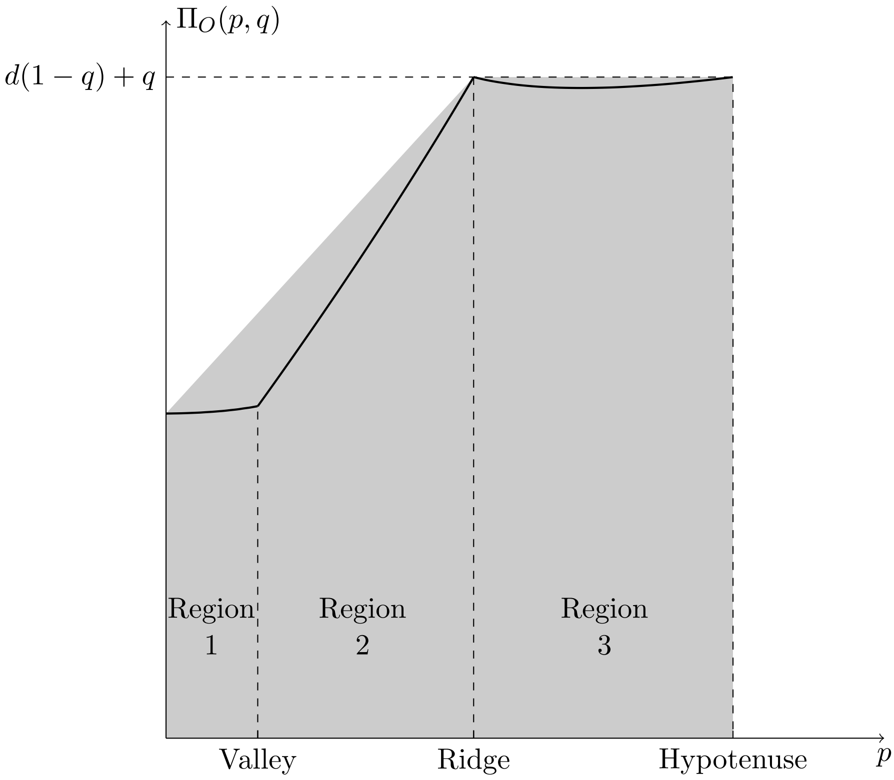

Figure 5: Cross Section (Fixed q Level)

Valley

Proposition 4.14. *When posterior locates in the valley, expected total effort function can be expressed as a function of p alone,*

$$\Pi_O(p, q_v(p)) = 1 - (d-1)p \tag{20}$$

We find out that boundary of region 1 and region 2 with Equation 10 is indeed a valley. If we take the cross section by fixing the value of q and connect the point with $p = 0$ and $p = p_r(q)$, the point located in the valley is below the line segment connecting q-axis and ridge, see Figure 5. Mathematically,

Lemma 4.15. *When posterior locates in the valley, expected total effort satisfies the following condition,*

$$\Pi_O(p_v(q), q) \leq \frac{p_r(q) - p_v(q)}{p_r(q)} \Pi_O(0, q) + \frac{p_v(q)}{p_r(q)} \Pi_O(p_r(q), q)$$

We also found that the lower bound of $\Pi_O(p, q)$ is obtained in the valley.

Proposition 4.16. $\Pi_O(p, q) \geq \frac{d}{2d-1}$. *This lower bound is achieved when $p = \frac{1}{2d-1}$ and $q = 0$.*

5 Main Results

5.1 An Illustrative Example

In a simultaneous all-pay auction contest, the prior type distributions for both players are independently and identically distributed,

$$v_i = \begin{cases} 1 & \text{w.p.} \quad 0.5 \\ 2 & \text{w.p.} \quad 0.5 \end{cases}$$

Use the terminology of this article, we have the ratio of winning value $d = 2$ and prior joint distribution $p_0 = q_0 = 0.25$.

In Lu, Ma and Wang (2018), which following Serena (2016) and followed by Chen, Kuang and Zheng (2017*b*), they characterizes four type-dependent disclosure policies. They only focus on anonymous disclosure policies. Ignoring the identities of the players, since each player has two types (v_H and v_L), the possible type profiles are $(v_H, v_H), (v_H, v_L)$ and (v_L, v_L), anonymous disclosure policy thus specifies for each profile whether to disclose it to both players (henceforth, D, for disclosure), or conceal it from both players (henceforth, C, for concealment). For instance, use (D, D, D) to denote this policy where the first D means disclosing when the profile is (v_H, v_H) the second D means disclosing when the profile is (v_H, v_L); and the third D means disclosing when the profile is (v_L, v_L). Recall that expected total efforts in those four policies are ranking as

$$(C, C, C) \succ (D, C, C) \succ (D, D, D) \succ (C, C, D)$$

with revenue showed in the following table.

Strategy	Expected total effort
(C,C,C)	5/4
(D,C,C)	7/6
(D,D,D)	9/8
(C,C,D)	3/4

However, contest organizer can design the following signal structure for better outcome. Assume $S = \{s_E, s_U\}$ where s_E stands for "Equal" and s_U stands for "Unequal". There are three combination of values, the conditional distribution of s depending on combination is shown in the following table.

	s_E	s_U
(v_H, v_H)	1	0
(v_H, v_L)	0.5	0.5
(v_L, v_L)	1	0

Contest organizer will reveal whether both contestants have the same winning value. If they are evenly matched, contest organizer will reveal the corresponding signal s_E. However, if their winning values are different, contest organizer will truthfully reveal this information by s_U with only half probability and misreport the signal s_E with remaining half probability. Therefore, the probabilities of receiving

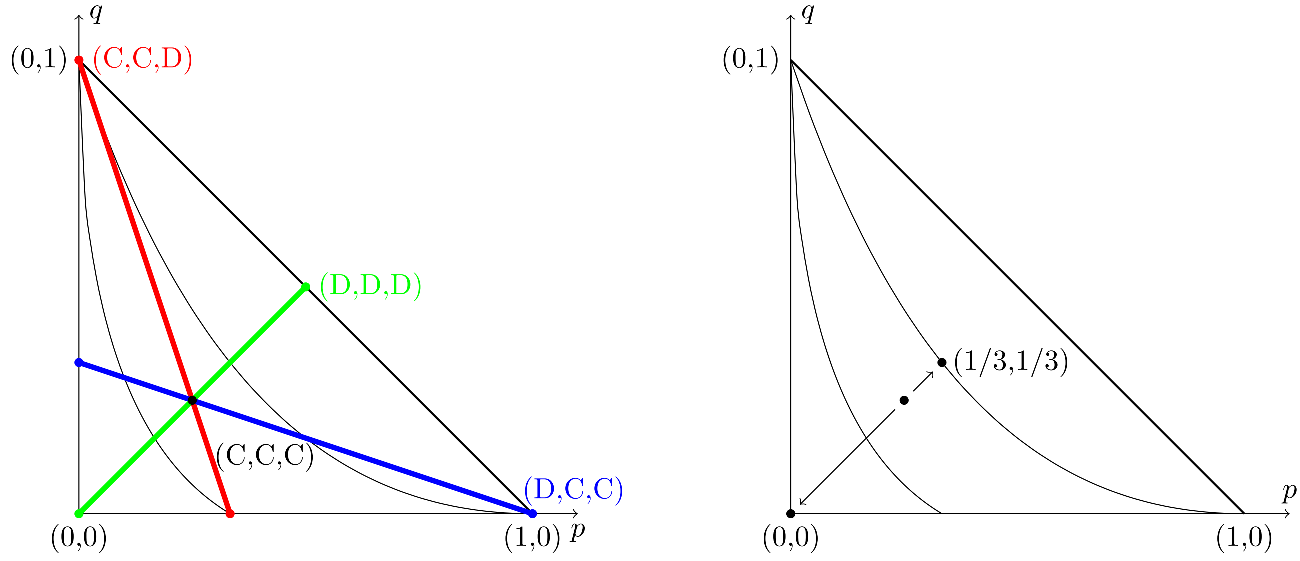

Figure 6: Strategy in Lu, Ma and Wang (2018) (left) and This Paper (right)

s_E and s_U are

$$\begin{aligned}\Pr(s_E) &= 3/4 \\ \Pr(s_U) &= 1/4\end{aligned}$$

When receiving s_E, the posterior distribution μ_E is

	High	Low
High	1/3	1/6
Low	1/6	1/3

When receiving s_U, the posterior distribution μ_U is

	High	Low
High	0	1/2
Low	1/2	0

Therefore, the expected total effort will be

$$\mathbb{E}_s(\Pi_O) = \Pr(s_E)\Pi_O(\mu_E) + \Pr(s_E)\Pi_O(\mu_E) = \frac{3}{4} \times \frac{5}{3} + \frac{1}{4} \times \frac{3}{4} = \frac{23}{16} > \frac{5}{4} = \frac{20}{16} \quad \text{(better than } (C,C,C))$$

In previous example, information design can generate higher expected effort than type dependent disclosure policy. Although the prior distribution is as simple as binary, it may be payoff-improving for the designer to apply information design. More specifically, information design is advantageous for almost every prior distribution.

The signal structure we mentioned above is not only beneficial for contest organizer compared with degenerated policy, but also optimal under the framework of Bayesian persuasion. In order to prove the optimality, we preview two results from later sections.

1. When prior $\mu_0 = (p_0, q_0)$ locates in region 1 or 2, then the optimal signal is associated with exactly two posteriors such that one locates at q-axis and the other locates on the ridge.

2. The optimal signal is unique and can be figure out efficiently by linear searching algorithm.

We now briefly discuss the process how we derive these two posteriors in these part. And we would present a general method and corresponding proofs in the next section. By the first result above, we could assume that there exists one posterior that located at q-axis, which is denoted as $(0, \bar{q})$. And there exists one posterior that located at ridge, which is denoted as $(\bar{p}, q_r(\bar{p}))$. The problem becomes searching for optimal $\bar{q}$ such that generates highest expected total effort.

We extend the straight line pass through points $(0, \bar{q})$ and $(0.25, 0.25)$ as

$$q = (1 - 4\bar{q})p + \bar{q} \tag{21}$$

denoted as line $\mathcal{L}$.

When $q > 0.36$, line $\mathcal{L}$ has no intersect points with ridge. When $q \leq 0.36$, we can solve for the intersection point $(\bar{p}, q_r(\bar{p}))$ as

$$(\bar{p}, q_r(\bar{p})) \equiv \left(\frac{2(1-\bar{q})}{3 - 3\bar{q} + \sqrt{9 - 34\bar{q} + 25\bar{q}^2}}, \frac{2(1-\bar{q})(1-4\bar{q})}{3 - 3\bar{q} + \sqrt{9 - 34\bar{q} + 25\bar{q}^2}} + \bar{q} \right)$$

See left panel of Figure 7 for graphical illustration.

If we linearly expand the prior distribution $(0.25, 0.25)$ into two posterior distributions $(0, \bar{q})$ and $(\bar{p}, q_r(\bar{p}))$, the resulting total effort at $(0, \bar{q})$ and $(\bar{p}, q_r(\bar{p}))$ are

$$\Pi_O(0, \bar{q}) = \frac{3 + 2\bar{q} - \bar{q}^2}{4} \tag{22}$$

$$\Pi_O(\bar{p}, q_r(\bar{p})) = \frac{(4 + \bar{q} - 5\bar{q}^2) + (2 - \bar{q})\sqrt{9 - 34\bar{q} + 25\bar{q}^2}}{3 - 3\bar{q} + \sqrt{9 - 34\bar{q} + 25\bar{q}^2}} \tag{23}$$

Hence, we can express the expected total effort after information design as a function of $\bar{q}$,

$$\begin{aligned}
g(\bar{q}) &= \frac{(\bar{p} - 0.25)\Pi_O(0, \bar{q}) + 0.25\Pi_O(\bar{p}, q_r(\bar{p}))}{\bar{p}} \\
&= \frac{\Pi_O(\bar{p}, q_r(\bar{p})) - \Pi_O(0, \bar{q})}{4\bar{p}} + \Pi_O(0, \bar{q}) \\
&= \frac{(4 + \bar{q} - 5\bar{q}^2) + (2 - \bar{q})\sqrt{9 - 34\bar{q} + 25\bar{q}^2}}{8(1 - \bar{q})} - \frac{(3 - 3\bar{q} + \sqrt{9 - 34\bar{q} + 25\bar{q}^2})(3 + 2\bar{q} - \bar{q}^2)}{32(1 - \bar{q})} + \frac{3 + 2\bar{q} - \bar{q}^2}{4} \\
&= \frac{(7 + 7\bar{q} - 11\bar{q}^2 - 3\bar{q}^3) + (5 - 6\bar{q} + \bar{q}^2)\sqrt{9 - 34\bar{q} + 25\bar{q}^2}}{32(1 - \bar{q})} + \frac{3 + 2\bar{q} - \bar{q}^2}{4} \\
&= \frac{(31 - \bar{q} - 35\bar{q}^2 + 5\bar{q}^3) + (5 - 6\bar{q} + \bar{q}^2)\sqrt{9 - 34\bar{q} + 25\bar{q}^2}}{32(1 - \bar{q})} \\
&= \frac{(31 + 30\bar{q} - 5\bar{q}^2) + (5 - \bar{q})\sqrt{9 - 34\bar{q} + 25\bar{q}^2}}{32}, \bar{q} \leq 0.36
\end{aligned}$$

Now the optimal information design problem for the designer becomes finding the optimal $\bar{q}^*$ such

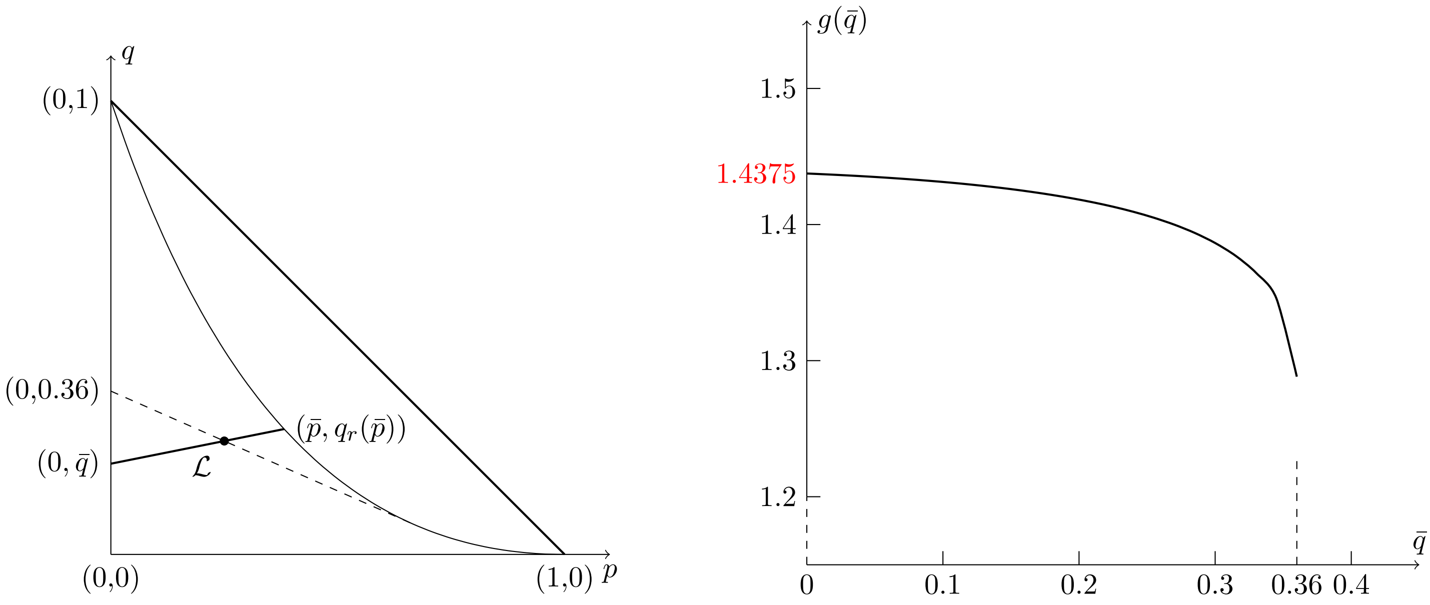

Figure 7: Example

that the resulting expected total effort after information design is maximized

$$\max_{\bar{q} \leq 0.36} g(\bar{q})$$

$g(\bar{q})$ is a quasi-concave function and hence solvable by linear searching. We can also prove the optimality of $\bar{q}^* = 0$ using the approach of calculus. We have the first order condition

$$\begin{aligned} g'(\bar{q}) &= \frac{(30 - 10\bar{q}) - (94 - 176\bar{q} + 50\bar{q}^2)(9 - 34\bar{q} + 25\bar{q}^2)^{-0.5}}{32} \\ g'(0) &= -\frac{1}{24} < 0 \end{aligned}$$

and the second order condition

$$\begin{aligned} g''(\bar{q}) &= \frac{-10 - (-176 + 100\bar{q})(9 - 34\bar{q} + 25\bar{q}^2)^{-0.5} + (94 - 176\bar{q} + 50\bar{q}^2)(-17 + 25\bar{q})(9 - 34\bar{q} + 25\bar{q}^2)^{-1.5}}{32} \\ &= -0.3125 - \frac{(-176 + 100\bar{q})(9 - 34\bar{q} + 25\bar{q}^2) - (94 - 176\bar{q} + 50\bar{q}^2)(-17 + 25\bar{q})}{32(1 - \bar{q})^{1.5}(9 - 25\bar{q})^{1.5}} \\ &= -0.3125 - \frac{14 + 1542\bar{q} - 2550\bar{q}^2 + 1250\bar{q}^3}{32(1 - \bar{q})^{1.5}(9 - 25\bar{q})^{1.5}} \\ &= -0.3125 - \frac{14 + \bar{q}(272 - 50\bar{q}) + 1250\bar{q}(1 - \bar{q})^2}{32(1 - \bar{q})^{1.5}(9 - 25\bar{q})^{1.5}} < 0 \end{aligned}$$

second order condition implies that $g(\bar{q})$ is a concave function. By first order derivative at $\bar{q}^* = 0$, we know that $g(\bar{q})$ is strictly decreasing and hence reaches its maximum at $\bar{q}^* = 0$. Please refer to right panel of Figure 7 for diagram of $g(\bar{q})$.

5.2 Summary

The main task of this section is twofold.

1. Providing the necessary and sufficient condition on when Bayesian persuasion is superior to full disclosure or full concealment. (subsection 5.3)

2. Characterizing the optimal Bayesian persuasion of the main model. (subsection 5.4 and subsection 5.5)

Our analysis on both problems are based on ridge phenomenon of positive correlation (Proposition 4.13), which divides the entire domain into two parts:

1. Sufficiently positively correlated (Region 3)

2. Sufficiently negatively correlated (Region 1) and mildly correlated (Region 2)

For the first question, we propose a new method to refine posterior distributions, called elimination of strictly dominated posteriors, which is further inherited by Kuang (2019). For some posterior distribution (p, q), if we can find a strictly better signal than full concealment, then this posterior distribution will never be contained in any optimal posterior profile. The signal we use here actually horizontally spreads the prior distributions into two posteriors locating on the boundary along the direction of p-axis. After refining the candidates of posterior distributions, there are only three curves remaining, q-axis, ridge and hypotenuse. See Figure 8, where blue line denotes inefficient remaining posteriors and red lines denote efficient remaining posteriors. For the second question, we show that for almost all prior distributions (except for those three colored curves), it is without loss of generality to consider only two posterior distributions that one of them is located on the ridge.

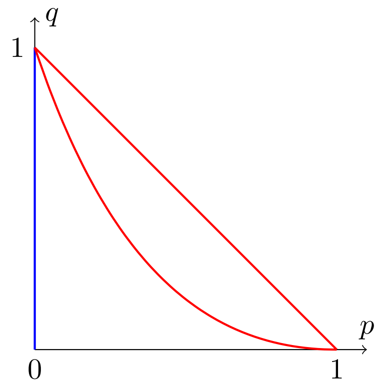

Figure 8: Remaining Posteriors

When prior (p_0, q_0) is located in region 3, we can draw a horizontal line through the prior that parallel to p-axis and take the intersection points with ridge and hypotenuse as two posterior distributions, $(0, q_0)$ and $(p_r(q_0), q_0)$ (Theorem 5.1). In both of posteriors, $\Pi_O = d(1 - q_0) + q_0$ (Proposition 4.12 and Proposition 4.13). Since this signal structure achieves the upper bound stated in Proposition 4.8, the signal is guaranteed to be optimal (Proposition 5.5). Last but not least, we show that as long as all posteriors are lying on ridge and hypotenuse, the corresponding signal is optimal. This finding has two implications, as shown in Corollary 5.7, (1) there are infinitely many optimal signals and (2) optimal may include any number of posteriors.

When prior (p_0, q_0) located in region 1 or 2, we can also draw a horizontal line that is parallel to p-axis and take the intersection points with ridge and q-axis as two posterior distributions (Theorem 5.3). This posterior profile that generated by the parallel line to p-axis through (p_0, q_0) could always provide a improved total effort directly from the fact that we use horizontal spread to eliminate interior distributions. So we can focus on the posteriors on the ridge and q-axis. Then we prove the number of posteriors on q-axis (Lemma 4.10) and the ridge (Lemma 5.10) is at most one. Furthermore, we prove that the uniqueness of the optimal signal in Theorem 5.12. Two associated posteriors are located on the ridge and q-axis respectively. Finally, we show that it is searching for the optimal signal is efficient because the objective function for searching algorithm is quasi concave (Corollary 5.14).

5.3 Elimination of Strictly Dominated Posteriors

5.3.1 Region 3

When prior distribution is located in region 3, we can always increase the expected total effort by applying the following signal.

Theorem 5.1. *For any prior $\mu_0 = (p_0, q_0)$ in region 3 (but not in boundary), the signal associated with two posteriors, $(p_r(q_0), q_0)$ and $(1 - q_0, q_0)$ is (strictly) better than full concealment. The resulting expected total effort after information design is $d(1 - q_0) + q_0$.*

This theorem is concluded immediately after Proposition 4.12 and Proposition 4.13. By Proposition 4.8, we have $\Pi_O(p, q) \leq d(1 - q) + q$ for arbitrary posterior distribution (p, q). When distribution is located inside region 3, then $\Pi_O(p, q) < d(1 - q) + q$.

Proposition 5.2. *All posteriors located inside region 3 will not be included in posterior profile.*

5.3.2 Region 1 and Region 2

Theorem 5.3. *For any prior $\mu_0 = (p_0, q_0)$ in region 1 or 2, there exists a signal associated with two posteriors, $(0, q_0)$ and $(p_r(q_0), q_0)$ is (strictly) better than full concealment.*

Mathematically, the above signal is strictly better than no disclosure if and only if

$$\Pi_O(\lambda p_r(q), q) < \lambda \Pi_O(p_r(q), q) + (1 - \lambda)\Pi_O(0, q), \forall \lambda \in (0, 1)$$

When fixing d and q, we can prove that $\Pi_O(p, q)$ is convex in region 1 with respect to p and convex in region 2 with respect to p as well. Hence, $\Pi_O(p, q_0)$ is piecewise convex. By Lemma 4.15, we have the the value of $\Pi_O(p, q)$ at turning point $p_v(q_0)$ is lies below the line connecting 0 and $p_r(v_0)$. This concludes the proof. Please refer to appendix for details.

This theorem states that for any prior in region 1 or 2, designer could always increase his surplus compared with no disclosure by horizontally spread prior distribution to two posteriors lying on q-axis and ridge respectively.

Proposition 5.4. *All posteriors located inside region 1, region 2 and valley will not be included in posterior profile.*

5.4 Information Design for Region 3

5.4.1 Optimal Public Persuasion

By Proposition 4.8, we have $\Pi_O(p,q) \leq d(1-q)+q$ for arbitrary posterior distribution (p,q). Hence, it is impossible to design a strictly better signal.

Proposition 5.5. *For any prior $\mu_0 = (p_0, q_0)$ in region 3 (but not in boundary), the signal interpreted in Theorem 5.1 is optimal among all possible signals.*

Recall that $\Pi_O(p,q) = d(1-q)+q$ for all distribution on the ridge or hypotenuse. If all posteriors associated with some signal are on the ridge or hypotenuse, contest organizer obtains $d(1-q_0)+q_0$ by implementing this signal.

Proposition 5.6. *When prior locates inside region 3, as long as all induced beliefs are at boundary of region 3 (including ridge and hypotenuse), the corresponding signal is optimal.*

Directly from the above proposition,

Corollary 5.7.

1. *When prior locates inside region 3, there exist infinite optimal signals.*
2. *When prior locates inside region 3, optimal signal may include any number of posteriors.*

Now, we present another example to visualize the results we have derived in region 3. We set $d = 2$ and $p = q = 0.4$.

	High	Low
High	2/5	1/10
Low	1/10	2/5

Expected total effort without information disclosure is 1.57143. Contest organizer can design the following signal structure for better outcome. Similar to subsection 5.1, we still let $S = \{s_E, s_U\}$. The conditional distribution of s depending on combination is shown in the following table.

	s_E	s_U
(v_H, v_H)	0.5	0.5
(v_H, v_L)	0	1
(v_L, v_L)	0.5	0.5

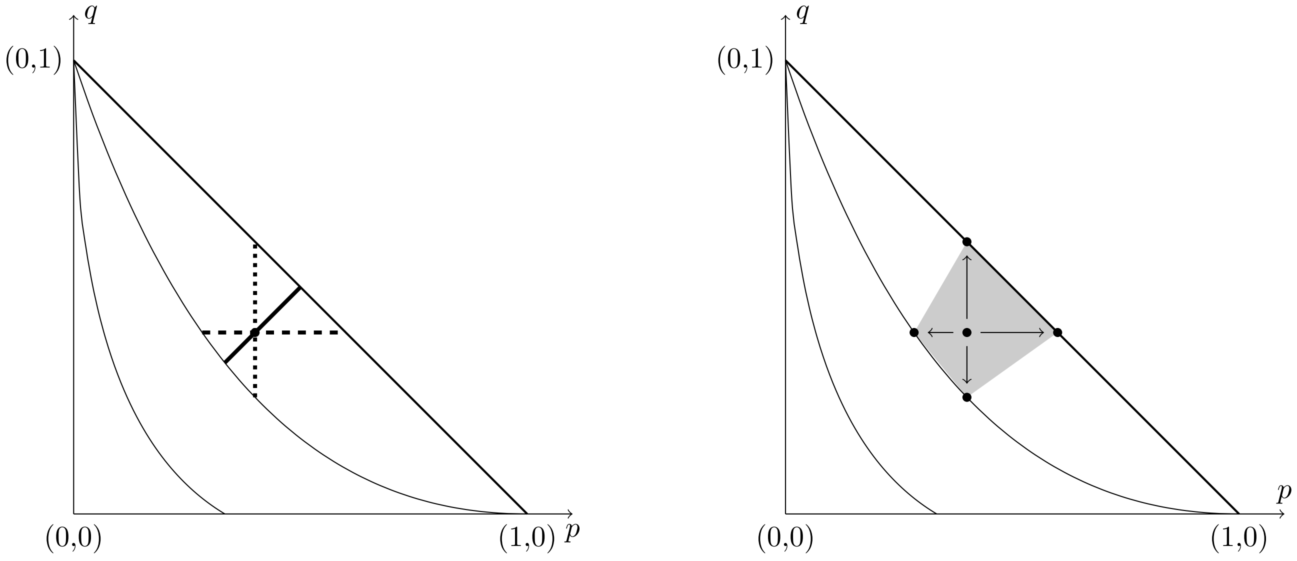

Figure 9: Corollary 5.7, first part (left) and second part (right)

Contest organizer will reveal whether both contestants have the same winning value. If their winning values are different, contest organizer will truthfully reveal the signal s_U. Otherwise, contest organizer output a signal s randomly. Therefore, the probabilities of receiving s_E and s_U are

$$\begin{aligned}\Pr(s_E) &= 2/5 \\ \Pr(s_U) &= 3/5\end{aligned}$$

When receiving s_E, the posterior distribution μ_E is

	High	Low
High	1/2	0
Low	0	1/2

When receiving s_U, the posterior distribution μ_U is

	High	Low
High	1/3	1/6
Low	1/6	1/3

Therefore, the expected total effort will be

$$\mathbb{E}_s(\Pi_O) = \Pr(s_E)\Pi_O(\mu_E) + \Pr(s_E)\Pi_O(\mu_E) = \frac{2}{5}\times\frac{3}{2} + \frac{3}{5}\times\frac{5}{3} = 1.6 > 1.57143 \quad \text{(better than full concealment)}$$

5.4.2 Surplus Triangle Analysis

In this part, we characterize what could happen to consumer and producer surplus for all possible signal structures. By Bayes plausible requirement, the set of possible surplus pairs should be convex. From the previous analysis, we know that at least two points will be attained. On one hand, if

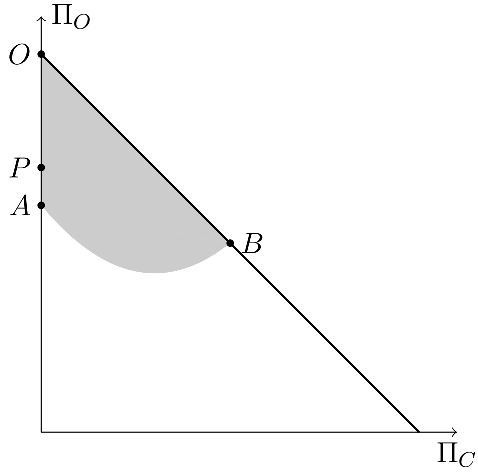

Figure 10: Surplus Triangle for Prior in Region 3

the contest organizer delivers no message to contestants, the posterior will remain at the prior. By Proposition 4.3, $\Pi_C = 0$. Meanwhile, organizer surplus does not reach the maximum. This is marked by point P in Figure 10 where P stands for "prior". On the other hand, if the contest organizer applies the optimal Bayesian persuasion signal in Theorem 5.1, organizer surplus reaches the maximum social surplus $d(1-q)+q$. This is marked by point O in Figure 10 where O stands for "optimal".

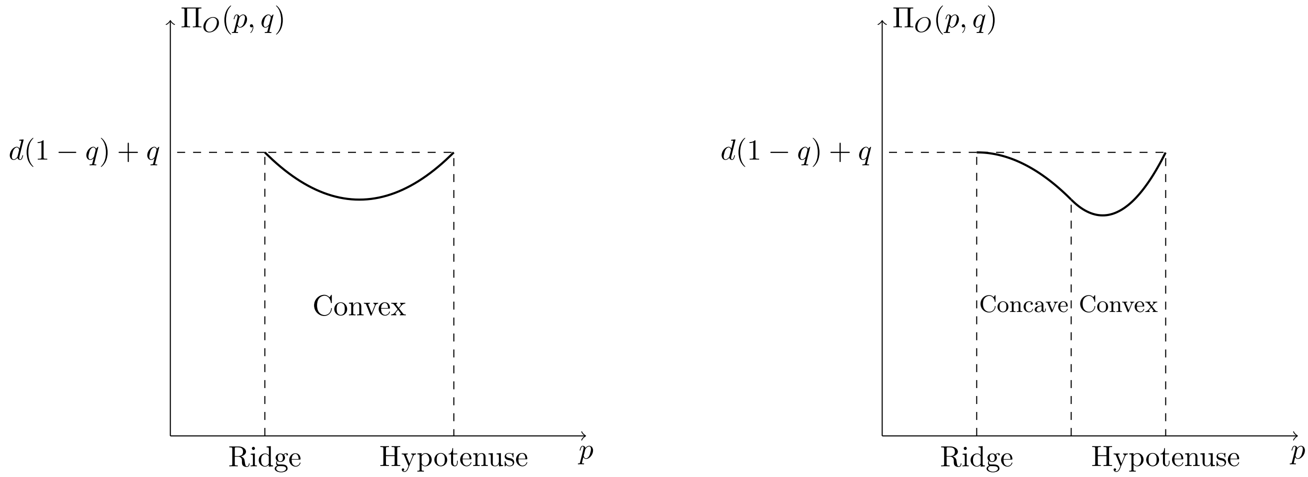

Figure 11: Cross Section, Convex (left) or Concave-Convex (right)

We are concerned with the welfare consequences of all possible signal structures, in addition to the two mentioned above. To begin with, we can identify some elementary bounds on organizer and contestants surplus for three edges of surplus triangle. The point marked A is where organizer surplus is minimized under exploitative condition. Recall that exploitative condition holds only in region 3. Point A must be attained using posteriors in region 3 only. Under some specific settings, point A lies below point P, indicating that we can find out a signal structure inferior to full concealment. Let us take the cross section by fixing the value of q, $\Pi_O(p,q)$ is either convex with respect to p or concave-convex with respect to p. When the latter case happens, we can find out a worse signal structure than full concealment. Convexification is required to solve for minimum organizer surplus.

The point marked B is where organizer surplus is minimized under efficient condition. Recall that efficient condition holds only in region 2 and the hypotenuse. As long as including posterior inside region 2, contestant surplus will be positive. At last, organizer surplus will always be strictly positive. Therefore, all posteriors inside the shaded area can be implemented.

5.5 Information Design for Regions 1 and 2

5.5.1 Optimal Public Persuasion

According to our results in subsection 5.3, only posteriors in q-axis, ridge and hypotenuse will be included in optimal posterior profile. Because any point (p, q) that located inside region 1 and region 2 and valley, we can design a better signal that shown in Theorem 5.3 to reach strictly higher utility. In order to satisfy the Bayes plausible condition, at least one of the optimal Bayesian persuasion signals should be located on the q-axis. Otherwise, by convexity of region 3, prior distribution must be in region 3, which creates contradiction. The following lemma further shows that for optimal signal, there exists exactly one posterior that located in q-axis.

Lemma 5.8. *In optimal signal, there exists at most one posterior that located at q-axis.*

By **??**, $\Pi_O(0, q)$ is a concave function with respect to q. If more than two posteriors are located at q-axis, we would assume that two of them are $(0, q_1)$ and $(0, q_2)$. By merging two posteriors into one according to their weight in posterior profile, contest designer can get strictly higher revenue.

Lemma 5.9. *In optimal signal, there exists no posterior that located at hypotenuse.*

We denote the unique posterior in q-axis as $(0, \bar{q})$. Assume there exists one posterior that located at hypotenuse, denoted as $(p_1, 1 - p_1)$. Let $(p_2, q_r(p_2))$ denote the intersection point of ridge and the line connecting $(0, \bar{q})$ and $(p_1, 1 - p_1)$. The following condition holds because of collinearity,

$$\frac{1 - p_1 - \bar{q}}{p_1} = \frac{q_r(p_2) - \bar{q}}{p_2}$$

The expected total effort for those three distributions are

$$\begin{aligned} (0, \bar{q}) &: \Pi_O(0, \bar{q}) \\ (p_1, 1 - p_1) &: dp_1 + 1 - p_1 \\ (p_2, q_r(p_2)) &: d(1 - q_r(p_2)) + q_r(p_2) \end{aligned}$$

We consider the following belief profile,

$$\begin{aligned} (0, \bar{q}) &\quad \text{with probability} \quad 1 - \frac{p_2}{p_1} \\ (p_1, 1 - p_1) &\quad \text{with probability} \quad \frac{p_2}{p_1} \end{aligned}$$

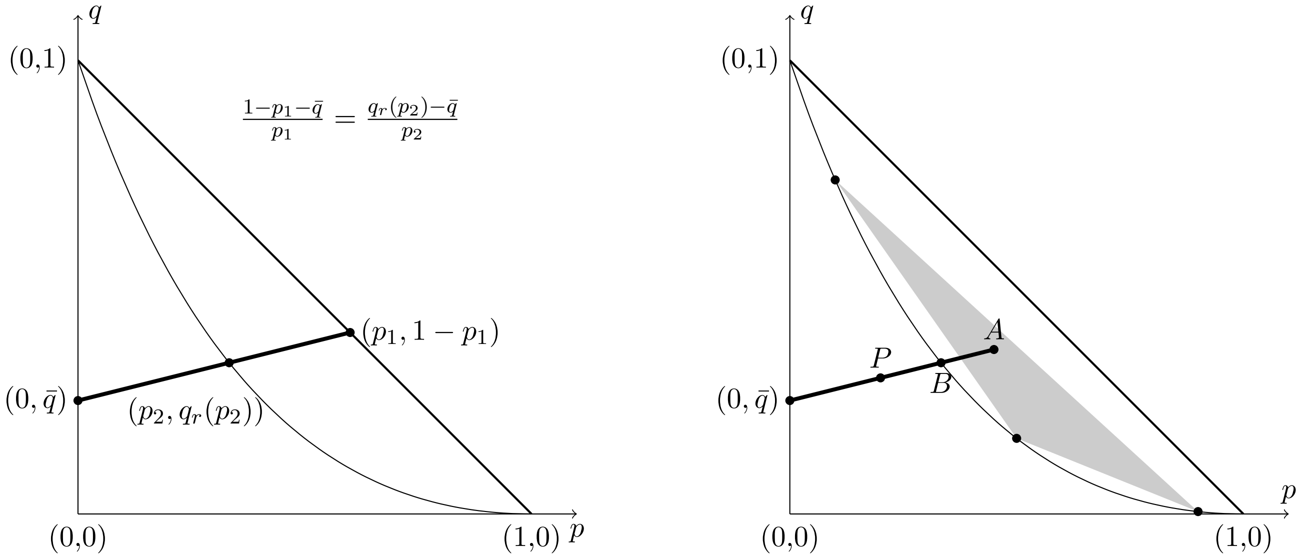

Figure 12: Lemma 5.9 (left) and Lemma 5.10 (right)

the expected total effort will be

$$(1-\frac{p_2}{p_1})\Pi_O(0,\bar{q}) + \frac{p_2}{p_1}(dp_1+1-p_1) < (1-\frac{p_2}{p_1})(d(1-\bar{q})+\bar{q}) + \frac{p_2}{p_1}(dp_1+1-p_1) = d(1-q_r(p_2))+q_r(p_2)$$

because $\Pi_O(0,\bar{q}) < d(1-\bar{q})+\bar{q}$. Therefore, posterior that located at hypotenuse is dominated when analyzing the optimal signal.

Lemma 5.10. *In optimal signal, there exists at most one posterior that located at ridge.*

We denote the prior distribution as (p_0, q_0). This is marked by point P in right panel of Figure 12. We denote the unique posterior in q-axis as $(0,\bar{q})$. Assume that for optimal signal, there exists multiple posteriors that located on ridge, denoted as $(p_i, q_r(p_i))_{i=1}^n$. Let w_i denote the weight of $(p_i, q_r(p_i))$, with $\sum_{i=1}^n w_i < 1$. The weight of $(0,\bar{q})$ is then $1-\sum_{i=1}^n w_i$ in belief profile. We define (p_c, q_c) as the weighted average of $(p_i, q_r(p_i))_{i=1}^n$ with

$$p_c \sum_{i=1}^n w_i = \sum_{i=1}^n w_i p_i \tag{24}$$

$$q_c \sum_{i=1}^n w_i = \sum_{i=1}^n w_i q_r(p_i) \tag{25}$$

by convexity of region 3 (Observation 4.6), we find out that (p_c, q_c) are located inside region 3. This is marked by point A in right panel of Figure 12. According to Bayes plausible condition, $\sum_{i=1}^n w_i = \frac{p_0}{p_c}$. Let $(\bar{p}, q_r(\bar{p}))$ be the intercept of ridge and line segment between $(0,\bar{q})$ and (p_c, q_c). This is marked by point B in right panel of Figure 12. We claim that signal associated with posteriors $(0,\bar{q})$ and $(\bar{p}, q_r(\bar{p}))$ is better than optimal signal.

The expected total effort under original belief profile is

$$\frac{p_c - p_0}{p_c}\Pi_O(0,\bar{q}) + \frac{p_0}{p_c}\Big(d(1-q_c)+q_c\Big) = d(1-q_0)+q_0 - \frac{p_c - p_0}{p_c}\Big(d(1-\bar{q})+q-\Pi_O(0,\bar{q})\Big) \qquad (26)$$

The expected total effort under new belief profile is

$$\frac{\bar{p} - p_0}{\bar{p}}\Pi_O(0,\bar{q}) + \frac{p_0}{\bar{p}}\Big(d(1-q_r(\bar{p}))+q_r(\bar{p})\Big) = d(1-q_0)+q_0 - \frac{\bar{p} - p_0}{\bar{p}}\Big(d(1-\bar{q})+q-\Pi_O(0,\bar{q})\Big) \quad (27)$$

New belief profile is preferred by contest organizer because $\bar{p} < p_c$. Combining the above three lemmas (Lemma 5.8, Lemma 5.9, Lemma 5.10), we can prove the following theorem directly..

Theorem 5.11. *When prior (p_0, q_0) is located at region 1 or 2 but neither q-axis nor the ridge, the optimal signal is associated with one posterior located at q-axis and the other located on ridge.*

The structure of the posterior profile corresponding to optimal persuasion signal has the form of linear expansion. In other words, prior and both posteriors are collinear. The following theorem told us the optimal signal exists uniquely.

Theorem 5.12. *The optimal signal for (p_0, q_0) in region 1 and 2 exists uniquely.*

Assume that there exists at least two possible optimal signals. Assume the first one is associated with posteriors $(0, q_1), (p_1, q_r(p_1))$ and the second one is associated with posteriors $(0, q_2), (p_2, q_r(p_2))$. By Bayes plausible condition, we have

$$\begin{aligned}
\frac{p_1 - p_0}{p_1}(0,q_1) + \frac{p_0}{p_1}(p_1, q_r(p_1)) &= (p_0, q_0) \\
\frac{p_2 - p_0}{p_2}(0,q_2) + \frac{p_0}{p_2}(p_2, q_r(p_2)) &= (p_0, q_0)
\end{aligned}$$

Therefore, for all $\lambda \in (0,1)$, the following Bayes plausible condition pins down another optimal signal,

$$\frac{\lambda(p_1 - p_0)}{p_1}(0,q_1) + \frac{\lambda p_0}{p_1}(p_1, q_1') + \frac{(1-\lambda)p_2 - p_0}{p_2}(0,q_2) + \frac{(1-\lambda)p_0}{p_2}(p_2, q_2') = (p_0, q_0)$$

However, we can use $\left(\frac{\lambda(p_1-p_0)}{p_1} + \frac{(1-\lambda)p_2-p_0}{p_2}\right)(0,q_3)$(a posterior on the q-axis generated by the two priors on the q-axis to replace $\frac{\lambda(p_1-p_0)}{p_1}(0,q_1) + \frac{(1-\lambda)p_2-p_0}{p_2}(0,q_2)$, which can generate strictly higher expected total effort for contest designer by Lemma 4.10. Contradiction happens.

The above analysis process takes advantage of the concavity of $\Pi_O(0,q)$ with respect to q. This property is also useful when proving the efficiency on searching optimal signal. For now on, we assume the signal satisfies Theorem 5.11. Let $g(\bar{q})$ denote the expected total effort for contest designer when the posterior at q-axis is $(0,\bar{q})$.

Lemma 5.13. *$g(\bar{q})$ is strictly quasi-concave.*

Let us look at Figure 13. Lemma 5.13 suggests that signal associated with solid line is better than signal associated with dashed line or the signal associated with dashed line. Since $g(\bar{q})$ is quasi

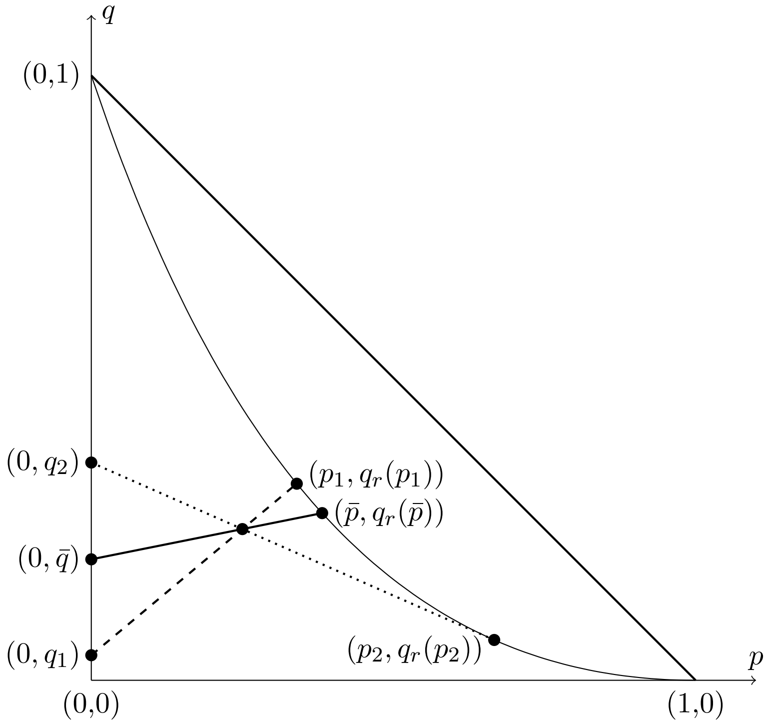

Figure 13: Quasi Concavity of $g(\bar{q})$

concave, the linear searching algorithm can always solve for $\bar{q}^* = \arg\max g(\bar{q})$ efficiently because $g(\bar{q})$ is monotonic increasing when $\bar{q} < \bar{q}^*$ and monotonic decreasing when $\bar{q} > \bar{q}^*$.

Corollary 5.14. *Searching for optimal signal is efficient.*

It is also worth noting that line segments determined by the optimal posterior profile will not intersect with each other. In other words,

Remark 5.15. *If for some prior distribution, the optimal signal corresponding to two posteriors* $(0, \bar{q}), (\bar{p}, q_r(\bar{p}))$*, then for all distribution collinear with* $(0, \bar{q})$ *and* $(\bar{p}, q_r(\bar{p}))$*, the optimal signal corresponding the same posterior profile.*

We have already present the example in region 2 (subsection 5.1). Here, we present another example to visualize the results we have derived in region 1. We set $d = 2$ and $p = q = 0.1$.

	High	Low
High	1/10	2/5
Low	2/5	1/10

Expected total effort without information disclosure is 0.72727. Contest organizer can design the following signal structure for better outcome. Similar to previous two examples, we still let $S = \{s_E, s_U\}$. The conditional distribution of s depending on combination is shown in the following table.

	s_E	s_U
(v_H, v_H)	1	0
(v_H, v_L)	1/8	7/8
(v_L, v_L)	1	0

Contest organizer will reveal whether both contestants have the same winning value. If their winning values are the same, contest organizer will truthfully reveal the signal s_E. Otherwise, contest organizer output correct signal s_U with probability less than 1. Therefore, the probabilities of receiving s_E and s_U are

$$\begin{aligned} \Pr(s_E) &= 3/10 \\ \Pr(s_U) &= 7/10 \end{aligned}$$

When receiving s_E, the posterior distribution μ_E is

	High	Low
High	1/3	1/6
Low	1/6	1/3

When receiving s_U, the posterior distribution μ_U is

	High	Low
High	0	1/2
Low	1/2	0

The two posteriors are exactly the same as subsection 5.1, consistent with Remark 5.15. Therefore, the expected total effort will be

$$\mathbb{E}_s(\Pi_O) = \Pr(s_E)\Pi_O(\mu_E)+\Pr(s_E)\Pi_O(\mu_E) = \frac{3}{10}\times\frac{5}{3}+\frac{7}{10}\times\frac{3}{4} = 1.025 > 0.72727 \quad \text{(better than full concealment)}$$

5.5.2 Surplus Triangle Analysis

When the prior distribution is in region 1, all three edges of the surplus triangle are unreachable. Only thing we know is that the set of possible surplus pairs is convex.

When the prior distribution is in region 2, both legs of surplus triangle are unreachable. From the previous analysis, we know that at least two points will be attained. On one hand, if the contest organizer delivers no message to contestants, the posterior will remain at the prior. By Proposition 4.2, surplus pair is located on efficient frontier. This is marked by point P in Figure 14 where P stands for "prior". On the other hand, if the contest organizer applies the optimal Bayesian persuasion signal, organizer surplus will increase while allocation becomes inefficient. This is marked by point O in Figure 14 where O stands for "optimal".

We are concerned with the welfare consequences of all possible signal structures, in addition to the two mentioned above. We need to characterize the potential outcomes within efficient frontier. The point marked A is where organizer surplus is maximized under efficient condition. Recall that efficient condition holds inside only region 2. Point A must be attained using posteriors in region 2 only. By Theorem 5.3, we know that $\Pi_O(p, q)$ is convex with respect to p inside region 2. Hence, point A is located at the top left of point P. The point marked B is where organizer surplus is minimized under efficient condition. The Hessian matrix of $\Pi_O(p,q)$ inside region 2 is

$$\frac{4}{(1-p+q)^3}\begin{bmatrix} 2q & 1-p-q \\ 1-p-q & -8(1-p) \end{bmatrix}$$

which implies that $\Pi_O(p,q)$ is convex with respect to q inside region 2. Hence, point B is located at the bottom right of point P. At last, organizer surplus will always be strictly positive. Therefore, all posteriors inside the shaded area can be implemented.

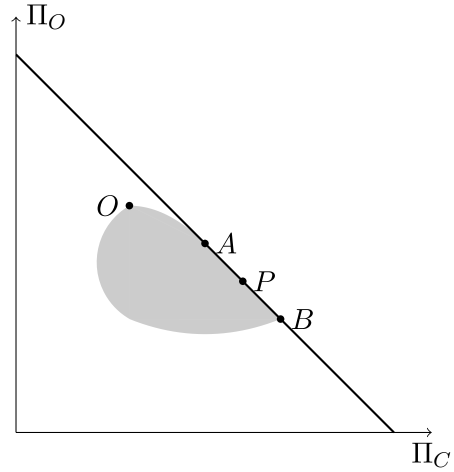

Figure 14: Surplus Triangle for Prior in Region 2

6 Extension I: No Private Information

In previous sections, we assume that players always recognize their true valuation before posterior contest game thus contest designer cannot distort the information about their own types. Nonetheless, if players do not know their realized type ex ante, should contest designer disclose type information to them? Typically, this type information has non-negative value for contest designer because he can reveal all information to guarantee the revenue of scheme $\mathbb{P}$.

Recall that the contest designer is able to commit to three different scenarios on players' information regarding their types

1. No Information. Denoted as $\mathbb{N}$. (Neither player knows his/her own winning value.)
2. Private Information. Denoted as $\mathbb{P}$. (Both players know their own winning values.)
3. Asymmetric Information. Denoted as $\mathbb{A}$. (Exactly one player know his/her own winning value.)

The private information scenario has been studied thoroughly in previous sections.

6.1 Public Persuasion with Uninformed Players, $\mathbb{N}$

The expected valuation for both players are $\frac{1+p-q}{2}d + \frac{1-p+q}{2}$. Hence, this circumstance is identical to complete information all-pay auction with winning valuation $v = \frac{1+p-q}{2}d + \frac{1-p+q}{2}$. The expected total effort for contest designer is hence

$$\Pi_O^{\mathbb{N}}(p,q) = \frac{(1+p-q)d + (1-p+q)}{2} = \frac{d+1+(d-1)(p-q)}{2} \tag{28}$$

which is linear with p and q.

Theorem 6.1. *Contest designer cannot be better off by public persuasion when both players have no private information.*

Next we will show if we allow for asymmetric posteriors, the result does not change.

6.2 Public Persuasion with Asymmetrically Informed Players, $\mathbb{A}$

Let $r = \frac{1+p-q}{2}$. The expected valuation for uninformed player is $v_A = \frac{1+p-q}{2}d + \frac{1-p+q}{2} = rd + 1 - r$ as well. And the realized distribution for informed player is

$$v_B = \begin{cases} d & w.p. \quad r \\ 1 & w.p. \quad 1-r \end{cases}$$

Proposition 6.2. *The above all-pay auction contest has the unique mixed strategy equilibrium such that*

$$F_A(x) = \begin{cases} x & x \in [0, V - rv_A] \\ \frac{x}{d} + 1 - \frac{V}{d} & x \in [V - rv_A, V] \end{cases} \tag{29}$$

$$F_L(x) = \frac{x}{(1-r)v_A} + \frac{1}{1-p} - \frac{V}{(1-r)v_A}, \quad x \in [0, V - rv_A] \tag{30}$$

$$F_H(x) = \frac{x}{rv_A} + 1 - \frac{V}{rv_A}, x \in [V - rv_A, V] \tag{31}$$

where $V = 1 + \frac{(d-1)rv_A}{d}$ *and* $v_A = (d-1)r + 1$.

Although by the general methodology that we mentioned before, we need to characterize the property of $\Pi_O^{\mathbb{A}}$ as a function of (p,q), denoted as $\Pi_O^{\mathbb{A}}(p,q)$. However, by variable substitution, we use $r = \frac{1+p-q}{2}$ instead of p, q, with range $r \in [0,1]$, denoted as $\Pi_O^{\mathbb{A}}(r)$. This abuse use of notation will not lead to confusion because two expressions contain different number of variables. According to the equilibrium characterization in Chen, Kuang and Zheng (2018). The resulting expected total effort is hence

$$\begin{aligned} \Pi_O^{\mathbb{A}}(r) &= \mathbb{E}(x_A) + r\mathbb{E}(x_H) + (1-r)\mathbb{E}(x_L) \\ &= \frac{rv_AV}{2d} + \frac{V - rv_A}{2} + \frac{V^2}{2v_A} \end{aligned} \tag{32}$$

The second order condition gives us that

Lemma 6.3. *The second derivative of* $\Pi_O^{\mathbb{A}}(r)$ *is always positive when* $r \in [0,1]$.

$$\frac{\partial^2 \Pi_O^{\mathbb{A}}(r)}{\partial r^2} > 0 \tag{33}$$

Hence, only posteriors locate on q-axis or p-axis may be included in optimal posteriors because other points are strictly dominated. What's more, we can get the second order condition for $f_A(p,q)$ by Lemma 6.3 as well as chain rule.

$$\begin{aligned} \frac{\partial^2 \Pi_O^{\mathbb{A}}(p,q)}{\partial p^2} &> 0 \\ \frac{\partial^2 \Pi_O^{\mathbb{A}}(p,q)}{\partial q^2} &> 0 \end{aligned}$$

which proves that only vertex posteriors[2] may be contained in optimal posteriors. We need further calculate the $\Pi_O^{\mathbb{A}}$ on those vertex posteriors,

$$\begin{aligned} \Pi_O^{\mathbb{A}}(0,0) &= 1 + \frac{(d-1)(3d^3+5d^2+17d-1)}{32d^2(d+1)} \\ \Pi_O^{\mathbb{A}}(1,0) &= d \\ \Pi_O^{\mathbb{A}}(0,1) &= 1 \end{aligned}$$

Theorem 6.4. *For prior* (p_0, q_0)*, when apply asymmetry disclosure policy, the contest designer will induce the following posterior profile,*

$$\begin{aligned} &(0,0) \quad \textit{with probability} \quad 1 - p_0 - q_0 \\ &(1,0) \quad \textit{with probability} \quad p_0 \\ &(0,1) \quad \textit{with probability} \quad q_0 \end{aligned}$$

6.3 Comparison

Theorem 6.5. *From the contest designer's perspective, for any prior* (p_0, q_0)*,*

$$\mathbb{N} \succ \mathbb{A}$$

When applying policy $i = \mathbb{N}$ or $\mathbb{A}$, the resulting expected total effort will lies on a plane that defined by $(0,0,\Pi_O^i(0,0)), (1,0,\Pi_O^i(1,0))$ and $(0,1,\Pi_O^i(0,1))$ where $i = \mathbb{N}, \mathbb{A}$. Since

$$\begin{aligned} \Pi_O^{\mathbb{A}}(1,0) &= \Pi_O^{\mathbb{N}}(1,0) \\ \Pi_O^{\mathbb{A}}(0,1) &= \Pi_O^{\mathbb{N}}(0,1) \end{aligned}$$

[2] (0,0),(1,0) and (0,1)

we need only to compare $\Pi_O^{\mathbb{A}}(0,0)$ and $\Pi_O^{\mathbb{N}}(0,0)$,

$$\Pi_O^{\mathbb{A}}(0,0) - \Pi_O^{\mathbb{N}}(0,0) = \frac{(d-1)(-13d^3 - 11d^2 + 17d - 1)}{32d^2(d+1)} < 0$$

Theorem 6.6. *The contest designer's preference over $\mathbb{P}$ and $\mathbb{N}/\mathbb{A}$ depends on the values of parameters. In other words, the following three circumstances are all possible:*

$$\mathbb{P} \succ \mathbb{N} \succ \mathbb{A}$$
$$\mathbb{N} \succ \mathbb{P} \succ \mathbb{A}$$
$$\mathbb{N} \succ \mathbb{A} \succ \mathbb{P}$$

7 Extension II: Ridge Phenomena with Discrete Distribution

Most of our results in the benchmark model depend on the ridge phenomenon of positive correlation. For both validity and optimality issues, we require one posterior located on ridge. For ridge distribution, resulting equilibrium is both efficient and exploitative, which helps contest organizer reaches the social maximum surplus. In this section, we extend this phenomenon from a symmetric binary joint distribution setup to a symmetric discrete joint distribution setup. We show that given the marginal distribution of the winning value, there exists unique joint distribution that satisfies the efficient condition and the exploitative condition.

The winning value of participant $\mathbb{P} = \text{I}, \text{II}$, denoted as v_i is now drawn from a finite set $\Omega_{\mathbb{P}}$. Both players share the same finite state space with N elements, $\Omega_{\mathbb{P}} = \Omega = \{v_1, v_2, \cdots, v_N\}$ where $v_1 < v_2 < \cdots < v_N$. We assume the joint distribution of $(v_{\text{I}}, v_{\text{II}})$ is symmetric,

$$\Pr(v_i, v_j) = \Pr(v_j, v_i)$$

Let $\Pr(v_1), \cdots, \Pr(v_N) > 0$ denote the marginal distribution of winning value. This is well defined when distribution is symmetric.

7.1 Monotonicity

Siegel (2014) analyzes the monotonicity of equilibrium for asymmetric all-pay auctions with interdependent valuations. Our setup is a special case where winning value of one contestant does not depend on the other. For individual value and symmetric distribution, **WM** condition (Siegel, 2014) is reduced to the following. $\Pr(v_{-\mathbb{P}}|v_{\mathbb{P}})v_{\mathbb{P}}$ weakly increases in $v_{\mathbb{P}}$ for every possible value of $v_{-\mathbb{P}}$ where $-\mathbb{P}$ denote the rival of participant $\mathbb{P}$. There are $N(N-1)$ inequality constraints in total.

$$\Pr(v_j|v_{i+1})v_{i+1} \geq \Pr(v_j|v_i)v_i, \forall i = 1, \cdots, N-1; j = 1, \cdots, N \tag{\textbf{WM}}$$

When all constraints are satisfied, the equilibrium strategy for player with winning value v_k is bidding uniformly within the following interval with probability density $\frac{1}{\Pr(v_k|v_k)v_k}$,

$$\left[\sum_{i=1}^{k-1}\Pr(v_i|v_i)v_i, \sum_{i=1}^{k}\Pr(v_i|v_i)v_i\right]$$

7.2 Exploitation Given Monotonicity

Given the monotonicity, equilibrium is exploitative when all types of player get zero utility in equilibrium. For player with winning value v_k, the utility of bidding $\sum_{i=1}^{k-1}\Pr(v_i|v_i)v_i$ is zero. Mathematically,

$$v_k \underbrace{\sum_{i=1}^{k-1}\Pr(v_i|v_k)}_{\text{Winning Probability}} = \sum_{i=1}^{k-1}\Pr(v_i|v_i)v_i$$

According to the (**WM**) condition, $\Pr(v_i|v_k)v_k \geq \Pr(v_i|v_i)v_i$ for $i < k$. Therefore, joint distribution must satisfy the following exploitative conditions. There are $\frac{N(N-1)}{2}$ equation constraints in total.

$$\Pr(v_i|v_k)v_k = \Pr(v_i|v_i)v_i, \forall k > i \tag{\textbf{E}}$$

For instance, when $N = 2$, (**E**) condition required that

$$\Pr(L|H)H = \Pr(L|L)L$$

which happens to be the parametric equation of ridge.

7.3 Conditional Distribution

In this part, we will show that the joint distribution achieveing both efficience and exploitation is uniquely determined given marginal distribution of winning value. Let us consider the problem of solving conditional probabilities of such joint distribution. Namely, independent variables are $\Pr(v_i|v_j)$, N^2 in total. To meet the symmetric condition, we have $\frac{N(N-1)}{2}$ equation constraints.

$$\Pr(v_i|v_j)\Pr(v_j) = \Pr(v_j|v_i)\Pr(v_i), \forall i \neq j \tag{\textbf{S}}$$

At last, we have N regularity constraint.

$$\sum_{i=1}^{N}\Pr(v_i|v_j) = 1, \forall j \tag{\textbf{R}}$$

Therefore, we have N^2 independent variables and N^2 linear equations, including (**E**)(**S**)(**R**). The solutions should be nonnegative, hence we define the positive condition and weakly positive condition.

$$\Pr(v_i|v_j) > 0, \forall i, j \tag{\textbf{P}}$$

$$\Pr(v_i|v_j) \geq 0, \forall i,j \tag{WP}$$

Proposition 7.1. *The unique solution of linear equation system* (**E**)(**S**)(**R**) *is given by*

$$\begin{aligned}\Pr(v_k|v_j) &= \Big(1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j)\Big)\Big(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\Big)^{-1} \frac{\Pr(v_k)}{v_k}, k \geq j \\ \Pr(v_k|v_j) &= \frac{\Pr(v_j|v_k)\Pr(v_k)}{\Pr(v_j)}, k < j\end{aligned}$$

The solution given by the above formula is defined by recursion. This recursion is well defined because when computing $\Pr(v_k|v_j), k \geq j$, all $\Pr(v_i|v_j)$ have already computed. The sequence of solving conditional probabilities are listed as follow,

Step 1. Solve for $\Pr(v_k|v_1), k \geq 1$ $\Rightarrow$ Compute for all $\Pr(v_1|v_j), j > 1$

Step 2. Solve for $\Pr(v_k|v_2), k \geq 2$ $\Rightarrow$ Compute for all $\Pr(v_2|v_j), j > 2$

$\ldots$

Step i. Solve for $\Pr(v_k|v_i), k \geq i$ $\Rightarrow$ Compute for all $\Pr(v_i|v_j), j > i$

$\ldots$

Step N. Solve for $\Pr(v_N|v_N)$

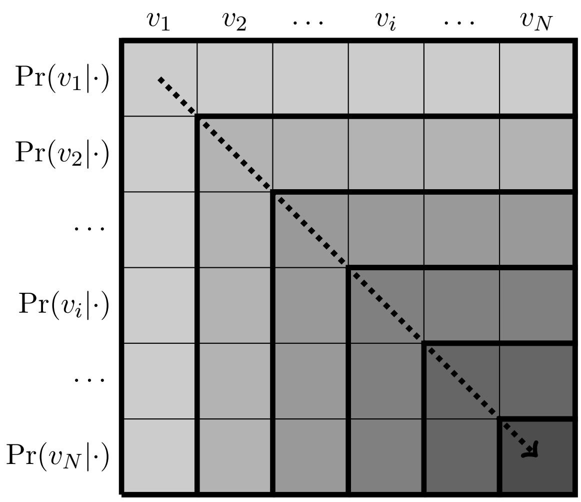

Figure 15: Iteration Process

This computation process is well defined because $1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j)$ is always positive by Lemma 7.2.

Lemma 7.2. *The solution provided in Proposition 7.1 satisfying the the following inequality* $1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j) \geq \left(\sum_{i=1}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1}\left(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\right)$.

To finish the proof of validity of solution given by Proposition 7.1, we need to verify that the solution meets all inequality constraints, (**P**)(**WM**). The proof process of both conditions depends on

Lemma 7.2.

Proposition 7.3. *The solution provided in Proposition 7.1 satisfying the inequality conditions* (**P**)(**WM**)*.*

7.4 Joint Distribution and Examples

Directly by Proposition 7.1,

Proposition 7.4. *The Bayes Nash equilibrium under the following joint distribution is both efficient and exploitative,*

$$\Pr(v_k, v_j) = \Big(\Pr(v_j) - \sum_{i=1}^{j-1} \Pr(v_i, v_j)\Big)\Big(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\Big)^{-1} \frac{\Pr(v_k)}{v_k}, k \geq j$$

Example 7.5. *We consider the ternary distribution case where* $(v_1, v_2, v_3) = (2, 3, 6)$*.*

We use $\begin{bmatrix} \Pr(v_1, v_1) & \Pr(v_1, v_2) & \Pr(v_1, v_3) \\ \Pr(v_2, v_1) & \Pr(v_2, v_2) & \Pr(v_2, v_3) \\ \Pr(v_3, v_1) & \Pr(v_3, v_2) & \Pr(v_3, v_3) \end{bmatrix}$ *to express the joint distribution.*

If $\Pr(v_1) = \Pr(v_2) = \Pr(v_3) = \frac{1}{3}$*, the joint distribution will be* $\begin{bmatrix} \frac{1}{6} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \frac{4}{27} & \frac{2}{27} \\ \frac{1}{18} & \frac{2}{27} & \frac{11}{54} \end{bmatrix}$ *by*

$$\begin{bmatrix} \frac{1}{6} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \Pr(v_2, v_2) & \Pr(v_2, v_3) \\ \frac{1}{18} & \Pr(v_3, v_2) & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{1}{6} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \frac{4}{27} & \frac{2}{27} \\ \frac{1}{18} & \frac{2}{27} & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{1}{6} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \frac{4}{27} & \frac{2}{27} \\ \frac{1}{18} & \frac{2}{27} & \frac{11}{54} \end{bmatrix}$$

If $\Pr(v_1) = \frac{1}{2}, \Pr(v_2) = \Pr(v_3) = \frac{1}{4}$*, the joint distribution will be* $\begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \frac{5}{54} & \frac{5}{108} \\ \frac{1}{18} & \frac{5}{108} & \frac{4}{27} \end{bmatrix}$ *by*

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \Pr(v_2, v_2) & \Pr(v_2, v_3) \\ \frac{1}{18} & \Pr(v_3, v_2) & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \frac{5}{54} & \frac{5}{108} \\ \frac{1}{18} & \frac{5}{108} & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{1}{3} & \frac{1}{9} & \frac{1}{18} \\ \frac{1}{9} & \frac{5}{54} & \frac{5}{108} \\ \frac{1}{18} & \frac{5}{108} & \frac{4}{27} \end{bmatrix}$$

If $\Pr(v_2) = \frac{1}{2}, \Pr(v_1) = \Pr(v_3) = \frac{1}{4}$*, the joint distribution will be* $\begin{bmatrix} \frac{3}{32} & \frac{1}{8} & \frac{1}{32} \\ \frac{1}{8} & \frac{3}{10} & \frac{3}{40} \\ \frac{1}{32} & \frac{3}{40} & \frac{23}{160} \end{bmatrix}$ *by*

$$\begin{bmatrix} \frac{3}{32} & \frac{1}{8} & \frac{1}{32} \\ \frac{1}{8} & \Pr(v_2, v_2) & \Pr(v_2, v_3) \\ \frac{1}{32} & \Pr(v_3, v_2) & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{3}{32} & \frac{1}{8} & \frac{1}{32} \\ \frac{1}{8} & \frac{3}{10} & \frac{3}{40} \\ \frac{1}{32} & \frac{3}{40} & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{3}{32} & \frac{1}{8} & \frac{1}{32} \\ \frac{1}{8} & \frac{3}{10} & \frac{3}{40} \\ \frac{1}{32} & \frac{3}{40} & \frac{23}{160} \end{bmatrix}$$

If $\Pr(v_3) = \frac{1}{2}, \Pr(v_1) = \Pr(v_2) = \frac{1}{4}$, *the joint distribution will be* $\begin{bmatrix} \frac{3}{28} & \frac{1}{14} & \frac{1}{14} \\ \frac{1}{14} & \frac{5}{56} & \frac{5}{56} \\ \frac{1}{14} & \frac{5}{56} & \frac{19}{56} \end{bmatrix}$ *by*

$$\begin{bmatrix} \frac{3}{28} & \frac{1}{14} & \frac{1}{14} \\ \frac{1}{14} & \Pr(v_2, v_2) & \Pr(v_2, v_3) \\ \frac{1}{14} & \Pr(v_3, v_2) & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{3}{28} & \frac{1}{14} & \frac{1}{14} \\ \frac{1}{14} & \frac{5}{56} & \frac{5}{56} \\ \frac{1}{14} & \frac{5}{56} & \Pr(v_3, v_3) \end{bmatrix} \Longrightarrow \begin{bmatrix} \frac{3}{28} & \frac{1}{14} & \frac{1}{14} \\ \frac{1}{14} & \frac{5}{56} & \frac{5}{56} \\ \frac{1}{14} & \frac{5}{56} & \frac{19}{56} \end{bmatrix}$$

For prior distribution that can be expressed as linear combination of posterior joint distributions that meet both monotonicity condition and exploitation condition, contest organizer can achieve first best outcome by information design alone.

8 Conclusion

In this paper, we study the information design problem and fully characterize the optimal information disclosure policy for the contest organizer in a simultaneous 2-player 2-type all-pay auction contest environment, where players have private information about their own types. We allow the players' ex-ante symmetric type distributions to be correlated, and the information disclosure policy to take the stochastic approach of Bayesian persuasion, which is a generalization of the traditional information disclosure policy. We focus on the effectiveness and optimality of Bayesian persuasion signal. The optimal signal, the structure of which depends on the degree of the correlation of players' types, is completed characterized under each setting and is shown to work better than the type-dependent discrete information disclosure policy proposed by Lu, Ma and Wang (2018).

The main results of this article are based on ridge phenomenon of positive correlation. It states that when distribution has moderate positive correlation, contest organizer may achieve first best in equilibrium. Social welfare is maximized and all surplus is attained by contest organizer. If winning values of both players are sufficiently positively correlated, the optimal design consists of two posteriors with one representing a perfect positive correlation and the other representing a positive correlation identified by a cutoff condition. If winning values are mildly correlated, the optimal design consists of two posteriors with one such that both players being high types is impossible and the other representing a positive correlation identified by the cutoff condition. The main insights behind our study is that partial information stimulates participants to devote more effort than full information and no information. On one hand, Bayesian persuasion can be viewed as a test or an investigation, hence contest designer will not design accurate test that trying to disclose type information completely. On the other hand, since type information is unobservable and often inferred from data, partial disclosure is more realistic than full disclosure. Consider a crowd sourcing game (or known as innovation contest), the platform may require each contestant to submit his profile and release the score of all profiles to all contestant before exerting effort. This methodology is also meaningful in various scenarios including grant applications..

Two extensions of the benchmark model are considered. First, we investigate the case of limited information on individual's own value. We allow the contest designer choose one of three schemes depending on the number of player having private information on their winning values, two, one or zero. By conducting welfare comparisons among three cases, we show that the setting where only one player knows his own type is dominated by the setting where no players know their own type, from the designer's perspective. Our second extension seeks to generalize ridge phenomenon of positive correlation to arbitrary discrete joint distribution. Given the marginal distribution, there exists unique joint distribution whose equilibrium satisfies both monotonicity and exploitation.

To the best of our knowledge, this is the first study on full characterization of public persuasion for games with two-sided incomplete information. We also make the first attempt to show new properties of the expected payoff function in two-player all-pay auction contests, which are essential in pining down the optimal information structure for the designer. The results that the optimal signals are generated by up to two posteriors and can outperform the traditional type-dependent discrete disclosure polices provide us with a better understanding of the optimal information design in the all-pay auction environment.

We notice a few directions to further generalize our work on optimal information design via public persuasion. Two key assumptions of benchmark model is the two player assumption and binary type assumption. For future studies, one is to extend the two player environment to the general multi player environment, and the other one is to extend the binary type setup to the general discrete type setup, as we did in Section 7. In addition, Cai, Kuang and Zheng (2019) studies a joint design problem for a similar all-pay auction contest situation by considering both information design and reserve price design.

Appendix A: Proofs

Proof of Observation 4.5. **Part I**. Take the partial derivative on d in Equation 10 with parameter q.

$$\begin{aligned}(4d-2)p\frac{\partial p}{\partial d} - 2dq\frac{\partial p}{\partial d} - 2d\frac{\partial p}{\partial d} &= 2p + 2pq - 2p^2 \\ (4dp - 2p - 2dq - 2d)\frac{\partial p}{\partial d} &= 2p(1+q-p) \\ \frac{\partial p}{\partial d} &= \frac{p(1+q-p)}{(2d-1)p - dq - d}\end{aligned}$$

the numerator is strictly less than 0 because $p_v(q) \le p_v(0) = \frac{1}{2d-1}$.

Part II. Take the partial derivative on d in Equation 12 with parameter q.

$$\begin{aligned}-2dp\frac{\partial p}{\partial d} + 2q\frac{\partial p}{\partial d} + 2d\frac{\partial p}{\partial d} &= 1 + p^2 - 2p - q^2 \\ (2q + 2d - 2dp)\frac{\partial p}{\partial d} &= (1-p-q)(1-p+q) \\ \frac{\partial p}{\partial d} &= \frac{(1-p-q)(1-p+q)}{2q+2d-2dp} > 0\end{aligned}$$

□

Proof of Observation 4.6. Equation 12 represents a quadratic curve. Intuitively, one quadratic curve and one straight line form a convex set. In details,

Case(1) When $d < 2$, then Equation 12 represents an ellipse.

Case(2) When $d = 2$, then Equation 12 represents a parabola.

Case(3) When $d > 2$, then Equation 12 represents a hyperbola. □

Proof of Proposition 4.14. When p, q satisfies Equation 10, according to Theorem 4.4, we have the expected total effort

$$\begin{aligned}\Pi_O(p, q_v(p)) &= q + pd + \frac{2q(1+p-q)}{1-p+q} \\ &= \frac{3q + pq - q^2 + pd - dp^2 + pqd}{1-p+q} \\ &= \frac{(d-1)p^2 - (d-1)pq - dp + q + 1}{1-p+q} \\ &= \frac{(d-1)p^2 - (d-1)pq - (d-1)p + 1 - p + q}{1-p+q} \\ &= 1 - (d-1)p\end{aligned}$$

□

Proof of Lemma 4.15. During the proof, we fix the q level. To prove this lemma, two steps are required. Firstly, calculate the slope connecting $(0, q, \Pi_O(0, q))$ and $(p_r(q), q, \Pi_O(p_r(q), q))$. Secondly, prove $\Pi_O(p_v(q), q) \le \Pi_O(0, q) + \text{slope} \cdot p_v(q)$.

Step (1) The difference between $\Pi_O(p_r(q), q)$ and $\Pi_O(0, q)$ is computed as follow,

$$\begin{aligned}\Pi_O(p_r(q), q) - \Pi_O(0, q) &= d(1-q) + q - \frac{(1+4q-q^2)d + (1-6q+q^2)}{(2+2q)d - 4q} \\ &= \frac{(2-2q^2)d^2 - (1+6q-7q^2)d - (1-6q+5q^2)}{(2+2q)d-4q}\end{aligned}$$

So the slope of the line segment between $(0, q, \Pi_O(0, q))$ and $(p_r(q), q, \Pi_O(p_r(q), q))$ is

$$\begin{aligned}\text{slope} &= \frac{\Pi_O(p_r(q), q) - \Pi_O(0, q)}{p_r(q)} \\ &= \frac{(2+2q)d^2 - (1+7q)d - (1-5q)}{2(d+dq-2q)^2}\left(d + q + \sqrt{4dq + (d-1)^2q^2}\right) \\ &= \frac{\Big((2+2q)d + (1-5q)\Big)(d-1)}{2(d+dq-2q)^2}\left(d + q + \sqrt{4dq + (d-1)^2q^2}\right)\end{aligned}$$

where $p_r(q) = \frac{(d+dq-2q)(1-q)}{d+q+\sqrt{4dq+(d-1)^2q^2}}$ is computed by solving Equation 12.

Step (2) Solving Equation 10, we have $p_v(q) = \frac{d+dq-\sqrt{(1+q)^2d^2-2(1-q)^2d+(1-q)^2}}{2d-1}$. According to Lemma 4.15,

$$\Pi_O(p_v(q), q) = 1 - (d-1)p_v(q)$$

As for $\Pi_O(0, q) + \text{slope} \cdot p_v(q)$, it equals

$$\underbrace{\frac{(1+4q-q^2)d + (1-6q+q^2)}{2(d+dq-2q)}}_{\Pi_O(0,q)} + \underbrace{\frac{(2+2q)d^2 - (1+7q)d - (1-5q)}{2(d+dq-2q)^2}\left(d + q + \sqrt{4dq + (d-1)^2q^2}\right)}_{\text{slope}} p_v(q)$$

Take the difference of $\Pi_O(p_v(q), q)$ and $\Pi_O(0, q) + \text{slope} \cdot p_v(q)$, and divide this difference by $d-1$, we have

$$\phi(q) = \frac{(2+2q)d + (1-5q)}{2(d+dq-2q)^2}\left(d + q + \sqrt{4dq + (d-1)^2q^2}\right)p_v(q) - \frac{(1-q)^2}{2(d+dq-2q)} + p_v(q)$$

This function is non-negative for all $d \geq 1$ and $q \in (0, 1)$. □

Proof of Theorem 5.3. Mathematically, the above signal is strictly better than no disclosure if and only if

$$\Pi_O(\lambda p_r(q), q) < \lambda \Pi_O(p_r(q), q) + (1-\lambda)\Pi_O(0, q), \forall \lambda \in (0, 1)$$

When fixing d and q, we can prove that $\Pi_O(p, q)$ is convex in region 1 with respect to p. Likewise, we can prove that $Pi_O(p, q)$ is also convex in region 2 with respect to p. Hence, $\Pi_O(p, q_0)$ is piecewise convex. By Lemma 4.15, we have

$$\Pi_O(p_v(q), q) < \frac{p_r(q) - p_v(q)}{p_r(q)}\Pi_O(0, q) + \frac{p_v(q)}{p_r(q)}\Pi_O(p_r(q), q)$$

when $\lambda p_r(q) < p_v(q)$,

$$\begin{aligned}\Pi_O(\lambda p_r(q), q) &< \frac{\lambda p_r(q)}{p_v(q)}\Pi_O(p_v(q), q) + \frac{p_v(q) - \lambda p_r(q)}{p_v(q)}\Pi_O(0, q) \\ &< \frac{\lambda p_r(q)}{p_v(q)}\Big(\frac{p_r(q) - p_v(q)}{p_r(q)}\Pi_O(0, q) + \frac{p_v(q)}{p_r(q)}\Pi_O(p_r(q), q)\Big) + \frac{p_v(q) - \lambda p_r(q)}{p_v(q)}\Pi_O(0, q) \\ &= \lambda\Pi_O(p_r(q), q) + (1-\lambda)\Pi_O(0, q)\end{aligned}$$

when $\lambda p_r(q) > p_v(q)$,

$$\begin{aligned}\Pi_O(\lambda p_r(q), q) &< \frac{(1-\lambda)p_r(q)}{p_r(q) - p_v(q)}\Pi_O(p_v(q), q) + \frac{\lambda p_r(q) - p_v(q)}{p_r(q) - p_v(q)}\Pi_O(p_r(q), q) \\ &< \frac{(1-\lambda)p_r(q)}{p_r(q) - p_v(q)}\Big(\frac{p_r(q) - p_v(q)}{p_r(q)}\Pi_O(0, q) + \frac{p_v(q)}{p_r(q)}\Pi_O(p_r(q), q)\Big) + \frac{\lambda p_r(q) - p_v(q)}{p_r(q) - p_v(q)}\Pi_O(p_r(q), q) \\ &= \lambda\Pi_O(p_r(q), q) + (1-\lambda)\Pi_O(0, q)\end{aligned}$$

Convexity in Region 1. In region 1, the second order derivative is expressed as

$$\frac{\partial^2 \Pi_O(p, q)}{\partial p^2} = \frac{4(d-1)^2 qB}{A^3}$$

where

$$\begin{aligned}A &= -2q(1+p-q) + d(1-2p+p^2-q^2) \\ B &= (1+p-q)^3 q + Cd \\ C &= (2-q)p^3 - 3q(1-q)p^2 - 3(1-q)^2(2+q)p + (1-q)^2(4+3q+q^2)\end{aligned}$$

We now analyze the sign of $\frac{\partial^2 \Pi_O(p,q)}{\partial p^2}$ step by step. In region 1, we have $\sqrt{p} + \sqrt{q} < 1$ and hence $2\sqrt{pq} < 1 - p - q$. We take the square of both side and get $4pq < 1 + p^2 + q^2 - 2p - 2q + 2pq$ and re-arrange as $1 + p^2 + q^2 - 2p - 2q - 2pq > 0$, implying that $A > \underbrace{1 + p^2 + q^2 - 2p - 2q - 2pq}_{A_{d=1}} > 0$.

Take the partial derivative of C on p, we have

$$\frac{\partial C}{\partial p} = 3(2-q)p^2 - 6q(1-q)p - 3(1-q)^2(2+q)$$

Since $p \leq 1 - q$,

$$\frac{\partial C}{\partial p} < 3(2-q)(1-q)^2 - 3(1-q)^2(2+q) = -6q(1-q)^2 \leq 0$$

C is monotonic decreasing when $p \in [0, 1-q]$,

$$\begin{aligned} C &= (2-q)p^3 - 3q(1-q)p^2 - 3(1-q)^2(2+q)p + (1-q)^2(4+3q+q^2) \\ &\geq (2-q)(1-q)^3 - 3q(1-q)^3 - 3(1-q)^3(2+q) + (1-q)^2(4+3q+q^2) \\ &= -(4+7q)(1-q)^3 + (1-q)^2(4+3q+q^2) \\ &= (1-q)^2\Big((4+3q+q^2) - (4+3q-7q^2)\Big) \\ &= 8(1-q)^2q^2 \geq 0 \end{aligned}$$

we can conclude that C is nonnegative inside entire domain. B is then nonnegative. Combining all the above formulas, we know that $\Pi_O(p, q)$ is a convex function in region 1 with respect to p.

Convexity in Region 2. In region 2, we can easily show that the second order derivative is positive,

$$\frac{\partial^2 \Pi_O(p, q)}{\partial p^2} = \frac{8q}{(1-p+q)^3} \geq 0$$

□

Proof of Lemma 5.13. We need to prove,

$$g(\bar{q}) > \min\{g(q_1), g(q_2)\}, \forall 0 \leq q_1 < \bar{q} < q_2 \leq 1$$

Assume the associated posteriors with $g(q_i)$ is $(0, q_i)$ and $(p_i, q_r(p_i))$,

$$g(q_i) = \frac{p_i - p_0}{p_i}\Pi_O(0, q) + \frac{p_0}{p_i}\Pi_O(p_i, q_r(p_i)), i = 1, 2$$

For any $\bar{q} \in (q_1, q_2)$, there exists $\lambda \in (0, 1)$ such that

$$\lambda\frac{p_1 - p_0}{p_1}q_1 + (1-\lambda)\frac{p_2 - p_0}{p_2}q_2 = (\lambda\frac{p_1 - p_0}{p_1} + (1-\lambda)\frac{p_2 - p_0}{p_2})\bar{q}$$

We consider the following belief profile $\mathbb{B}_0$ which yield expected total effort of $\lambda g(q_1) + (1-\lambda)g(q_2)$,

$$\begin{aligned} (0, q_1) &\quad \text{with probability} \quad \lambda\frac{p_1 - p_0}{p_1} \\ (p_1, q_r(p_1)) &\quad \text{with probability} \quad \lambda\frac{p_0}{p_1} \\ (0, q_2) &\quad \text{with probability} \quad (1-\lambda)\frac{p_2 - p_0}{p_2} \\ (p_2, q_r(p_2)) &\quad \text{with probability} \quad (1-\lambda)\frac{p_0}{p_2} \end{aligned}$$

By concavity of $\Pi_O(0, q)$ in q-axis (Lemma 4.10), the belief profile $\mathbb{B}_0$ is dominated by the following belief profile $\mathbb{B}_1$,

$$\begin{array}{rll} (0,\bar{q}) & \text{with probability} & \lambda\frac{p_1-p_0}{p_1}+(1-\lambda)\frac{p_2-p_0}{p_2} \\ (p_1,q_r(p_1)) & \text{with probability} & \lambda\frac{p_0}{p_1} \\ (p_2,q_r(p_2)) & \text{with probability} & (1-\lambda)\frac{p_0}{p_2} \end{array}$$

Let (p_c,q_c) be the weighted average of $(p_1,q_r(p_1))$ and $(p_2,q_r(p_2))$

$$\begin{aligned} \Big(\lambda\frac{p_0}{p_1}+(1-\lambda)\frac{p_0}{p_2}\Big)p_c &= p_0 \\ \Big(\lambda\frac{p_0}{p_1}+(1-\lambda)\frac{p_0}{p_2}\Big)q_c &= \lambda\frac{p_0}{p_1}q_r(p_1)+(1-\lambda)\frac{p_0}{p_2}q_r(p_2) \end{aligned}$$

By proof of Lemma 5.10, we can conclude that for a set of posteriors that include multiple points located on ridge with weighted central (p_c,q_c), we can increase expected total effort by introducing one posterior $(\bar{p},q_r(\bar{p}))$ in the ridge that collinear with (p_c,q_c) and $(0,\bar{q})$ to replace all points on the ridge. Now the belief profile $\mathbb{B}_2$ has revenue $g(\bar{q})$,

$$\begin{array}{rll} (0,\bar{q}) & \text{with probability} & \frac{\bar{p}-p_0}{\bar{p}} \\ (\bar{p},q_r(\bar{p})) & \text{with probability} & \frac{p_0}{\bar{p}} \end{array}$$

Therefore, on the behalf of contest designer,

$$\mathbb{B}_2 \succ \mathbb{B}_1 \succ \mathbb{B}_0$$

Rewrite the above preference using expected total effort, we have

$$g(\bar{q}) > \lambda g(q_1)+(1-\lambda)g(q_2) \geq \min\{g(q_1),g(q_2)\}$$

□

Proof of Theorem 6.6. **Case (1)** If prior distribution lies on the ridge, then

$$\mathbb{P} \succ \mathbb{N} \succ \mathbb{A}$$

For example, when $d=2, p_0=q_0=\frac{1}{3}$. The revenues generated by $\mathbb{P},\mathbb{N},\mathbb{A}$ are $1.667>1.500>1.400$ respectively.

Case (2) If prior distribution lies on q-axis, then

$$\mathbb{N} \succ \mathbb{A} \succ \mathbb{P}$$

For example, when $d=2, p_0=0, q_0=\frac{1}{2}$. The revenues generated by $\mathbb{N},\mathbb{A},\mathbb{P}$ are $1.250>1.100>0.938$ respectively.

Then by continuity, we can conclude that $\mathbb{N} \succ \mathbb{P} \succ \mathbb{A}$ is also possible. For example, when $d = 2, p_0 = q_0 = \frac{1}{4}$. The revenues generated by $\mathbb{N}, \mathbb{P}, \mathbb{A}$ are $1.500 > 1.438 > 1.350$ respectively. □

Proof of Lemma 7.2. Using induction method, when $j = 1$, $1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j) = 1 = \left(\sum_{i=1}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1} \left(\sum_{i=1}^{N} \frac{\Pr(v_i)}{v_i}\right)$. Assume the lemma holds for j case,

$$1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j) \geq \left(\sum_{i=1}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1} \left(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\right)$$

then for $j+1$ case, we have $\Pr(v_i|v_{j+1}) < \Pr(v_i|v_j)$ for $i \leq j$ because $\Pr(v_i|v_{j+1})v_{j+1} = \Pr(v_i|v_j)v_j$. Using the fact that $\Pr(v_j|v_j) = \left(1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j)\right)\left(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1} \frac{\Pr(v_j)}{v_j}$

$$\begin{aligned}
1 - \sum_{i=1}^{j} \Pr(v_i|v_{j+1}) &\geq 1 - \sum_{i=1}^{j} \Pr(v_i|v_j) = 1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j) - \Pr(v_j|v_j) \\
&= \left(1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j)\right)\left(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1}\left(\sum_{i=j+1}^{N} \frac{\Pr(v_i)}{v_i}\right) \\
&\geq \left(\sum_{i=1}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1}\left(\sum_{i=j+1}^{N} \frac{\Pr(v_i)}{v_i}\right)
\end{aligned}$$

The lemma holds for $j+1$ case. □

Proof of Proposition 7.3. By Lemma 7.2, $1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j) > 0$, (**P**) condition holds. We can verify that $\Pr(v_k|v_j)v_j < \Pr(v_k|v_{j+1})v_{j+1}$ for $k > j$.

$$\begin{aligned}
\Pr(v_k|v_{j+1})v_{j+1} &= \left(1 - \sum_{i=1}^{j} \Pr(v_i|v_{j+1})\right)\left(\sum_{i=j+1}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1} \frac{\Pr(v_k)v_{j+1}}{v_k} \\
&\geq \left(1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j)\right)\left(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1} \frac{\Pr(v_k)v_{j+1}}{v_k} \\
&> \left(1 - \sum_{i=1}^{j-1} \Pr(v_i|v_j)\right)\left(\sum_{i=j}^{N} \frac{\Pr(v_i)}{v_i}\right)^{-1} \frac{\Pr(v_k)v_j}{v_k} \\
&= \Pr(v_k|v_j)v_j
\end{aligned}$$

(**WM**) condition holds. □

References

Alonso, Ricardo, and Odilon Camara. 2016*a*. "Bayesian Persuasion with Heterogeneous Priors." *Journal of Economic Theory*, 165: 672–706.

Alonso, Ricardo, and Odilon Camara. 2016*b*. "Persuading Voters." *American Economic Review*, 106(11): 3590–3605.

Arieli, Itai, and Yakov Babichenko. 2019. "Private Bayesian Persuasion." *Journal of Economic Theory*, 182: 185–217.

Bardhi, Arjada, and Yingni Guo. 2018. "Modes of Persuasion Toward Unanimous Consent." *Theoretical Economics*, 13(3): 1111–1150.

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

Bergemann, Dirk, and Stephen Morris. 2016*b*. "Information Design, Bayesian Persuasion and Bayes Correlated Equilibrium." *American Economic Review*, 106(5): 586–591.

Bergemann, Dirk, and Stephen Morris. 2019. "Information design: A unified perspective." *Journal of Economic Literature*, 57(1): 44–95.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2015. "The Limits of Price Discrimination." *American Economic Review*, 105(3): 921–957.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2017. "First-Price Auctions With General Information Structures: Implications for Bidding and Revenue." *Econometrica*, 85(1).

Cai, Gaoyang, Qian Jiao, Jingfeng Lu, and Jie Zheng. 2019. "Disclosure Policy in All-Pay Contests with Reserve Prices." *Working Paper.*

Cai, Gaoyang, Zhonghong Kuang, and Jie Zheng. 2019. "Combining Information Design with Mechanism Design in All-pay Auctions." *Working Paper.*

Chan, Jimmy, Seher Gupta, Fei Li, and Yun Wang. 2019. "Pivotal Persuasion." *Journal of Economic Theory*, 180: 178–202.

Chen, Jian, Zhonghong Kuang, and Jie Zheng. 2017*a*. "Bayesian Persuasion in Sequential Tullock Contests." *Working Paper.*

Chen, Jian, Zhonghong Kuang, and Jie Zheng. 2017*b*. "Optimal Joint Design of Timing and Information Disclosure in Tullock Contest." *Working Paper.*

Chen, Jian, Zhonghong Kuang, and Jie Zheng. 2018. "Bayesian Persuasion in All-pay Auction Contests." *Working Paper.*

Chen, Zhuoqiong. 2017. "Information Disclosure in Contests: Private vs. Public Signals." *Working Paper.*

Fu, Qiang, Qian Jiao, and Jingfeng Lu. 2011. "On disclosure policy in contests with stochastic entry." *Public Choice*, 148(3-4): 419–434.

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

Hoshino, Tetsuya. 2017. "Using Strategic Uncertainty to Persuade Multiple Agents." *Working Paper.*

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

Hurley, Terrance M., and Jason F. Shogren. 1998*b*. "Effort levels in a Cournot Nash contest with asymmetric information." *Journal of Public Economics*, 69(2): 195–210.

Kamenica, Emir. 2019. "Bayesian Persuasion and Information Design." *Annual Review of Economics*, forthcoming.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Kartik, Navin, and Weijie Zhong. 2019. "Lemonade from Lemons: Information Design and Adverse Selection." *Working Paper.*

Kovenock, Dan, Florian Morath, and Johannes Mnster. 2015. "Information Sharing in Contests." *Journal of Economics & Management Strategy*, 24(3): 570–596.

Kuang, Zhonghong. 2019. "Bayesian Persuasion in One-sided Asymmetric All-Pay Auction Contests." *Working Paper.*

Laclau, Marie, and Ludovic Renou. 2016. "Public Persuasion." *Working Paper.*

Liu, Zhiyang, and Bo Chen. 2016. "A symmetric two-player all-pay contest with correlated information." *Economics Letters*, 145: 6–10.

Lu, Jingfeng, and Srgio O. Parreiras. 2017. "Monotone equilibrium of two-bidder all-pay auctions Redux." *Games & Economic Behavior*, 104.

Lu, Jingfeng, Hongkun Ma, and Zhe Wang. 2018. "Ranking Disclosure Policies in All-Pay Auctions with Incomplete Information." *Economic Inquiry*, 56(3): 1464–1485.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2019. "On Information Design in Games." *Journal of Political Economy.*

Michaeli, Beatrice. 2017. "Divide and Inform: Rationing Information to Facilitate Persuasion." *The Accounting Review*, 92(5).

Morath, Florian, and Johannes Munster. 2008. "Private versus complete information in auctions." *Economics Letters*, 101(3): 214–216.

Rayo, Luis, and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of Political Economy*, 118(5): 949–987.

Roesler, Anne-Katrin, and Balzs Szentes. 2017. "Buyer-Optimal Learning and Monopoly Pricing." *American Economic Review*, 107(7): 2072–2080.

Serena, Marco. 2016. "Harnessing Beliefs to Stimulate Efforts." *Social Science Electronic Publishing*.

Shimoji, Makoto. 2016. "Rationalizable Persuasion." *Working Paper*.

Siegel, Ron. 2014. "Asymmetric all-pay auctions with interdependent valuations." *Journal of Economic Theory*, 153(C): 684–702.

Song, Yangbo, and Mofei Zhao. 2019. "Designing Noise in Persuasion." *Working Paper*.

Wang, Yun. 2018. "Bayesian Persuasion with Multiple Receivers." *Working Paper*.

Warneryd, Karl. 2003. "Information in conflicts." *Journal of Economic Theory*, 110(1): 121–136.

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

Wu, Zenan, and Jie Zheng. 2017. "Information sharing in private value lottery contest ." *Economics Letters*, 157: 36–40.

Zhang, Jun, and Junjie Zhou. 2016. "Information Disclosure in Contests: A Bayesian Persuasion Approach." *Economic Journal*, 126: 2197–2217.