THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Optimal Punishment in Contests with Endogenous Entry

OPEN ACCESS

Optimal Punishment in Contests with Endogenous Entry*

Jonathan P. Thomas†   Zhewei Wang‡

November, 2012

Abstract

We study optimal punishment in an all-pay contest with endogenous entry, where the participant with the lowest performance may be punished. When a small punishment is introduced, the lowest ability players drop out and those of medium ability exert less effort, while only the highest ability players exert more effort. A sufficient condition is given for the optimal punishment to be zero if the objective is to maximize the expected total effort. As cost functions become more convex, punishment becomes less desirable. When the objective is to maximize the expected highest individual effort, a positive punishment is desirable under much weaker conditions. (JEL C72, D72, D82)

Key words: endogenous entry, punishment, total effort, highest individual effort

†Jonathan Thomas: School of Economics, University of Edinburgh, 31 Buccleuch Place, Edinburgh, EH8 9JT, UK, Tel: 44(0)1316504515, Fax: 44(0)1316504514, Email: jonathan.thomas@ed.ac.uk.

‡Zhewei Wang: School of Economics, Shandong University, 27 Shanda Nanlu, Jinan, CHINA, 250100; Tel: (86)53188369992, Fax: (86)53188571371, Email: zheweiwang@sdu.edu.cn.

1   Introduction

In daily life, "carrots and sticks" refers to a policy of offering a combination of rewards and punishments to induce some desired behavior. 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). A possible reason why punishments have received little theoretical attention is that if players have to participate in a contest then it is trivial that introducing a punishment will be effective in increasing effort. That is, for a given group of players who have to participate, punishing the player who exerts the lowest effort level will increase the total effort for sure. In fact, punishments should be made as large as possible in order to maximize effort. However, adding a punishment, especially when the punishment is large, may violate individual rationality constraints, i.e., a player can find that his expected utility in equilibrium is below his outside option.

In this paper we assume that potential players observe the reward/punishment scheme before deciding whether or not to participate in a contest. We call this type of contest an *open* contest and consider whether punishments are desirable in this context. For example, a profession in which low performers lose their jobs—which can be regarded as a punishment—may discourage entry; is it the case that this in turn may lead to less competition among those who do enter, and so negate the positive effect on effort of a punishment mentioned above? Should essay contests announce only the winners, or should they announce the entire ranking, subjecting the worst performers to potential humiliation? Should promotion contests where employees can choose whether or not to participate only announce the winner, or

1

would it lead to a better top candidate if the bottom candidates were penalized in some way?

To make progress on this sort of question, we analyze whether punishments are a useful incentive mechanism for increasing effort in an open, perfectly discriminating contest (where efforts exerted are perfectly observable to the contest designer), where players differ by ability (cost of effort), which is private information. We assume that there is a fixed prize for the highest effort, but that the contest designer can choose to impose a punishment on the lowest performer. The punishment neither consumes resources nor yields resources to the designer. We build on the seminal model of Moldovanu and Sela (2001) which explains prize structures in contests within the framework of private value all-pay auctions.

Our results can be summarized as follows. If the contest designer wants to maximize the total effort from all potential players, the optimal punishment will be zero for a wide class of cases (a positive optimal punishment can only occur when high ability players are relatively probable). As cost functions become more convex, starting from linear costs, the optimal punishment decreases, i.e., punishment becomes less desirable. If the contest designer seeks only to maximize the effort of the top player, a strictly positive punishment should be set under weaker conditions, and certainly if there are a sufficient number of players.

Our work is closely related to Moldovanu, Sela and Shi (2010), who also look at punishments in perfectly discriminating contests. In one section of their paper, a model in which players can choose whether or not to participate is also analyzed.[1]

---

[1]Moldovanu, Sela and Shi (2010) also consider a range of other scenarios, starting with situations where punishments can only be administered at a cost (subject to a fixed budget), and where players have to participate. If only punishments can be administered, they establish under a likelihood ratio condition that using all resources on a single punishment on the worst performer is optimal. If both rewards and punishments are feasible, then resources may be expended on a single

Their result appears to contradict the corresponding result in our paper: In order to maximize expected total effort, they find that a strictly positive punishment is always optimal while we find the optimal punishment is zero in a wide range of cases.[2] The reason for this difference lies in the assumption about the support of the distribution of the marginal cost of effort. Our paper follows the assumptions of Moldovanu and Sela (2001) in assuming that this distribution has positive and *bounded* support. By contrast, in Moldovanu, Sela and Shi (2010) the inverse of the marginal cost (denoted by a) is assumed to be distributed on the interval $[0, 1]$, so the marginal cost (i.e., $1/a$) is distributed on the interval $[1, +\infty)$. Our results show that with a bounded support $[\underline{s}, \overline{s}]$ (where $\overline{s} > \underline{s} > 0$), the desirability of a punishment depends critically on the shape of the cost parameter distribution.[3] In this respect, our results can be seen as complementary to theirs, and we would argue that in practice a bounded support is often realistic. For example, in contests involving professionals, the support of the ability distribution is typically bounded due to prior constraints on entry to the profession. Thus, in these situations, punishment is likely to be undesirable. This seems to be more consistent with what we observe in reality: explicit punishment is rarely used in open contests.

Intuitively, introducing a punishment has two effects. Firstly, a selection effect: some players will drop out, and these will be those towards the bottom of the ability range who are likely to lose anyway. This leads to the competition between the actual participants becoming less fierce since fewer players are involved. Those

---

prize/punishment depending on the distribution of abilities. Alternatively, if the size of punishment is fixed, and punishments are costless, they characterize the optimal number of punishments.

[2] See Proposition 3 and Example 1 for details.

[3] Replacing $[\underline{s}, \overline{s}]$ with $[1, +\infty)$, our model (with linear cost functions) would be exactly the same as that in Moldovanu, Sela and Shi (2010). We confirm this equivalence in Section 3.1 below. While Moldovanu, Sela and Shi (2010) focus on the case with linear cost functions, as already mentioned with convex cost functions we find that, in order to maximize expected total effort, punishment becomes less desirable when the cost function becomes more convex.

who participate but are near the nonparticipation threshold will put in less effort, since they anticipate being beaten by higher ability players (at the threshold, effort must be zero). Secondly, there is an incentive effect due to the desire to avoid the punishment. The two effects occur at the same time. We show that when a sufficiently small punishment is introduced, the low ability players drop out and the medium ability players exert less effort, while only the high ability players exert more effort. This explains our contrasting results. When a punishment is introduced, expected total effort is likely to fall because of the loss of the lowest ability players and the fact that the medium ability ones exert less effort. On the other hand because the highest ability players exert more effort, the expected highest individual effort will increase for a wider range of ability distributions.

An entry fee (or minimum-effort requirement) is in some respects similar to a punishment in that it also excludes low-ability players from a contest. Higgins, Shughart and Tollison (1985) study a contest where there is a fixed entry cost for everyone and contestants enter randomly in equilibrium. In an all-pay auction model, Kaplan and Sela (2010) provide a rationale for entry fees in contests by analyzing a two-stage model[4] with privately known entry costs. Fu and Lu (2010) investigate an imperfectly discriminating contest where the potential contestants bear fixed entry costs and the contest designer has a fixed budget with two strategic instruments: the prize purse and monetary transfers (subsidy/fee). Fu, Jiao and Lu (2011) study imperfectly discriminating contests with endogenous and stochastic entries.[5] They show that the designer may benefit from noisier contests and prefers to invite only a subset of potential contestants to participate. Finally there has

---

[4]In the first stage, potential players make entry decisions given entry costs being privately known; in the second stage, participants make efforts (bids) after finding out who else has entered.

[5]Myerson and Warneryd (2006), Munster (2006), Lim and Matros (2009) and Fu, Jiao and Lu (2010) also examine contests with stochastic participation.

been some recent experimental work on the contest entry decision, although not with explicit punishments, see Cason, Masters and Sherementa (2010) (where there is a positive outside option which is lost upon entry, so this is similar to an entry fee) and Morgan, Orzen and Sefton (2010).

The difference between an entry fee and a punishment should be emphasized. First, with an entry fee, all participants have to suffer some cost 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 is total effort maximizing among all feasible mechanisms that are incentive compatible and individual rational (Myerson, 1981; Riley and Samuelson, 1981), while in this paper, with no entry fee, we show that for the same objective function, a punishment on the worst performer is often not desirable.

In a seminal paper of a large literature on contests (or tournaments), Lazear and Rosen (1981) argue that rank-order contests help to solve a moral hazard problem. In a Lazear-Rosen contest, Nalebuff and Stiglitz (1983) discuss (among other matters) how punishments can affect the global incentive compatibility condition—ensuring that contestants are not better off choosing zero effort over the effort identified by the usual marginal conditions. More recently, Gilpatric (2009) considers how the balance of prizes and punishment affect risk-taking in a Lazear-Rosen tournament, showing that adding a punishment enables the contest designer to control contestants' incentives to exert effort and to alter output variance according to the designer's aims.[6] Akerlof and Holden (2010) extend Lazear and Rosen's (1981) analysis to the case

---

[6]For example, as here, Gilpatric (2009) analyzes two possible aims of the contest organizer: maximizing total effort when she values all contestants' effort equally, or maximizing highest individual effort when she only values the highest of the contestants' efforts.

with multiple prizes and show that it is often optimal to give rewards that differ between top performers by a smaller magnitude than the corresponding punishments to poor performers. We stress that the context of the above papers is very different from the one which we deal with: they focus on the symmetric case where all players are homogeneous but effort and performance is stochastically related, whereas we look at a perfectly discriminating contest with endogenous entry where heterogeneous contestants have private information on their abilities. In a setting of perfectly discriminating contests (as in Moldovanu and Sela, 2001, and this paper), Minor (2012) shows that with strictly convex costs, having an inverted reward structure—in which a larger prize goes to second place than to first place—may be optimal as the less able are more incentivized. Likewise punishment, considered here, leads to a "steep" reward structure and may create adverse incentives, the more so as convexity of costs increase. In this sense, Minor's results are consistent with ours.

The contest literature has mostly focussed on the case of maximizing expected total effort. However in practice, the contest designer may not value all contestants' efforts equally, and may care more about the performance of the top (one or several) contestants. Given this motivation, and as mentioned earlier, we also analyze what the optimal punishment would be when the contest designer seeks to maximize the expected highest individual effort. For example, in the research contests studied by Taylor (1995), the contest designer will only use the best submission from among all contestants. In sporting competitions, the contest designer may be interested only in the performance of the top player(s). Levitt (1995) argues that in many contexts where multiple players are assigned to a task, only one of their outputs will be used: This is especially true of creative endeavors such as the development of advertising campaigns. Another example is suggested by Gilpatric (2009): "If one considers

6

a group of junior faculty competing to win tenure, the department may value the output of all contestants, but the output of the winners may be valued more than that of the losers because the winners will be retained and their output will provide greater ongoing reputation benefits to the department."

The remainder of the paper proceeds as follows. In section 2, after setting up the model formally with a linear cost function, we derive a symmetric equilibrium where the effort levels of participants in the contest are characterized by an equilibrium effort function. By analyzing this equilibrium effort function, we elaborate on what happens when a small or large punishment is introduced. In sections 3 and 4, we discuss what the optimal punishment should be when maximizing the expected total effort and the expected highest individual effort respectively. In addition, the relationship of our work to that of Moldovanu, Sela and Shi (2010) is analyzed in section 3.1. In section 5, we extend our previous analysis (with linear cost functions) to the convex cost case. Concluding remarks are provided in section 6.

2   The Model

There are $k \geq 3$ potential players in a perfectly discriminating contest with a fixed[7] prize $V > 0$. Assuming there is at least one participant, 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, $0 \leq P \leq V$, which is a choice variable of the contest designer. If only one player participates in the contest, he receives the prize and the punishment

---

[7]We assume the prize is simply fixed in value, and it is indivisible. While we do not show that with divisibility it is still optimal to have a single prize, Moldovanu and Sela (2001) show that, with linear and concave cost functions, it is optimal to allocate the entire prize sum to a single first prize in order to maximize the expected total effort.

at the same time.[8]

Ex ante all potential players choose simultaneously whether or not to enter this contest, and (at the same time) conditional on entry, player i chooses an effort level x_i.[9] Effort level x_i causes player i a disutility of $c_i x_i$, where c_i denotes player i's (constant) marginal cost of effort, which is private information. Parameter c_i is also called the ability parameter of player i, a low c_i indicating a 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 player maximizes expected utility given the values of the prize and the punishment. We assume that if a potential contestant chooses not to enter the contest, he receives an outside option of 0. Thus, for each player, the participation constraint requires his (ex ante) expected utility to be non-negative. 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.[10]

2.1   The Objective Function and Entry Decision

Given the commonly known values of V and P, a participant (who chooses to enter the contest) with ability parameter c, solves the following problem by choosing effort

---

[8]If more than one player exerts the highest (lowest) effort, the prize (punishment) is randomly allocated among them. In the equilibrium we study this happens with zero probability.

[9]Take an essay contest for example: students have to submit their essays by the deadline, so they do not know the number of participants until after the deadline.

[10]We assume that the contest designer only focuses on effort levels and does not get any material benefit or cost directly from the prize or the punishment. The value of the punishment thus cannot be used to finance the prize.

level x:

$$Max_{x}\{V \times \Pr(x \text{ is the highest}) - P \times \Pr(x \text{ is the lowest}) - cx\}.$$

We look for an equilibrium such that players with $c \in [\underline{s}, e)$ participate in the contest and every player exerts effort according to a strictly decreasing differentiable equilibrium effort function $x = b(c)$ when $c \in [\underline{s}, e)$. Players with $c \in [e, \overline{s}]$ do not participate in the contest.

A player with $c = e$ is indifferent between participating in the contest or not; we refer to such a player as the *marginal player*. If he enters he will exert zero effort, $b(e) = 0$. This follows as the marginal player has the lowest effort of any entrant: in equilibrium he will lose against all other entrants with probability one, so if he was putting in positive effort a deviation to zero effort would be profitable. He anticipates being punished with probability one which is exactly offset by the chance he is the only entrant, in which case he would win the prize. So the marginal player's expected utility is:

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

which implies

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

Players with $c \geq e$ are indifferent about entering and setting $e = 0$; we consider only equilibria in which they do not enter.[11] Equation (1) implies that the larger P is, the smaller $F(e)$ is, and so the smaller e is, i.e., fewer players would enter the contest. In particular, if the contest designer sets the punishment to the same value as the prize, i.e., $P = V$, then from (1), $1 - F(e) = 1$ so that $F(e) = 0$ and $e = \underline{s}$.

---

[11]Note that if there was a positive measure of zero-effort entrants, increasing effort a tiny amount would be a profitable deviation, so we can rule out symmetric pure-strategy equilibria with higher cost agents $(c > e)$ also entering.

9

Consequently no player will enter.[12] Only when $P < V$ do potential entrants exist and exert positive effort.

If a player's ability parameter is c, the probability of another player's ability parameter being smaller than c is $F(c)$. Moreover by the fact that entrants' effort is strictly decreasing in c, a participant who makes an effort x in equilibrium has ability $c = b^{-1}(x)$. Then, given the equilibrium behavior of other competitors, a player who enters the contest solves the following problem:[13]

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

where $[1 - F(b^{-1}(x))]^{k-1}$ is the probability that all other potential players exert less effort than x and $[F(b^{-1}(x)) + (1 - F(e))]^{k-1}$ is the probability that all other players either exert more effort than x or do not participate in the contest.[14]

2.2   The Equilibrium

Proposition 1 *In a symmetric equilibrium with prize V and punishment P where $0 \le P \le V$, players with $c \in [e, \ \overline{s}]$ do not participate in the contest, while players with $c \in [\underline{s}, \ e)$ participate in the contest and exert effort according to the following strictly decreasing equilibrium effort function:*

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

where e satisfies (1).

---

[12]This is intuitive, since otherwise with the value of the punishment being equal to the value of the prize, by collecting the punishment from the bottom participant and awarding it to the top participant, the contest designer could get a positive total effort for no cost.

[13]Note that the objective function (2) is relevant only for types $c < e$, i.e., the players who actually participate in the contest.

[14]In equilibrium, for player i, $\Pr(x_i$ is the lowest) equals the probability that every other player's type resides on the interval $[\underline{s}, c_i) \cup [e, \overline{s}]$, i.e., every other player either participates in the contest with an effort higher than x_i, or does not participate in the contest.

Proof. See Appendix. ∎

From (3), we can get

$$b'(c) = -\frac{(k-1)f(c)}{c}\{V[1-F(c)]^{k-2} + P[F(c)+1-F(e)]^{k-2}\}. \qquad (4)$$

As $V, P > 0$ and $b'(c) < 0$, the equilibrium effort function is strictly decreasing in c, i.e., the more able a participant is, the higher the effort he exerts in equilibrium.[15]

2.3   Introducing a Small Punishment

With zero punishment, $e = \bar{s}$ and $F(e) = F(\bar{s}) = 1$, so by (4) we obtain

$$b'(c)|_{P=0} = -\frac{(k-1)f(c)}{c}V[1-F(c)]^{k-2}. \qquad (5)$$

When a punishment $P > 0$ is introduced, we can write (4) as

$$b'(c)|_{P>0} = b'(c)|_{P=0}\underbrace{-\frac{(k-1)f(c)}{c}P[F(c)+1-F(e)]^{k-2}}_{<0}. \qquad (6)$$

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

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

We can interpret the slope of $b(c)$ as the degree of *relative* competition between participants, so this shows that a positive punishment leads to more intense relative competition. This is what we referred to as an incentive effect in the introduction. However, it does not follow that participants will exert more effort than before. Marginal participants, that is with abilities close to e, will exert less effort, and

---

[15]From (4), it also follows that a consolation prize (a negative punishment, $P < 0$) for the bottom player will never be optimal as everyone exerts less effort compared to the case with $P = 0$ ($e = \bar{s}$ for $P \leq 0$, so there is no gain from increased participation). To the extent they exist in the real world, it could be argued that there may be a psychic loss for the bottom participant from being revealed as the loser; therefore, a consolation prize to cancel out this "punishment" would be optimal whenever $P = 0$ is optimal in the corresponding model with zero psychic costs.

it may even be that all players exert less effort because fewer players participate. The following proposition summarizes the relative competition effect, and conditions under which some players exert more effort and those under which all players exert less effort.

Proposition 2 *(a) The equilibrium effort function $b(c)$ becomes steeper as P increases, that is, $b'(c)|_{P_1} < b'(c)|_{P_2}$ for $P_1 > P_2$ and every c such that $b(c)|_{P_1} > 0$. (b) The two equilibrium effort functions $b(c)|_{P>0}$ and $b(c)|_{P=0}$ either cross once or do not cross at all. For P sufficiently small they will cross once, while for P sufficiently large they do not cross. When they cross once, say at point $c = c^*$, $b(c)|_{P>0} > b(c)|_{P=0}$ for every $c \in [\underline{s},\, c^*)$ and $b(c)|_{P>0} < b(c)|_{P=0}$ for every $c \in (c^*,\, \overline{s})$. When they do not cross, $b(c)|_{P>0} < b(c)|_{P=0}$ for every $c \in [\underline{s},\, \overline{s}]$.*

Proof. See Appendix. ■

In the proof of (b) it is shown that if the contest designer introduces a (sufficiently) small punishment into an open contest, players with the highest ability (lowest values of c) will increase their effort. Because it is always the case that some low ability (high c) players drop out when $P > 0$ and the effort function is steeper when positive, this means the effort functions must cross once. This is illustrated in Figure 1 (where $b_1(c)$ corresponds to $P = 0$ and $b_2(c)$ to some small $P > 0$). The players with $c \in [\underline{s},\, c^*)$, whom we call the high ability players, will exert more effort; the players with $c \in (c^*,\, e]$, whom we call the medium ability players, will exert less effort; and the players with $c \in [e,\, \overline{s}]$ whom we call the low ability players, will drop out.

However, when the punishment is (sufficiently) large, as in Figure 2 (where $b_1(c)$ corresponds to $P = 0$, and $b_2(c)$ now to some large $P > 0$), all participants will exert

less effort than before since too many players drop out, i.e., $b(c)|_{P>0} < b(c)|_{P=0}$ for all potential values of c.
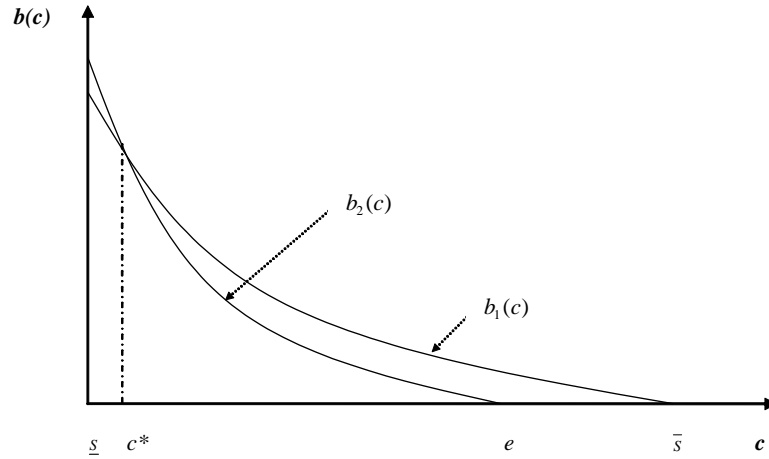

Figure 1 Effort functions when a small punishment is introduced
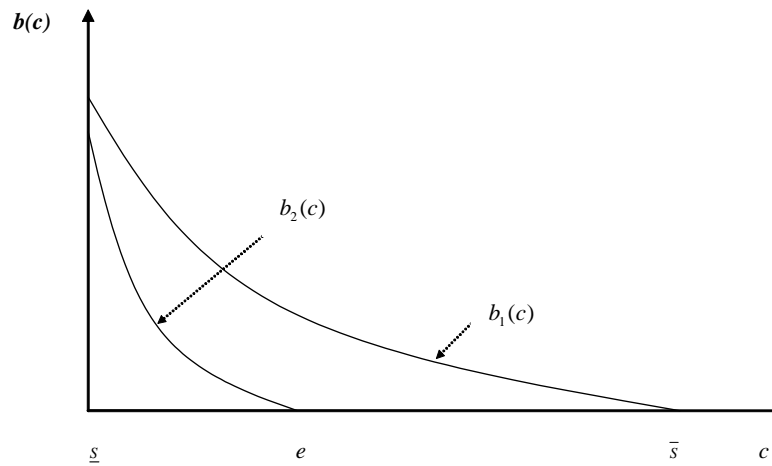

Figure 2 Effort functions when a large punishment is introduced

13

This characterization leaves open the question of whether a positive punishment is desirable. Even though the effort of the most able unambiguously rises with the introduction of a relatively small P, it does not follow that the *expected highest effort* rises since the most able may not be present in a given population of players. Moreover if the contest designer is interested in the *sum* of efforts, even if the most able are present, the fact that when $P > 0$ others reduce their effort or do not participate, implies that a positive punishment is even less likely to be desirable.[16] We now turn to analyze this question in more detail.

3   Maximizing Expected Total Effort

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

In equilibrium, the expected average effort (AE) of each potential player is given by

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

We have shown that there is an equilibrium effort function $x = b(c)$ which is strictly decreasing for participants with $c \in [\underline{s},\, e)$, and $b(c) = 0$ for all players with $c \geq e$. There are k potential players, so from (3) the expected total effort (TE) is

$$TE := k \times AE = \int_{\underline{s}}^{\overline{s}} b(c)f(c)dc = k(k-1)R_1, \tag{9}$$

---

[16] Although either way it is clear that a large punishment is never optimal.

where

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

Maximizing TE is equivalent to maximizing R_1. In the appendix, we prove the following result by analyzing (10):

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

Proof. See Appendix. ■

Proposition 3 states that a non-decreasing density constitutes a sufficient condition for optimal punishment being zero.[17] This is not a necessary condition, however, and the optimal punishment can be zero with a decreasing density distributions.

When $f(c)$ is non-decreasing (i.e., increasing or staying constant) with c, the contest designer anticipates relatively few high ability players. Then, adding even a small punishment, which will exert the low ability players drop out and medium players exert less effort, will decrease the expected total effort.

When $f(c)$ is decreasing with c, to maximize expected total effort, the optimal punishment may still be zero (see Example 1) or strictly positive (see Example 2). A decreasing density function implies that the contest designer expects there to be a relatively large number of high ability players. Since their effort levels respond positively to a small punishment, in order to maximize total effort punishment may be desirable.

---

[17]In this section, we focus on monotone density functions. We do not have general results with non-monotone density functions.
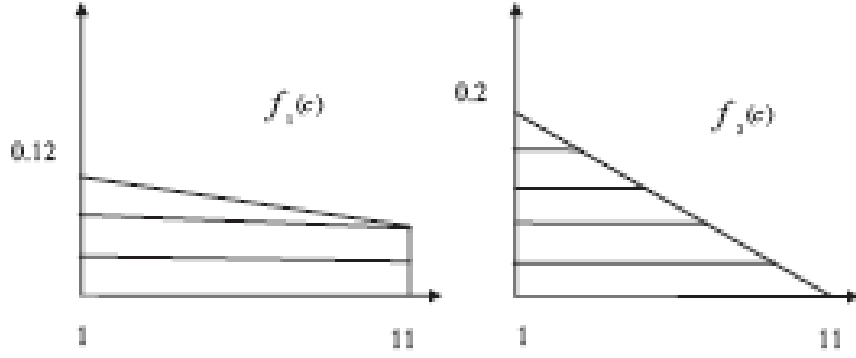
Figure 1: Figure 3 Density Functions in Example 1 [LEFT] and Example 2 [RIGHT].

Example 1 *Let $V = 1$, $k = 3$, $\underline{s} = 1$, $\bar{s} = 11$. and consider the (linear) density function $f_1(c) = (31 - c)/250$, which is strictly decreasing in c on $[1, 11]$. In this case $e = 31 - 10\sqrt{4 + 5P^{1/2}}$. It can be shown that $dTE/dP < 0$ for any $P \in [0, 1)$, and the optimal punishment is zero.*

Example 2 *Let $V = 1$, $k = 3$, $\underline{s} = 1$, $\bar{s} = 11$ and consider the (linear) density function $f_2(c) = (11 - c)/50$, which is also strictly decreasing with c on the interval $[1, 11]$. In this case $e = 11 - 10P^{1/4}$, and TE is maximized when $P \approx 0.011$. Thus, the optimal punishment is strictly positive.*

Note that in Example 2 (RIGHT in Figure 3) $f_2(c)$ is decreasing in c at a faster rate compared to $f_1(c)$ in Example 1 (LEFT in Figure 3), which is consistent with the intuition given above.[18]

---

[18]Examples 1 and 2 can be established analytically. In general with $k = 3$ and support $[1, 11]$, and assuming the density function is linear with a slope a (so that $a \in [-0.02, 0.02]$ which ensures that the density function is always strictly positive on the interior of the support $[1, 11]$), and grid step of 0.001, numerical simulations show that, the optimal punishment $P^* > 0$ when $-0.020 \leq a \leq -0.006$ and $P^* = 0$ when $-0.005 \leq a \leq 0.020$. Examples 1 and 2 are then two special cases of those simulations with $a = -0.004$ and $a = -0.02$ respectively.

3.1 Relationship to Moldovanu, Sela and Shi (2010)

Moldovanu, Sela and Shi (2010) analyze a similar situation to the above.[19] They prove that to maximize expected total effort the optimal punishment is always strictly positive. This seems to contradict our above result that when $f(c)$ is non-decreasing in c, the optimal punishment is zero, and even when $f(c)$ is decreasing in c, the optimal punishment may still be zero (see Example 1). As discussed in the introduction, however, translated into our model they assume that the density function for c must be positive everywhere on the interval $[1, +\infty)$.

If we let the support of F be $[1, +\infty)$ instead of $[\underline{s}, \overline{s}]$, our model would be the same as that in Moldovanu, Sela and Shi (2010), and we get the following:

Case 1 *(Proposition 7 of Moldovanu, Sela and Shi, 2010) When the support of F is $[1, +\infty)$,*

$$\frac{dTE}{dP}\Big|_{P=0} = k(k-1)\int_1^{+\infty} \frac{1}{t}[F(t)]^{k-1}dF(t) > 0.$$

Therefore the optimal punishment is strictly positive.

Proof. See Appendix. ∎

If the support of F is $[1, +\infty)$ then this *excludes* the possibility that the Proposition 3 condition holds that $f(c)$ should be non-decreasing with c on the support since $f(c)$ must be decreasing as $c \to +\infty$ given $\int_1^{+\infty} f(c)dc = 1$; consequently the two results are not in fact in conflict.

To get some rough intuition, consider starting with a finite support for F, and suppose a small punishment $\widetilde{P}$ is introduced. By (1), this fixes $F(e)$ and hence e. As argued above, the effect of introducing $\widetilde{P}$ is that this increases the effort

---

[19]See section 4.2 of their paper.

levels of the most able while reducing effort levels of those with costs close to (but below) e. Moreover all players with costs above e drop out. As we have seen, the benefit of introducing $\widetilde{P}$ may be positive or negative depending on F. Suppose now that we change F by increasing its support (letting $\bar{s}$ increase) and "stretching" the distribution across this wider support, but leaving F unchanged for $c \leq e$. Clearly, the equilibrium of the game with punishment $\widetilde{P}$ is unchanged as exactly the same players participate as before (e is unchanged). However the benefit of introducing $\widetilde{P}$ is different now: when $P = 0$ the players with $c > e$ are likely to exert very low levels of effort as they mostly have high values for c. Moreover even the players with c close to e will have very low levels of effort because there is effectively no competition from players with lower ability (see (3)). So when $\widetilde{P}$ is introduced, not only is the cost of players with $c > e$ dropping out very small, but also the drop in effort made by those close to e is also small. Effectively what we called the selection effect becomes insignificant, and the incentive effect of the punishment on the higher ability players dominates for a sufficiently stretched support. The benefit of introducing $\widetilde{P}$ will thus become positive.

In other words, c being distributed on $[1, +\infty)$ with $f(c) > 0$ implies, from the contest designer's point of view, the weakest (possible) players are always a group of extremely low ability players (with $c = 1/a \to +\infty$ as $a \to 0$), so starting from a situation without punishment, introducing a small punishment will make these extremely low ability players drop out and the high ability players exert more effort. Because those players with extremely low abilities exert little (almost zero) effort in the situation without punishment, the selection effect is dominated by the incentive effect. Therefore, the expected total effort increases after the introduction of an appropriately small punishment.

4 Maximizing Expected Highest Individual Effort

Instead of maximizing expected total effort, in this section, we focus on the case where the contest designer wants to elicit the highest individual effort. As we mentioned in the introduction, in many contexts, such as research contests and contests among creative endeavors, the contest designer may only care about the best submission from among all contestants, i.e., she seeks to maximize the expected highest individual effort. Or in our previous example, assume now the university only needs the best essay from its students, with all essays of a lesser quality than the best being of no interest. Even though we have seen that a positive punishment will raise the effort of the highest ability players, this does not mean that the expected highest effort will increase as it may be that all k players have abilities below the critical level above which effort increases (i.e., with c above c^* as defined in Proposition 2). Nevertheless given that it only the highest effort level that matters, we will find that there are more circumstances under which a positive punishment is called for compared to the previous case.[20]

Rank the players' ability parameters as follows: $c_1 < c_2 < ... < c_k$, so c_1 is the most able player. First consider $G_1(c)$, defined as the distribution function of c_1. The probability that all potential players are less able than type c, is $(1 - F(c))^k$, then the probability that at least one player is more able than c is $1 - (1 - F(c))^k$. Therefore,

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

---

[20]For a given P, the punishment is ex post beneficial if the most able player, type c_1, is more able than type c^*, i.e., $c_1 < c^*$. This is more likely to occur than the total effort being ex post higher (it follows from Proposition 2 that $c_1 < c^*$ is a necessary but not sufficient condition for the total to be higher). This suggest that the ex ante comparison will go the same way.

Hence, the probability density function of c_1 is

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

Therefore, the expected highest individual effort can be expressed as

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

where

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

(11)

Proposition 4 *In an open contest with $k \geq 3$ players, given a distribution function F, there exists a number k^* such that for any number of players $k > k^*$, the optimal punishment is always strictly positive when the contest designer's aim is to maximize the expected highest individual effort.*

Proof. See Appendix. ■

In Proposition 4, a sufficiently large k ensures that the optimal punishment is strictly positive. Intuitively, when the number of potential players is sufficiently large, the chance of the top player being a high ability player will be close to one, in which case a strictly positive punishment is optimal.

Allowing the density function $f(c)$ to take any form, the proposition gives a relatively strong condition on the number of potential players ($k > k^*$) to guarantee a positive optimal punishment. For specific forms of $f(c)$, k^* need not be large, and may not bind at all. For example:

Case 2 *In an open contest where abilities are drawn from a uniform distribution on $[\underline{s}, \overline{s}]$, i.e., $f(c) = 1/(\underline{s} - \overline{s})$, and when $\overline{s}/\underline{s} \geq 1.47$, then for any $k \geq 3$ it is optimal to set a strictly positive punishment to maximize the expected highest individual effort.*

20

Proof. See Appendix. ∎

The requirement that the most able player is at least 1.47 times as efficient as the least possible able player seems to be fairly mild for practical applications.[21]

Thus, when abilities are drawn from a uniform distribution and $\overline{s}/\underline{s} \geq 1.47$, to maximize the expected highest individual effort the optimal punishment is strictly positive, while by Proposition 3, to maximize expected total effort the optimal punishment is zero.

5   Strictly Convex Costs

So far we assumed a linear cost function. In this section, we look at the case with a strictly convex cost function.[22] This is arguably a more realistic assumption.[23] We look at the same model described in Section 2 with the only difference that we assume now that an effort x will cause a player with ability c a disutility of $c\gamma(x)$. Assume $\gamma(0) = 0$, $\gamma' > 0$ and $\gamma'' > 0$, so the cost function $c\gamma(x)$ is convex. Let g be the inverse function of γ, i.e., $g := \gamma^{-1}$, then it is straightforward to show that $g' > 0$ and $g'' < 0$. The following can be obtained by a simple transformation of the equilibrium strategies we found in the linear case.

Proposition 5 *In a symmetric equilibrium with prize V and punishment P where $0 \leq P \leq V$, players with $c \in [e, \overline{s}]$ do not participate in the contest, while players with $c \in [\underline{s}, e)$ participate in the contest and exert effort according to the following*

---

[21] When $\overline{s}/\underline{s} < 1.47$, the optimal punishment can be zero or positive depending on k.

[22] Henceforth we call this the convex case, and similarly we call the case with a linear cost function the linear case.

[23] Though Moldovanu and Sela (2001) consider linear, concave and convex cost functions, they argue that the convex case is the most applicable.

strictly decreasing equilibrium effort function:

$$B(c) = g[b(c)], \tag{12}$$

where e satisfies (1) and $b(c)$ is the equilibrium effort function in the linear case, which is defined by (3).

Proof. See Appendix. ∎

Given $g' > 0$, equation (12) implies that our previous results (in Proposition 2 (b)) on the ranking of effort functions as P changes still hold. In particular, $B(c)|_{P>0} >, =, < B(c)|_{P=0}$ when $b(c)|_{P>0} >, =, < b(c)|_{P=0}$.

Let TE^X denote the expected total effort in the convex case (where the superscript X refers to the case with *convex* cost functions). Thus,

$$TE^X = k \int_{\underline{s}}^{e} g(b(c))f(c)dc. \tag{13}$$

From (13),

$$\frac{dTE^X}{dP} = k \int_{\underline{s}}^{e} g' \frac{db(c)}{dP} f(c)dc. \tag{14}$$

By Proposition 2, there are two possible cases regarding the sign of of $\frac{db(c)}{dP}$: either $\frac{db(c)}{dP} < 0$ for all c or there exists a c^*, $\underline{s} < c^* < e$, such that $\frac{db(c)}{dP} > 0$ for small c ($c < c^*$) and $\frac{db(c)}{dP} < 0$ for large c ($e > c > c^*$). In the former case, $\frac{dTE^X}{dP}$ defined by (14) will be negative as $g' > 0$. In the latter case, $g'' < 0$ and $b(c)$ decreasing imply that the negative terms of $\frac{db(c)}{dP}$ in the integral defining $\frac{dTE^X}{dP}$ are multiplied by higher values of g' than the positive terms. Thus, other parameters held constant, $\frac{dTE^X}{dP}$ is negative if $\frac{dTE}{dP}$ (in the linear case) is negative, and $\frac{dTE^X}{dP}$ will be negative for some parameters even when $\frac{dTE}{dP}$ is positive. Thus Proposition 3 extends to the convex case, as asserted in Propositon 6 (i) below. However, since convexity of the cost function enlarges the set of parameters for which $\frac{dTE^X}{dP}$ is negative when $\frac{dTE}{dP}$ is

22

positive, it will be optimal to set $P = 0$ in more situations. Indeed, for γ sufficiently convex (e.g., take $\gamma(x) = x^\alpha$, $\alpha > 1$, and let $\alpha \to \infty$), the weight placed on (the negative value of) $\frac{db(c)}{dP}$ in a neighbourhood of $c = e$, relative to lower values of c, becomes arbitrarily large, and part (ii) of the proposition follows straightforwardly.

Likewise as $\gamma(\cdot)$ becomes more convex, i.e., if a strictly convex transformation is taken of γ, then again $\frac{dTE^X}{dP}$ will be negative for a wider constellation of parameters in the more convex case.[24] Finally, by the same logic, starting from a strictly positive optimal punishment, so $\frac{dTE^X}{dP} = 0$ at some $P > 0$, when the cost functions become more convex, ceteris paribus, $\frac{dTE^X}{dP}$ will become negative and so the optimal punishment will decrease, which justifies part (iii) of the following proposition.

Proposition 6 *(i) In an open contest with $k \geq 3$ players and strictly convex cost functions, if the density function $f(c)$ is non-decreasing in c on the interval $[\underline{s}, \overline{s}]$, it is optimal to set $P = 0$ in order to maximize expected total effort; (ii) For a given $f(\cdot)$, sufficient convexity of the cost function implies that it is optimal to set $P = 0$; (iii) Starting from a situation where the optimal punishment is strictly positive, when the cost functions become more convex (ceteris paribus), the optimal punishment will decrease.*

Roughly speaking, with convex cost functions, it becomes increasingly costly for a player to exert additional effort. Since, in equilibrium, a more able player exerts more effort than a less able player, the more able player is more discouraged by the increasing marginal cost of exerting effort. As P is increased, the extra effort exerted by higher ability players—the only ones who increase effort—is reduced relative to the reduced effort of the lower ability players. Consequently when cost functions

---

[24]From equation (20) in the Appendix, $\frac{db(c)}{dP}$ is decreasing in c, so that as costs become more convex an increasingly higher weight is placed on the more negative terms.

become more convex, total effort is more likely to fall when a punishment is either introduced or increased.

6 Concluding Remarks

We have studied a contest with a fixed prize where potential players can freely choose whether or not to enter. The contest designer can punish the bottom participant and we focused on the optimal punishment for maximizing either the expected total effort or the expected highest individual effort. By introducing a (sufficiently) small punishment, some low ability players drop out, medium ability players exert less effort and the highest ability players exert more effort. When the punishment is large enough, low ability players drop out and all participants exert less effort than without punishment. We further show that in order to maximize the expected total effort, punishment is guaranteed to be undesirable when the density function for the effort cost is nondecreasing—the contest designer expects there to be relatively few high ability players; on the other hand, to maximize the expected highest individual effort, punishment is considerably more likely to be desirable. In some circumstances there is a trade-off between maximizing the expected total effort and maximizing the expected highest individual effort. In addition, as cost functions become more convex, punishment becomes less desirable. Hence, depending on the objectives of the contest designer, the distribution of abilities and the convexity of costs, punishment may be part of the (optimal) answer.

In our model the prize is exogenously fixed, and we focussed on finding the optimal amount of punishment. It is however straightforward to see that if the exogenous prize becomes larger (smaller), the corresponding optimal punishment should be increased (decreased) by precisely the same proportion. When both the

prize and punishment are endogenously set, and increasing the prize is costly for the contest designer,[25] then the optimal prize (and corresponding punishment) will depend on the cost function of increasing the prize. This is beyond the scope of this paper and is left for future research.

We have maintained the assumption that the outside options of potential contestants are zero, so that there is no cost to staying out of the contest. However, it may not unreasonable to suppose that contestants have negative outside options. For instance, if an economics department increases its failure rate, students may have to suffer a cost in switching to a different course. This would allow a positive punishment to be introduced at no cost in terms of participation. Our model can be extended in a straightforward fashion to encompass such cases.

7   Appendix

7.1   Proof of Proposition 1

To maximize (2), the first-order condition is:

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

Rearranging:

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

$$(15)$$

Let y denote $b^{-1}(x)$. As in equilibrium $b(c) = x$, $c = b^{-1}(x) = y$. Then (15) can be written as

$$1 = -\frac{1}{y}(k-1)f(y)y'\{V[1-F(y)]^{k-2} + P[F(y)) + 1 - F(e)]^{k-2}\}. \qquad (16)$$

---

[25] This must be the case, otherwise the contest designer will want to set an infinite prize.

The marginal player with ability $c = e$ makes zero effort in equilibrium, this gives the boundary condition $y(0) = e$. The solution to the differential equation with the boundary condition is given by: $\int_x^0 dt = -G(y)$ where

$$G(y) = \int_y^e \frac{1}{t}(k-1)f(t)\{V[1 - F(t)]^{k-2} + P[F(t)) + 1 - F(e)]^{k-2}\}dt. \qquad (17)$$

Then we obtain that $x = G(y) = G(b^{-1}(x))$, therefore, $b \equiv G$, thus the effort function of every participant (who enters the contest actively) is given by

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

Thus,

$$b'(c) = -(k-1)\frac{1}{c}\{V[1 - F(c)]^{k-2} + P[F(c) + 1 - F(e)]^{k-2}\}f(c) < 0,$$

i.e., $b(c)$ is strictly decreasing and differentiable for $c \in [\underline{s}, e)$, as we assumed initially. Assuming other players with $c \in [\underline{s}, e)$ exert effort according to $b(c)$, we need to show that for any type c, the effort $b(c)$ maximizes the expected utility of that type. The necessary first order condition is satisfied by construction of $b(c)$. 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 an 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 $x = b(c)$. 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. Thus, by $\pi_{xc}(x,c) = -1 < 0$, we obtain $\pi_x(x,c) \geq \pi_x(x,\widehat{c})$. Since $x = b(\widehat{c})$, $\pi_x(x,\widehat{c}) = 0$ by the first-order condition, and therefore $\pi_x(x,c) \geq 0$ for every $x < b(c)$. A similar argument shows that $\pi_x(x,c) \leq 0$ for every $x > b(c)$.

26

7.2 Proof of Proposition 2

From (4), we derive

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

Recall that when P increases, e decreases, so $F(e)$ decreases and $P[F(c)+1-F(e)]^{k-2}$ increases. Thus $|b'(c)|$ gets larger as P increases and claim (a) follows. Hence $b(c)|_{P>0}$ is steeper than $b(c)|_{P=0}$. Thus, If $b(c)|_{P>0}$ and $b(c)|_{P=0}$ cross, they cannot cross more than once because $b(c)|_{P>0}$ is always steeper than $b(c)|_{P=0}$. Suppose they cross at point $c = c^*$; clearly $b(c)|_{P>0} > b(c)|_{P=0}$ for $c < c^*$ and $b(c)|_{P>0} < b(c)|_{P=0}$ for $c > c^*$. If they do not cross, $b(c)|_{P>0} < b(c)|_{P=0}$ for all c.

Next, we prove that when the punishment is sufficiently small, $b(c)|_{P>0}$ and $b(c)|_{P=0}$ will cross, or equivalently, that when the punishment is very small, $b(\underline{s})|_{P>0} > b(\underline{s})|_{P=0}$ (as $b(c)|_{P>0} = 0$ at $c = e < \bar{s}$). From (3):

$$\begin{aligned}
\frac{db(c)}{dP} = (k-1)\{&(\frac{de}{dP})\frac{1}{e}(V[1-F(e)]^{k-2} + P)f(e) \\
&+ \int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dt \\
&+ P(k-2)(-f(e))\frac{de}{dP}\int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)dt\}.
\end{aligned}$$

From (1),

$$P = (1-F(e))^{k-1}V, \tag{18}$$

so that

$$\frac{de}{dP} = \frac{-1}{(k-1)f(e)(1-F(e))^{k-2}V}. \tag{19}$$

Substituting (18) and (19) into the above equation:

$$\frac{db(c)}{dP} = (k-1)\{-\frac{1}{(k-1)e}(2-F(e))$$

$$+ \int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dt$$

$$+ \frac{(k-2)}{(k-1)}(1-F(e))\int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)dt\}. \qquad (20)$$

Let $P=0$ and $c=\underline{s}$; $P=0$ implies $e=\overline{s}$ so that $F(e)=F(\overline{s})=1$; thus

$$\frac{db(\underline{s})}{dP}\Big|_{P=0} = (k-1)\{\int_{\underline{s}}^{\overline{s}} \frac{1}{t}F(t)^{k-2}f(t)dt - \frac{1}{(k-1)\overline{s}}\}$$

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

$$= (k-1)\{\int_{\underline{s}}^{\overline{s}} (\frac{1}{t}-\frac{1}{\overline{s}})F(t)^{k-2}f(t)dt\} > 0.$$

Thus when a sufficiently small punishment is introduced, $b(\underline{s})|_{P>0} > b(\underline{s})|_{P=0}$ will hold and it follows that $b(c)|_{P>0}$ and $b(c)|_{P=0}$ cross once.

When $P \to V$, recall from (1) that $e \to \underline{s}$, so $b(\underline{s})|_{P>0} \to 0$ and consequently for large enough P we have $b(\underline{s})|_{P>0} < b(\underline{s})|_{P=0}$ and $b(c)|_{P>0}$ and $b(c)|_{P=0}$ do not cross.

7.3 Proof of Proposition 3

Recall that

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

Differentiating:

$$\frac{dZ}{dP} = \frac{de}{dP}(\frac{1}{e})[V(1-F(e))^{k-2} + P]f(e)f(c)$$

$$+ \int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dtf(c)$$

$$- (k-2)f(e)\frac{de}{dP}P\int_c^e \frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)dtf(c).$$

28

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:

$$\frac{dR_1}{dP} = \underbrace{\frac{de}{dP}(\frac{1}{e})[V(1-F(e))^{k-2} + P]f(e)\int_{\underline{s}}^{e} f(c)dc}_{(\alpha)}$$

$$+ \underbrace{\int_{\underline{s}}^{e}\{\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)dt\}f(c)dc}_{(\beta)}$$

$$\underbrace{-(k-2)\frac{de}{dP}f(e)P\int_{\underline{s}}^{e}\{\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)dt\}f(c)dc.}_{(\gamma)} \qquad (21)$$

Our aim is to prove that when $f(x)$ is non-decreasing in x, $\frac{dR_1}{dP} < 0$ for $0 \le P < V$, and thus the optimal punishment is zero. Substituting (18) and (19) into (α), we get

$$(\alpha) = -\frac{1}{(k-1)e}[2-F(e)]F(e). \qquad (22)$$

In (β), reversing the order of integration we can write

$$(\beta) = \int_{\underline{s}}^{e}\int_{\underline{s}}^{t}\frac{1}{t}[F(t)+1-F(e)]^{k-2}f(t)f(c)dcdt$$

$$= \int_{\underline{s}}^{e}\frac{F(t)}{t}[F(t)+1-F(e)]^{k-2}f(t)dt. \qquad (23)$$

By assumption $f'(t) \ge 0$. Let $g(t) := F(t)/t$, $h(t) := tf(t) - F(t)$, so that $h'(t) = tf'(t) \ge 0$; thus $h(t) > 0$ as $h(\underline{s}) = \underline{s}f(\underline{s}) > 0$. Consequently $g'(t) = (tf(t) - F(t))/t^2 = h(t)/t^2 > 0$. Hence for all $t < e$, $F(t)/t < F(e)/e$. Substituting into (23):

$$(\beta) = \int_{\underline{s}}^{e}\frac{F(t)}{t}[F(t)+(1-F(e))]^{k-2}f(t)dt$$

$$< \int_{\underline{s}}^{e}\frac{F(e)}{e}[F(t)+(1-F(e))]^{k-2}f(t)dt$$

$$= \frac{F(e)}{e(k-1)}[1-(1-F(e))^{k-1}] \le \frac{F(e)}{e(k-1)}. \qquad (24)$$

By a similar argument,

$$\int_{\underline{s}}^{e} \int_{c}^{e} \frac{1}{t} [F(t) + (1 - F(e))]^{k-3} f(t) f(c) dt dc < \frac{F(e)}{e(k-2)}. \tag{25}$$

Substituting (18), (19) and (25) into (γ), we derive

$$(\gamma) \leq \frac{F(e)(1 - F(e))}{(k-1)e} \tag{26}$$

(equality occurs when $P = 0$). From (22), (24) and (26), 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, $dR_1/dP < 0$ for all $P \in [0, V)$.

7.3.1   Proof of Claim in Case 1:

Substituting (22) and (23) into (21), and noting from substituting (18) and (19) into (21) that $(\gamma) = 0$ at $P = 0$,

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

For $\overline{s} = +\infty$ the first term on the R.H.S. is zero, and so $dR_1/dP|_{P=0} > 0$.

7.4   Proof of Proposition 4

Recall that

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

We get

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

30

Consequently,

$$\frac{dR_2}{dP} = \frac{de}{dP}X|_{c=e} + \int_{\underline{s}}^{e}\frac{dX}{dP}dc = \int_{\underline{s}}^{e}\frac{dX}{dP}dc$$

$$= \underbrace{\frac{de}{dP}(\frac{V}{e})[(1-F(e))^{k-2}+P]f(e)\int_{\underline{s}}^{e}(1-F(c))^{k-1}f(c)dc}_{(a)}$$

$$+ \underbrace{\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e))]^{k-2}f(t)(1-F(c))^{k-1}f(c)dtdc}_{(b)}$$

$$+ \underbrace{(-\frac{de}{dP})P(k-2)f(e)\int_{\underline{s}}^{e}\int_{c}^{e}\frac{1}{t}[F(t)+1-F(e)]^{k-3}f(t)(1-F(c))^{k-1}f(c)dtdc}_{(c)}.$$

Substituting (18) and (19) into (a) and (c), we get

$$(a) = -\frac{[2-F(e)][1-(1-F(e))^{k}]}{k(k-1)e}$$

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

When $P=0$, $e=\bar{s}$ and $F(e) = F(\bar{s}) = 1$, so that $(c) = 0$; thus

$$\frac{dR_2}{dP}|_{P=0} = (a) + (b) > 0$$

if and only if

$$\int_{\underline{s}}^{\bar{s}}\int_{c}^{\bar{s}}\frac{1}{t}F(t)^{k-2}f(t)(1-F(c))^{k-1}f(c)dtdc > \frac{1}{k(k-1)\bar{s}}. \tag{27}$$

We can change the order of integration so

$$LHS \ of \ (27) = \int_{\underline{s}}^{\bar{s}}\int_{\underline{s}}^{t}[\frac{1}{t}F(t)^{k-2}f(t)(1-F(c))^{k-1}f(c)dc]dt$$

$$= \frac{1}{k}\int_{\underline{s}}^{\bar{s}}\frac{1}{t}F(t)^{k-2}f(t)[1-(1-F(t))^{k}]dt.$$

So (27) holds if and only if

$$\int_{\underline{s}}^{\bar{s}}\frac{1}{t}[1-(1-F(t))^{k}]F(t)^{k-2}f(t)dt > \frac{1}{\bar{s}(k-1)}. \tag{28}$$

31

We can express $1/(\bar{s}(k-1))$ as $\int_{\underline{s}}^{\bar{s}} \frac{1}{\bar{s}} F(t)^{k-2} f(t) dt$. Thus (28) holds if and only if

$$\int_{\underline{s}}^{\bar{s}} \frac{1}{t} [1 - (1 - F(t))^k] F(t)^{k-2} f(t) dt - \int_{\underline{s}}^{\bar{s}} \frac{1}{\bar{s}} F(t)^{k-2} f(t) dt$$

$$= \int_{\underline{s}}^{\bar{s}} [(1 - \frac{t}{\bar{s}}) - (1 - F(t))^{k^*}] \frac{1}{t} F(t)^{k^*-2} f(t) dt > 0. \tag{29}$$

Consider the two terms inside the square brackets in (29), $(1 - \frac{t}{\bar{s}})$ and $(1 - F(t))^k$. On $(\underline{s}, \bar{s})$,

$$\frac{d(1 - F(t))^k}{dt} = -k(1 - F(t))^{k-1} f(t) < 0, \tag{30}$$

and moreover at $t = \bar{s}$,

$$\frac{d(1 - F(t))^k}{dt} = -k(1 - F(t))^{k-1} f(t) = 0. \tag{31}$$

Also $y = (1 - F(t))^k$ crosses the y and t axes at points $(t = \underline{s}, y = 1)$ and $(t = \bar{s}, y = 0)$ respectively and the linear function $y = (1 - \frac{t}{\bar{s}})$ crosses the y and t axes at points $(t = \underline{s}, y = 1 - \frac{\underline{s}}{\bar{s}})$ and $(t = \bar{s}, y = 0)$ respectively. Consider increasing k: the function $y = (1 - \frac{t}{\bar{s}})$ is unchanged but from (30) $(1 - F(t))^k$ is decreasing in k on $(\underline{s}, \bar{s})$ and converges to 0 as $k \to \infty$, with the two points $(t = \underline{s}, y = 1)$ and $(t = \bar{s}, y = 0)$ staying fixed. Thus for any $\varepsilon > 0$ and $t^* > \underline{s}$, there exists a k_1^* such that for $k > k_1^*$, $(1 - F(t))^k < \varepsilon$ and (using (31)) $(1 - F(t))^k < (1 - \frac{t}{\bar{s}})$ on $[t^*, \bar{s})$. Consequently $[(1 - \frac{t}{\bar{s}}) - (1 - F(t))^k] < 0$ on an arbitrarily small set close to $\underline{s}$.

Next consider $\frac{1}{t} F(t)^{k-2} f(t)$. Since $F(t)$ increases from 0 to 1 when t increases from $\underline{s}$ to $\bar{s}$, when k gets larger, $\frac{1}{t} F(t)^{k-2} f(t)$ will assign a relatively larger/smaller weight to $[(1 - \frac{t}{\bar{s}}) - (1 - F(t))^k]$ for a large/small t. It is then straightforward to show that by letting $\varepsilon \to 0$ and $t^* \to \underline{s}$, the last two facts together imply that there must exist a k_2^* such that (29) holds for all $k > k_2^*$. This completes the proof.

32

7.5 Proof of Case 2

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 then follows that

$$\frac{dY}{dP} = \frac{de}{dP}(\frac{V}{e})[(\overline{s} - e)^{k-2} + P(\overline{s} - \underline{s})^{k-2}](\overline{s} - c)^{k-1}$$
$$+ \int_{c}^{e} \frac{1}{t}(t + \overline{s} - \underline{s} - e)^{k-2}(\overline{s} - c)^{k-1} dt$$
$$+ (-\frac{de}{dP})P(k-2)\int_{c}^{e} \frac{1}{t}(t + \overline{s} - \underline{s} - e)^{k-3}(\overline{s} - c)^{k-1} dt.$$

Hence

$$\frac{dR_2}{dP} = \frac{1}{(\overline{s} - \underline{s})^{2k-1}}\{\frac{de}{dP}Y|_{c=e} + \int_{\underline{s}}^{e} \frac{dY}{dP} dc\} = \frac{1}{(\overline{s} - \underline{s})^{2k-1}} \int_{\underline{s}}^{e} \frac{dY}{dP} dc$$
$$= \frac{1}{(\overline{s} - \underline{s})^{2k-1}}\{\underbrace{\frac{de}{dP}(\frac{V}{e})[(\overline{s} - e)^{k-2} + P(\overline{s} - \underline{s})^{k-2}]\int_{\underline{s}}^{e}(\overline{s} - c)^{k-1}dc}_{(a_1)}$$
$$+ \underbrace{\int_{\underline{s}}^{e}\int_{c}^{e} \frac{1}{t}(t + \overline{s} - \underline{s} - e)^{k-2}(\overline{s} - c)^{k-1} dt dc}_{(b_1)}$$
$$+ \underbrace{(-\frac{de}{dP})P(k-2)\int_{\underline{s}}^{e}\int_{c}^{e} \frac{1}{t}(t + \overline{s} - \underline{s} - e)^{k-3}(\overline{s} - c)^{k-1} dt dc}_{(c_1)}\}.$$

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

$$P = (\frac{\overline{s} - e}{\overline{s} - \underline{s}})^{k-1}V; \tag{32}$$

$$\frac{de}{dP} = \frac{-(\overline{s} - \underline{s})^{k-1}}{(k-1)(\overline{s} - e)^{k-2}V}. \tag{33}$$

Substituting (32) and (33) into (a_1) and (c_1), we have

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

33

and

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

When $P = 0$, $e = \overline{s}$, and

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

$$(b_1) = \int_{\underline{s}}^{\overline{s}} \int_{c}^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1} dt dc;$$

$$(c_1) = 0 \times \int_{\underline{s}}^{\overline{s}} \int_{c}^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-3}(\overline{s}-c)^{k-1} dt dc = 0.$$

Thus we have

$$\frac{dR_2}{dP}\Big|_{P=0} = \frac{1}{(\overline{s}-\underline{s})^{2k-1}} \{ \int_{\underline{s}}^{\overline{s}} \int_{c}^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1} dt dc - \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}} \}.$$

Therefore, $dR_2/dP|_{P=0} > 0$ if and only if

$$\int_{\underline{s}}^{\overline{s}} \int_{c}^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1} dt dc > \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}}. \tag{34}$$

So the optimal punishment is strictly positive when (34) holds. We can change the order of integration, so

$$LHS \ of \ (34) = \int_{\underline{s}}^{\overline{s}} \int_{\underline{s}}^{t} \frac{1}{t}(t-\underline{s})^{k-2}(\overline{s}-c)^{k-1} dc dt$$

$$= \frac{1}{k} \int_{\underline{s}}^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-2}[(\overline{s}-\underline{s})^{k} - (\overline{s}-t)^{k}] dt.$$

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 $\underline{s} \le t \le \overline{s}$, $0 \le (\overline{s}-t)/(\overline{s}-\underline{s}) \le 1$, i.e., $0 \le v \le 1$. Notice that $v = 1$ when $t = \underline{s}$ and $v = 0$ when $t = \overline{s}$. Then we have

$$LHS \ of \ (34) = \frac{1}{k} \int_{\underline{s}}^{\overline{s}} \frac{1}{t}(t-\underline{s})^{k-2}[(\overline{s}-\underline{s})^{k} - (\overline{s}-t)^{k}] dt$$

$$= \frac{(\overline{s}-\underline{s})^{2k-1}}{k} \int_{0}^{1} \frac{(1-v)^{k-2}(1-v^{k})}{\overline{s}-v(\overline{s}-\underline{s})} dv.$$

34

We claim that for all $k \geq 3$, (34) holds if

$$\int_0^1 (1-v)^{k-2} \{ \frac{(1-v^3)}{1-v(1-(\underline{s}/\overline{s}))} - 1 \} dv > 0. \qquad (35)$$

This is true because

$$(35) \Rightarrow \frac{1}{\overline{s}} \int_0^1 (1-v)^{k-2} \{ \frac{(1-v^3)}{1-v(1-(\underline{s}/\overline{s}))} - 1 \} dv > 0$$

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

$$\Rightarrow \int_0^1 (1-v)^{k-2} \{ \frac{(1-v^k)}{\overline{s} - v(\overline{s}-\underline{s})} - \frac{1}{\overline{s}} \} dv > 0 \text{ (since } k \geq 3)$$

$$\Rightarrow \int_0^1 \frac{(1-v)^{k-2}(1-v^k)}{\overline{s} - v(\overline{s}-\underline{s})} dv > \frac{1}{\overline{s}} \int_0^1 (1-v)^{k-2} dv$$

$$\Rightarrow \int_0^1 \frac{(1-v)^{k-2}(1-v^k)}{\overline{s} - v(\overline{s}-\underline{s})} dv \geq \frac{1}{(k-1)\overline{s}}$$

$$\Rightarrow \frac{(\overline{s}-\underline{s})^{2k-1}}{k} \int_0^1 \frac{(1-v)^{k-2}(1-v^k)}{\overline{s} - v(\overline{s}-\underline{s})} dv \geq \frac{(\overline{s}-\underline{s})^{2k-1}}{k(k-1)\overline{s}}.$$

Let

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

then the *LHS of* (35) becomes

$$\int_0^1 (1-v)^{k-2} j(v) dv. \qquad (36)$$

We can see that the sign of $(1-v)^{k-2}j(v)$ is determined by $j(v)$ as $0 \leq v \leq 1$. Graphically, the value of (36) is equal to the area between the v axis and the curve $(1-v)^{k-2}j(v)$ on the interval $[0,1]$. From the expression for $j(v)$, we can prove that when $0 \leq v \leq 1$,

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

As k increases in (36), 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

35

become negative. Therefore, we conclude that if $\int_0^1 (1-v)^{k-2} j(v) dv|_{k=3} > 0$, then for all $k \geq 3$, $\int_0^1 (1-v)^{k-2} j(v) dv > 0$. We have:

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

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

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

By analyzing the above equation, it is easy to check that when $0 < (\underline{s}/\overline{s}) \leq 0.68$, i.e., when $(\overline{s}/\underline{s}) \geq 1.47$, $\int_0^1 (1-v)^{k-2} j(v) dv|_{k=3} > 0$. Thus, the optimal punishment is strictly positive for all $k \geq 3$.

7.6   Proof of Proposition 5

As now the cost function is $c\gamma(x)$ instead of cx, player $i's$ maximization problem becomes:

$$\underset{x}{Max}\{V \times \underbrace{[1 - F(B^{-1}(x))]^{k-1}}_{\Pr(x \text{ is the highest})} - P \times \underbrace{[F(B^{-1}(x)) + 1 - F(e)]^{k-1}}_{\Pr(x \text{ is the lowest})} - c\gamma(x)\}.$$

Let y be the inverse of B, i.e., $y(\cdot) = B^{-1}(\cdot)$. As $B(c) = x$, $c = B^{-1}(x) = y(x)$. Then the FOC can be written as

$$\gamma'(x) = -\frac{1}{y}(k-1)f(y)y'\{V[1-F(y)]^{k-2} + P[F(y)) + 1 - F(e)]^{k-2}\}.$$

Using boundary condition $y(e) = 0$ and integration, we can derive that $\gamma(x) = G(y)$ where $G(y)$ is defined exactly by (17). Thus, $x = \gamma^{-1}(G(y))$, then $B = x = g(G(y)) = g(b(c))$. The equilibrium effort function (12) is strictly decreasing since for all $c \in [\underline{s}, e)$, it can be shown that $\frac{dg}{dc} = g'b' < 0$. For the sufficient second-order condition we proceed exactly as in the proof of Proposition 1.

References

[1] **Akerlof, R. and Holden, R. (2010).** "The Nature of Tournaments," *Economic Theory*, forthcoming.

[2] **Baye, M.R., Kovenock, D. and de Vries, C.G. (1996).** "The All-Pay Auction with Complete Information," *Economic Theory,* 8(2), 291-305.

[3] **Cason, T.N., Masters, W.A. and Sheremeta, R.M. (2010).** "Entry into Winner-take-all and Proportional-prize Contests: An Experimental Study," *Journal of Public Economics*, 94, 604-611.

[4] **Fu, Q. and Lu, J. (2010).** "Contest Design and Optimal Endogenous Entry," *Economic Inquiry*, 48(1), 80-88.

[5] **Fu, Q., Jiao, Q. and Lu, J. (2010).** "On Disclosure Policy in Contests with Stochastic Entry.," *Public Choice*, forthcoming.

[6] **Fu, Q., Jiao, Q. and Lu, J. (2011).** "Contests with Endogenous and Stochastic Entry," working paper, National University of Singapore.

[7] **Gilpatric, S.M. (2009).** "Risk-taking in contests and the role of carrots and sticks," Economic Inquiry, 47(2), 266–277.

[8] **Glazer, A. and Hassin, R. (1988).** "Optimal Contests," *Economic Inquiry,* 26(1), 133-43.

[9] **Green, J.R. and Stokey, N.L. (1983).** "A Comparison of Tournaments and Contracts," *Journal of Political Economy,* 91(3), 349-64.

[10] **Higgins, R.S., Shughart W.F.II. and Tollison, R.D. (1985).** "Free Entry and Effcient Rent Seeking," *Public Choice*, 46, 247-258.

[11] **Kaplan, T.R. and Sela, A. (2010).** "Effective Contests," *Economics Letters*, 106, 38-41.

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

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

[14] **Levitt, S.D. (1995).** "Optimal Incentive Schemes when Only the Agents' "Best" Output Matters to the Principal." *RAND Journal of Economics*, 26, 744–60.

[15] **Lim, W. and Matros, A. (2009)**. "Contests with a Stochastic Number of Players," *Games and Economic Behavior*, 67, 584-597.

[16] **Minor, Dylan B. (2012).** "Increasing Revenue through Rewarding the Best Less," working paper, Kellogg School of Management, Northwestern University.

[17] **Moldovanu, B. and Sela, A. (2001).** "The Optimal Allocation of Prizes in Contest," *American Economic Review*, 91(3), 542-558.

[18] **Moldovanu, B. and Sela, A. (2006).** "Contest Architecture," *Journal of Economic Theory*, 126(1), 70-97.

[19] **Moldovanu, B. and Sela, A. and Shi, X. (2010).** "Carrots and Sticks: Prizes and Punishments in Contests," *Economic Inquiry*, forthcoming.

[20] **Morgan, J., Orzen, H. and Sefton, M. (2010)**. "Endogenous Entry in Contests," *Economic Theory*, forthcoming.

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

[22] **Myerson, R.B. and Warneryd, K. (2006).** "Population Uncertainty in Contests," *Economic Theory*, 27(2), 469-474.

[23] **Nalebuff, B.J. and Stiglitz, J.E. (1983).** "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *The Bell Journal of Economics*, 14(1), 21-43.

[24] **Riley, J.G. and Samuelson, W.F. (1981).** "Optimal Auctions," *American Economic Review,* 71(3), 381-92.

[25] **Shaked, M. and Shantikumar, G.J. (1994).** "Stochastic Orders and their Applications," San Diego, CA: Academic Press.

[26] **Taylor, C. R. (1995).** "Digging for Golden Carrots: An Analysis of Research Tournaments." *American Economic Review*, 85(4), 872–90.