Punishment and Accuracy Level in Contests

Zhewei Wang

Degree of PhD in Economics

The University of Edinburgh

2010

Contents

Declaration of Own Work

I hereby declare that this thesis is the result of my own work and does not contain any work done in collaboration with others, except as stated in the acknowledgements. References to the work of other people are explicitly stated in the text. I further declare that no part of this thesis has been submitted for any other degree or qualification at any other university.

Zhewei Wang

University of Edinburgh

February 2010

Acknowledgement

First and foremost, I would like to thank my two supervisors, Dr Ahmed Anwar and Professor Jonathan Thomas, for their advice, guidance and encouragement over my four years at Edinburgh. They have always willingly made time to listen to my thoughts and logic, offer constructive comments and read drafts of this thesis. I would also like to thank them for their trust they put in me at the beginning of my graduate studies and for the friendship that I hope will last beyond this work.

Many thanks to Professor Ed Hopkins, Dr. Kohei Kawamura and Professor David Ulph for their encouragement and valuable advice.

Special thanks to Dr Tatiana Kornienko for her very constructive suggestions and helpful comments.

Last but not least, I want to thank my parents, who always support me.

Abstract of the Thesis

In the literature on contests, punishments have received much less attention than prizes. One possible reason is that punishing the bottom player(s) in a contest where all contestants are not allowed to quit, while effective in increasing contestants' total effort, often violates individual rationality constraints. But what will happen in an open contest where all potential contestants can choose whether or not to participate? In chapter 1, we study a model of this type and allow the contest designer to punish the bottom participant according to their performances. We conclude that punishment is often not desirable (optimal punishment is zero) when the contest designer wants to maximize the expected total effort, while punishment is often desirable (optimal punishment is strictly positive) when the contest designer wants to maximize the expected highest individual effort.

In the literature on imperfectly discriminating contests, researchers normally assume that the contest designer has a certain level of accuracy in choosing the winner, which can be represented by the discriminatory power r in the Power Contest Success Function (the Power CSF, proposed by Tullock in 1980). With symmetric contestants, it is well known that increasing accuracy (r) always increases total effort when the pure-strategy equilibrium exists. In chapter 2, we look at the cases where the contestants are heterogeneous in ability. We construct an equilibrium set on $r > 0$, where a unique pure-strategy equilibrium exists for any r below a critical value and a mixed-strategy equilibrium exists for any r above this critical value. We find that if the contestants are sufficiently different in ability, there always exists an optimal accuracy level for the contest designer. Additionally, as we increase the difference in their abilities, the optimal accuracy level decreases. The above conclusions provide an explanation to many phenomena in the real world and may give guidance in some applications.

In chapter 3, we propose the Power Contest Defeat Function (the Power CDF)

which eliminates one player out at a time over successive rounds. We show that the Power CDF has the same good qualities as the Power Contest Success Function (the Power CSF) and is more realistic in some cases. We look at both the Power CSF mechanism (selecting winners in sequence) and the Power CDF mechanism (selecting losers in sequence) and show that punishments increase expected total efforts significantly. More interestingly, we also find that when the contestants' effort levels are different, the Power CDF mechanism is more accurate in finding the correct winner (the one who makes the greatest effort) and the Power CSF mechanism is more accurate in finding the correct loser (the one who makes the smallest effort).

Introduction to the Thesis

Theory of Contests

A contest is defined as a situation where contestants compete against each other to win a prize or multiple prizes. In reality, many types of interaction (in which players expend effort in trying to get ahead of their rivals) have been studied in the field of contest theory both within these specific contexts and at a higher level of abstraction. Such interactions include sports, rent-seeking for rents allocated by a public regulator, political competition, patent races, litigation, relative reward schemes in firms or schools, competition for jobs or promotions, and arm races, military combat or war, etc.

Among the above practical examples which can be seen or studied as contests, there are some conventional types of contests such as sports contests and tournament, but there are some situations which people normally do not take as contests at first sight, such as rent-seeking, arm races, war and litigation. Next, we try to look at those situations from a certain angle. A part of economics (e.g., general equilibrium) studies situations where property rights are defined clearly and agents voluntarily trade rights over goods or produce rights for new goods. This approach produce important insights into the role of markets in resource allocation such as the existence and efficiency of competitive equilibrium, the optimal specialization under international trade, the role of prices in providing information to the agents, etc. However, there are other situations where agents do not trade but rather fight over property rights. In these situations, agents can influence the outcome of the process by means of certain actions such as investment in military power, bribing judges or policy regulators, lobbying politicians, hiring lawyers, advertising, etc. These situations can also be seen and studied as contests besides the conventional types of contests.

In theory, two main branches can be distinguished in the literature on contests. Firstly, perfectly discriminating contests – effort is perfectly observable, the contestants make irreversible efforts and the one who makes the highest effort wins the prize for certain. Technically speaking, a perfectly discriminating contest is very similar to an all pay auction: the prize is like the object auctioned and a contestant's effort is like his bid. Therefore, a lot of techniques of auction theory can be applied to perfectly discriminating contests. Secondly, imperfectly discriminating contests – effort is not perfectly observable, so the contestant who expends the largest effort may not win the prize, but the probability of a particular contestant winning is increasing in his effort and decreasing in the effort of his opponents'. A critical component of a contest in the literature on imperfectly discriminating contests is the Contest Success Function (CSF), which provides each player's probability of winning for any given level of effort.

The literature on contests has developed rapidly from the seminal papers by Tullock (1967, 1980) which study rent-seeking. For example, in the literature on perfectly discriminating contests, Hillman and Riley (1989) study a two-contestant case and Baye, Kovenock and deVries (1996) look at the more than two-contestant case and prove the equilibrium is unique, Baye, Kovenock and deVries (1998) consider contests with a non-linear cost function, Che and Gale (1998) study a situation where contestants have constraints on effort. All the above papers consider cases with complete information. Perfectly discriminating contests with incomplete information have also attracted considerable interest in the literature, a selection of contributions are Glazer and Hassin (1988), Amann and Leininger (1996), Krishna and Morgan (1997), Baye, Kovenock and deVries (1998), Clark and Riis (2000), Moldovanu and Sela (2001), Moldovanu, Sela and Shi (2008). While in the literature on imperfectly discriminating contests, Hillman and Katz (1984), Hillman and Samet (1987), Skaperdas and Gan (1995), Konrad and Schlesinger

(1997) study cases with strict risk-averse contestants. Dixit (1987) analyses a Stackelberg formulation in which one player is able to precommit. Perez-Catrillo and Verdier (1992) explore the implications of a Contest Success Function (CSF) where r is an exogenous discrimination factor, Blavatsky (2004) considers CSF with the possibility of a draw. A large variety of different types of contests have been studied. For example, rent-seeking by Tullock (1980), Hillman (1989) and etc., conflict and appropriation by Garfinkel and Skaperdas (1996), R&D by Loury (1979), patent races by Nti (1997), nonprice competition by Huck et al (2001), the choice between lobbying and litigation by Rubin et al (2001), the periodic contests to host prestigious events like the Olympic Games by Corchon (2000) and Status games by Frank (1985), Frank and Cook (1995), multi-stage contests by Fu and Lu (2007), persuasion (as in advertising, litigation and political campaigning, etc) by Skaperdas and Vaidya (2007). The above is only a very small part of the literature[1].

Aims and Contributions

The focus of my Ph.D. thesis is on finding the optimal level of the contest designer's choice variable, such as punishment and accuracy level, to maximize the contestants' (expected) total effort or highest individual effort with heterogeneous contestants.

In daily life, the expression 'carrots and sticks' refers to a policy of offering a combination of rewards and punishments to induce some desired behaviour. In the literature on contests, the focus has been on the carrots – allocating prizes to the top players, with little attention paid to the sticks – punishing the bottom players.

[1] The theory of contests is a vast literature and a dynamic field, the interested reader can consult the surveys of Nitzan (1994), Konrad (2007) and Corchon (2007) for a general review of this literature.

However, in practice punishing the bottom players can be observed in a variety of circumstances. One important reason that punishments draw much less attention than prizes in the literature is that it is trivial that adding punishments is effective in increasing effort levels ignoring participation constraints. That is for a given group of contestants who can not quit the contest, punishing the bottom player who exerts the lowest effort level, will increase the total effort of the contestants for certain. In this case, punishments should be made as large as possible from the contest designer's point of view. However, adding a punishment, especially when the punishment is large, will often violate individual rationality constraints – a contestant may find that his expected utility in equilibrium is negative. So what would happen if we allow all contestants to freely choose whether or not to participate? The first chapter of this thesis attempts to take both prize and punishment into account in the literature on perfectly discriminating contests. I show that in an "open" contest where all potential contestants can freely choose whether or not to enter, it is optimal to set no punishment if the contest designer wants to maximize the expected total effort in most cases and it is optimal to set an appropriate amount of punishment if the contest designer wants to maximize the expected highest effort in most cases. Intuitively, some low ability players will drop out if a punishment is introduced, so the competition between the participants is likely to become less fierce from this perspective since fewer players are involved. However, the competition between the participants is also likely to become fiercer since the participants want to avoid the punishment. The overall effect of introducing a (small) punishment is that the low ability players drop out, the medium ability players make less effort but the high ability players make more effort. This is the reason for the difference between the two objectives.

In the Power CSF, the parameter r can be interpreted as the cognitive ability or the accuracy level of the contest designer. The greater r is, the higher the cognitive ability the contest designer has and the more accurate the contest designer is. In the previous literature on imperfectly discriminating contests, most researchers take r as an exogenous variable. One important reason might be that researchers focus on the symmetric case – with symmetric contestants, the higher r is, the greater the total effort elicited from the contestants. Therefore, it is widely believed that the contest designer always has an incentive to increase r, so from the contest designer's point of view, r has already been increased to the highest possible level. In the second chapter, we focus on the contest designer's accuracy level in choosing the winner in an imperfectly discriminating contest. We look at a model where the two contestants are heterogeneous in ability and construct an equilibrium set on $r > 0$, where a unique pure-strategy equilibrium exists for any r below a critical value and a mixed-strategy equilibrium exists for any r above this critical value. We find that if the contestants are sufficiently different in ability, there always exists an optimal accuracy level for the contest designer. In these circumstances to maximize total effort, r should be set at the optimal accuracy level instead of the highest possible level. Additionally, as we increase the difference in their abilities, the optimal accuracy level decreases. The above conclusions provide an explanation to many phenomena in the real world and may give guidance in some applications. For example, in recent years with the rapid development of technologies, some people argue that it is time to introduce high-tech into the sports (like tennis, football and basketball, etc) to make the games more accurate. However, our model shows that there is a reason for not using replays and other technologies, more accuracy may reduce effort and therefore reduce skill levels.

The Power CSF has been much used to select the winner or multiple winners in the literature on imperfectly discriminating contests. However, things become more complicated in a technical way when the contest designer wants to identify the bottom players in order to punish them. This is because we need the whole rank of all contestants to identify the bottom players (i.e., the losers) in the Power CSF mechanism. In the third chapter, we propose the Power Contest Defeat Function (Power CDF) which eliminates one player at a time over successive rounds. We show that the Power CDF has the same good qualities as the Power CSF and is more realistic in some cases. For instance, suppose several cities are in a competition to host the Olympic Games, one city will be eliminated in each round until only one city remains, which is the winner. We look at both the Power CSF mechanism (selecting winners in sequence) and the Power CDF mechanism (selecting losers in sequence) and show that punishments increase expected total efforts significantly. More interestingly, we also find that when the contestants' effort levels are different, the Power CDF mechanism is more accurate in finding the correct winner (the one who makes the greatest effort) and the Power CSF mechanism is more accurate in finding the correct loser (the one who makes the smallest effort). In other words, the multi-step mechanism provides more accuracy in finding the correct winner or loser.

1 Chapter 1: Punishment in an Open Contest

1.1 Introduction

In daily life, 'carrots and sticks' refers to a policy of offering a combination of rewards and punishments to induce some desired behaviour. In the literature on contests, focus has been on the carrots – allocating prizes to the top players, with little attention paid to the sticks – punishing the bottom players. However, in practice punishing the bottom players can be observed in a variety of circumstances. One important reason that punishments draw much less attention than prizes in the literature could be that it is trivial that adding punishments is effective in increasing effort levels ignoring participation constraints. That is for a given group of contestants who can not quit the contest, punishing the bottom player who exerts the lowest effort level, will increase the total effort of the contestants for certain. In this case, punishments should be made as large as possible from the contest designer's point of view. However, adding a punishment, especially when the punishment is large, will often violate the individual rationality constraints – a contestant can find that his expected utility in equilibrium is negative. So what would happen if we allow all contestants to freely choose whether or not to participate?

In this paper we assume there is no loss if contestants choose not to enter the contest, so all players can freely choose whether or not to participate; we call this type of contest an *open contest*. There are many practical examples of open contests, such as a photo contest in which all photographers satisfying a certain criterion can choose whether or not to participate; or an essay contest in which students can choose whether or not to participate; or a contest for promotion where all workers can choose to be entered or not, etc.

In the real world, the results of this paper provide the following insights: given an unchanged prize, if the contest designer wants to maximize the total effort from all potential players, no punishment should be set (on the worst performer among all participants) in most cases. However, if the contest designer wants to get the best work from the top contestant, a strictly positive punishment should be set in most cases. Intuitively, introducing a punishment will have two effects. Firstly, the selection effect: some players will drop out (the low ability players), and the competition between the participants will become less fierce since fewer players are involved, consequently, some participants will make less effort (the medium ability players). Secondly, the incentive effect: some participants (the high ability players) will make more effort to avoid the punishment. These two effects occur at the same time. Therefore, by introducing a punishment, the low ability players drop out and the medium ability players make less effort, while only the high ability players make more effort. This is the reason for the difference between the two objectives – punishment is often not desirable when the contest designer wants to maximize the expected total effort, while punishment is often desirable when the contest designer wants to maximize the expected highest individual effort.

According to the conclusion of our model, fewer players will participate if the contest designer increases the magnitude of punishment. This result points out one important cost of adding punishment – decreasing the number of participants. We believe that this form of cost, at least in a long run, is the major cost of adding punishment in a contest. For example, suppose the Economics School (at a university) announces that, from the next semester, half of the students will be failed at the end of each semester according to their exam results. One can be certain that the current students who are studying economics will work much harder since no one wants to be failed. However, in the long run, fewer students would choose to study economics because of the high probability of failure.

In a seminal paper of a large literature on tournaments, Lazear and Rosen

(1981) argue that rank-order tournaments help to solve a moral hazard problem faced by firms. This paper and the following papers have shown that rewarding players based on their work performances provides effective incentives in labor tournaments. Akerlof and Holden (2007) extend the analysis of Lazear and Rosen (1981) to the case with multiple prizes and show that it is generally optimal to give rewards to top performers that are smaller in magnitude than corresponding punishments to poor performers. Their model assumes that all players are homogeneous in abilities and effort and performance is stochastically related – which is the main difference between their works and ours.

The two papers most closely related to the present paper are Moldovanu and Sela (2001) and Moldovanu, Sela and Shi (2008). The first seeks to explain prize structures in tournaments within the framework of private value all-pay auctions. The model we use in this paper is based on that in Moldovanu and Sela (2001) with fixed first prize and linear cost functions – the main difference being that we allow the possibility of punishing the bottom participant in a contest where all potential players can freely choose whether or not to enter, with no cost incurred if they stay out of the contest. In one section of Moldovanu, Sela and Shi (2008), the case where contestants can exit the contest without cost is analyzed, this is very similar to what we analyze in this paper[2]. Interestingly, one important conclusion of this paper seems to contradict the corresponding result in Moldovanu, Sela and Shi (2008), where they find that when the contest designer wants to maximize the expected total effort, punishment is always optimal.[3] We will discuss the difference and the reasons for it later.

Minimum-effort requirement (or entry fee) can also be used to exclude low-

[2] The original version of this paper was produced and circulated in 2007, independently of Moldovanu, Sela and Shi (2008).

[3] In this paper, we also look at a case where the contest designer wants to maximize the expected highest individual effort, which Moldovanu, Sela and Shi (2008) did not cover.

ability players from the contest. Here, we want to emphasize the difference between minimum-effort requirement and punishment. Firstly, with minimum-effort requirement, all participants have to make at least a certain amount of effort to enter the contest, while in our model only the participant with the lowest effort will be punished by suffering a loss. Secondly, it has been proved that with linear cost functions, a contest with a single first prize and an (optimally set) entry fee (i.e., minimum-effort requirement) is total effort maximizing among all feasible mechanisms (that are incentive compatible and individual rational)[4], while in this paper we find that punishment is often not desirable to maximize total effort.

1.2 The Model

There are $k \geq 3$ potential players in a contest with a fixed first prize $V > 0$. [5] All potential players can freely choose whether or not to participate in this contest. Among all contestants who participate in the contest, the player with the highest effort will win the prize, and the player with the lowest effort will be punished by bearing a loss p, where $p \in [0, V]$ which is a choice variable of the contest designer.[6]

Each player, say contestant i simultaneously makes the participation decision and chooses an effort level x_i if he participates without knowing the decision of others. Take an essay contest for example: the students who have to submit their essays by the deadline do not know the number of participants until after the

[4] See Riley and William F. Samuelson (1981) for details.

[5] We assume the prize is simply fixed, i.e., it is indivisible. Moldovanu and Sela (2001) prove that, with linear cost functions, if the contest designer can award several prizes (without punishment), it is optimal to allocate the entire prize sum to a single first prize to maximize the expected total effort. This is a reason why we consider single prize in our model.

[6] Note here we assume if only one person participates in the contest, then he will get the prize and the punishment at the same time.

deadline. An effort x_i causes player i a disutility denoted by $c_i x_i$, where c_i refers to contestant i's (constant) marginal cost of effort, which is private information to himself. c_i is also called the ability parameter of contestant i, a low c_i indicates high ability and vice versa. Ability parameters are drawn independently of each other on the interval $[\underline{s}, \overline{s}]$ (where $\overline{s} > \underline{s} > 0$) according to a distribution function F that is common knowledge. We assume that F has a continuous density function $f = dF/dc > 0$.

Each contestant chooses his effort level in order to maximize his expected utility given the values of the prize and the punishment. The contest designer determines the size of the punishment in order to maximize the expected value of the sum of the efforts (i.e. $\sum_{i=1}^{k} x_i$) or the expected value of the highest individual effort.[7]

Notice that in this model we allow the possibility of punishing the bottom participant in a contest where all potential players can opt not to enter at no cost. Because of the existence of punishment, we have to consider contestants' participation constraints – each player's expected utility in equilibrium is non-negative, an issue that did not arise in Moldovanu and Sela (2001) since without punishment, every contestant's expected utility is always non-negative.

1.2.1 The Objective Function and Entry Decision

Given the commonly known values of the prize V and punishment p, to any contestant who decides to participate in this contest, a contestant with the ability parameter c, solves the following problem by choosing effort level x:

$$\underset{x}{Max}\{V \times \Pr(x \text{ is the highest}) - p \times \Pr(x \text{ is the lowest}) - cx\}.$$

Assume there is an equilibrium such that only contestants with $c \in [\underline{s}, e)$ participate in the contest and each contestant makes effort according to a strictly decreas-

[7] We assume that the contest designer only focuses on effort levels and he dose not get any material benefit directly from the prize or the punishment.

ing differentiable symmetric equilibrium effort function $x = b(c)$ when $c \in [\underline{s}, e)$. All contestants with $c \in [e, \overline{s}]$ do not participate in the contest; in other words, they make zero effort and their expected utility is zero.

The contestant with $c = e$ will just be indifferent between participating in the contest or not. In both situations, he will make zero effort, i.e., $b(e) = 0$ when he enters. By the assumption of the equilibrium effort function, he has the lowest effort of any entrant, there is no point in him putting in a positive effort as he will lose against all other entrants with probability one. He wants to enter the contest with zero effort (which guarantees being punished with probability one) because there is a chance he is the only entrant, in that case, he wins the prize. The marginal contestant's expected utility is[8]:

$$V \times \Pr(\text{effort is the highest}) - p \times \Pr(\text{effort is the lowest}) - e \times 0 = 0$$

$$\Rightarrow F(e) = 1 - (p/V)^{\frac{1}{k-1}}. \tag{1}$$

By looking at the marginal player (the contestant who is just indifferent about entering[9]) whose ability parameter $c = e$ where e satisfies (1), we can see that the larger p is, the smaller $F(e)$ is, and so the smaller e is, i.e., fewer players would participate in the contest. We can see if the contest designer sets the punishment to the same value as the prize, according to (1),

$$1 - F(e) = 1 \Rightarrow F(e) = 0 \Rightarrow e = \underline{s},$$

which means no player will participate in this contest, so total effort is always zero. Intuitively, this is because if it was not true, with the value of the punishment being

[8] Recall we assume if there is only one contestant who participates in this contest, he simultaneously gets the punishment and the prize.

[9] In the equilibrium, all players with $c \geq e$ are indifferent about entering. We make them not participate in our equilibrium. This may seem somewhat arbitrary. But adding a very small miminum effort if a player participates would make higher cost agents ($c \geq e$) strictly prefer to not participate.

equal to the value of the prize ($p = V$), by collecting the punishment from the bottom player and awarding it to the top player, the contest designer can get a positive total effort for free! Therefore, only when $p < V$, do potential entrants exist and make positive effort.

Because we assume $b(c)$ is strictly decreasing in c, this implies that if one contestant's ability parameter is c, the probability of one other contestant's ability parameter being smaller than c is $F(c)$.

$$x = b(c) \Rightarrow c = b^{-1}(x),$$

which means in equilibrium if a participant makes an effort x, by the equilibrium effort function, we can infer his ability by $c = b^{-1}(x)$. Then, given the equilibrium behaviour of other competitors, a player who enters this contest solves the following problem:

$$\underset{x}{Max}\{V \times \underbrace{[1 - F(b^{-1}(x))]^{k-1}}_{\Pr(x \text{ is the highest})} - p \times \underbrace{[F(b^{-1}(x)) + 1 - F(e)]^{k-1}}_{\Pr(x \text{ is the lowest})} - cx\} \tag{2}$$

where $[1 - F(b^{-1}(x))]^{k-1}$ refers to the probability that all other potential contestants make less effort than x and $[F(b^{-1}(x)) + (1 - F(e))]^{k-1}$ refers to the probability that all other contestants either make more effort than x or do not participate in the contest.

1.2.2 The Equilibrium

In Appendix 1A, we prove the following proposition:

Proposition 1 *Given V and $p \in [0, V]$, there exists a symmetric equilibrium: all players with $c \in [e, \overline{s}]$ do not participate in the contest; while all players with $c \in [\underline{s}, e)$ participate in the contest and make effort according to the following bidding function:*

$$b(c) = (k-1)\int_c^e \frac{1}{t}\{V[1 - F(t)]^{k-2} + p[F(t) + 1 - F(e)]^{k-2}\}f(t)dt, \tag{3}$$

where e satisfies (1).

We can split the equilibrium effort function into two parts:

$$b(c) = VA(c) + pB(c),$$

where

$$\begin{aligned} A(c) &= (k-1)\int_c^e \frac{1}{t}[1-F(t)]^{k-2} f(t)dt, \\ B(c) &= (k-1)\int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-2} f(t)dt. \end{aligned}$$

So then

$$\begin{aligned} A'(c) &= -(k-1)\frac{1}{c}[1-F(c)]^{k-2} f(c) < 0, \\ B'(c) &= -(k-1)\frac{1}{c}[F(c)+1-F(e)]^{k-2} f(c) < 0. \end{aligned}$$

Therefore

$$b'(c) = VA'(c) + pB'(c) < 0. \tag{4}$$

Thus, $b(c)$ is strictly decreasing and differentiable when $c \in [\underline{s}, e)$, which is consistent with what we assumed initially. This means the lower the ability parameter is (i.e., the more able the participant is), the more effort the participant is going to make in equilibrium.

1.2.3 Two Effects of Introducing a small Punishment

From (4), we can derive

$$b'(c) = -\frac{(k-1)f(c)}{c}\{V[1-F(c)]^{k-2} + p[F(c)+1-F(e)]^{k-2}\}. \tag{5}$$

We start a situation with no punishment,

$$p = 0 \Rightarrow e = \overline{s} \Rightarrow F(e) = F(\overline{s}) = 1,$$

thus

$$b'(c)|_{p=0} = -\frac{(k-1)f(c)}{c}V[1-F(c)]^{k-2}.$$

When a punishment $p > 0$ is introduced, then

$$(5) \Rightarrow b'(c)|_{p>0} = b'(c)|_{p=0} - \frac{(k-1)f(c)}{c}p[F(c)+1-F(e)]^{k-2} \tag{6}$$

Thus for every $c \in [\underline{s},\, e)$,

$$b'(c)|_{p>0} < b'(c)|_{p=0}. \tag{7}$$

By using (7), we derive the following proposition in Appendix 1A.

Proposition 2 *For two equilibrium effort functions $b(c)|_{p>0}$ (with a positive punishment) and $b(c)|_{p=0}$ (with no punishment): (a) $b(c)|_{p>0}$ is always steeper than $b(c)|_{p=0}$. (b) $b(c)|_{p>0}$ and $b(c)|_{p=0}$ at most cross once. If they cross at point $c = c^*$, then $b(c)|_{p>0} > b(c)|_{p=0}$ for $c < c^*$ and $b(c)|_{p>0} < b(c)|_{p=0}$ for $c > c^*$. If they do not cross, then $b(c)|_{p>0} < b(c)|_{p=0}$ for all c.*

Part (a) shows that when a punishment is introduced, the competition between participants will become relatively more fierce. Part (b) indicates that in an open contest without punishment, if the contest designer introduces a small punishment, $b(c)|_{p>0}$ and $b(c)|_{p=0}$ cross at point $c = c^*$, then the players with $c \in [\underline{s},\, c^*)$ whom we call the high ability players, will make more effort; the players with $c \in (c^*,\, e]$ whom we call the medium ability players[10], will make less effort; and the players with $c \in [e,\, \overline{s}]$ whom we call the low ability players, will drop out (see Figure 1.1).

Intuitively, introducing a small punishment has two main effects. Firstly, the selection effect: by adding a punishment, some players (the low ability players) will drop out, i.e., they decide not to participate. So there are fewer players involved, from this point of view, the competition between the participants will become less

[10]Notice that without punishment, all players with $c \in [\overline{s},\, \underline{s}]$ participate in the contest, while when a small punishment is introduced, only players with $c \in [\overline{s},\, e)$ enter.

fierce. This will cause some players (the medium players) to make less effort since it is easier to outbid the low ability players. Secondly, the incentive effect: by adding a punishment, some players (the high ability players) will make more effort to avoid the punishment. The reasons are, firstly, it costs them less compared with other players by putting in the same amount of extra effort; secondly, each of them will make more effort given that other high ability players making more effort. From this point of view, the competition among the high ability players will become more fierce.

However, when the punishment is too large, all participants will make less effort than before since too many players drop out in this situation (i.e., $b(c)|_{p>0} < b(c)|_{p=0}$ for all c, see Figure 1.2). Therefore, to maximize either the expected total effort or the expected highest individual effort, a large punishment is never optimal.

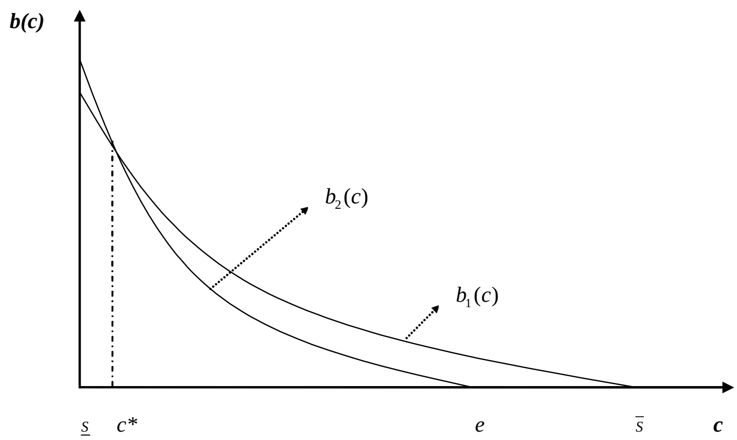

Figure 1.1

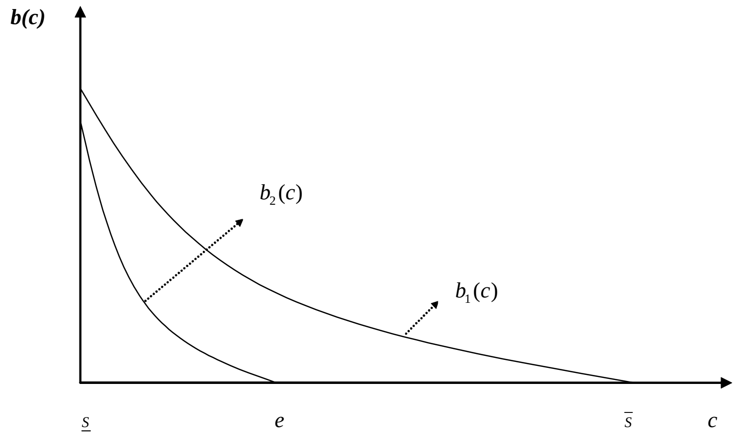

Figure 1.2

Figure 1.1 and Figure 1.2 describe the equilibrium effort functions where $b_1(c)$ with $p = 0$ and $b_2(c)$ with $p > 0$.[11] In Figure 1.1, we can see that with the introduction of a small punishment, the high ability players who have $c \in [\underline{s},\, c^*)$ will make higher effort, the medium ability players who have $c \in (c^*,\, e)$ will make less effort and the low ability players who have $c \in (e,\, \overline{s}]$ will drop out. However, in Figure 1.2, we see that when a large punishment is introduced, all participants make less effort than before.

1.2.4 Two Objectives of the Contest Designer

Maximizing Expected Total Effort In this section, it is assumed that the contest designer's aim is to maximize the expected total effort. Take an essay contest as an example: say a university wants to set an essay contest in some specific field to improve the overall academic level of all students in that field, so it wants all the students to contribute as much as possible, i.e., maximize the total effort.

In the equilibrium we characterize, from the contest designer's point of view,

[11] Note that p is small/large compared with V in Figure 1.1/Figure 1.2.

the average effort of each potential contestant is given by

$$AE = \int_{\underline{s}}^{\overline{s}} b(c) f(c) dc. \tag{8}$$

We have shown that the equilibrium effort function $x = b(c)$ is strictly decreasing for participants with $c \in [\underline{s},\, e)$, and from the contest designer's point of view, $b(c) = 0$ for all contestants with $c \geq e$ who stay out of this contest. There are k potential contestants, so the expected total effort (TE) is

$$TE = k \times AE = k(k-1)R_1 \tag{9}$$

where

$$R_1 = \int_{\underline{s}}^{e} \int_{c}^{e} \frac{1}{t} \{V[1-F(t)]^{k-2} + p[F(t)+1-F(e)]^{k-2}\} f(t) dt f(c) dc. \tag{10}$$

We can see that maximizing TE is equivalent to maximizing R_1. In Appendix 1A, we prove the following proposition by analyzing (10):

Proposition 3 *In an open contest, if the density function $f(c)$ is non-decreasing*[12] *with c on the interval $[\underline{s},\, \overline{s}]$, then it is always optimal to set $p = 0$ in order to maximize the expected total effort.*

When $f(c)$ is non-decreasing (i.e., increasing or staying constant) with c, intuitively, the contest designer expects it is very likely that there are only a few high ability players ($c \in [\underline{s},\, c^*)$) and the majority of the potential players are the low ability players ($c \in (e,\, \overline{s}]$) and the medium players ($c \in (c^*,\, e)$). Therefore, adding a punishment, which will make the low ability players drop out and medium players make less effort, will always decrease the expected total effort.

When $f(c)$ is decreasing with c on the interval $[\underline{s},\, \overline{s}]$, to maximize the expected total effort, the optimal punishment can still be zero (see Case 2 in Appendix 1B for

[12] In this section, we focus on monotone density functions because it is difficult to derive any general conclusions with non-monotone density functions.

an example) or strictly positive (see Case 3 in Appendix 1B for an example). Why is it possible that the optimal punishment can be strictly positive with decreasing density function? It is because the contest designer expects it is very likely to have a lot of high ability players who will make more effort with the introduction of a punishment, so he will put more weight on the high ability players when he maximizes the total effort, which makes it possible to make punishment desirable in this case.

The above result is consistent with what we observe in the real world quite often – in most open contests there is no punishment. Therefore in our essay contest example, the university should announce the student with the top quality essay will be allocated with a prize and no one will be punished at all due to their poor quality essays.

In one section of Moldovanu, Sela and Shi (2008), a very similar situation has been analyzed – they prove that the optimal punishment is always strictly positive when the contest designer wants to maximize the expected total effort. This contradicts our corresponding result: when $f(c)$ is non-decreasing in c, the optimal punishment is always zero and even when $f(c)$ is decreasing in c, the optimal punishment can still be zero (see Case 2 in Appendix 1B for an example). Next we analyze and discuss this difference.

The main difference between our model (which follows Moldovanu and Sela (2001)) and Moldovanu, Sela and Shi (2008) is on the distribution of the marginal cost of effort. The marginal cost of effort in our model (i.e., the ability parameter c), is replaced with $1/a$ in theirs where a is distributed on the interval $[0, 1]$ according to a distribution function Π that is common knowledge which has a continuous density function $\pi = d\Pi/da > 0$. Since in their model a is distributed on the interval $[0, 1]$, the marginal cost of effort $1/a$ is distributed on the interval $[1, +\infty)$. Because they assume $\pi > 0$, the density function in terms of $1/a$ must be always positive on the interval $[1, +\infty)$. Therefore, the real difference between our model and theirs is that the marginal cost of effort (represented by c in our model and $1/a$ in theirs) is distributed on $[\underline{s}, \overline{s}]$ in our model and $[1, +\infty)$ in theirs.

Theoretically, if we set $\underline{s} = 1$ and $\overline{s} \to +\infty$, our model will be the same as the one in Moldovanu, Sela and Shi (2008). In Case 1 in Appendix 1B, we look at the case where $\overline{s} \to +\infty$ and we derive the same result as theirs – the optimal punishment is strictly positive. However, this result does not contradict our conclusion. It is because in this paper we asserts that the optimal punishment is always zero when $f(c)$ is non-decreasing with c on the interval $[\underline{s}, \overline{s}]$, while when $\overline{s} \to +\infty$, the density function can not be always non-decreasing – $f(c)$ must be decreasing as $c \to +\infty$ given $\int_1^{+\infty} f(c)dc = 1$.

Therefore, their conclusion is based on the case where the marginal cost of effort is distributed on $[1, +\infty)$ with density function being always strictly positive. Intuitively, this assumption always allows a possibility that a group of very low ability players exists ($1/a \to +\infty$ as $a \to 0$), so starting from a situation without punishment, introducing an appropriate punishment will make these very low ability players drop out and the high ability players make more effort. Because these very low ability players only make little effort in the situation without any punishment, the selection effect is smaller than the incentive effect, therefore, the expected total effort must increase after the introduction of the punishment. This is why the optimal punishment is always strictly positive in their model. However, we argue that Moldovanu, Sela and Shi (2008) analyze a limiting case of our model that is when $\underline{s} = 1$ and $\overline{s} \to +\infty$. In this sense our model is more general.

Maximizing Expected Highest Individual Effort Instead of maximizing the expected total effort, in some cases the contest designer may only want to elicit the highest individual effort, i.e., the best work from all of the potential contestants. For example, for some reason a university may only need the best essay from its students, with all essays of a lesser quality than the best being of no interest to the university. In this section, we focus on the case when the contest designer wants to maximize the expected highest individual effort.

Rank the contestants' ability parameter as follows: $c_1 < c_2 < ... < c_k$, so c_1 is the most able player. First consider $G_1(c)$, which is the distribution function of c_1. The probability that all k potential players' ability parameters are bigger than c, i.e., all potential players are less able than type c, is

$$(1 - F(c))^k,$$

then the probability that at least one contestant is more able than c is

$$1 - (1 - F(c))^k.$$

Therefore,

$$G_1(c) = \Pr(c_1 < c) = 1 - (1 - F(c))^k,$$

hence, the probability density function of c_1 is

$$g_1(c) = G_1'(c) = k(1 - F(c))^{k-1} f(c).$$

Therefore, we write the expected highest individual effort

$$E[b(c_1)] = \int_{\underline{s}}^{\overline{s}} g_1(c) b(c) dc = k(k-1) R_2,$$

where

$$R_2 = \int_{\underline{s}}^{e} \int_{c}^{e} \frac{1}{t} \{V[1 - F(t)]^{k-2} + p[F(t) + 1 - F(e)]^{k-2}\} f(t)(1 - F(c))^{k-1} f(c) dt dc. \tag{11}$$

We can see that maximizing $E[b(c_1)]$ is equivalent to maximizing R_2. In Appendix 1A, we prove the following proposition by analyzing (11):

Proposition 4 *In an open contest with k players, there always exists a number of contestants k^* such that for $k > k^*$, for any form of the distribution density function $f(c)$, the optimal punishment is always strictly positive when the contest designer's aim is to maximize the expected highest individual effort.*

We have a relatively strong condition in Proposition 3: k must be large enough ($k > k^*$) to guarantee the optimal punishment being strictly positive. This is because we allow the density function $f(c)$ to take any form. In fact, for many forms of $f(c)$, we can relax this restriction on k. For example, in Appendix 1A, we also prove the following proposition:

Proposition 5 *In an open contest where abilities are drawn from a uniform distribution on $[\underline{s}, \overline{s}]$, i.e., $f(c) = 1/(\underline{s} - \overline{s})$, when $\overline{s}/\underline{s} \geq 1.47$, for any $k \geq 3$ it is always optimal to set a strictly positive punishment if the contest designer's aim is to maximize the expected highest individual effort.*

Note that when $\overline{s}/\underline{s} < 1.47$, the optimal punishment can be zero or positive. We want to emphasize that $\overline{s}/\underline{s} \geq 1.47$ refers to the case when the highest possible able player is at least 1.47 times as efficient as the least possible able player, which covers most common cases in reality. Therefore, when abilities are drawn from a uniform distribution on $[\overline{s}, \underline{s}]$, the optimal punishment is zero when the contest designer wants to maximize the expected total effort, while as long as $\overline{s}/\underline{s} \geq 1.47$, the optimal punishment is strictly positive when the contest designer wants to maximize the expected highest individual effort.

When the contest designer only wants to maximize the expected highest individual effort, there is a good chance that the most able player, type c_1, is more able than type c^*, i.e., $c_1 < c^*$, so the contest designer will put more weight on the high ability players. More intuitively, when we are maximizing the expected highest individual effort, the contest designer cares about the high ability players much more than other players because the top player of this contest is the one who has the highest ability. Introducing a punishment will increase the high ability players' effort which is very likely to increase the highest individual effort. That is why the contest designer often has an incentive to set a strictly positive amount of punishment to maximize the expected highest individual effort.

Therefore, in our essay contest example, when the university only wants to get the best essay, the university should set a strictly positive punishment in the contest, i.e., it should announce that the participant whose essay is considered to be the best will get a prize, while the participant whose essay is considered to be the worst will be punished in some way.

1.3 Concluding Comments

We study an open contest where all potential contestants (who have private information about their abilities) can choose whether or not to enter and allow the contest designer to punish the bottom participant according to their performance. We conclude that punishment is often not desirable (optimal punishment is zero) when the contest designer wants to maximize the expected total effort, while punishment is often desirable (optimal punishment is strictly positive) when the contest designer wants to maximize the expected highest individual effort, i.e., when the contest designer only cares about the performance of the top contestant. In many circumstances there is a trade-off between maximizing the expected total effort and maximizing the expected highest individual effort. Hence, depending on the objectives of the contest designer, punishment may be part of the (optimal) answer.

1.4 Appendix 1A

Proof of Proposition 1

To maximize (2), we write out the first-order condition as follows:

$$-(k-1)f(b^{-1}(x))\frac{db^{-1}(x)}{dx}\{V[1-F(b^{-1}(x))]^{k-2}+p[F(b^{-1}(x))+1-F(e)]^{k-2}\}-c=0,$$

then we derive

$$(\frac{db^{-1}(x)}{dx})^{-1}=-\frac{1}{c}(k-1)f(b^{-1}(x))\{V[1-F(b^{-1}(x))]^{k-2}+p[F(b^{-1}(x))+1-F(e)]^{k-2}\}. \quad (12)$$

In equilibrium,

$$b(c)=x\Rightarrow c=b^{-1}(x) \quad (13)$$

$$\Rightarrow\frac{db(c)}{dc}=\frac{dx}{db^{-1}(x)}=(\frac{db^{-1}(x)}{dx})^{-1}. \quad (14)$$

Substituting (12) and (13) into (14),

$$\frac{db(c)}{dc}=-\frac{1}{c}(k-1)\{V[1-F(c)]^{k-2}-p[F(c)+1-F(e)]^{k-2}\}f(c).$$

Given the boundary condition: $b(e)=0$, we derive

$$b(c)=b(c)-b(e)=-\int_c^e\frac{db(t)}{dt}dt=VA(c)+pB(c)$$

where

$$\begin{aligned}A(c) &= (k-1)\int_c^e\frac{1}{t}[1-F(t)]^{k-2}f(t)dt\\ B(c) &= (k-1)\int_c^e\frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dt.\end{aligned}$$

In the main text, we have shown that

$$b'(c)=VA'(c)+pB'(c)<0.$$

Therefore, as we assumed, $b(c)$ is strictly decreasing and differentiable when $c\in[\underline{s},\ e)$. Assuming except contestant i, all other players with $c\in[\underline{s},\ e)$ make effort according to $b(c)$, we need to show that for any type c of contestant i, the effort $b(c)$ maximizes the expected utility of that type. The necessary first order condition is clearly satisfied since this is how we assumed $b(c)$ to start with. Let

$$\pi(x,c) = V[1 - F(b^{-1}(x))]^{k-1} - p[F(b^{-1}(x)) + 1 - F(e)]^{k-1} - cx$$

be the expected utility of player i with type c that makes effort x. We will show that the derivative $\pi_x(x,c)$ is nonnegative if x is smaller than $b(c)$ and nonpositive if x is larger than $b(c)$. As $\pi(x,c)$ is continuous in x, this implies that $\pi(x,c)$ is maximized at $x = b(c)$. Note that

$$\begin{aligned}\pi_x(x,c) &= -(k-1)f(b^{-1}(x))\frac{db^{-1}(x)}{dx}\{V[1 - F(b^{-1}(x))]^{k-2} \\ &\quad +p[F(b^{-1}(x)) + 1 - F(e)]^{k-2}\} - c.\end{aligned}$$

Let $x < b(c)$, and let $\widehat{c}$ be the type who is supposed to bid x, that is $b(\widehat{c}) = x < b(c)$. Note that $\widehat{c} > c$ because b is strictly decreasing. Differentiating $\pi_x(x,c)$ with respect to c yields $\pi_{xc}(x,c) = -1 < 0$. That is, the function $\pi_x(x,.)$ is decreasing in c. Since $\widehat{c} > c$, we obtain $\pi_x(x,c) \geq \pi_x(x,\widehat{c})$.

Since $x = b(\widehat{c})$ we obtain $\pi_x(x,\widehat{c}) = 0$ by the first-order condition, and therefore that $\pi_x(x,c) \geq 0$ for every $x < b(c)$. A similar argument shows that $\pi_x(x,c) \leq 0$ for every $x > b(c)$.

Proof of Proposition 2

Because $b'(c) < 0$ and $b'(c)|_{p>0} < b'(c)|_{p=0}$, it follows that (a) $b(c)|_{p>0}$ is always steeper than $b(c)|_{p=0}$. To prove (b), suppose that $b'(c)|_{p>0}$ and $b'(c)|_{p=0}$ cross more than once, we can choose two points where they cross, say point c_1^* and c_2^* where $c_1^* < c_2^*$. Then there must exist a point $c^m \in (c_1^*,\, c_2^*)$ where at point c^m, the two equilibrium effort functions have the same slope, i.e., $b'(c^m)|_{p>0} = b'(c^m)|_{p=0}$, which contradicts (7). Therefore, $b(c)|_{p>0}$ and $b(c)|_{p=0}$ can not cross more than once. If they cross once at point c^*, $b'(c)|_{p>0} < b'(c)|_{p=0} < 0$, so $b(c)|_{p>0} > b(c)|_{p=0}$ for $c < c^*$ and $b(c)|_{p>0} < b(c)|_{p=0}$ for $c > c^*$. If they do not cross, since $b'(c)|_{p>0} < b'(c)|_{p=0}$ and $b(e)|_{p>0} = 0$, $b(e)|_{p=0} > 0$, it is trivial to see that $b(c)|_{p>0} < b(c)|_{p=0}$ for all c.

1.4.1 Proof of Proposition 3

Recall that

$$R_1 = \int_{\underline{s}}^{e} \underbrace{\int_{c}^{e} \frac{1}{t}\{V[1-F(t)]^{k-2} + p[F(t)+1-F(e)]^{k-2}\}f(t)dtf(c)dc}_{Z}.$$

We can derive

$$\begin{aligned}\frac{dZ}{dp} &= \frac{de}{dp}(\frac{V}{e})[(1-F(e))^{k-2}+p]f(e)f(c) \\ &+\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dtf(c) \\ &-(k-2)\frac{de}{dp}f(e)p\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)dtf(c).\end{aligned}$$

Thus,

$$\frac{dR_1}{dp} = \frac{de}{dp} \times Z|_{c=e} + \int_{\underline{s}}^{e}\frac{dZ}{dp}dc = \int_{\underline{s}}^{e}\frac{dZ}{dp}dc.$$

That is:

$$\begin{aligned}\frac{dR_1}{dp} &= \underbrace{\frac{de}{dp}(\frac{V}{e})[(1-F(e))^{k-2}+p]f(e)\int_{\underline{s}}^{e}f(c)dc}_{(\alpha)} \\ &\underbrace{+\int_{\underline{s}}^{e}\{\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dt\}f(c)dc}_{(\beta)} \\ &\underbrace{-(k-2)\frac{de}{dp}f(e)p\int_{\underline{s}}^{e}\{\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)dt\}f(c)dc.}_{(\gamma)} \quad (15)\end{aligned}$$

Our aim is to prove that when $f(x)$ is non-decreasing in x, $\frac{dR_1}{dp} < 0$ for $0 \le p < 1$, thus the optimal punishment is zero.

$$
\begin{aligned}
(\mathbf{??}) \quad &\Rightarrow \quad p = (1 - F(e))^{k-1} V && (16)\\
&\Rightarrow \quad \frac{de}{dp} = \frac{-1}{(k-1) f(e) (1 - F(e))^{k-2} V}. && (17)
\end{aligned}
$$

Substituting (16) and (17) into (α), we can derive

$$
(\alpha) = -\frac{1}{(k-1)e}[2 - F(e)]F(e). \tag{18}
$$

In (β), in a two-dimensional world, the area where $c \le t \le e$ and $\underline{s} \le c \le e$ can be expressed as the area where $\underline{s} \le t \le e$ and $\underline{s} \le c \le t$, so we write

$$
\begin{aligned}
(\beta) &= \int_{\underline{s}}^{e} \int_{\underline{s}}^{t} \frac{1}{t} [F(t) + 1 - F(e)]^{k-2} f(t) f(c) dc dt \\
&= \int_{\underline{s}}^{e} \frac{F(t)}{t} [F(t) + 1 - F(e)]^{k-2} f(t) dt. && (19)
\end{aligned}
$$

Because f is non-decreasing, we can infer that $f'(t) \ge 0$. Let

$$
\begin{aligned}
g(t) &= F(t)/t, \; h(t) = t f(t) - F(t) \\
&\Rightarrow h'(t) = t f'(t) \ge 0 \Rightarrow h(t) > 0 \text{ as } h(\underline{s}) = \underline{s} f(\underline{s}) > 0 \\
&\Rightarrow \frac{dg(t)}{dt} = \frac{t f(t) - F(t)}{t^2} = \frac{h(t)}{t^2} > 0.
\end{aligned}
$$

So $g(t) = F(t)/t$ is increasing with t; then for all $t < e$,

$$
\frac{F(t)}{t} < \frac{F(e)}{e}. \tag{20}
$$

Substituting (20) into (19)

$$
\begin{aligned}
(\beta) &= \int_{\underline{s}}^{e} \frac{F(t)}{t} [F(t) + (1 - F(e))]^{k-2} f(t) dt \\
&< \int_{\underline{s}}^{e} \frac{F(e)}{e} [F(t) + (1 - F(e))]^{k-2} f(t) dt \\
&= \frac{F(e)}{e(k-1)} [1 - (1 - F(e))^{k-1}] \le \frac{F(e)}{e(k-1)},
\end{aligned}
$$

i.e.,

$$(\beta) < \frac{F(e)}{e(k-1)}. \tag{21}$$

By using similar method, we can derive

$$\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}[F(t)+(1-F(e))]^{k-3}f(t)f(c)dtdc < \frac{F(e)}{e(k-2)}. \tag{22}$$

Substituting (16), (17) and (22) into (γ), we derive

$$(\gamma) < \frac{F(e)(1-F(e))}{(k-1)e}. \tag{23}$$

From (18), (21) and (23), we obtain

$$\frac{dR_1}{dp} = (\alpha)+(\beta)+(\gamma) < -\frac{(2-F(e))F(e)}{(k-1)e} + \frac{F(e)}{(k-1)e} + \frac{F(e)(1-F(e))}{(k-1)e} = 0.$$

Therefore, when $f(x)$ is non-decreasing in x, $dR_1/dp < 0$ for all $p \in [0,1)$.

Proof of Proposition 4

Recall that

$$R_2 = \int_{\underline{s}}^{e} \underbrace{\int_{c}^{e}\frac{V}{t}([1-F(t)]^{k-2} + p[F(t)+1-F(e)]^{k-2})f(t)(1-F(c))^{k-1}f(c)dt}_{X}dc.$$

We get

$$\begin{aligned}\frac{dX}{dp} &= \frac{de}{dp}(\frac{V}{e})[(1-F(e))^{k-2}+p]f(e)(1-F(c))^{k-1}f(c)\\ &+\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)(1-F(c))^{k-1}f(c)dt\\ &+(-\frac{de}{dp})p(k-2)f(e)\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)(1-F(c))^{k-1}f(c)dt.\end{aligned}$$

Consequently, we write

$$\begin{aligned}
\frac{dR_2}{dp} &= \frac{de}{dp}X|_{c=e} + \int_{\underline{s}}^{e}\frac{dX}{dp}dc = \int_{\underline{s}}^{e}\frac{dX}{dp}dc \\
&= \underbrace{\frac{de}{dp}(\frac{V}{e})[(1-F(e))^{k-2}+p]f(e)\int_{\underline{s}}^{e}(1-F(c))^{k-1}f(c)dc}_{(a)} \\
&\quad \underbrace{+\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e))]^{k-2}f(t)(1-F(c))^{k-1}f(c)dtdc}_{(b)} \\
&\quad \underbrace{+(-\frac{de}{dp})p(k-2)f(e)\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)(1-F(c))^{k-1}f(c)dtdc.}_{(c)}
\end{aligned}$$

Substituting (16) and (17) into (a) and (c), we can derive

$$\begin{aligned}
(a) &= -\frac{[2-F(e)][1-(1-F(e))^k]}{k(k-1)e} \\
(c) &= \frac{(k-2)(1-F(e))}{(k-1)V}\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)(1-F(c))^{k-1}f(c)dtdc.
\end{aligned}$$

When $p=0$, $e=\overline{s}$ and $F(e)=F(\overline{s})=1$, we derive

$$\frac{dR_2}{dp}|_{p=0} = (a)+(b)+(c)$$

$$= \int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}}\frac{1}{t}F(t)^{k-2}f(t)(1-F(c))^{k-1}f(c)dtdc - \frac{1}{k(k-1)\overline{s}}.$$

Therefore, $dR_2/dp|_{p=0} > 0$ if and only if

$$\int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}}\frac{1}{t}F(t)^{k-2}f(t)(1-F(c))^{k-1}f(c)dtdc > \frac{1}{k(k-1)\overline{s}}. \tag{24}$$

In a two-dimensional world, the area where $c \le t \le \overline{s}$ and $\underline{s} \le c \le \overline{s}$ is equivalent to the area where $\underline{s} \le c \le t$ and $\underline{s} \le t \le \overline{s}$, so mathematically we derive:

$$\begin{aligned}
LHS\ of\ (24) &= \int_{\underline{s}}^{\overline{s}}\int_{\underline{s}}^{t}[\frac{1}{t}F(t)^{k-2}f(t)(1-F(c))^{k-1}f(c)dc]dt \\
&= \frac{1}{k}\int_{\underline{s}}^{\overline{s}}\frac{1}{t}F(t)^{k-2}f(t)[1-(1-F(t))^k]dt.
\end{aligned}$$

So (24) holds if and only if

$$\int_{\underline{s}}^{\overline{s}} \frac{1}{t}[1-(1-F(t))^k]F(t)^{k-2}f(t)dt > \frac{1}{\overline{s}(k-1)}. \tag{25}$$

Because we can derive

$$RHS\ of\ (25) = \frac{1}{\overline{s}(k-1)} = \int_{\underline{s}}^{\overline{s}} \frac{1}{\overline{s}} F(t)^{k-2}f(t)dt.$$

Thus, (25) holds if and only if

$$\begin{aligned}
& \int_{\underline{s}}^{\overline{s}} \frac{1}{t}[1-(1-F(t))^k]F(t)^{k-2}f(t)dt - \frac{1}{(k-1)\overline{s}} \\
= & \int_{\underline{s}}^{\overline{s}} \frac{1}{t}[1-(1-F(t))^k]F(t)^{k-2}f(t)dt - \int_{\underline{s}}^{\overline{s}} \frac{1}{\overline{s}} F(t)^{k-2}f(t)dt \\
= & \int_{\underline{s}}^{\overline{s}} (1-\frac{t}{\overline{s}})\frac{1}{t}F(t)^{k-2}f(t)dt - \int_{\underline{s}}^{\overline{s}} (1-F(t))^k\frac{1}{t}F(t)^{k-2}f(t)dt > 0. \quad (26)
\end{aligned}$$

We claim that (26) always holds when $k > k^*$ where k^* satisfies[13]

$$\int_{\underline{s}}^{\overline{s}} [(1-\frac{t}{\overline{s}}) - (1-F(t))^{k^*}]\frac{1}{t}F(t)^{k^*-2}f(t)dt = 0. \tag{27}$$

By comparing $(1-\frac{t}{\overline{s}})\frac{1}{t}F(t)^{k-2}f(t)$ and $(1-F(t))^k\frac{1}{t}F(t)^{k-2}f(t)$, it can be seen that, for all k where $k > k^*$, when k gets larger, $(1-\frac{t}{\overline{s}})$ becomes larger compared with $(1-F(t))^k$ and $(1-\frac{t}{\overline{s}})\frac{1}{t}F(t)^{k-2}f(t)$ becomes relatively larger compared with $(1-F(t))^k\frac{1}{t}F(t)^{k-2}f(t)$. Therefore, when $k > k^*$, (26) always holds.

To sum up, the logic of the whole proof is that: when $k > k^*$,

$$\begin{aligned}
(26) \quad & \Rightarrow \int_{\underline{s}}^{\overline{s}} \frac{1}{t}[1-(1-F(t))^k]F(t)^{k-2}f(t)dt > \frac{1}{(k-1)\overline{s}} \\
& \Rightarrow \int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}} \frac{1}{t}F(t)^{k-2}f(t)(1-F(c))^{k-1}f(c)dtdc - \frac{1}{k(k-1)\overline{s}} > 0 \\
& \Rightarrow \frac{dR_2}{dp}|_{p=0} = (a)+(b)+(c) > 0
\end{aligned}$$

[13]Note here we assume $k^* \in R$, while $k \in N$.

Proof of Proposition 5

Substituting $F(t) = \dfrac{t-\underline{s}}{\overline{s}-\underline{s}}$ and $f(t) = \dfrac{1}{\overline{s}-\underline{s}}$ into (11), we have:

$$R_2 = \frac{1}{(\overline{s}-\underline{s})^{2k-1}} \int_{\underline{s}}^{e} [\underbrace{\int_c^e \frac{V}{t}(\overline{s}-t)^{k-2} + \frac{p}{t}(t-\underline{s}+\overline{s}-e)^{k-2}(\overline{s}-c)^{k-1}dt}_{Y}]dc.$$

It can be derived that

$$\begin{aligned}\frac{dY}{dp} &= \frac{de}{dp}(\frac{V}{e})[(\overline{s}-e)^{k-2} + p(\overline{s}-\underline{s})^{k-2}](\overline{s}-c)^{k-1} \\ &\quad + \int_c^e \frac{1}{t}(t+\overline{s}-\underline{s}-e)^{k-2}(\overline{s}-c)^{k-1}dt \\ &\quad + (-\frac{de}{dp})p(k-2)\int_c^e \frac{1}{t}(t+\overline{s}-\underline{s}-e)^{k-3}(\overline{s}-c)^{k-1}dt.\end{aligned}$$

Then we derive

$$\begin{aligned}\frac{dR_2}{dp} &= \frac{1}{(\overline{s}-\underline{s})^{2k-1}}\{\frac{de}{dp}Y|_{c=e} + \int_{\underline{s}}^{e}\frac{dY}{dp}dc\} = \frac{1}{(\overline{s}-\underline{s})^{2k-1}}\int_{\underline{s}}^{e}\frac{dY}{dp}dc \\ &= \frac{1}{(\overline{s}-\underline{s})^{2k-1}}\{\underbrace{\frac{de}{dp}(\frac{V}{e})[(\overline{s}-e)^{k-2} + p(\overline{s}-\underline{s})^{k-2}]\int_{\underline{s}}^{e}(\overline{s}-c)^{k-1}dc}_{(a_1)} \\ &\quad \underbrace{+\int_{\underline{s}}^{e}\int_c^e \frac{1}{t}(t+\overline{s}-\underline{s}-e)^{k-2}(\overline{s}-c)^{k-1}dtdc}_{(b_1)} \\ &\quad \underbrace{+(-\frac{de}{dp})p(k-2)\int_{\underline{s}}^{e}\int_c^e \frac{1}{t}(t+\overline{s}-\underline{s}-e)^{k-3}(\overline{s}-c)^{k-1}dtdc}_{(c_1)}\}.\end{aligned}$$

Using (1) and $F(e) = \dfrac{e-\underline{s}}{\overline{s}-\underline{s}}$, we derive

$$p = (\frac{\overline{s}-e}{\overline{s}-\underline{s}})^{k-1}V \tag{28}$$

$$\frac{de}{dp} = \frac{-(\overline{s}-\underline{s})^{k-1}}{(k-1)(\overline{s}-e)^{k-2}V} \tag{29}$$

Substituting (28) and (29) into (a_1) and (c_1), we have

$$(a_1) = -\frac{(\overline{s}-\underline{s})^{k-2}(2\overline{s}-\underline{s}-e)}{(k-1)e}\int_{\underline{s}}^{e}(\overline{s}-c)^{k-1}dc$$

$$(c_1) = \frac{(k-2)(\overline{s}-e)}{(k-1)V}\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}(t+\overline{s}-\underline{s}-e)^{k-3}(\overline{s}-c)^{k-1}dtdc.$$

When $p=0$, $e=\overline{s}$, we derive

$$\begin{aligned}
(a_1) &= -\frac{(\overline{s}-\underline{s})^{k-1}}{(k-1)\overline{s}}\int_{\underline{s}}^{\overline{s}}(\overline{s}-c)^{k-1}dc = -\frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}} \\
(b_1) &= \int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}}\frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1}dtdc \\
(c_1) &= 0\times\int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}}\frac{1}{t}(t-\underline{s})^{k-3}(\overline{s}-c)^{k-1}dtdc = 0.
\end{aligned}$$

Thus we write

$$\frac{dR_2}{dp}|_{p=0} = \frac{1}{(\overline{s}-\underline{s})^{2k-1}}\{\int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}}\frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1}dtdc - \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}}\}.$$

Therefore, $dR_2/dp|_{p=0} > 0$ if and only if

$$\int_{\underline{s}}^{\overline{s}}\int_{c}^{\overline{s}}\frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1}dtdc > \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}}. \tag{30}$$

So the optimal punishment is strictly positive when (30) holds. In a two-dimensional world, the area where $c \le t \le \overline{s}$ and $\underline{s} \le c \le \overline{s}$ is equivalent to the area where $\underline{s} \le c \le t$ and $\underline{s} \le t \le \overline{s}$, so mathematically we derive

$$\begin{aligned}
LHS\ of\ (30) &= \int_{\underline{s}}^{\overline{s}}\int_{\underline{s}}^{t}\frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1}dcdt \\
&= \frac{1}{k}\int_{\underline{s}}^{\overline{s}}\frac{1}{t}(t-\underline{s})^{k-2}[(\overline{s}-\underline{s})^{k}-(\overline{s}-t)^{k}]dt.
\end{aligned}$$

Let $v = \dfrac{\overline{s}-t}{\overline{s}-\underline{s}}$, then $t = \overline{s}-(\overline{s}-\underline{s})v$, so $dt = -(\overline{s}-\underline{s})dv$; since $\underline{s} \le t \le \overline{s}$, we can derive

$$0 \le \frac{\overline{s}-t}{\overline{s}-\underline{s}} \le 1,$$

i.e., $0 \le v \le 1$. Notice that $v=1$ when $t=\underline{s}$ and $v=0$ when $t=\overline{s}$. Then we write

$$\begin{aligned}
LHS\ of\ (30) &= \frac{1}{k}\int_{\underline{s}}^{\overline{s}}\frac{1}{t}(t-\underline{s})^{k-2}[(\overline{s}-\underline{s})^{k}-(\overline{s}-t)^{k}]dt \\
&= \frac{(\overline{s}-\underline{s})^{2k-1}}{k}\int_{0}^{1}\frac{(1-v)^{k-2}(1-v^{k})}{\overline{s}-v(\overline{s}-\underline{s})}dv.
\end{aligned}$$

We claim that for all $k \geq 3$, (30) holds if

$$\int_0^1 (1-v)^{k-2}\{\frac{(1-v^3)}{1-v(1-(\underline{s}/\overline{s}))} - 1\}dv > 0. \tag{31}$$

This is true because

$$\begin{aligned}
(31) \quad &\Rightarrow \quad \frac{1}{\overline{s}}\int_0^1 (1-v)^{k-2}\{\frac{(1-v^3)}{1-v(1-(\underline{s}/\overline{s}))} - 1\}dv > 0 \\
&\Rightarrow \quad \int_0^1 (1-v)^{k-2}\{\frac{(1-v^3)}{\overline{s}-v(\overline{s}-\underline{s})} - \frac{1}{\overline{s}}\}dv > 0 \\
&\Rightarrow \quad \int_0^1 (1-v)^{k-2}\{\frac{(1-v^k)}{\overline{s}-v(\overline{s}-\underline{s})} - \frac{1}{\overline{s}}\}dv > 0 \text{ (since } k \geq 3) \\
&\Rightarrow \quad \int_0^1 \frac{(1-v)^{k-2}(1-v^k)}{\overline{s}-v(\overline{s}-\underline{s})}dv > \frac{1}{\overline{s}}\int_0^1 (1-v)^{k-2}dv \\
&\Rightarrow \quad \int_0^1 \frac{(1-v)^{k-2}(1-v^k)}{\overline{s}-v(\overline{s}-\underline{s})}dv \geq \frac{1}{(k-1)\overline{s}} \\
&\Rightarrow \quad \frac{(\overline{s}-\underline{s})^{2k-1}}{k}\int_0^1 \frac{(1-v)^{k-2}(1-v^k)}{\overline{s}-v(\overline{s}-\underline{s})}dv \geq \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}}.
\end{aligned}$$

Let

$$j(v) = \frac{(1-v^3)}{1-(1-(\underline{s}/\overline{s}))v} - 1,$$

then the $LHS\ of$ (31) becomes

$$\int_0^1 (1-v)^{k-2} j(v)dv. \tag{32}$$

We can see that the sign of $(1-v)^{k-2}j(v)$ is determined by $j(v)$ as $0 \leq v \leq 1$. According to the definition of integration, graphically, the value of (32) is equal to the area between the v axis and the curve $(1-v)^{k-2}j(v)$ on the interval $[0,1]$. By looking at the expression for $j(v)$, we can prove that when $0 \leq v \leq 1$,

$$j(v)\left\{\begin{array}{c} > \\ = \\ < \end{array}\right\} 0 \text{ when } v \left\{\begin{array}{c} < \\ = \\ > \end{array}\right\} \sqrt{1-(\underline{s}/\overline{s})}.$$

As k increases in (31), more relative weight is put on $j(v)$ for lower values of v, and as $j(v)$ crosses the axis only once (when $v = \sqrt{1-(\underline{s}/\overline{s})}$), a positive integral cannot become negative. Therefore, we conclude that if

$$\int_0^1 (1-v)^{k-2} j(v) dv|_{k=3} > 0,$$

then for all $k \geq 3$

$$\int_0^1 (1-v)^{k-2} j(v) dv > 0.$$

Therefore, we write:

$$\int_0^1 (1-v)^{k-2} j(v) dv|_{k=3} = \int_0^1 (1-v)\{\frac{(1-v^3)}{1-[1-(\underline{s}/\overline{s})]v} - 1\} dv$$

$$= (1/12)[(\underline{s}/\overline{s})-1]^{-5}\{-3+28(\underline{s}/\overline{s})-30(\underline{s}/\overline{s})^2-6(\underline{s}/\overline{s})^3+17(\underline{s}/\overline{s})^4$$

$$-6(\underline{s}/\overline{s})^5+36(\underline{s}/\overline{s})^2\ln(\underline{s}/\overline{s})-36(\underline{s}/\overline{s})^3\ln(\underline{s}/\overline{s})+12(\underline{s}/\overline{s})^4\ln(\underline{s}/\overline{s})\}.$$

By analyzing the above equation, it is easy to check when $0 < (\underline{s}/\overline{s}) \leq 0.68$, in other words, when $(\overline{s}/\underline{s}) \geq 1.47$,

$$\int_0^1 (1-v)^{k-2} j(v) dv|_{k=3} > 0,$$

so when $(\overline{s}/\underline{s}) \geq 1.47$,

$$\int_0^1 (1-v)^{k-2} j(v) dv > 0 \text{ for } k \geq 3.$$

Thus, the optimal punishment is strictly positive in these cases. To sum up, the logic of the whole proof is that:

$$\int_0^1 (1-v)^{k-2} j(v) dv|_{k=3} > 0$$

$$\Rightarrow \int_0^1 (1-v)^{k-2} j(v) ds > 0 \text{ for } k \geq 3$$

$$\Rightarrow \int_{\underline{s}}^{\overline{s}} [\int_c^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-2}](\overline{s}-c)^{k-1} dt dc > \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}}$$

$$\Rightarrow \frac{dR_2}{dp}|_{p=0} = \frac{1}{(\overline{s}-\underline{s})^{2k-1}}\{(a_1)+(b_1)+(c_1)\} > 0.$$

1.5 Appendix 1B

Case 1 *(Proposition 6 of Moldovanu, Sela and Shi, 2008) Let the support of* F *is* $[a, +\infty)$, *i.e.,* $\overline{s} \to +\infty$. *Substituting* (18) *and* (19) *into* (15), *we derive*

$$\frac{dR_1}{dp}|_{p=0} = -\frac{1}{(k-1)\overline{s}} + \int_{\underline{s}}^{\overline{s}} \frac{1}{t}[F(t)]^{k-1} f(t)dt.$$

As $\overline{s} \to +\infty$, *we write*

$$\frac{dR_1}{dp}|_{p=0} = \int_{\underline{s}}^{+\infty} \frac{1}{t}[F(t)]^{k-1} dF(t) > 0.$$

Therefore, the optimal punishment is strictly positive when $\overline{s} \to +\infty$.

Case 2 *Let* $V = 1$, $k = 3$, $\underline{s} = 1$, $\overline{s} = 2$, $F(c) = -c^2 + 4c - 3$ *and* $f(c) = F'(c) = 4 - 2c$. *We can see the density function is strictly decreasing with* c *on the interval* $[1, 2]$. *According to* (9) *the expected total effort can be calculated as follows:*

$$TE = \int_1^e \int_c^e \frac{6}{t}\{(t-2)^2 + p[4t - t^2 - 3 + (e-2)^2]\}(4-2t)(4-2c)dtdc. \quad (33)$$

Substituting $F(e) = -e^2 + 4e - 3$ *into* (1), *we derive*

$$e = 2 - p^{\frac{1}{4}}. \quad (34)$$

Substituting (34) *into* (33), *we write*

$$dTE/dp = -172.4 - 49p^{\frac{3}{4}} - 12p - 3.6p^{\frac{5}{4}} + 108(2-p^{\frac{1}{2}})\ln(2-p^{\frac{1}{4}}) + 45/(2-p^{\frac{1}{4}}) + 96p^{\frac{1}{4}} + 96p^{\frac{1}{2}}.$$

It is easy to show that $dTE/dp < 0$ *for all* $p \in [0, 1)$. *Thus, the optimal punishment can still be zero when the density function decreases in* c.

Case 3 *Let* $V = 1$, $k = 3$, $\underline{s} = 1$, $\overline{s} = 11$, $F(c) = -0.01c^2 + 0.22c - 0.21$ *and* $f(c) = F'(c) = 0.22 - 0.02c$. *We can see the density function is strictly decreasing with* c *on the interval* $[1, 11]$. *According to* (9) *the expected total effort can be calculated as follows:*

$$TE = \int_1^e \int_c^e \frac{6}{t}\{(0.1t - 1.1)^2 + p[1 + 0.22(t-e) - 0.01(t^2 - e^2)]\}(0.22 - 0.02t)(0.22 - 0.02c)dtdc. \quad (35)$$

Substituting $F(e) = -0.01e^2 + 0.22e - 0.21$ *into* (1), *we derive*

$$e = 11 - 10p^{\frac{1}{4}}. \tag{36}$$

Substituting (36) *into* (35), *we can express* TE *as a function of* p. *It can be found that* TE *is maximized when* $p = 0.011$. *Thus, the optimal punishment can be strictly positive when the density function decreases in* c.

2 Chapter 2: The Optimal Accuracy level in Asymmetric Contests

2.1 Introduction

The International Table Tennis Federation (ITTF) changed the points scoring system for international matches from first to 21 to first to 11 in 2000. One reason for doing this is to reduce the accuracy level of the matches. The rationale for this is simple. The domination of China meant that there was little incentive for the other teams. Reducing the accuracy level increases the chance that a team other than China will win, thus inducing more effort from the other teams. This increased competition could in turn result in greater effort from the Chinese team. The greater all round effort in preparing for matches results in higher quality and more entertaining matches. However, with very little accuracy there is also little incentive for effort. In the extreme case we simply have a lottery where the probability of winning is independent of effort. We investigate the optimal level of accuracy in the face of these trade-offs.

A contest is a situation in which players compete against each other by making irreversible effort, often for a prize or multiple prizes[14]. In theory, two branches can be distinguished in the literature. Firstly, perfectly discriminating contests (very similar to first-prize all-pay auctions) – the effort is perfectly observable, the contestants make irreversible efforts and the player who makes the highest effort wins the prize for certain. Secondly, imperfectly discriminating contests – the effort is not perfectly observable, so the contestant who expends the largest effort

[14] Typical examples are various types of rent–seeking contests: competition among firms to win a monopoly rent allocated by a public regulator, litigation, beauty contests, patent races, research and development (R&D), political competition, competition to higher ranks, competition for jobs or promotions, arm races and sports events, etc.

may not win the prize, but the probability of a particular contestant winning is increasing in his effort and decreasing in the effort of his opponents'. A critical component of a contest in the literature on imperfectly discriminating contests is the Contest Success Function (CSF), which provides each player's probability of winning for any given level of effort. In this paper, we use the Power CSF which was proposed by Tullock[15] in 1980:

$$\begin{aligned} p_i(e) &= \frac{e_i^r}{\sum_{j=1}^n e_j^r} \qquad \text{if } \; max\{e_1, ...e_n\} > 0; \\ p_i(e) &= 1/n \qquad \text{otherwise,} \end{aligned} \tag{37}$$

where $e = (e_1, e_2, ..., .e_n)$ denotes a vector of efforts for the n players. There are n contestants, e_i refers to the effort contestant i makes and $p_i(e)$ refers to the probability with which contestant i wins the contest. Mathematically, the parameter r is the elasticity of the odds of winning for contestant i.[16] It is often interpreted as indicating returns to scale in efforts: if $r > 1$ ($r < 1, r = 1$), then returns to scale are increasing (decreasing, constant). So r can be seen as the discriminatory power of the Power CSF (note the Power CSF becomes perfectly discriminating as $r \to +\infty$).

More intuitively, we could interpret r as the cognitive ability or the accuracy level of the contest designer. The greater r is, the higher cognitive ability the contest designer has and the more accurate the contest designer is. In practice,

[15] The Power CSF has also been called the Tullock CSF. Skaperdas (1996) derives the Power CSF from easily interpretable axioms – the Power CSF is the **only** continuous success functional form which satisfies all the following axioms: (1) Imperfect Discrimination, (2) Monotonicity, (3) Anonymity, (4) Consistency, (5) Independence and (6) Homogeneity. This justifies the popularity of the Power CSF and is one important reason why we use it in this paper. For a review of the general properties of Contest Success Functions, see Skaperdas(1996).

[16] Note here $r = \dfrac{d\ln[p_i(e)/(1-p_i(e))]}{d\ln e_i}$ i.e., r measures percentage changes of $p_i(e)/(1-p_i(e))$ in response to a one percent change of e_i.

the contest designer often has ways to increase or decrease the accuracy level. For example, in an essay contest, the marker can spend more time on marking each essay and do more comparisons between essays, etc. By doing this, r will be greater. On the contrary, if he decreases the time on marking each essay, like he only gives himself few minutes to mark each essay, r must be lower.

In the previous literature on imperfectly discriminating contests, most researchers take r as an exogenous variable. One important reason might be that most researchers focus on the symmetric case – with symmetric contestants, the higher r is, the greater the total effort elicited from the contestants. Therefore, it is widely believed that the contest designer always has an incentive to increase r, so from the contest designer's point of view, r has already been increased to the highest possible level.

The two papers most closely related to the present research are Michaels (1988) and Nti (2004). Michaels (1988) finds that $r = 2$ is optimal for the symmetric valuations case when the contest designer can choose r to maximize total effort[17]. Nti (2004) looks at the asymmetric valuations case and finds that there exists an optimal r in the region where a pure-strategy equilibrium exists. The main differences between this paper and Nti's are two: firstly, contestants are heterogeneous in ability in our model while contestants have different valuations (to the same prize) in Nti's model, which leads to different conclusions and applications; secondly, except for the case where a unique pure-strategy equilibrium exists (for any $r \in (0, \overline{r}]$), we also look at the case where a mixed-strategy equilibrium exists (for any $r \in (\overline{r}, +\infty)$) while Nti does not.[18]

[17] With symmetric contestants, the maximum value of r is 2 to ensure a pure-strategy equilibrium exists.

[18] The parameter $\overline{r}$ will be explained later in detail. You will find that in this paper a unique pure-strategy equilibrium exists for any $r \in (0, \overline{r}]$, and at least one mixed-stategy equilibria exist for any $r \in [\ \overline{r}, +\infty)$.

In this paper, we build a model with two contestants who are heterogeneous in ability and show that when the contestants are sufficiently different in their abilities, there exists an optimal accuracy level, which maximizes total effort. So from the contest designer's point of view, it is not true that the total effort always increases with the accuracy level. Hence, if the contest designer's objective is to maximize the total effort, r should not be increased to the highest possible level. Additionally, we also find that the optimal accuracy level would decrease when the two contestants become more different in ability.

If the contestants are sufficient different in ability, when the contest designer increases the accuracy level, intuitively there are two effects. Firstly, it increases every contestant's incentive to invest more effort since the competition is more fierce as r increases. Secondly, the less able contestant realizes he is less likely to win as r increases – it decreases his incentive to invest more effort. So when r increases to some certain level, the less able contestant starts to invest less effort, in turn this will again cause the more able contestant to invest less effort. Therefore, after r reaches this level, total effort falls and this level of r is the optimal accuracy level. When the difference in their abilities becomes larger, the less able contestant becomes even weaker compared with his opponent, he will start to invest less effort earlier and so total effort falls earlier than in the previous case. Hence the optimal level of r decreases. Therefore, as a contest designer who wants to maximize total effort, he should set r at the optimal accuracy level instead of the highest possible level and the optimal accuracy level decreases when contestants become more different in ability.

Besides the table tennis example we gave at the beginning of this paper, our model can help us with explaining or understanding some other phenomena in the real world. In 1976, FIFA changed the tie-breaking method from replay to penalty-shootout. We believe that a shootout is more like a random draw compared with replays. So why did FIFA make football matches less accurate? In dive meets, the judges are not allowed to see the slow motion while the audience can see it right after each dive. Though all the judges are professional, no one can argue that letting them see the slow motion would help them in increasing their accuracy level. Then why not? Besides that, there are other examples which may look stranger. In a lot of sports, every athlete only has one shot, like the balance beam. Why not give them more chances considering it is very easy to have some bad miss on the balance beam even for a world-class athlete. In a long-jump match, every athlete only has three chances, why not give them more chances given it is almost costless to let them jump several more times. There are of course many reasons why these restrictions are a good idea. Our analysis highlights one, the contest designer may wish to reduce accuracy to increase total effort in some situations.

2.2 The Model

There are two risk-neutral contestants involved in a contest with a single prize V. Each contestant, say contestant i, has a linear cost function, $cost_i = c_i e_i$ $(i = 1, 2)$ where e_i refers to contestant i's effort level and c_i is contestant i's ability parameter. It can be seen that the more able the contestant is, the lower his ability parameter is. Contestant 1 has an ability parameter c_1 and contestant 2 has an ability parameter $c_2 = c * c_1$. Assume that contestant 1 is more able than contestant 2, so $c > 1$. The probability of winning is defined by the Power Contest Success Function (the Power CSF, see (37)). All the parameters (r, c_1 and c_2) are common knowledge.

Each contestant's aim is to maximize his expected profit, $\pi_i = P_i V - c_i e_i$ where P_i represents contestant i's probability of winning. Each contestant can freely choose to stay active in the contest by making a positive effort or stay inactive by making zero effort. The contest designer's aim is to maximize the expected total effort.[19]

The timing of the model is as follows. First the contest designer chooses an accuracy level r which he can commit to. Then, given that the values of c_1, c_2 and r are public information, contestants make their efforts. Finally, the contest designer chooses the winner by applying the Power CSF with the r he chooses in the first stage.

2.2.1 The Pure-strategy Equilibrium

The Equilibrium Total Effort In this section we assume that the contest designer can costlessly choose his accuracy level from the set $(0, \overline{r}]$, where the value $\overline{r}$ is determined to ensure that a pure-strategy equilibrium exists which we will explain in detail later.

Given contestant 1 makes an effort e_1 and contestant 2 makes an effort e_2, the expected profits for each contestant are:

$$\pi_1 = \frac{e_1^r}{e_1^r + e_2^r}V - c_1 e_1, \quad \pi_2 = \frac{e_2^r}{e_1^r + e_2^r}V - c_2 e_2. \tag{38}$$

Each contestant chooses his effort level to maximize his expected profit. In Appendix 2A, we show that subject to participation constraints, there always exists a unique pure-strategy equilibrium where contestant 1 makes an effort e_1^* and contestant 2 makes an effort e_2^*:

$$e_1^* = \frac{c_1^r c_2^r r V}{c_1(c_1^r + c_2^r)^2}, \quad e_2^* = \frac{c_1^r c_2^r r V}{c_2(c_1^r + c_2^r)^2}. \tag{39}$$

[19] In many cases, we may have aims other than maximizing total effort, like the contest designer may care about the social welfare instead of the effort level, he may care about waste from duplication like in an R&D contests where increased effort entails wasteful duplication. In some other cases the contest designer may just want to identify the most able agent as accurately as possible and does not really care about the effort level. However, in this paper we focus our attention on the cases where the contest designer's aim is to maximize the total effort.

Substituting $c_2 = c * c_1$ into (39), we have:

$$e_1^* = \frac{c^r rV}{c_1(1+c^r)^2}, \quad e_2^* = \frac{c^{r-1} rV}{c_1(1+c^r)^2}. \tag{40}$$

The contest designer's aim is to maximize the total effort:

$$TE = e_1^* + e_2^* = \frac{(c^{r-1} + c^r) rV}{c_1(1+c^r)^2}. \tag{41}$$

To make sure of the existence of the pure-strategy equilibrium, all contestants' participation constraints must hold, i.e., each contestant's expected profit should be greater than or equal to zero in equilibrium:

$$\pi_1|_{e_1=e_1^*;\ e_2=e_2^*} \geq 0, \ \ \pi_2|_{e_1=e_1^*;\ e_2=e_2^*} \geq 0. \tag{42}$$

In Appendix 2B we show that to make (42) hold, $r \in (0, \overline{r}]$ where $\overline{r}$ satisfies $c^{\overline{r}} = 1/(\overline{r} - 1)$. We also find that as c increases from 1 to $+\infty$, $\overline{r}$ decreases from 2 to 1. The following proposition summarizes the results so far.

Proposition 6 *For any $r \in (0, \overline{r}]$, there always exists a unique pure-strategy Nash equilibrium, where $\overline{r}$ satisfies $c^{\overline{r}} = 1/(\overline{r} - 1)$. As c increases from 1 to $+\infty$, $\overline{r}$ decreases from 2 to 1. The equilibrium effort levels are:*

$$e_1^* = \frac{c^r rV}{c_1(1+c^r)^2}, \ e_2^* = \frac{c^{r-1} rV}{c_1(1+c^r)^2}, \ TE = \frac{(c^{r-1} + c^r) rV}{c_1(1+c^r)^2}.$$

The Optimal Accuracy Level Hence

$$\frac{dTE}{dr} = \frac{de_1^*}{dr} + \frac{de_2^*}{dr}, \tag{43}$$

where

$$\frac{de_1^*}{dr} = \frac{c^r(1 + c^r - (c^r - 1) r \log c) V}{c_1(1+c^r)^3}, \tag{44}$$

$$\frac{de_2^*}{dr} = \frac{c^{r-1}(1 + c^r - (c^r - 1) r \log c) V}{c_1(1+c^r)^3}. \tag{45}$$

In Appendix 2C, we show that

$$\frac{dTE}{dr}(\frac{de_1^*}{dr}, \frac{de_2^*}{dr}) > 0 \text{ when } r < \widehat{r},$$
$$\frac{dTE}{dr}(\frac{de_1^*}{dr}, \frac{de_2^*}{dr}) = 0 \text{ when } r = \widehat{r},$$
$$\frac{dTE}{dr}(\frac{de_1^*}{dr}, \frac{de_2^*}{dr}) < 0 \text{ when } r > \widehat{r},$$

where $\widehat{r}$ satisfies

$$1 + c^{\widehat{r}} - (c^{\widehat{r}} - 1) \log c^{\widehat{r}} = 0 \ (\Rightarrow \widehat{r} \approx \log_c^{4.68}). \tag{46}$$

When $c \geq 4.68$, $\widehat{r} \leq 1 < \overline{r}$ since (46), it must be the case that $\widehat{r}$ locates in the region $(0, \overline{r}]$ where a unique pure-strategy equilibrium exists; when $c < 4.68$, $\widehat{r} > 1$ since (46), to ensure that $\widehat{r}$ locates in the region $(0, \overline{r}]$, we need $c^{\widehat{r}} < 1/(\widehat{r} - 1)$. In Appendix 2D, we show that

$$c^{\widehat{r}} < 1/(\widehat{r} - 1) \Rightarrow c > 3.5665.$$

Therefore, when $c > 3.5665$, $\widehat{r}$ locates in the region $(0, \overline{r}]$, so TE and e_i^* increase with r when $r < \widehat{r}$ and decrease with r when $r > \widehat{r}$; when $c \leq 3.5665$, TE always increases with r in the region $(0, \overline{r}]$.

From (46), it can be seen that $\widehat{r}$ decreases when c increases, which means the bigger the difference in the contestants' abilities, the sooner dTE/dr and de_i^*/dr would turn to negative when r increases. In Appendix 2D we also show that TE and e_i^* approach zero as r goes to zero. The following proposition summarizes these results.

Proposition 7 *(1) Total effort (TE) and individual effort (e_i^*) approach zero as r goes to zero: $TE \to 0$ and $e_i^* \to 0$ as $r \to 0$. (2) When $c > 3.5665$, in the region where $r \in (0, \overline{r}]$, the optimal accuracy level is $\widehat{r}$ where $\widehat{r}$ satisfies (46), and $\widehat{r}$ decreases when c increases. (3) When $c \leq 3.5665$, in the region where $r \in (0, \overline{r}]$, the optimal accuracy level is $\overline{r}$.*

From the above results, we can see when the accuracy level (r) is extremely low ($r \to 0$), no player is willing to put in any effort because the winning probability is always close to a half no matter how much effort he makes. In the region where $r \in (0, \widehat{r})$, with an increase of the accuracy level, both players would make more effort in equilibrium. This is because when r is small, both contestants' efforts are low, as r increases, by making one additional unit of effort, the marginal revenue is big compared with the marginal cost[20]. However, when r increases above $\widehat{r}$, both contestants have already made relatively large efforts, as r increases, by making one additional unit of effort, the marginal revenue is small compared with the marginal cost, and this is true especially for the less able contestant who has a bigger marginal cost. Hence the less able contestant will decrease his effort at the point $r = \widehat{r}$ where the marginal revenue equals marginal cost. The more able contestant will find that it is optimal to decrease effort at the same point given his opponent does so, so total effort (TE) falls after $\widehat{r}$. Therefore, $\widehat{r}$ is optimal for the contest designer who wants to maximize the total effort in the region where $r \in (0, \overline{r}]$. Intuitively, we can think of the less able contestant as first rising to and then shrinking from competition as r rises. So $\widehat{r}$ is the point where the less able contestant starts to shrink facing the competition. The result "the optimal accuracy level $\widehat{r}$ decreases as c increases" indicates that when the contestants become more different in ability, the optimal accuracy level decreases. Intuitively, this shows that when the less able contestant becomes even weaker compared with his opponent, he will shrink earlier than before. For example, when c goes to infinity which means the less able contestant is extremely weak compared with his opponent, $\widehat{r}$ approaches zero.

[20] Recall that the marginal cost is always c_1 for the more able contestant and $c_2 = c * c_1$ for the less able contestant where $c > 1$.

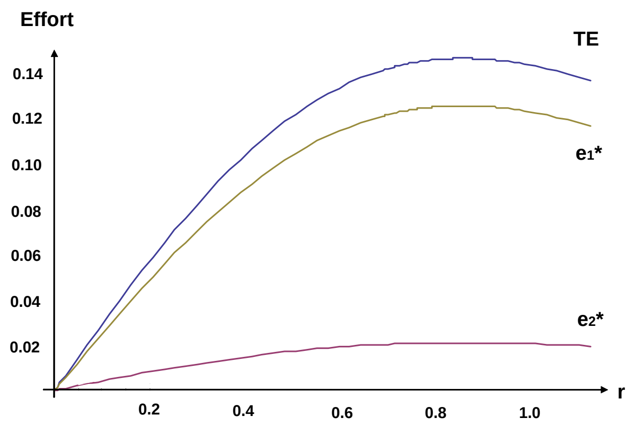

Figure 2

For any $r \in (0, \overline{r}]$, in equilibrium:

$$P_2 = \frac{e_2^{*r}}{e_1^{*r} + e_2^{*r}} = \frac{(\frac{e_2^*}{e_1^*})^r}{1 + (\frac{e_2^*}{e_1^*})^r} = \frac{(\frac{1}{c})^r}{1 + (\frac{1}{c})^r} = \frac{1}{1 + c^r}, \tag{47}$$

we can see that the probability of winning for the less able contestant always decreases with an increase of r, so decreasing the accuracy level actually increases the probability of winning for the less able contestant in a fair way. So when $r > \widehat{r}$, by reducing the accuracy level r, the less able contestant is encouraged, hence he increases his effort which causes the more able contestant to invest more effort as well.

Figure 2 depicts the case with $c_1 = 1$, $c_2 = 6$ where $r \in (0, \overline{r}]$. We can calculate that $\overline{r} \approx 1.132$ and $\widehat{r} \approx 0.861$. So when $r < 1.132$, there always exists a pure-strategy equilibrium where both contestants make positive effort. It is clear to see that total effort and each individual effort (TE, e_1^* and e_1^*) are maximized when r reaches the optimal level $\widehat{r}$. After that, TE, e_1^* and e_2^* are decreasing in r.

Other Findings in the Pure-strategy Equilibrium How will each contestant's expected profit change as r increases from 0 to $\bar{r}$. In Appendix 2E, we show that:

$$\pi_1^* = \frac{c^r(1+c^r-r)V}{(1+c^r)^2},\ \pi_2^* = \frac{(1+c^r-c^r r)V}{(1+c^r)^2}, \tag{48}$$

$$\frac{\partial \pi_1^*}{\partial r} = \frac{-c^r V[1+c^r-(1-r+c^r(1+r))\log c]}{(1+c^r)^3}, \tag{49}$$

$$\frac{\partial \pi_2^*}{\partial r} = \frac{-c^r V[1+c^r+(1+r+c^r(1-r))\log c]}{(1+c^r)^3}. \tag{50}$$

In Appendix 2E, we find the following results by analyzing the above equations.

(1) $\pi_1^* - \pi_2^* > 0$, which means the more able contestant always makes higher profit than the less able one.

(2) $\partial(\pi_1^* - \pi_2^*)/\partial r > 0$, which shows the difference in the two contestants' profits gets greater when r increases.

(3) $\partial \pi_2^*/\partial r < 0$, which indicates that the less able contestant's expected profit decreases from $V/2$ to 0 when r increases from 0 to $\bar{r}$.

(4) $\partial \pi_1^*/\partial r$ can be positive, negative or zero. When c is small, $\partial \pi_1^*/\partial r \leq 0$ for all r; when c is medium, $\partial \pi_1^*/\partial r \leq 0$ for small r and $\partial \pi_1^*/\partial r > 0$ for big r; when c is big, $\partial \pi_1^*/\partial r > 0$ for all r.

Intuitively, the less able contestant always prefers a smaller r, but the more able contestant's preference depends on the difference between the contestants' abilities – when the difference is small, the more able contestant's profit decreases as r increases; when the difference is big, the more able contestant's profit increases as r increases.

So far, we have analyzed the equilibrium effort levels when r changes given the ability parameters c_1 and c_2 unchanged. Next, we are going to analyze the case when c_1 or c_2 changes given that r is exogenous. From (39) and (40), it can be easily derived that:

$$\frac{\partial e_1^*}{\partial c_1} < 0,\quad \frac{\partial e_1^*}{\partial c_2} < 0,\quad \frac{\partial e_2^*}{\partial c_1} > 0,\quad \frac{\partial e_2^*}{\partial c_2} < 0.$$

It is not a surprise that $\partial e_1^*/\partial c_1 < 0$ and $\partial e_2^*/\partial c_2 < 0$, which means when a contestant is more able, he makes more effort. But $\partial e_1^*/\partial c_2 < 0$ and $\partial e_2^*/\partial c_1 > 0$ are very interesting. This tells us that the more able contestant always increases (decreases) his effort level when his opponent gets more (less) able but the less able contestant always decreases (increases) his effort level when his opponent gets more (less) able. So intuitively, we can think of this as an advantage of the more able contestant – he will always increase his effort when his opponent gets more able (as long as he is more able than his opponent), while the less able contestant always decrease his effort when his opponent gets more able.

2.2.2 Mixed-strategy Equilibrium when $r \in [\overline{r}, +\infty)$

In the previous discussion, we focus on the case when $r \in (0, \overline{r}]$ where a unique pure-strategy equilibrium exists. In this section, we are going to look at the case when $r \in [\overline{r}, +\infty)$ where the pure-strategy equilibrium does not exist, but one or multiple mixed-strategy equilibria exist.

The All-pay Auction Equilibrium when $r \in [2, +\infty)$ In the existing literature, there are only a few papers (Baye (1994) and Alcalde and Dahm (2007)) discussing the situation when $r \geq \overline{r}$ where one or multiple mixed-strategy equilibria exist. Alcalde and Dahm (2007) have proved that in Tullock's Rent-Seeking Game[21], when $r \geq 2$ there exists a mixed-strategy equilibrium which they call the all-pay auction equilibrium[22]. Alcade and Dahm derive an all-pay auction equilibrium[23] with asymmetric valuations $V_1 \geq V_2 = V$ from the Tullock's Rent-Seeking Game, i.e., the case with symmetric valuations $V_1 = V_2 = V$. We derive the

[21] Tullock's Rent-seeking Game in their paper is the same with our model in this paper with $c_1 = c_2 = 1$. Notice that $\overline{r} = 2$ in Tullock's Rent-seeking Game.

[22] We also refer this kind of equilibrium as all-pay auction equilibrium in this paper.

[23] See Alcade and Dahm (2007) for details.

following proposition by using similar techniques.

Proposition 8 *Let C^A be a two-player contest with asymmetric abilities $c_1 < c_2$ with $c_2/c_1 = c$. Let C^S be the same contest with symmetric ability $c_1^s = c_2^s = c_2$. If $\mu^* = (\mu_1^*, \mu_2^*)$ is a symmetric Nash equilibrium strategy profile to C^S in which the rent is completely dissipated in expectation, then the following strategy profile $v^* = (v_1^*, v_2^*)$ constitutes a Nash equilibrium to C^A: contestant 1 uses strategy $v_1^* = \mu_1^*$, contestant 2's strategy v_2^* is that he uses μ_2^* with probability $1/c$ and stays inactive (i.e. makes zero effort) with probability $1 - (1/c)$.*

Proof. Alcade and Dahm (2007) have proved that: when $c_1 = c_2 = 1$, for any $r \in [2, +\infty)$, there exists an all-pay auction equilibrium where the expected effort of each contestant is $V/2$. Based on the this, we can easily show that when $c_1 = c_2$, for any $r \in [2, +\infty)$, there exists an all-pay auction equilibrium where the expected effort of each contestant is $V/(2c_2)$.

In C^S, the symmetry of the game assures that on average each player wins half of the time, each contestant's expected revenue is:

$$E\Pi(\mu^*) = (V/2) - c_2 V/(2c_2) = 0.$$

So in C^S the rent is completely dissipated in expectation. Because $v_1^* = \mu_1^*$ and $c_2^s = c_2$ in C^S and C^A, any pure strategy for contestant 2 in C^A yields the same as that in C^S and contestant 2 obtains $E\Pi_2(v^*) = 0$. So in C^A, he is willing to stay inactive (i.e. make zero effort) with probability $1 - (1/c)$.

In C^S given contestant 2's strategy μ_2^*, contestant 1's pure strategy b_1^s in the support of μ^* maximizes

$$\begin{aligned} E\Pi_1(b_1^s, \mu^*) &= E[\Pr\{\text{player 1 wins} \mid b_1^s, \mu^*\}]V - c_1^s b_1^s \\ &= E[\Pr\{\text{player 1 wins} \mid b_1^s, \mu^*\}]V - c_2 b_1^s. \end{aligned} \tag{51}$$

In C^A, given contestant 2's strategy v_2^* (which is that he uses μ_2^* with probability $(1/c)$ and stays inactive with probability $1-(1/c)$), contestant 1's pure strategy b_1 maximizes

$$\begin{aligned} E\Pi_1(b_1, v^*) &= \frac{1}{c}E[\Pr\{\text{player 1 wins} \mid b_1, \mu^*\}]V + (1-\frac{1}{c})V - c_1 b_1 \\ &= \frac{1}{c}E[\Pr\{\text{player 1 wins} \mid b_1, \mu^*\}]V - c_1 b_1 + (1-\frac{1}{c})V \\ &= \frac{1}{c}(E[\Pr\{\text{player 1 wins} \mid b_1, \mu^*\}]V - c_2 b_1) + (1-\frac{1}{c})V. \quad (52) \end{aligned}$$

We complete the proof by noticing that the b_1^s which maximizes (51) must maximize (52), so contestant 1 does not have an incentive to deviate from v_1^*. ■

The expected total effort in the all-pay auction equilibrium is

$$E[TE]^{All-pay} = \frac{V}{2c_2} + (\frac{1}{c})\frac{V}{2c_2} = \frac{(c+1)}{2cc_2}V. \quad (53)$$

The Mixed-strategy Equilibrium when $r \in [\overline{r}, 2]$ So far, we have found a unique pure-strategy equilibrium for any $r \in (0, \overline{r}]$ and also constructed an all-pay auction equilibrium for any $r \in [2, +\infty)$. Next, we construct a mixed-strategy equilibrium for any $r \in [\overline{r}, 2]$ as follows. Given r fixed, contestant 1 always participates and makes an effort x^* (notice here contestant 1 adopts a pure strategy), while contestant 2 stays inactive with probability $1-p$ and makes an effort y^* with probability p (notice here contestant 2 adopts a mixed strategy). Note contestant 2 is indifferent between participating with effort y^* and staying inactive since we assume his expected profit is always zero.

In Appendix 2F, we prove the existence of this mixed-strategy equilibrium and we summarize the results into the following proposition.

Proposition 9 *For any* $r \in [\overline{r}, 2]$*, there always exists a mixed-strategy equilibrium in which contestant 1 always participates with effort level* x^**, contestant 2 stays inactive with probability* $1-p$ *and participates with an effort level* y^* *with probability*

p, *where*

$$x^* = \frac{1}{c_2}(r-1)^{-\frac{1}{r}}(1-\frac{1}{r})V, \quad y^* = \frac{1}{c_2}(1-\frac{1}{r})V, \quad p = \frac{1}{c}(r-1)^{-\frac{1}{r}},$$

$$E(TE)^m = x^* + py^* = \frac{1}{c_2}(r-1)^{-\frac{1}{r}}(1-\frac{1}{r})(1+\frac{1}{c})V. \tag{54}$$

According to (54),

$$E(TE)^m|_{r=2} = \frac{(c+1)}{2cc_2}V. \tag{55}$$

We have shown that for any $r \in [2, +\infty)$, there exists an all-pay auction equilibrium, using (53),

$$E[TE]^{All-pay} = \frac{(c+1)}{2cc_2}V = E(TE)^m|_{r=2}.$$

So the expected total effort when $r = 2$ from the mixed-strategy equilibrium coincides with that from the all-pay auction equilibrium. When $r = \overline{r}$ (then $c^{\overline{r}} = (\overline{r}-1)^{-1}$), in the mixed-strategy equilibrium we have:

$$\begin{aligned} E(TE)^m|_{r=\overline{r}} &= (\overline{r}-1)^{-\frac{1}{\overline{r}}}(1-\frac{1}{\overline{r}})(1+\frac{1}{c})\frac{1}{c_2}V \\ &= (1-\frac{1}{\overline{r}})(1+c)\frac{V}{c_2}. \end{aligned}$$

When $r = \overline{r}$, then we have $c^{\overline{r}} = (\overline{r}-1)^{-1}$, in the pure-strategy equilibrium, using (41),

$$TE|_{r=\overline{r}} = \frac{c^{\overline{r}}(1+c)\overline{r}V}{c_2(1+c^{\overline{r}})^2} = (1-\frac{1}{\overline{r}})(c+1)\frac{V}{c_2} = E(TE)^m|_{r=\overline{r}}.$$

So the mixed-strategy equilibrium when $r \in [\overline{r}, 2]$ connects the pure-strategy equilibrium when $r \in (0, \overline{r}]$ and the all-pay auction equilibrium when $r \in [2, +\infty)$ by making the expected total effort continuous in the region where $r \in (0, +\infty)$. Next we look at the expression of expected total effort when $r \in [\overline{r}, 2]$.

$$(54) \Rightarrow \frac{dE(TE)^m}{dr} = \frac{(c+1)}{cc_2r^3}V(r-1)^{1-\frac{1}{r}}\ln(r-1), \tag{56}$$

$$\overline{r} \le r \le 2 \text{ and } 1 < \overline{r} < 2 \Rightarrow \ln(r-1) < 0 \text{ and } (r-1)^{1-\frac{1}{r}} > 0.$$

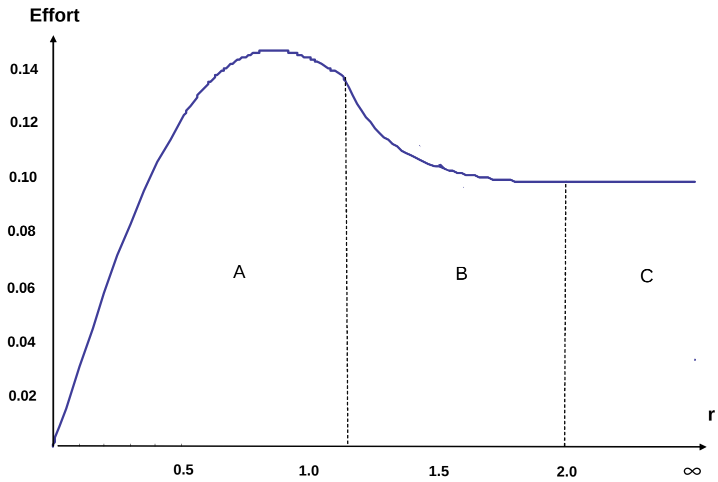

Figure 3

Then according to (56), we derive when $\overline{r} \leq r \leq 2$, $dE(TE)^m/dr < 0$. Therefore, in the mixed-strategy equilibrium when $r \in [\overline{r}, 2]$, the expected total effort always decreases with an increase of r.

To sum up, we have constructed the entire equilibrium set where $r \in (0, +\infty)$, which consists of region $(0, \overline{r}]$ where a unique pure-strategy equilibrium exists for any r in this region, region $[\overline{r}, 2]$ where a mixed-strategy equilibrium exists for any r in this region and region $[2, +\infty)$ where an all-pay auction equilibrium exists for any r in this region. We find that the expected total effort always decreases when $r \in [\overline{r}, 2]$ and stays constant when $r \in [2, +\infty)$. Therefore, in this equilibrium set, the optimal accuracy level in the region $(0, \overline{r}]$ is the optimal accuracy level in the entire region $(0, +\infty)$.[24]

Figure 3 depicts the case with $c_1 = 1$, $c_2 = 6$, where $r \in (0, +\infty)$. We can calculate that $\overline{r} \approx 1.132$ and $\widehat{r} \approx 0.861$. We can see that for any $r \in (0, 1.132]$ (in region A) a unique pure-strategy equilibrium exists, for any $r \in [1.132, 2]$ (in

[24] We have not been able to rule out other mixed-strategy equilibria when $r \in [\overline{r}, +\infty)$, so this is only a ranking of equilibria that have been characterized.

region B) a mixed-strategy equilibrium exists, and for any $r \in [2, +\infty]$ (in region C) an all-pay auction equilibrium exists. It is clear to see that the expected total effort is maximized when r reaches the optimal level $\widehat{r} \approx 0.861$. After that, TE decreases with r in region A and B and remains unchanged in region C.

2.3 Concluding Comments

In this paper, we interpret r in the Power CSF as the accuracy level of the contest designer. We claim that in many circumstances the contest designer has ways to change r, i.e., increase or decrease the accuracy level. The question we ask is, is it always better to be more accurate? With symmetric contestants the answer is quite simple, it has been shown that increasing accuracy (r) always increases total effort when $r < 2$ and when $r \geq 2$ the expected total effort stays constant[25] with an increase of r.

In this paper, we look at a model with two contestants who are heterogeneous in ability. We construct an equilibrium set on $r \in (0, +\infty)$, which consists of the pure-strategy equilibrium region when $r \in (0, \overline{r}]$, the mixed-strategy equilibrium region when $r \in [\overline{r}, 2]$ and the all-pay auction equilibrium region when $r \in [2, +\infty)$. We have shown that when the difference between the contestants' abilities is relatively small ($c \leq 3.5665$), total effort always increases with an increase of r when $r \in (0, \overline{r}]$ and stays constant after that. When the difference between the contestants' abilities is relatively big ($c > 3.5665$), the expected total effort increases with an increase of r until r reaches the optimal point $\widehat{r}$ where $\widehat{r} < \overline{r}$, after that total effort decreases in region $(\widehat{r}, 2]$ with an increase of r and stays constant in region $[2, +\infty)$. We also find that the optimal accuracy level ($\widehat{r}$) decreases when the contestants become more different in ability. Therefore, in a contest with players who are heterogeneous in ability, it is not always true that the more accurate the

[25] With symmetric contestants, $\overline{r} = 2$, the expected total effort is fully dissipated when $r \geq 2$.

better, especially when there is a big difference between the contestants' abilities.

Although it is reasonable to assume that there is always room for the contest designer to reduce his accuracy as long as $r > 0$, in many cases the contest designer's ability of increasing accuracy is bounded above. In cases where the feasible accuracy level is in the region $(0, \overline{r}]$, we can restrict attention to pure-strategies. In such cases we have a unique optimal accuracy level $\widehat{r}$. However, if accuracy levels above $\overline{r}$ are possible, we need to consider mixed-strategies. Although we still have the same optimal accuracy level for the equilibria we characterize, we are unable to rule out the existence of other mixed strategy equilibria. Therefore, our strongest results are for cases where there is this upper limit $\overline{r}$ on accuracy levels.

In recent years with the rapid development of technologies, some people have argued that it is time to introduce high-tech into sports (like tennis, football and basketball, etc.) to make the outcomes of the games a more accurate reflection of ability. In many cases, this has already been done. However, our model shows one reason for not using accuracy increasing technologies such as replays. If we think of sportsmen as having different inherent abilities then more accuracy may reduce the effort that they put into training and therefore reduce skill levels and the entertainment value of the games.

2.4 Appendix 2A

Assume there exists a pure-strategy equilibrium and in the equilibrium: contestant 1 makes an effort $e_1^* = a$ and contestant 2 makes an effort $e_2^* = b$ ($a > b$ since contestant 1 is more able). From contestant 1's point of view, given contestant 2 makes an effort b, he would want to choose an effort level x to maximize his expected profit π_1 where

$$\begin{aligned}\pi_1 &= \frac{x^r}{x^r + b^r}V - c_1 x, \\ \frac{d\pi_1}{dx} &= 0 \Rightarrow \frac{b^r x^{r-1} r V}{(b^r + x^r)^2} = c_1.\end{aligned}$$

In equilibrium, contestant 1 chooses $x = a$ to maximize his expected profit, then:

$$\frac{b^r a^{r-1} r V}{(b^r + a^r)^2} = c_1. \tag{57}$$

From contestant 2's point of view, given contestant 1 makes an effort a, he would want to choose an effort level y to maximize his expected profit π_2 where

$$\begin{aligned}\pi_2 &= \frac{y^r}{y^r + a^r}V - c_2 x, \\ \frac{d\pi_2}{dx} &= 0 \Rightarrow \frac{a^r y^{r-1} r V}{(a^r + y^r)^2} = c_2.\end{aligned}$$

In equilibrium, contestant 2 chooses $y = b$ to maximize his expected profit, then:

$$\frac{a^r b^{r-1} r V}{(a^r + b^r)^2} = c_2, \tag{58}$$

$$\frac{(57)}{(58)} \Rightarrow a = \frac{c_2}{c_1} b = cb.$$

Substituting $a = \frac{c_2}{c_1} b$ into (57), we have

$$\frac{b^r (\frac{c_2}{c_1} b)^{r-1} r V}{(b^r + (\frac{c_2}{c_1} b)^r)^2} = c_1,$$

$$\Rightarrow e_2^* = b = \frac{c_1^r c_2^r r V}{c_2(c_1^r + c_2^r)^2},$$

$$\Rightarrow e_1^* = a = \frac{c_1^r c_2^r r V}{c_1(c_1^r + c_2^r)^2}.$$

Next, we show the second order conditions are always satisfied when $e_1 = e_1^*$ and $e_2 = e_2^*$, i.e.,

$$\frac{d^2\pi_1}{de_1^2}|_{e_1=e_1^*,\ e_2=e_2^*} < 0; \quad \frac{d^2\pi_2}{de_2^2}|_{e_1=e_1^*,\ e_2=e_2^*} < 0.$$

Given contestant 2 makes an effort b, if contestant 1 makes an effort e_1, his expected profit would be:

$$\pi_1|_{e_2=b} = \frac{e_1^r}{e_1^r + b^r}V - c_1 e_1,$$

$$\Rightarrow \frac{d^2\pi_1}{de_1^2}|_{e_2=b} = \frac{b^r r V e_1^{-2+r}(b^r(r-1)) - (r+1)e_1^r)}{(e_1^r + b^r)^3}. \tag{59}$$

Substituting $e_1 = a = cb$ into (59), we have

$$\frac{d^2\pi_1}{de_1^2}|_{e_1=a,\ e_2=b} = -\frac{rVa^{-2+r}b^{2r}}{(a^r + b^r)^3}[(r+1)c^r - (r-1)].$$

$$\because \quad -\frac{rVa^{-2+r}b^{2r}}{(a^r + b^r)^3} < 0 \text{ and } [(r+1)c^r - (r-1)] > (1+r) + 1 - r = 2 > 0,$$

$$\therefore \quad \frac{d^2\pi_1}{de_1^2}|_{e_1=a,\ e_2=b} < 0.$$

Given contestant 1 makes an effort a, if contestant 2 makes an effort e_2, his expected profit would be:

$$\pi_2|_{e_1=a} = \frac{e_2^r}{e_2^r + a^r}V - ce_2 \Rightarrow$$

$$\frac{d^2\pi_2}{de_2^2}|_{e_1=a} = \frac{a^r r V e_2^{-2+r}(a^r(r-1)) - (r+1)e_2^r)}{(a^r + e_2^r)^3}. \tag{60}$$

Substituting $e_2 = b$ and $a = cb$ into (60), we have

$$\frac{d^2\pi_2}{de_2^2}|_{e_1=a,\ e_2=b} = \frac{a^r r V b^{-2+2r}}{(a^r + b^r)^3}[c^r(r-1)) - (r+1)],$$

$$
\begin{aligned}
\because \quad & \pi_2 \geq 0 \Rightarrow c^r(r-1) \leq 1 \\
\therefore \quad & c^r(r-1)) - (r+1) \leq -r < 0 \\
\Rightarrow \quad & \frac{d^2\pi_2}{de_2^2}|_{e_1=a,\ e_2=b} < 0.
\end{aligned}
$$

2.5 Appendix 2B

$$
\begin{aligned}
\pi_1|_{e_1=a;\ e_2=b} &= \frac{a^r}{a^r+b^r}V - c_1 a = \frac{c^r(1+c^r-r)V}{(1+c^r)^2} \geq 0 \\
&\Rightarrow c^r \geq r-1,
\end{aligned}
$$

$$
\begin{aligned}
\pi_2|_{e_1=a;\ e_2=b} &= \frac{b^r}{a^r+b^r}V - c_2 b = \frac{(1-c^r(r-1))V}{(1+c^r)^2} \geq 0 \\
&\Rightarrow c^r(r-1) \leq 1.
\end{aligned}
$$

therefore,

$$
\pi_1|_{e_1=e_1^*;\ e_2=e_2^*} \geq 0 \Rightarrow c^r \geq r-1, \tag{61}
$$

$$
\pi_2|_{e_1=e_1^*;\ e_2=e_2^*} \geq 0 \Rightarrow c^r(r-1) \leq 1. \tag{62}
$$

When $r \leq 1$, it is always the case that $c^r(r-1) \leq 1$ and $c^r \geq r-1$, thus the pure-strategy equilibrium always exists. When $r > 1$, to make sure the pure-strategy equilibrium exists, by (61) and (62) the following condition must hold:

$$
r-1 \leq c^r \leq \frac{1}{r-1}. \tag{63}
$$

If (63) holds, it must be the case that $r \leq 2$, otherwise $r-1 > 1/(r-1)$. When $r \leq 2$, we always have $r-1 \leq c^r$. So (63) changes to

$$
c^r \leq \frac{1}{r-1}.
$$

Given c constant, to make sure that the pure-strategy equilibrium exists, it must be the case that

$$
r \leq \overline{r}, \text{ where } \overline{r} \text{ satisfies } c^{\overline{r}} = \frac{1}{\overline{r}-1}.
$$

Let $f(c,\overline{r}) = c^{\overline{r}} - \frac{1}{\overline{r}-1} = 0$, we have

$$\begin{aligned}
\frac{\partial f}{\partial c} &= c^{\overline{r}-1}\overline{r} > 0 \text{ and } \frac{\partial f}{\partial \overline{r}} = \frac{1}{(\overline{r}-1)^2} + c^r \log c > 0 \\
df &= \frac{\partial f}{\partial c}dc + \frac{\partial f}{\partial \overline{r}}d\overline{r} = 0 \\
&\Rightarrow \frac{d\overline{r}}{dc} = -\frac{(\frac{\partial f}{\partial c})}{(\frac{\partial f}{\partial \overline{r}})} < 0. \\
\text{When } c &= 1,\, c^{\overline{r}} - \frac{1}{\overline{r}-1} = 1 - \frac{1}{\overline{r}-1} = 0, \\
&\Rightarrow \overline{r} = 2, \\
\text{When } c &\to +\infty,\, \lim_{c\to+\infty} c^{\overline{r}} - \frac{1}{\overline{r}-1} = 0 \\
&\Rightarrow \overline{r} \to 1.
\end{aligned}$$

So, we can safely conclude that when c increases from 1 to $+\infty$, $\overline{r}$ decreases from 2 to 1.

2.6 Appendix 2C

Let $f(r) = 1 + c^r - (c^r - 1)r\log c$, from (43),(44) and (45) we can derive:

$$\begin{aligned}
\frac{dTE}{dr} &= \frac{c^{r-1}(1+c)V}{(1+c^r)^3}f(r) \\
\frac{de_1^*}{dr} &= \frac{c^{r-1}V}{(1+c^r)^3}f(r) \\
\frac{de_2^*}{dr} &= \frac{c^r V}{(1+c^r)^3}f(r)
\end{aligned}$$

Thus, the signs of $\frac{dTE}{dr}$, $\frac{de_1^*}{dr}$ and $\frac{de_2^*}{dr}$ are all determined on the sign of $f(r)$.

$$\frac{df(r)}{dr} = [1 - c^r \log c^r]\log c,\quad \frac{d^2 f(r)}{dr^2} = -c^r(\log c)^2[1 + r\log c] < 0.$$

$f(r)|_{r=0} = 2 > 0$, $\frac{df(r)}{dr}|_{r=0} = \log c > 0$, and $\frac{df(r)}{dr}$ decreases as r increases (since $\frac{d^2 f(r)}{dr^2} < 0$). Let $r = \widehat{r}$ where

$$f(\widehat{r}) = 1 + c^{\widehat{r}} - (c^{\widehat{r}} - 1)\log c^{\widehat{r}} = 0 \ (\Rightarrow c^{\widehat{r}} \approx 4.68).$$

Therefore, we can conclude that

$$\begin{aligned} f(r) &> 0 \text{ when } r < \widehat{r}, \\ f(r) &= 0 \text{ when } r = \widehat{r}, \\ f(r) &< 0 \text{ when } r > \widehat{r}, \end{aligned}$$

i.e.,

$$\begin{aligned} \frac{dTE}{dr}(\frac{de_1^*}{dr}, \frac{de_2^*}{dr}) &> 0 \text{ when } r < \widehat{r} \\ \frac{dTE}{dr}(\frac{de_1^*}{dr}, \frac{de_2^*}{dr}) &= 0 \text{ when } r = \widehat{r} \\ \frac{dTE}{dr}(\frac{de_1^*}{dr}, \frac{de_2^*}{dr}) &< 0 \text{ when } r > \widehat{r} \end{aligned}$$

where $\widehat{r}$ satisfies

$$1 + c^{\widehat{r}} - (c^{\widehat{r}} - 1)\log c^{\widehat{r}} = 0 \ (\Rightarrow c^{\widehat{r}} \approx 4.68).$$

2.7 Appendix 2D

$$\begin{aligned} \text{When } \widehat{r} &\leq 1,\ c^{\widehat{r}} \approx 4.68 \Rightarrow c \geq 4.68; \\ \text{When } \widehat{r} &> 1,\ c^{\widehat{r}} \leq 1/(\widehat{r} - 1) \text{ and } c^{\widehat{r}} \approx 4.68 \\ &\Rightarrow 4.68 \leq 1/(\widehat{r} - 1) \\ &\Rightarrow \widehat{r} \leq 1.21368, \text{ and } c^{\widehat{r}} \approx 4.68 \\ &\Rightarrow c \geq 3.5665. \end{aligned}$$

So to make sure $\widehat{r}$ locates in the region $[0, \overline{r}]$, we need $c \geq 3.5665$. In other words, to ensure that $dTE/dr < 0$ happens in the region where $r \in [0, \overline{r}]$, we need $c > 3.5665$.

$$\lim_{r \to 0} e_1^* = \lim_{r \to 0} \frac{c^r r V}{c_1(1 + c^r)^2} = \frac{1 \times 0 \times V}{c_1(1 + 1)^2} = 0,$$

$$\lim_{r \to 0} e_2^* = \lim_{r \to 0} \frac{c^{r-1} r V}{c_1(1 + c^r)^2} = \frac{(1/c) \times 0 \times V}{c_1(1 + 1)^2} = 0,$$

$$\underset{r\to 0}{lim} TE = \underset{r\to 0}{lim} \frac{(c^{r-1}+c^r)rV}{c_1(1+c^r)^2} = \frac{((1/c)+c)\times 0 \times V}{c_1(1+1)^2} = 0.$$

2.8 Appendix 2E

Substituting (40) into (38), we have

$$\begin{aligned} \pi_1^* &= \frac{c^r(1+c^r-r)V}{(1+c^r)^2}; \\ \pi_2^* &= \frac{(1+c^r-c^r r)V}{(1+c^r)^2}. \end{aligned}$$

Then we further derive that

$$\frac{\partial \pi_1^*}{\partial r} = \frac{-c^r V[(1+c^r-(1-r+c^r(1+r))\log c]}{(1+c^r)^3},$$

$$\frac{\partial \pi_2^*}{\partial r} = \frac{-c^r V[1+c^r+(1+r+c^r(1-r))\log c]}{(1+c^r)^3}.$$

Since $c^r \geq r-1$, we can derive that

$$1+r+c^r(1-r) \geq 1+r+(r-1)(1-r) = (3-r)r > 0.$$

Therefore, $\frac{\partial \pi_2^*}{\partial r} < 0$. The sign of $\frac{\partial \pi_1^*}{\partial r}$ depends on the sign of $g(c,r)$ where

$$g(c,r) = (1-r+c^r(1+r))\log c - (1+c^r),$$

$$\frac{\partial g}{\partial c} = \frac{c^r-(r-1)+c^r r(1+r)\log c}{c} > 0 \text{ since } c^r \geq r-1.$$

So g increases as c increases. We also derive

$$\frac{\partial g}{\partial r} = (c^r(1+r)\log c - 1)\log c, \tag{64}$$

$$g|_{r=0} = 2(\log c - 1). \tag{65}$$

From (64) and (65), we can see that when $c \geq e \approx 2.718$,

$$g|_{r=0} = 2(\log c - 1) \geq 0 \text{ and } \frac{\partial g}{\partial r} = (c^r(1+r)\log c - 1)\log c > 0.$$

So when c is big ($c \geq e \approx 2.718$), $\partial\pi_1^*/\partial r \geq 0$ for all $r \in (0, \overline{r}]$. When $c < e \approx 2.718$, there are two possible cases. The first case is when c is small where $g \leq 0$ for all r. The second case is when c is medium, although $g|_{r=0} < 0$, g increases with r when $\partial g/\partial r = (c^r(1+r)\log c - 1)\log c > 0$ and $g > 0$ after r exceeds some certain value. To sum up, when c is small, $\partial\pi_1^*/\partial r \leq 0$ for all r, when c is medium, $\partial\pi_1^*/\partial r \leq 0$ for small r and $\partial\pi_1^*/\partial r > 0$ for big r, and when c is big $\partial\pi_1^*/\partial r \geq 0$ for all r.

2.9 Appendix 2F

From contestant 1's point of view, given his opponent's strategy is staying inactive with probability $1-p$ and biding y^* with probability p. His expected profit is

$$E\Pi_1 = \{p\frac{x^r}{(x^r + y^{*r})} + (1-p)\}V - c_1 x. \tag{66}$$

Since x^* is the maximizer of (66), then we have

$$c_1((x^*)^r + (y^*)^r)^2 = (y^*)^r(x^*)^{-1+r}prV. \tag{67}$$

Given contestant 1 bids x^*, contestant 2's expected profit is

$$E\Pi_2 = \frac{y^r}{(y^r + (x^*)^r)}V - c_2 y. \tag{68}$$

Since y^* is the maximizer of (68), then we have

$$c_2((x^*)^r + (y^*)^r)^2 = (x^*)^r(y^*)^{-1+r}rV. \tag{69}$$

Because contestant 2's expected revenue is always zero in equilibrium, then we have

$$E\Pi_2|_{y=y^*} = \frac{(y^*)^r}{((y^*)^r + (x^*)^r)}V - c_2(y^*) = 0. \tag{70}$$

From (67) and (69), we have

$$\frac{1}{c_1}(y^*)^r(x^*)^{-1+r}prV = \frac{1}{c_2}(x^*)^r(y^*)^{-1+r}V$$

$$\Rightarrow x^* = (\frac{c_2}{c_1})py^* = cpy^*. \tag{71}$$

Substituting (71) into (70), we have

$$\begin{aligned} y^* &= \frac{V}{c_2(1+(cp)^r)}, &(72)\\ x^* &= \frac{cpV}{c_2(1+(cp)^r)}. &(73) \end{aligned}$$

Substituting (71) into (69), we have

$$y^* = \frac{(cp)^r rV}{c_2(1+(cp)^r)^2} = \frac{(cp)^r r}{(1+(cp)^r)}(\frac{V}{c_2(1+(cp)^r)}) = \frac{(cp)^r r}{(1+(cp)^r)}y^*,$$

$$\Rightarrow \quad p^r c^r = \frac{1}{(r-1)}, \tag{74}$$

$$\Rightarrow \quad p = \frac{1}{c}(r-1)^{-\frac{1}{r}}. \tag{75}$$

Substituting (74) into (72) and (73), we have

$$\begin{aligned} y^* &= \frac{1}{c_2}(1-\frac{1}{r})V, \\ x^* &= (r-1)^{-\frac{1}{r}}\frac{1}{c_2}(1-\frac{1}{r})V. \end{aligned}$$

To make sure of the existence of the mixed-strategy equilibrium we proposed, we need to check the second order conditions when $r \in [\overline{r}, 2]$

$$(66) \Rightarrow \frac{d^2 E\Pi_1}{dx^2} = \frac{-prVx^{-2+r}y^r}{(x^r+y^r)^3}[(1+r)x^r - (r-1)y^r)].$$

In equilibrium $x = cpy$, $p^r c^r = (r-1)^{-1}$,

$$\begin{aligned} \because \quad & (1+r)x^r - (r-1)y^r \\ = \quad & y^r((1-r)+(1+r)p^r c^r) \\ = \quad & \frac{y^r r(3-r)}{(r-1)} > 0 \text{ since } 1 < \overline{r} \le r \le 2, \\ \therefore \quad & \frac{d^2 E\Pi_1}{dx^2}|_{x=x^*,\ y=y^*} < 0 \text{ when } \overline{r} \le r \le 2. \end{aligned}$$

$$
\begin{aligned}
E\Pi_2 &= \frac{y^r}{(y^r + x^r)}V - c_2 y \Rightarrow \\
\frac{d^2 E\Pi_2}{dy^2} &= \frac{rVy^{-2+r}x^r((r-1)x^r - (1+r)y^r)}{(x^r + y^r)^3},
\end{aligned}
$$

and also

$$
(r-1)x^r - (1+r)y^r = ((r-1)p^r c^r - (1+r))y^r = -ry^r < 0.
$$

$$
\therefore \frac{d^2 E\Pi_2}{dy^2}|_{x=x^*,\ y=y^*} < 0 \text{ when } \overline{r} \leq r \leq 2.
$$

Therefore, we can safely express the expected total effort in equilibrium as follows

$$
\begin{aligned}
E(TE)^m &= x^* + py^* \\
&= (r-1)^{-\frac{1}{r}}(1 - \frac{1}{r})(1 + \frac{1}{c})\frac{1}{c_2}V. \qquad (76)
\end{aligned}
$$

3 Chapter 3: The Power Contest Success Function and the Power Contest Defeat Function

3.1 Introduction

In the theory of contests, we can distinguish between two main branches in the literature. Firstly, perfectly discriminating contests – effort is perfectly observable and the contestants making irreversible effort are designated the winners according to their effort levels: the highest-effort contestant wins the first prize, the second highest-effort contestant wins the second prize and so on. Perfectly discriminating contests have been studied extensively by Hillman and Samet (1987), Hillman and Riley (1989) and Krishna & Morgan (1997) for the case of a single prize, and the cases of several prizes has also been studied by Clark and Riis(1998) and Moldovanu and Sela (2001), among others. Secondly, imperfectly discriminating contests – effort is not perfectly observable, so the contestant who expends the largest effort may not win the prize, but the probability of a particular contestant winning is increasing in his effort and decreasing in the effort of the opponents'. In either literature, the focus has been on the case in which a number of contestants compete to win prizes and much less attention is paid to punishing the bottom players. However, except rewarding top players (winners), punishing bottom players (losers) is also a popular way to motivate contestants to compete and actually occurs often in practice. For instance, when assigning University course marks, 5% of the students will be failed; or the worst interns will be fired at the end of their internship. What conclusions can we draw if we consider punishments in contests? In Chapter 1, we look at punishment in the literature on perfectly discriminating contests, this paper (Chapter 3) is an attempt to take both prizes and punishments into account in the literature on imperfectly discriminating contests.

The Power Contest Success Function (the Power CSF, proposed by Tullock in 1980), has been much used to select the winner or multiple winners in the literature on imperfectly discriminating contests. However, things become more complicated in a technical way when the contest designer wants to identify the bottom players in order to punish them. This is because we need the whole rank of all contestants to identify the bottom players (i.e. the losers) in the Power CSF mechanism. In this paper, we propose the Power Contest Defeat Function (Power CDF) which successively eliminates the loser at a time. We show that the Power CDF has the same good qualities as the Power CSF and is more realistic in some cases. For instance, suppose several cities are in a competition to host the Olympic Games, one city is going to be eliminated in each round until only one city remains, which is the winner – in the contest of hosting 2012 Olympic games, Moscow was eliminated in the first round, New York in the second round, and then Madrid in the third round, Paris was eliminated in the last round and so the winner was London.

In this paper we look at both the Power CSF mechanism (selecting winners in sequence) and the Power CDF mechanism (selecting losers in sequence) and show that punishments increase expected total efforts significantly. More interestingly, we also find that when the contestants' effort levels are different, the Power CDF mechanism is more accurate in finding the correct winner (the one who makes the greatest effort) and the Power CSF mechanism is more accurate in finding the correct loser (the one who makes the smallest effort). In other words, the multi-step mechanism provides more accuracy in finding the correct winner or loser.

The Power CSF mechanism and the Power CDF mechanism imply two different procedures and yield different results. In the mean time, compared with the Power CSF mechanism, the Power CDF mechanism is no more complicated and more realistic in some cases, then why hasn't the Power CDF been proposed and used before while the Power CSF has been studied extensively since it was proposed by Tullock in 1980? The main reason might be because in the literature on contests, the focus has been on the top players who can acquire prizes (i.e. the winners) and less attention has been paid to the bottom players who can be punished (i.e. the losers).

3.2 The Power Contest Success Function (CSF)

3.2.1 Definition of the Power CSF

In an imperfectly discriminating contest, all contestants compete for prizes by expending effort so as to increase their probability of winning. In this literature, a critical component of a contest is the Contest Success Function (CSF) which provides each player's probability of winning for any given level of efforts. The Power CSF[26] was first proposed by Tullock (1980) to study the problem of rival rent-seekers who expend resources to influence the policy outcome in their favor, and it has been widely employed and analyzed in research since then. The Power CSF can be expressed as follows:

$$
\begin{aligned}
p_i(e) &= \frac{e_i^r}{\sum_{j=1}^{n} e_j^r} && \text{if } max\{e_1, ...e_n\} > 0; \\
p_i(e) &= 1/n && \text{otherwise,}
\end{aligned}
$$

where $e = (e_1, e_2, ..., .e_n)$ denotes a vector of efforts for the n players. In the above Power CSF, there are n contestants, e_i refers to the effort contestant i makes and $p_i(e)$ refers to the probability that contestant i wins the contest. We assume that the parameter r is an exogenous variable where $r \geq 0$. Mathematically, r is the elasticity of the odds of winning[27] for contestant i. It is often interpreted as

[26] It also has been called the Tullock CSF.

[27] Note here $r = \dfrac{d \ln[p_i(e)/(1-p_i(e))]}{d \ln e_i}$, i.e., r measures the percentage changes of $p_i(e)/(1 - p_i(e))$ in response to one percent change of e_i.

indicating returns to scale in efforts: if $r > 1$, $r < 1$ or $r = 1$, then returns to scale are increasing, decreasing or constant. So r can be seen as the discriminatory power of the Power CSF. Note that as $r \to +\infty$ the Power CSF becomes perfectly discriminating.

What are the reasons why the Power CSF has been used extensively in the literature apart from its analytical convenience? Skaperdas (1996) derives the Power CSF from easily interpretable axioms, which justifies the Power CSF by showing that the Power CSF is the **only** continuous success functional form which satisfies all the following axioms[28].

1. $\sum_{i \in n} p_i(e) = 1$ and $p_i(e) \geq 0$.

2. For all $i \in n$, $p_i(e)$ is increasing in e_i and decreasing in e_j for all $i \neq j$.

3. For any two different contestants $i \neq j$ (given other contestants' efforts the same), $p_i(e) = p_j(e)$ if $e_i = e_j$.

4. Let N be the set of players who may participate in a contest. Denote by $p_i^M(e)$ the i^{th} contestant's probability of success who participates in a contest among the members of the subset M (*i.e.* $M \subset N$) which we assume to have at least two elements and $\sum_{j \in M} p_j(e) > 0$, $p_i^M(e) = p_i(e)/[\sum_{j \in M} p_j(e)]$.

5. $p_i^M(e)$ is independent of the efforts of the players who are not included in the subset M ($\subset N$).

6. Let $\lambda e = (\lambda e_1, \lambda e_2, ..., \lambda e_n)$, then $p_i(\lambda e) = p_i(e)$.

3.2.2 The Power CSF Mechanism

In the literature on imperfectly discriminating contests, most papers focus on the first-prize contest in which the winner is determined by applying the Power CSF

[28] For a review of the general properties of Contest Success Functions, see Skaperdas(1996).

once, there is relatively little on multi-prize context. Berry (1993) and Clark and Riis (1996) present a method, based on an imperfectly discriminating rent-seeking game, for distributing several homogeneous rents in which each player may win no more than once. Clark and Riis (1998 c) present a contribution to the analysis of imperfectly discriminating rent-seeking contests with several positive prizes.

In the imperfectly discriminating context with multiple prizes, the following procedure has been much used to select multiple winners, which we call the Power CSF mechanism in this paper. Given each contestant has made his effort, firstly, the contest designer[29] selects the winner (the one who gets the first prize) by applying the Power CSF once, then the contest designer takes the winner out and he selects the second winner (the one who gets the second prize) among the remaining contestants by applying the Power CSF again[30]. This procedure goes on until the last winner has been chosen. Intuitively, the Power CSF mechanism describes a procedure selecting winners in sequence, i.e., from the first player who gets the first prize to the player who gets the last prize.

Example 1 *In a contest with three contestants, with a Power CSF where* $r = 1$*, suppose the contestants' effort levels are* $e_1 = 1$*,* $e_2 = 2$*,* $e_3 = 3$ *respectively. Assume the contest designer allocates three prizes*[31] *where* $a_1 > a_2 > a_3$*. Each contestant has a certain probability of winning any of the prizes in the Power CSF mechanism. We list all the possible ranks and their corresponding probabilities in table 1.*

[29] It can also be called the contest administrator.

[30] Note that the Power CSF can be used several times in this context due to it satisfying Property 4 above (see Skaperdas (1996)).

[31] Here, the prizes do not necessarily have to be positive, some might be negative prizes, i.e., punishments to the bottom players.

Rank Order	Probability
123	4/60
132	6/60
231	15/60
213	5/60
312	10/60
321	20/60

Table 1

It can be seen that in each row of table 1, the left part refers to the rank of the contestants and the right part refers to the probability of this rank occurring. For example, the first row $\{123, 4/60\}$ indicates there is a 4/60 chance that this contest will end up with rank 123 (rank 123 means contestant 1 wins a_1, contestant 2 wins a_2 and contestant 3 wins a_3). The following is how we calculate this probability: given $e_1 = 1$, $e_2 = 2$ and $e_3 = 3$, there is a probability of $1/(1 + 2 + 3) = 1/6$ that contestant 1 is selected as the winner who gets a_1. Given a_1 is allocated to contestant 1, between contestant 2 and contestant 3, there is a probability of $2/(2 + 3) = 2/5$ that contestant 2 is selected as the winner who gets a_2. So the probability of rank 123 occurring is $1/6 \times 2/5 = 4/60$. By a similar procedure, we can calculate the probabilities of all other possible results.

3.2.3 The Power CSF Model

We state the assumptions of the Power CSF model as follows:

1. There are n contestants and n prizes: $a_1 \geqslant a_2 \geqslant a_3 ... \geqslant a_{n-1} \geqslant a_n$. Note here we do not assume the prizes must be positive, that is to say, there could be some zero prizes and negative prizes (i.e. punishments) for the bottom players.[32]

[32] The main differences between this model and that in Derek J. Clark and Chiristian Riis

2. Every contestant has the same linear cost function: $cost = x$, where x is the effort level; i.e., contestants are symmetric in ability.

3. Contestants are risk neutral and rational; the value of prizes (a_1, a_2, ..., a_n) and the discriminatory power (r) are public information.

4. The result of the contest is determined by the Power CSF mechanism.

We prove the following proposition in Appendix 3A:

Proposition 10 *In a unique symmetric pure-strategy equilibrium*[33] *of the Power CSF model, the average*[34] *effort (AE^s) which each contestant makes is:*

$$AE^s = \frac{r}{n}[(1-\frac{1}{n})a_1 + ... + c_i a_i + ... + (1-\frac{1}{n} - ... - \frac{1}{2} - 1)a_n] \tag{77}$$

where

$$c_i = 1 - \sum_{j=0}^{i-1} \frac{1}{n-j}. \tag{78}$$

From the expression of c_i, we can derive

$$c_1 = 1 - \frac{1}{n} \tag{79}$$

$$c_n = 1 - \sum_{j=0}^{n-1} \frac{1}{n-j} = -\sum_{t=2}^{n} \frac{1}{t}$$

$$\Rightarrow \mid c_n \mid = -c_n = \sum_{t=2}^{n} \frac{1}{t} \tag{80}$$

By analyzing (78), (79) and (80) in Appendix 3A, we prove the following results:

(1998 c) are: in their model with n contestants there are k prizes where $1 \leq k < n$ and all prizes must be strictly positive.

[33] Note that r has to be located in a certain region to ensure the existance of the symmetric equilibrium. Because the focus of this paper is on the analysis of the equilibrium rather than the conditions of its existance, we put all the relavant discussions and proof in Appendix 3B.

[34] Here "average" does not suggest there is a distribution on contestants' effort levels because every contestant makes the same effort in the symmetric equilibrium.

Proposition 11 *In a unique symmetric pure-strategy equilibrium of the Power CSF model: (a) when n is large, there exists $k^* \approx \lfloor 0.632n \rfloor$ where $c_1, c_2, ..., c_{k^*-1}$ are positive and $c_{k^*}, c_{k^*+1}, ..., c_n$ are negative*[35]*, (b) $c_i > c_{i+1}$, (c) $\sum_{i=1}^{n} c_i = 0$, (d) c_1 is smaller than $\mid c_n \mid$ for $n \geq 3$ (notice $a_1 = \mid a_n \mid$ for $n = 2$), (e) when n increases, $\mid c_n \mid$ increases faster than c_1.*

Among the above results, (a) indicates that, from a_1 to a_n, (approximately) the first 63.2% prizes have positive effect on increasing total effort while the other prizes have negative effect on increasing total effort. Notice that the prizes which locate behind the critical prize (a_{k^*}), i.e., $a_{k^*}, a_{k^*+1}, ..., a_n$, have negative effect on increasing average effort (since $c_{k^*}, c_{k^*+1}, ..., c_n$ are negative). We can treat these prizes as punishments by making them negative, which means if a contestant gets this position, he will be punished by suffering a loss.

(b) tells us that different prizes have different effects on increasing average effort. Among all prizes, an increase in the first prize has the biggest effect on increasing average effort, an increase in the second prize has the second biggest effect, ..., an increase in the i^{th} prize has the i^{th} biggest effect; and among all punishments, an increase in the last punishment[36] ($p_n = -a_n$) has the biggest effect and an increase in the second last punishment ($p_{n-1} = -a_{n-1}$) has the second biggest effect, ..., an increase in the last i^{th} punishment has the i^{th} biggest effect and so on. For instance, in the 6-contestant case, the average effort is:

$$x = \frac{r}{6}(\frac{50}{60}a_1 + \frac{38}{60}a_2 + \frac{23}{60}a_3 + \frac{3}{60}a_4 - \frac{27}{60}a_5 - \frac{87}{60}a_6),$$

[35] Firstly, note that $\lfloor x \rfloor$ or $floor(x)$ is the function that returns the highest integer less than or equal to x. For example, for $n = 10$, $k^* \approx \lfloor 0.632 \times 10 \rfloor = \lfloor 6.32 \rfloor = 6$. Secondly, note $k^* \approx \lfloor 0.632n \rfloor$ is a rough approximately true, but this equation becomes much more accurate as n gets large, where $k^* \approx \lfloor (1 - e^{-1})n \rfloor$.

[36] The last punishment refers to the punishment to a player who ranks last in a contest, and the second last punishment refers to the punishment to a player who ranks second last in the contest.

then

$$c_1 = \frac{50}{60}, c_2 = \frac{38}{60}, c_3 = \frac{23}{60}, c_4 = \frac{3}{60}, c_5 = -\frac{27}{60}, c_6 = -\frac{87}{60}.$$

Since c_5 and c_6 are negative, substituting p_5 and p_6 for $-a_5$ and $-a_6$ (p_i refers to the punishment a player receives if he ranks i^{th} among all contestants), let c_5^p and c_6^p be the coefficients of p_5 and p_6, so we write:

$$x = \frac{r}{6}(c_1a_1 + c_2a_2 + c_3a_3 + c_4a_4 + c_5^p p_5 + c_6^p p_6),$$

where

$$c_1 = \frac{50}{60}, c_2 = \frac{38}{60}, c_3 = \frac{23}{60}, c_4 = \frac{3}{60}, c_5^p = \frac{27}{60}, c_6^p = \frac{87}{60}.$$

We can see among the positive prizes:

$$c_1 = \frac{50}{60} > c_2 = \frac{38}{60} > c_3 = \frac{23}{60},$$

and among the punishments:

$$c_6^p = \frac{87}{60} > c_5^p = \frac{27}{60}.$$

So from (b) we can clearly see that: if there is a total prize (punishment) sum and the contest designer determines the distribution of the total prize (punishment) sum among the different prizes (punishments), as long as the participation constraints hold, in order to maximize total effort it is always optimal for the designer to allocate the entire prize (punishment) sum to a single prize (punishment), i.e., to make a_1 and p_n $(i.e. -a_n)$ as big as possible.

(c) simply states that if every prize is increased by the same amount, say Δa, the average effort does not change ($\Delta AE^s = \sum_{i=1}^n c_i \Delta a = \Delta a \sum_{i=1}^n c_i = 0$). Because in this model we assume there is a prize for each contestant, increasing every prize by the same amount means every contestant is better off by the same amount, however, the competition between all contestants is the same as before. Therefore, everyone makes the same effort in equilibrium after the change.

More surprisingly, (d) and (e) indicate that in the Power CSF mechanism with $n \geq 3$ contestants, the last punishment is always more effective in increasing average effort than the first prize, and if the number of contestants increases, the last punishment would become more and more effective in increasing average effort compared with the first prize. For instance: in a 3-contestant case:

$$x = \frac{r}{18c}(4a_1 + a_2 - 5a_3) = \frac{r}{18c}(4a_1 + a_2 + 5p_3).$$

By looking at the above equation, it can be seen that if the contest designer increases 1 unit (monetary material) on the first prize (i.e. a_1), holding other prizes fixed, the average effort will increase by $4r/18$ units. If the contest designer increases 1 unit on the last punishment (i.e. p_3), the average effort will increase by $5r/18$ units. This is also true in all n-contestant cases for $n \geq 3$.

Why is the last punishment more effective than the first prize on increasing the average effort? In our model, getting the last punishment means the bottom player will lose p_n and getting the first prize means the top player will gain a_1. According to the assumptions of our model, contestants are risk neutral, one might expect that if $p_n = a_1$, the incentive of a player trying to get the first prize should be as big as the incentive of a player trying to avoid the punishment, in other words, the first prize and last punishment should have the same effect in increasing average effort. But from result (b), the last punishment is always more effective than first prize in increasing average effort. How should we explain this paradox? The key to answering this question lies in the Power CSF mechanism we use. We will discuss this question later in this paper.

3.3 The Power Contest Defeat Function (CDF)

The Power CSF has been much used to select winners in the literature on imperfectly discriminating contests. However, things are more complicated in a technical way when the contest designer wants to identify the bottom players in order to punish them. This is because we need the whole rank of all contestants to identify the bottom players (i.e. the losers) in the Power CSF mechanism. Are there other mechanisms which can be used beside the Power CSF mechanism? If there are, will the results differ and can the difference help us to understand the paradox we raise before?

In this section we propose the Power Contest Defeat Function (CDF) and consider selecting losers in sequence in the Power CDF Mechanism. In practice, there are many applications which have a similar procedure on ranking contestants or candidates. For instance, suppose several cities are in a competition of hosting the Olympic Games, one city will be eliminated in each round until only one city remains, which is the winner. As in the contest of hosting 2012 Olympic games, Moscow was eliminated in the first round, New York in the second round, and then Madrid in the third round, Paris was eliminated in the last round and so the winner was London. There are other examples, as in some real beauty contests or TV shows, the judges don't select the winner directly, they eliminate one or several contestants in each round until one contestant is left, then the last contestant is the winner and the rank of all contestants can be derived from the order of elimination.

3.3.1 Definition of the Power CDF

We define the power Contest Defeat Function (the Power CDF) as follows:

$$p_i(e) = \frac{e_i^{-r}}{\sum_{j=1}^{n} e_j^{-r}} \qquad \text{if } \; min\{e_1, ...e_n\} > 0;$$

$$p_i(e) = \frac{1}{m} \text{ for } e_i = 0 \text{ and } \; p_i(e) = 0 \text{ for } e_i > 0 \qquad \text{otherwise,}$$

where $e = (e_1, e_2, ..., .e_n)$ denotes a vector of efforts for the n players and m is the number of contestants making zero effort.

In the above Power CDF, there are n contestants, e_i refers to the effort contestant i makes and p_i refers to the probability of contestant i losing. As in the Power CSF, the parameter r is an exogenous variable where $r \geq 0$ and r here can be interpreted as indicating returns to scale in efforts. If $r > 1$, $r < 1$ or $r = 1$, then returns to scale are increasing, decreasing or constant. So r is the discriminatory power of the Power CDF and as $r \to +\infty$ the Power CDF becomes perfectly discriminating. It is straightforward to show that the Power CDF satisfies all the following axioms (which are similar with the axioms that the Power CSF satisfies).

1. $\sum_{i \in n} p_i(e) = 1$ and $p_i(e) \geq 0$.

2. For all $i \in n$, $p_i(e)$ is decreasing in e_i and increasing in e_j for all $i \neq j$.

3. For any two different contestants $i \neq j$ (given other contestants' efforts the same), $p_i(e) = p_j(e)$ if $e_i = e_j$.

4. Let N be the set of players who may participate in a contest. Denote by $p_i^M(e)$ the i^{th} contestant's probability of losing who participates in a contest among the members of the subset M ($i.e.$ $M \subset N$) which we assume to have at least two elements and $\sum_{j \in M} p_j(e) > 0$, $p_i^M(e) = p_i(e)/[\sum_{j \in M} p_j(e)]$.

5. $p_i^M(e)$ is independent of the efforts of the players not included in the subset $M(\subset N)$.

6. Let $\lambda e = (\lambda e_1, \lambda e_2, ..., \lambda e_n)$, then $p_i(\lambda e) = p_i(e)$.

3.3.2 The Power CDF Mechanism

In a Power CDF mechanism contest designer selects losers in sequence, i.e., from the first loser to the last loser (it is obvious that the last loser is the winner). Consider again Example 1 when the Power CSF is replaced with the Power CDF (with $r = 1$). Each contestant has a certain probability of winning any of the prizes in the Power CDF mechanism. We list all the possible ranks and their corresponding probabilities in the following table.

Rank Order	Probability
123	40/660
132	45/660
231	144/660
213	80/660
312	135/660
321	216/660

Table 2

In each row of table 2, the left part refers to the rank of the contestant and the right part refers to the probability of this rank occurring. For example, the first row $(123, 40/660)$ means there is a $40/660$ chance that this contest will end up with rank 123. The following is how we calculate this probability. Given $e_1 = 1, e_2 = 2$ and $e_3 = 3$, there is a probability of $(\frac{1}{3})/(1 + \frac{1}{2} + \frac{1}{3}) = 2/11$ that contestant 3 is selected as the first loser (i.e. he gets a_3). Then among the remaining contestants, i.e., contestant 1 and contestant 2, there is a probability of $(\frac{1}{2})/(1 + \frac{1}{2}) = 1/3$ that contestant 2 is selected as the loser. So probability of rank 123 occurring is $(2/11) \times (1/3) = 40/660$. Similarly, we can calculate the probabilities of other possible ranks. Therefore, the Power CDF mechanism actually implies a procedure of finding the rank of the contestants by selecting losers in sequence.

3.3.3 The Power CDF Model

The assumptions of the Power CDF model are the same as those of the Power CSF model except that the result of the contest is determined by the CDF mechanism instead of the Power CSF mechanism. In Appendix 3A, we prove the following proposition.

Proposition 12 *In a unique symmetric pure-strategy equilibrium*[37] *of the Power CDF model, the average*[38] *effort (AE^l) which each contestant makes is:*

$$AE^l = \frac{r}{n}[(\frac{1}{n} + \frac{1}{n-1} + ... + \frac{1}{2} + 1 - 1)a_1 + ... + c_i a_i + ... + (\frac{1}{n} - 1)a_n], \quad (81)$$

where

$$c_i = -1 + \sum_{j=0}^{n-i} \frac{1}{n-j}. \quad (82)$$

By comparing the results between the Power CSF mechanism and the Power CDF mechanism, i.e., (77) and (81), we find that in a n-contestant case, the coefficients in the two expressions of the average effort are connected in the following forms:

$$c_i^l = -c_{n-i+1}^s,$$

where c_i^s represents c_i in the Power CSF mechanism and c_i^l represents c_i in the Power CDF mechanism.

Proof. From (78) and (82), we have

$$c_i^s = 1 - \sum_{j=0}^{i-1} \frac{1}{n-j} \text{ and } c_i^l = -1 + \sum_{j=0}^{n-i} \frac{1}{n-j},$$

it follows that,

$$-c_{n-i+1}^s = \sum_{j=0}^{n-i+1-1} \frac{1}{n-j} - 1 = \sum_{j=0}^{n-i} \frac{1}{n-j} - 1 = c_i^l.$$

■

Since $c_i^l = -c_{n-i+1}^s$, we can simply derive the following proposition from Proposition 11.

[37] Note that r has to be located in a certain region to ensure the existance of the symmetric equilibrium. See Appendix 3B for details.

[38] Here "average" does not suggest there is a distribution on contestants' effort levels because every contestant makes the same effort in the symmetric equilibrium.

Proposition 13 *In a unique symmetric pure-strategy equilibrium of the Power CDF model: (a)* $c_1, c_2, ..., c_{k^*-1}$ *are positive and* $c_{k^*}, c_{k^*+1}, ..., c_n$ *are negative, where* $k^* \approx \lceil 0.368n \rceil$ *when* n *is large,*[39] *(b)* c_i *decreases when* i *increases, (d)* $\sum_{i=1}^{n} c_i = 0$*,* c_1 *is always bigger than* $| c_n |$ *for* $n \geq 3$ *(*$c_1 =| c_n |$ *for* $n = 2$*), (e) when* n *increases,* $| c_n |$ *increases more slowly than* c_1*.*

Among the above results, (b) and (c) tell very similar intuitions to that in the Power CSF mechanism. However, (a) indicates that (approximately) the first 36.8% prizes have positive effect on increasing total effort while 63.2% in the Power CSF mechanism. (d) and (e) show that under the Power CDF mechanism, the first prize is always more effective than the last punishment, and if the number of the contestants increases, the first prize will become more and more effective compared with the last punishment. So the results of (a), (d) and (e) is just the reverse of what we find under the Power CSF mechanism. In the next section, we will try to investigate this difference from a different angle.

3.4 Selecting the Highest Ability

In the Power CSF model and the Power CDF model, we analyze a symmetric equilibrium where each contestant makes the same effort. In this section, we look at a different situation where contestants are heterogenous and each contestant has the same (very short) time constraint of making effort. For example, when a company wants to recruit new employees, the recruiter will set written exam or interview for each applicant. In the given time period of the exam or interview, each applicant will do their best to impress the recruiter, i.e., make the most

[39] Firstly, $\lceil x \rceil$ or $ceiling(x)$ is the function that returns the smallest integer not less than x. For example, when $n = 100$, $k^* \approx \lceil 0.368 \times 100 \rceil = \lceil 36.8 \rceil = 37$. Secondly, note that $k^* \approx \lceil 0.368n \rceil$ is only approximately true, this equation gets much more accurate when n gets larger, when n is very large, $k^* \approx \lceil (e^{-1})n \rceil$.

possible effort. Therefore, contestants make effort according to their abilities, i.e., more/less able contestant make more/less effort. Notice that each contestant will do his/her best in the fixed time period because the marginal revenue of making effort is always more than the marginal cost of making effort. We can see that in these cases, the contest designer (the recruiter in the example) wants to select the correct winner instead of maximizing the total effort. In the following discussion, we try to look at the accuracy level of selecting the correct winner/loser in the Power CSF mechanism and the Power CDF mechanism[40] given that the contestants are making different levels of effort.

3.4.1 An Example

In the contest of Example 1, table 1 and 2 can be summarized by the following table:

CSF	CDF
123......(44/660)	123......(40/660)
132......(66/660)	132......(45/660)
231.....(165/660)	231.....(144/660)
213......(55/660)	213......(80/660)
312.....(110/660)	312.....(135/660)
321.....(220/660)	321.....(216/660)

Table 4

We can derive table 5 from table 4:[41]

[40] We assume that the contest designer can choose between the two mechanisms with r being the same.

[41] $110/660 + 220/660 = 330/660 \quad 135/660 + 216/660 = 351/660$
$165/660 + 220/660 = 385/660 \quad 144/660 + 216/660 = 360/660$

	CSF	CDF
Probability of selecting the correct winner	330/660	351/660
Probability of selecting the correct loser	385/660	360/660

Table 5

From the comparison of the results in table 5, we make the following conjecture:

Conjecture 1 *Given that the contestants are making different levels of effort, the Power CSF mechanism is more accurate in finding the correct loser while the Power CDF mechanism is more accurate in finding the correct winner.*

Next, we attempt to establish the above conjecture in a more general context than the example.

3.4.2 3-contestant case

In a 3-contestant case with (very short) time constraint of making effort, assume that contestant 3 is the most able one and makes the highest effort, contestant 1 is the least able one and makes the smallest effort, i.e., $e_1 < e_2 < e_3$. Can we prove the probability of picking contestant 3 as the winner under the Power CSF mechanism is always smaller than that under the Power CDF mechanism and the probability of picking contestant 1 as the loser under the Power CSF mechanism is always bigger than that under the Power CDF mechanism (like we conjectured)? The answer is positive.

Claim 1 *In a 3-contestant model with r being the same in both mechanisms, when the contestants make different effort levels, the probability of selecting the correct winner in the Power CSF mechanism is always smaller than that in the Power CDF mechanism, and the probability of selecting the correct loser in the Power CSF mechanism is always bigger than that in the Power CDF mechanism.*

Proof. Let p^s and p^l be the probability that contestant 3 wins the first prize under the Power CSF mechanism and the Power CDF mechanism respectively, then we derive:

$$p^s = \frac{e_3^r}{e_1^r + e_2^r + e_3^r},$$

$$\begin{aligned} p^l &= p(\text{rank } 321) + p(\text{rank } 312) \\ &= \left(\frac{e_1^{-r}}{e_1^{-r} + e_2^{-r} + e_3^{-r}}\right)\left(\frac{e_2^{-r}}{e_2^{-r} + e_3^{-r}}\right) + \left(\frac{e_2^{-r}}{e_1^{-r} + e_2^{-r} + e_3^{-r}}\right)\left(\frac{e_1^{-r}}{e_1^{-r} + e_3^{-r}}\right) \\ &= \left(\frac{e_3^r}{e_1^r e_2^r + e_2^r e_3^r + e_1^r e_3^r}\right)\left(\frac{e_2^r e_3^r}{e_2^r + e_3^r} + \frac{e_1^r e_3^r}{e_1^r + e_3^r}\right), \end{aligned}$$

then

$$\begin{aligned} p^s - p^l &= \frac{e_3^r}{e_1^r + e_2^r + e_3^r} - \left(\frac{e_3^r}{e_1^r e_2^r + e_2^r e_3^r + e_1^r e_3^r}\right)\left(\frac{e_2^r e_3^r}{e_2^r + e_3^r} + \frac{e_1^r e_3^r}{e_1^r + e_3^r}\right) \\ &= \frac{-e_1^r e_2^r e_3^r (e_3^{2r} - e_1^r e_2^r)}{(e_1^r + e_3^r)(e_2^r + e_3^r)(e_1^r + e_2^r + e_3^r)(e_1^r e_3^r + e_2^r e_3^r + e_1^r e_2^r)}, \end{aligned}$$

$$\because 0 < e_1 < e_2 < e_3 \Rightarrow e_3^{2r} - e_1^r e_2^r > 0, \quad \therefore p^s - p^l < 0.$$

Therefore, in a 3-contestant case with contestants making different effort levels, the probability of selecting the correct winner in the Power CDF mechanism is always bigger than that in the Power CSF mechanism.

By a similar procedure, we can also prove that the probability of selecting the correct winner (contestant 1) in the Power CDF is bigger than that in the Power CSF mechanism. In other words, the Power CSF mechanism is more accurate in finding the correct loser and the Power CDF mechanism is more accurate in finding the correct winner. ■

3.4.3 n-contestant case

We have shown that in a 3-contestant case with contestants making different effort levels, the Power CDF (CSF) mechanism is more accurate in finding the correct winner (loser). To make this argument stronger, we are going to analyze a specific model with n contestants.

Claim 2 *In a n-contestant model (where $n \geq 3$) with r being the same in both mechanisms, suppose only one contestant, say contestant 1, makes an effort e while all other contestants make the same effort level which we normalize it to 1, $e \neq 1$. When $e > 1$, the probability of selecting the correct winner (i.e. contestant 1) in the Power CDF mechanism is always bigger than that in the Power CSF mechanism; while when $e < 1$, the probability of selecting the correct loser (i.e. contestant 1) in the Power CSF mechanism is always bigger than that in the Power CDF mechanism.*

Proof. When $e > 1$, for a 3-contestant case, it is always the case that

$$\begin{aligned} p_3^l - p_3^s &= (\frac{2}{e^{-r}+2})(\frac{1}{e^{-r}+1}) - \frac{e^r}{e^r+2} \\ &= \frac{e^r(e^r+1)}{(1+2e^r)(1+e^r)(2+e^r)} > 0. \end{aligned}$$

Next step, we want to prove that for $n \geq 3$,

$$p_{n+1}^l - p_{n+1}^s > 0 \text{ if } p_n^l - p_n^s > 0,$$

$$p_n^s = \frac{e^r}{e^r+n-1},\ p_{n+1}^s = \frac{e^r}{e^r+n}$$

$$\Rightarrow p_{n+1}^s = (\frac{e^r+n-1}{e^r+n})p_n^s,$$

and

$$p_n^l = (\frac{n-1}{e^{-r}+n-1})...(\frac{2}{e^{-r}+2})(\frac{1}{e^{-r}+1}),$$

$$p_{n+1}^l = (\frac{n}{e^{-r}+n})(\frac{n-1}{e^{-r}+n-1})...(\frac{2}{e^{-r}+2})(\frac{1}{e^{-r}+1})$$

$$\Rightarrow p_{n+1}^l = (\frac{n}{e^{-r}+n})p_n^l.$$

Given $p_n^l > p_n^s$, we can derive $p_{n+1}^l > p_{n+1}^s$ because:

$$\begin{aligned}
e^r &> 1 \text{ since } n \geq 3 \\
&\Rightarrow (n-1)e^r > n-1 \\
&\Rightarrow \frac{1}{ne^r+1} < \frac{1}{e^r+n} \\
&\Rightarrow 1-\frac{1}{ne^r+1} > 1-\frac{1}{e^r+n} \\
&\Rightarrow \frac{n}{e^{-r}+n} > \frac{e^r+n-1}{e^r+n} \\
&\Rightarrow (\frac{n}{e^{-r}+n})p_n^l > (\frac{e^r+n-1}{e^r+n})p_n^s.
\end{aligned}$$

So for any $n \geq 3$,

$$p_n^l - p_n^s > 0 \Rightarrow p_{n+1}^l - p_{n+1}^s > 0.$$

If we change the assumptions a bit by assuming that contestant 1 makes the lowest effort, i.e., $e < 1$ and all other contestants make the same effort 1. By a similar procedure, we can prove that the probability of finding the correct loser (contestant 1) in the Power CSF mechanism is always bigger than that in the Power CDF mechanism. ■

3.5 Concluding Comments

The above results, at least to some extent, confirm our conjecture – given that the contestants are making different levels of effort (according to their abilities), the Power CSF mechanism is more accurate in finding the correct loser and the Power CDF mechanism is more accurate in finding the correct winner. Why? In the Power CSF mechanism, the contest designer only uses one step (apply the Power CSF once) to select the winner among all contestants, while in the Power CDF mechanism (with n contestants) the contest designer needs to use multiple steps (apply the Power CDF $n-1$ times) to find the winner. According to our results, given r the same in both mechanisms, the multi-step mechanism (i.e. the Power

CDF mechanism) provides more accuracy in finding the correct winner; while when the aim is to find the correct loser, the Power CSF mechanism, which takes multiple steps to find the loser, is more accurate than the Power CDF mechanism which selects the loser in one step. Therefore, the multi-step mechanism has an advantage in finding the correct winner/loser. With the same level of r, it is not straightforward to see why the multi-step mechanism provides more accuracy. Intuitively, the reason is that in the multi-step mechanism, with the number of contestants decreasing, the contest designer is more likely to choose the correct contestant among all remaining contestants.

With the same r in both mechanisms, if the last punishment and the first prize are the same amount monetarily, a rational risk neutral contestant will take the last punishment more seriously in the Power CSF mechanism and take the first prize more seriously in the Power CDF mechanism. This is because the Power CSF mechanism is more accurate in punishing the one who makes the lowest effort and the Power CDF mechanism is more accurate in rewarding the one who makes the highest effort. This also gives an explanation why the last punishment is more effective than the first prize in increasing average effort in the Power CSF mechanism and the first prize is more effective than the last punishment in increasing average effort in the Power CDF mechanism.

3.6 Appendix 3A

Proof of Proposition 10

Our aim is to prove that in a symmetric pure-strategy equilibrium, the expression of average effort each contestant makes is AE^s. We prove it by mathematical induction.

First-step. In a 2-contestant case with prizes: $a_1 > a_2$, the Power CSF gives:

$$p_1 = \frac{e_1^r}{e_1^r + e_2^r},\ p_2 = \frac{e_2^r}{e_1^r + e_2^r},$$

where p_i is contestant i' probability of winning and e_i is contestant i' effort.

$$L_1 = p_1 a_1 + (1 - p_1)a_2 - e_1 = \frac{e_1^r a_1 + e_2^r a_2}{e_1^r + e_2^r} - e_1,$$

$$L_2 = p_2 a_1 + (1 - p_2)a_2 - e_2 = \frac{e_2^r a_1 + e_1^r a_2}{e_1^r + e_2^r} - e_2,$$

where L_i is contestant i's expected utility. Each contestant maximizes his expected revenue, in the symmetric pure-strategy equilibrium[42] (contestants 1 and 2 will make the same level of effort since they are symmetric), $e_1 = e_2 = x$, then:

$$\frac{\partial L_1}{\partial e_1} = 0 \Rightarrow x = \frac{r(a_1 - a_2)}{4}.$$

By substituting $n = 2$ into (77)

$$AE^s = \frac{r}{2}[(1 - \frac{1}{2})a_1 + (1 - \frac{1}{2} - 1)a_2] = \frac{r(a_1 - a_2)}{4} = x,$$

so equation (77) holds when $n = 2$. Existence (of the equilibrium) requires:

$$\frac{\partial^2 L_1}{\partial e_1^2} < 0 \Leftrightarrow a_1 > a_2.$$

Participation constraint requires:

$$L_{1,2} = \frac{(a_1 + a_2)}{2} - \frac{r(a_1 - a_2)}{4} \geq 0 \Rightarrow r \leq \frac{2a_1 + 2a_2}{a_1 - a_2}.$$

Second-step. Assume that equation (77) is true in a t-contestant case where $t \geq 2$,

$$AE_t^s = \frac{r}{t}[(1 - \frac{1}{t})a_1 + ... + c_i a_i + ... + (1 - \frac{1}{t} - ... - \frac{1}{2} - 1)a_t],$$

[42] In this paper, we restrict our attention to the unique symmetric equilibrium for simplicity, it remains to be investigated whether there could exist an asymmetric equilibrium.

where

$$c_i = 1 - \sum_{j=0}^{i-1} \frac{1}{t-j},$$

then our task is to prove in the $t+1$ case,

$$AE_{t+1}^{s} = \frac{r}{t+1}[(1-\frac{1}{t+1})a_1 + ... + c_i a_i + ... + (1 - \frac{1}{t+1} - \frac{1}{t} - ... - \frac{1}{2} - 1)a_{t+1}],$$

where

$$c_i = 1 - \sum_{j=0}^{i-1} \frac{1}{t+1-j}.$$

In the symmetric pure-strategy equilibrium of the t-contestant and (t+1)-contestant cases, from any contestant's point of view, say contestant 1, given all other contestants are making the same effort x,

$$U^t = L_1^t a_1 + L_2^t a_2 + ... + L_i^t a_i + ... + L_t^t a_t - e_1,$$

$$U^{t+1} = L_1^{t+1} a_1 + L_2^{t+1} a_2 + ... + L_i^{t+1} a_i + ... + L_{t+1}^{t+1} a_t - e_1,$$

where U^t and U^{t+1} are the expected utilities for contestant 1 in the t contestant case and (t+1) contestant case respectively; and L_i^t and L_i^{t+1} are the probabilities of winning the i^{th} prize (a_i) for contestant 1 in the t-contestant case and the (t+1)-contestant case respectively. Therefore, L_i^t or $L_i^{t+1}$are the probabilities that contestant 1 is chosen as the winner in the i^{th} "round" (which indicates contestant 1 has not been chosen as a winner in all previous $i-1$ rounds). So given that the other contestants are making effort x, we can derive:

$$L_i^t = \frac{P_{i-1}^{t-1} x^{(i-1)r} e_1^r}{[e_1^r + (t-1)x^r][e_1^r + (t-2)x^r]...[e_1^r + (t-i+1)x^r][e_1^r + (t-i)x^r]}, \tag{83}$$

$$L_i^{t+1} = \frac{P_{i-1}^{t} x^{(i-1)r} e_1^r}{[e_1^r + tx^r][e_1^r + (t-1)x^r]...[e_1^r + (t-i+2)x^r][e_1^r + (t-i+1)x^r]}, \tag{84}$$

where

$$P_{i-1}^{t-1} = \frac{(t-1)!}{(t-i)!} = (t-1)(t-2)...(t-i+1),$$
$$P_{i-1}^{t} = \frac{t!}{(t-i+1)!} = t(t-1)...(t-i+2).$$

Contestant 1 maximizes his utility,

$$\frac{\partial U^t}{\partial e_1} = 0 \Rightarrow \sum_{i=1}^{t} \frac{\partial L_i^t}{\partial e_1} a_i = 1. \tag{85}$$

Because (77) is true in the t-contestant case,

$$(77) \quad \Rightarrow \quad x = \sum_{i=1}^{t} \frac{r}{t}(1 - \sum_{j=0}^{i-1} \frac{1}{t-j}) a_i \tag{86}$$

$$\Rightarrow \quad \sum_{i=1}^{n} \frac{r}{tx}(1 - \sum_{j=0}^{i-1} \frac{1}{t-j}) a_i = 1, \tag{87}$$

$$(85) \text{ and } (87) \Rightarrow \sum_{i=1}^{t} \frac{\partial L_i^t}{\partial e_1} a_i = \sum_{i=1}^{t} \frac{r}{tx}(1 - \sum_{j=0}^{i-1} \frac{1}{t-j}) a_i. \tag{88}$$

For any given $\{a_1, a_2, ..., a_t\}$, (88) always holds if and only if

$$\frac{\partial L_i^t}{\partial e_1} = \frac{r}{tx}(1 - \sum_{j=0}^{i-1} \frac{1}{t-j}). \tag{89}$$

Therefore, we conclude (77) holds for $n = t$ given (89) holds. Similarly, we can derive that (77) holds for $n = t+1$ given

$$\frac{\partial L_i^{t+1}}{\partial e_1} = \frac{r}{(t+1)x}(1 - \sum_{j=0}^{i-1} \frac{1}{t+1-j}). \tag{90}$$

So the our aim is to show (90) holds if (89) holds. The following is the proof.

From (83) and (84), we have

$$L_i^{t+1} = \frac{P_{i-1}^t [e_1^r + (t-i)x^r]}{P_{i-1}^{t-1}[e_1^r + tx^r]} L_i^t = \frac{P_{i-1}^t}{P_{i-1}^{t-1}}(1 - \frac{ix^r}{e_1^r + tx^r}) L_i^t,$$

$$\Rightarrow \frac{\partial L_i^{t+1}}{\partial e_1} = \frac{P_{i-1}^t}{P_{i-1}^{t-1}}[\frac{rix^r e_1^{r-1}}{(e_1^r + tx^r)^2} L_i^t + (1 - \frac{ix^r}{e_1^r + tx^r})(\frac{\partial L_i^t}{\partial e_1})]. \tag{91}$$

Substituting $e_1 = x$ into (83):

$$L_i^t = \frac{(t-1)(t-2)...(t-i+1)x^{ir}}{t(t-1)...(t-i+1)x^{ir}} = \frac{1}{t}. \tag{92}$$

Substituting (92), into (91):

$$\begin{aligned} \frac{\partial L_i^{t+1}}{\partial e_1} &= \frac{t(t-1)...(t-i)(t-i+1)}{(t-1)(t-2)...(t-i)}[\frac{rix^r e_1^{r-1}}{t(e_1^r + tx^r)^2} + (1 - \frac{ix^r}{e_1^r + tx^r})(\frac{\partial L_i^t}{\partial e_1})] \\ &= (\frac{t}{t-i+1})[\frac{rix^r e_1^{r-1}}{t(e_1^r + tx^r)^2} + (1 - \frac{ix^r}{e_1^r + tx^r})(\frac{\partial L_i^t}{\partial e_1})]. \end{aligned} \tag{93}$$

Substituting $e_1 = x$ and (89) into (93), we derive

$$\begin{aligned}
\frac{\partial L_i^{t+1}}{\partial e_1} &= \frac{r}{x(t+1)}[\frac{1}{(t-i+1)} - \frac{1}{t+1} + (1 - \sum_{j=0}^{i-1} \frac{1}{t-j})] \\
&= \frac{r}{x(t+1)}[\frac{1}{(t-i+1)} - \frac{1}{t+1} + 1 - (\frac{1}{t} + \frac{1}{t-1} + ... + \frac{1}{t-i+1})] \\
&= \frac{r}{x(t+1)}[1 - (\frac{1}{t+1} + \frac{1}{t} + \frac{1}{t-1} + ... + \frac{1}{t-i+2})] \\
&= \frac{r}{x(t+1)}(1 - \sum_{j=0}^{i-1} \frac{1}{t+1-j}),
\end{aligned}$$

which is exactly (90) and completes the proof.

Proof of Proposition 11

Proof of (a):

$$c_{k^*+1} = 1 - \sum_{j=0}^{k^*} \frac{1}{n-j} \approx 0 \Rightarrow \sum_{j=0}^{k^*} \frac{1}{n-j} \approx 1,$$

since

$$\sum_{j=0}^{k^*} \frac{1}{n-j} = \sum_{j=0}^{k^*} \frac{1/n}{1-j/n} = \frac{1}{n} \sum_{j=0}^{k^*} \frac{1}{1-j/n},$$

when n is large, we have:

$$\frac{1}{n} \sum_{j=0}^{k^*} \frac{1}{1-j/n} \approx \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{k^*} \frac{1}{1-j/n} = \int_0^{k^*/n} \frac{1}{1-t} dt \approx 1,$$

$$\Rightarrow \frac{k^*}{n} \approx \frac{e-1}{e} \approx 0.632 \Rightarrow k^* \approx 0.632n.$$

Proof of (c):

$$\begin{aligned}
\sum_{i=1}^{n} c_i &= c_1 + c_2 + ... + c_n \\
&= (1 - \frac{1}{n}) + (1 - \frac{1}{n} - \frac{1}{n-1}) + ... + (1 - \frac{1}{n} - \frac{1}{n-1} - \frac{1}{2} - 1)n \\
&= n - \{n\frac{1}{n} + (n-1)(\frac{1}{n-1}) + (n-2)(\frac{1}{n-2}) + ... + 2(\frac{1}{2}) + 1\} \\
&= n - n = 0.
\end{aligned}$$

Proof of (d) and (e):

$c_1 = 1 - \frac{1}{n}$, when n increases from n to $n+1$, c_1 will increase by

$$(1 - \frac{1}{n+1}) - (1 - \frac{1}{n}) = \frac{1}{n(n+1)},$$

relatively, $-c_n$ will increase by

$$(\frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n} + \frac{1}{n+1}) - (\frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n}) = \frac{1}{n+1},$$

Since

$$\frac{1}{n+1} > \frac{1}{n(n+1)},$$

so when n increases, $\mid c_n \mid$ increases relatively faster than c_1. When $n = 2$, $c_1 =\mid c_n \mid$ and when $n = 3$,

$$c_1 = \frac{4}{18} <\mid c_n \mid= \frac{5}{18}.$$

Since we have proved that when n increases, $\mid c_n \mid$ increases relatively faster than c_1, we conclude that c_1 is always smaller than $\mid c_n \mid$ for $n \in [3, +\infty)$.

Proof of Proposition 12

We prove it by mathematical induction.

First-step. In a 2-contestant case with prizes $a_1 > a_2$, according to the Power CDF:

$$p_1 = \frac{e_1^{-r}}{e_1^{-r} + e_2^{-r}},\ p_2 = \frac{e_2^{-r}}{e_1^{-r} + e_2^{-r}},$$

where p_i is contestant i's probability of losing and e_i is contestant i's effort.

$$L_1 = (1 - p_1)a_1 + p_1 a_2 - e_1 = \frac{e_2^{-r}}{e_1^{-r} + e_2^{-r}} a_1 + \frac{e_1^{-r}}{e_1^{-r} + e_2^{-r}} a_2 - e_1,$$

$$L_2 = (1 - p_2)a_1 + p_2 a_2 - e_2 = \frac{e_1^{-r}}{e_1^{-r} + e_2^{-r}} a_1 + \frac{e_2^{-r}}{e_1^{-r} + e_2^{-r}} a_2 - e_2,$$

where L_i is contestant i's expected utility. Each contestant maximizes his revenue, in the symmetric pure-strategy equilibrium[43], player 1 and player 2 will make the

[43] In this paper, we restrict our attention to the unique symmetric pure-strategy equilibrium for simplicity, it remains to be investigated whether there could exist asymmetric equilibria.

same level of effort, i.e., $e_1 = e_2 = x$, by first order conditions:

$$x = \frac{r(a_1 - a_2)}{4},$$

so equation (81) holds when $n = 2$. We can see that the bigger $(a_1 - a_2)$ is, the more effort each contestant makes.

Existence requires

$$L^*_{1,2} = \frac{1}{2}(a_1 + a_2) - \frac{r(a_1 - a_2)}{4} \geq 0,$$

$$\Rightarrow r \leq \frac{2(a_1 + a_2)}{(a_1 - a_2)}. \tag{94}$$

So in the 2-contestant case to ensure the existence of the symmetric pure-strategy equilibrium, r has to be in certain region – (94) must hold.

Second-step. Assume that equation (81) is true in t-contestant case:

$$AE^l = \frac{r}{t}[(\frac{1}{t} + \frac{1}{t-1} + ... + \frac{1}{2} + 1 - 1)a_1 + ... + c_i a_i + ... + (\frac{1}{t} - 1)a_t], \tag{95}$$

where

$$c_i = -1 + \textstyle\sum_{j=0}^{n-i} \frac{1}{t-j}. \tag{96}$$

Then our task is to prove in $t + 1$ case,

$$AE^l = \frac{r}{t+1}[(\frac{1}{t+1} + \frac{1}{t} + ... + \frac{1}{2} + 1 - 1)a_1 + ... + c_i a_i + ... + (\frac{1}{t+1} - 1)a_{t+1}], \tag{97}$$

where

$$c_i = -1 + \textstyle\sum_{j=0}^{t+1-i} \frac{1}{t+1-j}. \tag{98}$$

In the symmetric pure-strategy equilibrium of the t-contestant case and (t+1)-contestant case, from any contestant's point of view, say contestant 1, given that all the other contestants are making effort x, his utilities are:

$$U^t = R^t_1 a_1 + R^t_2 a_2 + ... + R^t_i a_i + ... + R^t_n a_t - e_1,$$

$$U^{t+1} = R^{t+1}_1 a_1 + R^{t+1}_2 a_2 + ... + R^{t+1}_i a_i + ... + R^{t+1}_{t+1} a_t - e_1,$$

where U^t and U^{t+1} are the expected utilities for contest 1 in t-contestant case and (t+1) contestant case respectively. R_i^t and R_i^{t+1} are the probabilities of winning the i^{th} prize (a_i) for contestant 1 in a t-contestant case and (t+1)-contestant case respectively. Therefore, R_i^t is the probability that contestant 1 is chosen as the loser in the $(t+1-i)^{th}$ round (which means he has not been chosen as a loser in the previous $t-i$ rounds). So given all the other contestants making effort x, we derive

$$R_i^t = \frac{P_{t-i}^{t-1} x^{-(t-i)r} e_1^{-r}}{[e_1^{-r}+(t-1)x^{-r}][e_1^{-r}+(t-2)x^{-r}]...[e_1^{-r}+ix^{-r}][e_1^{-r}+(i-1)x^{-r}]}, \tag{99}$$

$$R_i^{t+1} = \frac{P_{t+1-i}^{t} x^{-(t+1-i)r} e_1^{-r}}{[e_1^{-r}+tx^{-r}][e_1^{-r}+(t-1)x^{-r}]...[e_1^{-r}+ix^{-r}][e_1^{-r}+(i-1)x^{-r}]}, \tag{100}$$

where

$$\begin{aligned} P_{t-i}^{t-1} &= \frac{(t-1)!}{(i-1)!} = (t-1)(t-2)...(i+1)i, \\ P_{t+1-i}^{t} &= \frac{t!}{(i-1)!} = t(t-1)...(i+1)i. \end{aligned}$$

In equilibrium, contestant 1 maximizes his utility,

$$\frac{\partial U^t}{\partial e_1} = 0 \Rightarrow \sum_{i=1}^{t} \frac{\partial R_i^t}{\partial e_1} a_i = 1. \tag{101}$$

In the t-contestant case,

$$\begin{aligned} (81) \quad &\Rightarrow \quad x = \sum_{i=1}^{t} \frac{r}{t}(-1+\sum_{j=0}^{t-i}\frac{1}{t-j})a_i \\ &\Rightarrow \quad \sum_{i=1}^{t} \frac{r}{tx}(-1+\sum_{j=0}^{t-i}\frac{1}{t-j})a_i = 1, \end{aligned} \tag{102}$$

$$(101) \text{ and } (102) \Rightarrow \sum_{i=1}^{t} \frac{\partial R_i^t}{\partial e_1} a_i = \sum_{i=1}^{t} \frac{r}{tx}(-1+\sum_{j=0}^{t-i}\frac{1}{t-j})a_i. \tag{103}$$

For any given values of $\{a_1, a_2, ..., a_t\}$, (103) always holds if and only if

$$\frac{\partial R_i^t}{\partial e_1} = \frac{r}{tx}(-1+\sum_{j=0}^{t-i}\frac{1}{t-j}). \tag{104}$$

Therefore, we conclude (81) holds for $n = t$ given (104) holds. Similarly, we can derive that (81) holds for $n = t + 1$ given

$$\frac{\partial R_i^{t+1}}{\partial e_1} = \frac{r}{(t+1)x}(-1 + \sum_{j=0}^{t+1-i} \frac{1}{t+1-j}). \tag{105}$$

So the our aim is to show (105) holds if (104) holds. The following is the proof.

From (99) and (100), we derive

$$R_i^{t+1} = \frac{P_{t+1-i}^t}{P_{t-i}^{t-1}}(\frac{x^{-r}}{e_1^{-r} + tx^{-r}})R_i^t$$

$$\Rightarrow \frac{\partial R_i^{t+1}}{\partial e_1} = \frac{P_{t+1-i}^t}{P_{t-i}^{t-1}}[\frac{e^{-1+r}rx^r}{(e^r t + x^r)^2}R_i^t + (\frac{x^{-r}}{e_1^{-r} + tx^{-r}})(\frac{\partial R_i^t}{\partial e_1})]. \tag{106}$$

Substituting $e_1 = x$ into (99):

$$R_i^t = \frac{(t-1)...ix^{-(t+1-i)r}}{t(t-1)...ix^{-(t+1-i)r}} = \frac{1}{t}. \tag{107}$$

Substituting (107), (104) and $e_1 = x$ into (106):

$$\begin{aligned}
\frac{\partial R_i^{t+1}}{\partial e_1} &= \frac{t(t-1)...i}{(t-1)(t-2)...i}[\frac{x^{-1+r}rx^r}{(x^r t + x^r)^2}(\frac{1}{t}) + (\frac{x^{-r}}{x^{-r} + tx^{-r}})\frac{r}{tx}(-1 + \sum_{j=0}^{t-i} \frac{1}{t-j})] \\
&= \frac{r}{x(1+t)}[\frac{1}{(1+t)} + (-1 + \sum_{j=0}^{t-i} \frac{1}{t-j})] \\
&= \frac{r}{x(t+1)}[-1 + \sum_{j=0}^{t+1-i} \frac{1}{t+1-j}],
\end{aligned}$$

which is exactly (105) and completes the proof.

3.7 Appendix 3B

(a) In the pure-strategy equilibrium of the Power CSF model,

$$\begin{aligned}
EU_i &= \frac{e_i^r}{e_i^r + (n-1)x^r}a_1 + \frac{(n-1)x^r e_i^r}{[e_i^r + (n-1)x^r][e_i^r + (n-2)x^r]}a_2 + ... \\
&+ \frac{(n-1)(n-2)...(n-t)x^{tr}e_i^r}{[e_i^r + (n-1)x^r][e_i^r + (n-2)x^r]...[e_i^r + (n-t-1)x^r]}a_{t+1} + ... \\
&+ \frac{(n-1)(n-2)...1 \times x^{(n-1)r}e_i^r}{[e_i^r + (n-1)x^r][e_i^r + (n-2)x^r]...[e_i^r + x^r]}a_n - e_i.
\end{aligned}$$

Defining $\lambda = \frac{e_i^r}{x^r}$, where

$$x = AE^s = \sum_{i=1}^n \frac{r}{n}(1 - \sum_{j=0}^{i-1} \frac{1}{n-j})a_i. \tag{108}$$

Then we can express EU_i as follows:

$$\begin{aligned} EU_i &= \frac{\lambda}{\lambda + (n-1)} a_{1+} \frac{(n-1)\lambda}{[\lambda + (n-1)][\lambda + (n-2)]} a_2 + ... \\ &+ \frac{(n-1)(n-2)...(n-t)\lambda}{[\lambda + (n-1)][\lambda + (n-2)]...[\lambda + (n-t-1)]} a_{t+1} + ... \\ &+ \frac{(n-1)(n-2)...1 \times \lambda}{[\lambda + (n-1)][\lambda + (n-2)]...[\lambda + 1]} a_n - e_i. \end{aligned}$$

Mathematically, EU_i can be written in the following form:

$$\begin{aligned} EU_i &= \{\frac{1}{n} \sum_{i=1}^n [(\lambda \prod_{t=0}^{i-1} \frac{n-t}{n-t+\lambda-1})a_i\} - e_i \\ &= \{\frac{1}{n} \sum_{i=1}^n [(\lambda \prod_{t=0}^{i-1} \frac{n-t}{n-t+\lambda-1})a_i\} - (\frac{e_i^r}{x^r})^{\frac{1}{r}} x \\ &= \{\frac{1}{n} \sum_{i=1}^n [(\lambda \prod_{t=0}^{i-1} \frac{n-t}{n-t+\lambda-1})a_i\} - \lambda^{\frac{1}{r}} x. \end{aligned}$$

Substituting (108) into the above equation, we have

$$EU_i = \frac{1}{n} \sum_{i=1}^n [(\lambda \prod_{t=0}^{i-1} \frac{n-t}{n-t+\lambda-1}) - r(1 - \sum_{t=0}^{i-1} \frac{1}{n-t})\lambda^{\frac{1}{r}}]a_i.$$

Hence, the expected utility is expressed as a function of λ. The first and second order conditions for a local interior maximum are the following. When $\lambda = 1$, the first order condition is:

$$\frac{1}{n} \sum_{i=1}^n [(\prod_{t=0}^{i-1} \frac{n-t}{n-t+\lambda-1})(1 - \sum_{t=0}^{i-1} \frac{\lambda}{n-t+\lambda-1}) - (1 - \sum_{t=0}^{i-1} \frac{1}{n-t})\lambda^{\frac{1}{r}-1}]a_i = 0.$$

But to make sure $\lambda = 1$ is a local interior maximum, we have to make sure the following condition holds:

$$sign\{\frac{\partial^2 EU_i}{\partial \lambda^2}|_{\lambda=1}\} = sign\{\sum_{i=1}^n a_i(1 - \sum_{t=0}^{i-1} \frac{\lambda}{n-t+\lambda-1})$$

$$\times[(1-\sum_{t=0}^{i-1}\frac{\lambda}{n-t+\lambda-1})-\frac{1}{r}-\frac{\sum_{t=0}^{i-1}\frac{\lambda}{n-t+\lambda-1}(1-\frac{\lambda}{n-t+\lambda-1})}{1-\sum_{t=0}^{i-1}\frac{\lambda}{n-t+\lambda-1}}]\}<0.$$

Therefore, to ensure the point $\lambda = 1$ is a local maximum,

$$r < \frac{\sum_{i=1}^{n} a_i(1-\sum_{t=0}^{i-1}\frac{1}{n-t})}{\sum_{i=1}^{n} a_i[(1-\sum_{t=0}^{i-1}\frac{1}{n-t})^2-\sum_{t=0}^{i-1}\frac{1}{n-t}(1-\frac{1}{n-t})]} = r^{\text{sec}} \tag{109}$$

if the denominator of r^{sec} is positive, and $r > r^{\text{sec}}$ when the denominator is negative[44].

Participation constraint requires

$$\begin{aligned} EU &= \frac{1}{n}\sum_{i=1}^{n} a_i - x \\ &= \frac{1}{n}\sum_{i=1}^{n} a_i - \frac{r}{n}\sum_{i=1}^{n} a_i(1-\sum_{t=0}^{i=1}\frac{1}{n-t}) \geq 0 \\ &\Rightarrow\ r \leq \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} a_i(1-\sum_{t=0}^{i-1}\frac{1}{n-t})}. \end{aligned} \tag{110}$$

To sum up, r has to be located in a certain region to ensure the existence of the symmetric pure-strategy equilibrium, i.e., (109) and (110) both hold simultaneously.

(b) In the pure-strategy equilibrium of the Power CDF model, by a similar procedure with (a), we show that to ensure the existence of the symmetric pure-strategy equilibrium, the following conditions must hold:

$$r < \frac{\sum_{i=1}^{n} a_{n+1-i}(-1+\sum_{t=0}^{i-1}\frac{1}{n-t})}{\sum_{i=1}^{n} a_{n+1-i}[(1-\sum_{t=0}^{i-1}\frac{1}{n-t})^2-\sum_{t=0}^{i-1}\frac{1}{n-t}(1-\frac{1}{n-t})]} = r^{\text{sec}} \tag{111}$$

[44] Notice that he numerator of r^{sec} is always bigger than zero since

$$\sum_{i=1}^{n} a_i(1-\sum_{t=0}^{i-1}\frac{1}{n-t}) = AE^s > 0.$$

if the denominator of r^{sec} is positive, and $r > r^{\text{sec}}$ when the denominator is negative[45]; and

$$r \leq \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} a_i[(\sum_{t=0}^{n-i} \dfrac{1}{n-t}) - 1]}. \tag{112}$$

[45]Notice that the numerator of r^{sec} is always bigger than zero since

$$\sum_{i=1}^{n} a_{n+1-i}(-1 + \sum_{t=0}^{i-1} \frac{1}{n-t}) = \sum_{i=1}^{n} a_i(-1 + \sum_{t=0}^{n-i} \frac{1}{n-t}) = AE^l > 0.$$

References

[1] Akerlof, Robert and Richard Holden (2007), "The Nature of Tournaments," working paper, Sloan School of Management, MIT.

[2] Alcalde, José and Matthias Dahm (2007), "All-Pay Auction Equilibria In Contests," Working Papers. Serie AD 2007-27, Instituto Valenciano de Investigaciones Económicas, S.A. (Ivie).

[3] Amann, Erwin and Wolfgang Leininger (1996), "Asymmetric all-pay auctions with incomplete information: The two-player case," Games and Economic Behavior, 14(1), 1-18.

[4] Amegashie, J. Atsu (2006), "A contest success function with a tractable noise parameter," Public Choice, 126, 135–144.

[5] Baik, Kyung Hwan (1994), "Effort levels in contests with two asymmetric players," Southern Economic Journal 61: 367–378.

[6] Baik, Kyung Hwan (1997), "Difference-form contest success functions and effort levels in contests," European Journal of Political Economy, 14, 685–701.

[7] Baik, Kyung Hwan (2004), "Two-Player Asymmetric Contests with Ratio-Form Contest Success Functions," Economic Inquiry, 42, 679-689.

[8] Barut, Y., Kovenock, D. (1998), "The symmetric multiple prize all-pay auction with complete information," European Journal of Political Economy, 14, 627-44.

[9] Baye, Michael R., Dan Kovenock and Casper G. de Vries (1994), "The solution to the Tullock Rent-Seeking Game when $R > 2$: Mixed-stategy equilibra and Mean Dissipation Rates," Public Choice, 81, 363-380.

[10] Baye, Michael R., Dan Kovenock and Casper G. de Vries (1996), "The All-Pay Auction with Complete Information," Economic Theory, 8(2), 291-305.

[11] Baye, Michael R., Dan Kovenock and Casper G. de Vries (1998), "A general linear model of contests," mimeo, Kelley School of Business, Indiana University, Bloomington.

[12] Berry, S. Keith. (1993), "Rent-seeking with multiple winners," Public Choice, 77, 437-443.

[13] Blavatsky, Pavlo (2004), "Contest Success Function with the Possibility of a Draw: Axiomation," IEW Working Paper No. 208.

[14] Che, Yeon-Koo and Ian Gale (1998), "Caps on political lobbying," American Economic Review, 88, 643-651.

[15] Che, Yeon-Koo and Ian Gale (2000), "Difference-form contests and the robustness of all-pay auctions," Games and Economic Behavior, 30, 22–43.

[16] Clark, Derek J. and Christian Riis (1996), "A multi-winner nested rent-seeking contest," Public Choice, 87, 177–84.

[17] Clark, Derek J. and Christian Riis (1996 a), "Competition over more than one prize," American Economic Review, 88, 276-89.

[18] Clark, Derek J. and Christian Riis (1996 b), "Contest success functions: an extension," Economic Theory, 11, 201-104.

[19] Clark, Derek J. and Christian Riis (1996 c), "Influence and the discretionary allocation of several prizes," European Journal of Political Economy, 14, 605-625.

[20] Clark, Derek J. and Christian Riis (1996 d), "On the win probability in rent-seeking games," University of Troms0 Discussion Paper in Economics.

[21] Clark, Derek J. and Christian Riis (2000), "Allocation efficiency in a competitive bribery game," Journal of Economic Behavior and Organization, 42(1), 109-24.

[22] Corchon, Luis C. (2000), "On the allocative effects of rent seeking," Journal of Public Economic Theory, 2, 483-92.

[23] Corchon, Luis C. (2007), "The theory of contests: a survey," Review of Economic Design, 11(2), 69-100.

[24] Corchon, Luis C. and Matthias Dahm (2007), "Foundations For Contest Success Functions," Economics Working Papers we070401, Universidad Carlos III, Departamento de Economía.

[25] Cornes, Richard and Roger Hartley (2005), "Asymmetric contests with general technologies," Economic Theory, 26(4), 923-946.

[26] Dasgupta, A. and K. O. Nti (1998), "Designing an optimal contest," European Journal of Political Economy, 14, 587–603.

[27] Dixit, Avinash Kamalakar (1987), "Strategic behavior in contests," American Economic Review, 77, 891-8.

[28] Frank, Robert H. (1985), "Choosing the Right Pond," Human Behavior and the Quest for status, Oxford University Press, Oxford.

[29] Frank, Robert H. and Philip. J. Cook (1995), "The Winner-take-all Society," Penguin Books.

[30] Fu, Qiang and JingFeng Lu (2007), "The Optimal Multi-Stage Contest," Working Paper Reviews with number 843644000000000387.

[31] Garfinkel, Michelle R. and Stergios Skaperdas (1996), "The Political Economy of Conflict and Appropriation," Cambridge University Press.

[32] Glazer, Amihai and Hassin, Refael (1988), "Optimal Contests," Economic Inquiry, 26(1), 133-43.

[33] Green, Jerry R. and Nancy L. Stokey (1983), "A Comparison of Tournaments and Contracts," Journal of Political Economy, 91(3), 349-64.

[34] Hillman, Arye L. and Eliakim Katz. (1984), "Risk-averse rent seekers and the social cost of monopoly power," Economic Journal, 94, 104-110.

[35] Hillman, Arye L. and Dov Samet (1987), "Dissipation of contestable rents by small numbers of contenders," Public Choice, 54, 63-82.

[36] Hillman, Arye L. (1989), "The political Economy of Protection," Harwood Academic Publishers.

[37] Hillman, Arye L., and John G. Riley (1989), "Politically contestable rents and transfers," Economics and Politics, 1, 17-40.

[38] Hirshleifer, J. (1989), "Conflict and rent-seeking success functions: Ratio and difference models of relative success," Public Choice, 63, 101–112.

[39] Huck, Steffen,.Kai A. Konrad and Mueller Wuller (2001), "Merger and collusion in contests," Journal of Institutional and Theroetical Economics, 158(4), 563-575.

[40] Konrad, Kai A. and Harris Schlesinger (1997), "Risk aversion in rent-seeking and rent-augmenting games," Economic Journal, 107, 1671-83.

[41] Konrad, Kai A (2007), "Strategy in Contests – An Introduction." WZB-Markets and Politics Working Paper No. SP II 2007-01.

[42] Krishna, Vijay and John Morgan (1997), "An analysis of the war of attrition and the all-pay action," Journal of Economic Theory, 72(2), 343-62.

[43] Lazear, Edward P and Sherwin Rosen (1981), "Rank-Order Tournaments as Optimum Labor Contracts," Journal of Political Economy, 89(5), 841-64.

[44] Loury, Glenn C. (1979), "Market structure and innovation," Quarterly Journal of Economics, 93(3), 395-410.

[45] Michaels, R. (1988), "The design of rent-seeking competitions," Public Choice, 56, 17-29.

[46] Moldovanu, Benny and Aner Sela (2001), "The Optimal Allocation of Prizes in Contest," American Economic Review, 91, 542-558.

[47] Moldovanu, Benny and Aner Sela (2006), "Contest Architecture," Journal of Economic Theory, 126(1), 70-96.

[48] Moldovanu, Benny; Aner Sela, and Xianwen Shi (2008), "Carrots and Sticks: Prizes and Punishments in Contests," April, 2008, CEPR Discussion Papers with number 6770.

[49] Myerson, Roger B. (1981), "Optimal Auction Design," Mathematics of Operations Research, 6(1), 58-73.

[50] Nitzan, S. (1994), "Modelling rent-seeking contests," European Journal of Political Economy, 10, 41–60.

[51] Nti, Kofi O. (1997), "Comparative statics of contests and rent-seeking games," International Economic Review, 38, 43-59.

[52] Nti, Kofi O. (1999), "Rent-seeking with asymmetric valuations," Public Choice, 98, 415-430.

[53] Nti, Kofi O. (2004), "Maximum efforts in contests with asymmetric valuations," European Journal of Political Economy, 20, 1059-1066.

[54] Perez-Castrillo, J. David. and Thierry Verdier (1992), "A general analysis of rent seeking games," Public Choice, 73, 335-350.

[55] Riley, John G. and William F Samuelson (1981). "Optimal Auctions," American Economic Review, 71(3), 381-92.

[56] Rosen, Sherwin (1986), "Prizes and Incentives in Elimination Tournaments," American Economic Review, 76(4), 701-15.

[57] Rubin, Paul H., Christopher Curran and John F. Curran (2001), "Litigation versus legislation: forum shopping by rent seekers," Public Choice, 107, 295-310.

[58] Shaked, Moshe and George J. Shantikumar (1994), "Stochastic orders and their applications," San Diego, CA: Academic Press, 1994.

[59] Skaperdas, Stergios and Li Gan (1995), "Risk aversion in contests," Economic Journal, 105, 951-62.

[60] Skaperdas, Stergios (1996), "Contest Success Functions," Economic Theory, 7, 283–290.

[61] Skaperdas, Stergios and Samarth Vaidya (2007), "Persuasion as a Contest," CESifo Working Paper Series, CESifo Working Paper No. 070809, CESifo GmbH.

[62] Stein, William E (2002), "Asymmetric Rent-Seeking with More Than Two Contestants," Public Choice, 113(3-4), 325-36.

[63] Szymanski, Stefan (2003), "The Economic Design of Sporting Contests," Journal of Economic Literature, 41(4), 1137-1187.

[64] Szymanski, Stefan and Tommaso M. Valletti (2005), "Incentive effects of second prizes," European Journal of Political Economy, 21, 467-481.

[65] Tullock, Gordon (1967), "The welfare cost of tariffs, monopolies, and theft," Western Economic Journal, 5, 224-232.

[66] Tullock, Gordon (1980), "Efficient rent seeking," In J.M. Buchanan, R.D. Tollison, and G. Tullock, Toward a theory of the rent-seeking society, 97–112. College Station: Texas A&M University Press.